International Symposium of Forecasting 2024

June 30-July 3, Dijon, France

General Chair: Mohsen Hamoudia, Program Chair: Laurent Ferrara

Book of Abstracts

Unveiling Clarity: Exploring Central Bank Communication And Public Emotions In Sub-Saharan Africa

Presenter: Keegan Chisha
Co-authors: Keegan Chisha;Luigi Gifuni

This study explores the communication clarity of 14 central banks across sub-Saharan Africa and its potential correlation with human emotions as reflected in country-specific leading newspapers. The study delves into the pivotal role of communication in central banking, particularly in a region marked by diverse economic landscapes and socio-political contexts. Through a comprehensive analysis of official communications from these central banks and sentiment analysis of newspaper articles, this research aims to shed light on the relationship between clarity in central bank communication and the emotional responses elicited among the populace. The findings of this study hold implications for policymakers, economists, and communication practitioners, offering insights into the effectiveness of central bank communication strategies in fostering transparency, trust, and public understanding.

Capturing International Influences In Us Monetary Policy Through A Nlp Approach

Presenter: Nicolas de Roux
Co-authors: Nicolas De Roux;Laurent Ferrara

The U.S. Federal Reserve has a statutory dual domestic mandate of price stability and full employment. In this paper, we question the role of the international environment in shaping Fed monetary policy decisions. In this respect, we apply natural language processing (NLP) techniques to minutes of the Federal Open Market Committee (FOMC) and construct indexes of the attention paid by U.S. monetary policymakers to the international economic and financial situation, as well as sentiment indexes. By integrating those text-based indicators into a Taylor rule, we derive various quantitative measures of the external influences on Fed decisions. Our results show that when there is a focus on international topics within the FOMC, the Fed's monetary policy generally tends to be more accommodative than expected by a standard Taylor rule. This result is robust to various alternatives that includes a time-varying neutral interest rate or a shadow central bank interest rate.

The Role Of Human Emotions In Commodity Markets

Presenter: Francesco Ravazzolo
Co-authors: Francesco Ravazzolo;Joseph Byrne;Luigi Gifuni

This paper explores the often under-theorised interrelation between human emotions and commodity risk within financial and macroeconomic markets. While traditional risk models focus on quantitative

factors, our study delves into the psychological aspects that shape market dynamics. By exploiting information embedded in economic leading newspapers, we use emotional sentiment as a feature to detect country-specific commodity risk. Understanding how human sentiment interplays with market trends provides valuable insights for risk management strategies. Our findings contribute to a more holistic approach to risk assessment, shedding light on the human element that underlies the complexities of commodity markets. This research has practical implications for investors, policymakers, and risk analysts seeking to enhance their understanding of the nuanced factors driving commodity price movements.

Spike-And-Slab Group Dirichlet-Laplace Priors For Sparse Shrinkages

Presenter: Deborah Gefang
Co-authors: Deborah Gefang

Dirichlet-Laplace (DL) priors of prior studies have proved powerful in variable selection and parameter estimations. However, similar to many other popular global-local shrinkage priors, DL priors usually produce posterior means (medians) that are close to zero but not exactly zero for true zero parameters, and underestimate the magnitudes of true non-zero parameters when the number of variables is large. Spike-and-slab group Dirichlet-Laplace priors is introduced to identify variable groups and sparsity both at the group level and within groups.

Forecasting Macroeconomic Data With Bayesian Vars: Sparse Or Dense? It Depends!

Presenter: Gregor Kastner
Co-authors: Gregor Kastner;Luis Gruber

Vectorautogressions (VARs) are widely applied when it comes to modeling and forecasting macroeconomic variables. In high dimensions, however, they are prone to overfitting. Bayesian methods, more concretely shrinking priors, have shown to be successful in improving prediction performance. In the present paper, we introduce the semi-global framework, in which we replace the traditional global shrinkage parameter with group-specific shrinkage parameters. We show how this framework can be applied to various shrinking priors, such as global-local priors and stochastic search variable selection priors. We demonstrate the virtues of the proposed framework in an extensive simulation study and in an empirical application forecasting data of the US economy. Further, we shed more light on the ongoing "Illusion of Sparsity" debate, finding that forecasting performances under sparse/dense priors vary across evaluated economic variables and across time frames. Dynamic model averaging, however, can combine the merits of both worlds.

Bayesian Bi-Level Sparse Group Regressions For Macroeconomic Forecasting

Presenter: Matteo Mogliani
Co-authors: Matteo Mogliani;Anna Simoni

In this paper we construct optimal forecasts for macroeconomic aggregates in presence of a large number of series that can be cast into groups. The aim is to provide policymakers with tools that accurately track economic conditions in real time. The variables in each group have strong covariation and common characteristics and patterns. The group structure is exploited by designing a convenient prior that induces a bi-level sparsity – at the group level and within group. Such a sparsity structure mirrors the fact that not all predictors are relevant to forecast the target series, conditional on the remaining groups and predictors. This is particularly true when there are many predictors weakly correlated with the target variable, as it is often the case with predictors arising from alternative sources. Under the assumption that the true data generating process exhibits this bi-level sparse structure, our posterior distribution is able to recover the optimal forecast asymptotically and its support is made of parameters with at most the same sparsity as the true sparse model. The rate of contraction of the posterior distribution is recovered. Finite sample

properties of our procedure are illustrated through Monte Carlo experiments. We illustrate the performance of our procedure with real data through a nowcasting exercise of the quarterly growth rate of the US GDP.

The Experience Formation Mechanism

Presenter: Alexander Glas
Co-authors: Alexander Glas;Christian Conrad;Zeno Enders

We show that inflation expectations of households do not depend on experienced inflation directly; they are rather influenced by remembered inflation. This makesa difference as remembered and actually experienced inflation generally differ. We therefore investigate how inflation experiences are formed and what they depend on.Our main findings are: (i) On average, households overestimate lifetime inflation. (ii) While higher remembered average inflation leads to higher expectations of futureinflation rates, there is no such effect for actual average inflation. (iii) Higher remembered peak inflation increases current perceived inflation. Moreover, the higher the remembered peak, the stronger respondents extrapolate from perceived current to expected future inflation.

Framing Effects In Consumer Expectation Surveys

Presenter: Lora Pavlova
Co-authors: Lora Pavlova

In a randomized experiment embedded in a survey, I test the effects of variations in question wording and format on consumer response behavior and the corresponding inflation expectations. To this end, survey participants from a representative sample of German consumers are broken down into four treatment groups and presented with different versions of a question asking for their subjective distribution for inflation over the next 12 months. As part of the experiment, two competing wordings, previously known from leading consumer surveys, are considered: (i) the change in prices in general or (ii) the inflation rate. In addition, I compare the responses to a question asking for consumers' probabilistic beliefs about future inflation, to those from a simpler one asking for the expected minimum, maximum, and most likely inflation rate over the short term. I find that response behavior varies strongly with framing. Simpler wording such as 'prices in general' and a less restrictive format lead to higher mean expected inflation, on average. While simpler wording increases individual uncertainty derived from the subjective histograms, asking for minimum, maximum and mode leads to lower uncertainty about expected inflation.

The Effects Of Interest Rate Increases On Consumers' Inflation Expectations: The Roles Of Informedness And Compliance

Presenter: James Mitchell
Co-authors: James Mitchell;Ed Knotek;Mathieu Pedemonte;Taylor Shiroff

We study how monetary policy communications associated with increasing the federal funds rate causally affect consumers' inflation expectations. In a large-scale, multi-wave randomized controlled trial (RCT), we find weak evidence on average that communicating policy changes lowers consumers' medium-term inflation expectations. However, information differs systematically across demographic groups, in terms of ex ante informedness about monetary policy and ex post compliance with the information treatment. Monetary policy communications have a much stronger effect on people who had not previously heard news about monetary policy and who take sufficient time to read the treatment, implying scope to increase the impact of communications by targeting specific groups of the general public. Our findings show that, in an inflationary environment, consumers expect that raising interest rates will lower inflation. More generally, our results emphasize the importance of measuring both respondents' information sets and their compliance with treatment when using RCTs in empirical macroeconomics, to better understand the well-documented evidence of heterogeneous treatment effects.

Near-Real-Time Analysis Of The German Economy Based On High Sampling Data

Presenter: Dirk Beinert
Co-authors: Dirk Beinert;Matthias Dorner

Analyzing and forecasting economic developments is a crucial aspect for managing companies individually and for decision making in politics. Most studies focus either on highly aggregated data of the GDP and respective derivatives, survey data or individual company data. With the support and consent of German tax consultants, we constructed a unique large scale data set of taxes, balance sheets and wages of more than 800,000 German companies across various industry sectors. This opens the possibility to analyze the German economy on a at least monthly interval across various dimensions.Accordingly, this talk presents current advances of financial process data analysis versus public surveys and panel data.We display the economic drift of German businesses with time series methods and compare it with official statistic news.

Analyzing The Effectiveness Of Deep Learning And Conventional Machine Learning Techniques In Time Series Forecasting Of Daily Demand For An Automotive Product

Presenter: Andreas Rükgauer
Co-authors: Andreas Rükgauer;David Tumma

This study explores the effectiveness of deep learning and conventional machine learning techniques for time series forecasting of daily demand for an automotive product. The research compares the performance of neural network architectures, including Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Convolutional Neural Networks (CNN), a hybrid model (CNN-LSTM), and traditional statistical approaches such as eXtreme Gradient Boosting (XGBoost), Autoregressive Integrated Moving Average (ARIMA), Holt-Winters exponential smoothing and Prophet. The study also investigates the impact of different data preparation approaches on forecasting accuracy. The results demonstrate that deep learning models, particularly LSTM and GRU, outperform conventional techniques in capturing complex patterns and dependencies in the time series data. The data preparation approach, which uses a sequence of past elements to predict the next step, proves to be the most effective for multi-step forecasting. The findings provide valuable insights for improving demand forecasting accuracy in the automotive industry and optimizing production planning processes.

Historical Consistent Neural Networks: From Theory To Practice

Presenter: Christian Menden
Co-authors: Christian Menden;Hans Georg Zimmermann

Time series modelling is often seen as a process in which one starts with a data set, identifying features of the data (e.g. periodicity) and go on to the model building. In this talk we will choose another approach – we start from first principles in the modeling of dynamical systems and work out a neural network design before we use the training data. Let's base on open recurrent networks and their universal approximation property to explain a complicated output as a superposition of an eigen-dynamic and external drivers. But in the future we do not have external drivers – so we have to embed the open into a closed dynamical system framework. Another point is the number of matrices we use in the modelling: optimal would be only one state transition matrix to avoid a different learning speed in the network parts. This is especially important in overparameterized models. And last but not least in which way can we scale models to large state or output dimensions. All of this is of interest before we fill in the data.

Spectral Signature Analysis In Large-Scale Spatio-Temporal Dynamics: Case Study In Forecasting Power Grid Load Profiles

Presenter: Heman Shakeri; Behnaz Moradijamei
Co-authors: Heman Shakeri;Ali Tavasoli;Behnaz Moradijamei

Electricity load forecasting is crucial for effectively managing and optimizing power grids. Over the past few decades, various statistical and deep learning approaches have been used to develop load forecasting models. We will present an interpretable machine learning approach that identifies load dynamics using data-driven methods within an operator-theoretic framework. We utilize the Koopman operator, inherent to the underlying dynamics, to represent the load data. By computing the corresponding eigenfunctions, we decompose the load dynamics into coherent spatiotemporal patterns, which are the most robust features of the dynamics. Each pattern evolves independently according to its single frequency, facilitating predictability based on linear dynamics. We highlight that these dynamics are constructed from coherent spatiotemporal patterns intrinsic to the system, encoding rich dynamical features across multiple time scales. These features stem from complex interactions over interconnected power grids and various exogenous effects. To implement the Koopman operator approach more efficiently, we cluster the load data using a kernel-based clustering method, identifying power stations with similar load patterns, especially those exhibiting synchronized dynamics. We evaluate our approach using a large-scale dataset from a renewable electric power system within the continental European electricity system, demonstrating that the Koopman-based approach surpasses a deep learning (LSTM) architecture in accuracy and computational efficiency.

Short-Term Ensemble Load Forecasting For Transmission Grid Congestions

Presenter: Jean Thorey
Co-authors: Jean Thorey;Ouassim Feliachi;Benjamin Donnot

The electric grid management is mainly driven by congestion forecasts. Grid operators take action to avoid congestion events, at a minimal cost, to keep the grid within safety margins.Each power line is given an intensity threshold. Its exceedance is a congestion. Given injections and grid topology, the load flow computes the flows (power, intensity) for every power line, and subsequently congestions.Congestion forecasts rely on several forecasts: load, intermittent generation, dispatchable generation, topology, and interconnections. Interestingly, the forecasts are local to provide nodal injections, but the load and intermittent generation forecasts show strong spatial and temporal correlations.The first part of the work focuses on a better understanding of uncertainty sources, depending on the lead time and the location. A naive but informative approach switches observations and forecasts and recompute flows. This approach gives answer to the following kind of interrogations. How much more accurate would power flow forecasts be if the electric load were perfectly predicted?An in-depth sensitivity analysis would rely on sampling the joint distribution of thousands of inputs. Think number of cities and renewable power plants in a country.Current operations for the French TSO RTE rely on deterministic forecasts only. The risk policy for network security is achieved through the so-called N-1 policy where the congestion constraints must be satisfied even under the disconnection of any power line. The second part of the work explores ensemble load forecasting, as a deterministic forecast perturbation with historical residuals. Ensemble load forecasting is a step towards probabilistic congestion forecasts. We investigate lead times up to a few hours.Several techniques including empirical copulas are compared under two validation paradigms. First, we check whether ensemble load forecasts are calibrated, also under a multivariate perspective. On the second hand, for the power lines which are highly sensitive to the load, we quantify the variability of the ensemble power flow forecasts obtained by running a load flow on each load forecast member.

Prosumer Electric Load Forecasting Using Machine Learning-Based Algorithms

Presenter: Kasım Zor
Co-authors: Kasım Zor

Due to the ascending penetration of renewable distributed generators into the modern electric power systems since the end of the second millennium, the traditional consumer concept in the electricity markets has been evolved into the contemporary concept of prosumer which can be referred to as an individual who consumes and produces electricity within the age of smart grid. In order to preserve the vital equilibrium between the generation and consumption of electricity, prosumer electric load forecasting has come into a prerequisite in energy management and planning of today's microgrids. Machine learning (ML)-based algorithms are frequently implemented in forecasting the electric loads whose characteristics can be defined as nonlinear and affected by a variety of factors such as seasonal effects and weather conditions. The aim of this study is to present a benchmark appertaining to prosumer electric load forecasting for a residential building located in New York, USA via using ML-based algorithms consisting of gene expression programming (GEP), gradient boosted decision trees (GBDT), and group method of data handling-type deep polynomial neural (GMDH) networks within the very short-term horizon. In addition to calendar and historical electrical variables, utilised data set has been enriched by introducing meteorological variables with a temporal granularity of 15-minute obtained from MERRA-2 database of NASA. The proposed models have been validated by simple random sampling and the benchmark has been meticulously performed with respect to several performance and error metrics. Consequently, the results have revealed that GBDT outperformed the others in terms of coefficient of determination (93.504%) and mean absolute error (0.468) under the scenario of selecting explanatory features according to the Pearson correlation. Additionally, it is thought that there is a gap in the literature for prosumer electric load forecasting, and this study will bridge this gap along with guiding the potential researchers in the field.

Detecting Bubbles In Financial Time Series Data Through Mixed Integer Programming

Presenter: Alexander Semenov
Co-authors: Alexander Semenov;Artem Prokhorov;Anton Skrobotov;Peter Radchenko

The detection of bubbles in financial time series data has been extensively studied in recent research. In this paper, we present an optimization model based on mixed integer optimization (MIO) for identifying such bubbles. Our model introduces a novel formulation to capture the dynamics of rational bubbles. Through empirical analysis, we demonstrate the effectiveness of our approach in detecting and characterizing rational bubbles, offering valuable insights for financial analysts and policymakers. We further validate our method through extensive numerical experiments, showing more accurate estimation of bubbles compared to popular non-MIO methods, and demonstrate its empirical applications.

An Early Warning System for Financial Markets: Entropy Meets Machine Learning

Presenter: Artem Prokhorov
Co-authors: Artem Prokhorov

We develop and apply a new online early warning system (EWS) for what is known in machine learning as concept drift, in economics as a regime shift and in statistics as a change point. The system goes beyond linearity assumed in many conventional methods, and is robust to heavy tails and tail-dependence in the data, making it particularly suitable for emerging markets. The key component is an effective change-point detection mechanism for conditional entropy of the data, rather than for a particular indicator of interest. Combined with recent advances in machine learn- ing methods for high-dimensional random forests, the mechanism is capable of finding significant shifts in information transfer between interdependent time series when traditional methods fail. We explore when this happens using simulations and we provide illustrations by applying the method to Uzbekistan's commodity and equity markets as well as to Russia's equity market in 2021-2023.

Adaptive Now- And Forecasting Of Global Temperatures Under Smooth Structural Changes

Presenter: Robinson Kruse-Becher
Co-authors: Robinson Kruse-Becher

Accurate short-term now- and forecasting of global temperatures is an important issue and helpful for policy design and decision making in the public and private sector. We compose a raw mixed-frequency data set from weather stations around the globe (1920-2020). First, we document smooth variation in average monthly and annual temperature series by applying a dynamic stochastic coefficient model. Second, we use adaptive cross-validated forecasting methods which are robust to smooth changes of unknown form in the short-run. Therein, recent and past observations are weighted in a mean squared error-optimal way. Overall, it turns out exponential smoothing methods (with bootstrap aggregation) often performs best. Third, by exploiting monthly data, we propose a simple procedure to update annual nowcasts during a running calendar year and demonstrate its usefulness. Further, we show that these findings are robust with respect to climate zones. Finally, we investigate now- and forecasting of climate volatility via a range-based measure and a quantile-based climate risk measure.

Catastrophe Duration And Loss Prediction Via Natural Language Processing

Presenter: Han Li
Co-authors: Han Li;Han Wang;Wen Wang;Feng Li;Yanfei Kang

Textual information from online news is more timely than insurance claim data during catastrophes, and there is value in using this information to achieve earlier damage estimates. In this paper, we use text-based information to predict the duration and severity of catastrophes. We construct text vectors through Word2Vec and BERT models, using Random Forest, LightGBM, and XGBoost as different learners, all of which show more satisfactory prediction results. This new approach is informative in providing timely warnings of the severity of a catastrophe, which can aid decision-making and support appropriate responses.

Machine Learning Method For Predicting The Impact Of Socioeconomic-Related Climate Change On Global Pm2.5 Air Quality

Presenter: Shan Shan
Co-authors: Shan Shan

Air pollution caused by fine particulate matter (diameter of 2.5 m or less) (PM2.5) is a major global concern, particularly for human health. This study used machine learning tools to uncover the social factors influencing PM2.5 emissions. Text mining techniques were employed to extract key variables from previous research databases related to the target variable PM2.5 air pollution and its determinants. Four important features were derived, encompassing a wide range of factors: PM2.5, air pollution levels, population metrics, gross domestic product per capita, military expenditure, health expenditure, and environmental features. This study identified significant changes in PM2.5, related to health expenditures and economic contributions, emphasizing the need for interdisciplinary efforts to address this global problem. From a technical standpoint, machine-learning feature extraction was deployed to identify four critical factors significantly influencing air quality. The support vector regression method stood out in terms of its model efficacy. This technique excels in producing accurate predictions and understanding the correlation between key social factors and global exposure to PM2.5. This study contributes to the understanding of the complex relationships between socioeconomic variables and PM2.5 and sheds light on the multifaceted factors influencing PM2.5 emissions.

Two Stage Volatility-Return Models Using Deep Learning Density Networks With Application To Cryptocurrencies

Presenter: Niya Chen
Co-authors: Niya Chen;Jennifer Chan

This paper introduces density networks structured around the two-stage volatility-return models that transition from volatility forecasting to return prediction. The initial stage harnesses the capabilities of deep learning density networks for volatility measures and incorporates the negative log-likelihood of Generalized Beta Type two (GB2) and Weibull distributions as the loss functions. Subsequently, the second stage return networks incorporate predicted volatilities from the volatility network and define the negative log-likelihood loss function using the normal and t-distributions. Our empirical examination focuses on daily returns and Parkinson volatility measures of Bitcoin and Ethereum, highlighting the critical influence of distributional assumptions on the accuracy of volatility and return quantile predictions. The findings underscore the superiority of our density network approach over traditional statistical models in out-of-sample forecasting performance.

Machine Learning Density Networks In Auto Insurance Claim Prediction Using Telematics Data

Presenter: Alice Xiaodan Dong
Co-authors: Xiaodan (Alice) Dong;Jennifer Chan;Farha Osman

Neural networks are instrumental in optimizing the precision and efficiency of insurance claim predictions, offering advanced models for comprehensive analysis and forecasting of claim outcomes. This research proposes a methodological refinement by introducing the Poisson mixture negative log-likelihood function into the estimation process of deep learning neural networks. This modification is specifically designed to address the inherent variability in risk profiles among drivers, particularly in instances marked by an abundance of zero occurrences in insurance claims.Moreover, an empirical investigation into the Deep Learning Neural Network (DLNN) performance is conducted through the application of ensembling methods. This examination seeks to identify and understand the contributing factors behind the model's superior predictive capabilities. A comparative analysis shows that the proposed model outperforms conventional models.

New Loss Reserve Models With Persistence Effects To Forecast Losses In Run-Off Triangles

Presenter: Jennifer Chan
Co-authors: Jennifer Chan;Farha Osman

Modelling loss reserve data in run-off triangles is challenging due to the complex but unknown dynamics in the claim/loss process. Popular loss reserve models describe the mean process through development year, accident year, and calendar year effects using the analysis of variance and covariance (ANCOVA) models. We propose to include in the mean function the persistence terms in the conditional autoregressive range model for modelling the persistence of claim across development years. In the ANCOVA model, we adopt linear trends for the accident and calendar year effects and a quadratic trend for the development year effect. We investigate linear or log-transformedmean functions and four distributions, namely generalised beta type 2, generalised gamma, Weibull, and exponential extension, with positive support to enhance the model flexibility. The proposed models are implemented using the Bayesian user-friendly package Stan running in the R environment. Results show that the models with log-transformed mean function and persistence terms provide better model fits. Lastly, the best model is applied to forecast partial loss reserve and calendar year reserve for three years.

When Will My Model Fail? A Method For Detecting Patterns Of Forecasting Failures

Presenter: Matthew Bobea
Co-authors: Matthew Bobea;Galit Shmueli

AI-capable smart devices now collect and forecast human condition and behaviors. Forecasting failures can be costly especially during peak/dip periods. We introduce TreeAlert, a novel approach for detecting temporal patterns of extreme predictive failures when applying time series forecasting algorithms. TreeAlert empowers users to anticipate failure periods, and analysts to evaluate, compare, and improve forecasting algorithms. Our approach leverages a unique application of a machine learning algorithm–regression trees–to a model's forecast errors combined with Large Language Models (LLMs) to translate the tree rules into human-understandable statements. This results in easily interpretable and actionable rules for a human decision maker. We illustrate the method on two real datasets: heart-rate data arising from mobile health devices and bike sharing usage from bike docking stations.

Evaluating Forecasts When The Outcome Is Uncertain

Presenter: Christopher Ferro
Co-authors: Christopher Ferro;Harris Sop Nkuiate

The true values of the outcomes that we try to forecast are often unknown. For example, the targets of weather forecasts are usually measured with error, and the targets of economic forecasts are revised with each data vintage. If we ignore such observation errors then we risk obtaining biased views of forecast performance and even of favouring worse forecasters. This paper illustrates the effects of observation errors on forecast evaluation and shows how these effects can be corrected.The gold-standard measures of performance for probabilistic forecasts are proper scoring rules. If outcomes are measured with error then these scores are biased, the sizes of associated statistical tests are inaccurate, and forecasters may be ranked incorrectly. This paper extends the work of Ferro (2017) who used observation error statistics to form unbiased estimates of the scores that would be obtained if there were no observation error. Both additive and multiplicative observation errors are considered for a variety of forecast distributions and scoring rules. In cases where unbiased scores are unavailable, a first-order correction is used to obtain approximately unbiased scores. A simulation study demonstrates that statistical tests based on the error-corrected scores have high power and a data example illustrates the benefit of using the scores to compare competing forecasting models.Ferro CAT (2017) Measuring forecast performance in the presence of observation error. Quarterly Journal of the Royal Meteorological Society, 143, 2665–2676.

Statistical Inference For Score Decompositions

Presenter: Marius Puke
Co-authors: Marius Puke;Timo Dimitriadis

The decomposition, $S = MCB - DSC + UNC$, of a forecast's average loss, S, into the interpretable components miscalibration (MCB), discrimination (DSC), and uncertainty (UNC), was recently gaining interest in the forecasting literature. Although this decomposition enables thorough forecast evaluation, a notable limitation of this decomposition is that there are no objective measures for how large or small these components should be. In this project, we heal that problem by proposing asymptotic theory for these components. In particular, we suggest Diebold-Mariano-style testing procedures, which allow to assess the null hypotheses of equal miscalibration and equal discrimination among two rival forecasts, respectively. Finally, we apply the newly proposed tests to mean forecasting exercises in risk management and inflation forecasting and demonstrate that these new tests allow for more informative evaluation of forecasts.

Optimizing Inventory In Manufacturing Industry

Presenter: Arpit Jain
Co-authors: Arpit Jain

In today's rapidly evolving manufacturing landscape, the ability to forecast accurately is crucial for success. However, the value of forecasts extends beyond mere accuracy; it lies in their ability to guide inventory management decisions effectively. This discussion explores the vital role of inventory optimization in the manufacturing sector, emphasizing its importance in maximizing value for businesses. By aligning forecasts with inventory strategies, companies can strike a balance between supply and demand, minimizing excess inventory while ensuring product availability. Through the integration of advanced forecasting techniques, data analytics, and strategic decision-making, organizations can streamline operations, reduce costs, and enhance overall efficiency.This talk will present a case study where SAS implemented an inventory optimization strategy for a manufacturing company. The strategy involved retroactively simulating historical demand forecasts and optimizing stock levels. Additionally, we'll discuss the application of A/B testing in real-time to assess the effectiveness of the implemented strategy.In summary, effective inventory optimization is essential for navigating the complexities of modern manufacturing. By leveraging advanced tools and methodologies, companies can enhance their ability to respond to market demands efficiently, ultimately driving greater value and competitiveness in the industry.

Bridging The Xai Gap: Where Data Science Meets Business Expertise

Presenter: Michał Kurcewicz
Co-authors: Michał Kurcewicz;Marcin Wzgarda;Julia Zabojnikova

This paper discusses a recent forecast explainability project implemented at Mondelez. We begin with providing a high-level overview of the current, complex multi-step demand forecasting step process which utilizes both statistical time-series forecasting and AI techniques. Next, we discuss the motivations for starting the e(X)plainable AI project. Following this, we introduce the chosen methodology for forecast explainability, which blends machine learning tools and business knowledge, and provide real-world examples how explainable forecasts can help make informed business decisions. We conclude discussing both achievements and challenges encountered during the adoption of the XAI process.

How To Measure The Impact Of A Time Series On Global Accuracy

Presenter: Martin Schorter
Co-authors: Martin Schorter

In Retail and CPG, forecasters have hundreds or even thousands of time series in their portfolios. This multiplied by number of periods (history and forecast horizon) makes millions of data points flowing into their brains. How not to drown ? Indeed, forecast engines do the job at scale. They analyze data, they model, they forecast and they produce KPIs … tons of KPIs. Nevertheless, they are not self-correcting yet. They don't report to managers yet. And they don't know yet about impacting informations flying in the air and not in systems. Does this means there is still room for some human cognition out there ?Like anyone, professionnal conscience of forecasters push them into exhaustivity. Like anyone, forecasters have to deal with a complex reality and a demanding management for simple reporting. Like anyone, forecasters want to be impactful in what they do without destroying their work-life balance. How to overview exhaustively, report simply and embed impactfully ? That's the subject of this presentation.

Sparse Multiple Index Models For High-Dimensional Nonparametric Forecasting

Presenter: Nuwani Palihawadana
Co-authors: Nuwani Kodikara Palihawadana;Rob J Hyndman;Xiaoqian Wang

High-dimensionality is a common phenomenon in real-world forecasting. Oftentimes, forecasts are

contingent on a long history of predictors, while the relationships between some predictors and the response of interest exhibit complex nonlinear patterns. In such cases, a nonlinear "transfer function" model with additivity constraints is a conspicuous choice to mitigate the issue of curse of dimensionality. Particularly, nonparametric additive index models offer a way to greatly reduce the number of parameters to be estimated compared to general additive models. In this paper, we present a novel algorithm, Sparse Multiple Index (SMI) Modelling, designed to estimate high-dimensional nonparametric additive index models, while highlighting predictor grouping and optimal predictor selection. The SMI Modelling algorithm uses an iterative approach based on mixed integer programming to solve an L0-regularised nonlinear least squares optimisation problem with linear constraints. We demonstrate the performance of the proposed algorithm through a simple simulation, along with two empirical applications in forecasting heat exposure related daily mortality and daily solar intensity, respectively.

Benchmarking Of Imputation Methods In Multivariate Time Series Analysis For Forecasting

Presenter: Guillaume Saës
Co-authors: Guillaume Saës;Julien Roussel;Nicolas Brunel

In this article, we compare several multivariate time series imputation methods using machine learning in the case of Missing Completely at Random (MCAR) data. Analyzed methods range from simple methods like median and interpolation to more complex algorithms such as MICE (Multiple Imputation by Chained Equations), EM (Expectation-Maximization), RPCA (Robust Principal Component Analysis) or TS-DDPM (Temporal Diffusion Model for Time Series). Our primary goal is to examine the impact of these methods on the forecasting models' performance (MAE/WMAPE), and how it depends on the time series characteristics such as seasonality, dimensionality, trend, and with consecutivity of missing data. To achieve this, we employed a benchmarking procedure based on cross-validation where missingess is either introduced in the train set or in the test set. Our findings indicate a significant variation in the performance of the tested imputation methods, with a special mention for MICE, which proved to be a useful tool in MCAR scenarios for seasonal time series. This observation underscores the importance of carefully selecting an appropriate imputation method to improve forecast accuracy, especially in fields where understanding seasonal patterns is crucial. We hope this analysis provides useful insights into the efficiency of different imputation methods in the specific context of seasonal time series and assists practitioners and researchers in navigating the choice of imputation methods for their forecasting models.

Bayesian Causal Prediction: Multivariate Graphical Dynamic Models

Presenter: Luke Vrotsos
Co-authors: Luke Vrotsos;Mike West

We discuss a novel Bayesian modeling framework for causal inference and forecasting with multivariate time series. This approach exploits simultaneous graphical dynamic models (SGDLMs) that allow analysts to specify a dynamic model for each unit-specific time series, incorporating customized predictors and trends. These series-specific models are united in a graphical model that defines sparse cross-series relationships. Observation of post-treatment outcomes enables inference about the time-varying individual treatment effects on treated units and prediction of potential outcomes. Parallel implementation using variational methods and importance sampling enables scaling to large numbers of time series. We also contribute to the literature on interference effects in time series data by defining estimands that permit more general patterns of interference. Illustration comes from a multivariate macro-economic study on German reunification, demonstrating the importance of sequential dynamic modeling for causal time series, computational advantages, and evaluation of interference effects using this new class of Bayesian models.

Nowcasting Inflation At Quantiles: Causality From Commodities

Presenter: Sara Boni
Co-authors: Sara Boni;Massimiliano Caporin;Francesco Ravazzolo

This paper proposes a non-parametric test for Granger causality in quantiles to detect causality from a high-frequency driver to a low-frequency target. In an economic application, we examine Granger causality between inflation, as a low-frequency macroeconomic variable, and a selection of commodity futures, including gold, oil, and corn, as high-frequency financial variables. We find that logarithmic returns on given commodity futures are a prima facie cause of inflation at the lower quantiles of the distribution and marginally around the median. In the context of a nowcasting exercise, we find that incorporating commodity futures in the model with a polynomial function enhances short-term forecasting accuracy, leveraging timely data for more precise nowcasting of inflationary trends.

Forecasting Spatio-Temporal Data With Clustered Factor Models

Presenter: Raffaele Mattera
Co-authors: Raffaele Mattera;Philip Hans Franses

This paper introduces a new procedure for clustering and forecasting a large number of georeferenced time series characterized by a grouped factor structure. The method automatically estimates the number of global factors, the clustering structure and the number of clustered factors. Moreover, our procedure enhances spatial clustering so that the nature of clustered factors reflects both the similarity of the time series in the time domain and their spatial proximity. The method is applied for modelling and forecasting house prices growth in U.S. states in the time period 1975-2023. The results demonstrate that forecasting approaches incorporating both global and clustered factors provide more accuracy than models using the global factors only and those without factorial structure.

Commodity Currencies Revisited: The Role Of Global Commodity Price Uncertainty

Presenter: Athanasios Triantafyllou
Co-authors: Athanasios Triantafyllou;Laurent Ferrara;Theodora Bermpei;Aikaterini Karadimitropoulou

Exchange rates of commodity exporting countries, generally known as commodity currencies, are often considered to be driven by some specific commodity prices. In this paper, we show that the uncertainty common to a basket of commodity prices is also a significant driver of exchange rate dynamics for a panel of commodity exporting countries. In particular, an increase on global commodity price uncertainty leads to a short-run depreciation of the effective exchange rate in commodity currency countries, followed by a medium-term rebound. We document that this pattern is specific to commodity currencies and is not visible on benchmark currencies like the euro or the U.S. dollar, the latter acting as a typical safe haven currency. We refer to this pattern as the "commodity uncertainty currency" property.

Do High Frequency Text Data Help Forecast Crude Oil Prices? Mf-Var Vs. Midas

Presenter: Luigi Gifuni
Co-authors: Luigi Gifuni

This paper investigates the predictability of monthly real oil prices when daily and weekly text data are combined with oil market fundamentals. Text information is retrieved from 6,447,630 full oil-related articles featured in The Financial Times, Thomson Reuters and The Independent from 1982M1 to 2021M6. I show that models containing high-frequency financial and commodity variables do not yield significant improvements on the no-change forecast. In contrast, when text data are used along with commodity variables and oil market fundamentals, the preferred models reduce the MSPE by 18%. However, despite

this marginal improvement, gains are low. Indeed, the corresponding models with variables observed at homogeneous frequency, generate similar out-of-sample forecasts in terms of accuracy. I thus conclude that variables sampled at different frequencies do not significantly improve the predictability of monthly real oil prices. This is true for point and density forecasts.

An Evaluation Of Us Gdp Growth Nowcasts and Forecasts During Covid-19

Presenter: Adel Abou Jaber
Co-authors: Adel Abou Jaber

I evaluate a subset of mixed-data sampling (MIDAS) model adjustments proposed by Foroni et al (2022) to improve point nowcasts & forecasts of US GDP growth during the Covid-19 recession and recovery period (2020Q1-2022Q4). Using information from the 2008 Great Recession, they compute (1) similarity-adjusted forecasts by re-estimating the model over 2002Q1-2013Q4 and (2) intercept-adjusted forecasts by correcting forecasts based on forecast errors from the Great Recession. I show that both the unadjusted and adjusted models do not forecast better than a benchmark AR(2) model by using a non-parametric test for forecast accuracy (Döhrn, 2019), which demonstrates good power insmall samples. Regardless of the adjustment, the forecasts underpredict the magnitude of the Covid-19 shock and overpredict its persistence. Updating the estimation sample to 2022Q4 and excluding the Covid-19 observations does not change the MIDAS lag structure for nowcasts obtained using industrial production data, indicating that the Covid-19 shock had a transitory effect on GDP growth. These results support Schorfheide and Song's (2020) suggestion to exclude those observations from the estimation sample.

Time-Varying Parameter MIDAS Models: Application to Nowcasting US Real GDP

Presenter: Aubrey Poon
Co-authors: Aubrey Poon;Joshua Chan;Dan Zhu

We propose a novel time-varying parameter mixed-data sampling (TVP-MIDAS) framework. Specifically, we decompose the MIDAS coefficients into a scalar parameter representing the overall impact of high-frequency variables and a vector of weights, allowing both features to vary over time. Our study applies this framework to real-time forecasting for US real GDP. Our analysis demonstrates that the TVP-MIDAS model specifications produce superior point forecasts and are particularly effective in capturing left tail risk compared to their time-invariant counterparts. Additionally, our in-sample analysis reveals a significant negative trend in the influence of the National Financial Conditions Index (NFCI) on US real GDP, suggesting a progressively adverse correlation over time. Conversely, the impact of the yield curve slope on US real GDP exhibits minimal variation over time.

Midas-Qr With 2-Dimensional Structure

Presenter: Tibor Szendrei
Co-authors: Tibor Szendrei;Arnab Bhattacharjee;Mark Schaffer

Mixed frequency data has been shown to improve the performance of growth-at-risk models. Nevertheless, in the literature most of the research has focused on only imposing structure on the high-frequency lags when estimating MIDAS-QR models. In this paper we extend the framework by introducing structure on both the lag dimension and the quantile dimension. In this way we are able to shrink unnecessary quantile variation in the high-freuqncy variables. This leads to more gradual lag profiles in both dimensions compared to the MIDAS-QR and UMIDAS-QR. We show that this proposed method leads to further gains in nowcasting and forecasting on a pseudo-out-of-sample exercise on US data.

A Mixed-Frequency Factor Model For Nowcasting French Gdp

Presenter: Marie Bessec
Co-authors: Marie Bessec;Julien André

This article presents a new nowcasting model for quarterly real GDP growth in France, developed at the Banque de France. The model belongs to the class of targeted factor models and it is estimated using the mixed-frequency three-pass regression filter. It is estimated on a dataset of 60 monthly predictors, including the Banque de France survey variables. We evaluate three distinct models designed to nowcast the first release of French GDP growth during each month of the quarter. Using a pseudo-real-time evaluation, we find that our new model performs well when compared to simple benchmarks and existing tools at the Banque de France, especially during the first two months of the quarter. Furthermore, by extending the formulae for the contributions of the predictors to the mixed-frequency case, we analyze the contributions of different groups of variables on the demand and supply side to French GDP growth. We find that all groups of variables have exerted a negative influence on French real activity since the outbreak of the COVID-19 pandemic in 2020.

Macroeconomic Survey Forecasting In Times Of Crises

Presenter: Philip Letixerant
Co-authors: Philip Letixerant;Robinson Kruse-Becher

Survey-based forecasts like the Survey of Professional Forecasters (SPF) are generally accurate for a variety of target variables and forecast horizons. Due to their high forecast accuracy, they are serving as a common benchmark when evaluating competing forecast models. In times of crises, accurate economic forecasts are particularly difficult to obtain, while being of utmost importance for both the private and public sector. The previous research has demonstrated that survey forecasts as the SPF tend to outperform model-based forecasts during crises. Our aim is to further improve the survey forecasts by exploiting historical forecast errors during similar times of turmoil. We thoroughly investigate whether the MSFE can be reduced by implementing a similarity-based intercept adjustment. We do so by adjusting the predictions by forecast errors from previous similar periods, where latter are found by a matching algorithm. To this end, we rely on existing Nearest Neighbours approaches and enrich them in various directions. For a set of key macroeconomic variables our results demonstrate improvements in times of crises as well as more stable times.

Do Professional Forecasters' Phillips Curves Incorporate The Beliefs Of Others?

Presenter: Shixuan Wang
Co-authors: Shixuan Wang;Michael Clements

We apply functional data analysis to survey expectations data, and show that functional principal component analysis combined with functional regression analysis is a fruitful way of capturing the effects of others' forecasts on a respondent's inflation forecasts. We estimate forward-looking Phillips curves on each respondent's inflation and unemployment rate forecasts, and show that for nearly a half of the respondents the forecasts of others are important. The functional principal components of the cross-sectional distributions of forecasts are shown to capture characteristics other than the mean or consensus forecast, and include forecaster disagreement.

Mixtures Of Predictive Densities For Colombian Expected Inflation

Presenter: Hector Zarate
Co-authors: Hector Zarate

Surveys of professional forecasters collect histogram-based density forecasts as a source to elicit expectations about inflation, a crucial factor that plays a role in decision-making impacting the financial

and macroeconomic sectors. Density forecast combinations provide a basis to characterize the shape and uncertainty around the entire probability distribution of expected inflation. We use data from the quarterly surveys of inflation expectations from 2016.01 to 2023.04 applied to Colombian experts at various horizons to assemble the data panel to compute measures of disagreement and uncertainty at the aggregate and individual levels. We construct mixtures of expected inflation densities based on shrinkage and regularization techniques that outperform individual forecasters. However, the Bayesian predictive synthesis framework allows nonlinear mixtures and time-varying weights in this small-time coverage empirical application.

Bimodality In The Probabilistic Expectations

Presenter: Xuguang Simon Sheng
Co-authors: Xuguang Simon Sheng;Antar Diallo;Brent Meyer;Emil Mihaylov

We document the presence of bimodality in the subjective probabilistic expectations of both business executives and professional forecasters. The share of probabilistic expectations that exhibit bimodality varies by type of respondent (business decision-makers versus professional forecasters), whether the forecast object is an aggregate or own-firm quantity, and by the form of the survey question itself, but we find it to be present in all the probabilistic expectations we investigate. We utilize a novel, flexible technique to parametrically estimate uncertainty in the presence of bimodality called the bimodal asymmetric power normal distribution. We apply this novel technique in investigating the uncertainty and characteristics of bimodal responses, namely their persistence, cyclicality, and state-dependence.

On-Line Conformalized Neural Networks Ensembles For Probabilistic Forecasting Of Day-Ahead Electricity Prices

Presenter: Alessandro Brusaferri
Co-authors: Alessandro Brusaferri;Andrea Ballarino;Luigi Grossi;Fabrizio Laurini

Probabilistic electricity price forecasting (PEPF) is subject of growing interest, following the demand for proper quantification of prediction uncertainty, to support the operation in complex power markets with increasing share of renewable generation. Ensembles of distributional neural networks (NN) parameterizing flexible Johnson's SU distributions have been recently proposed for this purpose, outperforming state of the art PEPF benchmarks including both conventional Gaussian forms and the widely applied quantile regression on NN ensembles (QRA). Despite the relevant improvements, it has been observed that such models still lack the required calibration capabilities, failing to pass the coverage tests at various hours on the prediction horizon. Conformal Prediction (CP) provides a principled framework to attain distribution free finite sample marginal calibration guarantees. However, CP has still attracted minor attention in the PEPF field. In this work, we extend the NN ensembles based PEPF methods by introducing conformal inference based techniques within an on-line recalibration procedure. An asymmetric CP formulation is deployed through a daily recalibrated multi-horizon time series setup to support flexible stepwise compensations of each upper/lower quantiles. The target prediction bands over the different coverage levels are then dynamically adjusted by tracking the related quantiles of the conformity score sequences through a Conformal Proportional-Integral technique. To estimate the quantiles to be calibrated, we explore both a quantile regression setup and samples from different parameterized distributional NNs to assess their potential impact on both sample efficiency and reliability. A uniform vincentization technique is exploited for ensembles aggregation. Moreover, a post-hoc sorting operator has been included before combination to achieve quantile non-crossing. The experiments are executed on the German market dataset as well as on the different regional bidding zones constituting the Italian day-ahead markets, providing a compelling setup for testing under heterogeneous conditions. A comparison against state of the art benchmarks is performed, including QRA, quantile regression NNs, distributional NNs, as well as conventional absolute score and normalized score based CP settings, showing the capability of the proposed approach to achieve day-ahead forecasts with improved hourly coverage and stable probabilistic scores.

Electricity Price Modeling And Forecasting With A Markov Switching Regime Model

Presenter: Frederic LANTZ
Co-authors: Frédéric Lantz;Kallia Charron

Energy transition policies in Europe focus on two main levers: increasing the share of renewables in the electricity mix and electrification, specifically electrification of transportation. Those dynamics will increase the constraints on the electricity system, as the transmission system operator will have to balance an increased electricity consumption with a more intermittent production. On the electricity market, demand response influences the equilibrium price determined by the merit order mechanism: power units are run from the lowest to the highest marginal cost. The priority of renewable energies (wind and solar units) sometimes leads to negative prices. It is cheaper to maintain the production of some conventional units rather than to make them stop and then restart (ramp-up time is an important problem for such unit). Subsequently, load shedding during hours of peak residual load (load minus the wind and solar production) can result in lowering the price, while load shifting to a time when renewables production is at its highest can limit negative pricing. In this context, this article aims at measuring this merit order effect of demand response and renewable supply on electricity market prices. Using an hourly data set over one year in France (2019), we point out the strong difference between peak and off-peak hours through a statistical analysis of the market prices, electricity supply and demand. Then we estimate a Markov switching model with two regimes to project the day-ahead hourly price on the French market. This enables us to estimate the variation in price induced by a variation of the load and a variation of the renewable supply in the electricity mix. A price transformation has been required to consider negative electricity prices with the hyperbolic sine function. We then run this model to assess how an increased share of renewables in the electricity mix can strongly affect this market value according to the daily period peak/off peak which is reflected in the Markov Switching model estimation.

Real-Time Electricity Price Forecasting With Attnet

Presenter: Kristen Schell
Co-authors: Kristen Schell;Haolin Yang

This work tackles the challenge of forecasting electricity price in the real-time market (RTM), which exists to correct imbalances in supply and demand that can be created by inaccurate bids in the prior day-ahead market (DAM). Severe imbalances in the DAM can result in high volatility in the RTM, where prices range from $30- to $4,000/MWh. The operating time interval of the RTM is five minutes, which results in high-frequency price fluctuations. These market characteristics result in high-noise, high-nonlinearity, and high-uncertainty, making forecasting for the RTM a challenge. Inaccurate forecasts pose risks and difficulties for several power system activities, including strategic bidding, generator rescheduling and renewable energy curtailment. Recently, deep learning (DL) has garnered interest as a promising approach to address the challenges presented. Nevertheless, significant gaps in the modeling frameworks exist. Currently, the majority of models prioritize hourly or longer time periods, with only a limited number of models designed for high-frequency forecasting. This study introduces a novel deep-learning neural network called ATTnet, which utilizes an attention mechanism to accurately anticipate electricity prices in the RTM. Additionally, it proposes an explainable model pipeline to enhance the interpretability of the predictions. The network is composed of a compact structure that includes a 5-head attention mechanism and gated recurrent units. These components have been specifically designed to capture the temporal relationships present in the market data. The importance of input features is studied in two manners: temporally through the attention scores derived from the input sequences, and globally through the feature Shapley values. The most significant factors in real-time electricity price forecasts are historical prices, temperature, hour, and zonal load. The proposed deep learning architecture underwent testing using real-time pricing profiles obtained from eight generators operating inside the New York Independent System Operator (NYISO) network. The model demonstrates a significant improvement in performance, with a 21% reduction in Mean Absolute Error (MAE) and a 22% reduction in Mean Absolute Percentage Error (MAPE) compared to the current

state-of-the-art approaches.

System Length And Price Forecasting In British Balancing Mechanism

Presenter: Klimis Stylpnopoulos
Co-authors: Klimis Stylpnopoulos;Jethro Browell;Wayne Jones;Ciaran Gilbert;Janine Illian

The electricity market in Great Britain has undergone many transformations driven by the need to integrate renewable energy sources and maintain security of supply. Integrating weather-dependent renewable energy into the electricity sector introduces uncertainty in short-term electricity markets. Consequently, stakeholders in the market need accurate forecasts to inform bidding strategies and mitigate risk. A continuous balancing mechanism provides a price signal to incentivise market participants to optimise their portfolios, as the Electricity System Operator takes corrective actions to address any energy imbalances.The research community has shown increased interest in forecasting electricity price imbalances due to the escalating costs and uncertainties inherent in real-time balancing. Imbalance price forecasting presents significant challenges owing to volatility, heteroscedasticity, and system length, which introduces distinct regimes depending on whether the market is in surplus of deficit. In our study, we utilise data from the GB electricity market covering multiple years of data from GB Balancing Mechanism and develop a method for predicting the imbalance price (relative to the known day-ahead price) two settlement periods ahead of delivery, on a rolling half-hourly basis. Our modelling strategy produces probabilistic forecasts in order to quantify forecast uncertainty. We compare two frameworks: the first is to forecast quantiles of the imbalance price and construct non-parametric predictive densities, and the second predicts the imbalance price as a mixture of parametric densities weighted by the predicted probability of the system being in deficit or surplus.We assess the effectiveness of these forecasting frameworks by evaluating its accuracy and reliability of our probabilistic forecasting results, and incorporate a range of covariates published by the GB system operator in real-time. Significantly, approaches that forecast the imbalance price as a mixture of weighted predictive densities exhibit better accuracy and reliability compared to methods that solely predict quantiles. Additionally, parametric methods that predict both the location and scale parameters of mixture components demonstrate consistent improvement over both non-parametric methods and parametric methods that assume a fixed variance.

Solnet: Open-Source Deep Learning Models For Photovoltaic Power Forecasting Across The Globe

Presenter: Joris Depoortere
Co-authors: Joris Depoortere;Hussain Kazmi;Johan Driesen;Johan Suykens

Deep learning models have gained prominence in the field of solar forecasting. One drawback of these models is that they require a lot of high-quality data. This is often not feasible, especially in solar power forecasting due to poor measurement infrastructure in legacy systems and the rapid build-up of new systems. This paper proposes SolNet: a novel, general-purpose solar power forecaster, which addresses these challenges by using transfer learning from abundant synthetic data generated from PVGIS. Using actual production data from hundreds of sites in Australia, Belgium and the Netherlands, we show that SolNet improves forecasting performance over data scarce settings as well as simple benchmark models. We observe this effect to be the strongest when only limited data is available, and show that seasonal patterns have a major impact on the results. The study also quantifies synthetic data requirements to achieve these improvements, as well as the effect of misspecifications

Quantile Additive Models For Forecasting Solar Electricity Production Using Gridded Numerical Weather Forecasts

Presenter: Ben Griffiths
Co-authors: Ben Griffiths;Matteo Fasiolo;Yannig Goude;Simon Wood;Abdelhadi El Yazidi

We are interested in probabilistic forecasting of regional solar farm production in Aquitaine, a region in South-Western France. In particular, we will show how probabilistic predictions can be obtained using quantile additive models (QGAMs), an extension of quantile regression which integrates the flexible additive structure of generalized additive models (GAMs). To tackle the complexity of the model and the size of the data considered in this application, we develop novel big data methods for fitting QGAM. The new methods are based on a covariate discretisation approach which considerably speeds up computations, while having little or no effect on the accuracy of the predictions.We compare the probabilistic predictive performance of several GAMs, QGAMs and GAMs for location, scale and shape (GAMLSS) models, the latter being an extension of GAMs beyond mean modelling. Given that solar production is strongly dependent on weather, the gridded output of a numerical weather model is an important input of all the models considered in this work. While the simplest approach to integrate a gridded weather forecast in an additive model is to reduce it to a scalar summary (e.g. the spatial average of solar irradiation over Aquitaine at a given time t), we show how functional smooth effects can be used to integrate raw gridded forecasts directly in a (Q)GAM model.The GAM and GAMLSS model considered in this work have been fitted using the mgcv R package, while the methods for building and fitting (big data) QGAMs are provided by the qgam R package.

Forecasting Renewable Energy Production In Pv And Wind Power Plants - A Stochastic Simulation Approach

Presenter: Gustavo Melo
Co-authors: Gustavo Melo;Fernando Oliveira;Paula Maçaira;Erick Meira

The increased participation of Variable Renewable Energy sources (VREs) in electrical matrices worldwide brings several challenges to the planning and operation of power systems, mainly due to the intermittency of VREs. In this context, the complementary effects between different intermittent sources have become increasingly important in the literature since such properties result in a combined power with less variability and intermittency, reducing the demand for storage and smoothing the operation of power systems. This work proposes a forecasting approach for renewable energy production based on a stochastic simulation framework that explores several complementary effects between wind and photovoltaic solar (PV) sources. The approach considers: (i) calculating energy production states (clusters) to simultaneously represent multiple renewable energy sources using clustering algorithms; (ii) deriving state transition probability matrices from historical data within the defined clusters; (iii) obtaining energy scenarios by applying Monte Carlo simulation from a uniform [0,1] distribution to randomly select states in the transition probability matrices. The proposed methodology is applied to two case studies in the Northeast region of Brazil, which has the highest potential for VREs production in the country. In the first case, we simulate the production of a hybrid plant, comprising both a wind and a PV plant, aiming to exploit the temporal complementary effects between the sources. In the other case, we simulate the combined production of two plants from different locations, one being a wind complex and the other a PV plant, thereby aiming to explore both temporal and spatial complementary effects between the VREs. The synthetic time series are subsequently assessed using statistical methods, such as probability distribution functions, descriptive statistics, and autocorrelation functions. The results indicate good adherence of the generated scenarios to the most stylized facts of the historical series, demonstrating the effectiveness of the proposed methodology for both simulating and forecasting renewable energy scenarios.

Day-Ahead Solar Irradiation Forecast Techniques And Outcomes

Presenter: Rodrigo Huber
Co-authors: Rodrigo Mendes;Laura Ferreira;Fernanda Araujo Baião;Reinaldo Castro Souza

This research focuses on stating the major differences in forecasting outcomes for a day-ahead solar irradiation forecast using three different methods. Three forecasting strategies were used in this research, their respective benefits, shortcomings, and pitfalls are compared as well as the pre-processing work and time demanded for each method. The results were analyzed through its errors and the true values observed. The three techniques used were a Climate-meteorological forecast method, a multi-season ARIMA

forecast method, and a Machine Learning method.This research paper focuses on the forecast of solar irradiation using machine-learning techniques models (XGBoost and Random Forest), a multi-seasonal ARIMA model, and a novel Climate-meteorological forecast method based on temperature measurements and its correlation between changes in clear-sky and overcast day index.This research aims to enhance the accuracy and efficiency of solar energy forecasting for the National Interconnected System of Brazil. The dataset used in this research was installed in Miranda do Norte, Brazil by the research team and it contains atmospheric pressure, ambient temperature, relative humidity, and solar irradiation measurements with a temporal granularity of one minute for 4 years straight.The methodology involved in this research includes preprocessing the dataset, the exploration of temporal patterns through decomposition techniques which facilitates the analysis and implementation of a multiple seasonality ARIMA model, the implementation of machine-learning models (XGBoost and Random Forest) optimized through Grid, Bayes, and Random Search and the implementation of the Climate-meteorological forecast method (used as a benchmark).Furthermore, the research highlights the significance of incorporating exogenous variables, particularly relative humidity, to improve forecasting accuracy. The comparison of model performances based on metrics such as (RMSE) and implementation work and time demanded reveals the strengths and limitations of each approach.Overall, the research gives valuable insights into the application of machine learning algorithms for solar energy forecasting, emphasizing the importance of model selection based on precision and simplicity. The findings elucidate the potential of machine learning models in enhancing solar irradiation forecasting accuracy, which is crucial for optimizing energy management and decision-making within the energy sector.

Investor Preferences In Xai-Enhanced Algorithmic Trading: An Empirical Lstm Study

Presenter: Stanisław Łaniewski
Co-authors: Stanisław Łaniewski;Robert Ślepaczuk

Exploring the intersection of machine learning and experimental economics, this study presents an innovative approach to understanding investor preferences in algorithmic trading. We develop three tailored investment strategies on main US stock indices (S&P, NASDAQ, Russell 2000) using Long Short-Term Memory (LSTM) networks and extensive hyperparameter optimization (such as but not limited to layers/nodes, loss function, learning rate, lookback). Each model has different characteristic: a complex high-performance model, a simpler low-volatility model, and a model enhanced with Explainable AI (XAI) techniques, particularly Shapley values and feature permutation importance. These strategies cater to distinct investor priorities: maximizing returns, minimizing risk, and providing transparent decision-making processes. Central to our research is an empirical survey, where participants evaluate and select their preferred investment approach from the three LSTM-derived strategies. The models are quantitatively assessed, presenting historic performance graphs and key metrics such as average profits, volatility, costs of each strategy. By analyzing preferences for return, risk aversion, or the clarity offered by XAI, we seek to uncover insights into the investor psyche in the digital age of finance. We find that WTP for XAI methods to be significant and positive and that the effect is stronger among people with higher risk-aversion. Our results also indicate that enhanced transparency through XAI increases trust in the model.Our findings deliver a dual impact: advancing the application of LSTM models in financial trading and enriching our understanding of investor behavior towards technology-driven investment tools. This research bridges the gap between algorithmic strategy development and user-centric financial product design, highlighting the need for XAI in the age of AI.

Transformer Deep Learning Models For Financial Assets Returns Prediction

Presenter: Jakub Michańków
Co-authors: Jakub Michańków;Paweł Sakowski;Robert Ślepaczuk

This research aims to implement a transformer deep learning model for predicting financial asset returns. The predictions generated by the model are used to generate signals employed in algorithmic

investment strategies. A comparative analysis is conducted between the performance of transformer networks and a previous version of a similar model known as LSTM networks. The research focuses on examining the robustness of the model by employing logarithmic returns from diverse financial asset classes. To mitigate data overfitting concerns, a strict backtesting strategy based on walk-forward validation is employed. Models are tested on log returns of S&P500, BTC, and GLD from 2004-01-02 to 2020-03-29. The output of the model is a single number predicting the next return value. Based on the sign of the predicted return value, signals of -1, 0, 1 (buy, hold, sale) are assigned, depending on the strategy. The adaptation of basic transformer model architecture to time series forecasting is successfully demonstrated. The transformer model outperforms B&H and LSTM based strategies for two assets: S&P 500 index and Bitcoin.

Mean Absolute Directional Loss: A New Loss Function For Machine Learning Models In Algorithmic Investment Strategies

Presenter: Pawel Sakowski
Co-authors: Paweł Sakowski;Robert Ślepaczuk;Jakub Michańków

We investigate the problem of choosing an adequate loss function in the optimization of machine learning models used in the forecasting of financial time series for the purpose of building algorithmic investment strategies (AIS). We propose the Mean Absolute Directional Loss (MADL) function, solving important problems of the classical forecast error functions in extracting information from forecasts to create efficient buy/sell signals in algorithmic investment strategies. Based on the data from two different asset classes (cryptocurrencies: Bitcoin and commodities: Crude Oil), we show that the new loss function allows to select better hyperparameters for the LSTM model and to obtain more efficient investment strategies,with regard to risk-adjusted return metrics on the out-of-sample data.

Supervised Autoencoder Mlp For Financial Time Series Forecasting

Presenter: Robert Ślepaczuk
Co-authors: Robert Slepaczuk;Bartosz Bieganowski

This paper investigates the enhancement of financial time series forecasting with the use of neural networks through supervised autoencoders, aiming to improve investment strategy performance. It specifically examines the impact of noise augmentation and triple barrier labeling on risk-adjusted returns using the Sharpe and Information Ratios. The study focuses on high-frequency data (5-, 15-, and 30 minutes) of the S&P 500 index, EUR/USD, and BTC/USD as the traded assets from January 1, 2010, to April 30, 2022. To address our research questions, we employ empirical methods, developing trading strategies based on various Supervised Autoencoder - Multi Layer Perceptron (SAE-MLP) model architectures. These models are tested using data collected at one-minute intervals. Our approach involves using the SAE-MLP model for price prediction, employing a walk-forward method alongside combinatorial purged cross-validation. This process begins with hyperparameters sourced from existing literature, followed by fine-tuning through hyperparameter optimization. We anticipate that models with algorithm-selected hyperparameters will show superior performance. Our methodology also includes transforming the return estimation problem from a regression model to a classification one, focusing on predicting price direction rather than the exact price. Through rigorous testing and sensitivity analysis, we demonstrated that employing supervised autoencoders significantly enhanced the performance metrics of the strategy. Our analysis suggested an optimal balance between the level of Gaussian noise added to the features and the size of the autoencoder's bottleneck. However, an increase in the level of noise and bottleneck size beyond certain thresholds led to a decrease in performance, presumably due to overfitting and the inclusion of more noise than signal, respectively. These findings underline the need for careful calibration of these parameters to ensure the most effective utilization of supervised autoencoders in this context. At the end, we conduct a sensitivity analysis to examine the robustness of our findings under different assumptions. The results of this study have substantial policy implications, suggesting that financial institutions and regulators could leverage techniques presented to enhance market stability and investor protection, while also encouraging more

informed and strategic investment approaches in various financial sectors.

Option Surface Prediction: Good Old Econometrics Rock

Presenter: Arnaud Dufays
Co-authors: Arnaud Dufays;Jeroen Rombouts

In this study, we address the prediction of implied volatility surfaces. Recent literature indicates that machine learning models, especially artificial neural networks (ANNs), outperform traditional regression and affine option pricing models. A common assumption for these comparisons is that parameter estimates follow a random walk, suggesting that the estimated surface for one day will persist unchanged in subsequent days. Using standard econometric tools, we challenge this assumption and demonstrate that incorporating traditional volatility models and econometric methods, such as change-point analysis, can enhance the prediction accuracy of all examined models (ANNs, regression, and option pricing). Our findings further reveal that the regression approach, when complemented with these tools, can exceed the performance of a purely machine learning-based approach.

Trending Time-Varying Coefficient Regression Models: Estimation And Prediction By Local Linear Smoothers Using Asymmetric Kernels

Presenter: Masayuki Hirukawa
Co-authors: Masayuki Hirukawa

In this paper, trending time-varying coefficient regression models are investigated. Time-varying coefficients are estimated nonparametrically by local linear ("LL") regression smoothing. Because the domain of varying coefficients is [0,1], nonstandard, asymmetric kernels ("AKs") that are free of boundary bias are employed. Among all such kernels, particular focuses are on the beta and gamma kernels due to their popularity in empirical studies in economics and finance. Convergence properties of AK-LL estimators for varying coefficients at a fixed design point and in the vicinity of 1, including their bias and variance approximations and asymptotic normality, are explored. Their finite-sample properties, as well as effectiveness of an implementation method, are examined via Monte Carlo simulations. Implications for prediction are also considered, in combination with long-run variance estimation for asymptotic variances of AK-LL estimators.

Change-Point Detection In Time Series Using Mixed Integer Programming

Presenter: Anton Skrobotov
Co-authors: Anton Skrobotov;Artem Prokhorov;Peter Radchenko;Alexander Semenov

We use recent advances in mixed integer optimization (MIO) methods to develop a framework for the identification and estimation of structural breaks in time series regressions. The framework requires a transformation of the the classical structural break detection problem into an Mixed Integer Quadratic Programming problem. We restate the problem as an l_0 penalized regression and compare it to the popular l_1 penalized regression (LASSO). MIO is capable of finding provably optimal solutions to this problem using a well-known optimization solver. The framework allows to determine the unknown number of structural breaks as well as the break locations. In addition to that, we demonstrate how to accommodate a specific number of breaks, or a minimal required number of breaks. We demonstrate the effectiveness of our approach through extensive numerical experiments obtaining much more accurate estimation of the number of breaks in comparison to the popular methods Bai and Perron (1998) and QIan and Su (2016).

Modeling Higher Moments And Risk Premiums For SandP 500 Returns

Presenter: Jeroen Rombouts
Co-authors: Jeroen Rombouts

Using joint estimation on a large sample of index option prices and the underlying returns, we study how multifactor models capture time-series and cross-sectional patterns in option prices through improved modeling of the dynamics of the first four moments of the return distribution. Including a second and especially a third stochastic volatility factor greatly improves option fit, and the resulting time series of skewness and kurtosis better match non-parametric benchmarks. The third volatility factor is critical in generating larger and more variable skewness and kurtosis risk premiums. Return jumps provide more modest improvements in option fit and a higher equity risk premium, but their impact on higher moment risk premiums is small. All models we investigate struggle to match the unconditional term structure of risk-neutral skewness and kurtosis.

Can Ai Predict The Price Of Gold?

Presenter: Angi Roesch
Co-authors: Angi Roesch;Harald Schmidbauer

In public opinion, AI is touted a paradigm-shifting technology challenging the capabilities of professionals, also in the field of forecasting. Naturally, it is tempting, for example, to ask AI, be it ChatGPT-4 or Google Bard, to predict the gold price.In our study we adopt neural network and model-based forecasting approaches and contrast them to AI gold price predictions on a common ground of information. The AI forecast provided used to be as sophisticated as is the manner of questioning, respectively, the amount of information given.

Leveraging Economic Indicators, Text And Open-Source Data To Forecast Cyberattacks

Presenter: Harald Schmidbauer
Co-authors: Harald Schmidbauer;Kamil Mizgier;Claudio Antonini

Cyberattacks pose a significant and growing threat to national security and economic well-being of organizations worldwide. However, accurately forecasting such attacks remains a challenge due to the limited availability of relevant data. Most companies and governments are hesitant to publicly disclose details of cyberattacks they experience.To address this shortcoming, our study utilizes the University of Maryland's Center for International and Security Studies (CISSM) Cyber Attacks Database as the primary data source. This rich dataset offers insights into various aspects of cyberattacks, including threat actors, motives, targets, and economic impacts. Among others, we hypothesize that incorporating economic indicators like economic growth, interest rates, inflation, and unemployment rates can act as valuable covariates alongside the autoregressive components of the models. To reduce the forecasting errors, in addition to quantitative variables, we add 'text-related' variables from Google Trends.By evaluating the suitability of various models, including neural networks, we aim to contribute to the development of more robust and informative forecasting methods in the cybersecurity domain, an area that is currently under-researched.

Modelling Artificial- And Human-Intelligence Versus Consciousness

Presenter: Hans Georg Zimmermann
Co-authors: Hans Georg Zimmermann

Artificial Intelligence (AI) and Human Intelligence (HI) try to solve similar problems: 'Perception, Understanding (=modelling), Action´. Obviously, the AI side, realizable on a computer, is accessible to mathematical modelling.In a second section we extend the AI mathematics to its human counterpart: the similarity of the tasks suggests similar approaches, e.g. the Mind can be seen as a simulation model. In detail we will find differences: In AI, Perception and Action are typically supervised learning tasks, while in HI they have to be described as unsupervised tasks.But can we extend the above guideline to a description

of Consciousness? In this talk I will explain why my answer is NO! Analog to the AI, HI tasks: 'Perception, Understanding, Action´ we will start with a conception of the Threeness: 'Insight, Awareness, Creativity ´. The non-emergent character of Consciousness limits the application of mathematics.Ongoing, we will discuss the relation of the pairs (Consciousness / Self), (Mind / Ego), (World / Body). We will work out a consistent description of a dualistic view between the above entities. Their alignment can be explained as the result of an identification between Self, Ego and Body together with learning processes without the need of a physical connection. The talk combines views from Computer Science, Mathematics and Philosophy. The gap between (AI) and Humans is not seen in the description of Intelligence but between Intelligence and Consciousness.

Successful And Unsuccesful Methods To Forecast With Large Language Models

Presenter: Claudio Antonini
Co-authors: Claudio Antonini

Due to the sudden availability of Large Language Models (LLMs), various applications are being considered by practitioners. Specialists go ahead with confidence, relying on the promise of potential outcomes while ignoring assumptions, particularities of the process, or poor demonstrated results.Dubious results are partly caused by fundamental differences with more traditional formal methods, such as:(1) Lack of a rigorous methodology to determine the forecasting suitability of events by a particular LLM, something that does not happen with more established methods for quantitative forecasting, for example, with regressions, where one can test if a time-series satisfies certain conditions before applying the method. (2) The hallucinations/delusions of the LLM answers, actual 'textual outliers' which require special monitoring or handlings not found in time-series—sometimes, arbitrary—that cannot be compared to well-studied processes applied to missing data or quantitative outliers.(3) Lack of repeatability of computed results (due to continuous and unannounced proprietary modifications done to the models, or the lack of features to create repeated runs).Still, after carefully consideration of the characteristics of the events and the particularities of the model, using LLMs to forecast is possible, and presents opportunities unavailable in conventional quantitative models.We will present two forecasting applications of LLMs. The first one, proposed by a Federal Reserve Bank, forecasts inflation. We will show some features of the LLM which make the methodology and the results questionable – different conclusions would be reached if a skillful human were analyzing the data. A second example, this time successful, allows to forecast the likelihood and impact of global risk events and compares those effects to surveys conducted on about a thousand professionals by the World Economic Forum (WEF). In this case, we will discuss the lack of continuity in results of the WEF along with their significant efforts to gather and sort the data, which makes the use of LLMs appealing.

Locating The Start Of Trend Breaks By Indicator Saturation

Presenter: Jurgen Doornik
Co-authors: Jurgen Doornik;Jennifer Castle;David Hendry

Sudden rapid trend changes, including tipping points, can be hard to discern initially and difficult to model in their early stages. We first investigate using a sequence of impulse indicators to characterize a succession of large same-sign 1-step-ahead forecast errors experienced as the forecast origin advances, indicating a break. We test replacing significant indicators by a new linear or log-linear trend, as well as against each other and step indicators. We also compare detection by trend indicator saturation. An application to the rapid increase in global sea level after 2011 illustrates these possibilities. Simulations of the first approach for detecting and forecasting sudden trend changes confirm its feasibility even after just 2 or 3 periods.

Predicting Hurricane Damages Now And In The Future

Presenter: Andrew Martinez
Co-authors: Andrew Martinez

I formulate a model of hurricane damages using machine learning methods. The final model includeskey measures of socioeconomic vulnerabilities and natural hazards that matter most for hurricane damages. Iuse the resulting model to predict damages in real-time both prior to landfall using weatherforecasts and immediately after landfall. The model nowcasts perform well against other catastrophemodels, especially in the first few days after landfall before high quality insurance claims are available. Finally, Iuse the model to predict hurricane damages in the distant future under alternative climate and economic scenarios.

Building Open-Source Empirical Models To Forecast Carbon Emissions

Presenter: Moritz Schwarz
Co-authors: Moritz Schwarz;Jonas Kurle;Andrew Martinez;Felix Pretis

Formulating and implementing climate policy requires a detailed understanding of likely pathways of future carbon emissions. However, most existing tools to provide forecasts for such emissions are limited. Projections are predominantly made for the long-term, lacking the necessary detail within the typical policy horizon of one to five years. Most short-term forecasting models are closed-source 'black-boxes' that do not reveal in detail how the obtained forecasts are generated. In this paper, we develop a generalised framework for modelling climate and environmental policies within an empirical macro-econometric forecasting approach. The resulting "Aggregate Model"' is an open-source econometric model builder that provides a platform to generate empirical forecasts of sectoral carbon emissions as well as the wider macroeconomy. The underlying estimated models are based on dynamic time series regressions allowing for model selection, outliers, in-sample structural breaks, and automatic forecast evaluation. We present an illustrative example providing the first short run sectoral greenhouse gas emission forecasts for Austria. The resulting model builder is open source, easily portable across countries, can be updated in real time when new data is released, and aims to improve transparency and flexibility in policy forecasts.

Forecasting Climate Change Using A Multivariate Cointegrated System

Presenter: Jennifer Castle
Co-authors: Jennifer Castle;David Hendry;Xinjia Hu

Forecasts of key climate variables are produced using a multivariate cointegrated VAR. The system augments carbon dioxide with methane and nitrous oxide in parts per million of atmospheric CO2 equivalents to model the effects of all major greenhouse gases (GHGs). Encompassing tests are used to establish the system model dominance over a model with radiative forcing. The model is augmented for natural variation including El Niño–Southern Oscillation and we include ice loss from cointegration between Greenland and Antarctic ice sheets plus changing albedo from the Arctic Ocean. The debate over the order of integration of the system (I(1) versus I(2) and multi-cointegration) plays little role in the forecast performance. Finally, we test the system for evidence of tipping points using a short sequence of impulse indicators to measure if large systematic one-sided 1-step-ahead forecast errors is occurring as the forecast origin advances. We propose replacing those indicators by the average of a broken linear and log-linear trend to forecast the evolution of a tipping point.

Evaluating C++ Computation Times With 'Rcpptimer'

Presenter: Jonathan Berrisch
Co-authors: Jonathan Berrisch

The ever-increasing number of complex forecasting methods strengthens the need for accurate evaluation. While these methods often outperform conventional forecasting approaches, they typically require

significantly more computation time. Thus, assessing both forecast performance (e.g., RMSE, CRPS) and the associated computational resources has become increasingly vital.Many implementations of these methods leverage low-level programming languages like C++ to minimize computation time. However, benchmarking low-level code presents challenges, particularly when accessed through higher-level languages such as R (via Rcpp) or Python (via pybind11). Moreover, utilizing parallelism with OpenMP further complicates benchmarking due to limited available benchmarking packages.We developed 'cpptimer', a straightforward tic-toc timer class for benchmarking C++ code to address this. Unlike existing solutions, 'cpptimer' supports overlapping timers and OpenMP parallelism. It also calculates summary statistics when benchmarking the same code segment multiple times. Being a header-only library, 'cpptimer' is easily bindable to higher-level languages. For instance, 'rcpptimer', an R package utilizing Rcpp, provides seamless timing of C++ code from R, automatically passing results into the user's R environment.In a demonstration, I will showcase the versatility of 'rcpptimer' across various Rcpp projects and explore its implementation intricacies. Key features such as OpenMP parallelism and automatic result return to R will be highlighted. Furthermore, I will discuss the rationale behind separating the core of 'rcpptimer' into the standalone project 'cpptimer'.

Light-Weight Pre-Trained Mixer Models For Effective Transfer Learning In Multivariate Time Series Forecasting

Presenter: Arindam Jati
Co-authors: Arindam Jati;Vijay Ekambaram;Pankaj Dayama;Nam H. Nguyen;Jayant Kalagnanam

Recently, Transformers have gained popularity in time series forecasting for their ability to capture long-sequence interactions. However, their high memory and computing requirements pose critical bottlenecks in enabling it for practical applications in resource-constrained settings. Additionally, the challenge of capturing cross-correlation across "variates" (also known as "channels") using traditional transformer architectures exacerbates GPU limitations, as we need to apply the attention mechanism across all channels, leading to very high memory consumption. In response to these challenges, we introduce lightweight time series models leveraging MLPMixer architectures. MLPMixer architectures circumvent the resource demands of the attention mechanism by employing dimension-aware MLP-based mixing modules, enabling effective capture of short- and long-term interactions within and across channels. Our contribution, PatchTSMixer, recently accepted at KDD 2023 and open-sourced via HuggingFace, surpasses existing transformer-based benchmarks while offering substantial resource savings of 2-3X.Furthermore, we extend our efforts by releasing various Mixer-style pre-trained models, facilitating effective zero/few-shot forecasting in data-constrained scenarios. This showcases the effectiveness of these models in transfer learning scenarios, especially when there is limited or no data available in the target domain.In this presentation, we aim to highlight the advantages of Mixer-style forecasting models over traditional transformer architectures, demonstrate their utility across diverse industrial settings using transfer learning, and share best practices for deploying our models in resource-constrained environments.

Sktime - A Collaborative Research Software Framework And Marketplace For Cutting-Edge Time Series Forecasting And Beyond

Presenter: Franz Kiraly
Co-authors: Franz Kiraly;Benedikt Heidrich

This presentation introduces sktime - the python framework for time series modelling - and how it can be used by applied scientists and methodology researchers to build, share, and benchmark algorithms for time series modelling, in particular forecasting.sktime is a library of algorithms, with a vision to integrate the entire time series modelling ecosystem. It is openly governed, by a distributed community of developers, users, and researchers.sktime is built with the methodology developer in mind, who can use one of the prefabricated extension templates to share cutting-edge research with the vast user base, construct novel algorithms out of primitive components, or set up experiments easily.sktime is built with the applied end user in mind, who can easily search the register of hundreds of algorithms crossing many individual

packages, and use lego-like composition pattern to combine transformation, feature extraction, auto-ML, or ensembling steps to easily deployable composite algorithms meeting their application needs.Each sktime-compatible algorithmic component (e.g., forecaster) is a mini-package, with individual owners, authors, taggable properties (e.g., capability to make probabilistic forecasts or in-sample forecasts) and python software dependencies.The presentation gives a general purpose introduction geared towards applied forecasters and developers of forecasting methodology, and then focuses on features of special interest to the community:- common abstraction interfaces for deep learning and the recent array of foundation models- reproducibility enabling patterns, such as standardized specification language for algorithms, composite models- marketplace features such as submitting algorithms to sktime and sharing them as part of the ever-groing index of models- setting up benchmarking studies such as M4 or M5, using objects in sktime such as forecasters, components, compositions, performance metrics- challenges, progress, and areas of development in probabilistic, global, and/or hierarcihcal forecasting enabled by the "ease of specification" that sktime providesPrior experience in python or software engineering is not required, this presentation is intended for an audience consisting of applied researchers, model developers, mathematicians, or methodologists. Pointers to further entry points on usage and contribution to the platform will be given as part of the presentation.

Fast Hampel: Enhancing Time Series Outlier Detection Performance Through A Novel Implementation Of The Hampel Filter

Presenter: Hongtao Hu
Co-authors: Hongtao Hu;Ran Bi

Time series outlier detection is a critical task in a variety of domains, including finance, health care, and industrial processes. The Hampel filter, renowned for its efficacy in identifying anomalies within time series data, uses a sliding-window approach of customizable dimensions. Several implementations of the Hampel filter exist, but they often exhibit suboptimal performance in terms of speed and handling missing values. This presentation introduces the fast Hampel method, a novel approach developed at SAS that takes advantage of the efficiencies of quick-sort and binary search algorithms. The new algorithm takes into account the correlation between multiple computational steps of the Hampel filter and uses binary search for global optimization. Our method outperforms existing implementations, reaching a time complexity of O(nm) compared to O(nmlog(m)) for traditional approaches, where n is the series length and m is the sliding window size. The implementations of our fast Hampel method demonstrate a 50% to 70% improvement in speed compared to existing implementations.

A Set Covering Model For Change Point Detection

Presenter: Vittorio Maniezzo
Co-authors: Vittorio Maniezzo

Detecting points of change in time series analysis is critical because it enables the identification of critical shifts or transitions within the data, facilitates the understanding of underlying patterns, trends, and anomalies, and produces actionable results, specifically for forecasting and missing value imputation. The segmentation of time series and the corresponding detection of points of change are therefore of significant theoretical and practical importance.The presentation proposes a Mixed Integer Programming (MIP) based approach, both a Lagrangian matheuristic and an exact model, to change point detection. The underlying mathematical model is able to adapt to different requirements on the resulting model, making it more tolerant to residuals or more representative of short trends. The analysis imposes no constraints on the model associated with each segment, which can be linear, polynomial, or defined on any distribution function of interest. It can even be a combination of different models, allowing for linearity on some segments and, for example, exponential increases on others.This flexibility comes at the cost of having to solve a NP problem on large size instances. The increase in effectiveness of general MIP solvers allows to consider the solution of real-world instances to optimality, and it also provides the basis for the design of mathematically grounded heuristics. We report preliminary results on sensor data series obtained in an

environmental monitoring, but also two experiments on financial and healthcare use cases.The results so far provide only a first indication of the possibilities offered by mathematical models, but also of the difficulties to be overcome. Besides the obvious limit imposed by the NP-hardiness of the problem (the "curse of dimensionality"), which is only partially alleviated by heuristic solving, there are problem-specific issues to be faced, such as which cost function to use or which constraint to impose on the runs. In fact, the final result is only indirectly determined by our model, and there are still cases where results are unsatisfactory to the eye, even though they are numerically satisfying.

A Multifrequency Shot-Noise Approach To Volatility Forecasting

Presenter: Jiawen Xu
Co-authors: Jiawen Xu;Laurent Calvet;Yapei Zhang

Empirical evidence shows that volatility jumps upwards due to shocks to uncertainty and then gradually decreases as the uncertainty resolves, a pattern often called the shot-noise effect. In this paper, based on the MSM model in Calvet and Fisher (2004), we develop a shot-noise volatility model (SN-MSM) that parsimoniously captures the jump-decay patterns and multifrequency properties of volatility. SN-MSM is obtained by introducing an asymmetric component into the transition matrix of a Markov-switching multifractal (MSM). SN-MSM generates extreme jumps and volatility decay patterns while preserving the fat tails and volatility clustering of MSM. We derive the closed-form likelihood function of our process and verify that maximum likelihood estimation produces accurate parameter estimates in Monte Carlo simulation. In an empirical application to four major currency series, the SN-MSM model exhibits superior performance than popular volatility models both in- and out-of-sample.

A Backcasting Approach For Anomaly Detection In Time Series Data

Presenter: Priyanga Dilini Talagala
Co-authors: Priyanga Dilini Talagala

This work introduces a novel framework for anomaly detection in time series data using backcasting. The proposed approach operates within a rolling window paradigm, where upon receiving new datasets, backcasts are iteratively generated. Anomalies are identified through significant deviations observed in the backcasted values compared to their predecessors within the rolling window. To accomplish this, we establish a typical distribution for the backcasted values using historical data. A significant deviation from this established backcasted distribution is marked as an outlier occurrence, providing precise localization of anomalies within the dataset. This methodology promises enhanced anomaly detection sensitivity and adaptability to dynamic data environments as it continuously updates the backcast model with incoming data. Key contributions of this framework include its ability to provide real-time anomaly detection, crucial for timely intervention in fluctuating datasets. Furthermore, the framework addresses the challenge of detecting anomalies amidst evolving data distributions, facilitating early anomaly detection and proactive decision-making. The effectiveness and versatility of the framework are demonstrated through comprehensive evaluations on synthetic and real-world datasets, including applications in outbreak detection scenarios. By analyzing real-world datasets encompassing disease incidence rates, hospital admissions, and environmental factors, the framework's ability to accurately identify anomalies indicative of potential outbreaks is assessed. This real-world application underscores the framework's practical utility in public health surveillance, enabling timely detection and response to emerging health threats for proactive intervention and mitigation strategies. Implementation of the proposed methodology is facilitated through open-source tools, ensuring accessibility and reproducibility for researchers and practitioners.

Measuring Cross-Country Spillovers Of Growth-At-Risk

Presenter: Giovanni Caggiano
Co-authors: Giovanni Caggiano;Natalia Bailey;Luyang Li

This paper uses a sample of 40 advanced and emerging economies and novel spatial econometric techniques to measure spillovers of downside risk to GDP growth across countries. We proceed in two steps. First, we estimate quantile regressions to construct a measure of downside risk to GDP growth, Growth-at-Risk (GaR), for 40 advanced and emerging market economies. Our measure captures the time-varying behaviour of the left tail of the conditional distribution of future GDP growth, estimated as a function of both country-specific and global economic and financial conditions. We then turn to the issue of modeling and quantifying the cross-country spillover effects of country-specific downside risks. For this purpose, we first disentangle the role of common shocks from that of idiosyncratic shocks in driving GaR. We then model the idiosyncratic GaRs (ie. the measure of GaR purged from the effect due to global shocks) using the Heterogenous Spatial AutoRegressive model (HSAR) developed by Aquaro et al. (2019). This model is particularly suited for our research question, because it relaxes the assumption of homogeneous spatial coefficients, which are allowed to differ across countries. Our empirical findings suggest that there exist significant spillover effects of GaR across countries. The strength of the spillovers is stronger for countries with tight economic links, especially those which are part of a formal economic union. In addition, we notice a high degree of heterogeneity in these spatial effects, with some countries (e.g. Australia and New Zealand) not being particularly affected by increases in economic risk elsewhere, while others (e.g. members of the EU) experiencing sizable impacts on risks to their economic activity from shocks originating in countries acting as their financial or trading partners. In terms of time dynamics, we find significant own-country feedback effects for most economies in our sample, while shocks to global financial conditions, monetary policy or oil markets pass through the system and die down within four quarters. These insights can prove useful to policymakers who are engaged in monitoring the degree and duration of exposure of their domestic economies to external shocks originating in neighbouring countries.

Term Spread Volatility As A Leading Indicator Of Economic Activity

Presenter: Anastasios Megaritis
Co-authors: Anastasios Megaritis;Dimitrios Bakas;Theodora Bermpei;Athanasios Triantafyllou

In this paper we examine the predictive power of the volatility of the US Treasury yield curve slope (term spread volatility) for economic activity. Our forecasting exercise shows that the US term spread volatility has significant forecasting power on various measures of US economic activity. The predictive power of the term spread volatility is higher for medium- and long-term forecasting horizons and remains robust to the inclusion of well-established predictors of economic activity, like interest rates, inflation, the term spread, stock market returns, credit spreads, and popular measures of economic uncertainty, like the VIX, and economic policy uncertainty index. Our results also show that the term spread volatility has statistically and economically differentiated forecasting power with that of other economic uncertainty measures. Moreover, the predictive power of the term spread volatility increases significantly after the 2008 Great Recession, showing that the linkages between uncertainty about macroeconomic expectations and macroeconomic performance have increased in the post-Great Recession period. Finally, our out-of-sample forecasting results show that the term spread volatility outperforms the term spread when forecasting economic activity in the longer-run.

Density Forecast Transformations

Presenter: Florens Odendahl
Co-authors: Florens Odendahl;Matteo Mogliani

The popular choice of using a direct forecasting scheme implies that the individual predictions do not contain information on cross-horizon dependence. However, this dependence is needed if the forecaster has to construct, based on the direct forecasts, predictive objects that are functions of several horizons; such as obtaining annual-average from quarter-on-quarter growth rates. To address this issue we propose to use copulas to combine the individual h-step-ahead predictive distributions into a joint predictive distribution. Our method is particularly appealing for practitioners for whom changing the direct forecasting specification is too costly. In a Monte Carlo study, we demonstrate that our approach leads to a better approximation

of the true predictive densities than an approach which ignores the potential dependence. We show the superior performance of our method in several empirical examples, where we construct (i) quarterly forecasts using month-on-month direct forecasts, (ii) annual-average forecasts using monthly year-on-year direct forecasts, and (iii) annual-average forecasts using quarter-on-quarter direct forecasts.

Testing Clustered Equal Predictive Ability With Unknown Clusters

Presenter: Oguzhan Akgun
Co-authors: Oguzhan Akgun;Alain Pirotte;Giovanni Urga;Zhenlin Yang

We develop tests for the clustered equal predictive ability (C-EPA) hypothesis in panels for the case when there is no particular information on the clusters for which the quality of competing forecasts may differ. Building on the recent literature on selective inference for clustering methods, we develop tests for the C-EPA hypothesis which correctly control for the Type I error rate. We compare the performance of the proposed methodology with a more straightforward set of split-sample tests. Our results show the asymptotic validity of the proposed statistics and that they have excellent small sample properties. The empirical relevance of the tests is illustrated in a model comparison exercise for a large data set of exchange rate forecasts.

Superior Predictive Ability In Unstable Environments With An Application To Downside Risk Forecasts

Presenter: Ignacio Crespo
Co-authors: Ignacio Crespo

This paper introduces the Fluctuant-SPA (FSPA) test, a methodology for evaluating superior predictive ability (SPA) in unstable environments. The highlight of the FSPA test is that it allows to detect superior predictive ability when it is time varying. We showcase our test using a comprehensive assessment of downside risk forecasts to the U.S. economy over a 45-year span, considering a large collection of forecast methodologies. A number of findings emerge from the empirical application. First, there is substantial heterogeneity in forecasting performance across time. Second, the quantile regression equipped with a financial conditions index – a major benchmark in this literature – outperforms its alternatives after the Global Financial Crisis (2007-2009), yet it is surpassed in the periods leading up to the crisis. These local findings contrast with recent global assessments where this benchmark was found inferior to several alternatives. Overall, the empirical application showcases that the FSPA is particularly useful for forecast evaluations in unstable environments.

Simultaneous Inference Bands For The Pit Histogram

Presenter: Matei Demetrescu
Co-authors: Matei Demetrescu;Felix Kießner;Malte Knüppel

The histogram of the Probability Integral Transforms (PITs) is a popular tool for the evaluation of forecast distributions. In the absence of systematic errors in those distributions, the PITs follow a uniform distribution over the interval [0,1]. Therefore, if the PIT histogram is not 'flat enough', systematic errors are likely to be present. We propose simultaneous inference bands, namely Bonferroni bands or asymptotically exact bands, to assess the flatness of the PIT histogram. In contrast to the often-used pointwise bands, simultaneous inference bands take the underlying multiple-testing problem into account. Their construction is straightforward when the PITs are iid. If the PITs are serially correlated – as in the case of multi-step-ahead forecasts – we construct both types of simultaneous inference bands by bootstrapping the PITs. We find that, in spite of their expected conservativeness, Bonferroni bands – possibly with adjustments when the series of forecasts is short – have good size and power properties. We use these Bonferroni bands to evaluate the Bank of England's inflation forecasts.

Forecasting With The Help Of Surveys Of Professionals: When Does It Improve (Point And Density) Accuracy?

Presenter: Federica Brenna
Co-authors: Federica Brenna

Professional forecasters' expectations are a well-studied and closely watched variable, given their importance for central banks' policy decisions. An open question, however, is how survey and model forecasts should be combined in order to improve forecasting performance, and to what extent adding survey information helps. The literature has proposed various methods to this end, from entropic tilting to optimal pooling, to other ex-post model adjustments. A standard result is that, on average, surveys improve, or at the least do not worsen, forecast accuracy. However, in some periods, particularly the more stable ones, the added information content of surveys is minimal or even detrimental to performance, as professionals may be overconfident and produce unrealistic density forecasts. In more turbulent times, on the other hand, their overconfidence and forward-looking ability can considerably improve forecast accuracy. In this paper, I exploit a new way (developed in own previous work, see Brenna and Budrys, 2024) of adding expectations to empirical models to produce forecasts, by augmenting a traditional VAR model with survey information and using the latter also to aid estimation. I use forecasts from the Survey of Professional Forecasters and real time macroeconomic aggregates for inflation, GDP and unemployment. The accuracy gains from adding SPF forecasts are limited on average over the sample, the more so the longer the forecast horizon. However, when analyzing sub-samples or rolling windows, the gains become apparent for periods of higher uncertainty. In a next stage of the project, I will investigate also the value added of density forecasts.

What Were They Thinking? Estimating The Quarterly Forecasts Underlying Annual Growth Projections

Presenter: Christian Hepenstrick
Co-authors: Christian Hepenstrick;Jason Blunier

Many prominent forecasters publish only annual projections. For applied work, however, an estimate of the underlying quarterly forecasts is often indispensable. We show that a simple state-space model can be used to obtain good estimates of the quarterly forecasts underlying annual projections. We validate the methodology by computing the annual projection implied by professional forecasts for quarterly GDP growth in the United States and then applying our imputation method. The imputed forecasts come close to the original quarterly forecast and perform as well as the original in a forecast evaluation. Applying the imputation methodology to Consensus forecasts for other advanced economies provides further evidence of the good performance of our proposed methodology.

Updating Survey Forecasts For Horizons Of Interest Using Other Survey Forecasts

Presenter: Malte Knüppel
Co-authors: Malte Knüppel

Survey forecasts are available for many different forecast horizons, and forecasts for some horizons might be updated more frequently than forecasts for other horizons. For variables like growth and inflation, the most common forecasts refer to the annual-average over annual-average growth rates (a-o-a growth rates) for the current and the next year. However, forecasts for the corresponding quarterly or monthly outcomes provide more information on the dynamics of the variable under study. For instance, in December, one might prefer obtaining forecasts for the next four quarterly growth rates to obtaining a single forecast for the a-o-a growth rate of next year. Yet, an update of these four quarterly growth rates might become available relatively late, while a new forecast for the respective a-o-a growth rate will usually become available in January. In this paper, we suggest updating forecasts of interest, which have a low frequency of updating, by using forecasts for a-o-a growth rates, which have a high frequency of updating. Employing

forecasts of Consensus Economics for inflation and GDP growth in several countries, we investigate different possibilities of updating forecasts for the next four quarters. The latter forecasts are updated once every three months, while the forecasts for a-o-a growth rates are updated each month. The resulting updating formulas are very easy to use in practice, and updating yields smaller forecast errors than sticking to the last published forecasts of the next four quarters for two more months.

Unlocking Wind Energy Potential: Improving Forecasting Accuracy In Brazil With Reanalysis Data

Presenter: Fernando Cyrino
Co-authors: Fernando Cyrino;Saulo Ferreira;Paula Maçaira

Unlocking Wind Energy Potential: Improving Forecasting Accuracy in Brazil with Reanalysis DataReanalysis data is increasingly valuable for forecasting in climate studies, especially in data-scarce regions. Wind speed time series from reanalysis datasets hold promise for forecasting applications such as wind energy assessment, yet their coarse spatial resolution poses challenges, notably in Brazil. This study evaluates the suitability of the MERRA-2 wind speed time series for forecasting Brazilian conditions. We aim to enhance their accuracy and align them with observed data through strategies like interpolation and bias correction. Our research bridges the resolution gap by assessing data quality, evaluating natural variability, introducing bias correction methods, and analyzing temporal and spatial scales. Insights into MERRA-2 data's utility for wind energy forecasting in Brazil will advance the reanalysis of the dataset used in climate and energy forecasting. Findings will inform decision-making on wind energy development and resource management in Brazil and beyond.

Seamless Short- To Mid-Term Probabilistic Wind Power Forecasting

Presenter: Gabriel Dantas
Co-authors: Gabriel Dantas;Jethro Browell

This work proposes a novel methodology for short- and mid-term wind power forecasting (i.e., up to 7 days ahead). Short- and mid-term Wind Power Forecasting (WPF) is crucial for reliably and cost-effectively integrating wind energy into existing power networks. Furthermore, probabilistic forecasts are increasingly adopted by users to support decision-making, including real-time operations, operational planning, and energy trading.Established methods for probabilistic short- and mid-term WPF, from hours to three days ahead, and from three days to seven days ahead, respectively, use Numerical Weather Prediction (NWP) as inputs to statistical and machine learning models. Methods powered by deterministic NWP perform all uncertainty estimation in post-processing, which is typically sufficient in the short-term; however, ensemble NWP is necessary to quantify weather forecast uncertainty in the mid-term. In both cases, the weather-to-power uncertainty is only modelled implicitly.To the best of the authors' knowledge, no previously proposed method explicitly models and combines the weather forecast uncertainty with the weather-to-power uncertainty. Thus, methods powered by ensemble NWP ignore a source of uncertainty and naively attempt to correct the forecasting via post-processing. As a result, they can only achieve accuracy comparable to the state-of-the-art in mid-range horizons, where NWP uncertainty becomes dominant.This work proposes a novel methodology that can be used for short- and mid-term wind power forecasting. It is based on probabilistic weather-to-power modelling associated with a novel ensemble post-processing procedure to produce forecasts. The proposed methodology is demonstrated on 113 wind farms in Great Britain using data collected over five years. This work uses the HRES deterministic NWP model and ENS probabilistic NWP from ECMWF. The results show that the proposed methodology outperforms the state-of-the-art in mid-term horizons where NWP uncertainty dominates. Furthermore, it achieves state-of-the-art accuracy at short-term horizons, significantly reducing the computational effort. We also show that while ensemble NWP adds value for onshore wind power forecasting beyond 72h ahead, this information significantly enhances forecasts for offshore wind farms from 12h ahead .

Wind Energy: Forecasting For Double-Bounded Stochastic Processes

Presenter: Pierre Pinson
Co-authors: Pierre Pinson;Amandine Pierrot

Wind energy forecasting is a very mature field within forecasting, both in terms of methodological developments and operational practice. Forecasts are provided for a very large part of the wind power generation capacities installed worldwide, based on different methodologies, input features, etc. while having different potential formats. We want here to look at one of the core features of wind power generation as a stochastic process, which is to be double-bounded between 0 and the nominal capacity of the wind farm, or portfolio of wind farms, at hand. On top of that, the upper bounded may actually vary in time owing to varying performance of the turbines, faults, maintenance, etc. This upper bound may not even be known in real time.The double-bounded nature of wind power generation as a stochastic process has an impact on forecast uncertainty and estimation in models used for forecasting, eventually. We will concentrate on the specific issues related to learning and forecasting double-bounded processes for which one of the bounds may vary in time. We will look at potential models and estimation approaches, as well as forecast verification aspects. The learning approach is developed in an online framework, i.e., by adaptively adapting model parameter estimates every time new data becomes available. Since the loss function may not be convex, but quasi-convex instead, we explore alternatives to the popular stochastic gradient descent. Illustrative results are discussed based on simulated data, as well as real-world data from large offshore wind farms.

Spatio-Temporal Probabilistic Forecasting Of Circular Variables: Enhancing The Value Of Wind Power Prospective Models By Incorporating Wind Direction Probabilistic Forecasts

Presenter: Mario E. Arrieta-Prieto
Co-authors: Mario Arrieta-Prieto;Kristen Schell

The pressing need to decarbonise the energy system to mitigate climate change has led to a shift in attention towards renewable sources of energy. In particular, energy derived from wind has increased its role in the market, becoming even a critical component of the supply portfolio for countries with high penetration of renewable sources, such as Canada, Denmark, Sweden, the US and India, due to its low marginal costs and potentially ubiquitous exploitability. However, its stochastic and climate-influenced variability makes it challenging to consider it a reliable source. Hence, cutting-edge models derived on historical data are key to a) understand and model the randomness involved in wind dynamics, b) quantify the uncertainty associated with the wind power conversion process; and finally, c) ensure wind energy's full utilization within the grid. To that purpose, several models have been proposed in the literature to characterize the evolution of the physical phenomenon of wind and its impact in the energy conversion process. These efforts, though, have focused mainly on studying the spatio-temporal evolution of wind speed's magnitude. Considerably fewer works have highlighted the importance of modeling wind direction and its variation, as an important predictor of wind power output. This work presents advances in elaborating the theoretical requirements that a data-driven forecasting model should satisfy, in order to accurately model the circular features of wind direction and its interrelationships with wind speed and wind power dynamics. The forecasting model developed is validated using high-resolution wind data, gathered by the West Texas Mesonet network.

Mstl-Nnar: A New Hybrid Model Of Machine Learning And Time Series Decomposition For Wind Speed Forecasting

Presenter: Mohammed Elseidi
Co-authors: Mohammed Elseidi

Wind speed forecasting is essential for various domains, such as renewable energy generation, aviation, agriculture, and disaster management. However, wind speed is a complex and stochastic phenomenon that exhibits multiple stochastic patterns. Therefore, forecasting methods that can account for the dynamics

and uncertainty of wind speed are required. The literature on wind forecasting classifies the time-scale into different categories, such as short-term, medium-term, and long-term. The data used for these categories is usually daily or sub-daily, such as hourly. This implies that high frequency data poses more challenges for the modeling. Machine learning (ML) methods have shown great potential in forecasting wind speed. However, ML methods may not be able to capture the complex and stochastic patterns of time series data, especially when there are multiple seasonal cycles. Therefore, it is often beneficial to use a preprocessing method like time series decomposition before applying ML methods. This article proposes a novel hybrid model that integrates machine learning and time series decomposition for forecasting wind speed. A multivariate seasonal trend decomposition method based on loess (MSTL) is used to preprocess the wind speed data before applying various forecasting methods. MSTL is a generalization of the popular STL method that can split the time series data into a trend component, multiple seasonal components, and a residual component. The seasonal components are considered as deterministic and forecasted using a seasonal naive method that repeats the most recent cycle. For the other non-seasonal components, the neural network autoregression (NNAR) method is employed. The same approach is followed for two common statistical methods: autoregressive integrated moving average and exponential smoothing state space model (ETS). This enables the comparison of the performance of machine learning and statistical forecasting methods based on time series decomposition. The MSTL-NNAR model is also contrasted with the NNAR model without preprocessing to assess the effectiveness of MSTL. The proposed model is applied to two wind speed datasets with different frequencies: daily and hourly. The model is evaluated for both long-term and short-term forecasts by comparing it with other models using various accuracy and bias measures. The results show that the hybrid model, MSTL-NNAR, outperforms all other models in terms of accuracy and bias for both types of

Wind Speed Forecasting By A Physics-Inspired Machine Learning Approach

Presenter: Martina Zannotti
Co-authors: Martina Zannotti;Carlo Lucheroni

The paper introduces a machine learning effective short-term multi-horizon autoregressive forecasting method for wind speed time series with one-minute time granularity, inspired by considerations about the phenomenon of turbulence.When used for forecasting wind speed at horizons of hours while using time lags at the hour or longer, this kind of minute data is difficult to manipulate because of the very large amount of time points available.However, because of turbulence and multiscale phenomena, blindly averaging or discarding data can mean losing too much information, at detriment of very short term forecasting quality. To this end, a diagonal recurrent neural network (DRNN), shallow and with a very small number of parameters, trained statically, topped by a recursive least squares (RLS) filter trained online, is used at once on multiple time averages. The purpose of the DRNN is that of capturing features at scales from half-months to hours, whereas the purpose of the RLS filter is that of including the effect of eddy phenomena at the minute.In this architecture, the DRNN can be seen as a pool of nonlinear recursive filtering modules each dedicated to a different scale, working in parallel, and delivering a backbone, long- to mid-term dynamic set of features to be processed by the RLS topping filter. This new architecture is shown to be very accurate in the test set, and it can be seen as a new way to look at recurrent networks. It consists of splitting the weight space into two sectors, one dedicated to extracting stable features and the other dedicated to more volatile features. This strategy is also inspired by the current debate on lifelong learning.Numerical experiments are presented based on NREL wind speed minute data, and compared on these data with complex state-of-the-art models like NBeats and Temporal Fusion Transformer.

Development Of A Pragmatic, Robust And Useful Macroeconomic Scalar To Adjust The Ifrs 9 Pd Forecast For Forward-Looking Information In Developing Countries

Presenter: Tanja Verster
Co-authors: Tanja Verster;Helgard Raubenheimer

We develop a probability of default forecast model to capture the causal link to the macroeconomic environment. The goal is to propose a methodology that can be utilised in ECL modelling (with a specific focus on PD) to capture logical links between the PD and the macroeconomic environment. IFRS 9 requires a forward-looking component (added to each risk parameter) when compared with the IAS 39 occurred losses. An alternative modelling approach for the case where there is no apparent link between macros and PDs in developing countries is proposed. Many forward-looking models tend to either severely over-or underpredict and there is no apparent linear relationship between the PD and macro variables when performing back-testing. Forward-looking models in developing countries usually use the following macroeconomic variables: Gross Domestic Product, Inflation, Foreign Exchange Rate, and the Central Bank Rate to forecast future PDs for 12 months and a lifetime. We developed a methodology that is robust, pragmatic and useful. This methodology will be illustrated in two developing African markets.

Currency Option Implied Volatility Networks And Geopolitical Risk

Presenter: Meltem Yagli
Co-authors: Meltem Yagli

This paper investigates the dynamic and directional network connectedness among implied volatility measures of twenty currency pairs from July 2004 to January 2023. Our data is collected from Bloomberg. Implied volatility measures are extracted from currency option prices in a model-free manner. These implied volatility measures are forward-looking and capture investors' expectations about the currency pairs' uncertainty over the next 30 days.We then study how shocks to implied volatilities propagate through the currency market, forming network structures over time. We construct measures of both aggregated and net directional networks following the methodology by Diebold and Yilmaz (2014). To analyse the transmission direction of shocks, we utilize a rolling window approach. Specifically, we employ rolling variance decompositions to track connectedness over time. This dynamic approach allows us to adapt the analysis window size based on evolving market conditions, providing real-time insights into the changing interconnectedness among currency pairs. By leveraging the simplicity and flexibility of the rolling window methodology, we gain valuable insights into market dynamics and inform risk management strategies effectively. Our aggregate network appears to spike in correspondence with main global economic and political events; the net directional networks present spikes related to country-specific distress.Furthermore, we study whether currency option implied volatility networks contain any predictive information for future geopolitical risk. Our findings reveal a significant predictive link between implied volatility networks and geopolitical risk. Our findings underscore the importance of information enclosed in currency uncertainty networks to improve the understanding of geopolitical risks. Our results open new avenues for research about potential monitoring tools for global geopolitical risk which can aid policy makers in their continuous effort to ensure financial stability worldwide.Keywords: network analysis, geopolitical risk, currency implied volatility, financial stability.

Forecasting The Success Of International Joint Ventures

Presenter: Aws YOUNES
Co-authors: Aws Hamo Younes;Konstantinos Nikolopoulos;Michel Phan

A joint venture is a business partnership where two or more companies work together on a specific project, sharing resources and risks. Each company keeps its own identity while collaborating to achieve a common goal. Partner selection can be defined as the process of seeking, evaluating, and finally choosing the right partner to achieve the firm's strategic growth objectives in a specific host country. Country Governance is defined as the traditions and institutions by which authority in a country is exercised. There are six dimensions of governance are constructed based on this definition; Voice and Accountability (VA), Political Stability (PV), Government Effectiveness (GE), Regulatory Quality (RQ), Rule of Law (RL) and Control of Corruption (CC). In an increasingly globalized world, firms aiming to stay competitive in international business must explore foreign markets and many multinational enterprises (MNEs) from developed countries use Foreign Direct Investment (FDI) as a strategy to enter markets in developing economies, benefiting from

local knowledge and connections. Due to external uncertainties and the challenges of operating in a foreign market, MNEs often prefer forming International JointVentures (IJVs) with local partners rather than wholly-owned subsidiaries (WOSs) to enhance the success of their international endeavours and reduce the risk of failure. This study focuses on analysing how weak country governance influences the criteria MNEs use to select local partners and develops a forecasting model for Forecasting the success of International Joint Ventures.

A Tale Of Dynamic Tail Carbon Beta

Presenter: Laura Garcia-Jorcano
Co-authors: Laura Garcia-Jorcano;Juan-Angel Jimenez-Martin;M.- Dolores Robles

Carbon-driven climate risk is an important concern for managers, who must assess the impact of transition risks in firms' business operations, and for investors, who need to incorporate climate risk exposures in their firm valuation models (e.g. Bolton and Kacperczyk, 2021). In this paper, we adopt this strategy to assess the exposure of industry portfolios to carbon-driven climate risk, which we measure from the CO2 emission allowance (carbon) prices. We analyze extreme stock returns conditional to extreme movements in carbon returns. We propose a dynamic tail carbon beta (TCB) that combines the Van Oort and Zhou (2019) tail "market" beta with the CAViaR model. We estimate TCB in different risk situations in the stock and carbon markets to analyze to what extent TCB depends on extreme conditions in both markets. For this purpose, we define low-probability brown and green states in the carbon market, and downside and upside risk in the stock market, which we define from the tails of the carbon and stock returns marginal distributions, respectively. We focus on the US equity market at the industry level and analyze the daily prices of 60 industrial indices from 2009 to 2023. We use daily carbon prices from the European Emissions Trading Scheme (EU-ETS) to measure carbon-driven climate risk, which allows for avoiding endogeneity concerns of carbon prices from US emissions markets for our analysis. We frame the analysis of the link between stock returns and carbon-driven transition risk in the Fama and French (2016) five-factor pricing model extended with the carbon risk factor. The study of our sample documents time-varying asymmetric tail dependence that is stronger when upside risk and brown state co-occur than when downside risk and green state do, and the Energy and Real Estate sectors show the highest size, variance, and the strongest persistence of the TCBs. Tail dependence in the other two scenarios is less asymmetric. We conclude that investors' worry about climate risk is influenced not only by the state of the climate but also by the conditions of the financial market, and this concern is heightened when they receive better investment returns.

An Expert Survey On The Perception Of Climate Geoengineering In South Korea

Presenter: Uijin Jung
Co-authors: Uijin Jung;Soonduck Yoo

This paper analyzed on the perception of climate geoengineering in South Korea. As climate change already worse than expected, the Earth's surface temperature has risen by 1.1 ℃ since 1850. Due to this severity, discussions about climate geoengineering have emerged. In the IPCC sixth Assessment report(2022) mentions the introduction of carbon dioxide removal(CDR) and solar radiation modification(SRM) technologies. However, these technologies are controversial due to their manipulation of the nature environment. Moreover, once the technology is applied to the Earth, it cannot be reversed to its previous state. The aim of this study is to determine the impacts of climate geoengineering and to understand the perception of each climate geoengineering technology among experts in South Korea. A perception survey was conducted targeting 320 experts in the field of science and technology. This study tested the following five hypothesis: (H1) Higher interest in climate change will be associated with a greater need for CDR and SRM technologies. (H2) A Higher perceived threat of climate change will be associated with a greater need for CDR and SRM technologies. (H3) Higher trust in technology will be associated with a greater need for CDR and SRM technologies.(H4) Those who believe that climate change is cause by human activities are more inclined to

consider the need for CDR and SRM technologies. (H5) Those who believe it is acceptable to tamper with nature, they are more likely to consider the need for CDR and SRM technologies. According to the results, several factors such as perceived threat of climate change, the cause of climate change and tampering with nature are significantly associated with CDR technologies. In the case of SRM technologies, a negative correlation was observed between a person's interest in climate change and their perceived necessity for SRM technologies. Even though there is a need for climate geoengineering technologies, the level of necessity varies for each technology.

Simplifying Random Forests By Sparsification

Presenter: Nils Koster
Co-authors: Nils Koster;Fabian Krüger

Since their introduction by Breiman, Random Forests (RFs) have proven to be useful for both classification and regression tasks. Lin and Jeon show that the RF prediction of a previously unseen observation can be represented as a weighted sum of all training sample observations.This nearest-neighbor-type representation is useful, among other things, for computing RF-based forecast distributions [Meinshausen]. In this paper, we consider simplifying non-sparse RFs by focusing on a small, sparse subset of nearest neighbors, while setting the remaining weights to zero, effectively sparsifying RFs.This step greatly improves the interpretability of (probabilistic or deterministic) RF predictions. It can be applied to any forecasting task without re-training existing RF models.In empirical experiments, we document that the simplified predictions can well be similar to or exceed the original ones in terms of forecasting performance. We explore the statistical sources of this finding via a stylized analytical model of RFs. The model suggests that simplification is particularly promising if the unknown true forecast distribution contains many small weights that are surrounded by estimation noise.

Probabilistic Forecasts For Global Models: Empirical Insights From Gradient Boosted Decision Trees.

Presenter: Filotas Theodosiou
Co-authors: Filotas Theodosiou;Yves R. Sagaert;Liselot De Vlieger

Global Learning has recently gained wide recognition, as it allows models to scale across multiple time series, significantly enhancing point forecasting performance. However, little research compares the performance of different probabilistic predictions for global models, in contrast to probabilistic forecasts for statistical models. Yet, probabilistic forecasting can enhance decision-making under uncertainty, crucial for optimizing business practices such as inventory management. In this work, we explore various methods for generating prediction intervals for Global Models, specifically emphasizing on Gradient Boosted Decision Trees (GBDTs). These models are known for their robust out-of-the-box performance across various forecasting scenarios, relaxing the need for excessive tuning and maintenance required by more complex Deep Learning architectures. We apply the proposed methods on different demand forecasting case studies in the sector of ultra-perishable food. Each case exhibits varying degrees of intermittency levels, reflecting the challenges encountered in practical settings. Additionally, we explore the inventory implications of the generated prediction intervals, providing insights into effective decision-making and highlighting the practical applications of our findings.

On The Diversity In Boosting Based Ensemble Learning

Presenter: Jue Wang
Co-authors: Jue Wang;Shuqin Liu;Sheng Cheng

The significance of diversity in ensemble learning has become a key focus of contemporary research due to its critical role in improving the generalization capabilities and forecasting accuracy of predictive models. In this study, we examine the importance of diversity and introduce a novel methodology

specifically designed to foster and enhance it. By embracing diversity in ensemble learning, we aim to elevate the performance and robustness of predictive models across various domains and applications.Firstly, we develop a novel diversity measurement for regression tasks, considering both the forecasting accuracy and bias direction rather than solely relying on the squared error. This helps in preventing the measurement from being solely influenced by the fraction of noisy samples that are difficult to predict among all samples.Secondly, we present a loss function penalized by diversity to balance forecasting accuracy and diversity. Incorporation of a diversity regularization term assists to achieve a balance between precise prediction and diversity by gradually shifting attention to samples that are difficult to accurately forecast. This mechanism is incorporated into the initialization stage of base models, enhancing the diversity among the base models.Lastly, we provide a two-dimensional attention (Bi-attention) mechanism that combines both sample and feature perturbations, resulting in superior diversity among individual models. In contrast to one-dimensional random sample perturbation, Bi-attention employs a supervised perturbation method. This method progressively improves the model's capacity to comprehend intricate relationships. Notably, it enables the assimilation of information from extreme samples and fully exploits the prominent role of high-dimensional features in different sample scenarios.To validate the efficacy of our proposed methodology, we conducted experiments using the WTI crude oil price dataset and UCI datasets. In comparison to the traditional boosting ensemble strategy, our approach achieved a substantial reduction in the maximum absolute percentage error (MAPE), with a maximum decrease of 68.4%. The consistent experimental results demonstrate the superior predictive performance of our method, underscoring the significance of diversity in ensemble learning and highlighting the effectiveness of our approach.

Tail Calibration Of Probabilistic Forecasts

Presenter: Sam Allen
Co-authors: Sam Allen;Jonathan Koh;Johan Segers;Johanna Ziegel

Probabilistic forecasts comprehensively describe the uncertainty in the unknown outcome, making them essential for decision making and risk management. While several methods have been introduced to evaluate probabilistic forecasts, existing techniques are ill-suited to the evaluation of tail properties of such forecasts. However, these tail properties are often of particular interest to forecast users due to the severe impacts caused by extreme outcomes. In this work, we reinforce previous results related to the deficiencies of proper scoring rules when evaluating forecast tails, and instead introduce several notions of tail calibration for probabilistic forecasts, allowing forecasters to assess the reliability of their predictions for extreme events. We study the relationships between these different notions, and propose diagnostic tools to assess tail calibration in practice. The benefit provided by these diagnostic tools is demonstrated in an application to European weather forecasts.

Generating And Evaluating Probabilistic Forecasts - A Tutorial

Presenter: Florian Ziel
Co-authors: Florian Ziel;Bahman Rostami-Tabar;Siddharth Arora

In many situations, probabilistic forecasting is crucial to inform policy and decision-making processes, especially in the face of uncertainty. While numerous research papers and books exist on this topic, a tutorial on probabilistic forecasting that provides the reader with a clear and concise introduction to the subject and signposts the advanced user to relevant resources is needed. Our objective is to provide a thorough tutorial on probabilistic forecasting that is appropriate for academics and practitioners. First, we briefly discuss different forecasting objectives, such as prediction intervals, distributional forecasting, and fully multivariate forecasts. While focusing on distribution forecasts, we cover major data, methods, and evaluation issues. We discuss time series and input data characteristics that support the choice of forecast modelling. We present state-of-the-art methods for distributional forecasting and discuss important properties, i.e., the ability to satisfy forecasting objectives (e.g. forecast horizon), coverage of relevant data characteristics, and computational cost. We will cover significant recent developments in data science, including forecasting using high-dimensional statistics, decision tree learning, deep learning and large

language models. Finally, we discuss reporting options for distributional forecasts and their evaluation in terms of calibration and sharpness, especially using scoring rules. To enhance reproducibility and facilitate the adoption of different applications and intuitive understanding, we provide the data, R and Python code, and the entire paper written in Quarto. All materials required to reproduce this paper will be accessible via a public GitHub repository.

Uncertainty Quantification In Forecast Comparisons

Presenter: Tanja Zahn
Co-authors: Tanja Zahn;Marc-Oliver Pohle;Sebastian Lerch

Comparing competing forecasting methods via consistent scoring functions is the cornerstone of forecast evaluation. Skill scores enhance interpretability in that they indicate the relative improvement of a forecasting method over a competitor. We introduce simultaneous confidence bands for expected scores and skill scores to quantify and communicate sampling uncertainty in forecast comparisons. The confidence bands are versatile in that they allow for joint inference over multiple forecast horizons, variables, forecasting methods and/or locations. Their validity in such very common multivariate settings is in contrast to pointwise confidence bands or pairwise Diebold-Mariano forecast accuracy tests, which are invalidated by multiple comparison problems. The confidence bands are applicable for any type of forecast, from mean over quantile to distributional forecasts. We provide a bootstrap implementation, where the scores are resampled via a moving block bootstrap, and an alternative one not relying on a bootstrap. We show that our bands have exact asymptotic coverage under multivariate extensions of the assumptions underlying the classical univariate Diebold-Mariano test. In two case studies we apply our approach to compare data-driven and physical models for probabilistic weather forecasting and to quantify the benefits of time-varying parameter models for macroeconomic forecasting.

Semiconductor Ageing Forecasting Through Temporal Fusion Transformers. Short- And Long-Term Forecasting Horizon Comparatives With Classical Forecasting Methods

Presenter: Jose Aizpurua
Co-authors: Jose Aizpurua;Adrian Villalobos;Iban Barrutia;Rafael Peña-Alzola

Metal-oxide-semiconductor field-effect transistor (MOSFETs) are ubiquitous semiconductor devices used in widespread power electronic applications. Generally, they operate in switching mode with different working profiles, which leads to repeated temperature fluctuations that cause damage in the interfaces between different material layers. This damage is expressed mainly as bond wire lift-off, which is empirically characterized by an exponential grow in the on resistance of the MOSFET. However, tracking and forecasting the ageing evolution of MOSFETs is complex and influenced by different ageing processes that vary with cycling and temperature. Short-term forecasts have gained the interest for online health monitoring applications, whereas long-term ageing forecasts tend to be challenging, as the ageing mechanisms can be diverse and may change abruptly. With the emergence of powerful machine learning (ML) based forecasting methods based on transformers, there is an opportunity to cover this challenging research area. Accordingly, this work presents a comprehensive comparison of different forecasting methods for different forecasting horizons tested and validated on a publicly available dataset with four run-to-failure experiments. The short and long-term forecasting performance of classical tracking and statistical forecasting algorithms such as Kalman filters and ARIMA filters is analysed. The results are compared with respect to ML based forecasting algorithms, including ensemble Extreme Learning Machines (ELM), and transformer-based Temporal Fusion Transformers (TFTs). Short-term forecasting is focused on 1, 2 and 4 steps ahead predictions and long-term forecasting is focused on 30, 50 and 70 steps ahead predictions. Different TFT architecture configurations have been designed and tested for different forecasting horizons. Results show that for short-term predictions, the best options were ARIMA and ELM. For long-term predictions, TFTs were the best option by a large margin. In the context of few run-to-failure trajectories, it has been observed that the forecasting ability of the TFT for short term predictions is limited and very sensitive. Finally, it is

shown that for long-term forecasting, it was necessary to design a synthetic covariate function.

Commodity Price Forecasting In Procurement

Presenter: Nico Beck
Co-authors: Nico Beck;Hans Georg Zimmermann

Commodity price forecasting has gained attention in forecasting, since it promises a huge benefit in procurement. Therefore, forecasters work on improving their models to enhance the accuracy. However, in areas like stock or price forecasting, which are very volatile, even small improvements compared to the naïve (no-change) forecast can be seen as a success. In general, it is unclear whether the information gained from the forecast induces a benefit in the actual procurement.To assess the actual value of commodity price forecasting, we predict prices of four steel types, that are highly relevant for large industry sectors, on a weekly basis for a horizon of twelve weeks. Afterwards, we simulate eight half year periods of procurement for each of the steel types, where the procurement decisions are fully based on the forecasts. We compare ARIMA, ARIMAX, ETS and Historical Consistent Neural Networks regarding their forecast accuracy and the costs they produce in the procurement simulation.We show that a good model forecast accuracy does not necessarily produce a low cost in the procurement simulation. Nevertheless, all forecast approaches can reduce costs compared to a baseline procurement policy, where the best approach outperforms the latter on 20 of 32 sub-periods.

The Digital Transformation Journey Of Sanofi Forecasting: A Blend Of Quantitative + Judgmental Forecasting In The Pharma Supply Chain

Presenter: Giorgia Felline
Co-authors: Giorgia Felline;Gráinne Costigan;Jamal Akhiad;Alex Pedurand

In this presentation, we want to introduce to you the mission of the SANOFI Statistical Forecasting Team operating under the Global Supply Chain Analytics Center of Excellence in the pharmaceutical sector. A multidisciplinary team with both technical and business expertise, we are dedicated to fostering innovation and ensuring continuous development.Embark our Digital Transformation journey!We will share insights into our demand forecasting processes, ways of working, and best practices. Learn about the generation of our statistical baseline and our close collaboration with counterparts in the Markets to gather their market intelligence, which further enriches our forecasts. Additionally, you will learn how we ensure the robustness and stability of these forecasts across the entire Supply Chain.Deep dive into our in-house developed and open-source codebase, the core engine of our Forecasting activities. After significant enhancements, this is now highly flexible and ready to adapt to urgent business needs.We will illustrate the methodologies employed in our data pipeline for signal cleansing and identifying demand drivers, along with the forecasting models currently in production. This encompasses both traditional statistical algorithms and our ongoing adoption of machine learning.We'll also highlight the evolution from local to cloud compute, discussing the challenges, limitations from the business and valuable lessons learnt along the way.Finally, we will dive into the strengths and weaknesses of our official metrics, wMAPE (Weighted Mean Absolute Percentage Error) and SPA (Sales Plan Adherence), alongside our recent push for adoption of FVA (Forecast Value Added) to monitor the overall judgmental process performance.This journey concludes by showcasing our achievements in terms of forecast accuracy and the savings generated for the company, demonstrating a successful transition from a 100% human-based approach to a collaborative one which integrates both quantitative baselines and judgmental adjustments.Lastly, we'll unveil the next steps and challenges of our journey.Get ready for an inspiring exploration of innovation within the Pharma Supply Chain!

Advancing Time Series Forecasting: A Comparative Analysis Of Bootstrapping Methods On Enbpi Performance

Presenter: NA
Co-authors: Benedikt Heidrich;Sankalp Gilda;Franz Kiraly

In the evolving landscape of time series forecasting, the Enhanced Non-exchangeable Bootstrap Prediction Interval (EnbPI) method stands out for its innovative approach to constructing distribution-free prediction intervals without relying on the exchangeability assumption. This assumption, often untenable in time series analysis due to the inherent order and correlation of data points, poses a significant challenge in traditional probabilistic forecasting methods. EnbPI, by integrating with any bootstrap ensemble estimator, offers a promising solution, especially for non-stationary time series data. This paper embarks on an empirical investigation to assess the efficacy and robustness of EnbPI across diverse time series datasets and contrasting bootstrapping methods, including block bootstrapping, residual bootstrapping, and Markov bootstrapping. Our research aims to delineate how varying bootstrap methodologies influence EnbPI's performance in terms of prediction accuracy, interval coverage, and computational efficiency. Through a comprehensive analysis spanning multiple real-world datasets, including renewable energy estimation and urban mobility forecasting, we evaluate EnbPI's adaptability and its capacity to maintain accurate marginal coverage under mild assumptions on time-series stochastic errors and regression estimators. Our findings contribute to the broader discourse on conformal prediction in time series analysis, highlighting EnbPI's versatility and potential for application across a spectrum of domains faced with the challenge of forecasting under uncertainty. This study not only enriches the theoretical understanding of EnbPI but also guides practitioners in selecting appropriate bootstrapping methods to enhance predictive performance in non-stationary environments.

Lag-Llama: Towards Foundation Models For Time Series Forecasting

Presenter: Kashif Rasul
Co-authors: Kashif Rasul

We present here our work on Lag-Llama, a general-purpose univariate probabilistic time-series forecasting model trained on a large collection of time-series data. The model shows good zero-shot prediction capabilities on unseen "out-of-distribution" time-series datasets, outperforming supervised baselines. We also fit the "smoothly broken power-laws" to investigate the scaling behaviour of this model as we increase the parameters of the model as well as data-points during training.

How To Bootstrap Time Series Without Attracting Attention Of Statisticians

Presenter: Ivan Svetunkov
Co-authors: Ivan Svetunkov

Bootstrap is extensively used in statistics and machine learning for cross-sectional data to account for uncertainty about the data, model form, and parameter estimates. However, conventional methods may not be suitable for time series data due to autocorrelation and specific dynamic structures. Over the years, various approaches have been developed to address this issue. Some assume specific models (e.g., STL), while others are non-parametric (e.g., Maximum Entropy Bootstrap, MEB). However, the former can be overly restrictive, while the latter may not perform well in case of outliers and external drivers. To address these issues, we propose a non-parametric bootstrap approach inspired by MEB, which does not assume any structure in the data yet creates reasonable copies of existing time series of different nature. These copies can be utilised in bagged ETS/ARIMA or any other approach involving small sample uncertainty. We demonstrate how the proposed bootstrap works using real-time series examples and assess improvements it brings in terms of forecasting accuracy compared to conventional approaches.

Speaking Of Inflation:the Influence Of Fed Speeches On Expectations

Presenter: Eleonora Granziera
Co-authors: Eleonora Granziera;Vegard Larsen;Greta Meggiorini

This paper examines the ability of the Federal Reserve (Fed) to influence expectations of economic

agents via speeches of FOMC members and regional Fed presidents. Using textual analysis, we extract an inflationary pressure index from the speeches and show that soft information about inflation impacts expectations of households, professional forecasters, and market participants. This effect is stronger after the Great Financial Crisis. We compute a measure of hawkishness of the FOMC members based on their speeches and find that professional forecasters anticipate inflation to be lower when FOMC speakers are perceived to be more willing to fight inflation. In contrast, households increase their inflation expectations when FOMC members talk about rising inflation regardless of the policy preference of the speaker.

Maximally Forward-Looking Core Inflation

Presenter: Karin Klieber
Co-authors: Karin Klieber;Philippe Goulet Coulombe;Maximilian Göbel;Christophe Barrette

Timely monetary policy decision-making requires timely core inflation measures. We create a new core inflation series that is explicitly designed to succeed at that goal. Precisely, we introduce the Assemblage Regression, a generalized nonnegative ridge regression problem that optimizes the price index's subcomponent weights such that the aggregate is maximally predictive of future headline inflation. Ordering subcomponents according to their rank in each period switches the algorithm to be learning supervised trimmed inflation—or, put differently, the maximally forward-looking summary statistic of the realized price changes distribution. In an extensive out-of-sample forecasting experiment for the US and the euro area, we find substantial improvements for signaling medium-term inflation developments in both the pre- and post-Covid years. Those coming from the supervised trimmed version are particularly striking, and are attributable to a highly asymmetric trimming which contrasts with conventional indicators. We also find that this metric was indicating first upward pressures on inflation as early as mid-2020 and quickly captured the turning point in 2022. We also consider extensions, like assembling inflation from geographical regions, trimmed temporal aggregation, and building core measures specialized for either upside or downside inflation risks.

Has The Phillips Curve Flattened?

Presenter: Barbara Rossi
Co-authors: Barbara Rossi;Atsushi Inoue;Yiru Wang

We contribute to the recent debate on the instability of the slope of the Phillips curve by offering insights from a flexible time-varying instrumental variable approach robust to weak instruments. Our robust approach focuses directly on the Phillips curve and allows general forms of instability, in contrast to current approaches based either on structural models with time-varying parameters or instrumental variable estimates in ad-hoc sub-samples. We find evidence of a weakening of the slope of the Phillips curve starting around 1980. We also offer novel insights on the Phillips curve during the recent pandemic: The flattening has reverted and the Phillips curve is back. Our results are important for forecasting inflation.

Forecasting Core Inflation And Its Goods, Housing, And Supercore Components

Presenter: Saeed Zaman
Co-authors: Saeed Zaman;Todd Clark;Matthew Gordon

This paper examines the forecasting efficacy and implications of the recently popular breakdown of core inflation into three components: goods excluding food and energy, services excluding energy and housing, and housing. A comprehensive historical evaluation of the accuracy of point and density forecasts from a range of models and approaches shows that a BVAR with stochastic volatility in aggregate core inflation, its three components, and wage growth is an effective tool for forecasting inflation's components as well as aggregate core inflation. Looking ahead, the model's baseline projection puts core inflation at 2.6 percent in 2026, well below its 2023 level but still elevated relative to the Federal Reserve's 2 percent objective. The probability that core inflation will return to 2 percent or less is much higher

when conditioning on goods or non-housing services inflation slowing to pre-pandemic levels than when conditioning on these components remaining above the same thresholds. Scenario analysis indicates that slower wage growth will likely be associated with reduced inflation in all three components, especially goods and non-housing services, helping to return core inflation to near the 2 percent target by 2026.

Impact Of Autocorrelation On The Smoothness Of The Estimated Trend Of A Time Series Where The Noise Follows An Autoregressive Process Of Order One

Presenter: Daniela Cortés Toto
Co-authors: Daniela Cortés Toto;Víctor Manuel Guerrero Guzmán

First, a smoothness index that controls the smoothness estimated in a time series with autoregressive errors of order one is presented; to estimate this trend, we use the Penalized Least Squares method. Then, we show the results of a simulation experiment, in which the impact of autocorrelation on the controlled smoothness when estimating the series trend is measured, for different series sizes and different levels of smoothness in the estimated trend. From this study we suggest how to appropriately choose the amount of smoothness when estimating the trend, such that the correlation coefficient and error variance parameters are always included in the interval of two standard errors around each parameter. Finally, the use of the smoothness index and the algorithm to estimate autocorrelation is illustrated with an empirical example of the estimation of the quarterly GDP trend of México.

Factor-Augmented Vars With Noisy Factor Proxies

Presenter: Soroosh Soofi Siavash
Co-authors: Soroosh Soofi Siavash;Emanuel Moench

In factor-augmented vector autoregression (FAVAR) models, some of the factors are treated as observable while the remaining factors are latent and need to be estimated from a large cross-section of time series. Given that economic concepts such as inflation or output can often be proxied by different variables and that macroeconomic time series are commonly subject to substantial data revisions, the assumption that some factors are perfectly observable appears unnecessarily strong. We propose to extend FAVARs to include observable variables which are noisy measures of the true underlying factors. We show that in this framework, the factors can be estimated by iterative principal components combined with a reduced-rank regression step. The estimator has superior performance particularly in the presence of weak noisy factors. We further illustrate its usefulness in estimating impulse response functions in an application to oil shocks.

Aggregating Interval Forecasts Viewed As Bivariate Data

Presenter: James Taylor
Co-authors: James Taylor

This paper considers interval forecast aggregation when there are many forecasters, and a record of past forecast accuracy is not available. For such cases, the median and trimmed means have been proposed as simple and robust alternatives to the mean. These robust approaches consider each interval bound separately. If each interval is viewed as a bivariate point, this amounts to finding the median and trimmed means separately with respect to the horizontal and vertical dimensions. If instead of analysing lower and upper bounds, we consider interval location and width, we are essentially viewing the bivariate points on axes rotated by 45 degrees. The natural extension of this is to view the bivariate points on axes rotated by other angles between zero and 90 degrees. However, another perspective is provided by the literature on multivariate statistical depth, which provides methods for ordering points in terms of centrality. The deepest point can be viewed as the median interval forecast, and the depth of each point can be used as the basis for trimming. In addition to the commonly used halfspace and simplicial depth measures, we consider weighted versions of each. We also describe how statistical functional depth can be used within approaches to aggregating forecasts of probability distributions. We provide empirical illustration using data from

surveys of professional macroeconomic forecasters.

Forecasting Using A Random Coefficient Autoregression Of Order P

Presenter: Philip Hans Franses
Co-authors: Philip Hans Franses;Dmitriy Knyazhitskiy

We consider forecasting economic time series using a random coefficient autoregression of order p. We present a simple diagnostic test to indicate proper model specification, where this test method also suggests an OLS-based estimation method. We show however that a Maximum Likelihood estimation method performs better in practice. Applications to forecasting quarterly US unemployment and annual US inflation shows the potential merits of this model.

Threshold Midas Forecasting Of Inflation Rate

Presenter: Chaoyi Chen
Co-authors: Chaoyi Chen;Yao Rao;Yiguo Sun

We propose several threshold mixed data sampling (TMIDAS) autoregressive modelsto forecast the Canadian inflation rate using predictors observed at different frequencies.These models take two low-frequency variables and a high-frequency index asa threshold variable. We compare our TMIDAS models to commonly used benchmarkmodels, evaluating their in-sample and out-of-sample forecasts. Our results demonstratethe good forecasting performance of the TMIDAS models. Particularly, thein-sample results highlight that the TMIDAS model using the high-frequency indexas the threshold variable outperforms other models. Through unconditional superiorpredictive ability (USPA) and conditional superior predictive ability (CSPA) tests forout-of-sample evaluation, we find that no single model consistently outperforms theothers, although at least one of our TMIDAS models remains competitive in mostcases.

Nowcasting Macroeconomic Variables With A Sparse Mixed Frequency Dynamic Factor Model

Presenter: NA
Co-authors: Domenic Franjic;Karsten Schweikert

We propose a two-step estimation procedure for sparse dynamic factor models (SDFM) in a mixed frequency setting. A sparse factor loadings matrix is estimated from sparse principal components analysis, which is then used to initialise the Kalman filter and smoother. We generalise existing theoretical results on the properties of the sparse principal components estimator and are able to show that our estimation procedure provides a consistent estimator of an orthogonal transformation of the true model parameters. Further, we employ a simulation study and real world application to investigate the nowcasting performance of SDFMs estimated in such a way. It is found that the cross-validated SDFM is able to outperform the benchmark model in most cases, especially when the measurement errors are cross-correlated.

Factor Augmented Forecasting Subject To Structural Breaks In The Factor Structure

Presenter: Ze Yu Zhong
Co-authors: Ze-Yu Zhong;Xu Han

This paper investigates the impact of structural breaks in the factor structure on factor-augmented forecasting. We decompose the break in the factor loading matrix into rotational and shift components. To effectively utilize the pre-break data and maintain robustness against shift breaks, we propose a novel factor estimator that minimizes the L2 distance between pre- and post-break loading matrices through the rotation of factor estimates. We call this estimator the ‘’rotated factors” and analyze its the asymptotic

properties, along with two competing factor estimators, in the presence of different types of breaks. To leverage the respective advantages of each factor estimator in an automatic data driven way, we introduce a method that averages over sets of factor estimates using a leave-h-out cross-validationcriterion. Simulations demonstrate that combining different factor estimates through the proposed cross-validation averaging approach leads to improved forecasting performance compared to existing methods. Furthermore, we evaluate the effectiveness of our methods in an empirical application with US macroeconomic data and emphasize the importance of incorporating structural breaks into factor-augmented forecasting models.

A Benchmark Of Deep Learning- And Tree-Based Methods For Prosumer Electric Load Forecasting

Presenter: Barış Aydın
Co-authors: Barış Aydin;Kasım Zor

Over the past decade, the increasing prevalence of renewable-based decentralised power plants in the modern electric power systems has turned the term of a classical consumer into an innovative prosumer which may defined as an individual both producing and consuming electricity. This change in the course of unidirectional to bidirectional power flow in the active distribution networks has brought prosumer electric load forecasting to the forefront in energy management and planning of the ubiquitous microgrids. Deep learning (DL)- and tree-based methods are frequently employed for the prediction of nonstationary electric loads owing to the fact the aforementioned methods provide higher accuracies along with producing lower error metrics. This paper aims to present a benchmark of DL- and tree-based methods, namely convolutional neural networks (CNN) and extreme gradient boosted decision trees (XGBoost) for an hour-ahead prosumer electric load forecasting residing in California, USA. Additionally, climatological data of the prosumer building have been acquired from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) database of NASA. The obtained results unveiled that CNN model showed slightly better performance against XGBoost model in terms of coefficient of determination (R2: 91.181% vs 90.819%) and root mean squared error (RMSE: 0.195 kW vs 0.199 kW). Despite the fact that both models yielded very close R2 and RMSE values, the XGBoost model surpassed the CNN model with respect to computational time (0.08 s vs 57.67 s). The findings of this paper contribute to the existing literature on prosumer electric load forecasting and provide valuable insights for researchers and practitioners in the field.

News And Load: Quantification Of Economic And Social Drivers Behind Nodal Electricity Demand

Presenter: Yun Bai
Co-authors: Yun Bai;Simon Camal;Andrea Michiorri

Modelling the accuracy and uncertainty of nodal demand contributes to matching the supply and demand of electricity and reducing energy waste and carbon emissions. Electricity demand is impacted by the national economic and social environment, which is innovatively quantified and applied to the forecasting system in this study. We take five regions in the UK and the neighbouring country Ireland as the case study. By experimenting with the deterministic and probabilistic scenarios at multiple horizons, we found that built external factors improved forecasts from 5% to 10%. The improvements vary due to the development of the economy and energy industry, and the level of Internet penetration in the region. Then, we identified the economic-beneficial regions as East Midlands and Northern Ireland, and the social-beneficial regions as West Midlands and South West. The influential social factors are military conflicts, transportation, the global pandemic, regional development, and the international energy market.Keywords: Power systems, nodal electricity demand, economy and society, multi-horizon forecasting, probabilistic forecasting

Weather Effects In Energy Seasonal Adjustment: An Application To France Energy Consumption

Presenter: Marie Bruguet
Co-authors: Marie Bruguet;Arthur Thomas

In the context of climate change, the Paris Agreement aims to reduce global emissions, and thus one of France's objectives is to limit energy consumption.To this purpose, the French government has been engaged in several policies with both efficiency and sufficiency goals. Assessing changes in energy consumption, to studying the impact of climate-related policies are complex as many factors contribute to them, such as socio-economic and meteorological factors. To make reliable comparisons over time and assess the impact of socio-economic factors alone, it is necessary to correct observed consumption for seasonal and weather variations.This paper introduces a robust statistical methodology for determining the base temperature of Heating Degree Days (HDD) at an aggregated level. The approach extends to selecting the most reliable meteorological indicator for aggregated analysis. Applied to France, the methodology yields a country-specific base temperature and explores alternative meteorological indicators. It also confirms that the base temperature is not static over time or space, emphasizing the need for adaptive parameter adjustments. The proposed methodology employs regSARIMA modelling and incorporates a two-stage process, based on self-extracted threshold (SETAR) and penalized regression methods (LASSO), for selecting meteorological vectors among high-frequency meteorological data. This adaptive approach ensures a more precise determination of the aggregate HDD base temperature, enhancing the robustness of energy consumption adjustments.

Scalable And Efficient Mlp-Based Fully Parameterized Quantile Forecast With Uncertainty Estimates For Power Load Forecasting

Presenter: Anthony Faustine
Co-authors: Anthony Faustine;Lucas Pereira

The power grid is undergoing a fundamental transformation from centralized to decentralized structures, driven by the integration of Renewable Energy Sources, Energy Storage Systems, and the electrification of sectors like Electric Vehicles. This shift, while reshaping how power is generated and consumed, also presents significant challenges for managing and operating the grid, particularly due to the growing uncertainty in load demands. Accurate short-term forecasts are essential for maintaining the stability of the power grid, especially considering the increasing uncertainty in load demand. Driven by this growing uncertainty, there has been a noticeable shift towards using probabilistic load forecasting methods. However, many machine learning-based forecasting techniques, particularly those utilizing neural networks often overlook the importance of computational efficiency and scalability. These factors are paramount in reducing computation costs when deploying forecasting models at scale. To address the need for accurate, efficient, and scalable forecasting solutions, this work presents the Parameterized Quantile Forecast using simple yet powerful Multilayer Perceptron's (MLPQF). The proposed method leverages multilayer perceptron ability to capture complex non-linear relationships between historical data and meteorological variables addressing the need for accurate forecasting in dynamic power systems. Additionally, unlike traditional quantile regression, feedforward neural networks are employed to learn quantile probabilities and their corresponding distributions, providing valuable insights into the uncertainty and variability inherent in load demand predictions.The efficacy of the proposed architecture is demonstrated through empirical experimentation on three different use-cases: substations net-load forecasting, load demands forecasting, and PV-generation forecasting. The results indicate that the MLPQF approach generates accurate and reliable forecasts, capturing short-term and medium-term patterns. By utilizing NRMSE as a measure of forecasting accuracy, it is shown that the proposed approach achieves substantial improvement, ranging between 12% to 30% when compared to baselines and state-of-the-art models. In addition to forecasting accuracy, the, our approach approximates the speed of a persistence naïve model in term of inference speed , being 300 times faster compared to state-of-the-art benchmarks such as LSTM.

Application Of Data-Driven Weather Models In Power Systems

Presenter: Ada Canaydin
Co-authors: Ada Canaydin;Hussain Kazmi

The increasing integration of variable renewable energy resources, such as wind and solar, in the power systems has brought substantial changes in the dynamics of modern power systems, making them increasingly reliant on weather conditions. In response to this transformation, accurate weather forecasting has become vital for energy planning and grid management. These forecasts serve as an essential input for downstream energy forecast models, empowering grid operators to anticipate fluctuations in renewable energy generation, maintain grid stability, optimize economic efficiency within energy markets, and effectively manage renewable energy resources. Traditional weather prediction models, such as numerical weather prediction (NWP) model, rely on physical and mathematical principles to simulate future weather conditions, enabling forecasts up to 15 days in advance; however, their accuracy diminishes significantly for shorter timescales (less than a few hours) due to computational limitations and data assimilation delays.In recent years, with the rise of machine learning in weather forecasting, data-driven weather models leveraging reanalysis data have emerged as a potential avenue to address these challenges in the evolving energy landscape. The ability of these models to offer frequent weather updates at a low computational cost and to capture the nonlinear and complex nature of relationships holds promise for enhancing the predictive capabilities of downstream energy tasks, including renewable energy forecasting.In this study, we investigate the input data requirements, computational-cost, and accuracy of two data-driven weather forecasting models, namely Pangu-Weather and GraphCast, in capturing meteorological variables crucial for short-term renewable energy forecasting. Subsequently, we demonstrate the practical utility of these weather forecasts in renewable energy forecasting by (1) enhancing timeliness and (2) analyzing the correlation between input weather variables and the forecasted output of renewable energy generation. To validate our findings, we conduct a case study using real-world wind generation data from the Belgium transmission system operator. Our results illustrate the effectiveness of data-driven weather models in enhancing the accuracy and timeliness of renewable power forecasts.

Lies, Damn Lies, And An Illusionary Measure Of Renewable Energy Predictability: The Case Of Solar Energy Generation Forecasting In Great Britain

Presenter: Kevin Forbes
Co-authors: Kevin Forbes

The generally accepted metric of the forecast error in the renewable energy sector is calculated by dividing the forecast's mean absolute forecast error or RMSE by the capacity of the equipment used to generate the renewable energy. While no peer-reviewed literature supports this approach to measuring the error, renewable energy forecasters embrace this method because it makes the error seem small. Indicative of this, the forecasts of solar energy generation in Great Britain have an error of about 2% or so, which is amazing given that researchers who are strictly guided by science have conceded that their skill in forecasting solar radiation, a key driver of solar energy generation, is very low. This paper demonstrates that a forecast error metric weighted by capacity can create an illusion of predictability. With this result in mind, this paper presents a statistical methodology to improve the predictability of solar energy generation using data from the British power grid with the hope that the results will help dispel the illusion of predictability that currently prevails. The paper's analysis indicates that the existing forecasts do not fully reflect expected meteorological conditions. It is further observed that solar energy generation is highly volatile at times but also has a significant diurnal autoregressive pattern. An ARCH/ARMAX time series model is formulated based on those properties. One of the key modeling innovations is the implementation of ARCH-in-mean effects, which boosts predictive accuracy by capturing the information in the conditional variance. The model is estimated using 30-minute data from Jan 1, 2017, through Dec 31, 2022. The model is evaluated using out-of-sample data from Jan 1, 2023, to Dec 31, 2023. The period-ahead out-of-sample predictions have a weighted-mean-absolute-percentage-error (WMAPE) of about 4.6 %, substantially less than the approximately 11.1 % WMAPE associated with the solar energy forecasts used by the system operator over

the same period. While this finding may not be fully effective in dispelling the myth that the existing solar forecasts are accurate, it may serve the larger purpose of informing the broader forecasting community that the error metrics reported by renewable energy forecasters cannot be trusted.

Forecasting Austrias Small Hydropower Sector: A Transmission System Operators Perspective On Hydropower

Presenter: Claude Klöckl
Co-authors: Claude Klöckl

In alpine countries small hydro power units constitute a significant share of their electricity production capacity. In contrast, to large hydro power units, these are typically exempt from reporting their future production. Nonetheless, their production needs to be considered when planning the operation of the electricity grid. Forecasting small hydro is a largely uncovered forecast topic that mandates balancing relatively little local knowledge with the need provide accurate predictions over a wide range of geographical sites.For instance, this necessitates the Austrian Transmission System Operator APG perform daily 48h forecasts of all Austrian small hydro power units production in 15 min resolution. This poses a challenging forecasting problem, that requires a delicate modelling of an entire nations hydrological conditions, its production capabilities and the energy economic data available.We state the problem of predicting small hydropower from the perspective of a TSO, inform about traditional heuristic prediction methods and comment on our recent efforts to modernize the forecast system. We comment on possible avenues for improvement, based (a) on the forecast methodology and (b) on the inclusion of additional input data. Methodologically, we experiment with methods based on classical machine learning (XGBoost) and neuronal networks (LSTM). Furthermore, we discuss the problem of finding the relevant prediction inputs that range from generation at a given site, generation at sites upstream or weather data.

Physics-Informed, End-To-End Learning For Hierarchical Forecasting Of Renewable Energy Production

Presenter: Chotiya Mahittigul
Co-authors: Chotiya Mahittigul;Akylas Stratigakos;Mark O'malley

The integration of variable renewable energy (VRE) sources into power systems is one of the most effective strategies to combat climate change and decarbonize the electricity grid. Installing massive amounts of VREs, however, introduces a lot of technical challenges, one of them being the stochastic and intermittent nature of their output. This variability leads to a need for highly accurate look-ahead models, i.e. forecasting models, for VRE production at a future moment in time in the electricity grid. Having access to higher-quality forecasts can help inform system operators to manage the frequency response of a power system reliably and cost-effectively. Unfortunately, many of the forecasting models in the power grid are isolated, leading to a lack of coordination in their output predictions across both time and location. As such, there is a need for a singular set of forecasting models across the electricity grid that can coordinate with each other, hence, we propose a hierarchical forecasting model with nodes at different spatial locations and levels. The hierarchical model will provide a VRE production forecast at a local solar/wind farm, at a regional level with multiple solar/wind farms, and finally at a global level. A major challenge in implementing such a hierarchical structure is ensuring the coherency of the forecasts across different levels of aggregation. To this end, we will develop deep learning (DL) models that utilize implicit layers. The implicit layer embeds a Euclidean projection step that ensures forecast coherency and the full model is trained end-to-end. To ensure that our forecasts are aligned with the downstream task, we will also leverage physics-informed loss that models the grid physics. We will benchmark our forecasts with naive hierarchical aggregation methods and non-DL models. This work can be useful for many stakeholders who require forecasts at various spatial levels and can be extended beyond VRE production forecasts to power line congestion forecasts.

Evaluating Financial Tail Risk Forecasts With The Model Confidence Set

Presenter: Lukas Bauer
Co-authors: Lukas Bauer

This paper is the first to provide results on the finite sample properties of the Model Confidence Set (MCS) by Hansen et al. (2011) applied tothe asymmetric loss functions specific to financial tail risk forecasts, such as Value-at-Risk (VaR) and Expected Shortfall (ES). In this paper, wefocus on statistical loss functions that are strictly consistent in the sense of Gneiting (2011a). Our comprehensive simulation results show that, first,the MCS test keeps the best model more frequently than the confidence level 1 − in most settings. Second, it eliminates few inferior models forout-of-sample sizes of up to four years. Third, the MCS test shows little power against models that underestimate tail risk at the extreme quantilelevels p = 0.01 and p = 0.025, while the power increases with the quantile level p. Our findings imply that the MCS test may be suitable to narrowdown a set of competing models, but that it is not appropriate to test if a new model beats its competitors due to the lack of power.

An Optimized Algorithm For Multi-Period Forecasting Of Volatility And Value-At-Risk: Insights From Nifty Banks Index With Non-Performing Assets Impact

Presenter: Kunal Rai
Co-authors: Kunal Rai;Abhinav Anand;Soudeep Deb

In today's competitive and uncertain landscape, organizations strive to gain strategic advantage and ensure stability by forecasting across multiple horizons. This study investigates the impact of non-performing assets on the volatility and Value-at-Risk of the NIFTY Banks Index, an index that comprises of the most liquid and large capitalized Indian banking stocks, employing multi-period forecasting strategies – Direct (Dir) and Recursive (Rec). We employ the ARMA-GARCH forecasting technique with external regressors and with a skewed generalized error distribution to address the instability. A comprehensive exploration of the parameter space for ARMA-GARCH is conducted to identify optimal models for daily, weekly, and monthly volatility forecasts. These models are then utilized to project future log-return series and volatilities, enabling prediction of Value-at-Risk at the 90th, 95th, and 99th percentiles of the NIFTY Banks Index.To enhance computational efficiency without compromising accuracy, we selectively incorporate only those banks exhibiting similar behavior to the target bank during the training period as external regressors. This is achieved through time series clustering of all NIFTY banks, followed by dimensionality reduction via principal component analysis within each cluster, minimizing computational overhead.Our proposed approach, through brute search, exhaustively explores the parameter space, albeit at the cost of time. To address this limitation, we next introduce an optimization scheme inspired by the Greedy Randomized Adaptive Search Procedure (GRASP), which efficiently identifies optimal models, demonstrating comparable accuracy while significantly reducing computation time.Comparative analysis reveals that employing different models for distinct forecasting horizons outperforms a single-model approach in terms of accuracy, as corroborated by the Diebold-Mariano test. Furthermore, our findings indicate that NIFTY banks exhibit greater stability during periods of low non-performing assets.This study contributes to the limited literature on multi-period forecasting of volatility and Value-at-Risk, particularly within the context of the NIFTY Banks Index. The methodology presented herein holds promise for extension to other domains such as tourism and rainfall forecasting.

Forecasting Tail Risk Via Neural Networks With Asymptotic Expansions

Presenter: Yuji Sakurai
Co-authors: Yuji Sakurai;Zhuohui Chen

We propose a new machine-learning-based approach for forecasting Value-at-Risk named CoFiE-NN where a neural network (NN) is combined with Cornish-Fisher expansions (CoFiE). The new approach has two advantages. It can capture non-linear dynamics of high-order statistical moments thanks to flexibility of a NN while maintaining interpretability of the outputs by explicitly linking moments with the percentile

of distribution via Cornish-Fisher expansion. First, we explain the details of CoFiE-NN and discuss several potential applications. Next, we compare the performance of VaR forecasting based on CoFiE-NN with the conventional models using both Monte Carlo simulation and real data. We employ Long Short-Term Memory (LSTM) as a specification of a NN because LSTM has been successful in time series modeling. We show that CoFiE-NN shows the better performance when the sample period is shorter compared to the EGARCH-t model using simulated data. We then apply the CoFiE-NN for 30 assets across different asset classes, with a special focus on foreign exchange markets where high-order dynamics matter. We find that the CoFiE-NN tends to outperform the EGARCH model in the real data. Finally, we discuss how the forecast of VaR based on CoFiE-NN can be used for constructing an empirical proxy for tail risk. We discover that the only 20 percent of tail risk dynamics across 22 currencies is explained by one common factor. This is contrasting to the fact that 60 percent of volatility dynamics across the same set of currencies is explained by one common factor.

Boosted Value At Risk And Expected Shortfall – Application Of Gradient Boosting Machine Learning Models In Market Risk Estimation

Presenter: Michał Woźniak
Co-authors: Michał Woźniak

Effective estimation and management of market risk, particularly through metrics like Value at Risk (VaR) and Expected Shortfall (ES), are crucial for financial institutions, especially amidst global economic crises. Traditional statistical models often prove inadequate in such volatile environments, prompting the need for more robust methods. This study delves into the intersection of machine learning and statistics, focusing on Gradient Boosting Machines (GBM) and various probabilistic learning architectures to address this challenge.Divided into three phases, the research first identifies optimal probabilistic learning approaches for market risk models, spanning (1) quantile e.g. Quantile LightGBM and Quantile CatBoost; (2) distribution e.g. Natural Gradient Boosting for Probabilistic Prediction (NGBoost) and XGBoost for Location, Scale, and Shape (XGBoostLSS); and (3) volatility e.g. ARCH-like LightGBM. Subsequently, techniques such as transfer learning, application of multiparametric probability distributions, and extensive hyperparameter tuning are employed to enhance the selected models. Finally, the effectiveness of these models is compared against classical econometric approaches.The study's results are evaluated using a custom market risk benchmark encompassing 375 validation paths, with two-thirds dedicated to VaR and one-third to ES. The benchmark covers the following validation techniques: fulfilling regulatory requirements, forecasting adequacy, and capital effectiveness. It consists of five asset categories, each represented by five assets, across five testing periods of 250 days each. Testing horizons are one-step-ahead, with VaR evaluated at 1% and 2.5%, and ES at a 2.5% confidence level.The findings reveal that among tested approaches, the distributional forecast is the best probabilistic learning architecture for market risk modeling. Furthermore, the model utilizing the GBM concept – XGBoostLSS – is the top-performing choice among the considered models. The utilization of advanced machine learning techniques such as transfer learning, the introduction of highly skewed probability distributions (like Johnson's SU-distribution), and extensive tuning of hyperparameters have led to significant improvements in the machine learning models for VaR and ES. Finally, the shallow statistical learning models developed in this study have demonstrated their capability to outperform state-of-the-art econometric models in terms of effectiveness.

Forecasting To Target A Queue's Blocking Probability

Presenter: Casey Lichtendahl
Co-authors: Casey Lichtendahl

We introduce an approach for setting a queue's capacity when there are long lead times for expanding its capacity. A target for the system's blocking probability is the service level the system operator intends to deliver. We provide an approximate formula for the quantile of the demand distribution that delivers on the operator's service level objective. This quantile depends on the uncertainty in the offered load on the system and on the uncertainty in demand given the offered load. We use Borovkov's limiting

distribution for demand in an infinite-capacity queue to characterize the uncertainty in demand given the offered load. The marginal distribution of demand turns out to be a normal mean-variance mixture.

More Accurate, Interpretable Forecasting With Large-Scale, Pretrained, Hybrid Ml/Stat Models

Presenter: Pablo Montero-Manso
Co-authors: Weijie Shen, Casey Lichtendahl, Steve Thomas, Haoyun Wu, Chris Fry

We present a pretrained neural network that learns popular statistical modeling elements. By pre-training on a large synthetic dataset (trillions of data points), we show that the network achieves near optimal forecasting performance under a large variety of data generating processes, including ARIMA, Exponential Smoothing and Combinations. The network even outperforms the known estimators for these popular models. Additional experiments on real data further show the value of pretraining on large-scale, diverse synthetic series. These pretrained models are attractive for use in practice because they work off-the-shelf, they are blazingly fast, require no hyperparameter tuning and produce results that are more accurate but as interpretable as our well-regarded statistical models. When these properties are taken together, pretrained hybrid models have the potential to become the de-facto way of forecasting.

IJF Archivist: A GenAI chat interface to the IJF corpus through Google Cloud Platform

Presenter: Weijie Shen
Co-authors: Hussain Chinoy, Skander Hannachi, Chris Fry

Since ChatGPT, the Large Language Model (LLM) has changed how people access information due to its ease of use and vast knowledge base. In this experiment, we use Google Cloud API to build a conversational interface on top of the corpus of recent IJF articles, called IJF Archivist. Users with IJF access could use it as a way to learn state-of-the-art forecasting methodologies, summarize controversial topics with diverse opinions, and search for exact definitions of key concepts and formulas. We will discuss how the model is trained and served, and showcase its capabilities, user journeys, and future plans.

Starry-Net On Vertex Ai Platform

Presenter: Steve Thomas
Co-authors: Chris Fry, Weijie Shen, Casey Lichtendahl, Haoyun Wu, Jill Zhou, Ella Liu, Pablo Montero Manso

We introduce Starry-Net, a hybrid neural-network/statistical model for forecasting time series, on Google Cloud's Vertex AI Platform and demonstrate how to use it. This demonstration will cover the workflow lifecycle of training, inference, evaluation, and interpretation. Importantly, the training, inference, and evaluation stages are codeless in this environment. The user stores and points to their data in BigQuery or Google Cloud Storage and then launches the pipeline with a few clicks. Because Starry-Net learns the TBATS model and other well-known statistical elements, the pipeline yields traditional decomposition plots which many practitioners rely on when interpreting forecasts. This feature has helped us gain trust internally and convince partner teams to adopt Starry-Net. We think Starry-Net on Vertex AI will be an ideal tool for Vertex AI users who are familiar with statistical models but would like to have the accuracy gains that come from neural networks.

Efficient Modeling Of Seasonal Patterns In Global Neural Networks

Presenter: Anastasios Kaltsounis
Co-authors: Anastasios Kaltsounis;Evangelos Spiliotis;Vassilios Assimakopoulos

Machine learning and especially neural networks (NNs) have dominated the field of forecasting in a series of different applications. In most cases, the data used for the networks' training originate from the same source (e.g. SKUs, Electricity Consumption Data, etc.) and therefore share similar seasonal patterns that can be easily identified by a NN. However, in some applications forecasts are required for a set of time series of different domains, sources, and frequencies, rendering the modeling of seasonality challenging. Moreover, seasonalities may often be expressed at multiple levels, complicating things further. In order to facilitate the learning process of NNs, researchers often process the data before importing them to the networks. This solution is based however on assumptions and can result in suboptimal results in terms of forecast accuracy. Motivated by this limitation, we investigate approaches which automatically account for the seasonalities of each time series. Insights of this work can be used for the development of fundamental models in the field of forecasting.

Beyond Numbers: Forecasting Stock Volatility Through Image-Based Deep Neural Networks

Presenter: Artemios Anargyros Semenoglou
Co-authors: Artemios-Anargyros Semenoglou;Javier Bas;Adam Clements

This study introduces a novel method for predicting stock index volatility direction using image representations of historical open/high/low/close OHLC price data combined with advanced Machine Learning (ML) techniques. Driven by the need for accurate volatility forecasts in financial risk management, portfolio optimization, or derivatives pricing, our approach seeks to address the relatively unexplored area of directional volatility prediction. Despite the effectiveness of traditional models like the Heterogeneous AutoRegressive (HAR) model in identifying volatility patterns, the advent of ML has opened new possibilities for enhancing forecast precision. Moreover, our proposed framework is novel in that it shifts from traditional numerical estimates to leveraging image representations of volatility data. These images are processed through ensembles of deep Convolutional Neural Networks (CNN) to forecast the direction of stock index volatility. This method not only transcends the limitations of conventional data analysis but also introduces a pioneering perspective, leading to more accurate forecasts. Preliminary results show that forecasts from CNN based on the visual representation of historical volatility provide more accurate forecasts than those based on the traditional numerical volatility data. In light of these results, our approach offers promising avenues for both academic research and practical applications in financial analysis and risk management. This paper contributes to the literature by presenting a novel methodology for stock index volatility forecasting. It unveils the untapped potential of image-based data analysis in financial markets, setting a new benchmark for future research and practice in volatility forecasting.

The Impact Of Dataset Similarity And Diversity On Transfer Learning Success

Presenter: Claudia Ehrig
Co-authors: Claudia Ehrig;Catherine Cleophas;Germain Forestier

Models pretrained on specific similar or generally diverse source datasets have become pivotal in enhancing the efficiency and accuracy of time series forecasting on target datasets through transfer learning. While benchmarks validate the model's generalization performance on various target datasets, little structured research provides similarity and diversity measures that explain which characteristics of the source and target data lead to transfer learning success.Our study pioneers in systematically evaluating the impact of source-target similarity and source diversity on forecasting outcomes in terms of accuracy, bias, and uncertainty estimation. We investigate these dynamics using pretrained neural networks across five public source datasets, applied in zero-shot and fine-tuned forecasting modalities on five distinct target datasets. The target datasets include sales data from two real-world wholesalers. We assess dataset similarity through two feature-based and one shape-based approaches, while evaluating diversity using feature-based measures and visualizing it in PCA space.Our findings reveal that similarity and diversity metrics significantly enhance forecasting accuracy, and similarity helps reduce bias. This research underscores the strategic advantage of transfer learning in time series forecasting, showcasing its capacity to harness complex data

relationships for improved predictive outcomes.

Infinite Forecast Combinations Based On Dirichlet Process

Presenter: Feng Li
Co-authors: Feng Li;Yinuo Ren;Yanfei Kang;Jue Wang

Forecast combination integrates information from various sources by consolidating multiple forecast results from the target time series. Instead of the need to select a single optimal forecasting model, this paper introduces a deep learning ensemble forecasting model based on the Dirichlet process. Initially, the learning rate is sampled with three basis distributions as hyperparameters to convert the infinite mixture into a finite one. All checkpoints are collected to establish a deep learning sub-model pool, and weight adjustment and diversity strategies are developed during the combination process. The main advantage of this method is its ability to generate the required base learners through a single training process, utilizing the decaying strategy to tackle the challenge posed by the stochastic nature of gradient descent in determining the optimal learning rate. To ensure the method's generalizability and competitiveness, this paper conducts an empirical analysis using the weekly dataset from the M4 competition and explores sensitivity to the number of models to be combined. The results demonstrate that the ensemble model proposed offers substantial improvements in prediction accuracy and stability compared to a single benchmark model.

Dynamic Forecast Combination Using Point Or Density Forecasts

Presenter: Maddie Smith
Co-authors: Maddie Smith;Nicos Pavlidis;Adam Sykulski

It is often the case that decision makers are presented with multiple forecasts for the same variable, produced perhaps from different forecasting models or experts. While attempting to identify a single 'best forecast' does offer a valid approach, favourable performance is often achieved through combining the available forecasts in some way. As such, forecast combination presents a ubiquitous problem for decision makers in a wide array of fields, ranging from economics to environmental applications and epidemiology.Although perhaps a seemingly simple problem, the combination of forecasts can prove difficult due to issues such as correlation, little historical data, and social influence. In this talk, we propose and contrast two Dynamic Linear Model (DLM)-based forecast combination methods; one for the combination of point forecasts, and the other utilising methodology from the density forecast combination literature. The proposed methods are online and dynamic, allowing combination weights to evolve as more observed data become available. We utilise simulations to demonstrate how the methods deal effectively with highly correlated forecasters and changing forecaster quality throughout time. Both of the proposed approaches can be extended to deal with the often less explored but highly relevant topic of missing forecaster data. In the point forecast combination framework, we introduce a discount factor parameter and discuss the possible role of adaptive discounting to deal with such cases. We compare this with the density combination approach, and discuss the merits and challenges of the two. Our methods are compared with a selection of benchmark forecast combination methods from the point and density forecast combination literature, and are shown to exhibit superior forecasting performance in both simulation studies and an environmental application.

On The Economic Benefit Of Complete Subset Regression For The Multivariate Har Model

Presenter: Andrey Vasnev
Co-authors: Andrey Vasnev;Adam Clements

Forecasts of the covariance matrix of returns is a crucial input into portfolio construction. In recent years, multivariate versions of the Heterogenous AutoRegressive (HAR) models have been designed to utilise realised measures of the covariance matrix to generate forecasts. This paper shows that combination

forecasts using a subset regression approach provide more stable coefficients estimates, and as a result more stable forecasts and lower portfolio turnover. The economic benefits of the combination approach become crucial when transactions costs are taken into account. This combination approach also provides benefits in the context of direct forecasts of the portfolio weights. Economic benefits are observed at both 1-day and 1-week ahead forecast horizons.

Forecast Using A Machine Learning System To New Products In A Portuguese Brewery

Presenter: Ricardo Galante
Co-authors: Ricardo Galante;Teresa Alpuim

The introduction of new products is an important point of growth for any brewery, yet the inherent uncertainty of consumer preferences poses a significant challenge. Traditional forecasting methods may struggle to accurately predict demand in this dynamic market. This research explores the potential for machine learning (ML) systems to enhance demand forecasting for new products within the Portuguese brewery industry.We propose a framework utilizing historical sales data, market trends, and relevant external factors such as seasonality and economic indicators. A suite of ML algorithms, including cluster analysis, regression models, decision trees, and potentially neural networks, will be evaluated for their predictive performance. The study aims to:• Identify the most important predictors of demand for new brewery products.• Compare the accuracy of various ML algorithms in this forecasting context.• Develop a practical ML-based forecasting system tailored to the Portuguese brewery sector.This research provides breweries with data-driven insights into demand for new products, aiding in decision-making, production planning and, ultimately, improving resource allocation and profitability.Keywords: New Product Forecasting, Cluster Analysis, Gradient Boosting, Demand Forecasting, Machine Learning

Navigating The Future In E-Commerce: An Ml-Based Approach To Sales Forecasting

Presenter: Eryk Lewinson
Co-authors: Eryk Lewinson

In today's ever-changing business world, predicting future trends is crucial for businesses. In e-commerce, accurate forecasting allows companies to adapt to market changes and shifting consumer preferences by, for example, optimizing supply chains, anticipating seasonal trends, adjusting pricing, reducing costs, and managing risks. It also helps to offer the customers speedy delivery of the goods they need and a person to pick up their phone call if something does ever go wrong.

Forecasting Walmart Ecommerce Demand

Presenter: Slawek Smyl
Co-authors: Slawek Smyl;Johann Posch

We will report on a new neural network-based forecasting system for Ecommerce demand in Walmart. We will describe methods we used to deal with several challenges, including massive amount of data and its prevailing sparsity.We will also describe computational frameworks we developed to manage development (in particular, hyperparameter tuning) and the large scale serving/inference.

On Forecast Stability

Presenter: Christoph Bergmeir
Co-authors: Christoph Bergmeir;Rakshitha Godahewa;Zeynep Erkin Baz;Chengjun Zhu;Salvador Garcia;Dario Benavides

Forecasts are typically not produced in a vacuum but in a business context, where forecasts are generated on a regular basis and interact with each other. For decisions, it may be important that forecasts do not change arbitrarily, and are stable in some sense. However, this area has received only limited attention in the forecasting literature. In this paper, we explore two types of forecast stability that we call vertical stability and horizontal stability. The existing works in the literature are only applicable to certain base models and extending these frameworks to be compatible with any base model is not straightforward. Furthermore, these frameworks can only stabilise the forecasts vertically. To fill this gap, we propose a simple linear-interpolation-based approach that is applicable to stabilise the forecasts provided by any base model vertically and horizontally. The approach can produce both accurate and stable forecasts. Using N-BEATS, Pooled Regression and LightGBM as the base models, in our evaluation on four publicly available datasets, the proposed framework is able to achieve significantly higher stability and/or accuracy compared to a set of benchmarks including a state-of-the-art forecast stabilisation method across three error metrics and six stability metrics.

Predict. Optimize. Revise. On Forecast And Policy Stability In Energy Management Systems

Presenter: Evgenii Genov
Co-authors: Evgenii Genov;Christoph Bergmeir;Julian Ruddick;Thierry Coosemans

This paper addresses the challenge of integrating forecasting and optimizationin energy management systems, focusing particularly on the impacts of switchingcosts and the role of forecast accuracy and stability. It introduces a novelframework for analyzing online optimization problems considering forecastaccuracy and stability of both deterministic and probabilistic approaches. Thestudy explores the concept of time-coupling in decision-making and performanceimplications of switching costs, and forecast errors, conversely termed as"switching incentives". Through empirical evaluation and theoretical analysis,the research reveals the nuanced balance between forecast accuracy, forecaststability, and switching costs in shaping policy performance in energymanagement. It proposes a metric for evaluating probabilistic forecast stabilityand examines the effects of forecast accuracy and stability on optimizationoutcomes using a case study from the Citylearn 2022 competition. The findingssuggest that switching costs significantly influence the trade-off betweenforecast accuracy and stability, highlighting the importance of integratedsystems that facilitate collaboration between forecasting and operational unitsfor improved decision-making. It is shown that for particular cases it isadvantageous to commit to a policy for a longer period, than to update it atevery time step. Results also indicate a correlation between forecast stabilityand policy performance, suggesting that stable forecasts can mitigate the impactof switching costs.

Optimizing For Forecast Stability In Distribution-Free Probabilistic Forecasting

Presenter: Jente Van Belle
Co-authors: Jente Van Belle;Honglin Wen;Wouter Verbeke;Pierre Pinson

Rolling origin forecast instability can be defined as the variability in forecasts for a specific time period induced by updating the forecast for this period when new observations become available, i.e., as time passes. Forecast updates are considered to result in both benefits and costs. The benefits are caused by an increase in forecast quality due to a shorter forecast horizon, while the costs stem from induced forecast instability, which may lead to costly changes to plans formulated based on the forecasts. Additionally, forecast instability can erode trust in the forecasting system, potentially prompting unwarranted judgmental adjustments by users.Methodologies exist to enhance forecast stability without compromising forecast quality in both point and Gaussian probabilistic forecasting settings. Furthermore, application of these methodologies can lead to improvements in both forecast stability and quality, suggesting that they can also serve as time-series-specific regularization mechanisms.In this paper, we aim to integrate forecast stability alongside quality into the optimization of distribution-free probabilistic time series forecasts to extend the above findings to the distribution-free setting. To achieve this goal, we propose a method to generate stabilized forecasted conditional quantile functions. These quantile functions are modeled using

linear isotonic regression splines, with the parameters learned through training a neural network using a discretized approximate version of the continuous ranked probability score as the loss function. Unlike approaches based on parametric probability density functions or those forecasting only a fixed set of quantiles, our proposed method offers the flexibility to produce full density forecasts for various output distributions without explicit specification, while also preventing quantile crossing. Furthermore, the conditional quantile function approach for characterizing density forecasts provides the essential flexibility to optimize forecast stability alongside quality. Specifically, it enables us to control the level of dissimilarity allowed at each forecast update, while also offering the flexibility to place varying importance on different parts of forecast distributions (e.g., central part vs. tails).We empirically demonstrate the effectiveness of our proposed approach on multiple datasets that exhibit different statistical properties.

Optimal Forecasting Under Parameter Instability

Presenter: Wenying Yao
Co-authors: Wenying Yao;Yu Bai;Bin Peng;Shuping Shi

This paper addresses three issues associated with local estimator in forecasting models that are affected by parameter instability. We first demonstrate the consistency of the local estimator under various types of parameter instability. Then, we analyze the choices of weighting function and tuning parameter associated with the local estimator. We propose a selection procedure for the tuning parameter and prove its asymptotic optimality of the tuning parameter selection procedure. Lastly, we provide an analytical criterion on the choice of weighting function. The theoretical results are examined through an extensive Monte Carlo study and four empirical applications on forecasting inflation, growth and inflation shocks, house price changes and bond returns.

Forecasting Quarterly National Accounts: A Comparative Analysis Of Institutional And Model-Based Forecasts

Presenter: Matteo Neufing
Co-authors: Matteo Neufing;Katja Heinisch

Annual growth forecasts are important figures for policy-making and serve as a benchmark for many forecasters. However, these forecasts typically rely on quarterly estimates, which do not get much attention and are hardly known. Therefore, this paper provides a detailed analysis of multi-period ahead quarterly GDP growth forecasts for German national accounts with respect to first-release and current-release data. Furthermore, we shed some light on the forecasting performance with respect to GDP components, such as private consumption or investment. The professional forecasters' predictions are evaluated against mean forecasts and those based on traditional autoregressive models as well as vector-autoregressive models. In the later case, higher-frequency indicators are used to predict GDP.Based on a novel dataset comprising quarterly institutional forecasts, our findings indicate that institutions consistently outperform empirical models in short-term forecasting, but there's no significant improvement beyond two quarters. Overall, forecast revisions and forecast errors are analyzed and the results show that the forecasts are not systematically biased. The forecast performance also varies across institutions and GDP components. The paper provides evidence that there is still room for improvement in forecasting techniques both for nowcasts but also forecasts up to eight quarters ahead.

Nowcasting Gdp: What Are The Gains From Machine Learning Algorithms?

Presenter: Rolf Scheufele
Co-authors: Rolf Scheufele;Milen Arro-Cannarsa

We compare several machine learning methods for nowcasting GDP. A large mixed-frequency data set is used to investigate different algorithms such asregression based methods (LASSO, ridge, elastic net), regression trees (bagging, random forest, gradient boosting) and SVR.As benchmarks we use univariate

models, a simple forward selection algorithm anda principle components regression. The analysis accounts for publication lags andtreats monthly indicators as quarterly variables combined via blocking.Our data set consists of more than 1100 time series. For the period after the GreatRecession, which is particularly challenging in terms of nowcasting, we find that allconsidered machine learning techniques beat the univariate benchmark and the forward subset selection algorithm up to 28 % in terms of out-of-sample RMSE. Ridge, elastic net and SVR are the most promising algorithms in our analysis and they outperform principle components regression on a significant level.

Enhancing Quarterly Gdp Growth Forecasting Across Sectors With General Regression Neural Networks

Presenter: Katja Heinisch
Co-authors: Katja Heinisch;Boris Kozyrev

Neural networks offer a valuable complement to traditional time series econometric forecasting methods, providing enhanced predictive capabilities. One drawback often associated with neural networks is the interpretation of specific forecasts. However, a specific type of neural network known as the General Regression Neural Network (GRNN) addresses this limitation by offering intuitive interpretation and ease of estimation.The GRNN approach leverages patterns from the most recent realizations to predict future outcomes. By identifying similar historical patterns and extrapolating those into the near future, the GRNN can effectively forecast the next realization. This paper aims to expand the GRNN approach to nowcast quarterly growth rates for German gross domestic product (GDP) and all gross value added (GVA) sectors. We use a large data set of higher-frequency indicators to predict the current state of the economy from different perspectives.The GRNN approach can be extended to incorporate information from multiple time series, enhancing its forecasting accuracy. Hence, individual forecasts based on indicators are then aggregated to derive the final sectoral growth rate estimate. Different forecast combination techniques are used to aggregate the best forecast efficiently. Furthermore, besides forecasting individual sectors and the total GVA directly, we adopt a bottom-up approach to synthesize the forecasts for individual sectors into the assessment of the total GVA.An empirical evaluation comparing the GRNN forecasts with those generated by autoregressive models, mean forecasts, and professional forecasters' predictions, for the period 2015Q1 to 2023Q4, reveals a superior performance of the sectoral GRNN models across most sectors.For enhancing the interpretability and explainability of GRNN forecasts, employing Shapley value decomposition reveals the individual contribution of each variable to the final prediction, and helps policymakers to assess the current state of sectors.

Business Cycle Dynamics After The Great Recession: An Extended Markov-Switching Dynamic Factor Model

Presenter: Catherine Doz
Co-authors: Catherine Doz;Laurent Ferrara;Pierre-Alain Pionnier

As illustrated by the Great Recession and the Covid pandemic, macroeconomists need to account for sudden and deep recessions, as well as fluctuations in macroeconomic volatility and trend GDP growth. In this respect, we put forward an extended Markov-Switching Dynamic Factor Model by incorporating switches in volatility and time-variation in trend GDP growth. We show that volatility switches improve the detection of business cycle turning points, that the Great Recession led to a temporary increase in volatility, and that the US trend GDP growth has been declining since the early 2000s. Information criteria, marginal likelihood comparisons and improved real-time performance support our model. Our results show that the model accommodates well the pre- and post-Covid times.

Frequency Identification In Singular Spectrum Analysis

Presenter: Gabriel Martos Venturini
Co-authors: Gabriel Martos Venturini;Diego Fresoli;Pilar Poncela Blanco

Singular Spectrum Analysis (SSA) is a nonparametric technique for signal extraction. The goal of SSA is the decomposition of one or several time series in their underlying unobserved components: Trend, cycles, possibly modulated amplitude, and noise. A widely recognized tool for extracting these components is the Singular Value Decomposition (SVD) of the trajectory matrix associated to the time series. However, once the different components are extracted, we need to identify their frequency of oscillation as the aim of the SVD is not linked to frequency identification but to maximize variability. This paper proposes a new twist to SSA changing the framework from the time domain to the frequency domain when analyzing time series data. In this way, we are able to tackle frequency identification and variability maximization simultaneously. We compare our method to existing alternatives through simulations and real data examples.

Vulnerable Regional Growth: The Case Of Spain

Presenter: Eva Senra
Co-authors: Eva Senra;Martín Llada;Pilar Poncela

Regional heterogeneity induces differences in growth and renders some regions more vulnerable to downside risks. Economic policy requires the assessment of vulnerability at a regional level to mitigate the effects of business cycle downturns. On the basis of monthly regional indicators and quantile regressions, this paper quantifies the risks of downside shocks associated to the business cycle, uncovering vulnerable aspects of regional growth.

To Forecast Or Not To Forecast: Evaluating The Predictive Power Of A Random Forest Against A Gravity Model

Presenter: Costanza Bosone
Co-authors: Costanza Bosone;Paolo Giudici

The age-old debate between Trees and Linear models sets the stage for our study, where we aim to compare the predictive power of a gravity model vis-à-vis a random forest in forecasting trade flows. The gravity model, a milestone in trade analysis, is explicitly designed to analyse trade flows between i and j as a result of their masses and frictions. On the other hand, the random forest harnesses the power of machine learning, employing a series of splitting rules to construct decision trees that are then aggregated to form a predictive model. Our research builds upon the established framework of traditional gravity settings, employing trade flows as the target variable in both models. We conduct out-of-sample analyses for the years 2020, 2021 and 2022 to assess the comparative performance of the random forest against the gravity model. Through the Diebold-Mariano test, we show that there is no statistically significant difference between the predictive performances of these approaches in any of the three out-of-sample tests. Our findings affirm that the random forest can forecast trade flows with the same level of accuracy as the gravity model. This discovery holds promising implications. The random forest showcasing superior predictive capabilities might usher in tangible benefits by streamlining data input processes, bypassing the need for bilateral data entry inherent in the gravity model. Moreover, the adoption of the random forest in Python offers a cost-effective alternative to other resource-intensive software solutions.

Optimal Predictor And Transformation Selection For Macroeconomic Forecasting Using Variable Importance In Random Forests

Presenter: Maurizio Daniele
Co-authors: Maurizio Daniele;Philipp Kronenberg;Tim Reinicke

In this paper, we propose a novel recursive group variable importance measure in random forests (RF) to select the most relevant indicators for predicting key macroeconomic variables. In contrast to existing RF based importance measures, our method enhances the modeling flexibility by accounting for

general types of time series structures in economic data.In an out-of-sample forecasting experiment using a large dimensional macroeconomic dataset based on the FRED-MD database, we illustrate significant improvements in forecasting US inflation, when employing our RF based selection approach for extracting the optimal predictors and data transformations compared to existing selection methods relying on conventional regularization techniques, e.g. the lasso and elastic net. Moreover, our findings reveal that optimal variable transformations uniformly enhance the predictive accuracy of various modeling approaches, including regularization methods, (dynamic) factor models, neural networks, and random forests. The observed forecasting improvements highlight the importance of considering alternative transformations beyond the conventional choices recommended in the FRED-MD dataset. Furthermore, we provide theoretical insights on our RF based selection criterion in an additive model framework.

Boosting Xgboost. Using The Panel Dimension To Improve Machine-Learning-Based Forecasts In Macroeconomics

Presenter: Johannes Frank
Co-authors: Johannes Frank;Jonas Dovern

The short time dimension of commonly used macroeconomic data sets poses a challenge for estimating machine learning (ML) models to monitor business cycles in real-time. This is unfortunate because they are (potentially) great tools to model non-linearities and use very large data sets for business cycle monitoring. Existing studies that use ML algorithms for macroeconomic forecasting mostly focus on time-series applications and do not consider panel data to increase the amount of information available for the training of the models. In this paper, we use data for individual states and the whole US economy to increase the training data set. The idea is that dynamics between variables and across time at state levels are similar to each other and to the dynamics at the national level. Using extreme gradient boosting (XGBoost), a ML algorithm for tree boosting, we analyze if moving to the panel setup improves forecast accuracy. We use data pooling in combination with weight sharing that allows for some degree of heterogeneity in the cross section. This allows for regularization of parameters/weights while at the same time guarding against overfitting. We find that such "soft" pooling approach helps improving forecast performance at the national level and reducing the variance as well as the average of the distribution of RMSEs across states. Thus, by using information on individual regions in a panel data setup along with appropriate regularization approaches, the issue of data scarcity can be overcome. In the further course of our project, we will verify the robustness of our results through various checks and use state-of-the-art statistical tests of relative predictive accuracy to test the statistical significance of our results. In addition, we aim to investigate whether using disaggregated Google Trends data can improve forecast performance further in the panel data setup. We also plan to implement other ML algorithms that might be more suitable to use information from the panel data efficiently.

Simulation Of Electricity Forward Curves

Presenter: Marina Dietze Monteiro
Co-authors: Marina Dietze Monteiro;Alexandre Street;Davi Michel Valladão;Stein-Erik Fleten

Hedging against spot price volatilities becomes increasingly important in deregulated power markets. Therefore, being able to model electricity forward prices is crucial in a competitive environment. Electricity differs from other commodities due to its limited storability and transportability. Furthermore, its derivatives are associated with a delivery period during which electricity is continuously delivered, implying on referring to power forwards as swaps. These peculiarities make the modeling of electricity contract prices a non-trivial task, where traditional models must be adapted to address the mentioned characteristics. In this context, we propose a novel semiparametric structural model to compute a continuous daily forward curve of electricity through maximum smoothness criterion. In addition, elementary forward contracts can be represented by any parametric structure for seasonality or even for exogenous variables. Our framework acknowledges the overlapped swaps and allows an analysis of arbitrage opportunities observed in power markets. With the estimated daily elementary prices, we perform a principal component analysis and than

generate scenarios multiple steps ahead for each PCA component. With the forecasts for the components, we can obtain scenarios for the elementary prices and than for the traded swaps. This information can be crucial in the decision process of multiple agents and addresses a complex problem of the electricity sector.

Kernel-Based Approach For Very Short Term Forecasting Of Electricity Prices In The German Continuous Intraday Market

Presenter: Andrzej Puć
Co-authors: Andrzej Puć;Joanna Janczura

The European continuous intraday electricity markets have experienced a period of dynamic growth in recent years. However, this growth did not directly translate into the increase in scientific works on forecasting in those markets. While methods of point forecasting of the prices on auction power markets are widely known, the continuous intraday price point forecasting seems to be a niche task. This task was also shown to be challenging, as the naïve forecast, being the last known price, is a strong benchmark in the very short term forecasting tasks. To fill this gap, we develop a kernel-based algorithm that allows us to perform a point forecast of continuous market prices. Method is based on kernel regression which uses a combination of Laplace and Gaussian kernels. The combination allows to put emphasis on the most important variables and is fine-tuned to fit the needs of a very short term forecasting task. The forecasting is performed for the price at a given horizon before the delivery, with a forecasting horizon extending from hours to minutes before trading. We validate our approach in a case study for the German continuous intraday market. Validation window comprises all trades performed in 2020, so the year impacted by the COVID-19 pandemic. We show that such a method outperforms the naïve forecast and its low computational complexity makes it feasible for a very short term forecasting. We also compare the suggested kernel method with solutions classically used in forecasting: LASSO and Random Forest. Suggested method can also be applied to any continuous power market and extended to generate probabilistic forecasts, which can then be applied for building short-term strategies.

Forecasting Supply And Demand Curves: A Functional L1 Approach

Presenter: Nabangshu Sinha
Co-authors: Nabangshu Sinha;Florian Ziel

Electricity market curve forecasting is crucial for market stability and efficiency. Accurate predictions inform decisions on energy production, distribution, and pricing. Demand forecasting analyzes historical consumption patterns and factors like weather, economics, and demographics. Supply curve forecasting predicts electricity availability from various sources, optimizing resource allocation and grid management. Accurate forecasting mitigates risks and optimizes investments.Forecasting electricity market curves presents several challenges, mainly due to the heterogeneous nature of the curves. The curves exhibit variability in the number of bid points and the curves' respective supports across different hours and days. Previous methods, such as the X-model and FAR, have been utilized for curve forecasting, with a primary focus on price prediction. However, these approaches tended to overlook local curve details, particularly around the market equilibrium price, as they prioritized price forecasting. Additionally, these methods typically involve preprocessing to standardize the curves into a common functional framework. In our study, we forecast these curves without explicit transformations. Essentially, our models utilize the raw curve data, without any preprocessing or parameterization, as the input to our models.We introduce a custom error function, employing the L1 loss function to quantify the area between curves. It computes loss only within a narrow vicinity of the intersection price, crucial from market participants' perspective. Two autoregressive models segment and forecast future curve segments, reconstructing the future curve. The "2-piece model" divides the curve into two segments: one spanning from negative infinity to -500 (inclusive), and the other from -500 to positive infinity. While the "3-piece model" segments it into three parts: from -infinity to -500, -500 to 0, and 0 to +infinity. Comparative analysis shows their superiority over naive models. To enhance interpretability, our models determine autoregression coefficients with respect to load, solar output, and wind output variables, enhancing. Although our results are primarily for supply curves, the same framework

can be used for modelling demand curves too.

"One Out Of Many": Consolidating A Long-Term Trend Forecast For Investing In Energy Commodities

Presenter: Fernanda Diaz-Rodriguez
Co-authors: Fernanda Diaz-Rodriguez;Maria-Dolores Robles-Fernandez

This paper adds to the few studies on the annual horizon of energy price forecasting. Accurately predicting trends in energy commodity prices is vital for economic and financial stability. However, the trend itself is difficult to capture, making single forecasts unreliable. This study proposes a novel approach combining multiple trend forecasts for Crude Oil prices and assesses its effectiveness in long-term investment strategies. Combining forecasts is known to improve accuracy and reduce risk compared to relying on a single method. This study compares five individual trend estimation methods and seven combination methods. The results show that combining forecasts with a novel method called METS performs best, minimizing errors and remaining stable across market conditions. Furthermore, using estimated trends outperforms using actual prices for long-term forecasting, highlighting the limitations of short-term price predictions. Denoising the data as in hybrid forecasting methods may not be necessary, especially for long-term forecasts.Keywords: Energy price, Crude oil, Commodities, Forecasting, Forecasts combination, Long-term trend, Meta-prediction, Ensemble, Wavelet decomposition, Feedforward neural network, PCA, Machine learning, Optimal weights, Time series.

Structured Additive Stacking For Forecasting In The Energy Domain

Presenter: Euan Enticott
Co-authors: Euan Enticott;Matteo Fasiolo;Christian Capezza;Yannig Goude;Biagio Palumbo

Ensemble techniques, such as stacking regressions, are routinely used to improve accuracy in many predictive applications. However, most existing ensemble methods do not adaptively adjust the experts' weights based on relevant covariates. We develop a framework for stacking predictive probability densities, where the weights are controlled via non-linear smooth effects to produce an adaptive predictive mixture distribution. To improve forecasting accuracy while providing statistical and computational scalability in the size of the ensemble, we propose weighting structures that exploit context-specific relationships between the experts. Further, we develop methods to fit the model and perform inference in an empirical Bayesian framework, which enables us to quantify uncertainty and perform model selection using standard tools. This work was motivated by the need to provide accurate probabilistic forecasts of electricity demand at a low level of aggregation and price predictions during unstable periods. In particular, electricity demand at the individual household level is characterised by a low signal-to-noise ratio and abrupt changes which are difficult to capture with a single model. Similarly, electricity spot prices can transition between calm and turbulent regimes, due to technical, socio-economical or weather-related factors. In this talk we will show how the new stacking methods exploit the information provided by exogenous factors and recent demand or price history to adapt the weights of an ensemble of models, thus improving forecasting performance.

Parsimonious And Unconstrained Dynamic Covariance Matrix Modelling

Presenter: Xinyue Guan
Co-authors: Xinyue Guan;Matteo Fasiolo;Haeran Cho

The effective management of renewable electrical power sources, such as solar and wind, is heavily depend on forecast accuracy.Such forecasts are most useful when provided in a probabilistic format that can support decision-making under uncertainty. In particular, when dealing with multiple forecasts corresponding to different locations or times, it is critical to assess whether forecasting errors are likely to compound or mitigate each other. For instance, if a forecast error persists at one wind farm, can we anticipate a corresponding error of similar magnitude and direction at a neighboring wind farm? This

question serves as a primary motivation for our research into dynamic covariance matrix modeling.Two primary obstacles are encountered in covariance matrix modeling: the quadratic growth in the number of matrix elements that must be modelled as the number of response variables increases, and the positive definiteness constraint. In our approach, we fulfill the positive definiteness constraint by employing unconstrained parameterizations obtained through matrix decomposition techniques, such as the modified Cholesky decomposition. Additionally, to manage the number of variables involved in modeling, we adopt covariance functions that can be controlled via a limited number of parameters. The flexibility of such parsimonious covariance matrix models is then enhanced by modelling their parameters via non-linear functions of explanatory variables, such as forecast lead time and location.

The Importance Of Correct Model Specification: A Regime Switching Garch Midas Approach

Presenter: Jie Cheng
Co-authors: Jie Cheng

Events such as pandemics, changes in government policies and wars result in structural breaks in many areas including oil markets. RS GARCH MIDASmodels, which consider both structural changes and macroeconomic factors affecting the oil prices have been studied by very few authors where theyassumed innovations are normally distributed. In this article, we consider different error distributions to analyse how effective they are in capturingthe characteristics of oil returns compared to Normal innovations. In a Monte Carlo simulation, we investigate how model misspecification affectsthe estimation results and find that misspecified models have greater bias, overestimation in the long-term component and problems with identificationof two volatility regimes. The results obtained in simulation are also confirmed in an empirical application to WTI crude oil returns. Finally, theforecast performance of RS GARCH MIDAS-t model is compared with various competitor models.

Deciphering Market Moods: The Conditional Impact Of Investor Sentiment On Financial Market Through A Quantile Lens

Presenter: Szymon Lis
Co-authors: Szymon Lis;Robert Ślepaczuk;Paweł Sakowski

This study delves into the intricate dynamics between investor sentiment and its impacts on stock returns and trading volumes, challenging the efficient market hypothesis (EMH) by integrating behavioral finance insights. Utilizing Fama-MacBeth and quantile regression analyses over a dataset spanning from May 1998 to March 2022, we investigate the multifaceted effects of investor sentiment, captured through various measures including the VIX, Consumer Confidence (CC), and the Baker-Wurgler Index (BW), on market outcomes. Our findings reveal that investor sentiment significantly influences both stock returns and trading volumes, with its impact varying across different market conditions and quantiles, underscoring the non-linear and state-dependent nature of sentiment's effects. Specifically, we find that sentiment indicators predict shifts in market dynamics, albeit with nuanced relationships across different financial metrics such as SMB, HML, RMW, and CMA factors. Furthermore, the study highlights the conditional influence of sentiment, demonstrating its varying impact across the distribution of returns and volumes, thereby providing insights into sentiment-driven market anomalies. These results not only contribute to the academic discourse on market efficiency and behavioral finance but also offer practical implications for portfolio management and investment strategies, suggesting avenues for future research in exploring the underlying mechanisms of sentiment's market impact.

Alternative Trend-Cycle Decomposition Methods: Real-Time Performance And Forecasting

Presenter: Gian Luigi Mazzi
Co-authors: Gian Luigi Mazzi;Ferdinando Biscosi;Stefano Grassi;Francesco Ravazzolo;Rosa Ruggeri Can-

nata;Piotr Ronkowski

This paper critically analyses three methods, Hodrick and Prescott, Christiano and Fitzgerald approximation, and Harvey and Trimbur model, for calculating trend and cycle estimates of various euro area indicators over the past three decades, inclusive of the COVID-19 pandemic era. The study incorporates an additive outlier correction to manage extreme observations during the pandemic. The research recreates the estimations over 30 successive periods from January 2020 to May 2022 as part of a distinct case study. Results reveal similar GDP and employment estimates across all three filters for the euro area. However, the Harvey and Trimbur model shows a higher trend in periods of Industrial Production Index (IPI) expansion and a lower one during contraction phases such as the financial crisis and COVID-19 pandemic. Furthermore, a novel methodology focusing on predicting trends, cycles and the aggregate variable is proposed as a distinguished factor for decomposition methods. Nevertheless, all three methods display equivalent accuracy, suggesting the absence of a superior method.

Predicting China's Outbound Travel With Multi-Source Data

Presenter: Wang Yongjing
Co-authors: Yongjing Wang;Xinyan Zhang;Haiyan Song

With the development of information and communication technologies, more and more high-frequency internet big data has become available for tourism demand forecasting. Traditional forecasting methods may face challenges in tackling multi-source, different frequency data, and nonlinear relationships. This study contributes to the research area of tourism demand forecasting with multi-source internet big data by including the tourism sentiment data together with the search engine data in the forecasting of China's outbound travel activities. Aiming to improve the demand predictions, we integrate various frequencies and different data sources, including economic variables, Baidu search query volume, and online social media data from sources such as Ctrip, Xiaohongshu and Weibo. The state-of-the-art Generalized Dynamic Factor Model with Mixed Data Sampling (GDFM-MIDAS) is employed to combine multi-source data with nonlinear relationships. Experimental results show different influences created by different data on the improvement of forecasting accuracy.

The Development Of A Self-Adaptive Tourism Demand Forecasting Platform

Presenter: Haiyan Song
Co-authors: Haiyan Song;Ying Liu;Anyu Liu;Tianran Qiu;Xinyan Zhang;Eden Jiao

This study introduces the methods and process in the development of a self-adaptive tourism demand forecasting platform, which encompasses key functions of data management, short-, medium-, and long-term forecasting, sentiment analysis, and interactive scenario forecasts. Employing an interdisciplinary approach that integrates well-established theories in economics, tourism management, and computer science, the platform enables tourism businesses and destination managements to respond more effectively to market dynamics. Additionally, the forecasting results of the platform provides valuable insights for policy-makers to design effective tourism strategies for sustainable tourism development in the selected destinations.

The Anatomy Of Machine Learning-Based Portfolio Performance

Presenter: David Rapach
Co-authors: David Rapach;Philippe Goulet Coulombe;Erik Christian Montes Schütte;Sander Schwenk-Nebbe

The relevance of asset return predictability is routinely assessed by the economic value that it produces in asset allocation exercises. Specifically, out-of-sample return forecasts are generated based on a set of predictors, increasingly via "black box" machine learning models. The return forecasts then serve as inputs for constructing a portfolio, and portfolio performance metrics are computed over the forecast

evaluation period. To shed light on the sources of the economic value generated by return predictability in fitted machine learning models, we develop a methodology based on Shapley values—the Shapley-based portfolio performance contribution (SPPC)—to directly estimate the contributions of individual or groups of predictors to portfolio performance. We illustrate the use of the SPPC in an empirical application measuring the economic value of cross-sectional stock return predictability based on a large number of firm characteristics and machine learning.

Bayesian Estimation Of Panel Models Under Potentially Sparse Heterogeneity

Presenter: Boyuan Zhang
Co-authors: Boyuan Zhang;Hyungsik Roger Moon;Frank Schorfheide

We incorporate a version of a spike and slab prior, comprising a pointmass at zero ("spike'') and a Normal distribution around zero ("slab") into a dynamic panel data framework to model coefficient heterogeneity. In addition to homogeneity and full heterogeneity, our specification can also capture sparse heterogeneity, that is, there is a core group of units that share common parameters and a set of deviators with idiosyncratic parameters. We fit a model with unobserved components to income data from the Panel Study of Income Dynamics. We find evidence for sparse heterogeneity for balanced panels composed of individuals with long employment histories.

From Reactive To Proactive Volatility Modeling With Hemisphere Neural Networks

Presenter: Philippe Goulet Coulombe
Co-authors: Philippe Goulet Coulombe;Mikael Frenette;Karin Klieber

We reinvigorate maximum likelihood estimation (MLE) for macroeconomic density forecasting through a novel neural network architecture with dedicated mean and variance hemispheres. Our architecture features several key ingredients making MLE work in this context. First, the hemispheres share a common core at the entrance of the network which accommodates for various forms of time variation in the error variance. Second, we introduce a volatility emphasis constraint that breaks mean/variance indeterminacy in this class of overparametrized nonlinear models. Third, we conduct a blocked out-of-bag reality check to curb overfitting in both conditional moments. Fourth, the algorithm utilizes standard deep learning software and thus handles large data sets – both computationally and statistically. Ergo, our Hemisphere Neural Network (HNN) provides proactive volatility forecasts based on leading indicators when it can, and reactive volatility based on the magnitude of previous prediction errors when it must. We evaluate point and density forecasts with an extensive out-of-sample experiment and benchmark against a suite of models ranging from classics to more modern machine learning-based offerings. In all cases, HNN fares well by consistently providing accurate mean/variance forecasts for all targets and horizons. Studying the resulting volatility paths reveals its versatility, while probabilistic forecasting evaluation metrics showcase its enviable reliability. Finally, we also demonstrate how this machinery can be merged with other structured deep learning models by revisiting Goulet Coulombe (2022)'s Neural Phillips Curve.

When Is A Forecast Not A Forecast?

Presenter: Gráinne Costigan
Co-authors: Grainne Costigan;Jordi Agut;Pol Riba;Kamil Bilicki;Paula Lopez;Cristina Gil

Within supply chain an accurate forecast can be used to limit safety stock, optimize distribution lines, in commercial it can be used to set sales targets and finance set global company long term goals. A forecast optimized for one, more often than not, is not fit another. Due to the digital transformation that Sanofi is undergoing, the discrepancies between these forecasts are becoming more apparent. However, this digitalization means they can also be swiftly identified and resolved systematically, a step which is becoming a key part of the transformation process. Automating our supply chain forecast has meant very low

adoption of the baselines in the past (full pipeline to be presented in presentation on our transformational process). Our continued focus is deriving an accurate baseline on short to midterm for supply chain factory production and distribution channel using statistical and machine learning models. However, to increase our accuracy across the time horizons and increase the adoption of the statistical forecast we are expanding our models. Changes in the market dynamics will affect our demand signal and long term company vision and product plans have to be accounted for. We have built a platform to systematically capture and market intelligence from those functions working closely with it and integrated it into our baselines to correct for changes in markets. Furthermore, we are deriving long term statistical forecast to meet the strategic vision of the company and individual products. Merging these forecasts means it is no longer a purely statistical forecast but becomes a hybrid forecast which includes objectives and opportunities, meeting the different stakeholder requirements across different time horizons. We take you through the development of these pipelines and how we have engaged and worked with the business to maximize the adoption and smooth the transformation process. Bringing this workflow to production we are capable of reducing the fine tuning of forecasts by hand, focus resources and supply further automation across the supply chain and neighboring functions.

Leading Indicators In Hierarchical Forecasting With Inventory Evaluation

Presenter: Yves R. Sagaert
Co-authors: Yves R. Sagaert;Nikolaos Kourentzes

Inventory management decisions have a direct impact on service levels and working capital, making them of paramount importance to a company's operations. These decisions are made at the stock keeping unit (SKU) level, typically relying on extrapolative forecasts. The noise of these time series makes including market information a challenge. However, information from leading indicators has shown to improve higher level of demand forecasts. Our proposed methodology takes advantage of the hierarchical structure of the problem. As a result, we provide probabilistic forecasts enriched with leading indicator information at SKU level, as input for inventory management. In this study, we present the results of a business case on forecasting accuracy and inventory evaluation, for both the backordering and lost sales scenarios. We consider statistical learning and machine learning models. Furthermore, we present in detail the source of improvement, whether it arises from the hierarchical alignment, the use of leading indicator information, or a combination of these elements.

Automated Demand Forecasting In Small- To Medium-Sized Enterprises

Presenter: Thomas Gärtner
Co-authors: Thomas Gärtner;Thomas Kluth;Antje Heine;Christoph Lippert;Stefan Konigorski

In response to the critical need for accurate demand forecasts to optimize production, purchasing, and logistics, this study introduces an automated forecasting pipeline designed to level the playing field for small- to medium-sized enterprises (SME). SMEs face unique challenges and typically do not have the same in-house expertise to specify forecasting tools to their individual needs such as in large corporations. We have developed a comprehensive data analysis pipeline that automates time series sales forecasting, encompassing data preparation, model fitting, and model selection based on an internal validation step, thereby addressing the specific needs of SMEs.The development process of the pipeline was divided into two phases: prediction model evaluation and prototype testing. In the first phase, we evaluated a large variety of state-of-the-art prediction models on the data of one SME for an 18-month prediction horizon. Based on the results, we identified six models as suitable candidates for the pipeline, including ARIMA, SARIMAX, Holt-Winters Exponential Smoothing, Regression Tree, Dilated Convolutional Neural Network, and Generalized Additive Model. Notably, Long-Short-Term-Memory models did not achieve the desired prediction accuracy, and Facebook Prophet, despite its robust predictions, failed to meet our production environment's implementation criteria. These models were implemented in an automated prototype pipeline, which recommends the best model based on a validation step. In the second phase, the prototype was tested with a group of SMEs which revealed variable model performance across different companies. While

one company did not benefit due to its project-based work nature, all other SMEs experienced realistic and superior sales forecasts compared to current practice. This underscored the importance of model diversity and automated internal validation in catering to the unique needs of each SME.Our findings suggest that no single prediction model is universally superior; rather, an array of models, when integrated within an automated validation framework, can significantly enhance forecasting accuracy for SMEs. This research contributes to the field by giving access to state-of-the-art sales forecasting tools, enabling SMEs to make data-driven decisions with improved efficiency.

Does Forecast Accuracy Even Matter? - Findings From A Retail Dataset

Presenter: Jim Hoover
Co-authors: Jim Hoover

Several researchers and practitioners have recently investigated the emphasis placed on forecast accuracy in practice. Many previous articles in business journals and consulting white papers make claims of large improvements in business outcomes as forecast accuracy improves. Our research utilizing a large dataset of daily demand data for stock keeping units (SKUs) at the store level indicates that this relationship is not linear and given the standard parameters of a demand planning system does not always improve business outcomes as accuracy improves. In this presentation, we will discuss the results of the research and its implications for key forecast practitioner activities, such as forecast process improvement.

Demand Forecasting For German Pharmacies

Presenter: Julia Schemm
Co-authors: Julia Schemm;Julian Stengl;Luca Rendsburg

Inventory management in pharmacies is characterized by a broad and diverse product range, while storage capacities are limited and purchasing conditions are complex. The manual order planning that predominates in German pharmacies is therefore very time-consuming. This poses a major problem in view of the shortage of skilled staff. Software-supported forecasts that take into account product diversity and demand uncertainty are one part of our solution to shorten planning time, meet future demand and avoid overstocking. The other part is a mathematical optimization model that uses the demand forecasts as input. In this submission, we focus on the forecast development and evaluation.In a case study with three German pharmacies, we have developed a demand forecasting approach that aims to achieve the aforementioned goals. By clustering the sales time series into four categories according to their intermittency and variance and training one DeepAR model per cluster, similar time series share information. We also evaluate the integration of exogenous features, including seasonal patterns, item-specific data and hierarchical dependencies as well as weather and pollen data. As local benchmark models, we use the moving average (industry standard) and an ets model.Since the demand forecast is supposed to be an input for a stochastic optimization model, which in turn outputs recommendations for orders, we are not only interested in the accuracy of the point forecast, but also in modeling the forecast uncertainty. Therefore, we compare the ets approach, in which samples are drawn from the assumed error distribution, to the DeepAR approach where the likelihood at each timestep is parametrized by the output of an autoregressive recurrent network.

Model Averaging For Decomposed Data

Presenter: Yuying Sun
Co-authors: Yuying Sun;Yongmiao Hong;Shouyang Wang;Wencan Lin

The decomposition-ensemble algorithm has received increasing attention in forecast and related fields, especially in capturing the nonlinear and nonstationary characteristics of time series data. A conventional strategy involves decomposing the target time series into various oscillation modes from the frequency domain and assigning equal weights to all decomposed modes for aggregated prediction. However, disparities in forecasting performance arise among different decomposed modes due to their distinct attributes and

forecast horizons. This paper proposes a novel forward-validation model averaging approach to combine decomposed modes with appropriate weights, thereby enhancing the accuracy of the target time series forecast. It is shown that the proposed model averaging estimator is asymptotically optimal in the sense of achieving the lowest possible quadratic prediction risk. The rate of the selected weights converging to the optimal weights to minimizing the expected quadratic loss is established. Simulation studies and empirical applications to consumption and exchange rate forecasting highlight the merits of the proposed method.

Forward Stagewise Linear Regression For Ensemble Methods

Presenter: Daniel Uys
Co-authors: Daniel Uys

In supervised learning, the forward stagewise regression algorithm is considered a more constrained version of forward stepwise regression. In its turn, the forward stagewise regression algorithm can be refined to produce the incremental forward stagewise regression model. In the latter model, the idea of slow learning is introduce where the residual vector and the appropriate regression coefficient are updated in very small steps at each iteration.Ensemble methods combine a large number of simpler base learners to form a collective model that can be used for prediction. Learning methods such as Bagging, Random Forests, Boosting, Stacking and Regression Splines amongst others, can all be regarded as ensemble methods. In these methods, the linear model is expressed as a linear combination of these simpler base learners, where the coefficients of the base learners are to be estimated by least squares. Since a large number of base learners is typically involved, the residual sum of squares of the linear combination of base learners has to be penalised by, for example, the lasso penalty. However, the large number of base learners complicates the minimisation of the coefficients in the penalised residual sum of squares criterion. By using the iterative forward stagewise linear regression algorithm for ensemble methods, which includes the idea of slow learning and closely approximate the lasso, estimators of the coefficients of the base learners can be obtained. In the talk, the performance of various ensemble methods is evaluated. This is done by applying the forward stagewise linear regression algorithm for ensemble methods to simulated and real life datasets.

Forecast Linear Augmented Projection (FLAP): A Free Lunch to Reduce Forecast Error Variance

Presenter: Yangzhuoran Fin Yang
Co-authors: Yangzhuoran Fin Yang;Rob J. Hyndman;George Athanasopoulos;Anastasios Panagiotelis

A novel forecast linear augmented projection (FLAP) method is introduced, which reduces the forecast error variance of any unbiased multivariate forecast without introducing bias. The method first constructs new component series which are linear combinations of the original series. Forecasts are then generated for both the original and component series. Finally, the full vector of forecasts is projected onto a linear subspace where the constraints implied by the combination weights hold. It is proven that the trace of the forecast error variance is non-increasing with the number of components, and mild conditions are established for which it is strictly decreasing. It is also shown that the proposed method achieves maximum forecast error variance reduction among linear projection methods. The theoretical results are validated through simulations and two empirical applications based on Australian tourism and FRED-MD data. Notably, using FLAP with Principal Component Analysis (PCA) to construct the new series leads to substantial forecast error variance reduction.

Co-Explosiveness Of Corporate Credit Spreads

Presenter: Marco Kerkemeier
Co-authors: Marco Kerkemeier

Financial crisis and exuberances are a companion of financial markets ever since. At latest since the global financial crisis 2007/09 (GFC) there is a growing awareness for the importance of systemic risk

in financial markets. In this vein and as a major contributor to the GFC, understanding of dependence structures within the fixed income market in a more detailed way is required. A special situation that demands attention is co-explosiveness of credit risk. In such a situation, two or more explosive time series share the same autoregressive regime. Under such circumstances, there is a significantly increased systemic risk that may threaten the stability of the financial system. This research deals with the co-explosiveness of different bond market spreads (especially those accounting for credit risk and the flight-to-quality behaviour of institutional investors during turmoil) within a country and across different countries. The six major markets analysed are Australia, Canada, Euro Area, Japan, UK and the US. For all these markets, weekly credit spreads between the 5th January 2001 and 29th September 2023 are analysed. Explosive periods are identified in all 24 analysed spreads and the date stamping of explosiveness reveals a clustering of those periods around the GFC and the Corona pandemic. To test them for co-explosiveness, a bivariate KPSS-based test, which is able to identify synchronous and asynchronous (lagged) co-explosiveness, is applied. So far, I find asynchronous co-explosiveness which points towards a connectedness between the explosive credit spread periods of the different markets. Most periods of credit spreads explosiveness are driven by the explosiveness of the US-market. Understanding such behaviour is important for policymakers, central bankers and investors because the explosiveness in the US-market can serve as a warning sign for upcoming explosiveness in the other markets.Additionally - as a by-product - I investigate the influence of outliers and jump-like situations on the size and power characteristics of the (co)explosiveness-tests. Unlike stock and precious metal prices, corporate bond spreads can behave much more extreme with fast ups and downs. This can e.g., be driven by a rebalancing of the underlying bonds of the index.

On The Stability Of Key Relations Between Us Interest Rates

Presenter: Tamas Kiss
Co-authors: Tamas Kiss;Pär Österholm;Sune Karlsson;Hoang Nguyen

In this paper we assess whether key relations between US interest rates have been stable over time. This is done by estimating trivariate hybrid time-varying parameter Bayesian VAR models with stochastic volatility for the three-month Treasury bill rate, the slope of the Treasury yield curve and the corporate bond-yield spread. As a methodological contribution, we also allow for disturbances with heavy tails. We analyse monthly data from April 1953 to February 2023 both within- and out-of-sample. Our results indicate that the relations have not been stable; more specifically, there is evidence that the equation of the corporate bond yield spread is subject to time-variation in its parameters. We also find that an increase in the corporate bond yield spread decreases the risk free rate. Finally, we note that while allowing for heavy tails receives a fair amount of support within sample, it appears to be of more limited importance from a forecasting perspective.

Forecasting The Yield Curve: The Role Of Additional And Time-Varying Decay Parameters, Conditional Heteroscedasticity, And Macro-Economic Factors

Presenter: Esther Ruiz
Co-authors: Esther Ruiz;Joao Caldeira;Werley Cordeiro;André Santos

In this paper, we analyse the forecasting performance of several parametric extensions of the popular Dynamic Nelson-Siegel (DNS) model for the yield curve. Our focus is on the role of additional and time-varying decay parameters, conditional heteroscedasticity, and macroeconomic variables. We also consider the role of several popular restrictions on the dynamics of the factors. Using a novel dataset of end-of-month continuously compounded Treasury yields on US zero-coupon bonds and frequentist estimation based on the extended Kalman filter, we show that a second decay parameter does not contribute to better forecasts. In concordance with the preferred habitat theory, we also show that the best forecasting model depends on the maturity. For short maturities, the best performance is obtained in a heteroscedastic model with a time-varying decay parameter. However, for long maturities, neither the time-varying decay nor the heteroscedasticity plays any role, and the best forecasts are obtained in the basic DNS model with the shape of the yield curve depending on macroeconomic activity. Finally, we find that assuming

non-stationary factors is helpful in forecasting at long horizons.

Credit Market Sentiment And Stock Returns

Presenter: Gergely Ganics
Co-authors: Gergely Ganics;Chaoyi Chen;Zhou Ren

Credit market sentiment is characterized by mean reversion and has been shown to possess significant predictive power for future economic activity through a dynamic mechanism known as "diagnostic expectations." Typically, this relationship has been analyzed using linear regressions with two predictors of future changes in credit spreads: the level of credit spreads and the share of high-yield bonds issued. However, it remains unclear whether the reversal of prior sentiment holds significant predictive power for changes in credit spreads under both high and low sentiment conditions. In this study, we find that diagnostic expectations exhibit asymmetry by examining quantile regressions. We then propose an entropy-based measure of positive sentiment to capture the asymmetric reversion. This measure not only performs well in cross-sectional asset pricing regressions but also delivers meaningful out-of-sample gains when predicting stock returns. Our findings highlight the importance of considering asymmetric reversion in credit market sentiment when forecasting economic activity and asset prices.

Sectoral Producer Price Inflation Connectedness Across The North American Trade Agreement Region

Presenter: Lenin Arango-Castillo
Co-authors: Lenin Arango-Castillo;María José Orraca;Regina Briseño

Since the North America Free Trade Agreement was signed in 1994, input-output linkages between the countries in the region have strengthened, leading to the development of regional manufacturing production networks. The connection through production links may facilitate the transmission of price shocks across the region. Using the Diebold-Yilmaz framework, we analyze the connectedness of producer price annual inflation of manufacturing industries in Canada, Mexico, and the United States from 1994 to 2019. We find that the connectedness of inflation between Mexico and the United States is stronger than between these countries and Canada. This could be pointing to a lower integration level of the latter with the region's production or to idiosyncratic factors being relatively more important for driving inflation in Canada. Hence, the influence of inflation of its trading partners is less relevant. We also find that the total connectedness of the system is stronger for a shorter period (2010-2019). In particular, the fraction of the variance of inflation of the different trade partners that the others can account for increases. This could be rationalized if, supported by the agreement, regional production networks strengthened over time. For example, suppose the participation of North American content in each other's production has increased. In that case, the influence of the trading partners' sectoral producer price inflation on production costs, and thus, their influence on inflation, should be larger. We also extended the analysis to 2023 to explore the changes when the pandemic period is included. We find an increase in the connectedness of the system, which is consistent with the view that global factors were driving inflation globally during the most recent period. Our results are robust when using monthly and quarterly inflation rates with lags congruent with annual variations.

Graph Neural Network Framework Based Economic Forecasting

Presenter: Htoo Wai Aung
Co-authors: Htoo Wai Aung;Ying Guo;Jiaming Li;Geoffrey Lee;Zili Zhu

The well-known cascading economic forecasting model first proposed by Wilkie is an ARIMA model of a cascading structure representing the inter-relationship between major economic variables. The hierarchical relationship structure was formulated from empirical observations and established economic theories. In this paper, we explore using machine-learning techniques to reveal inter-relationship structures

between these major economic variables without relying on any prior economics knowledge. For this paper, we use Graph Neural Network (GNN) framework to automatically generate a hierarchical structure of eleven major economic variables from the historical time-series data. As benchmark, we compare the outcome from the GNN framework with a cascading economic forecasting SUPA model (which is an extension of the original Wilkie model). An adjacency matrix maps out the interconnectedness among these economic variables in the GNN model. For the GNN model presented, we have implemented three distinct, data-drive methods for adjacency matrix construction: PC (Peter-Clark), Greedy Equivalence Search (GES), and EGLT (Economic Graph Lottery Ticket). The PC and GES methods are rooted in causal graph discovery techniques. The EGLT algorithm, an innovative strategy for forming the adjacency matrix within GNN, drawing inspiration from the Unified Sparsification GNN (UGS) model. For the eleven major Australian economic variables, after training with data of 1992-2016, and testing with data of 2017-2019, the GNN approaches surpass the SUPA model in reducing the overall Root Mean Square Error (RMSE), even at the level of individual variable forecasts. The EGLT method achieved the most efficient adjacency matrix at a 40.50% density by iteratively reducing connections from 100% to 9.92%. On the other hand, with the eleven economic variables, the PC approach generates a hierarchical inter-relationship structure that is most consistent with the current economic theory and is also very similar to the cascading structure of the SUPA model. In summary, the paper will demonstrate that Graph Neural Network framework with the PC algorithm can reliably be used to reveal the complex inter-relationship structure among key economic variables and trained GNN model can be used to provide more accurately economic forecasting.

Forecasting Of Regional Economic Resilience To Crisis Shocks

Presenter: Nadezda Kanygina
Co-authors: Nadezda Kanygina

In recent years, the world economy has been constantly experiencing the consequences of various types of crises. There are many studies devoted to predicting crises, analysis of their causes and consequences at the country level, but the internal regional effects of shocks have been little studied. However, a single shock can have a different impact on the economies of the regions, subsequently leading to an increase in the gap in their development and a slowdown in the rate of economic growth of the country as a whole. In this regard, it is strategically extremely important to study the crisis dynamics within the country, the degree of resistance of regions to shocks and the speed of recovery of regional economies, taking into account their interaction. The impact of crisis shocks on regions is directly related to the theory of resilience. For empirical research data for Russia over the period 2004-2021 are taken to investigate the economic resilience of its regions to absolutely different shocks (world financial crisis 2008-2009, economic crisis 2015, COVID-19) with the help of panel data models and spatial econometric models. In comparison to traditional approaches several types of geographically weighted regressions (GWR) were considered which turned out to provide with more precise estimates. A kind of a new type - geographically weighted error correction model was also derived that allowed to distinguish long term and short-term effects varying in space. All the regions were clustered according to the space and time varying marginal effects on local resilience and recommendations for policy makers were developed. Based on the examined experience of recoverability from these previous shocks the consequences of current crisis 2022 were estimated for regions.

Uk Regional Forecasting Model With Global Vector Autoregressive Approach

Presenter: Jeremy Kwok
Co-authors: Jeremy Kwok

This paper builds a UK regional economic model and measures the effects of various shocks on the UK economy. The paper aims to construct a model that can analyse monetary, fiscal and oil shocks to the regional economies of the UK. The methodology employs the Global vector autoregressive (GVAR) approach which links different UK regions by their distances and linkages to the dominant economy of London. Regional housing price index (HPI), economic activities and fiscal variables. The model also contains the UK interest rate and oil price as national variables. This paper has found evidence of heterogeneous

responses to various shocks in different regions, particularly in a region that is further away from London such as Northern Ireland and Scotland.

Nowcasting China's Price Indexes Based On The Factor Model And Machine Learning

Presenter: Qin BAO
Co-authors: Qin Bao;Yangyang Zheng;Yan Yu

Inflation indicating by aggregated price indexes are important reference indicators for economic analysis and policies. This paper proposes a three-step nowcasting model that combines econometric models and machine learning methods. Firstly, we identify the transmission relationship between different price sub-items; Secondly, we establish a dynamic factor model with conduction constraints to extract low dimensional factors from high-dimensional missing sample data; Finally, three machine learning methods are used to establish nonlinear models between factor variables, conduction variables, and dependent price variables. This method well integrates the economic logic information of price transmission into econometric models, and based on dynamic factor models, fully utilizes timely indicator information to obtain low dimensional factors, and explores the nowcasting performance of machine learning methods on price indices. The empirical results show that the dynamic factor model based on price transmission has higher nowcasting accuracy than traditional prediction models, and machine learning methods further improve the nowcasting performance of CPI, with support vector machine regression performing the best. In addition, the prediction error of the model for the major price indices gradually decreases with the increase of data in the current month. The acquisition and application of real-time data are of great significance in the nowcasting of low-frequency macroeconomic indicators. This paper proposes a nowcasting method that combines linear factor models with nonlinear machine learning models, whose performance in nowcasting of China's major price indices is explored and analyzed.

Forecasting Korean Inflation: A Case Study On Relationship Between Time Series Persistence And Machine Learning Forecasting

Presenter: Heejoon Han
Co-authors: Heejoon Han

This paper considers inflation forecasting in South Korea using various machine learning methods in a data-rich environment and investigates relationships between time series persistence and machine learning forecasting. We consider both the month-on-month inflation rate and the year-on-year inflation rate. The former is less persistent, and its autoregressive coefficient is 0.297. Whereas the latter is quite persistent, and its autoregressive coefficient is 0.964. It is not all, but some unit root tests indicate that the year-on-year inflation rate is a unit root process. We use 93 macroeconomic/financial variables as explanatory variables and adopt thirteen models. The samples are from September 2004 to June 2023. For the month-on-month inflation rate forecasting, one model outperforms the rest models from one-month to twelve-month ahead forecast horizons, which consists of the following two steps: 1) Select important variables based on the Boruta algorithm, 2) Using only those selected variables, implement the random forest and produce a forecast (henceforth BS-RF model). The tests by Giacomini and White (2006) and Hansen et al. (2009) show that the BS-RF model provides significantly better forecasts than the rest models. In particular, the Boruta algorithm selected total economically active population, total employed persons, BSI, and house price as important explanatory variables for Korean inflation forecasting. However, for the year-on-year inflation rate forecasting, the forecasting performances of machine learning methods are very poor, and the random walk model exhibits the smallest forecast losses for most forecast horizons. This may be originated from the fact that the year-on-year inflation rate is quite persistent, and it changed rapidly during the forecast period from June 2015 to June 2023. We conduct an additional study to investigate whether time series persistence matters in machine learning forecasting. We first produce forecasts of less persistent month-on-month inflation rates by using various machine learning methods and next convert them into forecasts of the year-on-year inflation rate by a simple calculation. This way turns out to produce smaller

forecast losses compared to each method that directly forecasts the year-on-year inflation rate. Moreover, it could provide significantly better forecasts than the random walk model that directly forecasts the year-on-year inflation rate.

Pure Inflation: The Case Of Colombia

Presenter: Juan Julio
Co-authors: Juan Julio

We estimate the pure component of inflation in Colombia following Reis and Watson (2010). These authors identify pure inflation as the absolute-price change component that is orthogonal to relative-price and idiosyncratic sectoral inflation variation. For this, we employ a dynamic factor model for the inflation rates of the totality of quarterly Colombian CPI classes. Since Colombian inflation is a process that follows a near unit root or unit root behaviour (due to smooth and sudden transitions between steady states), we propose an adaptation of the model in Reis and Watson's paper to deal with the arising non stationarity. We also address the aggregation issues that arise in these models that were pointed out by Ahn and Luciani (2021). We estimate our model through the EM algorithm we develop in the paper. We relate our inflation components's estimates to other variables at business cycle frequencies, emphasising on real activity, that is on the Phillips correlation, and assess the forecasting accuracy of inflation models based on this decomposition with those related to other decompositions.

Instantaneous Inflation As A Predictor Of The Inflation

Presenter: Wilmer Martínez-Rivera
Co-authors: Wilmer Martinez Rivera;Edgar Caicedo Garcia;Juan Bonilla Perez

Instantaneous Inflation is a transformation of year-on-year Inflation, which provides different monthly weights; the last months usually receive higher weights. This transformation allows us to find a variable potentially leading to the inflation dynamic. We explored the leading between instantaneous inflation and inflation by using the coincident profile developed by Martinez, Nieto, and Poncela (2016). Based on these results, we fit ARIMA, VECM, and ARIMAX models to evaluate the forecast performance and the usefulness of instantaneous inflation and its lead to inflation.

Rolch: Regularized Online Learning For Conditional Heteroskedasticity

Presenter: Simon Hirsch
Co-authors: Simon Hirsch;Jonathan Berrisch;Florian Ziel

Large-scale streaming data are common in modern machine learning applications and have led to the development of online learning algorithms. Many fields, such as supply chain management, weather and meteorology, energy markets, and finance, have pivoted towards using probabilistic forecasts, which yields the need not only for accurate learning of the expected value but also for learning the conditional heteroskedasticity. Against this backdrop, we present a methodology for online estimation of regularized linear models for conditional heteroskedasticity. The proposed estimation algorithm is based on a combination of the iteratively reweighted least squares estimation, which has been successfully used for the regularized estimation of ARCH processes with high-dimensional covariate spaces for the conditional variance and recent developments for the online estimation of LASSO models for the conditional mean. We provide simulation results showing the stability of online LASSO estimation and the behavior of the conditional variance estimates. Our algorithms are implemented in a computationally efficient C++ package with bindings to python and R.

Bootstrap-Based Forecasts In Battery Charging Strategies

Presenter: Tomasz Weron
Co-authors: Tomasz Weron;Katarzyna Maciejowska

Nowadays more and more renewable energy sources (RES) emerge across Europe. Their most important advantage over the conventional ones – little to no environmental pollution. What is healthy for our planet however, is not necessary beneficial for energy markets. With their volatility and uncertainty of generation, RES greatly destabilize power grids. To address this issue energy storage systems, such as batteries, are becoming more common. Such installations not only provide safety but offer new trading possibilities as well. Although they main role is to ensure stability, batteries can also be used to buy energy when the price is low, store it, and then sell it – when high. In this study, we propose a new autoregressive bootstrap-type approach to provide probabilistic forecasts of day-ahead market prices. Utilizing them, we develop several trading strategies. We use German, Spanish and Danish markets' data to evaluate our methods. We show that our approach can outperform simple autoregressive models in terms of Value-at-Risk and average profit per trading day.

Probabilistic Forecasting With A Hybrid Factor-Qra Approach: Application To Electricity Trading

Presenter: Katarzyna Maciejowska
Co-authors: Katarzyna Maciejowska;Tomasz Serafin;Bartosz Uniejewski

In this paper, we describe a novel application of a Factor-QRA method. It extends the PCA forecast averaging approach of Maciejowska et al. (2020) to probabilistic forecasting. In this research, Pincial Component method is used to aggregate information included in a rich panel of 673 point predictions. The factors are next used as an input to Quantile Regression. Unlike in previous works, the point forecasts, which constitute the panel, are subjected to standardization. Moreover, the standardization is implemented across the cross-sectional rather than the time dimension. It is also applied to the dependent variable, requiring a subsequent transformation of the final outcomes. As a result, it allows to incorporate the impact of point forecast heterogeneity, measured by a forecast variance, on the width of the prediction intervals. Although the outcomes suggest that the standardization has a crucial impact on forecast performance, its significance has not been studied in the literature.The forecast accuracy of the proposed method is evaluated with datasets from two European power markets: the German EPEX SPOT and the Polish Power Exchange (TGE). The data for the EPEX day-ahead market covers the tranquil period of 2015-2020, the COVID-19 pandemic and the turbulent period of the Ukrainian war. It encompasses different dynamics of the electricity prices and hence allows for a comprehensive comparison of different forecasting methods. The results show that (i) factor-based averaging schemes provide more accurate probabilistic forecasts than their QRA/QRM counterparts, (ii) the empirical coverage of the new methods is close to the nominal level and passes the Christoffersen (1998) test in most cases, (iii) methods using standardized panels are more accurate than their not standardized counterparts.Finally, to assess the economic value of the proposed forecasting schemes, an experiment is conducted that mimics a real-world trading problem. We consider a decision problem of a medium-sized energy storage unit. The results demonstrate that data-driven trading strategies outperform the unlimited-bids benchmark with respect to the average income per 1 MWh.

Integrating Probabilistic Forecasts Of Fundamental Variables For Enhanced Electricity Price Forecasting

Presenter: Bartosz Uniejewski
Co-authors: Bartosz Uniejewski

Electricity price forecasting plays a critical role in the efficient operation of power systems and market participants decisions. This paper explores a novel approach to improve the accuracy of electricity price forecasts by incorporating probabilistic inputs of fundamental variables. Instead of relying solely on point forecasts of exogenous variables such as load, wind, or solar generation forecasts, we propose to incorporate quantile forecasts of these variables. We conduct empirical tests on the German electricity market using recent data. Our results show that the integration of probabilistic forecasts significantly improves the accuracy of point forecasts of electricity prices. In particular, the largest improvements in

accuracy are attributed to the probabilistic forecast of wind power generation. This research contributes to the advancement of electricity price forecasting methodology and underscores the importance of accounting for uncertainty in fundamental variables to improve forecasting performance in energy markets.

Detecting Concept Drift In Household Consumption Time Series

Presenter: Eline Michiels
Co-authors: Eline Michiels;Lola Botman;Sibren Lagauw;Bart De Moor

Effective operation of smart grids and integration of renewable energy sources crucially depend on accurate household-level electricity demand forecasting. One of the challenges in achieving precise forecasting lies in the dynamic nature of electricity demand, characterized by a phenomenon known as concept drift. In the context of power consumption, concept drift is the change in statistical properties of the electricity demand over time, complicating the task of forecasting with traditional static models.Our research explores various concept drift detection methods such as Adaptive Windowing and the Page-Hinckley test, and delves into probabilistic forecasting methodologies. Probabilistic forecasting, a key focus of this study, predicts future electricity demand while quantifying uncertainty, employing both local and global methods. Local methods customize predictions for individual households based on specific characteristics, whereas global methods use a unified model for broader trend analysis across all households. By evaluating these probabilistic approaches, the study aims to identify the most effective strategies for adapting to the dynamic nature of electricity usage patterns. Benchmarking these methodologies against baseline models is an integral part of the research, enhancing our understanding of forecasting in the face of concept drift. Building on these insights, the most promising local and global forecasting models are integrated into a singular, robust hybrid model, employing the concept drift detection techniques to switch from a local method to a global method when drift is detected, to ensure its adaptability and resilience over time. The proposed methodology is validated by means of case study, which evaluates the performance of the hybrid model and conducts an extensive comparative analysis against the established baseline models. Two open datasets are used: the Low Carbon London dataset and the Commission for Energy Regulation Irish dataset.

Forcasting Electricity Consumption Of Chinese Nonferrous Metal Industry Based On Public Opinion Mining

Presenter: Wei Shang
Co-authors: Wei Shang;Xuerong Li;Zhengjun Zhang

Nonferrous metal is an essential natural resource for national economy and industry production. Nonferrous metal smelting and pressing industry is one of the most energy consuming industry. This research collects nonferrous market news and discussions from the Internet and builds related market opinion indices to understand the supply, demand, and price changes in nonferrous metal market. A self-learning graph neural network model-based sentiment index network model is proposed to measure the market sentiment-related features, addressing the issue of data missing caused by discontinuities in domain-specific news. Incorporating both statistical market indicators and the Internet sentiment indices, a Granger Causality-based ARDL Model (G-ARDL) and a Factor Analysis-based Transformer Model (F-Transformer) are designed to forecast the nonferrous metal industry electricity consumption. And then a bagging random sampling integration prediction framework is proposed to obtain the weight coefficients to integrate the forecasting results. Empirical study shows that the proposed overall market demand opinion index reveals a one-month leading pattern with the nonferrous metal electricity usage, and can serve as early signals of demand, supply, and prices change. Forecasting performance of the proposed bagging integration method has better in-sample fitting as well as higher out-sample accuracy. This research contributes to a novel methodology of using public opinion mining in electricity forecasting for a specific industrial sector.

Prosumer Net Consumption Forecasting: Investigating Novel Approaches Addressing Behind-The-Meter Self-Consumption

Presenter: Yuliia Siur
Co-authors: Yuliia Siur;Novin Shahroudi;Jean-Baptiste Scellier

In recent years, the adoption of renewable energy sources has shown a significant increase. Solar photovoltaic (PV) panels are gaining widespread popularity, particularly among private households. Residential rooftop PV systems enable private households to generate and utilize their own electricity. Private households with a dual role in electricity production and consumption, also known as "prosumers", establish direct market relationships with energy companies, facilitating the sale of surplus energy to the grid and purchasing when energy production is insufficient. The boost of prosumer-driven energy generation shifts energy companies' electricity flow management. Traditional Consumption metering is no longer sufficient, as it fails to capture prosumers' interactions with the grid. Instead, energy companies adopt Net Metering Systems, which monitor the difference between electricity injected into and withdrawn from the grid. The adoption of Net Metering Systems induces the development of novel methodologies for forecasting Net Consumption. However, the task of forecasting Net Consumption presents challenges arising from the three primary factors: a) the diverse behavioral consumption patterns exhibited by private households; b) the inherent variability of solar energy production, which is influenced by fluctuations in weather patterns and solar positioning across various time frames; c) the nature of Net metering does not consider the actual values of production and consumption but rather accounts for energy injected to and withdrawn from the grid. The discrepancy between actual and monitored values equals the prosumers' self-consumption of generated energy, which remains unmonitored, thereby increasing the complexity of modeling relationships indicated in a) and b). In this study, we focus on advancing one-day-ahead Net energy forecasting techniques using Estonian prosumers as a case study. We introduce novel types of additive and integrated models alongside advanced data engineering approaches, all aimed at mitigating uncertainty originating from unmetered self-consumption through improved capture of the complex relationships between predictors and the target. Experimental results exhibit the superior performance of the proposed models' enhancements and data engineering techniques compared to their unaltered counterparts, suggesting an increased ability to address self-consumption-related uncertainty and establishing a foundation for further research.

A Wavelet Analysis Of Turkish Electricity Consumption Cyclicality During The Month Of Ramadan

Presenter: Erhan Uluceviz
Co-authors: Erhan Uluceviz

Turkish domestic electricity consumption generally follows a certain type of cyclicality. Some periods that affect people's daily routines, such as weekends, holidays or months of religious activities, lead to changes in this cyclicality. Standard econometric methods are often unable to, precisely, detect these changes. However, the wavelet power of electricity consumption series decomposed in time and period (or frequency) domain provides this information. Wavelet transformation can decompose the time series into periods along the time axis and also allows one to reconstruct the original series with the time series of the wavelet power using selected periods. In this analysis, two periods are analyzed: (i) 35-day period from 2 days before hijri Ramadan month to the end of Ramadan feast in 2023 and (ii) the same calendar days (in gregorian calendar) of 2019 as selected in 2023.Wavelet transformation is applied to electricity consumption series in those periods and it is found that the 12, 24 and 168 hour periods are significant (i.e. have high wavelet power). Then, using wavelet power series for these periods, the hourly electricity consumption series are reconstructed. It is observed that these reconstructed series move significantly in tandem with the original series. Finally, the average hourly consumption for the days of the week of the reconstructed series and the original consumption series are compared and it is found that the average hourly electricity consumption from Thursday to Sunday in 2023 decreases significantly compared to the other days of the week. This finding indicates that the month of Ramadan has significant effects on the cyclicality of electricity consumption in Türkiye which could have implications regarding economic activity.

Newswire Tone-Overlay Commodity Portfolios

Presenter: Ana-Maria Fuertes
Co-authors: Ana-Maria Fuertes;Joelle Miffre;Adrian Fernandez-Perez;Nan Zhao

This paper contributes to the scarce literature on the cross-sectional predictive ability of sentiment in commodity futures markets. It puts forward a novel method to embed the sentiment extracted from commodity-specific newswires into extant commodity futures trading strategies. Implementing the newswire tone-overlay strategy on a cross-section of 26 commodities using fundamental commodity futures characteristics – basis, momentum, hedging pressure, relative basis or convexity, skewness, basis-momentum and liquidity – individually and combined, it documents through out-of-sample tests that a significant additional premium can be captured. The findings survive the consideration of transaction costs and are not spurious as suggested by a placebo test. Thus, given the relative trader composition of commodity futures markets, it provides indirect evidence to suggest that sentiment also influences the trading decisions of institutional investors. Furthermore, the paper uncovers recession and limits-to-arbitrage risks as two complementary channels for the transmission of newswire tone into commodity futures prices, and endorses the nexus between limited investor attention and news salience. Lastly, the paper additionally documents the superior efficacy of the tone-overlay approach proposed versus extant style-integration and sequential-sorting methods that could be used to embed the commodity newswire tone signal into a fundamental commodity characteristic. This largely stems from the flexibility of the tone-overlay strategy as it allows portfolio managers, for instance, to increase the signal-to-noise ratio by focusing on the salient (optimistic or pessimistic) newswire tone identified in a time- and commodity-specific manner trough the historical newswire tone distributions. The commodity newswire tone-overlay portfolios proposed can be a potentially useful diversification tool for equity investors given their sizeable performance gains in high inflation periods and their relatively low correlations with the broad equity market. Lastly, the results serve to validate proprietary news analytics software that assigns sentiment scores to commodity newswires.
Keywords: Newswire tone; Sentiment; Commodity futures; Tactical allocation; Salience; Cross-sectional predictability.

Large-Scale Portfolio Selection Via Regularization Ensemble

Presenter: Bonsoo Koo
Co-authors: Bonsoo Koo;Hong Wang

We propose a novel Regularization ENsemble (REN) approach to tackle the portfolio selection problem in finance. While most existing regularization methods directly regularize the weight allocation vector, REN first exploits the correlation structure of assets, which results in smaller target portfolios with lower within-protfolio correlations. The selection process achieves better diversification by sequentially selecting and including other assets using regularization, and the final portfolio is an ensemble of the constructed sub-portfolios. The proposed approach successfully achieves stable out-of-sample performance in the presence of strong correlations between assets. We conduct extensive experiments on four widely used datasets, including the Fama French portfolios and U.S equity indices. In particular, the Sharpe ratio of REN outperforms those of the state-of-the-art methods on all the four test samples.

Optimal Hedging Of Equity Index Options Portfolios

Presenter: Maciej Wysocki
Co-authors: Maciej Wysocki;Robert Ślepaczuk

This paper presents a comparative analysis of quantitative finance models employed in the valuation and hedging of S&P500 index options and an extensive backtest of systematic option-writing strategies. The comparison is conducted utilizing established performance measures with transaction costs. We use the Black-Scholes-Merton (BSM) model incorporating implied volatility and the Variance-Gamma (VG) model based on L´evy processes. The study leverages 1-minute S&P500 index option prices and index quotes from 2018 to 2023. The findings suggest the feasibility of outperforming a buy-and-hold (B&H) benchmark

through systematic investment strategies, with the BSM generally yielding more favorable hedging results than the VG model.

Dynamic Parametric Portfolio Policies

Presenter: Dick van Dijk
Co-authors: Dick Van Dijk;Rasmus Lonn;Bram Van Os

We put forward a Dynamic Regularized Parametric (DRP) approach for active portfolio policies. We build upon the parametric policy framework of Brandt et al. (2009) that directly links the portfolio weights to a limited set of asset characteristics. This yields a parsimonious specification that avoids modeling the joint distribution of returns, and as such remains applicable for large asset universes. We relax the assumption that policy coefficients are constant over time, to accommodate that the relevance of specific characteristics for future asset performance may vary. Dynamic policy coefficients are obtained by maximizing the conditional expected utility for each time period, with transaction costs being limited through a trading regularization. This dual-objective optimization problem results in an elegant filter to update the policy coefficients, balancing between adapting to valuable new, yet inherently noisy, information and providing a stable strategy that avoids costly re-balancing. We demonstrate that for a mean-variance utility investor, our framework yields an intuitive analytical solution. In an empirical application using the full universe of stocks from the NYSE, AMEX and Nasdaq, we find that the DRP approach produces substantial gains in out-of-sample portfolio performance, where both incorporating dynamics and regularization are important to achieve this.

Does Google Analytics Improve Prediction Of Tourism Demand Recovery?

Presenter: Andrea Saayman
Co-authors: Andrea Saayman;Ilse Botha

Recent literature shows that Google Trend indices are useful in improving the tourism demand forecasts relative to autoregressive baseline models. Furthermore, online news media coverage regarding a destination can affect the sentiment of source markets towards a destination and influence the number of tourist arrivals. The sentiment of online news media is often viewed as a leading indicator of actual tourism demand. Given the impact that the recent pandemic had on the tourism industry, this may prove to be an important predictor of tourism recovery in countries that are still struggling to recover. One such a country is South Africa, where tourism demand is still well below its pre-pandemic levels. The purpose of this paper is firstly to build on previous research that indicates Google Trends have the potential to improve tourism demand forecasting by testing this within the context of tourism recovery after a shock. Secondly, this paper aims to examine if changes in sentiment of the major source markets of South Africa possess information to predict actual tourist arrivals. This paper analyses Google trend indices based on country searches, including worldwide searches. This information gives insight into the search behaviour of the major overseas source markets to South Africa, where recovery is slow and lagging. Due to the differing data frequencies of tourist arrivals and big data, the MIDAS modelling approach is appropriate. These MIDAS models including web sentiment indicators are compared to typical time-series and naïve benchmarks to determine if online news media coverage can improve forecasts of tourist arrival recovery from key overseas source markets. The findings in this paper are essential to understand the concerns of tourists regarding events at the destination since it impacts destination choice and influences tourism recovery after shocks.

Pandemic And War: An Econometric Assessment Of Tourism Export Losses

Presenter: Egon Smeral
Co-authors: Egon Smeral

The shocks of the COVID-19 outbreaks and containment measures have caused severe damage to both society and the economy, including the tourism industry. Tourism has been slow to recover from the

negative effects of COVID-19 and has been hit again by the economic consequences of the Russian invasion of Ukraine. The conflict between Russia and Ukraine has increasingly disrupted critical links in the supply chain. The upward pressure on prices has led to losses in real incomes and thus reduced private consumption and travel demand. To analyse the costs of the pandemic and the war, we used an annual data sample of selected EU countries from 2000-2019 and modelled real tourism exports for each country, estimating them as a function of tourism export prices, export prices of all countries and real GDP of the EU countries included. To estimate the losses between 2020 and 2022, we projected real exports based on the available information for 2019 published by the OECD and the IMF. The first analysis showed that all actual country-specific tourism exports in 2019 were lower than the projected values for 2022. We interpreted the differences between actual and hypothetical performance between 2020-2021 largely as losses caused by the COVID-19 pandemic, while the impact for 2022 was mixed between the waning of the pandemic and the negative economic impact of the war. Comparing the country performance with the average performance, Austria, France, Croatia, Slovenia and Sweden have lower losses than all countries in both periods. Only in 2020-2021 do Belgium, Germany and the Netherlands also perform better than the average. A separate analysis of 2022 shows that Ireland, Italy and Portugal also outperform all countries. The diversity of country-specific recovery outcomes can be explained by differences in the potential impact of key resilience factors of socio-economic systems, such as economic, social and tourism diversity and human and social capital.

Tourism Demand Forecasting In A Post-Pandemic World: A Tvp-Lasso-Midas Model Approach

Presenter: NA
Co-authors: Long Wen;Ying Liu;Yanting Cai

The global reopening presents both great opportunities for tourism recovery and considerable challenges for decision-makers in accurately capturing trends and fluctuations in tourist arrivals. Traditional time series and econometric models are inadequate in providing reliable and real-time forecasts due to their reliance on low-frequency official statistics with long release delays. In contrast, high-frequency big data, such as user-generated content and search index, offer a more timely and accurate reflection of tourist behavior.While the mixed data sampling (MIDAS) model is suitable for dealing with mixed-frequency data, the information extracted from big data is often high-dimensional, and it cannot be directly accommodated into the MIDAS model. In addition, the MIDAS model assumes static parameters which may fail to capture the dynamic nature of the relationship between the low-frequency and high-frequency variables. To address this, this study proposes a new model, by integrating least absolute shrinkage and selection operator (LASSO) and the time-varying parameter (TVP) techniques into the MIDAS model, named TVP-LASSO-MIDAS. This new model offers the advantages provided by both LASSO and TVP, which can reduce dimensionality with minimal loss of information and accommodate dynamic relationships. This study uses big data to forecast tourism demand in a city, and the findings suggest that TVP-LASSO-MIDAS can outperform various baseline models used in the literature.

Tourism Demand Forecasting For Long-Haul And Short-Haul Source Markets In Europe Based On Bayesian Spatial Midas Model

Presenter: Peihuang WU
Co-authors: Peihuang Wu;Anyu Liu;Han Liu

In the evolving field of tourism demand forecasting, the integration of spatial econometric models has proven to be a significant leap forward, particularly with the inclusion of spatial lag terms that account for tourists' propensity to visit multiple destinations within a single journey. This innovative approach is premised on the observation that spatial dependencies significantly influence tourists' travel behaviors, a fact supported by numerous studies. Yet, despite these advancements, two critical gaps persist in the literature: the challenge of data frequency mismatch in spatial modeling and the scant empirical evidence on the relationship between travel distance and the intensity of tourist spatial spillovers.Addressing these gaps,

this study introduces a pioneering model—Bayesian Spatial Mixed-data Sampling (BS-MIDAS)—which represents the first attempt to synergize spatial econometrics, mixed-data sampling, and Bayesian inference within the realm of tourism demand forecasting. This model is meticulously applied to forecast quarterly inbound tourism demand for fifteen European destinations, classified into short-haul and long-haul categories based on the source market's travel distance. The explanatory variables for this model include quarterly GDP (as a proxy for tourists' income), monthly relative tourism pricing between destinations and their source markets, and weekly high-frequency search engine data.Our comprehensive comparative analysis pits the BS-MIDAS model against several benchmark models across different frequency domains, including time series models, mixed frequency models, and spatial models, to rigorously evaluate its forecasting prowess. The empirical findings of this study are multi-fold: spatial models outperform non-spatial counterparts in most scenarios; models utilizing original frequency data eclipse those aggregating data into a single frequency; Bayesian estimation significantly enhances the spatial models' forecasting accuracy; and notably, the impact of spatial dependence on tourism demand forecasting is more pronounced for long-haul destinations, suggesting a stronger inclination among tourists to explore neighboring areas.By fusing spatial econometrics with mixed-data sampling and Bayesian techniques, this study not only bridges significant research gaps but also sets a new benchmark in tourism demand forecasting, offering invaluable insights for both academics and practitioners in the field.

Zero-Shot Forecasting With Natural Language Contextualization

Presenter: Arjun Ashok
Co-authors: Arjun Ashok

Previous work on transformer-based forecasting models has culminated in the development of "general-purpose models" that can handle several stylized facts of real-world time series problems, such as arbitrarily complex data distributions, heterogeneous and irregular sampling frequencies, missing data etc. However, these models rely solely on historical data, and require a huge amount of it for effective training. In contrast, many forecasting applications have very few historical data points with which accurate forecasts need to be produced, a setting formally known as "zero-shot forecasting". We put forward several desiderata that a zero-shot forecasting model should satisfy, such as the ability to produce accurate forecasts on "unseen" time series that the model has not been trained on, the ability to condition prediction on unseen covariates, the ability to accurately predict and leverage multivariate dependencies between an unseen set of time series etc., all with no historical data. Acknowledging that this is an extremely difficult problem to solve, we hypothesize that the model could benefit from contextual and environmental information from other modalities. Specifically, we focus on the use of natural language to flexibly encode diverse kinds of contextual information into the model, such as constraints (such as non-negativity or other domain-specific boundaries) that the model should adhere to, information about the unseen covariates, meta-information about each time series that may aid the inference of multivariate dependencies, etc. Such an interface closes the gap between the human experimenter and the model, facilitating effective zero-shot forecasting.

Size And Speed Matter: Accurately Forecasting 10^9 Time Series Leveraging Zero-Shot Architectures

Presenter: Azul Garza
Co-authors: Azul Garza;Max Mergenthlaer;Cristian Challu

In this paper, we propose a novel approach for unprecedented large-scale forecasting tasks. By leveraging the inference-only, also known as zero-shot inference, nature of foundation models and new developments in distributed GPU-based computation, we explore large-scale forecasting and anomaly detection in massive datasets. We benchmark speed, efficiency, and accuracy on a dataset of one billion unique time series of varying lengths, frequencies, and properties. This work contributes to expanding the boundaries previously set by local and global models that followed the one or many models per dataset approach, and opens the door for new and exciting applications for massive data and extremely fast forecasting needs. The code and dataset used in the experiments will be made public to the community

Chronos: Learning The Language Of Time Series

Presenter: Lorenzo Stella
Co-authors: Lorenzo Stella;Caner Turkmen

We introduce Chronos, a simple yet effective framework for pretrained probabilistic time series models. Chronos tokenizes time series values using scaling and quantization into a fixed vocabulary and trains existing transformer-based language model architectures on these tokenized time series via the cross-entropy loss. We pretrained Chronos models based on the T5 family (ranging from 20M to 710M parameters) on a large collection of publicly available datasets, complemented by a synthetic dataset that we generated via Gaussian processes to improve generalization. In a comprehensive benchmark consisting of 42 datasets, and comprising both classical local models and deep learning methods, we show that Chronos models: (a) significantly outperform other methods on datasets that were part of the training corpus; and (b) have comparable and occasionally superior zero-shot performance on new datasets, relative to methods that were trained specifically on them. Our results demonstrate that Chronos models can leverage time series data from diverse domains to improve zero-shot accuracy on unseen forecasting tasks, positioning pretrained models as a viable tool to greatly simplify forecasting pipelines.

Predicting Clinical Events From The Underlying Patient Latent State Trajectory Using Collaborative Learning

Presenter: Hollan Haule
Co-authors: Hollan Haule;Javier Escudero

Predicting disease episodes from physiological time series recorded from patients in Intensive Care Units (ICU) is a crucial goal in healthcare research. This aims to ensure that, through appropriate prompt personalised treatment, patients maintain a stable condition and maximize their chances of recovery. The extensive, high-dimensional data produced in the ICU presents an opportunity to employ machine learning (ML) techniques for understanding individual disease trajectories and supporting clinical research and eventually decision-making. Existing ML techniques for clinical event prediction typically employ models that directly map from observation space (observation window of history time series) to event space (class label of the observation window at some time horizon). While this approach works well in building event detection systems, we hypothesize that learning from the underlying disease trajectory and leveraging patient similarities will enhance our understanding of diseases and result in more accurate predictions. Here, we introduce a novel approach that first transforms physiological time series from the observation space into a latent embedding space. This process is designed to learn patient similarities before projecting onto the event space. This step is based on our very recent work on collaborative learning and detection of episodes from multivariate time series, which maximizes similarities between embeddings of multivariate time series samples from common disease states across patients. Next, we employ LSTM (Long Short-Term Memory networks) for multistep forecasting on the learned embeddings. Finally, we project the forecasted embeddings to event space and evaluate the performance at different time horizons. We test our approach on the task of predicting intracranial hypertension for a range of prediction horizons in paediatric patients with Traumatic Brain Injury using KidBrainIT dataset, which was routinely collected across several paediatric ICU. Overall, our novel approach predicts clinically relevant episodes leveraging similarities in the trajectory of the underlying latent space learnt to reflect similarities across multivariate time series observations between subjects. Our approach can potentially be applied to any task involving prediction of binary states.

A Multi-Region Discrete Time Chain Binomial Model For Infectious Disease Transmission

Presenter: Siuli Mukhopadhyay
Co-authors: Siuli Mukhopadhyay;Pallab Sinha

Conventional mathematical models of infectious diseases frequently overlook the spatial spread of the disease concentrating only on local transmission. However, spatial propagation of various diseases have been noted between geographical regions mainly due to the movement of infectious individuals from one region to another. In this work, we propose a multi-region discrete time chain binomial framework to model dependencies between the multiple infection time series from neighbouring regions. It is assumed that the infection counts in each region at various time points is not only governed by local transmissions but also by interactions of individuals between spatial units. Effect of intervention strategies like vaccination campaigns used in disease control and various other socio-demographic factors like, live births, population density, vaccination coverage, disruption in disease surveillance due to covid have been taken into account while modelling the multiple infection time series. For estimating this multi-region chain-binomial model with tunable sparsity, an appropriate likelihood function maximization approach regularized with an l1-type constraint is proposed. Simulation results considering effects of seasonal pattern in disease outbreaks, out of sync outbreaks in connected geographical regions, variations in reporting rates in connecting regions; have been considered to depict realistic disease scenarios. Forecasting of infections and effect of spatial heterogeneity on future infections have also been studied. A real world application based on measles counts from adjoining spatial regions is presented to motivate the proposed modelling approach.

Analysis Of The Economic Impact Of Chronic Diseases In The United States: A Study Of Costs And Economic Consequences

Presenter: Odra Saucedo-Delgado
Co-authors: Odra Saucedo-Delgado;Maria Rosa Nieto

Progress in health care and public health has been a cornerstone in increasing global life expectancy. However, challenges remain. This study focuses on the United States undergoing a rapid epidemiologic transition. Despite medical advances, the healthcare system faces significant hurdles adapting to changing demands and providing quality care for chronic diseases. These degenerative conditions place a significant burden on healthcare systems and society at large. Therefore, understanding the relationship between healthcare expenditures, reported cases, and their economic impact is essential for informed health policy decisions. The primary objective is to examine the economic dimension of chronic disease in the United States. The study illustrates the relationship between spending on chronic degenerative diseases, the number of individuals seeking care, and reported cases. A projected increase in cases is expected to impact overall costs significantly. In addition, the future impact of spending and cases of chronic degenerative diseases is expected to exceed the projected GDP in 2030. The use of national data from respected organizations such as the World Health Organization (WHO), the American Heart Association, the Centers for Disease Control and Prevention (CDC), the National Cancer Institute, and the Medical Expenditure Panel Survey (MEPS) is integral to this analysis. Preliminary findings indicate that increased attention to health care correlates with a decline in real treatment costs. One notable observation is that increased access to chronic disease management reduces marginal costs. While total expenditures are attributed to doctor visits, drug purchases, and age, specific conditions such as chronic obstructive pulmonary disease (COPD), diabetes, and osteoarthritis show a pronounced economic escalation. In summary, the forecasting models confirm that even without accounting for the total economic impact, direct spending on these diseases in real terms has a growing and substantial impact on the economy as measured by GDP.

Nonlinear Dynamic Bayesian Modeling For Disease Outbreak Forecasting

Presenter: Spencer Wadsworth
Co-authors: Spencer Wadsworth;Jarad Niemi

The annual influenza outbreak leads to significant public health and economic burdens making itdesirable to have prompt and accurate probabilistic forecasts of the disease spread. The United StatesCenters for Disease Control (CDC) hosts annually a national flu forecasting competition which hasled to the development of a variety of flu forecast modeling methods. For several seasons the targetto be forecast was weekly influenza-like illness (ILI) but in 2021 the target was changed to weeklyhospitalizations. Modeling

hospitalization data alone is limited to the relatively small amount of data,but may be supplemented by existing ILI data which has been reported, published at the state level,and successfully forecast for nearly a decade. In this manuscript we introduce a two componentmodeling framework for forecasting weekly hospitalizations. The first component is for modeling ILIdata using dynamic nonlinear Bayesian hierarchical model. The second component is for modelinghospitalizations as a function of ILI. ILI forecasts are then used to inform hospitalization forecasts.We compare hospitalization forecasts from our model to other related models submitted to the CDCforecasting competition

Strategic Insights For Retail Growth: Forecasting Retail Trade Sales In North Carolina's Dynamic Market

Presenter: Imran Arif
Co-authors: Imran Arif

This study focuses on predicting retail trade sales in North Carolina, acknowledging the sector's substantial contribution to the state's economy, which comprises approximately $62.9 billion or 9% of the total North Carolina GDP and accounts for 0.9 million or 20% of direct jobs in the state. Utilizing monthly historical sales data from January 2012 to December 2023 obtained from the North Carolina Department of Revenue, the study forecasts retail sales for the next 12 months. The empirical methodology employs univariate and multivariate analytical methods, considering demographic trends and macroeconomic indicators to project North Carolina's retail trade sales trajectory for the upcoming 12 months. Consequently, it provides actionable insights for strategic planning, inventory management, and marketing strategies for the retail trade sector, policymakers, and investors.

Hybrid Deep Neural Networks With Skip Connections And Feature-Based Clustering For Supply Chain Demand Forecasting

Presenter: Bilel Abderrahmane BENZIANE
Co-authors: Bilel Abderrahmane Benziane;Benoit Lardeux;Maher Jridi;Ayoub Mcharek

Enhancing precision in demand forecasting within the supply chain is essential for optimizing business costs by mitigating missed sales opportunities and minimizing excess expenses incurred through overstocking. The contemporary trend in the literature leans towards the integration of neural networks, particularly due to their resilience in handling time series data. To increase the accuracy of forecasting, there is a growing reliance on hybrid models that combine various deep learning modules. However, according to the recent stat of the art, the prevalent issue with these models lies in their serial architectures, leading to challenges such as overfitting, vanishing gradients, and exploding gradients, due to an increasing number of layers. This study introduces a novel hybrid deep neural network, encompassing diverse architectures such as convolutional neural networks (CNN), recurrent neural networks (RNN), attention mechanisms, and feed-forward neural networks (FFNN). To deal with the aforementioned challenges, the model incorporates residual connections, effectively addressing overfitting, vanishing gradients, and exploding gradients. A second key contribution involves the implementation of feature-based clustering, organizing the time series database into more homogenous groups. Each group undergoes separate model training, fostering accuracy in cases where the time series dataset is heterogeneous. The effectiveness of this approach is comprehensively assessed through a comparison with state-of-the-art models, utilizing a set of real demand data. The results underscore the robustness of the proposed methodology, highlighting significant improvements. When compared with FFNN, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional LSTM (BiLSTM), Boosting-CNN-LSTM-FFNN, Attention, Median, and Deep Reinforcement Learning models, the Symmetric Mean Absolute Percentage Error (SMAPE) of the proposed approach decreases by 37%, 41%, 42%, 40%, 44%, 43%, 36%, and 43%, respectively.

Hierarchical Neural Additive Models For Interpretable Demand Forecasts

Presenter: Leif Feddersen
Co-authors: Leif Feddersen;Catherine Cleophas

Demand forecasting is a critical element in business decision-making, influencing numerous business decisions from inventory control to strategic planning. The adoption of machine learning (ML) techniques has been particularly useful for this task, especially when dealing with covariate-rich datasets that traditional statistical methods may struggle to process effectively.However, the lack of interpretability and user acceptance of ML models remains a significant challenge. Addressing this issue, we introduce Hierarchical Neural Additive Models (HNAM) for time series forecasting. Expanding upon Neural Additive Models (NAM), HNAM incorporates a time-series-specific additive model that accounts for trend and interactions between covariates based on a user-defined interaction hierarchy. Feature interactions are only allowed with covariates that are lower in the hierarchy.For example, HNAM may estimate the effects of weekdays independently of other covariates while holiday effects also depend on the weekday, and promotional effects on both former covariates.This structure lets analysts clearly discern and interpret individual covariate effects. Our evaluation compares HNAM against both state-of-the-art neural networks and conventional statistical models using three real-world retail datasets. HNAM demonstrates competitive predictive performance relative to more complex and slower-training neural networks, and significantly surpasses traditional statistical methods. The model's interpretability facilitates plausible explanations for covariate influences on forecast outcomes.HNAM represents a step forward in achieving the dual goals of interpretability and predictive accuracy in demand forecasting. We acknowledge potential limitations and suggest avenues for future research, including behavioral studies to validate the model's effectiveness in practical decision-support scenarios.

Demand Forecasting With Ml: A Deep-Dive Into The Various Loss Functions

Presenter: Rijk van der Meulen
Co-authors: Rijk Van Der Meulen

Machine Learning models are used more and more for supply chain demand forecasting. How well these models work in practice depends on factors like feature quality, hyperparameters, and the loss function chosen. In this talk we'll focus on the latter; exploring different loss functions, like RMSE, MAE, Tweedie, and Huber. We'll dive into how they work and in which situations they might be useful.

Penalized Linear Regression Models For Interval-Valued Data

Presenter: Haowen Bao
Co-authors: Haowen Bao

With the advent of complex information systems to collect massive data, interval-valued data modeling and forecasting have drawn increasing attention in economics and statistics. However, few works have considered interval-valued variable selection, especially when the number of potential predictors is divergent. We propose a penalized interval regression model via adaptive LASSO to select important interval-valued variables and estimate coefficients simultaneously, where adaptive weights penalize different coefficients in the L1-penalty. We find that the penalized interval regression estimators enjoy the oracle properties under the regularity conditions. We further extend our concern to the divergent-dimensional interval regression and establish associated asymptotic theories. Simulation studies show that the proposed model has good finite sample properties and outperforms the full model. Empirical applications to interval-valued crude oil prices forecasting and interval-based S&P100 index-tracking highlight the merits of the proposed method relative to other competing approaches.

Probabilistic Forecasting Of Intermittent Time Series With Tweedie Likelihoods

Presenter: Stefano Damato
Co-authors: Stefano Damato;Dario Azzimonti;Giorgio Corani

Intermittent time series are common in applications such asretail chain supply and spare parts replenishment. They are characterisedby a large number of zeros; it is therefore necessary forecasting both theoccurrence and the level of the demand. Most approaches for forecastingintermittent time series only provide point forecasts.The Tweedie loss is suitable for modelling zero-inflated data. Withinthe M5 competition, it has been used for training tree-based models withgood results.We propose to use the Tweedie distribution as a likelihood functionin a Gaussian Process (GP) model optimised with variational inference.This model produces probabilistic forecasts which are more insightful thansimple point forecasts; this is especially true for intermittent time series,whose predictive distribution is asymmetric and potentially multimodal.While most methods for intermittent time series predict separatelythe occurrence and the level of the demand, the Tweedie distribution al-lows jointly modelling both the occurrence and the level of the demand.The Tweedie distribution is characterized by compound Poisson gammadistributions and thus is zero-inflated. In particular, we choose a param-eterization where both the probability at zero and the mass at positivevalues are controlled as functions of the same parameters, thus makingthe model more flexible. We first discuss the choice of prior distributionsfor parameters through prior predictive checks and then show, on sev-eral standard datasets, that our model achieves competitive performancesversus state-of-the-art techniques.

Arma Estimation Of Intermittent Data

Presenter: Giacomo Sbrana
Co-authors: Giacomo Sbrana

When forecasting intermittent, or count data, it is relevant to ensure that the forecasts are non-negative. Despite their wide spread, in general, ARMA models do not ensure non-negativity. By addressing this issue we provide sufficient conditions that allow a generic ARMA model to produce non-negative filters/forecasts. Moreover, by enforcing non-negativity, we show that the predictive accuracy may be greatly improved. Indeed, using Walmart sales data, as used in the M5 competition, we show that a simple ARMA(1,1) clearly outperforms all standard benchmarks, as well as their combinations, scoring among the top teams, both in point forecasts and prediction intervals. This impressive result confirms the strong ability of ARMA models to predict also intermittent data, provided that non-negativity is respected.

Enhancing Inventory Management Through Tobit Exponential Smoothing And Time Aggregation

Presenter: Diego J. Pedregal
Co-authors: Diego J. Pedregal;Juan R. Trapero

This paper introduces the Tobit Exponential Smoothing model (TETS) with time aggregation constraints, providing a flexible framework for handling censored time series data, particularly useful in demand forecasting scenarios. The TETS model, a variant of the Tobit Innovations State Space system, is formulated to address censoring from above in observed output time series, such as sales data, where a known censoring level is imposed. The model's versatility extends to accommodating various modifications, including exogenous regression terms and heteroscedastic noise. Furthermore, the paper proposes a method for incorporating time aggregation into Tobit Innovations State Space systems, enabling the handling of aggregated data without loss of information. The aggregation period is defined, and the state equation is augmented to include observation equations reflecting cumulative aggregated data. This enables the modelling of aggregated censored variables while maintaining the underlying structure of the state space framework. Case studies demonstrate the efficiency of the proposed approach in practical scenarios. In a simulated demand setting, the TETS model effectively captures the true demand, even under stringent censoring conditions. Comparisons with standard methods highlight the TETS model's ability to mitigate

bias and produce more accurate forecasts, crucial for inventory management decisions. Moreover, the paper addresses challenges such as the spiral-down effect, where biased estimation leads to suboptimal inventory policies over time. Strategies for mitigating this effect, including replacing lost sales observations and employing the TETS model, are discussed. Results indicate that the TETS model significantly outperforms conventional methods in terms of minimizing lost sales and excess stock, thus offering substantial cost savings. In conclusion, the proposed Tobit Exponential Smoothing model with time aggregation constraints presents a robust framework for handling censored time series data, particularly in demand forecasting applications. Its flexibility and effectiveness make it a valuable tool for improving inventory management practices and decision-making processes.

Demand planning and the role of judgment in the new world of AI/ML

Presenter: Conor Doherty
Co-authors: Conor Doherty;Sven Crone;Alexey Tikhonov;Nicolas Vandeput;Robert Fildes;Paul Goodwin

This panel of speakers from industry and academia are all participants in the AI/ ML revolution that is affecting the supply chain and in particular, demand planning. But do the claims made by some as to the ability of AI/ML to automatically improve the forecasting accuracy and performance of demand planners stack up? Can AI/ ML outperform statistical methods even when enhanced by the judgmental interventions of the demand planners? Are demand planners becoming redundant? Following on from the earlier presentation by Fildes and Goodwin on forecast value added, this panel discussion will start with brief remarks from Crone, Tikhonov and Vandeput offering their perspectives on this important problem. After questions from the audience, the session will conclude with a brief retrospective from Fildes and Goodwin as to ways forward to achieve consistent improvements in accuracy and bias.

Bloody January Again

Presenter: Roy Batchelor
Co-authors: Roy Batchelor

Shocks to the global economy from covid, conflict and commodity prices have led to uncertainty about the levels at which inflation and growth will stabilize in the long run, and doubts about the credibility of tight central bank targets and official projections for public debt in developed economies. This paper uses a simple method to infer forecasters' steady state expectations for inflation and growth from their one- and (newly revealed in January) two-year ahead predictions published in the Consensus Economics service. Indicator saturation and structural models for UK data suggest dates for breaks in these steady states, and the possible causes of breaks - effectively testing whether inflation targets have proved credible in anchoring long term expectations (yes), and whether growth expectations have been permanently dented by major recessions and other shocks, including Brexit (yes), Covid and its aftermath (no). The models suggest that for the UK official growth forecasts do not influence private sector forecasts, but inflation forecasts are strongly tied to central bank targets, even in the most recent high inflation years.

An Investigation Into The Uncertainty Revision Process Of Professional Forecasters

Presenter: Michael Clements
Co-authors: Michael Clements;Robert Rich;Joseph Tracy

Following Manzan (2021), this paper examines how professional forecasters revise their uncertainty (variance) forecasts. We show that popular first moment "efficiency" tests are not applicable to study variance forecasts and instead employ monotonicity tests developed by Patton and Timmermann (2012). We find strong support for the Bayesian learning prediction of decreasing patterns in the variance of fixed-event density forecasts and their revisions as the forecast horizon declines. We explore the role of financial conditions indices in variance forecasts and document their predictive content for the revision

process of US professional forecasters, although the evidence is weaker for euro area forecasters.

Time-Varying Us Government Spending Anticipation In Real-Time

Presenter: Pascal Goemans
Co-authors: Pascal Goemans;Robinson Kruse-Becher

Due to legislation and implementation lags, forward-looking economic agents might anticipate changes in fiscal policy variables before they actually occur. The literature shows that this foresight poses a challenge to the econometric analysis of fiscal policies. However, most of the literature uses fully revised data to quantify the degree of fiscal foresight. We, in contrast, use the Survey of Professional Forecasters (SPF) and the Real-Time Data Set for Macroeconomists to analyze government spending anticipation. We also consider relative fiscal foresight compared to other national account variables. We find that real-time data matters for the predictability of federal and state & local government spending. We document remarkable time-variation in the foresight of the SPF and its information advantage against (augmented) autoregressive models over different US presidencies. This finding highlights the relevance of policy communications.

Corporate Profits As A Share Of Gdp: Forecasts And Data Revisions

Presenter: Dean Croushore
Co-authors: Dean Croushore

A primary driver of stock prices is corporate profits. If stock-market investors are rational, their demand for corporate stocks should depend on their forecasts of corporate profits. So, in this paper, I look at the forecasts of corporate profits made by professional forecasters to test their quality. Corporate profits are quite volatile over the business cycle, making the forecasts subject to substantial error, especially longer-term forecasts. In addition, data on corporate forecasts are revised substantially, making the series even more difficult to forecast. In this paper, I will establish facts about the quality of the forecasts and the nature of the data on corporate profits. I find that forecasts of corporate profits, using the Survey of Professional Forecasters (SPF) as a share of GDP are biased in some dimensions. However, the main source of the bias is data revisions, especially benchmark revisions of the National Income and Product Accounts. A forecast-improvement exercise shows no ability for anyone to make the forecasts better in real time.

Oil Price Dynamics

Presenter: Sasheendran Gopalakrishnakone
Co-authors: Sasheendran Gopalakrishnakone;Venkata L. Raju Chinthalapati

This paper provides a fivefold contribution to the literature: First, univariatewavelet decomposition of oil price time series data reveals the dynamics at scalelevel. Filtering into low medium and high frequency components differentiatesbetween economic fundamentals and financial speculative shocks driving the oilprice. Second, Multiresolution analysis (MRA) in the frequency domain showsthat medium to long term investment drives prices in contrast with short termspeculative trading activity as is commonly thought. Identification of the D8wavelet component (20-40 year time cycle) in oil time series data as significantcontributor to the overall oil price variation thus confirming Kuznets infrastruc-ture cycles hypothesised in previous work. Third, the time varying influencesof factors are revealed through specific frequencies when used in conjunctionwith linear regression models. Proxies for global economic activity specificallyinfrastructure-related commodities are the most significant determinants of oilprice variation. Fourth, transfer function shows that infrastructure commoditiesare also significant in forecasting oil prices. Fifth, steel prices show significancein all the models and is an important indicator of global real economic activity.

What Drives The European Carbon Market? Macroeconomic Factors And Forecasts

Presenter: Elisabetta Mirto
Co-authors: Elisabetta Mirto;Andrea Bastianin;Luca Rossini;Yan Qin

We tackle the issue of producing point, sign, and density forecasts for the monthly real price of carbon within the European carbon market, EU ETS. We show that a Bayesian Vector Autoregressive (BVAR) model, augmented with factors based on economic fundamentals, provides accuracy gains over a set of benchmark forecasts, including survey expectations and forecasts released by data providers. We also consider verified emissions and demonstrate that adding stochastic volatility can further improve the forecasting performance of a single-factor BVAR model. We rely on forecasts to build market monitoring tools that track demand and price pressure in the EU ETS.

High-Frequency Density Nowcasts Of Co2 Emissions In U.s. States

Presenter: Andrey Ramos
Co-authors: Andrey Ramos;Ignacio Garrón

This paper introduces a novel approach to panel density nowcasting, employing high-frequency data to yield timely predictions of CO2 emissions and energy consumption growth for all U.S. states. Using weekly state-level economic conditions and adopting a mixed-frequency model, the economic data are initially utilized to forecast energy consumption growth. Subsequently, this information is used to predict CO2 emissions growth quantiles, accounting for cross-sectional dependence across states through estimated factors. To capture both asymmetry and excess kurtosis, a skew-t distribution is fitted. The performance of the models is rigorously assessed through an out-of-sample forecasting study, revealing that the weekly nowcasts offer early insights into CO2 emissions and energy consumption growth across U.S. states. Our study contributes to the existing literature on several ways. First, the quantile approach adopted to produce CO2 emissions growth nowcasts allows us to provide a complete description of the uncertainty associated with the predictions. Related studies as Bennedsen et al. (2021) and Fosten and Nandi (2023) only produce point predictions at the national and state level, respectively. Second, the use of the weekly state-leveleconomic indicators of Baumeister et al. (2022) implies a higher frequency in the nowcasting exercise compared to the quarterly and monthly frequencies of Bennedsen et al. (2021) and Fosten and Nandi (2023), respectively. The information produced by our analysis is useful to inform policies aiming to reduce regional environmental degradation. The nowcasting exercise is particularly important in this context, given the large publication lags in the U.S. state-level energy consumption and emissions data.

Understanding The Future Of Critical Raw Materials For The Energy Transition: Svar Models For The U.s. Market

Presenter: Ilenia Gaia Romani
Co-authors: Ilenia Gaia Romani;Chiara Casoli

We examine the impact of energy transition policies on the U.S. markets of three critical minerals used for batteries, namely cobalt, lithium and nickel. To achievethis, we estimate three Structural Vector Autoregressive models, disentangling supply and demand shocks at the aggregate and mineral-specific level. We then performa structural forecast analysis to study mineral price patterns under various demand and supply scenarios up to 2030. Specifically, we investigate the implications of theU.S. Inflation Reduction Act (IRA) and the associated policies aimed at boosting the domestic production of these critical minerals, combining them with variousdemand projections. Our findings suggest that, whereas cobalt and lithium prices could decrease conditional on the successful implementation of energy transitionpolicies in the U.S., nickel price most likely will remain high.

Deep Learning The Persistence Structure Of Economic Variables

Presenter: Luboš Hanus
Co-authors: Lubos Hanus;Jozef Barunik;Lukas Vacha

We develop a novel deep learning method tailored for forecasting a time series. The newly proposed approach explores the dynamics of the heterogeneously persistent structure of the shocks driving the time series. We identify the time-varying structure and persistence of time series with different degrees and provides more reliable and explainable forecasts. We will use different machine learning techniques to account for recurring and non-linear structures of time series. As a showcase, we demonstrate the performance of our dynamically informative predictors against state-of-the-art benchmarks on US inflation data through an out-of-sample forecasting exercise.

Multi-Period Forecasting Of Macro Variables At Risk Using Parsimonious Neural Networks

Presenter: Sicco Kooiker
Co-authors: Sicco Kooiker;Julia Schaumburg;Lukas Hoesch

We propose to forecast the quantiles of the conditional distribution of macroeconomic variables at multiple horizons for multiple countries using flexible dynamic encoder-decoder models. By parameter-sharing between encoders of different countries we provide a parsimonious model that is applicable in typical sample sizes of macroeconomic data sets. The encoder part of the model, which is identical for each of the countries, captures nonlinearities in the data and allows for the extraction of shared dynamics, while the decoder is country-specific and aims to explain the country-individual effects. Our estimation routine combines the temporal convolutional neural network (TCN) as an encoder with an output layer that enforces non-crossing quantile estimates by penalization. The TCN provides a parameter-efficient model architecture for extracting complex temporal dependencies between variables. The multi-quantile, multi-horizon loss function is minimized using a stochastic gradient descent algorithm. In the simulation study, we open the black box of the neural network by comparing isolated model variants and establishing the hyperparameter importance. We study the finite sample properties by considering a range of linear and nonlinear data generating processes. The simulation study demonstrates that our neural network design outperforms the linear quantile regression, even in small sample sizes. We further find that the encoder-decoder structure and parameter-sharing lead to a more efficient modeling approach for (non)linear dependencies between predictors and quantiles. In an empirical illustration, we analyze the out-of-sample performance of the model on inflation data. We find that the method outperforms linear quantile regression in predicting the economic vulnerability of 21 EU countries for 6, 9, and 12-month-ahead forecasting during the pseudo-out-of-sample period 2014-2023.

Enhancing Short-Term Inflation Forecasts: Bottom-Up Approach With Machine Learning

Presenter: Dongjae Lee
Co-authors: Dongjae Lee

In recent years, the application of machine learning techniques to forecast macroeconomic indicators, including inflation and GDP, has gained prominence. These methods have consistently outperformed traditional time series models, particularly in capturing rapid economic shifts observed during recent pandemics. Standard machine learning forecasting models for aggregate variables have predominantly focused on the relationships and inherent dynamics within aggregated macroeconomic data. This approach, however, may obscure the intricate dynamics at the item level. Importantly, recent trends in inflation have increasingly been influenced by idiosyncratic shocks specific to certain industries or product categories, diverging from the influences of the broader economic cycle. This shift poses substantial challenges for predictive models that rely solely on aggregated data.Addressing this complexity, our research introduces a

novel bottom-up forecasting approach. This method utilizes a machine learning framework, incorporating the Boruta feature selection algorithm and random forest models, to predict the individual components of Korea's Consumer Price Index before their aggregation. Empirical tests on historical data reveal that our bottom-up approach secures modest yet significant improvements in forecasting aggregate inflation in Korea, surpassing the conventional method of using machine learning to directly forecast aggregated variables. Our findings highlight the potential of this short-term forecasting method to refine the prediction of macroeconomic variables, providing detailed insights into the dynamics at the sub-level, such as GDP.

Satellites Turn "Concrete": Tracking Cement With Satellite Data And Neural Networks

Presenter: Baptiste Meunier
Co-authors: Baptiste Meunier;Benjamin Lietti;Jean-Charles Bricongne;Alexandre D'aspremont;Simon Ben Arous

This paper exploits daily infrared images taken from satellites to track economic activity in advanced and emerging countries. We first develop a framework to read, clean, and exploit satellite images. Our algorithm uses the laws of physics (Planck's law) and machine learning to detect the heat produced by cement plants in activity. This allows us to monitor in real-time whether a cement plant is working. Using this information on around 500 plants, we construct a satellite-based index tracking activity. We show that using this satellite index outperforms benchmark models and alternative indicators for nowcasting the production of the cement industry as well as the activity in the construction sector. Comparing across methods, we find neural networks yields significantly more accurate predictions as they allow to exploit thegranularity of our daily and plant-level data. Overall, we show that combining satellite images and machine learning allows to track economic activity accurately

Beyond Day-Ahead With Econometric Models For Electricity Price Forecasting

Presenter: Paul Ghelasi
Co-authors: Paul Ghelasi;Florian Ziel

The recent surge in global power prices has heightened interest in mid to long-term forecasting for hedging and valuation. While short-term predictions like intra-day and day-ahead forecasts benefit from well-researched econometric models with established structures and reliable regressors, extending them to longer horizons presents challenges. Over time, short-term variables lose their predictive power, which can lead to inconsistencies in model coefficients. These inconsistencies may manifest as spurious effects and unexpected coefficient signs that are not aligned with fundamental economic theory.This study delves into power price predictability across various time horizons, from one day to one year ahead. To overcome the limitations of econometric models, we impose fundamental constraints on coefficients, derived from energy economics principles. This approach achieves more stable coefficient estimates and enables us to identify the dynamics of key influencing factors for different forecasting horizons. Next, we propose a method to integrate short-term regressors such as infeed from renewables and load into a long-term framework by separating model estimation from forecasting. By training the model on same-day data and forecasting expected variable levels, we capture same-day effects in-sample, while out-of-sample forecasts rely on average effects. We then examine the unit root behavior of power prices and its implications. Differentiation methods do not enhance predictive power when commodity prices are used as regressors, indicating the influence of commodity prices on power price unit root behavior. Furthermore, unit-root regressors are more prone to generating spurious effects.We conduct a forecasting analysis on hourly German day-ahead power prices, employing regularized regression methods and generalized additive models. Initial findings suggest an enhanced forecasting accuracy compared to simple benchmarks, though diminishing returns are anticipated with increasing forecasting horizons.

Managing Market Risk For A Small Renewable Energy Generator Using Leave-K-Out Sampling-Based Forecasts

Presenter: Weronika Nitka
Co-authors: Weronika Nitka;Katarzyna Maciejowska;Tomasz Weron

In this article, a leave-k-out sampling method for a joint probabilistic forecasting of a set of variables is proposed. The approach allows to predict a multidimensional distribution that maintains the correlation structure of the selected variables. Its advantages include a relatively low computational cost and a capability to be used jointly with a variety of point forecasting approaches.The accuracy of the novel forecasting method is evaluated with short-term forecast study of data describing German electricity spot market: day-ahead and intraday prices together with market fundamentals such as load, generation from renewable energy sources or residual demand. The results indicate that the proposed sampling method provides probabilistic predictions that are highly accurate and more reliable than a well-established quantile regression benchmark. The accuracy gains are the largest when functions of forecasted variables, such as price spread, residual load or trading revenue, are considered.Finally, the method is used to support a decision-making process of a small renewable energy farm that sells energy in either a day-ahead or an intraday market. The utility makes decisions under high uncertainty as it knows neither the future production level nor the prices. We show that joint forecasting of both market prices and fundamentals can be used to predict the distribution of the utility's revenue and hence helps to design a trading strategy that balances levels of profits and market risk.

Day-Ahead Electricity Price Forecasting Based On Residual Load Predictions For The Merit-Order Model In The German Market.

Presenter: Johannes Wagner
Co-authors: Johannes Wagner;Jan Koltermann;Sebastian Mieck

In the evolving landscape of energy markets, accurate price forecasting is critical for stakeholders to make informed decisions. Our modelling approach aims to address the volatility and unpredictability of energy prices, which are influenced by various factors including supply and demand dynamics, renewable energy integration, and market regulations.The merit-order effect plays a crucial role in determining energy prices in competitive markets. It ranks energy sources based on their marginal costs, with the least expensive sources being utilized first. Traditional models often fall short in accurately predicting prices due to the complex interplay of market forces and the increasing share of intermittent renewable energy sources. We present a merit-order based modeling approach, derived from the EWI Merit Order Tool, shifting the predictive modeling from the electricity price itself to the residual load, i.e. the total demand after subtracting renewable energy production, and conventional production forecasting. This approach resembles the actual pay-as-clear market mechanism. Our model combines the merit-order approach with machine learning algorithms to forecast short-term day-ahead electricity prices. We employ publicly available German power-plant availability, fuel price, weather and other online data sources. The methodology is validated through backtesting with historical market data, comparing the model's predictions with actual prices to assess accuracy. Preliminary results demonstrate that our model performs well, especially with high shares of renewable energy.

Probabilistic Forecasting Of Electricity Prices With Isotonic Distributional Regression

Presenter: Arkadiusz Lipiecki
Co-authors: Arkadiusz Lipiecki;Bartosz Uniejewski;Rafał Weron

Forecasting electricity prices is an essential tool to support decision making by electricity market participants. However, simple point forecasts provide only limited information. Therefore, probabilistic forecasts, represented as distributions of future prices, are gaining more attention as they allow to assess

and mitigate the risks associated with the high volatility of electricity prices. In my talk, I will discuss models for probabilistic forecasting of day-ahead electricity prices that rely on point forecasts as input. This approach allows us to borrow the complexity of expert point forecasting models and build relatively simple probabilistic forecasting models. In particular, I will focus on the application of stochastic order and discuss isotonic distributional regression, which has not been previously studied in electricity markets.

Enfobench: A Community-Driven Energy Forecasting Benchmark

Presenter: Attila Balint
Co-authors: Attila Balint;Johan Driesen;Hussain Kazmi

Forecasting plays a crucial role in the energy domain, including energy management, renewable energy integration, and grid stability. Despite the increasing prevalence of forecasting competitions aimed at benchmarking models and gaining knowledge, they often face challenges in properly evaluating model performance due to limited forecast horizons and testing periods. Moreover, these often focus solely on accuracy and give little consideration to other aspects such as computational costs or downstream utility.In this talk, we introduce EnFoBench, a novel open-source benchmarking toolkit designed to address these challenges and facilitate comprehensive evaluation of energy forecasting models in a transparent and reproducible manner. EnFoBench offers a dynamic, open-source framework that enables researchers and practitioners to benchmark forecasting models across various scenarios, with an initial emphasis on electricity load forecasting. The framework contains publicly accessible curated and quality controlled training data, along with readily available exogenous features. Additionally, EnFoBench provides a continuously updated dashboard featuring the latest metrics and models openly accessible to the public.This talk outlines the design principles underpinning EnFoBench and details its evaluation process, which focuses on the entire forecasting pipeline rather than just the model itself. We also present the list of available metrics, allowing for the evaluation of model performance, computational efficiency, and relative performance compared to baseline models in greater detail than in previous works. Finally, this talk also present results from the first round of evaluated models, sourced from our previous work, popular forecasting frameworks, previous competitions and recent foundational models.

A Decision Support System For Evaluating Smart Plug Forecasting Pipelines

Presenter: Giulia Rinaldi
Co-authors: Giulia Rinaldi;Lola Botman;Oscar Mauricio Agudelo;Bart De Moor

This research aims to design a Decision Support System (DSS) to help users evaluate, assess, and compare Artificial Intelligence (AI)-based pipelines for electricity consumption forecasting within a building premises, utilizing time series data from smart plugs. The objective is to support energy data analysts in making informed decisions concerning the adoption of optimal forecasting solutions, particularly pertinent to challenges such as resource allocation in industrial settings, plug scheduling, and overall electrical building efficiency integration within AI systems. In this study, the conventional formulation of a DSS, which comprises an interface, data management subsystem, model management subsystem, and knowledge management subsystem, has been tailored for evaluating time series pipelines. Particular focus was on enhancing the knowledge management subsystem by leveraging domain-specific expertise, statistical metrics, and the system's ability to provide accurate and detailed insights. In addition, the created DSS not only uses classical accuracy evaluations but also inspects the complexity of the pipeline, the computational efficiency, the training time, and the adaptability rules to evaluate both the outcome and the usability of the pipeline on a potential AI system.Additionally, to enhance the quality of the outcomes of the assessment process, the designer's insights are collected through a short questionnaire. The user-system interaction has been considered by incorporating a user-friendly interface. Once the DSS has analyzed the information, it produces a report with its findings and assessment. Presently, the envisaged DSS framework is undergoing its final implementation phase within a software demonstration, slated to be subjected to testing using real-life application scenarios pertinent to plug-level load forecasting. The dataset used is sourced from the University of California, San Diego, comprising readings obtained from over 150 smart plugs deployed

across various office buildings spanning a duration exceeding one year post the COVID era. With the proposed framework, this research is believed to contribute to the field of energy management by providing a comprehensive DSS for industrial energy consumption forecasting and by supplying decision-makers with an evaluation tool which helps them select appropriate, easy-to-incorporate and ready-to-use pipelines in energy organizations to optimize energy usage.

Electricity Market Supply And Demand Curves Forecasting Based On Functional Analysis, Isplines, And Machine Learning

Presenter: Carlo Lucheroni
Co-authors: Carlo Lucheroni;Florian Ziel;Nabangshu Sinha;Martina Zannotti

This paper introduces and explores a functional data analysis framework for forecasting electricity supply and demand curves, and for using these forecasts to better point-forecast electricity prices. Training and test data come from the financial Italian Day Ahead Market (DAM) IPEX.The proposed framework is based on two components. The first component is a suitable embedding space for the data, in the form of a Hilbert space related to coordinates with respect to an ispline basis. In this embedding, supply and demand curves can be represented in a manageable and explainable way as dynamic combinations of the ispline basis elements, which play the role of features, fixed in time. The second component is a forecasting engine, to be applied directly to the dynamic coordinates, in the feature space. This engine can include linear vector autoregressions or simple vector neural networks, which can forecast the dynamics of the curves in many comparable ways.Within this framework, dynamic representations of time series of market curves can be interpreted geometrically as paths on a hypersphere, and this geometric interpretation can help guide towards the selection of the best forecasting engine. The framework is compared with (and benchmarked to) an interesting linear functional autoregression model called FAR, introduced not long ago in the electricity market literature.

Moneyness Anomalies Of Bitcoin Options

Presenter: Bastien Buchwalter
Co-authors: Bastien Buchwalter;Jean-Michel Maeso;Vincent Milhau

This paper investigates the rational pricing of European call and put options of Bitcoin. Utilizing a comprehensive dataset of more than 12 million trades, our analysis assess whether the pricing of these options aligns with the theoretical expectations posited by rational pricing frameworks. We uncover notable moneyness anomalies for call options. Approximately 20% of the observed (valuation date, maturity) pairs demonstrate that the ratio of price to strike price does not increase as expected with moneyness - defined as the ratio of spot price to strike price. Similarly, the price to strike price for put options does not decrease with increasing moneyness contrary to theoretical predictions. However, this only occurs for 3% of put options. Further we introduce the concept of "cross-sectional implied volatility," which we define as the volatility parameter that most accurately aligns Black-Scholes model prices with observed market prices for a given pair. Our findings raise critical questions about the applicability of traditional rational pricing models to the pricing of Bitcoin options and suggest potential avenues for further research into the anomalies observed in the cryptocurrency options market. This study contributes to the burgeoning literature on cryptocurrency markets by providing empirical evidence of deviations from standard pricing models thus offering insights into the unique dynamics of this novel financial landscape.

Forecasting Fear And Greed In The Bitcoin Market Using Extreme Gradient Boosting

Presenter: Atikur Khan
Co-authors: Atikur Khan;Arifur Rahman;Milind Tiwari;Kuldeep Kumar

Trading decision of a trader largely depends on price dynamics and emotional intelligence at the

time of abrupt price changes and in the presence of buying and selling pressures. Many factors affect the trading decision-making process of heterogeneous traders and trading behavior can be greatly influenced by the psychological resistance of traders under different states in market dynamics. Given the fear and greed states of market dynamics are known, traders are likely to have less resistance in trading decision-making. This paper develops a buy and sell pressure index, return neutrality index, and external information search index to predict the fear and greed states in the Bitcoin market. We find that the buy and sell pressure induced features are the most dominant exogenous features and applications of extreme gradient boosting and deep neural network with these features provide improved forecast accuracies with excellent discriminatory power. Methods developed in this paper will enable a trader to obtain a highly accurate forecast of fear and greed states in the Bitcoin market and will equip a trader with a higher level of emotional intelligence in making trading decisions. Moreover, the indices developed have implications beyond individual trading decisions by offering an insight into an approach with a potential to be utilized for monitoring and combatting financial crimes in the cryptocurrency market.

Tone of the Cryptosphere: Anticipating Contagion Risk in Crypto Markets

Presenter: Nicolas Magner
Co-authors: Nicolas Magner;Aliro Sanhueza

We designed a measure of cryptocurrency market tone to predict the contagion risk of this market. We use generative AI to improve available dictionaries of financial terms and measure the tone of this dictionary on major cryptocurrency news sites using web scraping and text mining. We conduct extensive in- and out-of-sample tests to measure the predictive power of crypto pitch, including Newey West regressions, break analysis, quantile regression, impulse response analysis, variance decomposition, and ENCNEW tests. The most important result is that the crypto tone manages to anticipate variations in the risk of contagion in the cryptocurrency market in one day. Specifically, when the crypto tone becomes more negative, the risk of contagion increases the next day. These results are maintained when we include control variables associated by the literature with the risk of contagion, such as the fear and greed index, keyword searches on Google, such as inflation and recession, and other variables, such as dominance and the ratio of market value over the realized value of bitcoin.Additionally, our cryptocurrency tone measure does better at anticipating changes in contagion risk than all of the variables above in both our in-sample and out-of-sample analyses. This work is novel from two perspectives. First, we use a much-cited dictionary of positive and negative words developed by Loughran and MacDonald to measure the tone of corporate reporting and use artificial intelligence to improve it. Second, we measure contagion risk using correlation and graph analysis methodologies, specifically MSTL, a complementary measure to classical variance decomposition and spillover index methods. Our results are essential for investors and regulators better to understand the risk of contagion in the cryptocurrency market and improve their decisions in the face of variations in the contagion probabilities of adverse events.

Informed Decision And Optimal Policy For Cryptocurrency Portfolios

Presenter: Josué Thélissaint
Co-authors: Josué Thélissaint;Franck Martin

We apply the factor modeling to the Cryptomarkets with two main concerns in the spirit of the Maxi-mally Predictable Portfolio framework. Firstly, we address the arbitrage in the decision process whichconsists in weighted choice between the classical framework of rational expectations and the strategic-complementarity regarding allocation decisions for cryptocurrency portfolios. While the former re-flects the well proven policy in asset management when decisions are made based on macroeconomicfundamentals, the latter embodies a new trend of decision process driven by the collective psychologyof investors through emotional expressions via social networks. Secondly, we investigate the existenceand the behavior of common latent factors which drive the observed dynamics of the cryptocurrencydaily returns. Hence, our target of predictability maximization has a double interest. From a practicalpoint of view, it is motivated by the need to have optimal policy for tacticalallocations. From a theoretical perspective, it is essentially

about informational efficiency and the pos-sible empirical foundations for asset pricing theory dedicated to cryptocurrencies. The comprehensionof transmission mechanisms of aggregated fluctuations from the monetary policy or traditional assetsto cryptomarkets serves both theoretical and practical perspective. After a cluster-based dimensionreduction, we investigate the conditional latent factors. The Generalized Dynamic Principal Compo-nent (GDPC) outperforms the simple PCA-based approaches. Specifically in terms of R squared, withfewer factors GDPC explains 35-40% more of variability. Moreover, lag GDPCs remain significant andenable Conditional Factor Modeling which is more convenient when it comes to forecast markets. Ourfindings provide useful insights for practitioners regarding tactical allocations and the rationalizationof strategic complementarity behavior.Keywords: Cryptomarkets, Sentiments Analysis, Factor Modeling, Markets Prediction, MaximallyPredictable Portfolio

Effective Forecasting Techniques For Hotel Revenue Management Across Multiple Seasons

Presenter: Apostolos Ampountolas
Co-authors: Apostolos Ampountolas

Effective demand forecasting is crucial for revenue managers, particularly in industries like hospitality, which face significant uncertainties due to demand and supply dynamics worsened by events such as the recent global pandemic. This study investigates various techniques to improve demand forecasting in the hotel industry, focusing on capturing daily occupancy and average daily rate (ADR) seasonalities. Methods such as TBATS, Multiple STL Decomposition (MSTL), STL Decomposition, and Linear Regression are compared regarding their ability to model complex time series data. Using a five-year dataset from an Upper Upscale branded property, the study utilizes in-sample data for model development. It employs a rolling window approach for testing based on two scenarios. Results highlight the robust performance of TBATS and MSTL across different forecasting horizons, consistently outperforming Seasonal-Trend Decomposition (STLF) and linear regression. TBATS and MSTL maintain accuracy even in the test set, while STLF struggles with longer-term patterns. Although Linear Regression shows moderate performance, it falls short compared to more advanced methods. Furthermore, the study presents comparative analyses for occupancy and ADR forecasting, emphasizing the superiority of TBATS and MSTL. Scenario-based evaluations underscore the stability and accuracy of TBATS and MSTL, reaffirming their suitability for demand forecasting in the hotel industry. These findings highlight the limitations of Linear Regression, which, despite its simplicity, fails to match the performance of more advanced methods. Therefore, revenue managers are encouraged to consider advanced techniques like TBATS and MSTL for accurate and robust demand forecasting, with TBATS excelling particularly in capturing short-term seasonality patterns. Understanding the strengths and limitations of each method is crucial in selecting an appropriate forecasting approach, considering factors such as dataset characteristics and forecast horizon requirements. TBATS and MSTL emerge as top performers, providing valuable insights for revenue management strategies in the hospitality sector. Therefore, this study contributes to understanding demand forecasting methodologies and offers practical insights that can directly impact revenue optimization and strategic decision-making. The robust performance of TBATS and MSTL underscores their importance in enabling revenue managers to make informed decisions.

Daily Sharing Accommodation Demand Forecasting Based On Multimodal Data Fusion

Presenter: Mingming Hu
Co-authors: Mingming Hu;Yuling Ye;Yushan Lan;Na Dong

With the development of the sharing economy, sharing accommodations have become another choice alongside hotels due to their advantages of cost-effectiveness and unique living experience. Sharing accommodations share the same perishable characteristic with hotel rooms. It is valuable to forecast daily demand for sharing accommodations and to provide the corresponding supply, enabling effective pricing strategies. Demand forecasting in hospitality has been primarily focused on hotels, but forecasting room

demand for shared accommodations has not been extensively studied. Simultaneously, big data is widely used in tourism demand forecasting, encompassing different modalities such as numeric, text, and images. How to fuse these data and use for forecasting tourism demand is also necessary to explore. This study aims to forecast daily sharing accommodation room demand based on multimodal data fusion. Airbnb home demand on Shanghai is used in empirical study. Variables extracted from online reviews of Qunar and Ctrip and search engine are divided into two categories: online attention and review sentiment. Search queries and the number of online reviews indicate online attention, while online review ratings, textual sentiment, and image aesthetic reflect review sentiment. The roles of online attention and review sentiment in tourism demand forecasting are evaluated by SARIMAX and BPNN model. Results indicate that (a) Both online attention and review sentiment are useful in improving room demand forecasting accuracy on the sharing accommodation market; (b) Among two variables of online attention, the number of online reviews outperforms search queries in sharing accommodation room demand forecasting; (c) Of the three variables on review sentiment, online review ratings outperforms other variables in short-term forecasting, whereas text and images demonstrate superior effectiveness in long-term. This study contributes the literature in two aspects: (1) It's first paper to forecast room demand in sharing accommodation market, especially for the daily room demand. (2) This paper developed two variables, online attention and review sentiment from multimodal data, which will enrich the theory of big data tourism demand forecasting. Hosts can formulate pricing strategies according to different room demands, and the platform can develop a marketing strategy based on online attention and review sentiment.

The Impact Of Announcements Of Airbnb Regulations

Presenter: Miriam Scaglione
Co-authors: Miriam Scaglione;Martin Falk

With the end of the pandemic city, tourism is coming back, and so is Airbnb. However, locals are increasingly concerned that Airbnb leads to a displacement of long-term tenants and rising rents. Recently, many more cities have announced new regulations for Airbnb. However, it often takes some time for the regulation to come into force. In a Referendum in Bern (Switzerland) in February 2022, the citizens voted for regulations against Airbnb. Depending on the zone, the top floor and the floors from the second floor up will be reserved for permanent tenants in the future. The regulations are expected to come into force in 2024. This delay can also be observed for another city (Lucerne). In the canton (State), Vaud, whose capital is the city of Lausanne on the Geneva Lake, was introduced in June 2022. Short-term rentals of 90 days or more must register their activities with the authorities and apply for a permit. This paper analyses the impact of the announcement of new regulations on the supply and demand for Airbnb in Swiss cities. The general hypothesis is that the announcement of new regulations already has a negative impact on Airbnb hosts. We compare the announcement effect with the actual legal regulations. The data is based on the total number of Airbnb listings in Swiss cities, with about 30,000 properties per month in the 20 major cities for period 2022 and 2023. The method is a probit model estimating the presence of a listing and Poisson estimations for the performance indicators. A difference-in-differences approach with spatial effects is used to estimate whether the Airbnb listings are declining after the announcement or regulations. The explanatory variables are location (city), located in the city (centre), characteristics of the property (number of beds), and type of listing (flat or house). The results show that the announcement effect is negative and not so different from the actual regulations. The explanation for this is that enforcement of regulations could be stronger. This paper contributes to the growing literature on the impact of Airbnb regulations on short-term rentals.

Let's Pay By Card! Determinants Of Visitor Expenditure During The Venice International Film Festival Within A Spatial Panel-Data Framework

Presenter: Bozana Zekan
Co-authors: Bozana Zekan;Dario Bertocchi;Ulrich Gunter

The focus of this study is in understanding the behavior of visitors and the economic impact of

their visit during the Venice International Film Festival. In doing so, daily credit card expenditure data (i.e., total spending, total number of transactions, and the average value of a receipt) from Mastercard for the period of the 2022 edition of the festival will be analyzed. Notably, this anonymized, clustered, and indexed dataset contains expenditure metrics per card origin (domestic or international), card type (business or consumer), and type of expenditure (accommodation, culture, shopping, etc.). In addition to quantifying the marginal contributions of these attributes on visitor expenditure within an econometric framework, one major aim of this study is to investigate the temporal and spatial dependencies of all variables within a unified approach. Hence, a spatial panel-data econometric model represents a suitable statistical framework. Moreover, the temporal dimension of the data will allow the authors to use the estimated model to produce year-on-year forecasts for future editions of the festival.

Predicting And Optimizing The Fair Allocation Of Donations

Presenter: Lauren Davis
Co-authors: Lauren Davis;Nowshin Sharmile;Steven Jiang;Funda Samanlioglu;Carter Crain

Non-profit hunger relief organizations primarily depend on the benevolence of donors to help them alleviate hunger in their communities. However, the quantity and frequency of donations they receive may vary over time, thus making fair distribution of donated supplies a challenging task. This paper presents a hierarchical forecasting methodology to determine the quantity of food donations received per month in a multi-warehouse food aid network. We further link the forecasts to an optimization model to identify the fair allocation of donations, taking into consideration the network distribution capacity. The results indicate which locations within the network are under-served and how donated supplies can be allocated to minimize the deviation between overserved and underserved counties.

Ai-Driven Prediction And Explanation Of Future Student Performance

Presenter: Matthias Deceuninck
Co-authors: Matthias Deceuninck;Yves R. Sagaert;Benedikt Sonnleitner;Tom Madou;Filotas Theodosiou

We explore the potential of forecasting student performance midway through the semester to enable lecturers and students to take targeted measures for improving final exam outcomes. Using data extracted from an online learning platform, we evaluate 74 different features on their effectiveness to forecast student performance. Our analysis reveals that only a subset of these features demonstrate significant predictive capabilities. In addition we benchmark a lasso regression approach against different Machine Learning pipelines, including Bayesian Optimization, feature selection, and different non-linear regression models. For all, we evaluate a classification setting, predicting the performance category of each student, and a regression setting, where we directly predict final exam grades. Along with introducing an interesting applied forecasting task, we provide some insight on features that can drive efficient learning behavior

Rating Of Players By Laplace Approximation And Dynamic Modeling

Presenter: Ruby Weng
Co-authors: Ruby Weng;Hsuan-Fu Hua;Ching-Ju Chang;Tse-Ching Lin

The Elo rating system is a simple and widely used method for calculating players' skills from paired comparisons data. Many have extended it in various ways. Yet the question of updating players' variances remains to be further explored. In this paper, we address the issue of variance update by using the Laplace approximation for posterior distribution, together with a random walk model for the dynamics of players' strengths, and a lower bound on players' variances. The random walk model is motivated by the Glicko system, but here we assume nonidentically distributed increments to take care of player heterogeneity. Experiments on men's professional matches showed that the prediction accuracy slightly improves when the variance update is performed. They also showed that new players' strengths may be better captured with the variance update.

Navigating Through Shifting Seasons And Customer Behaviors: Innovations In Demand Forecasting

Presenter: Bor-Chau Juang
Co-authors: Bor-Chau Juang;Eyal Shafran;Yaxian Li;Shashank Shashikant Rao

This study explores multiple challenges associated with forecasting daily customer call data, a crucial factor in determining downstream staffing requirements for call center support agents. The ultimate aim is to strike an optimal balance between customer satisfaction and staffing cost; thus, forecasting precision is critical. The presented research provides an insightful, practical industry perspective on designing demand forecasting models that can accommodate diverse demand patterns, seasonal changes, and external disturbances.Our call demand data is riddled with challenges such as multiple seasonalities, varying time-series lengths, abnormal data occurrences, and alterations in customer behavior prompted by external factors, such as the pandemic. Additional intricacies emerge from floating holidays, unforeseen changes in product call volume, and sudden peaks linked to special events (e.g., significant holidays, product incentives). The initial modeling phase was further complicated due to its development in the immediate aftermath of the pandemic, which saw an unprecedented fluctuation in seasonal behaviors.In response to these challenges, we introduce a novel long-horizon forecasting framework. Its strength lies in the ability to function across multiple, interrelated time-series, whilst directly addressing the aforementioned data challenges. The approach integrates pre-processing, time-series decomposition, model ensembling, and post-processing stages.Our unique decomposition method partitions the data into seasonality and volume components. To maximize learning from extended demand patterns, we implement multiple temporal aggregations. A separate model incorporating cross-learning techniques is considered to capture the more refined temporal patterns. We demonstrate that segregating the data using similarity scores prior to cross-learning results in improved performance. We suggest the construction of relative features capable of generalizing unusual patterns that alter the data's seasonality. This adaptability enables the framework to accommodate historical data characterized by suddenly fluctuating seasonalities.Our findings validate that our forecasting framework consistently produces robust and precise predictions. Moreover, experimentation shows our framework outperforms other modeling methods. An ablation study of the framework illuminates the influence and contribution of each individual step.

Combination Of Experts For Sales Forecasting Of New Products

Presenter: Raphael Nedellec
Co-authors: Raphaël Nedellec;Anthony Vromant

Forecasting sales of new products has always been a challenge in the retail industry for the supply chain. It is especially true for a company such as Decathlon where thousands of new products are manufactured and launched every year. To ensure the delivery of the right amount of goods at the right time in the right place, users need to forecast both the volume and the seasonality of new products. Typically, the users need to forecast at a weekly frequency thousands of products up to two years' horizons. Given the scale of our problems, it is really challenging to automatically identify the best model or algorithm to perform the forecasts for each of our different products' families.We introduce an approach built on an online combination of experts to tackle this challenge. By aggregating dynamically different forecasts built on several scenarios or different models, we are able to quickly adjust our forecast to the real situation. This strategy greatly improves the forecasts accuracy and our ability to adjust our forecasts at scale. After a quick review of the literature, we will present the application developed at Decathlon to solve this problem.

Scenario-Based Forecasts In The Presence Of Product Promotions: A Laboratory Experiment On Judgmental Adjustments

Presenter: Niles Perera
Co-authors: Niles Perera;Binura Jayakody;Dilek Onkal;Dilina Kosgoda

Forecast accuracy is a critical factor in organizational planning, and the use of information systems in forecasting is common in the modern business environment. Demand planning professionals frequently make judgmental adjustments to system-generated forecasts due to contextual factors. Adjustments are commonplace in the fast-moving consumer goods (FMCG) sector especially to calibrate market trends and consumer behavior, despite the uncertainty of changing consumer preferences and market dynamics. Scenario-based thinking is a valuable approach for improving judgmental forecasting in the FMCG industry by developing plausible future scenarios considering both positive and negative circumstances. In this study, scenarios and judgmental forecasting are deployed together to understand behavior under different conditions and the potential impact of scenarios on demand forecasting in the FMCG industry.The Extreme World Method was chosen for scenario development which involves creating a best-case and worst-case scenario. The experiment is designed as a within-subject experiment, allowing participants to be exposed to various scenarios and make their judgmental adjustments for the system-generated forecasts. The experiment includes three conditions: the control condition (no scenarios are provided), the best-case scenario, and the worst-case scenario. Product promotions was used as the variable when developing scenarios, and the two scenarios were developed while highlighting the sales increments under product promotion and the demand drop due to a negative impact of the promotional campaign.150 bachelors and masters candidates specializing in supply chain management were recruited as participants for this experiment. As a preliminary result, the analysis showed that the availability of different scenarios affects people's judgmental forecasting patterns. In the best-case scenario, people generally anchor forecasts in a higher range compared to the other two conditions. In the worst-case scenario, they tend to finalize their forecasts closer to the system-generated forecast. An additional finding from the study was that participants gut feelings have a considerable impact on their judgmental forecasts.

Predicting Market Diffusion Of New Drugs Using The Dynamic Patient-Share Method

Presenter: Christian Schäfer
Co-authors: Christian Schäfer;Stephan Brebeck

When a life sciences company launches a new compound, it needs to assess the uptake as early and as accurately as possible, not only for demand planning but also for the allocation of promotional resources. We present the Dynamic Patient Share (DPS) methodology, which enables accurate and rather simple prediction of a drug's uptake curve. It eliminates the need for traditional but more time-consuming methods based on patient flow analysis and analog research. The DPS method is based on classic patient flow analysis, but uses basic time series theory to improve prediction accuracy. In addition, it facilitates the estimation of the impact of promotional activities on market share acquisition. We show how the DPS method is derived from classical patient flow analysis and evaluate its predictive accuracy based on actual sales performance.The DPS method can predict both, peak market shares and drug uptake based on the dynamics of a specific market, while considering the complete competitive picture. The dynamics that lead to this number of patients is often called "patient flow" or "source of business". The theoretical basis of the DPS method reflects the actual market dynamics when a company invests in advertising. A marketing campaign aims to influence the dynamics of a drug's dynamic patient share, the naive and switching patients. It does not directly affect the total patient share (as other methods suggest). In this way, the model is transparent to management, a clear advantage over other approaches to critical decision making. In any market, a new product gains market share by attracting new customers and trying to limit the loss of existing customers to competitors. Because this concept is broadly applicable, the DPS method should have the potential to monitor the success of a new product launch in near real time and to predict a product's future brand share at an early stage, allowing brand managers to take early marketing action if necessary.

Optimal Design Of Acceptance Sampling For Sequential Tests At Consecutive Times

Presenter: Hugalf Bernburg
Co-authors: Hugalf Bernburg;Stefan Ankirchner;Katy Klauenberg

We consider a population of N devices which were installed at time t_0. To test if the population meets reliability targets at time $t_i, i = 1..s$, a sequential acceptance sampling procedure is conducted. The test determines whether to accept or reject the population. In the rejection case, the whole population must be replaced. In the other case, the population remains, and a further test is possible at the next testing time t_{i+1}. We consider two adversarial parties: the consumer and the producer. The consumer defines acceptance numbers for the test, which are based on their Bayesian analysis with usually weak or skeptical prior knowledge. The producer knows these acceptance numbers and applies their own model, prior knowledge and costs to determine rejection numbers minimizing the expected total costs. Both parties may account for prior knowledge from all previous sequential test results.To the best of our knowledge, we are the first to determine optimal sequential acceptance sampling plans for consecutive points in time. To do so, we apply Bellman's optimality principle on costs depending on the age of the devices, the sample sizes, and the test results. Such determination of optimal sampling plans is generally applicable for repeated acceptance sampling. We demonstrate the optimal design of these sampling plans using utility meters subject to German regulations. For an example population, we conduct a sensitivity analysis on the expected costs and the corresponding optimal strategy w.r.t. model, prior and other assumptions.

Modeling Temporal Networks Of Events With Dependent Excitations And Spectral Clustering

Presenter: Subhadeep Paul
Co-authors: Subhadeep Paul;Kevin Xu;Lingfei Zhao;Hadeel Soliman

Temporal networks observed through timestamped relational events data are commonly encountered in applications, including online social media, human mobility, financial transactions, and international relations. Such datasets consist of directed interaction events among entities at specific time points. Temporal networks often exhibit community structure and strong dependence patterns among node pairs. We introduce generative models combining high-dimensional, mutually-exciting Hawkes processes with the stochastic block model to model community structure and node pair dependence. We obtain an upper bound on the misclustering error of spectral clustering of the event count matrix as a function of the number of nodes and communities, time duration, and a quantity measuring the amount of dependence in the model. The theoretical results provide insights into the effects of dependencies in the mutually-exciting Hawkes processes on the accuracy of spectral clustering. We assess the prediction and forecasting ability of the models using predictive loglikelihood and AUC for dynamic link prediction tasks. We empirically demonstrate our methods are computationally scalable, capable of modeling networks with millions of events, and provide excellent predictive ability on several real datasets.

Ensemble Learning With Latent Predictive Synthesis

Presenter: Joseph Rilling
Co-authors: Joseph Rilling;Kenichiro Mcalinn;Kōsaku Takanashi;Junpei Komiyama

We discuss the ensemble of multiple forecasts under bias, misspecification, and dependence. While linear density combination strategies are ubiquitous in the literature, we show that this strategy is fundamentally flawed, being underparameterized when all forecasts are misspecified. To develop a method to overcome this, we propose a novel theoretical strategy based on stochastic processes that identifies a more general class of ensemble methods, which we call latent predictive synthesis. Examining the predictive properties of this new class, we identify the conditions and mechanism for which latent predictive synthesis improves over linear combinations of densities, in terms of expected squared forecasts error. We present an extensive simulation study, as well as two real applications, to demonstrate that this class improves over existing ensembling strategies. A python package is included that can apply the discussed method.

Path Prediction Of Anticipative Alpha-Stable Moving Averages Using Semi-Norm Representations

Presenter: Arthur Thomas
Co-authors: Arthur Thomas;Sébastien Fries;Gilles De Truchis

This paper studies the conditional distribution of future paths of a two-sided alpha-stable moving average, given a piece of the observed trajectory and when the process is far from its central values. In this framework, vectors of the observed trajectory (t-m to t) and the forecast path up to a horizon h are multivariate alpha-stable, and the dependence between the past and future components is encoded in their spectral measures. A new representation of stable random vectors on unit cylinders and an appropriate semi-norm are proposed to describe the tail behaviour of the vectors when only the first m+1 components are assumed to be observed and large in norm. Not all stable vectors admit such a representation, and the process must be "anticipative enough" to admit one. The conditional distribution of future paths can then be explicitly derived using the regularly varying tails property of stable vectors and has a natural interpretation in terms of pattern identification. Through Monte Carlo simulations we develop procedures to forecast crash probabilities and crash dates and demonstrate their finite sample performances. As an empirical illustration, we estimate probabilities and reversal dates of El Niño and La Niña occurrences.

Impact Of Food Inflation On Core Inflation In Brazil: A Time-Varying Parameter Approach

Presenter: Diego Ferreira
Co-authors: Diego Ferreira;Andreza Palma;Ana Kreter;José Ronaldo C. Souza Júnior

This paper examines the influence of food inflation on core inflation in Brazil by employing a VAR model with time-varying parameters and stochastic volatility. The estimation is conducted through Bayesian simulation spanning the period from January 1999 to February 2023. In light of the recent upswing in commodity prices and the inflationary repercussions of the pandemic, a thorough evaluation of the impact of food prices on core inflation becomes imperative. Food inflation can have both direct and indirect effects on overall inflation (Cecchetti and Moessner (2008)). The direct effect occurs because food is part of the consumption basket used in price indices calculation. The indirect effect can occur through its impact on inflation expectations and also through inertial effects. Therefore, even though core inflation excludes food items, they can still have a relevant impact through indirect effects. These impacts may vary over time, depending, for example, on the state of the economy (recession, expansion) or the level of inflation (high, low). As pointed out by Ha, Ivanova, Montiel, and Pedroni (2019), core inflation in low-income countries responds more strongly to global food inflation than does core inflation in the other country groups. Our findings suggest a substantial element of uncertainty during the pandemic. Furthermore, the impact of food inflation on core inflation is significant and statistically meaningful, particularly in the recent period. This impact is slightly more pronounced in the current period compared to the pre-pandemic phase. Given the elevated levels of food inflation in the recent economic scenario and the recent trajectory of Brazilian inflation, these results bear strong relevance for policymakers. Moreover, this result can be explored for the definition of new measures of core inflation or in the discussion of whether the central bank should react to food prices or not. Food inflation poses a big challenge, especially for Central Banks in emerging economies, so these findings carry important implications for policymakers.

Quantifying Uncertainty Under Local Instability: A Dynamic Conformal Approch To Electricity Price Forecasting

Presenter: Alessandro Giovannelli
Co-authors: Alessandro Giovannelli;Tommaso Proietti;Andrea Cerasa;Fany Nan

This paper introduces a novel methodology, Dynamic Conformal Prediction (DPC), for the construction of prediction intervals with improved coverage in the presence of parameter instability. DCP allows the

adaptation of both estimation and calibration windows. Through a Monte Carlo analysis, we demonstrate the effectiveness and validity of our procedure. In particular, using several data generating processes that differently account for parameter instability, we highlight the ability of our procedure to generate forecast intervals with a coverage that is close to the nominal one. Then, DCP will be employed for constructing prediction intervals for the time series of the single national price of the Italian Electricity Spot market in the short run, i.e., for forecast horizons that are not larger than 14 days ahead. The forecasting methods used encompass different assumptions (reduced forms, structural decompositions, nonlinearity) and degrees of mean reversion.

Earth, Wind, Fire And Grocery Bills: Exploring The Forecasting Power Of Weather Variables For Euro Area Food Inflation

Presenter: Chiara Osbat
Co-authors: Chiara Osbat;Friderike Kuik;Ignacio Vidal-Quadras Costa

Food is a very important part of consumer baskets and its price tends to be more volatile than average inflation. Weather conditions are an important input in its production and With climate change set to increase the frequency of extreme weather events as well as their intensity, their impact on food price inflation will gain more and more centrality in the analysis of inflation.We perform a forecast horse-race to evaluate the predictive advantage in using weather data in addition to commodity prices, both in a linear context using ridge regression and in a nonlinear one, using random forests. We find that random forests can beat linear forecasts of food inflation, especially at shorter horizons. Including weather variables improves the forecasting performance: directly for unprocessed food, which is more subject to the gyrations of weather, and indirectly in a bottom-up forecasting exercise where the forecast of food inflation is performed by aggregating those for processed and unprocessed food. We also characterise the uncertainty around our nonlinear forecasts, and document that while random forests can beat linear benchmarks in recent years, they do not manage to bring the error down to its historical average. Finally we study the relative importance of each variable, using Shapley values. Focusing on unprocessed food and shorter horizons, where weather variables matter, we find that EU weather matters most, and the magnitude and sign of its impact changes each month. In future versions of the paper we will include a more complete characterisation of variable importance across horizons, both on average and for predicting around specific episodes. With a view to improving the prediction at times of extreme inflation realisations, we will also include results from a modification of the algorithm that attempts to mitigate the extrapolation problem of random forests.

Cyclical Fluctuations In The U.s. Real Gdp And National Income And Product Accounts Aggregates

Presenter: Baoline Chen
Co-authors: Baoline Chen;Kyle Hood;Tucker Mcelroy;Thomas Trimbur

This paper provides a new set of empirical regularities describing the U.S. macroeconomy, focusing on cyclical relationships between real GDP and eleven major aggregates from the U.S. National Income and Product Accounts (NIPAs). Patterns of cyclical movements are assessed using adaptive model-based filtering techniques that adapt to the properties of the series being studied. We employ a comprehensive dataset that includes the Great Recession caused by the 2008 financial crisis and the ensuing recovery, and cyclical components of the series exhibit highly diverse properties. Using trend and cycle estimates from two model-based filters, we aim to 1) examine lead-lag relationships via cross-relations between aggregate cycle in real GDP and the cyclical movements in each NIPA aggregate; 2) investigate econometrically the inter-linkage or predictability between the cyclical movement in real GDP and that in each of the NIPA aggregate by way of the Granger-causality test; and 3) evaluate the capability or predictive power of each NIPA aggregate to forecast real GDP growth via a two-step structural time series model (STM) which separately forecasts the trend and cycle growth before combining them to obtain total growth in real GDP, a single-equation STM which tracks marginal contributions of trend and cycle components of a NIPA aggregate to real GDP growth, and an AR(p) model which directly forecasts GDP growth using

unfiltered time series. Trend and cycle estimates from the HP and BK filters are also used in the analysis for comparison. The results show that the lead-lag and the Granger-causal relationships estimated using smoothed cycle estimates from the model-based filters are highly consistent with empirical observations. The two-step STM using trend and cycles estimates to forecast real GDP growth outperformed all other models for all forecasting horizons (i.e., with minimum RMSFEs), and the model-based filters are shown to be the most desirable filters for forecasting. The basic conclusion from the empirical analysis is that adaptive model-based filters have demonstrated advantages over the commonly used HP and BK filters for business cycle analysis across very diverse economic time series from the U.S. national accounts.

Time-Series Evidence On The Influence Of The Choice Of Seasonal Adjustment Method On Forecasting Accuracy

Presenter: Robert Kunst
Co-authors: Robert Kunst;Adrian Wende;Martin Ertl

With subannual (often quarterly or monthly) time series, seasonally adjusted data are routinely used in applied research, particularly in empirical economics. For the most part, two methods of seasonal adjustment are used today: The moving-average X-11 method and the SEATS method that is based on tentatively fitted ARIMA models. We study which of the two methods (and their variants) yields more accurate forecasts of annual targets after temporal re-aggregation, and when it is better not to seasonally adjust the data at all.We investigate these issues both empirically and with Monte Carlo simulations. For empirical applications, we study UK and Austrian macroeconomic time series from national accounts and related sources. The UK and Austria are among the few countries, where both adjusted and unadjusted versions of quarterly national accounts are readily available. For the simulations, we consider data-driven time-series models, both univariate and multivariate generating processes, ARIMA and VAR models among others.

Online Monitoring Of Policy Optimality

Presenter: Bjarni Einarsson
Co-authors: Bjarni Einarsson

Early detection of optimization failures in the stance of monetary policy is of crucial importance for policy makers. Since economic data relevant to the policy decisions is released every week, we present a framework for online monitoring of the stance of policy that incorporates the information contained in the constant inflow of new data. This framework depends on the causal effects of policy instruments on the target variables along with conditional expectations of the target variables given a choice for the policy instruments. The real-time aspect of the proposed framework has to do with the updating of the conditional paths of the target variables at the end of each week given all information available.To implement the framework, we need estimates of the causal effects of the policy tools on the target variables, inflation and unemployment. This involves the estimation of structural impulse response functions for which we will use penalized local projections with external instruments.Additionally, we need the conditional expectations for the paths of inflation and unemployment that incorporate the real-time inflow of data. Given the mixed frequency nature of the inflow of data, a nowcasting approach lends itself naturally to generating forecasts of the variables of interest. We use a high-dimensional mixed frequency BVAR to generate these nowcasts. The use of a VAR allows for rich interdependence among the model variables which is particularly important for the present paper since we seek a model that can jointly nowcast both the inflation rate and the unemployment rate.However, as the horizon over which monetary policy decisions are made extends well beyond what is usually considered in the nowcasting literature, we will augment the long end of the forecasts with long-run forecasts from the Survey of Professional Forecasters using relative entropy.In a retrospective analysis of the Fed's monetary policy decisions in the lead up to the Great Recession we find that we can reject the optimality of the policy stance as early as the beginning of February 2008. This early detection stems from the timely nowcasting of the deteriorating unemployment outlook.

Evidence And Insights From The State Of Britain's Foresight-Based Scientific Advice In Policymaking: Are We Fully Prepared For The Next Emergency Crisis?

Presenter: Yuna Lee
Co-authors: Yuna Lee

Britain's foresight program was created in 2001 when Sir David King, the Government's Chief Scientific Advisor (GCSA) at that time, called on the government to prepare for the foot-and-mouth epidemic. Later in 2004, Sir David King and Sir Keith Peters launched the new Council for Science and Technology (CST). This study delves into the 'Foresight-led scientific evidence for long-term policy decisions' in the UK with the question "How successful has science-oriented foresight been in bridging the gap with the dynamic and complex policy ecosystem, and what lessons we have learned from the government's response to the COVID-19 pandemic?". Many studies have proved the importance of scientific advice in policymaking based on national foresight program cases with different frameworks. However, in addition to the lack of studies of UK cases post-2010, the reason for self-reflection of scientists and politicians that they failed to respond on time to an actual pandemic, despite the UK government having groups of experts who provide advice and evidence from the best scientific and technological perspectives for mid-to-long-term thinking policies, as well as for emergencies, is also understudied. Adopting a mixed method, the data is collected through a review of the literature, policy documents and semi-structured interviews. The main participants of the interview were scientific advisors and civil servants involved in the Foresight and Horizon Scanning Centre (HSC) programmes and the Scientific Advisory Group for Emergencies (SAGE) between 2001 and 2023. Here, this study shed light on efforts, the practical impacts and enhanced integration by government departments to mitigate potential conflicts between the scientific evidence and policy and to provide trustworthy knowledge to the public. Further, apart from bias awareness and undesirable impacts, new limitations due to institutional and cultural differences with the US and other European foresight programs are identified in the interactions among policymakers, scientific advisers, and the public. By comprehensively and multifacetedly examining the behind-the-scenes and longitudinal pathways of Britain's foresight programme, and its impact on British society and international relations, this paper offers significant ramifications about the impact of policy-oriented scientific strategic foresight that can respond to complex and existential threats closely related to human health and security.

Populism And Covid-19 Syndemic. Checking A Forecast

Presenter: Eduardo Loría
Co-authors: Eduardo Loría

In 2022, I published the article Mexico: The Populism/Covid-19 Syndemic [International Review of Applied Economics, 36(5-6)] in which I econometrically proved that countries led by populist governments had the worst results in Case Fatality Ratio and Mortality Rate, caused by COVID-19.I found that structural variables (healthcare infrastructure, comorbidities, poverty and HDI) were not the main ones to explain these disastrous results, but rather crisis response variables, such as fiscal support, health policy, and, above all, populist government discourses and actions.With Narrative Economics Theory (Shiller, 2019) and Behavioral Economics (Banerjee, 1992, Chetty, 2015; Kahneman, 2003, Akerlof and Kranton, 2010, Thaler and Ganser, 2015, Akerlof and Shiller, 2015), and with the estimation of eight cross-section models for 31 countries, I concluded that this would cause them to have the worst economic and development recoveries in the following years. In 2023, those results were empirically validated through the HDI, Life Expectancy at Birth and Per capita GDP indicators.This is because the leaders of populist countries downplayed the severity of the pandemic, followed anti-scientific policies, and politicized public health measures, prioritizing their short-term political-electoral interests.

Probabilistic Forecasting Of Weather-Driven Faults On Electricity Distribution Networks

Presenter: Daniela Castro Camilo
Co-authors: Jethro Browell;Daniela Castro-Camilo

Electricity networks are exposed to the weather, and severe weather can cause faults that result in power cuts. Predicting the occurrence of faults in each region of these networks on time scales from hours to days ahead can increase preparedness and accelerate the response to weather-related faults, and ultimately reduce the duration of power cuts. Furthermore, these predictions should quantify uncertainty so that planners can assess risk and distribute limited resources accordingly. Here, we present a method for probabilistic fault prediction that leverages ensemble numerical weather prediction and flexible non-parametric regression, and a case study based two distribution networks in the UK over a 12-year period.Data describing network topology and vulnerability, such as elevation and proximity to vegetation, are combined with meteorological data to model the occurrence of faults, which is stochastic and may be heavy-tailed, resulting in sparse data in regions of interest as the most severe weather events occur infrequently. In addition, forecasts of future weather conditions are required; associated uncertainty is quantified via ensemble numerical weather prediction, which requires statistical post-processing. The fault forecast is therefore constructed by aggregating multiple density forecast across weather scenarios (ensemble members). Finally, we will address communication of the resulting forecast information, and forecast evaluation, which must be effectively communicated to a wide range of decision-makers and stakeholders.

Ai-Based Hybrid Modeling Approach For Predicting Solar Energy Production, Unifying Cnn, Lstm, Transformer, Xgboost, And Chatgpt

Presenter: Hae Rim Kim
Co-authors: Hae Rim Kim;Arnie De Castro;Mirim Yu;Jeongmin O;Ye Rin Kim;Tae Yoon Lee

Keywords: Solar Energy, 5G, COVID-19, CNN-LSTM-Transformer, ChatGPTAfter the COVID-19 pandemic, the acceleration of 5G and digital transformation led remote and non-face-to-face activities by many organisations. It is presumed that this has changed the pattern of energy consumption. With the recent surge in generative AI, data centers required for AI training and services have been constructed worldwide. Currently, about 8,000 data centers are in operation. Particularly, AI data centers consume more than twice the power of conventional data centers, thus significantly burdening the national power grid. According to the Ministry of Trade, Industry, and Energy, it is expected that the power demand in Korea will surge from 1,762 MW in 2022 to 49,397 MW by 2029. This suggests that Korea could face negative power grid issues within a few years. To prepare for such demand, the production of renewable energy sources, like solar energy, instead of traditional fossil fuels is considered important. The advancement of artificial intelligence technology has led to the utilization of various modeling techniques for predicting the production of renewable energy. However, it is difficult to verify if the performance of these models is maintained due to COVID-19, 5G, and new digital developments. This study aims to enhance the performance of models for predicting solar energy production. Since solar energy depends on the climate, weather data has been added. As changes in sunlight throughout the seasons affect solar energy production, date-related data has also been included. Through comprehensive analysis and utilization of these data, the variables required have been identified and implemented in the solar energy production prediction model. It is hoped that optimisation of the model developed may improve the accuracy of identifying solar energy production. For prediction, a hybrid model combining CNN, LSTM, Transformer, XGBoost regression, along with ensemble, stacking, and Boosting techniques is planned to be constructed. In the process of implementing the code, an attempt to use OpenAI's ChatGPT, which is currently activated, to further increase the predictive power is made. In this way, an improvement in the predictive power through the ensemble of machine learning, deep learning models, and ChatGPT models is expected.

Using Peak Load Forecasts For Demand Response Programs

Presenter: Tao Hong
Co-authors: Tao Hong;Shreyashi Shukla

Many power companies use demand response (DR) programs to save money when the demand charges are high. Although the many research papers have been devoted to load forecasting, and specifically peak load forecasting, few really connect these load forecasts to the practical use for demand response. In this presentation, we will discuss how to predict and reduce the monthly peak demand through demand response programs. The research problem to investigate is when to call the next DR event. The objective is to capture the timing of monthly peak demand while minimizing the "false alarm" days.

Enhancing Reliability In Prediction Intervals Using Point Forecasters: Cornish-Fisher Quantile Regression Averaging And Width-Adaptive Conformal Inference

Presenter: Carlos Sebastián
Co-authors: Carlos Sebastián;Carlos E. González-Guillén;Jesús Juan

Building prediction intervals for time series forecasting problems presents a complex challenge, particularly when relying solely on point predictors, a common scenario for practitioners in the industry. While research has primarily focused on achieving increasingly efficient valid intervals, we argue that traditional measures alone are insufficient. There are additional crucial characteristics: the intervals must vary in length, with this variation directly linked to the difficulty of the prediction, and the coverage of the interval must remain independent of the difficulty of the prediction for practical utility. We propose the Cornish-Fisher Quantile Regression Averaging (CFQRA) model and the Width-Adaptive Conformal Inference (WACI) method, providing theoretical coverage guarantees, to overcome those issues, respectively. The methodologies are evaluated in the context of Electricity Price Forecasting and Wind Power Forecasting, representing complex scenarios in time series forecasting. The results demonstrate that CFQRA and WACI not only improve or achieve typical measures of validity and efficiency but also successfully mitigate the commonly ignored mentioned problems.

Enhancing Natural Gas Demand Forecasting By Reconciling Incoherent Data Hierarchies

Presenter: Colin Quinn
Co-authors: Colin Quinn;Richard Povinelli

Time series reconciliation involves aligning a set of incoherent, independently generated base forecasts according to a predefined set of linear constraints. These linear constraints are typically based on inherent characteristics of the hierarchy being reconciled. If significant incoherence is present in the in-sample training data, out-of-sample forecasts may be reconciled under inaccurate constraints. This study investigates the effectiveness of applying a weighted reconciliation preprocessing technique to natural gas consumption data with significant in-sample incoherence prior to calculating the out-of-sample base forecast.Natural gas consumption in a single service area is often measured by multiple time series spanning various temporal resolutions and geographical regions. While these time series lack coherence when organized hierarchically, we demonstrate that relevant information can still be extracted from gas consumption data to enhance out-of-sample forecast reconciliation accuracy. Incoherence in consumption data occurs when one consumption signal does not align with others; for instance, an hourly consumption series not summing to its daily series counterpart. Adjustments to the incoherent consumption data are made by comparing aggregated versions of lower hierarchical levels to their higher counterparts, or vice versa.Historical heating season hierarchies are used to calculate a distribution for identifying appropriate reconciliation constraints for the out-of-sample forecast. A new hierarchical node is introduced to contain the preprocessing "coherency error", addressing the industry's "Lost and Unaccounted (LAUF)" gas found in the training data. Error from both the aggregated-consumption comparison and LAUF is propagated

throughout the in-sample consumption observations using Minimum Trace Reconciliation.Our in-sample reconciliation preprocessing technique is tested across three temperature-sensitive gas operating areas and results in a 9.6% reduction in MAPE in the included case studies. Through this analysis, we underscore the advantages of adopting hierarchical time series reconciliation techniques in gas demand forecasting and provide insights into adaptability of such techniques to incoherent data sets.

Modelling And Forecasting Supply Networks Using Functional Time Series And Mathematical Programming

Presenter: Nazgul Zakiyeva
Co-authors: Nazgul Zakiyeva;Milena Petkovic

We study a network functional autoregressive model for balanced energy network time series, where the demand and supply are balanced within one gas day. We approach the estimation of the proposed model using a Mixed Integer Optimization method. The inclusion of the high dimensional curves under balance constraint differentiates the proposed model from the classic functional autoregression models, as the proposed model captures both serial dependence in the functional time series and cross-sectional dependence in network by keeping the demand and supply curves in balance. We illustrate our methodology on natural gas network data set and show that our model provides more accurate 2 days-ahead hourly out-of-sample forecasts of the gas in- and out-flow compared to benchmark models.

Probabilistic Forecasts For Anomaly Detection

Presenter: Rob Hyndman
Co-authors: Rob Hyndman

When a forecast is very inaccurate, it is sometimes because a poor forecasting model is used, but it can also occur when an unusual observation occurs. I will discuss the latter situation, where a good forecasting model can be used to identify anomalies.The approach taken is to use a probabilistic forecast, and to compute the "density scores" equal to the negative log likelihood of the observations based on the forecast distributions. The density scores provide a measure of how anomalous each observation is, given the forecast density. A large density score indicates that the observation is unlikely, and so is a potential anomaly. On the other hand, typical values will have low density scores. A Generalized Pareto Distribution is fitted to the largest density scores to estimate the probability of each observation being an anomaly. Applications to tourism numbers and mortality data will be used to illustrate the ideas using the weird R package.

Generalized Linear Pools For Combining Probabilistic Forecasts

Presenter: Xiaochun Meng
Co-authors: Xiaochun Meng;James Taylor;James Curtis

For many applications, combining the individual probabilistic forecasts can improve their accuracy. The existing literature has extensively explored linear pools of forecasts of cumulative distribution functions or quantile functions. A general framework of combining methods is proposed, which encompasses the existing linear pools. We analyse the statistical properties of the proposed generalized linear pools. The framework and theoretical findings enable the provision of recommendations regarding the choice of combining methods and scores to use in practice. An empirical illustration is provided on simulated and real data.

Distribution-Free Uncertainty Quantification For Multivariate Functional Time Series, With An Application To The Italian Gas Market

Presenter: Matteo Fontana
Co-authors: Matteo Fontana;Jacopo Diquigiovanni;Simone Vantini

Probabilistic forecasting, especially when dealing with complex data characterized by non-trivial dependence structures, represents a topic of great methodological importance in statistics and data science, with potential wide-ranging applications in many business and public policy fields.We are pushed by novel advancements in the literature concerning distribution-free uncertainty quantification. Namely, we focus our attention on Conformal Prediction and its extensions to tackle functional data and time-series data. Our contribution consists in proposing a blending of these two advancements, thus introducing a scalable procedure that outputs closed-form simultaneous prediction bands for multivariate functional response variables in a time series setting, based on a Conformal Prediction framework. Such procedure is able to guarantee performance bounds in terms of unconditional coverage and asymptotic exactness, both under some regularity conditions. After evaluating its performance on synthetic data, the method is applied to build multivariate prediction bands for daily demand and offer curves of the Italian gas market. This application allows traders to directly evaluate the impact of their own actions on the market, providing an effective tool for business purposes.

Closer I Am To Fine: Using Pseudo-Earth Mover Divergence To Improve Policy-Relevant Probabilistic Forecasts Of Fatalities From Political Violence

Presenter: Michael Colaresi
Co-authors: Michael Colaresi

There is a growing trend towards forecasting political violence at increasingly fine-grained spatial and temporal resolutions in both academia and in government (Hegre, et al 2024). Useful predictions can guide the deployment of peacekeepers, details of evacuation plans, and the logistics of aid deliveries. Many of these benefits depend not on being (impossibly) exactly accurate about the placement and timing of the outbreak of conflict, but instead on distributing plausibility of higher (or lower) fatality levels as close as possible with the current state of technology to the actual outbreaks. Previous work (Colaresi, et al 2023, Hegre, et al 2022) has illustrated the benefits of using a new cross-bin performance metric to score point forecasts/expected values — pseudo-Earth Mover Divergence (pEMDiv). pEMDiv expands on the well-known Earth Mover Distance calculations but corrects for both mis-calibrated forecasts, disconnected spatial areas, and the asymmetries of time (early is a different type of mistake than late). However, policy-makers need probabilistic predictions such that they can query many different questions from a common modeling apparatus depending on dynamic political and security contexts (Brandt and Freeman 2002). In this paper, we illustrate how pEMDiv can be extended to score probabilistic predictions through the expansion of a general network representation of the optimal transport problem into a multiplex network. We use predictions from the Violence Early Warning System (ViEWS) (Hegre, et al 2021, Hegre, et al 2022) to illustrate the benefits and computational costs of this approach. We also illustrate a simpler to compute metric that also works by aggregating mass across ever-wider bins instead of transporting it across fixed nodes. Our work illustrates how cross-bin evaluation metrics such as pEMDiv and CRISS uniquely illuminate useful model traits and guide the creation of more policy-relevant and impactful ensembles.

Investigating The Practical Utility Of Conflict Early Warning Models: A Decision-Maker Perspective On Conflict Forecasting

Presenter: Paul Flachenecker
Co-authors: Paul Flachenecker

Over recent decades, violent conflicts have caused immense human suffering, mass displacement, and economic shocks. The looming climate crisis threatens to exacerbate this trend. In response, governments

and international organizations are investing significantly in conflict forecasts, harnessing big data and artificial intelligence (AI) to predict and prevent further outbreaks. However, despite these efforts, current methodologies struggle to lead to effective conflict warnings, thus failing to bridge the well-documented warning-response gap. Contemporary scholarship's emphasis on enhancing the predictive accuracy of conflict forecasts does not suffice to overcome this issue, as it fails to incorporate the crucial interaction between forecasts and decision-makers. This oversight represents a critical gap in research, as the effectiveness of early warnings hinges upon the perception and acceptance of their end-users. This paper addresses the lacuna in contemporary scholarship by advancing a two-pronged approach. First, it introduces a theoretical framework that explains the impact of the design of forecasting algorithms on warning success and introduces the decision-making level as a moderating variable. Second, this research leverages empirical evidence gathered from interviews with decision-makers to evaluate the practical utility of different design choices for warning effectiveness. The semi-structured interviews with decision-makers in German ministries and EU institutions are scheduled for April and May 2024. The results of this research are expected to shed light on how forecasting design choices, such as the selection of the dependent variable, model-fit metric, or data source, influence the effectiveness of conflict forecasts for decision-makers. These insights are anticipated to contribute to a little-explored research area that incorporates end-user inputs into the creation of conflict forecasts. Importantly, these findings are not limited to conflict forecasting but are likely to apply to other areas of 'algorithm in the loop' decision-making processes. Ultimately, this research aims to advance the field of forecasting toward a more responsive and impactful paradigm capable of anticipating crises and triggering timely responses.

Exploiting Systemic Biases Of Equity Analysts With Pre-Trained Neural Forecast Models

Presenter: Cristian Challu
Co-authors: Cristian Challu;Max Mergenthaler;Azul Garza;Roberto Gómez-Cram

This work proposes a novel methodology for enhancing financial forecasting accuracy and investment strategy formulation by leveraging the systematic bias of earning calls predictions. Our approach learns the future distribution of biases on earning calls to inform an investment strategy based on the expected relative performance of stocks. We investigate a range of established deep-learning methods and transfer pre-trained models and demonstrate the latter excels in predicting the median relative errors of analysts' forecasts. When these error predictions are incorporated into the revenue forecasting process, we observe a substantial improvement over conventional analyst predictions, with a 12% reduction in mean absolute error (MAE) and a 10% reduction in root mean square error (RMSE). An investment strategy informed by these refined forecasts generates alpha values exceeding 1%, significantly outperforming the market over a two-week investment cycle. These findings support the integration of pre-trained foundation model-driven error forecasting into financial analysis to enhance the precision of investment decisions.

Forecasting With The Latest In Ml: Automl, Pretrained Models, And Beyond

Presenter: Caner Turkmen
Co-authors: Caner Türkmen;Lorenzo Stella

In this talk, we will be covering a range of recent developments in machine learning for forecasting. Specifically, we will give an overview of pretrained time series models built on large neural network architectures for zero-shot forecasting. We will then provide a summary of the latest developments in automated machine learning (AutoML) for time series forecasting. Finally, we will show how these can be brought together to build a highly accurate and highly robust forecasting pipeline for a variety of real-world use cases.

Baseline As A Foundation For Demand Planning And Promotion Evaluation

Presenter: Tilak Raj Singh
Co-authors: Tilak Raj Singh;Anu Thomas;Nigel Nicholson

In the Consumer Packaged Goods (CPG) sector, the strategic planning of promotions is a critical success factor for maintaining competitive advantage. Effective evaluation of the incremental sales that promotions drive requires CPGs to calculate baseline sales or what the sales would have been had the promotion not run. A good baseline should adjust for a range of factors including product trends and seasonality, product lifecycle, price changes, changes in store distribution and cannibalization from other products in the family. A good baseline should also not react to short-term sales disruptions and spikes. The complexity of pre and post-promotion analysis arises due to the diverse range of mechanics and tactics employed and the variations in spend, duration and frequency. To tackle this challenge effectively, establishing a stable baseline that encompasses both historical and future periods is fundamental.Establishing and agreeing on a common baseline is crucial for an effective S&OP process. Different supply chain stakeholders: Supply Chain, Customer Development, Marketing and Finance need to collaborate effectively to agree promotion plans and long-term sales targets.We have developed an approach calculating robust baselines that is being rolled out across different geographies. This efficiently generates baselines for over 10,000 products and hundreds of customers. The approach leverages techniques from time series analysis, optimization, and hierarchical modelling. We aim to share our findings on several key aspects: (i) determining the appropriate hierarchy level and granularity for baseline estimation, (ii) interpreting the baseline, (iii) establishing the baseline for heavily promoted or new products, (iv) ensuring cycle-on-cycle stability when refreshing baselines, (v) identifying performance measures for evaluation, (vi) addressing product cannibalization, and (vii) conducting large-scale computational experiments, results and learnings.

The Value Of Coherent Forecasts For Decision Making

Presenter: Benedikt Sonnleitner
Co-authors: Benedikt Sonnleitner;Simon Hoelck;Nikolaos Kourentzes

Decision making in enterprises is often done across different levels, forming a hierarchy of decisions, that is served by a hierarchy of forecasts. These forecasts are by default not aligned, they are incoherent. This is commonly addressed by forecast reconciliation, which aligns forecasts on different levels. We evaluate the cost of forecast incoherency, and the value of forecast reconciliation in a predict then optimize scheme for different news-vendor inspired optimization set ups, and a staff scheduling Mixed Integer Problem, using workload data from 42 cross docks. We find that forecast incoherency is a substantial driver for actual cost in decisions, especially when planning is done upon non-aligned point forecasts, and when reconciliation is not done on the optimization side. In these cases, coherency gains by forecast reconciliation also translate to reduction in actual cost.

Controlling Forecast Bias In Demand Forecasting

Presenter: Iman Vasheghani Farahani
Co-authors: Iman Vasheghani Farahani

Retailers rely on demand forecasts to optimize their inventory and replenishment processes. While this underscores the importance of forecast accuracy, many other factors affect business operations. Two common factors are forecast accuracy and bias. Sometimes retailers decide to over-forecast demand to prevent stockouts, essentially accepting a positive forecast bias to lower the risk of lost sales. Consequently, the question arises: How can retailers generate forecasts with acceptable bias while targeting optimal accuracy?In this talk, we propose a two-stage heuristic to address this question. Our approach is tested across two model families: autoregressive integrated moving average (ARIMA) and exponential smoothing (ESM). In the initial stage, the best ARIMA or ESM model is identified along with its parameter estimates. If the resulting forecast bias is deemed unacceptable, a constraint specifying the desired level of bias is introduced to the parameter estimation optimization program and solved in the second stage. It is assumed

that the structure of the ARIMA or ESM model remains intact, while only the model parameters are adjusted to meet the specified constraint on forecast bias. The acceptable level of forecast bias serves as a hyperparameter in this framework. By varying this hyperparameter, decision makers have the flexibility to generate multiple forecasts and identify the one that offers an acceptable balance between accuracy and bias.

Beyond Accuracy: Navigating The Association Between Hierarchical Forecasting And Decision Making

Presenter: Mahdi Abolghasemi
Co-authors: Mahdi Abolghasemi

Hierarchical forecasting methods promise to generate coherent forecasts that can be used for consistent decision-making across different levels of a hierarchy. While reconciliation offers forecast coherency and often improves the average forecast accuracy (though it can compromise forecast accuracy at certain levels and nodes), it may not be necessary for decision-making in practice. Moreover, it is not evident if consistent forecasts lead to consistent decisions. Studies show that forecasts do not translate directly to decisions. We show how reconciled forecasts can lead to various decisions in hierarchical time series. Since forecasts across different levels serve different purposes and may have different business values for organizations, we discuss the criteria and how we can consider them in hierarchical forecasting. We aim to develop decision-informed forecasts for hierarchical time series by incorporating decision information from across the hierarchy.

Probabilistic Reconciliation Of Mixed-Type Hierarchical Time Series

Presenter: Lorenzo Zambon
Co-authors: Lorenzo Zambon;Dario Azzimonti;Nicolò Rubattu;Giorgio Corani

In hierarchical forecasting, a joint predictive distribution is coherent if it assigns positive probability only to points that satisfy the summing constraints. Probabilistic reconciliation is a post-processing step that adjusts the base incoherent forecast distributions, making them coherent. Current methods focus on the reconciliation of Gaussian or discrete base forecasts. In many applied scenarios, such as retail sales forecasting, the disaggregated time series have low-count values, while the aggregated ones are smooth and thus modeled as continuous. In this case, the base forecasts are Gaussian for the aggregated levels, and discrete for the bottom level. There are currently no methods for the reconciliation of such mixed hierarchies. This is a recognized open problem; we propose two approaches.We call the first mixed conditioning: it creates a mixed joint distribution of all the base forecasts, which are discrete for the bottom time series and Gaussian for the upper. The joint predictive distribution is then conditioned on the hierarchy constraints, yielding a coherent reconciled distribution. This method extends previous work on probabilistic reconciliation via conditioning.We call the second approach top-down conditioning; it works in two steps. First, the upper base forecasts are reconciled via conditioning, using only the hierarchical constraints between the upper: since they are Gaussian, this can be done analytically. Then, the reconciled joint bottom distribution is computed in a probabilistic top-down fashion.We present experiments on real time series datasets, showing that mixed conditioning is effective in case of moderately sized hierarchies, while top-down conditioning is better suited for hierarchies with thousands of time series, such as the retail time series of the M5 competition.

Sub-Hierarchical Forecasting

Presenter: Fotios Petropoulos
Co-authors: Fotios Petropoulos;Ross Hollyman

Hierarchical forecasting has been a prominent research topic for the last 15 years mainly due to its practical relevance. Various reconciliation methods have been proposed, and these methods offer coherent point and probabilistic forecasts across the various aggregation levels coupled with improved forecasting

performance across the hierarchy. However, one main issue of such reconciliation methods is their limited applicability on large hierarchies due to the computational complexity related with the matrix calculations involved. We address this issue by proposing an overarching approach to forecast reconciliation methods that is based on the construction of sub-hierarchies. Our approach can be applied in conjunction with any known forecast reconciliation method, either for point or for probabilistic forecasts. Sub-hierarchical forecasting not only renders the reconciliation calculations possible for hierarchies of any size but also results in robust improvements over the existing reconciliation methods. We apply our proposed approach on a large hierarchical dataset in the retail context and we showcase its value in practice.

Inflation Expectations And Real-Time Inflation Models

Presenter: Nicolas Bonino-Gayoso
Co-authors: Nicolas Bonino-Gayoso;Monica Correa-Lopez

Can we robustly and systematically identify the type of inflation expectations that matter the most to euro area inflation dynamics? This paper explores this question by analyzing the predictive power and forecasting stability in real time of open-economy inflation models of the (NK)PC tradition. We use vintages of euro area quarterly data starting in 2009:Q1 since the real-time update of inflation expectations contains information at each quarter t about the (perceived) future, and since a single snapshot is not enough for identification. In a thick modeling approach with rolling regressions, we assess the results of a (pseudo) out-of-sample conditional forecasting exercise by means of meta regressions. We compare the results obtained under the NKPC framework against those produced by a univariate model (Stock and Watson, 2007).

Inflation Trends In Colombia: Detecting Breaks And Forecasting Inflation

Presenter: NA
Co-authors: Norberto Rodriguez;Héctor Zárate

For a developing country, it is important to keep inflation rates at low levels and under control. In the Colombian case, its annual inflation reached 13.34% in the first quarter of 2023, the highest rate since the start of the inflation targeting regime for monetary policy around 2000. However, some groups of the basket show signs of lower inflation, while others show higher inflation. The persistence of this trend is a matter of active debate that involves analysing the trend component of both year-to-year and month-to-month changes in the price indices: in this paper, we use Casini and Perron (2018) ["Structural Breaks in Time Series", Oxford Research Encyclopedia of Economics and Finance, pp 1-37] method to identify shifts in inflation trends based on the 188 price indexes that make up the basket. Additionally, we employed those models with breaks to forecast the total and group inflation for 2024 and 2025, finding that inflation will decline for most groups in 2023 but maintain or even increase for some key groups in 2024. In this application we find out that the bottom-up strategy attains good results compared against non-disaggregation or non-breaks considerations. The middle-up strategy is also evaluated.

Forecasting Inflation With Supply Shocks

Presenter: Alessandro Barbarino
Co-authors: Alessandro Barbarino

The traditional take on inflation is that it is very hard to predict and it is very difficult to beat a random walk. Phillips curves have flattened. Not only, the pass-through from wages has declined. The postpandemic runup in inflation has the potential to offer the second important source or variation after the Great Inflation to study inflation dynamics. I summarize inflation dynamics by major components and I show that supply side shocks explain well both the runup in core inflation and the ensuing surpising deflation. I also construct indicators of wages that are tailored to the components of inflation that I study and show the strong relationship between them. My wage aggregate is significant and helps to forecast

inflation and improve upon a random walk. Also the Phillips curve forecasts significantly help improve the forecast above a random walk.

Forecast Combination And Interpretability Using Random Subspace

Presenter: Boris Kozyrev
Co-authors: Boris Kozyrev

This paper investigates forecast aggregation via random subspace regressions (RS) and explore the potential link between RS and the Shapley value decomposition (SVD) using the US GDP growth rates. This combination of techniques enables handling high-dimensional data and reveals the relative importance of each individual forecast. First, I demonstrate that in certain practical instances, it is possible to enhance forecasting performance by randomly selecting smaller subsets of individual forecasts and obtaining a new set of predictions based on a regression-based weighting scheme. The optimal value of selected individual forecasts is also empirically studied. Then, I propose a connection between RS and the SVD, enabling the examination of each individual forecast's contribution to the final prediction, even when the number of forecasts is relatively large. This approach is model-agnostic (can be applied to any set of forecasts) and facilitates understanding of how the aggregated prediction is obtained based on individual forecasts, which is crucial for decision-makers.

Random Factor Model Forecasts

Presenter: Rachida Ouysse
Co-authors: Rachida Ouysse;Andrey Vasnev

The literature on combination forecasts suggests that, in empirical situations involving actual forecasts, simple averages of predictions often outperform more complex statistical combination methods as demonstrated by prior studies (Bates and Granger (1969), Granger and Ramanathan (1984)). In another strand of the literature, dense methods like factor model forecasts, which consolidate large information, have been demonstrated to be more accurate than many single econometric models (Stock and Watson (2004)). This paper investigates a scenario where multiple forecasters have access to subsets of the complete information set. We propose a strategy which entails combining a large number of partial information forecasts using a random factor model forecat combination. A key new finding of this paper is that combination of forecasts from partial information sets outperforms a full information forecast. An application to forecasting the monthly growth rate of U.S. industrial production, where the full set of predictors consists of 130 economic indicators, shows that RFF with shrinkage weights outperforms the full information factor forecasts

International Inflation Vulnerability

Presenter: C. Vladimir Rodriguez Caballero
Co-authors: C. Vladimir Rodriguez-Caballero;Esther Ruiz;Ignacio Garron

Policymakers need to understand the sources of international movements in inflation and how they affect domestic inflation. Also, there is an increasing interest in assessing the level of exposure of domestic prices to small probability but potentially catastrophic international scenarios. In this paper, we propose the construction of domestic conditional densities for inflation based on factor-augmented quantile regressions estimated using international factors extracted by fitting a multi-level Dynamic Factor Model (ML-DFM) to a broad set of international inflations. Forecasts of the domestic densities are used to analyze the forecast accuracy across countries, finding interesting heterogeneous patterns. Furthermore, we measure the vulnerability of domestic inflation by constructing scenarios under stressed international factors. The choice of severe yet plausible stress scenarios for the factors is based on the joint probability distribution of the underlying factors driving inflation. The economic shock brought by the Ukraine war provides a natural environment in which to assess the vulnerability of inflation with the proposed methodology

Evaluating Leading And Coincident Indicators Of Regional Business Cycles, A Closer Look Into The Spanish Scenario

Presenter: Aránzazu de Juan Fernández
Co-authors: Aránzazu De Juan Fernández;Marco Aurelio Pérez Navarro

Using the Linear Dynamic Harmonic Regression algorithm developed by Bujosa, García-Ferrer and Young (2007), we obtain the Composite Leading Indicator (CLI) and Composite Coincident Indicator (CCI) for the Spanish Regions (Comunidades Autónomas) using the same variables used for the CLI and CCI for the Spanish economy as a whole (Bujosa et al, 2013 and Bujosa et al, 2020) and we analyze if there exists a pattern in leading the Spanish recession periods. We find that almost all Spanish regions with a coastline show a similar behavior. Specifically, they lead more than a year of the recession troughs. Two autonomous communities deviate from this behavior: on the one hand, Castilla y León, which is on average two and a half years ahead of the peaks of the recessions and more than one year ahead of the troughs of the recessions. The other autonomous community is the Canary Islands, which is almost three years leading on average the peaks of the Spanish economic cycle but coincides on average with the troughs of the recessions. We also analyzed whether there are differences in the behavior of the CLI and CCI of each of the Spanish regions in each recession period. The results for the CLI of the Spanish regions are used to obtain predictions of annual GDP growth and annual regional GDP growth whenever possible.

Time-Varying Impacts Of Monetary Policies On State-Level Housing Markets: Evidence From The Covid-19 Period

Presenter: MeiChi Huang
Co-authors: Meichi Huang

This study investigates the time-varying influences of monetary policies on state-level housing markets, utilizing a time-varying parameter vector autoregression model with stochastic volatility (TVP-VAR-SV). The results suggest that time-varying state-level volatilities in housing price returns and permits growths are evident, and some of them are higher in the Covid-19 period 2020-22 than the 2007-9 housing crisis. Monetary policies exert less persistent but stronger effects on housing quantities than housing prices. The findings provide supportive evidence that the Covid-19 pandemic enhances monetary-policy influences, and aggressive contractionary monetary policies in 2022 play vital roles in driving state-level housing markets. The differences and commonalities across state-level housing markets yield new implications for policy-making and risk diversification.

Tracking Sectoral Economic Conditions

Presenter: Daan Opschoor
Co-authors: Daan Opschoor

We construct a novel set of monthly U.S. sector-level economic conditions indices from a small but diverse set of sectoral economic indicators using mixed-frequency dynamic factor models. The resulting indices are driven by a balanced mix of the underlying indicators and display considerable heterogeneity, particularly in the depths, timing and duration of their downturns. Moreover, the sectoral economic conditions are driven by a common factor that explains most fluctuations in the overall economy and is closely related to aggregate production. Meanwhile, the service-providing sectors are additionally driven by a correction factor that handles the heterogeneous impacts of the financial crisis and covid pandemic. Lastly, sector-level GDP growth nowcasts are constructed, which are found to consistently outperform a simple autoregressive benchmark for almost all sectors, especially during the covid pandemic.

Modeling And Forecasting Racial Disparities In The Us Labor Market

Presenter: Meil Ericsson
Co-authors: Neil Ericsson;Fabian Leal;Kaythari Maw

The COVID-19 pandemic resulted in the most abrupt changes in US labor force participation and unemployment since the Second World War, with consequences differing markedly by race. Assessing these changes in the labor market and understanding their causes is central to implementing the Fed's dual mandate of maximum employment and stable prices.To quantify and interpret the pandemic's effects, we first model the pre-pandemic US labor market. Using machine learning for model selection and indicator saturation, we formulate well-specified dynamic cointegrated vector autoregressions that capture the short- and long-run relationships between disaggregated unemployment rates and labor force participation rates (LFPRs) for monthly data over 1980–2019. Our analysis establishes three types of long-run relationships across Whites, Blacks, Hispanics, and Asians: persistent gaps in LFPRs, persistent gaps in unemployment rates, and discouraged-worker or added-worker effects that relate the LFPR to the unemployment rate for a given race. Dynamic behavior is numerically, statistically, and economically highly significant and differs by race, with marked effects from the economic business cycle. We then use these models to forecast ex ante into the pandemic to understand the pandemic's labor market consequences, treating these forecasts as being from an alternative scenario in which the pandemic didn't occur. This approach—common in policy analysis—turns forecasting on its head, with the observed large forecast errors reflecting both behavioral and policy responses to the pandemic. Heterogeneity across race is particularly prominent at the pandemic's outset. For LFPRs, initial recovery from the pandemic is slower and more prolonged for Blacks relative to Whites, whereas recovery is relatively rapid for Hispanics and Asians. Intercept correction quantifies these effects. Unemployment rates for all demographic groups spiked early on in the pandemic, returning to near pre-pandemic levels by late 2021. The pre-pandemic relationships between and among the LFPRs and unemployment rates persist into the pandemic, with a prolonged shift in the equilibrium LFPRs, and significant dynamic adjustments back towards equilibrium for the unemployment rates.

Additive Covariance Matrix Models: Modelling Regional Electricity Net-Demand In Great Britain

Presenter: Matteo Fasiolo
Co-authors: Matteo Fasiolo;Vincenzo Gioia;Ruggero Bellio;Jethro Browell

Forecasts of regional electricity net-demand, consumption minus embedded generation, are an essential input for reliable and economic power system operation, andenergy trading. While such forecasts are typically performed region by region, operations such as managing power flows require spatially coherent joint forecasts, which account for cross-regional dependencies. Here, we forecast the joint distribution of net-demand across the 14 regions constituting Great Britain's electricity network. Joint modelling is complicated by the fact that the net-demand variability within each region, and the dependencies between regions, vary with temporal, socio-economical and weather-related factors. We accommodate for these characteristics by proposing a multivariate Gaussian model based on a modified Cholesky parametrisation, which allows us to model each unconstrained parameter via an additive model. Given that the number of model parameters and covariates is large, we adopt a semi-automated approach to model selection, based on gradient boosting. In addition to comparing the forecasting performance of several versions of the proposed model with that of two non-Gaussian copula-based models, we visually explore the model output to interpret how the covariates affect net-demand variability and dependencies. Methods for fitting multivariate Gaussian GAMs, where both the mean and the covariance matrix can depend on covariates via additive models, are provided by the SCM R package available at https://github.com/VinGioia90/SCM/subsection%7BAdaptive Gams For Extreme Quantile Forecasting}Presenter: Yannig Goude
Co-authors: Yannig Goude;Omar Himich;Amaury Durand

Uncertainty quantification is essential for the efficient and reliable operation of electricity systems. Probabilistic load forecasting is a critical aspect of this quantification. Recently, electricity consumption has been affected by unexpected or extreme events, requiring adaptive forecasting methods. Much work has been done on quantile forecasting and adaptive quantile forecasting models, among which GAMs and quantile GAMs have been good candidates, achieving a favourable trade-off between interpretability and efficiency. Recent developments in the machine learning community have been made for modelling rare

events in the regression setting. We propose to extend these approaches to the adaptive setting and present some results on electricity demand data.

Efficient mid-term forecasting of hourly electricity load using generalized additive models

Presenter: Monika Zimmermann
Co-authors: Monika Zimmermann;Florian Ziel

Accurate medium-term (several weeks to one year) hourly power load forecasts are essential for strategic decision-making in power plant operation, maintenance scheduling, and grid stability. However, while numerous models effectively predict short-term (a few days) hourly load, medium-term forecasting solutions remain scarce. In addition to daily, weekly, and annual seasonal and autoregressive effects, the main modeling challenges in medium-term load forecasting are the capture of weather and holiday effects, as well as socio-economic nonstationarities in the data. To address these challenges, we propose a novel forecasting method using Generalized Additive Models (GAMs) built from interpretable P-splines and enhanced with autoregressive post-processing. This model uses smoothed temperatures, Error-Trend-Seasonal modeled non-stationary levels, a nuanced representation of holiday effects with weekday variations, and seasonal information as input. The proposed model is evaluated on load data from 25 European countries. This analysis shows that the model not only achieves superior forecasting accuracy compared to all benchmarks, but also provides valuable insights into the impact of individual components on the predicted load, as the model is fully interpretable. With fast computation times of a few seconds for several years of hourly data, and with significantly improved forecasting accuracy compared to state-of-the-art benchmark models.

Detrending Vs. Trend Modeling For Natural Gas Demand Forecasting

Presenter: Richard Povinelli
Co-authors: Richard Povinelli;Ronald Brown

This study proposes a novel detrending algorithm to enhance short-term (daily) natural gas demand forecasts. Traditional forecasting methods trained on short-term data may lack diverse consumption patterns. While longer datasets offer a wider range of customer behaviors, their historical trends deviate from recent patterns. The proposed algorithm addresses this by detrending long-term data (yearly) instead of directly modeling trends within the forecasting model. The approach involves learning a five-parameter model for the current year and a historical year, estimating adjustment factors using the difference in model coefficients between these years. This process is repeated using a year-long sliding window with a stride of one month. The resulting sequence of coefficients are smoothed and then used to detrend past years, making them resemble current consumption patterns. Evaluated across 172 gas demand time series (7-28 years), the algorithm demonstrates improved performance. Models trained with our algorithm on 10 years of data achieved a lower Weighted Mean Absolute Percentage Error (WMAPE) of 6.5% compared to models without detrending (6.7% WMAPE) and those incorporating trend terms (6.9% WMAPE).

Enhancing Contraceptive Demand Forecasting: A Probabilistic Approach For Improved Family Planning Supply Chain Management

Presenter: Harsha Chamara
Co-authors: Harsha Ruwan Chamara Halgamuwe Hewage;Bahman Rostami-Tabar;Aris Syntetos;Federico Liberatore

Ensuring an effective family planning supply chain is crucial for maintaining a reliable provision of contraceptives. However, in developing countries, inefficiencies in the supply chain often lead to stockout situations within family planning health systems. This contributes significantly to the increase in unmet contraceptive needs over time, resulting in dropouts and unintended pregnancies with serious societal implications. Accurate demand forecasts are essential for making reliable inventory and replenishment

decisions for contraceptives. Unfortunately, practitioners at the operational level often rely on simplistic methods that may not perform well in all scenarios. Recognizing this challenge, United States Agency for International Development (USAID) initiated a competition aimed at developing intelligent forecasting methodologies. Despite these efforts, there remains a lack of capability to generate forecast distributions. However, decision-makers in the field are looking for dynamic estimates over time as a risk management tool to cope with the uncertainty of the estimations. To address this gap, we developed a probabilistic forecasting framework using data from the Logistics Management Information System (LMIS) of Cote d'Ivoire. Our study employed both univariate and machine learning models. Our findings indicate significant improvements in forecasting model performance compared to submissions to USAID. By ensembling these models, we demonstrated that both point forecasts and forecast distributions outperformed any individual model. Furthermore, we present a probabilistic forecasting framework that can be generalized to any domain using any forecasting model. This framework holds promise for enhancing forecasting accuracy and reliability across various sectors.

Deep Learning Forecasting Model For Kidney Function Post-Transplant Exploiting Missing Values

Presenter: Zahra Sajjadifar
Co-authors: Zahrasadat Sajjadifar;Axel Geysels;Alexandre Arnould;Maarten Naesens;Oscar Mauricio Agudelo Manozca;Bart De Moor

While regular and complete time series with evenly spaced timestamps enhance the potential for analysis and modeling, in real-world scenarios irregular time series with missing values are common which pose challenges in effective data handling and modeling. For solving the missing data problem in time series analysis, there are two main approaches either to impute the data before starting the prediction tasks or modify and apply changes directly to the predictive algorithms in a way that they become aware of irregularity in time series and be able to handle the missing values. In our research, we focus on clinical biomarkers time series, particularly the estimated glomerular filtration rate (eGFR) monitored after kidney transplantation for each recipient. The level of eGFR can be interpreted as an alert for performing a transplant biopsy if the values are not in a normal range. However, this normal range is patient-specific and there is no global threshold to determine the abnormal range. The clinically relevant task is to predict future values in a specific time interval from past observations. The forecasted future eGFR values for each graft can be considered as a patient-specific expected range and the newly measured eGFR values can be benchmarked against the predicted ones to evaluate variances from the expected range and this is going to prompt the patient-specific alarm in clinical practice. Given the frequent tracking of eGFR after transplant, an irregular time series emerges. The missingness in such time series is not random but it is systematic since the missing rate increases over time resulting in larger gaps due to the decrease in the follow-up rate of patients. The contribution of this study is implementing a one-stage sequence-to-sequence forecasting model modified to handle missing values using the gated recurrent unit with decay rate (GRU-D) to exploit the systematic missingness in the predictive model and compare the results with two-stage imputation and forecasting models based on the conventional autoregressive integrated moving average (ARIMA) and sequence-to-sequence recurrent neural network based on the gated recurrent unit (GRU).

Predictive Analytics In Public Health: Forecasting Malaria Incidence In Maharashtra

Presenter: Adithya Somaraj
Co-authors: Adithya B. Somaraj;Praveen D. Chougale;Usha Ananthakumar

Malaria remains a significant public health concern, particularly in regions with high transmission rates, such as Maharashtra. Accurately predicting malaria cases is crucial to the effective control and management of malaria. This study looks at malaria incidence in Maharashtra from 2012 to 2019, and it extends its methodology to predict malaria cases in specific districts of Maharashtra with high malaria burden. Regression models, such as Generalized Linear Models (GLMs) with linear, Poisson, and negative

binomial distributions, as well as ensemble machine learning and advanced deep learning models, were used to capture and predict the patterns in monthly malaria case counts accurately. Critical predictive variables that are important for understanding the determinants of malaria incidence are integrated into the study, such as population density and meteorological data, including temperature, humidity, and rainfall. By conducting a thorough comparative analysis of the results derived from these models, more precise and dependable forecasts are achieved.Through rigorous validation and analysis, this research presents a multidisciplinary strategy that incorporates statistics, machine learning, and deep learning techniques to predict the prevalence of malaria effectively. The applications of these methodologies offer a nuanced prediction capability, which is essential for the identification of potential outbreaks and the formulation of proactive intervention strategies. This research can make a significant contribution to the efforts to control the malaria burden in Maharashtra and minimize its effect on public health and economies in both urban and rural areas.

Evaluating The Randomness Of The M6 Competition

Presenter: Matt Schneider
Co-authors: Matthew Schneider;Rufus Rankin;Prabir Burman;Alexander Aue

The M6 Competition assessed the performance of competitors using a ranked probability score and an information ratio (IR). While these metrics do well at picking the winners in the competition, crucial questions remain for investors. To address these questions, we compare the performance of the competitors to a number of conventional (long-only) and alternative indices using standard industry metrics. To evaluate the randomness of the competitors, we use formal statistical tests on coefficients from factor models and a simulation study involving 10,000 randomly selected portfolios. Our simulation study shows that most competitors have lower risk-adjusted returns and lower maximum drawdowns than portfolios chosen at random, and that the majority of competitors were unable to generate significant out-performance. We also introduce two new strategies by investing in the top 10 performing superstars and bottom 10 performing superlosers from the prior month. Interestingly, we discover mean reversion behavior in that the portfolio of superlosers gains value while the portfolio of superstars loses value. We also discuss challenges that would face an investor seeking to allocate to one or more of these strategies.

The Impact Of Engagement And Consistency In The M6 Forecasting Competition

Presenter: Evangelos Spiliotis
Co-authors: Evangelos Spiliotis;Anastasios Kaltsounis;Evangelos Theodorou;Vassilios Assimakopoulos

The M6 financial competition attracted 226 teams and involved 12 monthly submission points for forecasts and investment decisions. Although teams could update their entries every month, they also had the option to retain a previous submission for one or multiple future rounds. Consequently, the frequency and average interval of submissions varied significantly across teams. Additionally, the investment performance of the teams, including the top performers, varied greatly across the submission rounds. Drawing from these observations, we investigate the effect of engagement in winning the M6 competition and, consequently, the value that regular portfolio updating can add to investment management. Moreover, we examine whether teams that consistently performed well managed to distinguish themselves from teams that have sporadically performed exceptional, thus testing the robustness of the M6 results. We find that team engagement and performance consistency are sufficient but not necessary conditions for constructing efficient portfolios and identify some additional parameters affecting the results.

The Value Of Avoiding Overconfidence: Evidence From The M6 Competition

Presenter: Spyros Makridakis
Co-authors: Spyros Makridakis;Evangelos Spiliotis;Maria Michailidis

One of the major objectives of the M6 competition was to identify accurate ways of forecasting asset prices and use such forecasts to improve the return of investment decisions. The competition was live, lasting one year and involving twelve submission points. In the forecasting track of M6, the participants were asked to estimate the probability that each of the 100 selected assets would be ranked within the first, second, third, fourth or fifth quintile in terms of their relative percentage returns. The results of the competition highlighted the challenges of the task with less than 25% of the teams estimating the requested probabilities more precisely than the benchmark (assigning equal probabilities to the five quintiles), while those that did so reporting inconsistent performance across the twelve submission points. Moreover, their forecast accuracy improved by less than 2.5% whereas there was zero correlation between overall forecasting accuracy and return on investment. Trying to investigate the factors contributing to these results, we find that accounting for price volatility is critical for improving performance and that, due to said volatility, assuming symmetric probabilities for quantiles 2 and 4, and especially 1 and 5, can lead to higher forecasting accuracy. We also demonstrate that when these symmetries are considered, even simple methods can outperform the benchmark and demonstrate that the corresponding forecasts can result to more profitable investment decisions.

Probabilistic Forecast Reconciliation Targeting Quantiles Using Bilevel Optimisation

Presenter: Anastasios Panagiotelis
Co-authors: Anastasios Panagiotelis;Hossein Alipour;Nam Ho-Nguyen;George Athanaspoulos

Collections of time series where some series are aggregates of one another, are known as hierarchical time series. When forecasting hierarchical time series, the linear constraints due to this aggregation structure may not hold. Forecast reconciliation allows for the adjustment of such forecasts ex post, to ensure that aggregation constraints are satisfied, i.e. forecasts are coherent. In the probabilistic setting, this implies that regions of points that do not satisfy the constraints are assigned zero probability, or alternatively, that a sample from the predictive distribution only contains coherent points. In this work, an algorithm for forecast reconciliation is proposed targeting optimality with respect to a given quantile level. This framework builds upon the score optimisation framework introduced by Panagiotelis et al (2023) but uses the pinball loss as the objective function. Due to the fact that the reconciled quantiles are themselves the solution to an optimisation involving pinball loss, the problem become one of bilevel optimisation. While we show that the problem can be solved by mixed integer linear programming, this proves to be slow to compute even with modern solvers and moderately sized hierarchies. Therefore we propose an approximate technique based on a smooth version of the pinball loss function. By exploiting a lemma that allows gradients to be found in bilevel optimisation problems, we show that the problem can be tackled using gradient based methods such as stochastic gradient descent. The proposed method will be demonstrated with an application to Australian tourism data.

Combining Probability Distribution Forecasts Within A Random Forest

Presenter: Siddharth Arora
Co-authors: Siddharth Arora;James Taylor

Random forests (RF) are one of the most commonly used off-the-shelf machine learning approaches. For regression, RF estimates the conditional mean of the response variable by combining individual forecasts from multiple regression trees, where each tree is grown using a bagged version of the training data and a random set of input features for split point selection. Thus, RF is an ensemble learning approach that utilizes the concept of 'wisdom of the crowd'. Quantile regression forests (QRF) were proposed as a generalization of RF to estimate the full conditional distribution of the response variable. Specifically, QRF combines the forecasts of the cumulative probability distributions generated from the different regression trees. The combination is performed using an equally-weighted linear opinion pool, which can be visually interpreted as the vertical average of the probability distribution forecasts from each tree. If the averaging is performed horizontally, this amounts to aggregating quantile forecasts from the different trees. Recently,

as an alternative to vertical and horizontal combining, the idea of angular combining has been proposed, where the forecasts of probability distributions are combined at an angle. In this study, we propose a generalization of QRFs, where the distributional forecasts obtained from each regression tree is permitted to be combined vertically, horizontally or at an angle, with the choice dictated by the continuous ranked probability score. For vertical combination, the approach defaults to QRF. We evaluate the out-of-sample probabilistic forecasting performance using a variety of different datasets.

Distribution-Free Conformal Joint Prediction Regions For Neural Marked Temporal Point Processes

Presenter: Souhaib Ben Taieb
Co-authors: Souhaib Ben Taieb;Victor Dheur;Tanguy Bosser;Rafael Izbicki

Sequences of labeled events observed at irregular intervals in continuous time are ubiquitous across various fields. Temporal Point Processes (TPPs) provide a mathematical framework for modeling these sequences, enabling inferences such as predicting the arrival time of future events and their associated label, called mark. However, due to model misspecification or a lack of training data, these probabilistic models may provide a poor approximation of the true, unknown underlying process, with prediction regions extracted from them being unreliable estimates of the underlying uncertainty. This paper develops more reliable methods for uncertainty quantification in neural TPP models via the framework of conformal prediction. A primary objective is to generate a distribution-free joint prediction region for the arrival time and mark, with a finite-sample marginal coverage guarantee. A key challenge is to handle both a strictly positive, continuous response and a categorical response, without distributional assumptions. We first consider a simple but overly conservative approach that combines individual prediction regions for the event arrival time and mark. Then, we introduce a more effective method based on bivariate highest density regions derived from the joint predictive density of event arrival time and mark. By leveraging the dependencies between these two variables, this method excludes unlikely combinations of the two, resulting in sharper prediction regions while still attaining the pre-specified coverage level. We also explore the generation of individual univariate prediction regions for arrival times and marks through conformal regression and classification techniques. Moreover, we investigate the stronger notion of conditional coverage. Finally, through extensive experimentation on both simulated and real-world datasets, we assess the validity and efficiency of these methods.

Information-Theoretic Properties Of Score-Driven Models

Presenter: Ramon de Punder
Co-authors: Ramon De Punder;Timo Dimitriadis;Rutger-Jan Lange

Score-driven models have been highly successful in the last decade, used in over 300 published articles. Much of this literature relies on the purported optimality result in Blasques et al (2015), which (roughly) states that score-driven updates are unique in locally reducing the Kullback-Leibler (KL) divergence relative to the true density with probability one in the observation. This result is at odds with other well-known optimality results; the Kalman filter, for example, is optimal in a mean-squared error sense, but not with probability one in the observation (occasionally, it moves in the wrong direction). Here, we show that score-driven filters are, similarly, not guaranteed to improve the local KL divergence at every time step. The discrepancy with the seemingly stronger result in Blasques et al. (2015) derives from their use of an improper scoring rule. Even as an almost sure improvement at every time step is unattainable, we prove that score-driven filters are unique in reducing the local KL divergence relative to the true density in expectation. This positive (albeit weaker) result justifies the continued use of score-driven filters and places their optimality on solid footing.

Counterfactual Predictions In Shared Markets: A Global Forecasting Approach With Deep Learning And Spillover Considerations

Presenter: Klaus Ackermann
Co-authors: Klaus Ackermann;Priscila Grecov;Christoph Bergmeir

We introduce a novel forecasting method employing global deep learning models for estimating the causal effects of interventions across multiple units, incorporating counterfactual and synthetic control for policy evaluation in shared markets. This approach addresses potential spillover effects and leverages time series data for identification. We redefine causal effect estimation as predicting outcomes without intervention, first estimating counterfactual outcomes using high-dimensional time series data. This process utilizes cross-correlation in time series, employing an autoregressive recurrent neural network with parameter sharing. The second stage estimates and tests the average treatment effect on the target variable for statistical significance. Demonstrated through simulations and empirical studies, our method uniquely estimates effects using pre-treatment data in scenarios where traditional control unit assumptions fail. An empirical example estimates the impact of promotional deals on US grocery store sales, showcasing the method's applicability and contribution to existing literature.

Forecasting With Artificial Neural Networks For Sparse Data - An Empirical Evaluation

Presenter: Sven F. Crone
Co-authors: Sven F. Crone

Deep Artificial Neural Networks succeed in pattern recognition in text, image, and speech data, which has rekindled interrest for other data types including univariate and multivariate time series data. Long-Short term Memory Recurrent Networks (Ma et al, 2015), Deep Belief Networks (Kuremoto et al., 2014), and Structural Autoencoders (Gensler et al., 2016) are but few of the deep architectures promising increases in accuracy over the more established multilayer perceptrons (Crone, 2010). However, a recent survey in forecasting practice indicated that only few companies employ Deep Neural Networks in forecasting, and that over 50% of all AI forecasting projects in industry fail (Crone, 2022). This empirical study seeks to assess this discrepancy between the hype and lack of implementations of neural networks by assessing the empirical accuracy of different architectures on a real world industry dataset in FMCG, using reliable error metrics, fixed multi-step horizons and multiple rolling time origins. We compare accuracy against established benchmarks from statistics (ets, autoarima, theta), data science (Facebook's Prophet, Google's bsts), and machine learning methods (XGBoost, random forests). Our experiments confirm prior studies (e.g. Markidakis and Spiliotis, 2018), which show that the standard implementations of both deep and shallow neural networks architectures fail to outperform established statistical benchmarks of ets, arima on the monthly industry datasets still widely used in practice. However, when presented with carefully engineered feature sets, applying time series feature creation, transformation and selection, and providing algorithm parameter turning for these features, both the shallow and deep networks can be customised to outperform statistical, ml and data science benchmarks methods. We conclude that using „vanilla" deep or shallow neural network implementations from R or python packages yields poor results, but careful customisation of features can significantly increase their efficacy.

Univariate And Multivariate Probabilistic Forecasting Using Quasi-Randomized Neural Networks

Presenter: T. Moudiki
Co-authors: T. Moudiki

I present a family of machine learning models for univariate and multivariate time series forecasting based on quasi-randomized 'neural' networks (QRNN). The specificity of QRNNs is that they rely on QR numbers instead of backpropagation for constructing hidden layers. This makes them fast to train, and not

subject to exploding or annealing gradients.

Probabilistic Forecasting Of Intermittent Time Series: A Tweedie Distribution Head For Neural Network Architectures

Presenter: Nicolò Rubattu
Co-authors: Nicolò Rubattu;Stefano Damato;Dario Azzimonti;Giorgio Corani

Intermittent time series are characterised by recurring periods of zero values. They are frequently encountered across diverse domains, such as retail, environmental monitoring, and epidemiology.The traditional statistical distributions used in probabilistic forecasting of these time series are typically unimodal, assume continuous or discrete data alone, and do not natively handle a massive probability at zero or varying dispersion. In contrast, the Tweedie distribution is zero-inflated and handles positive values by exploiting the properties of the continuous Gamma distribution. It is suited for modelling intermittent demand time series, as it can jointly model the occurrence, magnitude, and variance of the target variable.Over the past few years, neural network models for probabilistic forecasting have shown impressive performances, especially when trained as global models. Instead of computing point forecasts solely, such models learn the parameters of the predictive distribution from which samples can be obtained. So far, the Tweedie distribution has only been used as a loss function for models that return point forecasts solely. Since naive or zero forecasts often dominate intermittent demand forecasting point metrics, a predictive distribution is crucial to assess forecast quality.We propose implementing the Tweedie as distributional output head in state-of-the-art recurrent neural network (RNN) and Transformer-based architectures. We focus on count time series and compare the performance against other well-known distributional outputs, such as Poisson and Negative Binomial.A comprehensive analysis demonstrates improvements in point forecast accuracy and uncertainty estimation. Indeed, we show how the Tweedie parametrisation offers flexibility, from heavily zero-inflated outcomes to more regular ones, thereby enhancing forecasting models for intermittent demand time series.

Model-Agnostic Bayesian Optimization For Computationally Tractable Value-Oriented Forecasting

Presenter: Hussain Kazmi
Co-authors: Hussain Kazmi;Maria Paskevich;Attila Balint;Joris Depoortere

Forecasts underpin most modern planning activities. However, an age-old question persists in their creation: should forecast models prioritize accuracy or downstream utility? This question has been debated for well over a century, with diverging opinions in the community, spawning a plethora of evaluation metrics. Probabilistic forecasts can enable planners to make risk-informed decisions, but often do so at elevated computational (and conceptual) costs and are often not supported by planning tools.Forecast+Optimize (or Predict+Optimize) workflows have witnessed renewed attention in recent years to address this issue. These methods integrate downstream task information in the prediction loop to adapt the forecast model's outputs in a principled manner. Recent approaches have attempted to do this through direct loss function augmentation during the forecast model creation process, i.e. when the model parameters are estimated using data.Even though this approach can improve downstream utility, it comes at much greater computational cost since an expensive optimization problem (i.e. the downstream task) has to be solved during each iteration of the learning algorithm (e.g. a boosted tree or neural network). As modern algorithms often require hundreds or thousands of iterations (epochs) to properly tune their parameters, this quickly becomes intractable. Hyperparameter search to identify the best models further exacerbates the situation. Mechanistically, direct loss function augmentation also requires algorithm and library specific tweaks, although this is slowly starting to change with modern implementations.In this work, we propose and utilize a Bayesian Optimization based framework which can perform the Forecast+Optimize task in a computationally efficient and model-agnostic manner. By integrating the downstream function primarily in the hyperparameter search routine, we show it is possible to (1) significantly improve model accuracy and downstream performance in a model-agnostic manner, and (2) reduce the computational complexity

of the Forecast+Optimize operation byseveral orders of magnitude. We validate these findings using two different case studies in energy and retail supply chain optimization. Our results show that the proposed methodology works well for both the case when accuracy is well correlated with accuracy, as well as when it is not.

Extending Privacy-Preserving Regression Trees To A Value-Oriented Forecasting Approach For Trading Der Resources

Presenter: Lukas Stippel
Co-authors: Lukas Stippel;Simon Camal;Georges Kariniotakis

In the state of the art, it has been shown that in several applications like spatiotemporal forecasting of wind or solar power, accuracy in the next hours can be improved if data from neighbor plants are used as input. However, this is not always feasible due to confidentiality constraints that prevent data sharing. This required privacy-preserving forecasting methods which are also applicable to other use-cases like EV-load demand. We apply vertical federated learning, leveraging different sources of information. The Energy forecasts are used to optimize decisions in applications. In cases like energy trading, recent advances have shown that decisions and the resulting value (i.e., revenue on energy markets) can be enhanced if forecasting models are optimized for their value in the application instead of accuracy. We combine these two aspects in order to maximize the potential of forecasting models: 1) confidentiality preserving data sharing and 2) a forecasting model optimization based both on accuracy and value. We focus on the EV charging stations' demand forecasting.We will focus on tree-based models for a combined approach for the following reasons. First, as shown in real-world applications, trees with their non-parametric design are suited to complex tasks with unclear relationships. For vertical federated learning, gradient-boosted trees fit the design naturally. For value-oriented approaches, single-decision trees or Random Forests are more favored due to their natural fit. In previous works, we developed a privacy-preserving multivariate tree model. We propose its adaption with a loss function surrogate of SPO+. Additionally, we propose adapting SPO-trees to a multivariate data-sharing setting. We evaluate both approaches and apply secret-sharing to the better-performing approach. Secret-sharing is a lossless, efficient encryption technique.We benchmark the models on two European EV charging station data sets, Dundee and Paris. We evaluate the revenue of the single-price market mechanism, where each charging station owner acts as a buyer. We use a symmetric approach as if an owner was a seller. We evaluate the benefits of data-sharing vs. a non-data-sharing application.The results indicate that the combination of approaches outperforms each individual approach and higher efficiency.

Decision-Focused Forecast Combination

Presenter: Akylas Stratigakos
Co-authors: Akylas Stratigakos;Salvador Pineda;Juan Miguel Morales

Advanced data-driven methods, such as machine learning, time series forecasting, and optimization, hold significant potential to improve decision-making under uncertainty across several industries. Moving from data to decisions is typically a two-step process, which involves a forecasting model estimating a conditional distribution of uncertainty and then solving a constrained optimization problem. In many industries, such as power systems, decision-makers often have access to multiple probabilistic forecasts for the same unknown quantity, such as renewable production. Combining forecasts from different experts has long been known to lead to increased forecast quality, as measured by scoring rules in the case of probabilistic forecasts. However, increased forecast quality does not always translate into lower decision costs in the downstream problem. This work proposes a novel decision-focused approach for probabilistic forecast combination that embeds the downstream optimization problem and explicitly minimizes decision costs. Specifically, we propose a linear pool of probabilistic forecasts where the respective weights are learned by minimizing the expected decision cost of the induced combination, which is formulated as a nested optimization problem. Two methods are proposed for its solution: a gradient-based approach that utilizes differential optimization layers and a cross-validation, performance-based weighting approach. For

experimental validation, we consider two integral problems associated with renewable energy integration in modern power systems and compare them with well-established combination methods. Specifically, we consider the problem of an aggregator participating in an electricity market participation under stochastic solar production and a grid scheduling problem under stochastic wind production. Our results highlight that higher forecast quality, as measured by scoring rules, does not always translate into better decisions. Notably, a combination of decision cost and standard scoring rules consistently leads to better decisions while also maintaining high forecast quality.

Decision-Focused Forecasting In Real-Time Electricity Markets

Presenter: Jean-François Toubeau
Co-authors: Jean-François Toubeau;Ruben Smets

In the field of energy systems, many decision-makers are confronted with uncertainty surrounding parameters like energy prices, load and renewable energy generation. This problem is of great importance for market actors whose profitability directly depends on prices differences across various periods of the day. In particular, a lot of value can be generated in real-time balancing markets by compensating the system imbalance between load and generation, but such a strategy is very risky due to the high volatility of real-time prices. To obtain those prices, we propose to move towards value-oriented forecasting, which incorporates the downstream (optimization) problem within the learning phase of the forecaster. Existing techniques, however, often rely on a specific implementation that is dependent on the downstream problem structure or limit the forecaster to an overly simplified model. In this work, we propose a universally applicable value-oriented methodology for training time series forecasters of any complexity. This is achieved by introducing a generalized loss function that enables the time series forecaster to capture variability, and using the downstream regret as the selection criterion in the hyperparameter tuning step. The proposed methodology is tested on a case study considering different types of energy storage systems participating in the Belgian balancing market. The method is benchmarked against other forecasting techniques including a neural network trained to minimize the mean squared error, and a forecaster relying on a fundamental market model. Using real-life data over a test set of two months, we show that the methodology outperforms those traditional techniques in terms of ex-post out-of-sample profit with 13% to 176%, depending on the specific type of storage asset.

Demand Forecasting In The Presence Of Supply Chain Disruptions

Presenter: Ritika Arora
Co-authors: Ritika Arora;Anna-Lena Sachs;Ivan Svetunkov;John E. Boylan

Recent events such as Brexit, the Russian-Ukrainian war, and the Covid-19 pandemic have highlighted how challenging it is for retailers to accurately forecast demand during disruptions and afterwards. Forecasting methods, which typically perform well under normal circumstances, may not provide accurate forecasts during disruptions because demand is distorted. As a result, decision makers often end up relying on their judgement in demand planning rather than applying statistical forecasting methods. However, accurate demand forecasting is essential for companies during and even after disruptions. We propose a shock-smoothing algorithm which is a modification of the single source of error state-space model underlying exponential smoothing (ETS) to include an additional component that captures the disruption. Using simulated and real data, we compare our approach with existing methods. Our results show that the proposed shock smoothing model is robust in many cases and is suitable for assisting demand planners in making data-driven decisions during disruptions and afterwards.

Inventory Control With Fifo And Lifo Picking Behavior: The Roles Of Sustainability Messages And Price Discounts

Presenter: Anna-Lena Sachs
Co-authors: Anna-Lena Sachs;Thomas Vogt;Ulrich Thonemann;Ben Lowery

Customer picking behaviour plays an important role in retail inventory management. Standard inventory models usually distinguish between picking the newest items, i.e. Last-in-first-out (LIFO), or the oldest items first, i.e., First-in-first-out (FIFO). We analyze how LIFO and FIFO picking behaviour affects inventory management in retailing, and whether sustainability messages and price discounts can change customer picking behaviour to increase sales of earlier expiring items and reduce food waste in retailing. We conducted an online experiment where subjects chose between buying FIFO and LIFO using monetary and non-monetary incentives. Understanding how different customer types respond to incentives helps retailers to offer them only to customers who most likely respond with buying expiring items. We evaluate the effect of these findings on inventories using a periodic review model for perishable items with age-dependent lifetimes.

Intermittent Or Not? How To Tell The Difference

Presenter: Anna Sroginis
Co-authors: Anna Sroginis;Ivan Svetunkov

Intermittent demand forecasting has been considered a challenging task for many years. The conventional models typically do not work as intended in such situations, and different ones need to be used to capture the tendencies in such data. Yet, the problem of classifying demand into intermittent or not has not been fully resolved. Practitioners tend to use simple heuristic rules, deciding that demand is intermittent if the number of zeroes in the data exceeds some threshold. Academic literature has not thoroughly studied the problem, assuming that any number of zeroes in the data already implies intermittence. However, reality is typically not as straightforward, and zeroes can appear in demand for a variety of reasons. In this presentation, we propose a new model-based classification scheme that separates demand into finer groups, acknowledging different possible structures in the data. We demonstrate how the proposed scheme can be efficiently used in a controlled environment and how it can be applied to a real dataset.

Combining Probabilistic Forecasts Of Intermittent Demand

Presenter: Yanfei Kang
Co-authors: Yanfei Kang;Shengjie Wang;Fotios Petropoulos

In recent decades, new methods and approaches have been developed for forecasting intermittent demand series. However, the majority of research has focused on point forecasting, with little exploration into probabilistic intermittent demand forecasting. This is despite the fact that probabilistic forecasting is crucial for effective decision-making under uncertainty and inventory management. Additionally, most literature on this topic has focused solely on forecasting performance and has overlooked the inventory implications, which are directly relevant to intermittent demand. To address these gaps, this study aims to construct probabilistic forecasting combinations for intermittent demand while considering both forecasting accuracy and inventory control utility in obtaining combinations and evaluating forecasts. Our empirical findings demonstrate that combinations perform better than individual approaches for forecasting intermittent demand, but there is a trade-off between forecasting and inventory performance.

Further Developments In Regression-Based Cross-Temporal Forecast Reconciliation

Presenter: Daniele Girolimetto
Co-authors: Daniele Girolimetto;Tommaso Di Fonzo

Forecast reconciliation is a post-forecasting approach to ensure the coherence of forecasts across a variety of constraints (usually linear, not just simple aggregation). It harmonizes individual predictions to meet predefined relationships, leading to a consistent and comprehensive picture. This can include ensuring power generation for different photovoltaic plants sum up to the Independent System Operator (ISO), or

guaranteeing some property (e.g. non negativity). By incorporating these constraints, reconciliation can also improve forecast accuracy by leveraging the individual strengths. In this talk, we address some open-issues related to the relationships between sequential, iterative, and optimal combination cross-temporal forecast reconciliation. We discuss the conditions under which a sequential (either first-cross-sectional-then-temporal, or first-temporal-then-cross-sectional) approach is equivalent to a fully (i.e., cross-temporally) coherent iterative heuristic. We also show that, for specific patterns of the error covariance matrix of the regression model on which the optimal combination approach grounds, iterative reconciliation "converges" to the optimal combination solution. The reduction of the computing effort is evaluated in an experiment on the SPDIS, an hourly photovoltaic power generation dataset, using the R package FoReco.

Dynamic Forecast Reconciliation At Scale

Presenter: Ross Hollyman
Co-authors: Ross Hollyman

We introduce a dynamic approach to probabilistic hierarchical forecasting at scale. Our model differs from the existing literature in this area in several important ways. Firstly we explicitly allow the weights allocated to the base forecasts in forming the combined, reconciled forecasts to vary over time. Secondly we drop the assumption, near ubiquitous in the literature, that in-sample base forecasts are appropriate for determining these weights, and use out of sample forecasts instead. Most existing probabilistic reconciliation approaches rely on time consuming sampling based techniques, and therefore do not scale well (or at all) to large data sets. We address this problem in two main ways, firstly by developing a closed from estimator of covariance structure appropriate to hierarchical forecasting problems, and secondly by decomposing large hierarchies in to components which can be reconciled separately.

Feature Based Graph Pruning For Improved Forecast Reconciliation

Presenter: Mitchell O'Hara-Wild
Co-authors: Mitchell O'hara-Wild;Rob Hyndman;George Athanasopoulos

Large collections of related time series often use attributes that identify their relation to other series. These attributes typically relate to what is being measured, such as product categories or store locations for the sales of a product over time. Aggregating across these attributes produces additional time series that offer useful overviews of the data. Forecast reconciliation uses these aggregations to improve forecasting accuracy. When there exists many attributes for time series data, the number of series in the collection quickly grows. This presents many problems for forecasting, since producing many forecasts can be computationally infeasible and the forecast accuracy for aggregated series of interest can worsen.To overcome these problems I propose using time series features to identify noisy, uninformative, or otherwise unwanted series and leveraging graph representations of aggregation constraints to safely remove them while preserving coherency structures. In this talk I will introduce pruning coherency constraints using graphs and demonstrate how graph search algorithms can be used to efficiently identify and remove uninformative subgraphs of time series while maintaining coherency. Pruning subgraphs of time series from the collection can substantially reduce the number of series to forecast, while retaining most of the information. This helps limit the computational complexity of forecasting, while improving forecast accuracy for aggregated series due to reduced model misspecification in more disaggregated series.

Optimal Forecast Reconciliation With Time Series Selection

Presenter: Xiaoqian Wang
Co-authors: Xiaoqian Wang;Rob Hyndman;Shanika Wickramasuriya

Forecast reconciliation ensures forecasts of time series in a hierarchy adhere to aggregation constraints, enabling aligned decision making. While forecast reconciliation can enhance overall accuracy in hierarchical or grouped structures, the most substantial improvements occur in series with initially poor-performing base

forecasts. Nevertheless, certain series may experience deteriorations in reconciled forecasts. In practical settings, series in a structure often exhibit poor base forecasts due to model misspecification or low forecastability. To prevent their negative impact, we propose two categories of forecast reconciliation methods that incorporate time series selection based on out-of-sample and in-sample information, respectively. These methods keep "poor" base forecasts unused in forming reconciled forecasts, while adjusting weights allocated to the remaining series accordingly when generating bottom-level reconciled forecasts. Additionally, our methods ameliorate disparities stemming from varied estimates of the base forecast error covariance matrix, alleviating challenges associated with estimator selection. Empirical evaluations through two simulation studies and applications using Australian labour force and domestic tourism data demonstrate improved forecast accuracy, particularly evident in higher aggregation levels, longer forecast horizons, and cases involving model misspecification.

Nowcasting Inflation Expectations With News.

Presenter: Thomas Chuffart
Co-authors: Thomas Chuffart;Cyril Dell'eva

The goal of this paper is to evaluate the informational content extracted from news articles about monetary economics. We propose an attention index constructed on french corpus of news concerning the European Central Bank (ECB) from Latent Dirichlet Allocation topi modeling. Our dataset includes six large French newspapers, for a total of over twenty thousand articles. We apply our measure to inflation expectations nowcasting. Our findings suggest that news help to nowcast inflation expectations. We check for robustness by giving to our index a polarity though sentiment analysis using well known deep learning method FinBert.

Forecasting Inflation: A Comparative Study Of Temporal Fusion Transformers And Traditional Machine Learning Models

Presenter: Ayesha Patnaik
Co-authors: Ayesha Patnaik;Yuri Lawryshyn

Forecasting inflation has been a challenging task due to the dynamic nature of the market, and yet it is pivotal in shaping economic policies. Forecasts allow central banks to formulate monetary policies and help businesses and investment firms in mitigating risk. Historically, inflation has been forecasted using econometric analysis, time series analysis and analyzing macroeconomic indicators. Recently, researchers have employed machine learning models for time series analysis among which Random Forest has proven to outperform most traditional models. Subsequently, the focus has shifted to evaluating deep learning models such as Long Short-Term Memory (LSTM) models. However, the vanishing gradient problem in LSTM raises doubts of their suitability for longer horizons of data. To bridge this gap, transformers have demonstrated potential in managing longer data horizons due to their distinct "self-attention" mechanism. Currently, Temporal Fusion Transformers (TFTs) are the state-of-the-art attention-based models which show promise in multi-variate and multi-horizon forecasting. This study aims to utilize TFTs for forecasting inflation indicators for the US, namely Consumer Price Index (CPI) and Personal Consumption Expenditure (PCE). Our preliminary results suggest that TFTs demonstrate strong predictive power for long horizons. Experiments are being carried out to compare the performance of TFTs against several models such as simple exponential smoothing, autoregressive (AR), autoregressive integrated moving average (ARIMA), XGBoost, random forest, and LSTM. The experiments also include hyperparameter tuning, a part of which will be the evaluation of different training methodologies, such as the rolling window and expanding window. Furthermore, the study will aim to determine the ideal historical data depth for predictive accuracy, particularly in the context of inflation forecasting.

Econometric Forecasting Using Ubiquitous News Text: Text-Enhanced Factor Model

Presenter: Beomseok Seo
Co-authors: Beomseok Seo

The use of news text as a novel source for econometric forecasting is gaining increasing attention. This paper revisited the way of incorporating narrative information into econometric forecasting by effectively quantifying sector-specific textual information without requiring training data. We exploit Theme Frequency Indices(TFI) utilizing domain-specific subject-predicate patterns to gauge public's perception of the economy. TFIs of 15 sectors, including production, inflation, employment, capital investment, stock and house prices, and others, were examined and integrated into Text-enhanced Factor Model(TFM) using latent factor structures. Empirical analysis, based on over 18 million news articles in Korea, reveals that TFM enhances the accuracy of near-term GDP forecasts, demonstrating simple text-mining techniques along with domain knowledge are capable to leverage qualitative information without costly training. The proposed method is applicable to a wide range of subjects for utilizing narrative information of the economy, offering a rapid and cost-effective approach.

Maximum Entropy Bootstrap Structure Determination And Parameters' Estimation For Exponential Smoothing Models

Presenter: Livio Fenga
Co-authors: Livio Fenga;Luca Biazzo

Widely used in the field of time series analysis for a variety of tasks (e.g., forecasting and simulation), Exponential Smoothing models are recognized as a powerful tool adopted in many contexts (applied research, official bureaus, public and private companies) by various actors (e.g., statisticians, econometricians, and practitioners). Regardless of the purpose Exponential Smoothing models are built for, their usefulness greatly depends on the accuracy of their parameter estimates. The related inference and structure determination procedures are, in many instances, carried out under the Maximum Likelihood paradigm, which, unfortunately, can be heavily impacted by different sources of errors induced by bias components and uncertainty. This paper outlines a computer-intensive procedure aimed at attenuating the effects of such errors. The proposed approach, based on a bootstrap scheme of the Maximum Entropy type (Vinod et al., 2009), will be theoretically discussed and empirically evaluated within the MAICE (Minimum Akaike Information Criterion Expectation) framework, using the 366 monthly time series from the "M3 2010 Tourism Forecasting Competition" dataset (Athanasopoulos et al., 2011) – freely and publicly available in the R® package Tcomp.

Fast Gibbs Sampling For The Local And Global Trend Bayesian Exponential Smoothing Model

Presenter: Xueying Long
Co-authors: Xueying Long;Daniel Schmidt;Christoph Bergmeir

In Smyl et al. [Local and global trend Bayesian exponential smoothing models. arXiv preprint arXiv:2309.13950, 2023.], a generalised exponential smoothing model was proposed which is able to capture strong trends and volatility in time series. This method achieves state-of-the-art performance in many forecasting tasks, but its fitting procedure, which is based on the NUTS sampler, is very computationally expensive. In this work, we propose several modifications to the original model, as well as a bespoke Gibbs sampler; these changes improve sampling time by an order of magnitude, thus rendering the model much more practically relevant. The new sampler is evaluated on the M3 dataset and demonstrates competitive accuracy, and superiority in terms of time complexity, in comparison with the original implementation.

Forecast Congruence, Accuracy, And Shrinkage Estimators For Exponential Smoothing

Presenter: Kandrika Pritularga
Co-authors: Kandrika Pritularga

Exponential smoothing is widely used in practice because it is easy to use, intuitive, and widely used in practice and academics. However, due to limited sample sizes, the estimated parameters become inefficient and harm the forecast accuracy. Shrinkage estimators are introduced into the model, and it shows improvement in forecast accuracy. This approach results in a specific forecast trajectory, called forecast congruence. Forecasts are congruent if the point forecasts on a specific date from different origins have low variability, meaning that the forecasts are less 'jittery' across origins. A study finds that congruence is useful to stabilise the variability of order sizes in inventory management. It is a complement feature of 'good' forecasts, for example, if we have two sets of accurate forecasts, we choose the most congruent ones.The previous study used a shrinkage parameter to control the level of congruence and it could potentially result in overly congruent forecasts, which harm the forecast accuracy. This study revisits the current approach to finding the 'optimal' shrinkage parameter by understanding the relationships between accuracy, congruence, and a shrinkage parameter. The simulation study shows that the relationship between congruence and accuracy is controlled by the shrinkage parameter and its value depends on the level of integration of the time series. If a time series is non-stationary, the relationship is convex, and it is easy to find the optimal shrinkage parameter. If it is stationary, a smaller parameter suffices to achieve accuracy and congruence. Based on these findings, this study proposes various bi-objective loss functions that combine parameter shrinkage, congruence, and accuracy and test the forecasting performance of each loss function with simulated and real data.

A Sequential Monte Carlo Approach To Adaptive Exponential Smoothing

Presenter: Alisa Yusupova
Co-authors: Alisa Yusupova;Nicos Pavlidis

Exponential smoothing (ES) remains one of the "workhorses" of business forecasting. ES models used in practise assume that the characteristics of the time series are constant over time. This is not necessarily true however, and the disruptions caused by the COVID pandemic and the more recent war in Ukraine have caused a renewed interest in developing models that "optimally" adapt to events like persistent or transient changes, while making best use of the available historical data. The present paper aims to develop principled probabilistic frameworks to accommodate a dynamic data generating process.Broadly speaking two types of approaches have been proposed to accommodate time series whose characteristics vary over time. The first is agnostic as to the type of change that occurs, and focuses on rendering the smoothing parameters adaptive over time. Notable works in this approach include Trigg and Leach (1967); Snyder (1988); Pantazopoulos and Pappis (1996), and Taylor (2004). All of these works propose heuristics to sequentially tune the smoothing parameter for the level, alpha, based on past forecast errors. None of these approaches proposes a probabilistic model for the evolution of the smoothing parameter(s), or offers interpretation as to the type of change that occurs. Insights into the latter can only be indirectly inferred by observing the evolution of alpha, but the extent to which this is informative is not clear.The second approach due to Koehler et al. (2012) relies on the state space formulation of ES (Hyndman et al., 2008). Instead of adapting the smoothing parameters it accommodates changing time series dynamics by introducing "events". In our view this approach is clearly superior to designing rules to adapt alpha, because each event is characterised, and the impact of events is quantified in a principled manner. However, because no probabilistic model for the occurrence of events is proposed, events can only occur in the in-sample period. We propose an approach that attempts to alleviate these shortcomings by utilising ideas from sequential Monte Carlo (SMC) with the ES framework. The approach we propose allows to constantly update the estimated model and favours models that can explain better the observed data. It combines the benefits of using maximum likelihood to obtain estimates of the smoothing parameters and the benefits of MC to infer quantities like the probability of an event at a given time point.

Can Quality Of Marketer-Generated Videos Enhance Tourism Demand Prediction? A Multimodal Deep Learning Framework

Presenter: Li Xin
Co-authors: Xin Li;Chengyuan Zhang;Shouyang Wang

The burgeoning volume of online content, particularly on social media platforms, has significantly influenced consumer behavior across various sectors, including the tourism industry. This study is anchored in the context of the increasing importance of multimodality in online-generated content, which offers new avenues for predictive analytics in tourism demand forecasting. Drawing upon dual-coding theory, this research highlights the unique potential of video content in tourism demand forecasting. Videos, with their dynamic and comprehensive portrayal of temporal and contextual details, engage the nonverbal system more effectively than static images or text, offering richer insights for predictive analysis.Despite the acknowledged influence of user-generated content on consumer decisions, a critical gap remains in understanding the specific role and predictive power of high-quality marketer-generated contents (e.g., videos) within tourism demand forecasting. Moreover, existing literature has predominantly focused on numerical, textual, and image data, thereby overlooking the potential of video content in tourism forecasting. Employing a multimodal deep learning framework, this study assesses the quality of marketer-generated videos on the Douyin platform, extracts the predictive features based on technical and aesthetic perspectives, and further evaluates the predictive power of online contents generated from both attractions and consumers. Our findings demonstrate that the quality of marketer-generated videos significantly bolsters the accuracy of tourism demand forecasts in predicting daily arrivals to the selected tourist attractions in Beijing. The study also highlights the aesthetic features outperform technical features in terms of predictive capability. This study not only bridges the theoretical gap by applying dual-coding theory to tourism marketing but also introduces methodological advancements through the application of multimodal deep learning techniques.The implications of this research are twofold. Theoretically, it enriches the understanding of the interplay between different types of online content and their impact on tourism demand forecasting. Practically, it offers valuable insights for tourism marketers on the strategic use of high-quality video content to improve predictive analytics and decision-making processes. Ultimately, this study paves the way for further exploration of multimodal data analysis in enhancing the precis

Measuring And Tracing The Tail Risk Of Tourism Economy Growth With Mixed-Frequency Model

Presenter: Han Liu
Co-authors: Han Liu;Xinyu Guo

As a cornerstone of Hong Kong's economy, the tourism industry plays a vital role in its economic prosperity and social stability. However, it faces significant vulnerabilities to external shocks, underscoring the importance of understanding and managing its tail risks. This study innovates by applying the G@R framework, traditionally used in macroeconomic risk analysis, to the tourism sector, creating a tailored TG@R framework. By incorporating both high-frequency and traditional data sources, including search query data, and employing Quantile Regression (QR) and Mixed Frequency Data Sampling Quantile Regression (MIDAS-QR) models, the research provides a nuanced analysis of the tail risks from nine major tourist origin countries to Hong Kong. Preliminary findings suggest that the mixed-frequency approach enhances measurement accuracy and robustness, offering a more scientific basis for decision-making. The study reveals the heterogeneity in risk factors across different stages of tourism economic cycles, with political and safety concerns exacerbating risks during downturns. This pioneering work not only extends the applicability of the G@R framework to tourism but also highlights the significance of diverse data sources in assessing risk, offering invaluable insights for policymakers to bolster Hong Kong's tourism resilience. Despite limitations related to data uncertainty and scope, the research sets a foundation for further exploration into the micro-level factors affecting tourism's economic risk profile.

Forecasting Seasonal Time Series Using Random Forests

Presenter: Karsten Reichold
Co-authors: Karsten Reichold

This paper shows that random forests are a promising approach to forecast (detrended) time series with strong seasonal patterns. Focusing on a recursive forecasting scheme for different forecasting horizons, a carefully conducted simulation study reveals that random forests using lags of the time series to be forecasted as predictors are able to capture complex and potentially time varying seasonal patterns and allow to handle anomalies – like, e.g., crisis periods – in the training sample. Furthermore, we show that leveraging knowledge from calendar variables, i.e., deterministic variables which are known in advance for the whole forecasting horizon, further improves forecasting performance of random forests, especially in case the number of lags included as predictors is small or the forecasting horizon is large. Finally, an empirical application to monthly tourism demand in Austria demonstrates the usefulness of random forests to forecast seasonal time series in practice and quantifies the contributions of different groups of predictor variables on tourism demand forecasts by means of Shapley values.

Passenger Flow Prediction In Urban Public Transit Systems For Effective Operations Management

Presenter: Sushil Punia
Co-authors: Sushil Punia

This work focuses on medium-term and short-term ridership forecasting in public transit systems (PTSs). Typically, in PTSs, econometric models have been used to estimate ridership (or demand) for long-term decisions (5-15 years). Then these models are adapted to generate short-to-medium-term forecasts. These models are too complex to adapt and handle for effective operational decisions. In recent years, researchers and practitioners alike have been working with statistical forecasting models for PTSs. However, things are still in the initial stages, and available models are inadequate to inform complex operations decisions. In this context, this work proposes to develop a decision support system for ridership forecasting in PTSs. The proposed ridership forecasting model consists of time series, machine (and deep) learning, and other methods. The proposed decision support system incorporates cross-sectional and temporal hierarchical information from data to generate coherent forecasts for effective operational decision-making at various levels in PTSs.For the conference presentation, an empirical study part of the proposed ridership forecasting models will be presented. The study uses data from the Istanbul urban PTS that consists of hourly data for a one-year duration, i.e., 8760 observations per time series from 01 January 2022 to 31 December 2022, for all passengers boarding on different routes on different transit modes. Further, parameters like passenger boarding are categorized into normal and transfer passenger categories. Considering the combinations of transit types, transit modes, and types of passengers, the study consist of a set of 1837 time series. The times-series, machine (and deep) learning methods were applied to the data and performance of these methods has been evaluated using mean error, mean absolute errors and root mean squared errors for one-day, one-week and one-month ahead planning horizons. In short term forecasting, STLF is outperforming the other methods but for longer planning horizons in the study, the LSTM is working better than other methods. *The work is part of the project supported by the "International Institute of Forecasters and SAS® Award" to support research on forecasting in practice/management categories for the year 2022.

Forecasting Using Social Media Data: Quant-Qual Model Of Container Throughput Forecasting

Presenter: Sonali Shankar
Co-authors: Sonali Shankar

In the recent past, social media data has been incorporated into the forecasting models to improve

the forecasting accuracy or to determine the additional explanatory variable for the forecasting model. In addition, it was observed that business-to-business (B2B) companies were not very keen on their social media presence. Whereas, there is a change in social media usage trajectory observed in B2B companies during and post Covid19 era. Despite increasing social media usage by these companies, there is limited literature available on how social media data can be incorporated by B2B companies in decision making - demand forecasting. In this study, container supply chain is used as the field of study to forecast container throughput at the port of Singapore (PSA). PSA is one of the busiest ports in the world. In the literature, econometric and machine learning forecasting methods have been applied for container throughput along with some hybrid models. However, there is a lack of literature harnessing the power of social media to forecasting techniques and models. This research is an attempt to bridge this gap and provide social media and machine learning based container throughput forecasting using the proposed novel quantitative-qualitative (Quant-Qual) forecasting model. Social media data is divided into four categories: demographic information, engagement matrix, trend, and sentiments of the actors/users in a network at a given location of the destination port. The forecasting is conducted using this data and machine learning methods. The proposed model is benchmarked with widely used econometric and machine learning models using error metrics. This study addressed two research questions: How can social media data help in improving the forecasting performance of the container throughput? and; to what extent can a mixed method forecasting model be applied to improve the forecasting performance of the container throughput?

Data Driven Approaches For Predicting Air Cargo Turnaround Times

Presenter: Sarah Van Der Auweraer
Co-authors: Sarah Van Der Auweraer;Daniel Dobos;Sandria Ruessmann;Anne Lange

Cargo airlines typically operate on a tight schedule to maximize flight hours and thereby aim to minimize the time spent on ground, which is also known as the turnaround time. The turnaround process comprises all operations to prepare the aircraft for its next flight. When the actual turnaround time deviates substantially from the time originally scheduled, the consequences can be severe; the departure of the aircraft may be delayed, and this delay can further propagate to an entire route of consecutive journeys with the same aircraft, resources will be used inefficiently, and operational costs can increase. Creating prediction models to forecast the aircraft turnaround time accurately can help to improve decision making, thus staying within the crew duty time limits as well as adhering to airport timeslots. In the case of passenger aircraft, the turnaround process is well studied in academic literature. However, despite its importance, air cargo turnarounds have thus far received much less attention. The aim of our research is therefore to develop a better understanding of cargo aircraft turnaround times and to explore the potential use of data driven approaches to predict the duration of future turnarounds. This research is carried out in close collaboration with Cargolux Airlines International, a leading all-cargo carrier based in Luxembourg. The company provided a data set of approximately 75,000 turnarounds performed at various stations throughout the carrier's network from 2019 to 2022. This context-specific data is combined with factors known from the passenger airlines literature, such as airport congestion and weather data. Using this data, we test different Machine Learning and Deep Learning regression methods and provide an estimate of turnround times for airfreight logistics.

Algorithm Aversion In Time Series Forecasting: Effects Of Outcome Feedback And Guidance

Presenter: Nigel Harvey
Co-authors: Nigel Harvey;Shari De Baets

Most studies of algorithm aversion have focussed on cue-based forecasting where the future value of one variable is predicted from a previous value of each of a number of different variables. Given the choice of using an algorithmic or a judgmental forecast, people prefer the algorithmic forecast before receiving outcome feedback (algorithm appreciation) but prefer the judgmental forecast after doing so (algorithm aversion). It is assumed that this is because people are less tolerant of errors made by algorithms than of

those made by human beings. In time series forecasting, the opposite pattern of results has been reported. People are more influenced by advice from a human judge before feedback is given (algorithm aversion) but by advice from an algorithm after it is given (algorithm appreciation). Our experiments suggest that this reversal arises because a preference paradigm has been used in studies of cue-based forecasting but an advice-taking paradigm has been used in studies of time-series forecasting. When a preference paradigm is used in time series forecasting, the same pattern of results previously reported for cue-based forecasting is obtained.

Title Consensus Forecasting With Large Language Models

Presenter: Oliver Schaer
Co-authors: Oliver Schaer;Matt Schneider;Nikoalos Kourentzes

Demand forecasting is typically based on a combination of model-based forecasts and judgemental interventions, where analysts incorporate contextual information that is otherwise absent or accounts for known deficiencies of the model. However, often with such contextual information, there might be inherent bias in how some types of information are communicated and recorded (e.g., "this is a very successful product"). We propose LLMs to update the original contextual information, eliciting different responses from the analysts to improve judgemental adjustments.

Enhancing Accuracy In Business Forecasting With Behavioural And Decision Science

Presenter: Francesca Tamma
Co-authors: Francesca Tamma

Whether predicting the value of an investment in a new technology to enhance an existing business model, or the long-term strategic need of specific skillsets in a firm, accurate forecasting is key to many businesses. Supported by the digitalisation of operations and evolving data science approaches, predictions for core business processes are increasingly driven by sophisticated quantitative forecasting models. However, in contexts of high uncertainty and complexity, research has shown human expert judgement can still outperform quantitative forecasting models, or offer a valuable complementary perspective. Behavioural economics studies suggest, however, that the quality of expert judgement can be undermined by cognitive biases and inconsistencies (noise). While several methodologies, including 'estimate-talk-estimate' approaches (using the widely recognised Delphi approach), have been tested for their contributions to enhancing forecasting accuracy, applications by businesses to their actual forecasts have been more limited to date. Here, I will present the evaluation of the adoption of one such methodology by an international financial services firm. The results include: noise audits, online experiments, and sentiment analysis to assess the prevalence of cognitive biases and noise in key business forecasting activities. Informed by these diagnostics, a structured approach to expert judgement, specifically aimed at enhancing the business decision makers' ability to better navigate cognitive biases and minimise inconsistency, was devised and rolled out by the financial services firm. An econometric evaluation finds supportive evidence that gains of 7% in forecasting accuracy, as measured by the mean absolute percentage error, are possible using these techniques. Limitations to the internal validity relate to the nature of the business roll-out of this methodology, which was not designed as a randomised control trial, and robustness checks are presented showing consistent results. Lastly, implications for the concrete implementation of best practices into workflows, processes and procedures are discussed.In conclusion, my research suggests valuable improvements in forecasting accuracy are possible through best practice approaches for enhancing expert judgement, helping to improve business expert judgement accuracy.

A Triumph Of Experience Over Hope? An Experiment To Test A New Theory Of Forecast Revision In Groups.

Presenter: Fergus Bolger
Co-authors: Fergus Bolger;Iain Hamlin;Gene Rowe

Different group-judgment elicitation protocols have been proposed to reap benefits from groups' increased knowledge, while minimizing individual and group biases. These protocols are being applied to important decision problems where hard data is scarce, but the findings of the few existing validation studies have been weak and contradictory. One reason for these inconclusive findings is that key features of tasks and experts have neither been defined nor controlled. In addition, theoretical consideration of how judgments might improve in a group setting is scarce. 128 online participants were trained through 40 practice trials where they made forecasts using three leading indicators of economic growth for timeseries with known cue-criterion relationships – outcome feedback was given after each forecast. Individual participant linear-regression judgment models were then elicited from forecasts on 20 further trials without outcome feedback. True levels of expertise and task difficulty could therefore be determined.In a set-up where experts first make a judgment independently, receive feedback from other experts, then potentially revise their judgment (e.g., Judge-Advisor Systems, or the Delphi technique) we propose that the degree of revision is dependent on the experts' confidence in their own initial judgment and the strength of the 'pull' away from the initial judgment towards cues in the feedback to the location of the 'truth'. To test this 'Pull Theory' – and hypotheses about the relative pull of cues – after training, we asked participants to make GDP forecasts individually and express confidence in them. Next participants received one of three types of feedback from four other forecasters: the median of the forecasts; each judge's confidence in their forecast; a brief rationale for forecasts. After the feedback participants were invited to revise their initial forecasts and restate their confidence. To manipulate expertise in the final test phase, half the 20 sampled stimulus series were from the same environment as the training set, and half were from an environment with different relationships between variables. All stimulus series were simulated, but were representative of distinct periods of economic activity in the UK during the last 20 years, thereby maintaining a degree of ecological validity. Results will be discussed.

Forecasting Demand For New Contraceptives In Low Income Countries

Presenter: Fred Church
Co-authors: Fred Church

Studies have shown that many women in low and lower-middle income countries do not use contraception. Unintended pregnancies are common, and can result in negative health and economic outcomes. New technologies offer women longer-lasting and more convenient options which have the potential to increase contraceptive usage.The Bill & Melinda Gates Foundation funds the development of new contraceptive technologies. A number of new contraceptive products are in the pipeline for potential further investment. The foundation needed advanced demand forecasting to predict user demand for each new product so that the strongest candidates for development can be prioritized.With funding from the foundation, Format Analytics used quantitative survey research and advanced modeling to generate pre-launch demand forecasts for seven new contraceptive products in multiple low and lower-middle income countries: Kenya, Nigeria, Ethiopia, Burkina Faso, Zambia, and Pakistan. The seven new products included a 6-month injection, a biodegradable implant, a monthly pill, a microneedle patch, an on-demand pill, a hormonal IUD, and a vaginal ring.Format Analytics partnered with local fieldwork agencies in each country, maximizing knowledge among researchers who best understand the conditions in their own part of the world. Survey data was collected among approximately 6000 geographically-dispersed women and 1800 healthcare providers in each country. In the survey interviews, respondents were presented with profiles of each of the seven new products, and stated their usage intent by answering standardized questions. Importantly, the demand models use validated algorithms to adjust for the overstatement of usage interest that is inherent in all survey research. This enables the models to accurately predict true demand.Forecasts were delivered in the form of a powerful interactive simulator which enables the variation of important demand factors such

as product features, launch year, and levels of market access. Variations in product features were quantified in the survey research via a discrete choice experiment.Presentation of the results of this research will include a summary of the products that achieved the strongest demand forecasts, comparison of findings between the six different countries, and a demonstration of the interactive demand forecasting simulator.

Life Expectancy Recovery By Sex And Age Groups After Selected Catastrophic Events

Presenter: NA
Co-authors: Eliud Silva;José Manuel Aburto

We aim to estimate the time of life expectancy (e_0) recovery after mortality crises and to quantify which age groups drive observed post-crises trends. We focused on major European pandemics and wars during the 19th and 20th centuries. Data was retrieved from the Human Mortality Database (HMD). Regarding descriptive statistics across time, we identify the largest ten losses in e_0. Afterward, auto ARIMA's models for every selected case are used to forecast e_o had the crisis did not happen and we analyze the time it takes to recovery. The events are divided into Pandemics and Non-pandemics, and several statistical tests are carried out. World Wars were the events that caused the largest losses. Statistical terms show no significant differences by kind of event and sex. Finally, children are the primary age group contributing to recovering life expectancies.

Moving Aggregate Modified Autoregressive Copula-Based Time Series Models (Magmar-Copulas)

Presenter: Sven Pappert
Co-authors: Sven Pappert

We introduce the Moving Aggregate (MAG) modified Autoregressive (AR) copula-based time series model (MAGMAR-Copula) and employ it for probabilistic forecasting of (univariate) time series with various (different) types of temporal dependencies. In the MAGMAR-Copula model, a MAG-Copula is introduced to the copula-based time series model equation, enabling the model to incorporate persistency. This approach is similar to the introduction of the moving average to autoregressive time series. The model can properly adapt to non-linear, as well as long-term temporal dependencies of a time series while allowing for flexible marginal modeling. If both the AR and the MAG-copula are the Gaussian copula and the marginal distribution is the normal distribution, the classical (linear) ARMA model is be recovered. Hence, the MAGMAR-Copula model is not only a generalization of the classical copula-based time series model, but also of the ARMA model. In this work, we investigate the probabilistic forecasting performance of the MAGMAR-Copula model. We perform probabilistic forecasting studies for different time series, exhibiting different types of temporal dependencies. In particular, we forecast US inflation, energy commodities' prices (natural gas, oil, coal and European allowances), as well as wind turbine power production. These time series have different characteristics. While e.g. quarterly US inflation has a relatively strong first moment autocorrelation and is dominated by an asymmetric temporal dependence structure, energy prices have stronger second moment autocorrelation and are dominated by a heavy-tailed temporal dependence. We employ different types of copulas and marginal distributions to properly account for these characteristics. The predictive performance of the model is compared with benchmarks from the literature. The probabilistic forecasts are evaluated by the CRPS and by extracting confidence intervals from the probabilistic forecasts and measuring width, coverage rate and interval score.

Analyzing 167 Years Of Technological Convergence In Oecd Total Actor Productivity: A Semiparametric Heterogeneous Panel Model Approach

Presenter: Param Silvapulle
Co-authors: Param Silvapulle;Sium Hannadige

Recently, several studies have utilized the additive fixed effect (AFE) panel model and its extensions and showed that knowledge transfer, research and development (R&D), and human capital are the key drivers that explain cross-country differences in total factor productivity (TFP). In this paper, we introduce an innovative semiparametric panel model for the TFP of 21 OECD countries, employing 167 years of data spanning 1850-2016. The proposed model has three attractive features: (i) it accounts for cross-section dependence; (ii) it allows for heterogeneity in the model parameters; and (iii) it includes country-specific technological advances that are assumed to be unknown nonlinear functions of time. Moreover, the model includes domestic and foreign R&D capital stocks and education attainment (Edu) as determinants of TFP. We find evidence that technologies exhibit a positive trend for all 21 countries and σ-convergence. Domestic and foreign R&D have positive and significant impacts, whereas education has a negative effect on TFP. In the pre-war (1950) period, the impacts of domestic R&D and Edu were positive, and that of R&D spillovers was insignificant. The reverse was true for the post-war period. The spill-over R&D is larger in G7 than in other OECD countries. A noteworthy result is the σ-convergence of technologies of 21 countries, indicating technology catch-up due to significant (direct & indirect) knowledge and technology transfers among the 21 OECD nations.

MOMENTUM IS STILL THERE CONDITIONAL ON VOLATILITY-AMPLIFIED PESSIMISM

Presenter: Jack Strauss
Co-authors: Jack Strauss;Soroush Ghazi;Mark Schneider

We introduce an index of volatility-amplified pessimism (VAP) constructed from a representative agent asset pricing model with probability weighting. The model predicts momentum returns decrease in market volatility and pessimism, and predicts the opposite for the equity premium. Our real-time trading strategy uses the index to switch across different VAP states, and generates a large spread between momentum and market returns. In contrast, other momentum strategies have recently disappeared net of transaction costs. We find that momentum is still there, conditional on the interaction between the representative agent's pessimism and market volatility.

Forecasting With Random Arima: A Machine Learning Perspective

Presenter: NA
Co-authors: Devon Barrow;Nikoloas Kourentzes;Yves Sagaert;Ivan Svetunkov

ARIMA has been used in forecasting for decades in many contexts. The original methodology for order selection was developed by Box & Jenkins in the 70s and was considered as a gold standard for many years. This methodology was shown to work poorly in practice because of the complexity of the original model and the large variety of potential models that might be appropriate for any given data. Over the years, it has become clear that automated procedures for order selection, like one developed by Hyndman & Khandakar (2006), do better than the Box-Jenkins methodology, employing ACF/PACF analysis. Yet the question of order selection still has not been resolved entirely because of the huge parameter space that ARIMA can have, especially on seasonal data. In this presentation, we take a different view on ARIMA modelling and use Machine Learning principles and techniques to overcome these challenges. Instead of trying to find the best appropriate model for the data, we revert to a combination of ARIMAs, randomly selecting the orders based on a set of rules. We show how the approach works in principle, conduct a simulation experiment to demonstrate how it performs in a controlled environment and then provide a demonstration on a real time series. We demonstrate the gains in accuracy for the Random ARIMA in comparison with the conventional approaches.

Foundational Models For Time Series, A Comprehensive Evaluation.

Presenter: Max Mergenthaler Canseco
Co-authors: Max Mergenthaler Canseco;Cristian Challu;Azul Garza

The popularity of large language models (LLMs) has contributed to the interest in developing "foundation models" for time series. Traditionally, deep learning pipelines for time series forecasting operated within a single domain or data set, limiting the ability to leverage the potential of large pre-trained models. The concept of universal or foundational forecasting, which consists of pre-training a model on a diverse collection of datasets and transferring it to unseen domains, has recently gained popularity. The field has witnessed the proposal of numerous approaches with various architectures and sometimes opposing principles, ranging from pure temporal modules to reprograming existing LLMs. We identify the need to create a systematic method for benchmarking time series foundation models and other statistical and deep learning approaches. In this paper, we present a benchmark dataset for forecasting and anomaly detection, consisting of more than 300,000 series, and compare open and closed-source foundational models. The results provide evidence of the advantages in terms of accuracy and computational efficiency of the various proposed principles but it also helps to identify current limitations and shortcomings of these novel approaches.

Stylised Facts Of Cryptocurrency Markets

Presenter: Nursultan Abdullaev
Co-authors: Nursultan Abdullaev;Rustam Ibragimov

In this paper, we present a detailed statistical and econometric analysis of the main stylised facts of cryptocurrency markets. Our study focuses on three key properties of cryptocurrency price and return time series: (i) heavy tails, indicating, in particular, that large price/return downfalls and fluctuations are more common than might be expected under a normal distribution; (ii) absence of autocorrelations, implying that return time series are to some extent are unpredictable and do not exhibit linear dependence over time; and (iii) volatility clustering, where periods of high volatility tend to be followed by similar periods and likewise for low volatility, implying nonlinear dependence in return time series. The presence of and inference on these stylised facts provide crucial insights for econometric modelling of the cryptocurrency market, and have important implications for market participants, risk management, and policy formulation.

Intervention Aware Forecasting For Process Control With Sparse Data

Presenter: Sumanta Mukherjee
Co-authors: Sumanta Mukherjee;Chandramouli Kamanchi;Pankaj Dayama;Vijay Ekambaram;Arindam Jati;Kameshwaran Sampath

In recent decades, there has been a significant rise in the adoption of data-driven predictive models in various industrial application scenarios, especially industrial process control. The dynamics of an industrial process depend on its current state, leading to different responses to the same intervention. It is well known that forecasting models are a common choice for state-dependent modeling. To accommodate the process control scenario, the predictive model should provision for control/exogenous variables (active forecasting). Model predictive process control (MPC) uses the output of the predictive model for the system response estimation. The control system decides the optimal action by varying the future values of exogenous/control variables. A gap in predictive model estimates adversely impacts the control decisions.In practice, process operations follow standard operating instructions, leading to biased (correlated) changes in the control variables. Moreover, in many real-world setups, evidence of control set point changes is sparse. As data-driven modeling relies on data diversity, the sparse and biased control evidence poses a challenge in training a reliable model that correctly captures the intervention effect.We propose to utilize qualitative responses to interventions that are known to the field expert, to overcome the lack of diverse intervention (control) evidence. This knowledge of qualitative responses can be represented as a causal graph, which is further exploited in the data augmentation process. In this presentation, we discuss a novel approach to infuse such expertise into the predictive model through guided training. The guidance uses data augmentation, wherein the expert knowledge is incorporated into the model via contrastive training. The trained model exhibits intervention response behavior that is inline with the guidance provided by the

expert, without compromising the point forecast accuracy.

Transforming Support To Displaced Populations: Predicting Movement Patterns Through Sequential Modelling

Presenter: William Low
Co-authors: William Low;Auke Tas

The escalation of violence in Nigeria since 2014 has resulted in mass displacement and deprivation, with more than 2.2m people currently displaced in the states of Adamawa, Borno and Yobe. New incidences of conflict-induced displacement are chronic, and humanitarian needs are significant and widespread. Children are disproportionately affected, with greater exposure to protection risks and long term impact on their health and education – particularly for girls. Save the Children provides aid to these populations through its country office and local partners. However, access constraints and overall poor quality data on displacement flows has historically limited the effectiveness of our responses. Save the Children is now seeking to remedy these key data gaps through its Predictive Displacement project. Partnering with the University of Virginia and Brunel University, we have established sequential models to predict the scale, demographics and geography of conflict-induced displacement. The first element is a machine learning model that predicts the amount of displacement that will occur as a consequence of conflict events, based on the attributes of past conflicts and historical first-time displacement patterns. The respective scale of displaced populations from conflict events is then used as an input to the principle model. This second model is an agent-based model that predicts the movement of displaced people across a pre-defined network. It takes as input a location graph that approximates the areas of interest, and a population of agents based on the conflicts that occur during the simulation period and the prior model. The movement of agents across the network is controlled by key decision rules based on quantitative and qualitative analysis of the country and regional context. The outputs of this model provides an indication of where the weight of new displacements will fall. This data then forms part of the humanitarian response planning process. Where sufficient data exists, demographic attributes of agents such as age and sex can be tracked through the simulation, allowing further disaggregation of displaced populations and better targeting of humanitarian aid. In this paper we summarise this process, showing how data is sourced and applied to the models, and how its predictions are processed and communicated for use by Save the Children's humanitarian responses.

Climate Induced Migration In The Western Hemisphere – Forecasting Encounters At The Darien Gap And Us Southern Border

Presenter: Logan Stundal
Co-authors: Logan Stundal;David Leblang

Global international migration has significantly increased in the post-Covid period, particularly south-to-north migration as northern economies have experienced greater post-pandemic economic growth. However, rates of irregular migration – migration through illegal or unauthorized pathways – has also greatly expanded during this period. Along the US southern border, customs and border protection agents have reported record numbers of irregular migrant encounters while similar patterns have emerged in the Darien – a remote expanse of territory in Panama along the Colombian border. In this paper we investigate drivers of irregular migration through these two key corridors with a particular emphasis on the role of climate change in migrant country of origin as a key driver to explain variation in monthly encounters. We then examine how encounters in the Darien Gap can serve as an early leading forecasting indicator for irregular migrant arrivals along the US southern border. To achieve this we fit two categories of models in the paper. First, we explore the relationship between climate drivers and encounters using panel and time-series regression models. Here we also assess how conflict and political violence interacts with climate drivers to exacerbate or mitigate the flows of migrants through these corridors. Following this we fit two classes of machine learning forecasting models. To forecast US southern border encounters using Darien encounters and climate drivers as leading indicators we fit a long short-term memory (LSTM) recurrent neural network model and assess predictive performance. Finally we examine routes migrants take from the Darien to

the US southern border by employing a spatio-temporal graph convolutional neural network (ST-CGN) along with a panel of migrant encounters disaggregated by migrant origin country and subnational Mexican administrative divisions. This fine-grained subnational spatial analysis allows us to examine how climate change and political violence influences migration routing preferences as well as leakage of irregular migrants terminating their migration early. We conclude by assessing the performance of Darien gap crossings to forecast US encounters as well as evaluate the overall effectiveness of using climate indicators to predict future irregular migration flows.

Optimizing an integrative forecasting approach to forecast unauthorized immigration

Presenter: Douglas Baals; Justin Schon
Co-authors: Douglas Baals;Justin Schon;Nadwa Mossaad

In this paper, we explore how judgment can be used to improve forecasts of unauthorized immigration into the United States. We compare the performance of a statistical model with a naïve average of judgmental adjustments to the statistical model and weighted averages based on forecaster experience or rationales. Our study evaluates forecasts made on a monthly basis from February 2023 to present. Results suggest that blending a weighted average of judgmental adjustments to the statistical model with the unadjusted statistical model yields the most accurate forecasts. Moreover, the simple average of judgmental adjustments can perform as well as or better than the statistical model alone, provided there are enough people contributing their judgment to the forecast effort. Judgmental adjustment is particularly valuable when unauthorized immigration levels change more sharply, due to the challenge of specifying a reliable statistical model that can forecast large changes. Lastly, this study highlights the reality that forecaster rationales may not provide useful criteria to weight judgmental adjustments if the human subject matter experts are unable to dedicate substantial time to the forecasting process.

Mecovma-Framework: Implementing Machine Learning Under Macroeconomic Volatility For Marketing Predictions

Presenter: Manuel Johannes Muth
Co-authors: Manuel Muth

In light of the current focus on Machine Learning (ML) for forecasting in management disciplines and the parallel challenge of the volatile market environment, structured procedural guidelines are required. The MECOVMA framework, a new methodological framework for implementing ML under macroeconomic volatility for marketing forecasts, is intended to provide such guidance. It addresses 4 key gaps in existing research frameworks: (1) individualization to the requirements of ML application for marketing-specific forecasting; (2) relevance, by responding to the current volatile environment and macroeconomic disruptions; (3) consolidation, necessitating the integration of isolated approaches into a coherent system; and (4) interdisciplinarity, to bridge the terminologies and criteria of different disciplines.The study follows a two-phase approach for framework development based on McMeekin (2020), starting with data synthesis to define the overarching process steps of the MECOVMA framework, followed by the substantive concretization of these steps. The first phase comprises the systematic analysis of existing frameworks and the evaluation of their relevance in relation to the three thematic focal points of the research question. The second phase integrates the findings from a systematic literature review into the defined process steps to shape the content of the framework with scientific evidence. Particular attention is paid, for example, to the selection and preparation of relevant data in a volatile macro-environment, the application and evaluation of ML forecasting models adequate in this context, and the effective implementation in marketing practice. The framework thus contributes to the methodological advancement of marketing forecasting and provides a structured approach to the application of ML – recognizing the prevailing business conditions.\subsection{The Dna Of 100% Customer Service}Presenter: Klaus Spicher
Co-authors: Klaus Spicher

The DNA of 100% Customer Service The (standard) ERP-based S&OP approach utilizes high quality market fore-casts as input for operational (inventory) planning. Generally, high Customer Service represents the target of the planning process. High forecast-quality is a key requirement. – The DNA-approach is based on planning necessary warehouse-entries (Supply Function) for meeting 100% Customer Service (CS) at a pre- determinable in-ventory level. The DNA-profile provides all inventory levels for achieving 100% service. So, Customer Service becomes the intended and planned result of the planning process. In other words, the DNA approach can be seen as a "reverse ERP-process" – not depending on high forecast quality.The DNA-profile also provides the selected Supply Function for achieving 100% service. So, the DNA-approach enables to decide, on which inventory level to achieve 100% service. – The three key features of the DNA-approach are: (1) the method works well with low levels of forecast quality and (2) no safety stock is needed and (3) the DNA represents a high-level benchmark for inventory level assessments. But the application on the shop floor without selected safety stock is for practi-cal risk reasons not recommended.AAS = Average Annual Stock.The DNA-Profile shows the inventory levels allow-ing for 100% Customer Service. The planners can decide about the intended inventory level (AAS) from the DNA-profile. The related Supply Function provides timely the warehouse entry quantities. – The Min(AAS) can be seen as Benchmark for inventories.Related Display not accepted In case of highly random-infected time series the forecast quality is limited. So, the input for planning does not meet the necessary requirements, resulting in high – at least in scarcely manageable – reasonable inventories levels. - The DNA-approach solves the problem. Actual research deals with the extension for CS $<$ 100%. Another field of research is the comparable application with Stand-ard ERP-forecasting-based solutions for identifying the "value" of the DNA ap-proach.

Causal Forecasting For Pricing

Presenter: Johannes Stephan
Co-authors: Johannes Stephan;Douglas Schultz;Julian Sieber;Trudie Yeh;Patrick Doupe;Tim Janushowski

In many practical applications time series forecasts feed into downstream decision problems. With this work, we consider the case of an online fashion retailer, where demand forecasts are consumed by an algorithm for price optimization. Thus, modeling the causal relationship between price and demand is crucial when it comes to setting prices in a profit optimal manner.To this end, we present a novel approach for demand forecasting that combines Double Machine Learning (DML) methodology for causal effect estimation with state-of-the-art forecasting models. In particular, we use the classical DML split into outcome-, treatment- and effect model, whereas each individual learner is a transformer-based neural network. Moreover, we deviate from the established use of DML as an effect estimation method, and also use it as a means of forecasting demand at different price points.In extensive empirical experiments, we show on the one hand, that the resulting method estimates the causal effect better in a fully controlled setting via synthetic, yet realistic data. On the other hand, we demonstrate on real-world data that our method outperforms other forecasting methods in off-policy settings (i.e., when there's a change in the pricing policy in the prediction horizon), while only slightly trailing in performance in the on-policy setting.

How To Publish In Foresight

Presenter: Michael Gilliland
Co-authors: Michael Gilliland

This session is for anyone interested in publishing in Foresight: The International Journal of Applied Forecasting, which is IIF's quarterly journal oriented toward forecasting practitioners. Foresight publishes several types of content including articles, commentaries, book and software reviews, interviews, tutorials, and opinion-editorial pieces. All content, including articles based on academic research, focus on the practical "takeaways" that forecasters can put to use in their daily roles. All submissions are reviewed by the editorial staff for topic relevance and are edited for clarity of presentation. This session will describe Foresight's mission and scope, the manuscript submission and editorial review process (including timelines), and the kinds of topics most relevant to our readers. Audience questions will beaddressed by Foresight's

EiC and members of the editorial staff.

Fuzzy Hierarchical Forecasting

Presenter: Olivier Sprangers
Co-authors: Olivier Sprangers

In hierarchical forecasting we aim to create forecasts for a set of time series that are aggregated according to a cross-sectional and/or temporal hierarchy. Commonly, the cross-sectional and/or temporal hierarchy is pre-defined and considered constant throughout the forecasting process. However, the hierarchy might be misspecified, for example in the case of a retail clothing product such as a pair of shorts that is mislabelled as belonging to the product group 'hats' instead of the product group 'shorts'. Also, the hierarchy might be uncertain or ambiguous, for example in the case of a retail clothing product such as a pair of shorts that might be considered part of both the running department and the football department. If these departments are at a similar level in the hierarchy, where do we allocate the product to when performing hierarchical forecasting? In this work, we investigate the impact of both hierarchy misspecification and hierarchy uncertainty when performing hierarchical forecasting in an end-to-end manner. First, we show how the forecasting performance deteriorates when we increase the degree of misspecification in the hierarchy. Then, we define the concept of Fuzzy Hierarchical Forecasting, in which each time series can be part of multiple nodes at the same level in the hierarchy, by relaxing the constraints of the cross-sectional or temporal aggregation matrix. We demonstrate how using fuzzy aggregation matrices can improve forecasting performance as compared to the misspecification setting. A surprising finding is that forecasting performance can be improved by using fuzzy aggregation matrices, even compared to the 'oracle' setting where we assume perfect knowledge of the 'true' hierarchy. The lesson for practitioners is that it may be useful to employ fuzzy aggregation matrices when performing end-to-end hierarchical forecasting.

Augmenting Hierarchical Time Series Through Clustering: Is There An Optimal Way For Forecasting?

Presenter: Bohan Zhang
Co-authors: Bohan Zhang;Anastasios Panagiotelis;Han Li

Forecast reconciliation has attracted significant research interest in recent years, with most studies relying on pre-defined hierarchies constructed with time series metadata. With the goal of improving forecast accuracy in mind, we extend and contribute to the emerging research on the clustering-based reconciliation method by proposing a novel framework for hierarchy construction. This framework offers three approaches: cluster hierarchies, random hierarchies, and combination hierarchies. Utilizing the proposed approaches, we investigate the individual contributions of two primary factors, namely "grouping" and "structure", to the performance of forecast reconciliation. Through a simulation study and experiments on two real-world datasets, we demonstrate the practical efficacy of different hierarchy construction approaches. Our findings provide new insights into the dynamics between "grouping" and "structure", which lead to an improved understanding of forecast reconciliation.

A combination perspective of temporal hierarchy forecasting methods: what are the properties of such forecasts?

Presenter: Nikolaos Kourentzes
Co-authors: Nikolaos Kourentzes;George Athanasopoulos

Temporal hierarchy forecasting (THieF) has gained some momentum in its use for improving forecast accuracy in the literature. By contrasting their properties to alternative methods that rely on the use of multiple temporal aggregation levels, we show how the modelling problem of temporal hierarchies can be seen solely as a forecast combination modelling question. By doing so, we demonstrate that temporal hierarchies can be understood as a more general problem than hierarchical forecasting. Beyond any empirical

accuracy benefits, seeing temporal hierarchies as a forecast combination, helps us to derive some of their theoretical properties with respect to the expected (1) variance and (2) accuracy of the forecasts, explaining current empirical findings in the literature, and (3) provides a theoretical motivation for eliminating some of the temporal hierarchy levels, which we show is necessary for the approach to be always competitive with standard time series modelling. Finally, we demonstrate substantial estimation efficiency gains. We conclude by discussing whether these properties and efficiency gains apply to cross-sectional hierarchical forecasting.

Hierarchical Forecasting: The Role Of Information

Presenter: Farshid Vahid-Araghi
Co-authors: Farshid Vahid-Araghi;Minh Nguyen

In hierarchical forecasting, the process of forecast reconciliation transforms a set of "raw" forecasts that do not satisfy the hierarchical aggregation constraints in the real data into a set of "coherent" forecasts that do satisfy those constraints. The reconciliation algorithms are mostly variants of either the ordinary least squares algorithm by Hyndman et al (2011), which can be applied when there is no history to estimate the accuracy of raw forecasts, or the minimum trace algorithm by Wickramasuriya et al. (2019), which can be applied when there is sufficient history to estimate the accuracy of raw forecasts.The academic literature provides ample simulation evidence and real-world data scenarios to demonstrate the value of forecast reconciliation in improving hierarchical time series forecasts. This is attributed to the value of imposing the aggregation constraints. However, this evidence is derived from raw forecasts, each generated using a distinct information set, usually the univariate information set corresponding to each time series. Since reconciliation algorithms combine forecasts, it is difficult to identify to what extent the improvement is due to imposing a true constraint or to combining the information carried by different forecasts.In this paper, we demonstrate that reconciliation algorithms adjust raw forecasts by adding a proportion of their "incoherency" to them, and naturally, if the raw forecasts are already coherent, then these algorithms do not modify the raw forecasts at all. However, if each forecast is based on a distinct information set, and we have access to historical data to estimate the accuracy of raw forecasts, there is scope for improving raw forecasts by combining the information that each one carries, even when they are already coherent. We provide simulation evidence to illustrate the role of information in forecasting. We then propose an algorithm that involves a combination of information step prior to reconciliation. We apply this algorithm to datasets that have been used in the literature and discuss the results.