International Symposium of Forecasting 2021

June 27-30, Virtual + Local hubs

General Chair: George Athanasopoulos, Program Chair: Anastasios Panagiotelis

Book of Abstracts

Accuracy, Explainability, and Trust in Business Forecasting

Presenter: Simon Spavound
Co-authors: Simon Spavound; Nikos Kourentzes

Many businesses rely extensively on their forecasting systems in order to make decisions which concern the future. Given the importance that businesses put on making the correct decisions, a disproportionate emphasis is placed on the statistical accuracy of these systems, with little thought for the wider implementation challenges and overall utility within business processes. These issues can create larger barriers to implementation success and crucially user acceptance. This problem can be unwittingly exacerbated by the increasing algorithmic specialisation of data scientists in organisations, who are service providers to business users. With the advent of Machine Learning based forecasting approaches, and business decisions being made off the back of them, trust and the use of these systems by individuals and teams will become ever more important. This talk will discuss these issues, the wider issues of implementation, and argue for the need for trustworthy forecasting" and considering forecasting in context, rather than solely as a modelling problem.

An Analysis of Merchandise Allocator Adjustments to Sales Forecasts in Apparel Retailing

Presenter: Gökhan Sürmeli
Co-authors: Gökhan Sürmeli;

Large scale retailers in the apparel industry often face the challenging task of accurately forecasting demand for a large number of store-product combinations and varying horizons. It is common for merchandise allocators to go through the automatically generated statistical forecasts and revise them in order to ensure that the right amount of goods is available at the right stores at the right time. However, even though there may be cases where manual adjustments add value to the statistical forecasts, it is documented in the literature that they can also have an overall damaging impact on the accuracy. It is important for retailers to understand the conditions where manual overrides are beneficial and where they are not in order to determine the appropriate set of forecast adjustment rules or nudges. In this study, I share insights from the forecast adjustment practices of merchandise allocators based on statistical analysis of forecasts at a large multinational apparel retailer and particularly discuss how different aspects such as adjustment size, direction, forecast horizon and forecastability relate to whether the adjustments improve or damage the initial system forecast. Merchandise allocators are found to be biased towards positive adjustments to a large extent and generally decrease forecast accuracy. Even though negative adjustments are rare, they are more likely to add value to the statistical forecasts. Furthermore, I argue that merchandise allocators should be communicated regularly with regards to the product groups, time periods, and stores that systematically

have poorforecasting accuracy, and they should be nudged towards manual adjustments where statistical forecastability is low.

Local market information and decision support in retailing

Presenter: Anna-Lena Sachs
Co-authors: Anna-Lena Sachs; Michael Becker-Peth; Stefan Minner; Ulrich Thonemann

Retail managers make order decisions for many products every day. The task of balancing supply with demand is especially challenging for perishable products with short shelf lives as excess inventory has to be discarded. Central optimization tools can provide decision support for local decision makers. However, if local market information is not accessible for the centralized system, it might be beneficial to allow store managers to make decentralized decisions. Current behavioral operations literature suggests that human decision makers are biased. Therefore, it is not clear whether local market information compensates the potential efficiency loss by human decision makers. In a field experiment at a large European retail chain, we explore whether and when local store managers are able to incorporate local knowledge into ordering decisions efficiently. We compare the performance to a central optimization tool. The store managers determine the order quantity for breads, rolls and pastries. We randomly assign stores to one control and three treatment groups. We provide different levels of support to the store managers depending on their group. Besides a control group where decisions were made directly by the central system (group 4), the level of support ranged from sales data as it is readily available from standard point-of-sale scanner systems (Group 1) over providing lost sales estimates for stock-out days (Group 2) to recommending model-based order quantities (Group 3).We find that store managers in Group 1 and 2 achieve a service level that is significantly pulled-to-center. However, store managers can still outperform the central system in terms of total cost. Service levels are higher for Group 3 but still below the target service level of 70%. Interestingly, store managers struggle to incorporate local knowledge when provided with order quantity recommendations in Group 3. The improvement to the central benchmark (controlling for achieved SL) is significantly smaller for group 3 compared to Group 1 and 2. This suggests that providing more sophisticated support does not always improve total cost although it increases service levels. Consequently, retailers should aim for finding ways to combine the best of both worlds, i.e., local market information and optimization tools.

Multi-day-ahead Electricity Price Forecasting: A Comparison of fundamental, econometric and hybrid Models

Presenter: Arne Vogler
Co-authors: Arne Vogler; Philip Beran; Christoph Weber;

The optimization, trading and subsequent operation of energy companies' generation assets requires taking a view on future electricity prices. To adequately account for the technical restrictions of generation units and to serve the various electricity markets, the decision-making problems have to be considered over and beyond the day-ahead horizon. Consequently, electricity price forecasts have to be provided over multi-day-ahead horizons as well.

The majority of contemporaneous studies on short-term electricity price forecasting (EPF) has considered the day-ahead forecasting horizon. Thus, very little work has been done on assessing models for forecasting hourly day-ahead prices over longer horizons. We consider well-established econometric models and a representative of fundamental models that are rather scarce in context of short-term EPF, as well as their various combinations, so called hybrid models. These models are configured both as recursive and direct variants and we examine and compare all model specifications for the individual forecasting horizons t+1 to t+7 as well as over this entire period using a case study of the German day-ahead market for the year 2016. Our work thus contributes to the scarce literature on short-term fundamental EPF models and provides empirical evidence on the forecasting accuracy of popular EPF models over horizons beyond the standard day-ahead horizon.

We conclude that the best models across the individual horizons and across all horizons jointly are hybrid model approaches. They incorporate the common autoregressive elements of state-of-the-art EPF models and pair them with fundamental information. The purely fundamental models are outperformed over all horizons by the autoregressive and hybrid models, thus confirming the well-known difficulties of fundamental models with regard to the reproduction of extreme prices, price volatility and non-linear relationships. Although the forecasting errors of all models increase with forecasting horizon, our results show that the hybrid model approaches have the lowest errors and significantly outperform the other models. They combine the strengths of autoregressive models in terms of capturing daily structures with the immediate reactions of fundamental models to short-term events or fundamental changes in the market.

Price forecasting in electricity markets: evidence from the last decade

Presenter: Hussain Kazmi
Co-authors: Hussain Kazmi; Siqi Liu; Johan Driesen

Electricity markets remain an essential cog in the liberalization process that European governments initiated in the late 20th century. Over time, this has seen complex multi-level markets emerge, including futures, day-ahead, intraday and real-time (imbalance) markets. The rationale for such electricity markets is two-fold: they provide feedback to market participants on how to modulate their energy demand and generation in the short term. Theoretically, they also provide signals on long term investment opportunities and requirements.However, in the intervening years, the economic goal of market reform through liberalization has been superseded by the societal goal of decarbonization. On the one hand, this has led to a (comparatively) rapid increase in the amount of renewables installed in most countries. At the same time, it has seen demand in many places stagnate due, at least in part, to efficiency improvements. With the increasing emphasis on electric transportation and heating, demand growth is expected to pick up again in the coming years. In the case of Belgium, the subject of this study, the solar and wind capacity monitored by Elia, the transmission system operator, has grown by roughly 60% and 500% respectively since 2013. The load meanwhile has not changed much. Coupled with similar trends in neighbouring countries, this is increasingly leading to more volatility in electricity prices, including positive and negative price spikes. Pandemic induced lockdowns have further exacerbated the situation.As the proliferation of renewables continues, forecasting market behaviour will become increasingly important to ensure stable, sustainable grid operation. Therefore, we attempt to address a number of major research questions relevant to price forecasting in electricity markets. First, can electricity market prices be forecast accurately, and to what extent can knowledge of exogenous variables, such as renewable generation and demand, improve them. Second, how have errors made by the system operator in exogenous variable forecasts evolved over time, and whether (or how) they have affected market prices. Finally, we consider whether data-driven forecasting models can provide some information about the future direction of electricity markets, based on different scenarios for future energy demand and generation mix in Belgium.

Optimal bidding of large volumes on two day-ahead electricity price auctions

Presenter: Michal Narajewski
Co-authors: Michal Narajewski; Florian Ziel

Electricity exchanges offer several trading possibilities for market participants: starting with futures products through the spot market consisting of the auction and continuous part, and ending with the balancing market. This variety of choice creates a new question for traders — when to trade to maximize the gain. This problem is not trivial especially for trading larger volumes as the market participants should also consider their own price impact. The following paper raises this issue considering two markets: the hourly EPEX Day-Ahead Auction and the quarter-hourly EPEX Intraday Auction. We consider a realistic setting which includes a forecasting study and a suitable evaluation. For a meaningful optimization many price scenarios are considered that we obtain using bootstrap with models that are well-known and

researched in the electricity price forecasting literature. The own market impact is predicted by mimicking the demand or supply shift in the respectful auction curves. A number of trading strategies is considered, e.g. minimization of the trading costs, risk neutral or risk averse agents. Additionally, we provide theoretical results for risk neutral agents. Especially we show when the optimal trading path coincides with the solution that minimizes transaction costs. The application study is conducted using the German market data, but the presented methods can be easily utilized with other two auction-based markets. They could be also generalized to other market types, what is discussed in the paper as well. The empirical results show that market participants could increase their gains significantly compared to simple benchmark strategies.

Multi-asset return predictability using VARs

Presenter: Sofia Monteiro
Co-authors: Sofia Monteiro; Nuno Silva; Helder Sebastian

Financial markets exhibit high levels of instability and uncertainty that challenge investors to find more robust predictive methods and to use more sensible estimation models. In this study, we analyze the predictive accuracy of different multivariate-VAR prediction models and time-varying Bayesian models with dynamic model averaging/ selection when jointly forecasting the returns of stocks, bonds and, REITs, for the period of January 1976 to December 2017. We conclude that Bayesian-based approaches bring the largest gains in terms of statistical predictability, as measured by the R2 out-of-sample and the MSFE-adjusted statistic of Clark and West (2007), in terms of economic performance, which we quantify using certainty equivalent returns, Sortino ratios, Sharpe ratios. Lastly, we compare the performance of the models before and after the subprime turmoil period and conclude that the Bayesian approaches are able to accommodate market instability.

Incorporating bank-level financial variables into forecasting of Russian macroeconomic series

Presenter: Stas Tatarintsev
Co-authors: Stas Tatarintsev; Alexei Ponomarenko; Ramis Khabibullin

In this paper we incorporate individual level Russian bank's data to backcast aggregate real variables (industrial production, unemployment, etc.) after the outbreak of the Great Financial Crisis (GFC) in the August 2008. Several authors (Buch et al, 2010; Dave et al, 2013) Â showed the significance of the link between the bank-level financial data and macro-level variables. The majority of these papers are focused only on the bank-level loans data. Along with the total loans we consider each bank's share of low quality loans in the total loans. This variable potentially could be a good predictor of Â credit and business cycles, which could be helpful for forecasting macro-level variables. Specifically, using this variable we try to model macro-effect of worsening of individual bank financial position. We backcast Â Russian real variables after the GFC. We combine monthly microdata for Russian banks' loans with Russian monthly macroeconomic dataset from December 2008 to December 2020. We split this sample into the train sample from January 2011 to December 2020 and test sample from December 2008 to December 2010. We estimate several Bayesian Vector Autoregression models (BVAR) with optimized hyperarameters similar to GLP (Gianonne et al, 2015). Firstly, we estimate BVAR using macro-level data only, then we add bank-level variables to the set of endogenous variables. We also estimate regularized FAVAR. Finally, we estimate non-linear models: Threshold BVAR and Markov-switching BVAR. We show that the inclusion of bank-level variables in the conventional macro-level dataset improves the accuracy of real variables forecasts.ReferencesBuch, C. M., Eickmeier, S., & Prieto, E. (2010). Macroeconomic factors and micro-level bank risk.Dave, C., Dressler, S. J., & Zhang, L. (2013). The bank lending channel: a FAVAR analysis. Journal of Money, Credit and Banking, 45(8), 1705-1720.Giannone, D., Lenza, M., & Primiceri, G. E. (2015). Prior selection for vector autoregressions. Review of Economics and Statistics, 97(2), 436-451.

Dynamic relationship between Stock market and Bond market: A GAS MIDAS copula approach

Presenter: Hoang Nguyen
Co-authors: Hoang Nguyen;

There is evidence that macroeconomic variables influence the relation among financial variables, however they are sampled at different frequencies. This study proposes generalized autoregressive score mixed frequency data sampling (GAS MIDAS) copula models to analyze the dynamic relationship between Stock returns and Bond returns. A GAS MIDAS copula decomposes their relationship into a short term dependence and a long term dependence. While the long term effect is updated at a lower frequency using a MIDAS regression, the short term effect follows a GAS process. Asymmetric dependence at different quantiles are taken into account. The model helps to improve the in-sample goodness of fit and the out-of-sample forecast.

Multivariate Forecasting using Artificial Neural Networks

Presenter: Alina Taenzer
Co-authors: NA

In the recent years, machine learning has gained more and more importance in the field of macroeconomic forecasting. While there exists already selective research on forecasting with neural networks compared to linear vector autoregressive models, this paper contributes a multivariate forecasting comparison between structural models and ML-based tools. Specifically, a fully connected feed forward non-linear autoregressive neural network (ANN) is contrasted to the well-known dynamic stochastic general equilibrium (DSGE) model by Del Negro et al. (2015) and a bayesian vector autoregression (BVAR) using optimized priors as in Giannone et al. (2015). Model estimation and forecasting is based on quarterly U.S. data from 1964Q2 to 2019Q4. Using real-time data for 8 macroeconomic time series (GDP, inflation, federal funds rate, spread, consumption, investment, wage, hours worked), a forecasting comparison based on expanding window estimations is conducted. The results show that there is an overall gain from using ANN, while also supporting evidence that BVAR improves forecasts compared to DSGE. The superiority of the ANN varies with the out-of sample periods to be forecasted. Averaging over forecasts between 1974 and 2019, the long-term predictions (6 to 8 quarters ahead) by ANN for GDP, inflation and the federal funds rate are more precise. When forecasting periods are restricted to 1999:2019, the superiority expands to 3 up to 8 quarter ahead forecasts. Focusing on post-financial-crisis times, the ANN yields the lowest RMSFE, averaged over all 8 variables, with forecasting horizons 3 to 8. Inflation forecasts by the ANN even outperform conventional methods over all horizons. These results prove that non-linear data-driven ANNs are a useful method when it comes to macroeconomic modelling and forecasting.

A Neural Network Ensemble Approach for GDP Forecasting

Presenter: Luigi Longo
Co-authors: Luigi Longo; Massimo Riccaboni; Armando Rungi

We propose an ensemble learning methodology to forecast the future US GDP growth release. Our approach combines a Recurrent Neural Network (RNN) with a Dynamic Factor model accounting for time-variation in mean with a Generalized Autoregressive Score (DFM-GAS). The analysis is based on a set of predictors encompassing a wide range of variables measured at different frequencies. The forecast exercise is aimed at evaluating the predictive ability of each model's component of the ensemble by considering variations in mean, potentially caused by recessions affecting the economy. Thus, we show how the combination of RNN and DFM-GAS improves forecasts of the US GDP growth rate in the aftermath of

the 2008-09 global financial crisis. We find that a neural network ensemble markedly reduces the root mean squared error for the short-term forecast horizon.

Benchmarking pre-structured recurrent neural network architectures for macro-economic problems

Presenter: Nico Beck
Co-authors: Nico Beck; Julia Schemm; Claudia Ehrig

Accurate data-driven forecasts can provide a crucial advantage in many application areas. One of the methods with the most promising results in forecasting time series are neural networks. However, especially in macro-economic applications, it can be difficult and time-consuming to adapt state-of-the-art neural network architectures in a way that leads to satisfying results. For instance, the final prices of materials and stocks result from a highly complex interplay between supply and demand. Additionally, there is often only one (albeit long) historical time series available for training which makes correlations in the data difficult to detect. Under these circumstances, applying state-of-the-art neural networks architectures successfully poses a great challenge.Pre-structuring the models can solve this problem. For this purpose, Zimmermann, Tietz and Grothmann (Neural Networks: Tricks of the Trade, 2012) propose recurrent architectures for various time series problems that help recognize correlations. They recommend Error-Correction Neural Networks (ECNNs), Historical-Consistent Neural Networks (HCNNs) and Causal-Retro-Causal Neural Networks (CRCNNs). One of the main ideas of the pre-structuring is embedding the model in a larger architecture in order to use the past prediction errors for predicting the next time step. The three approaches mentioned use this idea and apply it in different settings.So far, the proposed architectures are not publicly available in common machine learning frameworks. Consequently, it has not been possible to evaluate and compare them to other recurrent networks like Long Short-Term Memory models (LSTM) or Gated Recurrent Unit models (GRU). Therefore, although they achieved remarkable results in diverse applications like forecasting the copper price (Zimmermann, Tietz, Grothmann et al, Recurrent Neural Networks for Industrial Procurement Decisions, 2012), an informative benchmark of the architectures is still missing.In order to close this research gap, we have implemented the models in PyTorch. This way, we can easily test them on diverse datasets. To evaluate further the performance, we benchmark these methods on datasets stemming from diverse areas of application. Here we compare the HCNNs, ECNNs and CRCNNs with state-of-the-art machine learning methods and classical forecasting approaches to show their capability.

New Product Demand Forecasting for Fashion Retail

Presenter: Vijay Ekambaram
Co-authors: Vijay Ekambaram; Kushagra Manglik; Sumanta Mukherjee; Surya Shravan Kumar Sajja; Satyam Dwivedi; Vikas Raykar

Fashion is considered to be the world's second-largest polluter, after oil and gas, as it comes with severe environmental costs in managing the unsold dead inventory. Fear of counterfeiting and loss of brand-value force fashion houses to take unsustainable measures like destroying unsold inventory. At the heart of this problem, lies the mismatch between supply and demand. Inaccurate demand forecasts and the urge to avoid stock-outs lead to overproduction. Unlike other retail industries which have rich historical time series sales data for forecasting, the fashion industry is heavily trend-driven; whereby, most products are new designs (i.e. no historical data to forecast).Â In this paper we propose forecasting algorithms for new product demand forecasting based on available product attributes, images and external factors. Existing works on new product demand forecasting are mostly based on K-Nearest Neighbor(KNN) approaches that lack the ability to model complex non-linear relations between multi-modal data sources, sales, and external factors. To address this, we propose and empirically evaluate different multi-modal attention-based encoder-decoder models that can effectively forecast the sales time-series of new products in the fast fashion

domain. We also study the impact of various multi-modal fusion techniques in new product time-series forecasting, which enables effective gradient flow towards data-sources leading to more information gain. To overcome the black-box nature of our models, we incorporate self-attention and cross-attention techniques and empirically validate their efficacy to enable effective explanations. We conduct experiments on a large-scale fashion data set (comprising of 10,290 products distributed across 45 categories) and report results and interesting findings to illustrate the benefits of modeling new product time-series forecast as a multi-modal encoder-decoder sequence problem as opposed to the conventional KNN approaches.

Selection of the number of neighbours for k-NN retail promotional forecasting

Presenter: Carlos-Eduardo Rodriguez-Calderon
Co-authors: Carlos-Eduardo Rodriguez-Calderon; Sven F. Crone; Anna-Lena Sachs

The k-Nearest Neighbours (k-NN) method has been used for classification, clustering, regression, and recently to perform numerical forecasting for time series. k-NN operates by calculating local averages, depending on the k closest observations (neighbours). This can be an advantage over conventional forecasting approaches for which the modeller would have to define the functional form of the forecasting model. Instead, the modeller needs to select the distance metric, the number of neighbours, and the type of average used, for the k-NN. Both inappropriately small or large number of neighbours can harm the forecast accuracy, and yet there are no theoretical guidelines how to choose the number of neighbours. Prior literature in k-NN classification recommends calculating the number of neighbours using a fixed value or using cross-validation by defining a search scope for k. This paper evaluates these common approaches for the task of promotional forecasting. We consider using a fixed number of k and cross-validation to define an optimal k for each item. Furthermore, we propose a novel heuristic approach, where the maximum k is restricted by the features in the observed data. For example, it restricts the generation of promotional forecasts to use at maximum a number of neighbours equal to the number of promotional periods in the past, if not less, in contrast to fixed or cross-validation identified k that can exceed that and therefore mixing promotional and non-promotional periods with adverse effects on forecast accuracy. We also recommend how to define boundaries in searching for k with cross-validation.

Univariate forecasting methodologies for forecasting medicine demand in hospital pharmacies

Presenter: Devon Barrow
Co-authors: Devon Barrow; Nikolaos Kourentzes

Within hospital pharmacies, forecasts of medicine demand are needed for accurate and effective management of the associated inventory. Accurate inventory levels should lead to reduced costs, release valuable resources which may otherwise be tied up in expensive medicine stock. The objective of this work is to provide a systematic evaluation of recent univariate forecast innovations to support operations management in this critical sector. We evaluate a broad range of univariate techniques for forecasting drug demand for a set of UK-based hospital pharmacies. We explore a range of standard to more recent techniques capable of responding to the unique features of the dataset. This includes an exploration of the cross-temporal structure of the dataset, with demand typically viewed daily and weekly across different hospital sites and wards, as well as the intermittent nature of drug orders at the daily level for individual wards.

Exploring the social influence of Kaggle virtual community on the M5 competition

Presenter: Yun Bai (BUAA)
Co-authors: Yun Bai; Xixi Li; Yanfei Kang

One of the most significant differences of M5 over previous forecasting competitions is that it was held on Kaggle, an online community of data scientists and machine learning practitioners. On the Kaggle platform, people can form virtual communities such as online notebooks and discussions to discuss their models, choice of features, loss functions, etc. This paper aims to study the social influence of virtual communities on the competition. We first study the content of the M5 virtual community by topic modeling and trend analysis. Further, we perform social media analysis to identify the potential relationship network of the virtual community. We find some key roles in the network and study their roles in spreading the LightGBM related information within the network. Overall, this study provides in-depth insights into the dynamic mechanism of the virtual community's influence on the participants and has potential implications for future online competitions.

Forecasting Principles from Experience with Forecasting Competitions

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

Economic forecasting is difficult, largely because of the many sources of nonstationarityinfluencing observational time series. Forecasting competitions aim to improve the practice of economic forecasting by providing very large data sets on which the efficacy of forecasting methods can be evaluated. We consider the general principles that seem to be the foundation for successful forecasting, and show how these are relevant for methods that did well in the M4 competition. We establish some general properties of the M4 data set, which we use to improve the basic benchmark methods, as well as the Card method that we created for our submission to that competition. A data generation process is proposed that captures the salient features of the annual data in M4.

From learning analytics to forecast learning: using students' digital footprint to improve learning

Presenter: Yves Sagaert
Co-authors: Yves R. Sagaert; Stefaan Haspeslagh; Liam Bossant; Tom Madou

The use of digital tools for learning has increased exponentially in recent decades. In fact, the recent COVID19 pandemic is expected to accelerate this transition. While the potential of these digital tools is generally recognised, their effectiveness remains unclear. Learning Management Systems (LMS) such as Blackboard or Moodle are established software applications that, in addition to other administrative tasks, facilitate the delivery of online educational resources (e.g. documents, screencasts, or quizzes) and user interaction (e.g. online forums). While most Learning Management Systems facilitate the analysis of user data to some extent, this practice is not well established among teachers and as a result, online learning remains a 'black box' operation in many cases. Combined with a strong reduction of real-life face time, this makes it difficult for teachers to assess learning results during a learning process. In-process evaluations of students' academic progress are referred to as formative assessment and are important to adjust the learning environment. Tracking students' online actions and using them as an input for teacher-student interaction is expected to benefit the students' learning curves. However, in order to take action on this analysis, it is imperative that the forecast results come early in the semester. To predict student performance early in the term, we use all registered student actions (the digital footprint) on the online learning platform. We collect user data from 184 students with a level of detail exceeding the standard LMS learning analytics modules using the xAPI eLearning standard and Learning Locker, an open source Learning Record Store. We analyse this data to identify key performance indicators facilitating the early forecasting of learning results with future use of the online learning tools of the LMS.

Talking in a language that everyone can understand? Transparency of speeches by the ECB Executive Board

Presenter: Lena Sophia Müller
Co-authors: Lena Sophia Müller; Alexander Glas

Using novel data on speeches held by members of the European Central Bank's Executive Board, we investigate whether monetary policy transparency has increased over time. With respect to the general public as the target audience, our findings suggest that the European Central Bank successfully improved the frequency and clarity of information provision since its inception. The increase in transparency is gradual, rather than being induced by changes in the Executive Board's composition or major economic events such as the Great Recession. However, the clarity of speeches in recent years is still fairly low. Moreover, our findings indicate that clarity decreased under Christine Lagarde's presidency following the outbreak of the Coronavirus pandemic. We conclude that while the European Central Bank was able to increase transparency over time, further improvements in clarity are required to make monetary policy truly accessible to the broad public.

Inflation dynamics and forecast: frequency matters

Presenter: Fabio Verona
Co-authors: Fabio Verona; Manuel Martins

We show that the new Keynesian Phillips Curve (NKPC) valuably forecasts U.S. inflation once frequency-domain information is taken into account. We do so by decomposing the time series (of inflation and its predictors) into several frequency bands and forecasting each frequency component of inflation. The larger statistically significant forecasting gains are achieved with a model that forecasts the lowest frequency component of inflation (corresponding to cycles above 16 years) flexibly using information from all frequency components of the NKPC inflation predictors. Its performance is particularly good in the turning from the Great Recession to the recovery.

Inflation forecasting performance of a developing economy central bank- The case of Ghana

Presenter: Stephen Opata
Co-authors: Stephen Opata;

This paper assessed the inflation forecast performance of Bank of Ghana (BoG) using "Mincer-Zarnowitz" regressions and "step indicator saturation" (SIS) methodology. The "forecast performance" of the BoG was compared to a benchmark forecast using the "random walk model" used in (Atkeson & Ohanian, 2001) and the IMF WEO forecasts. The BoG's forecast outperformed the other forecasts and the "Diebold Marino" test further supported the conclusion that its' "one-step ahead" forecast was superior to the random walk forecast. For robustness checks, the forecast encompassing test concluded that the BoG's forecasts reflected all information embedded in the random walk forecast but the same could not be said of the random walk forecast. The Bank of Ghana's one-quarter ahead inflation forecast exhibited stronger efficiency when SIS variables were incorporated in the forecast and demonstrates the importance in addressing "outliers" and "structural breaks" in evaluating macroeconomic forecasts especially in developing economies.JEL Classification: C51; C22; C53Keywords: Step indicator saturation, Mincer-Zarnowitz regression, forecast error, random walk forecast, inflation forecasting

Are political risk and policy uncertainty indicators useful to nowcast macro-financial developments? An application for Latin American using text-based variables

Presenter: Javier J. Perez
Co-authors: Javier J. Perez;

We develop political risk and economic policy uncertainty indicators for the main Latin American economies (Brasil, Mexico, Argentina, Peru, Colombia, Chile) using supervised learning techniques, by means of computational text analysis applied to a wealth of Spanish newspaper sources. The so-obtained indicators contain useful information to anticipate changes in the stock market value of Spanish multinational companies most exposed to Latin American markets. In addition, using mixed-frequency, time-series models, we also show that country-specific indicators are useful to nowcast country-specific macroeconomic variables in the selected Latin American economies.

The value of the online sentiment proxies in realized volatility forecasting: a large scale longitudinal evaluation

Presenter: Ping LIN
Co-authors: Ping Lin; Shaohui Ma; Robert Fildes

Although some studies show that online sentiment proxies can help to improve the prediction accuracy of stock market volatility, most of them lack out-of-sample validations, had inadequate validation samples, nor did they use proper controls in their empirical models. In this paper, we use a data sample consisting of 300 randomly selected stocks, 9 years of their daily realized volatilities, collected from the Shanghai and Shenzhen Stock Exchanges of China. Employing both local and global predictive methods, and adopting trading sentiment controls, we investigate the predictive value of online sentiment proxies extracted from three sources of online "˜big data'. We find that (1) traditional local linear regression models can not reveal the predictive value of online sentiment, but Random Forest with data pooling performs much better; (2) the readily available trading sentiment proxies have even more predictive power than online sentiment proxies; (3) when employing trading sentiment proxies as control variables, though the extra accuracy gains due to online sentiment proxies is statistically significant, the magnitude of the gains is in general very limited; (4) Online sentiment proxies are particular effective for stocks with large market value, and during the bear periods in the stock market.

Measure market vulnerability on market sentiments

Presenter: Michal Chojnowski
Co-authors: Michal Chojnowski;

Author analyses the impact of market sentiments on given European markets. The article proposes a novel method both in observing market sentiments and in evaluating its impact on the market. The method uses mixed LSTAR-ANN model with Google Trend as an indicator for market sentiments.The method is based on works "Sentiments" (Angeletos, La'O, 2013) and "Sentiments and aggregate demand fluctuations" (Benhabib, Wang, Wen, 2015) where market sentiments are treated as economic sunspots "" an exogenous variable which does not influence economic fundamentals, but might impact their expectations. In above-mentioned models market sentiments are affecting one's beliefs about the economy situation. To illustrate that, author uses LSTAR model, where customer confidence index is used as threshold variable.Artificial Neural Network is used to gather and aggregate Google Trends queries for more insight of market sentiment components. Each query is assigned to one of the chosen topic groups to determine which information is most influential for given markets. Author proposes a method, which combine LSTAR and ANN into one neural network, therefore both LSTAR coefficients and neural network weights are

computed together. With rolling window approach, it is possible to show evolution of topic impact on market sentiments.Paper focuses on 6 European countries: Germany, Spain, France, the Netherlands, Poland and the United Kingdom on 7 markets: beverages, tobacco, textiles, paper products, electrical equipment, furniture and electricity & gas. The analysed period is 2004-2020.

The effect of the forecasting metric choice in retail inventory management applications

Presenter: Satyam Dwivedi
Co-authors: Satyam Dwivedi; Ali Koc; Brian Quanz; Pavithra Harsha; Dhruv Shah; Mahesh Ramakrishna

In retail settings, many business decisions are based on forecasts from observed sales history time-series data. Thus being able to accurately forecast future sales from historical data gives businesses an edge over competition. However, how to best quantify "accuracy" of a forecasting algorithm is not clear. Literature suggests many different metrics to evaluate forecasts given the actual observations. This may give conflicting scores for the same forecast and observations. Further, study on the impact of metric choice on down-stream business usage of forecasts such as in inventory management applicationsÂ is lacking.Â This ambiguity suggests a clear need to study how the choice of forecasting metric affects different business use-cases.In this work, we study the effect of different forecasting metrics on overall business output in multiple simulated retail settings. Since the ultimate goal for every business is to maximize their profits, we use gross profits as a key business measure to compare different forecasting metrics. To eliminate the choice of the forecasting algorithm biasing our findings, we simulate demand and compare metric minimizers for the true demand distributions. We model demand for multiple products at multiple stores as count distributions matching typical retail patterns over time and the retail network, and simulate demand realizations from these distributions. We then obtain the forecasts for each of the forecasting metrics by optimizing the metric over the known distribution. These forecasts, along with the simulated actuals are then passed to an inventory management system which simulates real world transactions and calculates the profit and other KPIs for the simulated time period. Similar experiments are run for multiple retail settings and different inventory management approaches by passing different parameters to the demand simulation and inventory planner. These results are then used for a thorough comparative study on the effect of different forecasting metrics on actual business profits in various different settings.Â

A model of cannibalization for E-commerce demand forecasts

Presenter: Rémy Garnier
Co-authors: Rémy Garnier;

Multiple times series are sometimes in competition with one another. It is the case for the sales of the different products of the same type on a marketplace on an E-commerce website, with their changing prices and search engine ranking. This phenomenon is called cannibalization. However, these time series are subject to the same global effect, such as trend or seasonality. In this presentation, I will provide a predictive model of cannibalization for multiple time series. For this purpose, a "competitiveness" function is introduced and computed using external covariates and a deep learning approach. The sales of each product are distributed using this competitiveness function. An application of this model to E-commerce data is also proposed

Demand forecasting in times of COVID

Presenter: Michal Kurcewicz
Co-authors: Michal Kurcewicz;

This paper discusses the impact of COVID on the operations of a large-scale forecasting and replenishment system at a major Polish retailer. COVID had an unprecedented impact on the retailer's supply chain, changing both overall customer demand and its structure and affecting product availability at suppliers. We analyze the timeline of events and present process changes introduced to safeguard the forecasting process. We describe and evaluate changes to the statistical forecasting process including retuning of machine learning models and strategic use of forecast overrides. Also discussed are operational and organizational aspects that improved process resilience.

Testing the predictive accuracy of COVID-19 forecasts

Presenter: Laura Coroneo
Co-authors: Laura Coroneo; Alessia Paccagnini; Paulo Santos Monteiro; Fabrizio Iacone

We test the predictive accuracy of forecasts for the number of COVID-19 fatalities produced by several forecasting teams and collected by the United States Centers for Disease Control and Prevention (CDC), both at the national and state levels. We find three main results. First, at short-horizon (1-week ahead) no forecasting team outperforms a simple time-series benchmark. Second, at longer horizons (3 and 4-weeks ahead) forecasters are more successful and sometimes outperform the benchmark. Third, one of the best performing forecasts is the Ensemble forecast, that combines all available forecasts using uniform weights. In view of these results, collecting a wide range of forecasts and combining them in an ensemble forecast may be a safer approach for health authorities, rather than relying on a small number of forecasts.

The stock markets of the BRICS: How contagious is Covid-19?

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

The coronavirus pandemic has given a new quality of contagion to economies, plunging their financial markets in a state of uncertainty and fear. We investigate the impact of the coronavirus pandemic on the network of the BRICS (Brazil, Russia, India, China, South Africa) stock markets plus the US, with a focus on the distributional characteristics of the returns of their respective stock indices.To this end, the financial markets' perception of the pandemic is tracked on a daily basis from two perspectives: in terms of a monitoring of smoothed new Covid-19 cases and in terms of forecast errors, i.e. deviations between actual and forecasted new Covid-19 cases. The latter are obtained by rolling forecasts of growth rates for each country in the network. It turns out that OLS regression of stock index returns on market sentiment does not lead to significant results. A more promising approach is to link the returns' distributional behavior and the pandemic using multivariate quantile regression. Among our findings is that Russian Covid-19 cases, and the corresponding forecast errors, have a significant effect on upper return quantiles in the network, while Indian and Chinese cases tend to affect the return volatility.

Forecasting COVID-19: A large-scale comparison of alternative models.

Presenter: Christos Emmanouilides
Co-authors: Christos Emmanouilides;

The paper assesses the performance of alternative model classes in forecasting the COVID-19 pandemic evolution. A large-scale forecast comparison experiment is performed using data up to early 2021 on COVID-19 epidemic waves for 138 countries. Short to long-range point forecasts of the epidemic are generated from non-parametric regression, exponential smoothing, time-series, and growth curve models. The models are estimated recursively and in rolling windows of varying length. A series of statistical tests on several performance metrics formally compare the models' forecasting accuracy. The study provides a

substantial amount of evidence that forecasts from non-parametric regression models, estimated in about one-month-long rolling windows, are overall more accurate than point forecasts from other models.

Kalman fillter and smoothing factor extraction: Does the specification matter?

Presenter: Pilar Poncela
Co-authors: Esther Ruiz; Pilar Poncela; Karen Miranda

Dynamic Factor Models (DFMs), which assume the existence of a small number of un- observed latent factors that capture the comovements in a system of variables, are very popular among empirical macroeconomists as a procedure to reduce the dimension and ex- tract factors with an economic interpretation. Factors can be extracted using the Kalman lter and smoothing (KFS) procedures that are efficient and can cope with missing and mixed-frequency data, time-varying parameters, non-linearities and non-stationarity among many other stylized facts often observed in real systems of economic variables. This paper analyses the empirical consequences on KFS factor estimation and forecasting of using al- ternative estimators of the parameters under various sources of potential misspecification. In particular, we consider factor extraction when assuming different number of factors and different factor dynamics. The factors are extracted from a popular data base of US macroe- conomic variables that has been widely analyzed in the literature without consensus about the most appropriate model specification. We answer the question about whether this lack of consensus matters.

Nowcasting Spanish regional GDP by means of factor models

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

This paper estimates a multi-frequency factor model to nowcast regional GDP in Spain on a high frequency basis. This nowcasting exercise is of great interest due that regional GDP is estimated at an annual basis with a delay of 12 months after the end of the reference year. The lack of timely releases of regional GDP makes difficult the assessment of the regional business cycle, specially in moments of greater volatility. The model considers regional and national information and also several alternatives to mix data of different frequencies. Special attention is devoted to the coherence between regional and national forecasts. Forecast evaluation is carried out with data from 2015 to 2019 at the regional level and, since national GDP is known for 2020, also at the aggregate level for this last year. We provide forecasts the regional GDP for 2020 and 2021 and assess the heterogeneous impact of the COVID recession.

Commodity price uncertainty co-movement: Does it matter for global economic growth?

Presenter: Laurent Ferrara
Co-authors: Laurent Ferrara; Aikaterina Karadimitropoulou; Athanasios Triantafyllou

Global economic activity is surrounded by increasing uncertainties from various sourcess. In this paper, we focus on commodity prices and estimate a global commodity uncertainty factor by capturing co-movement in volatilities of major agricultural, metals and energy commodity markets through a specific Dynamic Factor Model. Then, by computing impulse response functions estimated using a Structural VAR model, we find that an increase in the common commodity price uncertainty results in a substantial and persistent drop in investment and trade for a set of emerging and advanced economies. We show that a global commodity uncertainty shock is more detrimental for economic growth than usual financial and economic policy uncertainty shocks, for both short and long horizons. Last, when examining the macroeconomic effect of commodity-specific uncertainty shocks, we find that, once we control for global

commodity uncertainty, a shock in uncertainty of metal prices possesses a much higher recessionary impact compared to agricultural and energy shocks.

Are search queries on neighboring destinations useful in tourism demand forecasting?

Presenter: Mingming Hu
Co-authors: Mingming Hu; Haishan Feng; Mengqing Xiao

The tourism industry is inherently vulnerable (Chandra and Menezes, 2001). It is essential to accurately forecast future tourist volumes to guide tourism practitioners' and destination managers' industry planning, policy formulation, and operations (Lin et al., 2015; Li et al., 2006). The growth of the Internet enables prospective tourists to obtain travel-related information via search engines and plan their itineraries in advance. Search engine traces reflect the public's attention to destination-related features, helping scholars and practitioners better understand the tourism market (Li et al., 2018). Numerous scholars have therefore adopted search engine data relating to destination to forecast tourism demand and have demonstrated the utility of these search query data in enhancing forecasting accuracy (Li et al., 2016; Zhang et al., 2017).Bao and Mckercher (2008) conducted a survey in Hong Kong and found that more than 50% of surveyed tourists from the United States, Germany, Indonesia, and Korea were taking multi-destination tours. Tourists visiting several destinations use search engines to research both primary and neighboring destinations; thus, their search traces appear for their primary destination as well as nearby locations. This study explores whether search queries on neighboring destinations can improve the performance of tourism demand forecasting. We take Singapore and United States as source markets with Hong Kong as the main destination, while Taiwan and Macau are taken as neighboring destinations. The out-of-sample one-step-ahead forecasting results of our autoregressive integrated moving average with exogenous input (SARIMAX) model reveal that (1) compared with time series models, models incorporating search engine big data can enhance tourism demand forecasting accuracy; and (2) compared with tourism demand forecasting based on destination search volumes, the forecasting performance based on destination and neighbor search volumes is significantly better. This study contributes to the literature in two aspects. First, it presents a new group of datasets with which to forecast tourism demand. Although search query data have been widely adopted in research, our work marks an initial attempt to integrate search queries on neighboring destinations in tourism demand forecasting. Second, this study is the first to compare search query performance for two neighboring destinations.

Tourism Demand Forecast Comparisons in Unstable and Complex Environments

Presenter: Han Liu
Co-authors: Han Liu; Yongjing Wang

Two tricky typical facts exist when forecast tourism demand in unstable and complex environments, such as COVID-19, oil crisis, financial and economic crisis, natural disasters, terrorist attacks, etc. The first is that the forecast results are unstable and will change over time. The second is that the in-sample goodness of fit of the prediction model cannot ensure good out-of-sample forecasting performance. This paper takes tourism demand and its comprehensive range of mixed frequency high-dimensional factors as the research object and uses robust mixed frequency Granger-causality tests to find out the helpful factors for tourism demand forecasting. Then, the fluctuation test and one-time reversal test proposed by Giacomini and Rossi (2010) to evaluate the performance of tourism demand forecasting. Finally, the error of tourism demand forecasting is decomposed into independent predictive content, over-fitting, and time-varying forecasting ability factors to diagnose the forecast breakdown of out-of-sample tourism demand forecasting in unstable and complex environments. The empirical results show that: (1) The typical facts of tourism demand forecasts do exist in unstable and complex environments; (2) The results of out-of-sample forecast error decomposition suggest that the reasons for the forecast breakdown in tourism demand forecasting are as

follows: the predictive content problem caused by the changing environment, the time-varying forecasting ability in the model specification, the over-fitting factor, etc. (3) The mixed frequency high-dimensional model can solve the problem of forecast breakdown caused by over-fitting factors of tourism demand forecasting when compared with other forecasting models. The comparison of tourism demand forecasting performance and the exploration of the sources of the forecast breakdown can clarify the reasons for the forecast break of tourism demand forecasting, which is helpful to improve the forecasting performance of the model under unstable and complex environments.

Forecasting tourism demand cycles: A dynamic Markov-Switching Approach

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

The effect that the business cycle has on tourism demand has received increasing attention in recent research. Smeral (2017) showed that income elasticities of tourism demand vary across the business cycle, while Croes and Ridderstraat (2018) indicate that business cycles explain tourism demand cycles and that these effects are asymmetric. Given the effect of COVID-19 on economies around the globe, understanding how tourism behaves during economic downswings have become more important since it also sheds light on tourism recovery. This paper extends current research by modelling and forecasting tourism demand cycles to five destinations, namely Australia, Canada, South Africa, Sweden and the USA. The two main approaches used to analyse cyclical patterns are spectral analysis and regime-switching analysis. While spectral analysis has been applied in tourism demand modelling and forecasting, the use of regime-switching models is not common. This paper aims to fill this void by modelling and forecasting tourism demand cycles using a dynamic Markov-Switching Approach. The cyclical components of tourism demand, GDP and prices are extracted using a double HP filter. Rolling forecasts at four different time horizons using the Markov-Switching model in a dynamic specification are compared to traditional models, such as the ARIMA and naÃ¯ve forecasts.

Intermittent Demand forecasting with Neighbour-based Multilayer Percpetrons

Presenter: Morteza Khani
Co-authors: Morteza Khani; Kamran Rismanchi; Sven F. Crone

Forecasting of intermittent time series poses particular challenges to forecasting algorithms. Traditionally Croston's method and extensions thereof have been employed, separating the time series into an estimate of the interdemand interval and a separate estimate of the demand volume, which are then recombined into a demand rate forecast. Similarly, machine learning and artificial neural network models have been applied predict separated time series based on autoregressive information of each series (Kourentzes, 2013). In contrast, k-nearest neighbours algorithms attempt to find similar time series patterns in historic time series data without creating any functional model, with some promising results (Nikolopoulos et al., 2016). However, the forecast is constrained to averages of historic patterns, and cannot extrapolate to previously unseen lengths of interdemand intervals or demand values. In this study we propose to combine the function approximation of a multilayer perceptron with the properties of k-nearest neighbours. The input-output-mapping is not based on autoregressive lags but rather explanatory variables to characterise interdemand-periods, with outputs representing the future predictions of the trace forecasts. Out approach shows promising accuracy and efficiency in predicting intermittent time series en par with alternative approaches.

Inventory Control for Periodic Intermittent Demand

Presenter: Sarah Van der Auweraer
Co-authors: Sarah Van der Auweraer; Thomas van Pelt; Joachim Arts

Intermittent demand is difficult to forecast, as many periods have no demand at all. Forecasting methods for such demand usually create separate estimates for the time between demand occurrences and the size of a demand occurrence. These methods implicitly assume that the time between demand occurrences is memoryless. Data from practice, however, indicates that the times between demand events is often not memoryless but ”“contrary to implicit forecasting model assumptions'displays periodicity. Consequently, the time since the last demand occurrence is an important predictor for future demand. We propose a demand model and a forecasting method that accommodate periodic demand. We benchmark the performance of our method in terms of the performance of inventory policies induced by the forecasts.

Re-Analysis of Intermittent Demand Forecasting Methods

Presenter: John Boylan
Co-authors: John Boylan; Zied Babai; Mona Mohammadipour; Aris Syntetos

The importance of replication (of the same method) in forecasting research is well recognized, but there has been less emphasis on re-analysis (on the same data) of different implementations of the same method. In the M5 competition, statistical forecasting methods were implemented on a series by series basis, with no opportunity of ”˜cross-learning' from other series. This invites the question of the extent to which the implementation of a forecasting method affects its accuracy. The opportunity for cross-learning is of particular interest when the data for an individual series is sparse, for example when the observed data is intermittent. Cross-learning opens up the opportunity to estimate seasonal patterns that would be difficult to identify from an individual series and to optimize common parameters more accurately. In this research, we re-analyse M5 competition data, and assess the impact on accuracy of different implementations of statistical forecasting methods, taking into account cross-learning opportunities.

Effect of meteorological variables in electric load forecasting: Spanish insular systems

Presenter: Eduardo Caro
Co-authors: Eduardo Caro; Jesus Juan; Shadi Nouhitehrani

This work analyses the impact of meteorological variables (such as maximum temperature, minimum temperature, and solar radiation) over the electric load forecasting for Spanish insular systems. The demand forecasting method is based on an ARIMA time series model with a significant number of regressors which model the particularities of each day.In order to evaluate the methodology's performance, a real data set from the Spanish electricity market has been used and the developed algorithm has been tested employing the RMSE as the accuracy metric.

Impact of Temperature Modeling over The Electric Load Forecasting Accuracy

Presenter: Shadi Nouhitehrani
Co-authors: Shadi Nouhi Tehrani;

This work proposes alternative models for the temperature effect over the Spanish electric energy load, and they are implemented in hourly consumption forecasting software. Some case studies have been analyzed, such as: considering both maximum and minimum daily temperature, changing the location of weather stations, the inclusion of interaction between weekdays and weekends, among others. An exhaustive analysis is presented concerning both the estimation results and the forecasting accuracy.

Forecasting of load profiles by time series feature extraction

Presenter: Carlos Quesada
Co-authors: Carlos Quesada; Diego Casado-Mansilla; Pablo Montero-Manso; Noah Pflugradt; Cruz Enrique Borges

Energy System Models (ESMs) are tools used to project the future energy supply and demand of a region. ESMs help energy analysts, planners and policy makers to simulate and evaluate the long-term impacts of different energy scenarios. On the supply side, ESMs have become a common means of obtaining useful projections, since they use detailed white-box models. However, on the demand side, they lack the degree of accuracy required to adequately characterize, among others, the energy consumption at the household level, since they still use high-level black-box models. Indeed, factors such as the diversity of dwelling types, socioeconomic conditions of the family units, and behavioral-related consumption patterns pose a challenge for the characterization of the residential energy demand that cannot be easily addressed by traditional ESMs. As a first step in building a more precise model for the residential sector, an extensive analysis of existing data sets has been carried out in this work. In order to characterize the types of households that exist in terms of energy consumption habits, a segmentation of electrical load profiles has been carried out. Different data sets, obtained from smart meters from public and private sources, have been analyzed using time-series feature extraction. Thus, a set of time-series features that best describe the load profiles (such as statistical moments, quantiles, seasonal aggregates, load factors, autocorrelations, etc.) has been selected. Then, these features have been computed for all data sets, creating a feature space smaller than the original time-series space. Finally, a clustering has been performed in the feature space using unsupervised Machine Learning techniques. Different patterns of energy behavior can be distinguished from the resulting automatic classification. The presented work describes the methodology used to achieve the segmentation, introduces the early results obtained and discusses their implications, which will be an important step towards the construction of better residential load models.

Chaoticity versus Stochasticity on financial markets: Are daily SandP500 return dynamics chaotic?

Presenter: Markus Vogl
Co-authors: Markus Vogl; Peter Gordon Roetzel

In this study, we present a combinatory chaos analysis of daily wavelet-filtered (denoised) S&P500 returns (2000-2020) compared with respective surrogate data sets, Brownian motion returns and a Lorenz system realisation. We show the dynamics of the S&P500 return series to consist of an almost equally divided combination between stochastic- and deterministic chaos. The strange attractor of the S&P500 return system is graphically displayed via Takens embedding and by spectral embedding in combination with Laplacian Eigenmaps. For the field of nonlinear- and financial chaos research, we present a bibliometric-paired with citation network analysis. We critically discuss implications and future prospects.

CRPS-Learning

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

Combination and aggregation techniques can improve forecast accuracy substantially. This also holds for probabilistic forecasting methods where full predictive distributions are combined. There are several time-varying and adaptive weighting schemes like Bayesian model averaging (BMA). However, the performance of different forecasters may vary not only over time but also in parts of the distribution. So one may be more accurate in the center of the distributions, and other ones perform better in predicting the distribution's tails. Consequently, we introduce a new weighting procedure that considers both varying performance across

time and the distribution. We discuss pointwise online aggregation algorithms that optimize with respect to the continuous ranked probability score (CRPS). After analyzing the theoretical properties of a fully adaptive Bernstein online aggregation (BOA) method, we introduce smoothing procedures for pointwise CRPS learning. The properties are confirmed and discussed using simulation studies. Additionally, we illustrate the performance in a forecasting study for carbon markets. In detail, we predict the distribution of European emission allowance prices.

Measurement error sensitivity of loss functions for distribution forecasts

Presenter: Onno Kleen
Co-authors: Onno Kleen;

I examine the sensitivity of loss functions equivalently called scoring rules'for distribution forecasts in two dimensions: sensitivity to linear rescaling of the data and the influence of measurement error on the forecast evaluation outcome. First, I show that all commonly used scoring rules for distribution forecasts are robust to rescaling the data. Second, it is revealed that the forecast ranking based on the continuous ranked probability score is less sensitive to measurement error than the ranking based on the log score. The theoretical results are complemented by a simulation study aligned with frequently revised quarterly US GDP growth data and an empirical application forecasting realized variances of S&P 100 constituents. In line with its proven gross-error insensitivity, the ranking of the continuous ranked probability score is the most consistent between evaluations based on the true outcome and the observations with measurement error.

The Tensor Auto-Regressive Model

Presenter: Chelsey Hill
Co-authors: Chelsey Hill; James Li; Matthew Schneider; Martin Wells

We introduce the Tensor Auto-Regressive (TAR) model for modeling time series data which is robust to model misspecification, seasonality and non-linear trends. We develop a parameter estimation algorithm for the proposed model by using the ð ‘¡-product, which allows us to model a three-dimensional block of parameters. We use the Fast-Fourier Transform, which allows for efficient and parallelizable computations. We use a combination of simulated data and an empirical application to: (i) validate the model, including seasonal and geometric trends, model misspecification analysis and bootstrapping to compute standard errors; (ii) present model selection results; and (iii) demonstrate the performance of the proposed model against benchmarking and competitive forecasting methods. Our results indicate that our model performs well against comparable methods, is robust and computationally efficient.

Superiority of simple average to equally weighted, variance-covariance and OLS combination forecasts

Presenter: Sevket Gunter
Co-authors: Sevket Gunter; Celal Aksu

We develop the expected ex-ante risk of simple average (SA) combination forecasts under a squared-error loss criterion and provide new analytical insights into their superiority to equally weighted, variance-covariance, and OLS combinations in existing analytical, empirical and simulation studies. Sufficient and necessary conditions under which the ex-ante risk of SA combination forecasts dominate those of these methods are first derived. The ranges of the hypothesis specification errors within which the SA combinations are expected to dominate these more sophisticated combinations are then analytically characterized. All research we examined that assumes stationarity and exogenous forecasts were either corroborated, or exactly

matched, or generalized using our risk functions framework, except for two specific existing findings which were negated by our results. We also found that whether the SA combination will dominate these more sophisticated combination methods additionally depends on the level of noise of the combination model itself, the sheer magnitude of the elements of the SSCP matrix of the forecasts, and the number of times a combination model will be repeatedly used without being re-estimated.

Autoregressive Denoising Diffusion Models for Multivariate Probabilistic Time Series Forecasting

Presenter: Kashif Rasul
Co-authors: Kashif Rasul; Calvin Seward; Ingmar Schuster; Roland Vollgraf

In this work, we propose TimeGrad, an autoregressive model for multivariate probabilistic time series forecasting which samples from the data distribution at each time step by estimating its gradient. To this end, we use diffusion probabilistic models, a class of latent variable models closely connected to score matching and energy-based methods. Our model learns gradients by optimizing a variational bound on the data likelihood and at inference time converts white noise into a sample of the distribution of interest through a Markov chain using Langevin sampling. We demonstrate experimentally that the proposed autoregressive denoising diffusion model is the new state-of-the-art multivariate probabilistic forecasting method on real-world data sets with thousands of correlated dimensions. We hope that this method is a useful tool for practitioners and lays the foundation for future research in this area.

Boosting the Wisdom of Crowds Within a Single Judgment Problem: Weighted Averaging Based on Peer Predictions

Presenter: Asa Palley
Co-authors: Asa Palley; Ville Satopaa

Combining point estimates from multiple judges often provides a more accurate aggregate estimate than using a point estimate from a single judge, a phenomenon called "the wisdom of crowds." However, if the judges use shared information when forming their estimates, the simple average will end up over-emphasizing this common component at the expense of the judges' private information. A decision maker could in theory obtain a more accurate estimate by appropriately combining all information in the judges' opinions. Although this information is embedded within the judges' individual estimates, it is typically unobservable and thus cannot be directly aggregated by a decision maker. In this article, we propose a weighting of judges' individual estimates that appropriately combines their collective information within a single estimation problem. Judges are asked to provide both a point estimate of the quantity of interest and a prediction of the average estimate that will be given by all other judges. Predictions of others are then used as part of a criterion to determine weights that are applied to each judge's estimate to form an aggregate estimate. Our weighting procedure is robust to noise in the judges' responses and can be expressed in closed form. We use both simulation and data from six experimental studies to illustrate that our procedure outperforms existing averaging-like methods.

Decomposing the Effects of Crowd-Wisdom Aggregators: The Bias-Information-Noise (BIN) Model

Presenter: Ville Satopää
Co-authors: Ville Satopaa; Marat Salikhov

Aggregating predictions from multiple judges often yields more accurate predictions than relying on a single judge: the "wisdom-of-the-crowd" effect. This aggregation can be conducted by different methods,

from simple averaging to complex techniques, like Bayesian estimators and prediction markets. This work applies a broad set of aggregation methods to subjective probability estimates from a series of geopolitical forecasting tournaments. It then uses the Bias-Information-Noise (BIN) model to disentangle three mechanisms by which each aggregation method improves accuracy: the tamping down of bias and noise and the extraction of valid information across forecasters. Averaging works almost entirely via noise reduction whereas more complex techniques, like prediction markets and Bayesian aggregators, work via all three BIN pathways: better signal extraction and noise and bias reduction.

Election forecasts can be improved using wisdom-of-crowds methods

Presenter: Henrik Olsson
Co-authors: Henrik Olsson; Wandi Bruine de Bruin

We show that election forecasts in 2018 and 2020 US elections are improved by adding wisdom-of-crowds questions to election polls. Wisdom-of-crowds questions ask poll participants about the percentage of their social contacts (social-circle question) or people in their state (state-winner question) who might vote for different candidates. These questions can be used to forecast elections on their own. They can also be integrated with responses to traditional poll questions that ask participants who they themselves will vote for (own-intentions question). We explore a new Bayesian bootstrap method for election forecasts that combines traditional polling questions about people's own intentions with their expectations about how others will vote. We compare different methods on national probabilistic samples (N=4000+) within the USC Understanding America Study panel. The bootstrap forecast outperforms aggregate national polls in the 2020 U.S. election, as well as the forecasts based on traditional polling questions posed on large national probabilistic samples before the 2018 and 2020 U.S. elections. The bootstrap forecast puts most weight on people's expectations about how their social contacts will vote, which might incorporate information about voters who are difficult to reach or who hide their true intentions. Beyond election polling, the new method is expected to improve the validity of other social science surveys.

Combining Aggregated Judgments: An Effective Method for Improving Accuracy by Stacking Multiple Weighting Models

Presenter: Shu Huang
Co-authors: Shu Huang; Stephen Broomell

Decision makers often seek a consensus prediction for an uncertain quantity of interest by aggregating judgments from a crowd of forecasters, taking advantage of the "wisdom of crowds." A variety of aggregation procedures (e.g., simple averaging, sub-crowd selection, and optimized weighted averaging) have been proposed in this literature to estimate weights that should be placed on each individual forecast to improve accuracy (e.g., to minimize mean squared error). However, it's challenging to find a single weighting method that outperforms others across different data contexts. We consider applying the wisdom of crowds to the problem of determining how to extract the wisdom of a crowd. We propose an ensemble weighting method based on the stacking technique to combine multiple different weighting schemes to improve the accuracy of the aggregate forecast. Applying our stacking method to three real-world datasets, we find that it outperforms the component weighting methods in most situations. We conclude that stacking multiple weighting models is an empirically validated, effective way to harness the wisdom of crowds.

Online short-term forecasting of aggregated heat load

Presenter: Li Bai
Co-authors: Li Bai; Pierre Pinson

Day-ahead short-term heat load forecasting for distribution heating systems is indispensable for optimizing operations. Heat load exhibits a nonlinear and non-stationary behaviour, which limits the predictive capabilities of the conventional off-the-shelf approaches that relied on batch learning. Instead, online learning approaches can continuously adapt and accommodate the time-varying properties in the heat load, as well as its dependencies on relevant exogenous processes. They are additionally computationally efficient in view of operational requirements. This allows considering a number of additional features to enhance their performance, apart from the commonly used meteorological features, e.g., outdoor temperature. For example, the stochastic characteristics of the heat load are highly affected by social behaviours of different types of consumers. Consequently, dedicated features that are consumer-specific ought to be accommodated within relevant online-learning approaches, as a basis to forecasting.In our work, we propose an approach to online short-term probabilistic forecasting of the heat load in a distribution heating system, considering feature extractions for individual types of users. The heat load is an aggregated load for over 4000 users in the same distribution heating system, including single homes and commercial buildings. The approach relies on quantile regression after relevant transformation of the heat load, to accommodate its non-Gaussian characteristics. The approach is benchmarked against a number of relevant naÃ¯ve approaches, e.g., probabilistic persistence, similar-day approach, etc. within a rigorous probabilistic forecast verification framework. This includes diagnostic tools like reliability diagrams, skill scores like the log-score and the CRPS, etc.

Predicting Energy Production by HPP Using Machine Learning Algorithms with Priority Weights

Presenter: Eralda Gjika
Co-authors: Eralda Gjika Dhamo; Enzo D.Lamberti; Lule Basha Hallaci

In this work we are testing machine-learning algorithms to help understanding and predicting the monthly energy production by HPP. Our target variable is the monthly energy produced in river cascade with 3 important hydropower's (Drin cascade in Albania) and considered variables are: average temperature, average precipitations, water inflow. Our results show that machine learning algorithms are adequate models which outperform classical time series models such as: naÃ¯ve methods, ARIMA, exponential smoothing etc. To further improve the forecast we have also developed a hybrid model by combining the models with high accuracy performance on predicting monthly energy produced in the cascade. Predictive performance of the model is evaluated by several measures such as: MSE, RMSE, MAPE, AIC, BIC etc. Linear correlation is used as another measure of goodness of fit between real time series and the fitted time series obtained from the models. Using a threshold of the correlation we select those models which predictions are highly correlated with the real observation. The proposed hybrid model is a combination of these models. The weights are estimated automatically from an algorithm which use least square support vector machine (LSSVM) model to give priority weights to those models showing high forecasting accuracy. In terms of accuracy and bias, our approach shows to exceed the performance of classical time series models and is comparable with machine learning models.

Outlier detection in water related demand-driven time series data based on a clustering approach

Presenter: Jens Kley-Holsteg
Co-authors: Jens Kley-Holsteg; Florian Ziel

The water sector is gathering vast amounts of data, needed for manual and automatic decision making with an increasing focus on near real-time applications. To enable reliable decision making and smooth operations of infrastructure and plants the data must be consistent and trustworthy. Unfortunately, this cannot always be guaranteed in practice as technical failures in the measurement and transmission process regularly occur. Hence, data pre-processing and in particular outlier handling are essential tasks before the

data can be processed in more specialized fields as forecasting. In this research a model-based clustering algorithm is introduced, which is based on finite Gaussian mixture modelling. By applying features tailored to demand related time series data, the proposed approach can detect complex anomalous structures by utilizing a pure data-driven method. Moreover, as the approach is based on statistical models the probability of each observation being an outlier can be revealed. Likewise, the cause for the assignment to the outlier cluster in terms of a specific feature or a combination of several features can be traced back. From a practical perspective these are valuable properties as collected process data is usually not labelled and the unambiguous identification of outliers caused by a technical failure or by an extreme but truly observable behavioural issue is often not easily retrievable. So, beside identifying outliers, the approach can reproduce the outlier assignment procedure to provide transparent and traceable results. To handle unlabelled and complex historical data sets but also data accruing in real-time, the procedure is optimized in terms of computational time and robustness.

Probabilistic Intraday Wastewater Treatment Plant Inflow Forecast Utilizing Rain Forecast Data and External Flow Measurements

Presenter: Björn Sonnenschein
Co-authors: Björn Sonnenschein; Florian Ziel

Increasing availability of internal and external data allows to furtherimprove the efficiency of waste water treatment plants. In recent years,forecasts of the waste water inflow into a waste water treatment plantproved to be of high value to practitioners, enabling smoother operationof treatment processes. However, it has been shown that especially theprediction of days with rainfall is still challenging. Here, theinclusion of external data sources such as weather forecasts and theconsideration of data from sensors inside the sewer system showed to bepromising to further improve the accuracy and reliability of forecasts.However, in order to improve the forecasting performance, dealing withthe efficient usage of the increasing amount of data, the modelcomplexity and non-linear effects, caused by the sewer network'shydrodynamic properties, is necessary.In this research, a seasonal probabilistic time series model is proposedfor modelling the intraday waste water inflow accurately, especially incase of rain events. Non-linear effects of the exogenous variables areapproximated by piecewise linear functions, accounting for the flow-ratedependent temporal correlation structure of precipitation quantity,external flow measurements and the inflow at the plant. Due to theresulting high-dimensional parameter space Lasso is used for featureselection and regularization during parameter estimation. Forforecasting performance evaluation the continuous ranked probabilityscore is utilized.The model is applied to quarter-hourly data from two German wastewatertreatment plants of different size in order to assess and demonstratepractical suitability. Non-linearities arising in the sewer system arereasonably well approximated, resulting in a reasonable improvement inforecast accuracy, whereas the regularization allows to deal with hugefeature space efficiently. Uncertainty quantification is provided forthe full predictive distribution.

When (where and why) forecasters get it wrong? An analysis of growth forecast errors

Presenter: Umberto Collodel
Co-authors: Umberto Collodel; Prakash Loungani; Zidong An;

We construct a database that contains short-term growth projections from all major institutions and private sector and examine the quality of these forecasts for a large sample of countries over the last three decades. Our findings are three-fold. First, the inability to forecast recessions and in particular financial crises determine the optimism usually ascribed by news agencies and governments alike to growth forecasts. Second, different forecasts resemble each other, exhibiting a high degree of collinearity. Third, IMF-supported program forecasts tend to be more optimistic for large programs. Nevertheless, a comparison with Consensus forecasts thwarts the hypothesis of a deliberate bias. Our findings suggest that some caution

is in order when consulting growth forecasts and have first-order implications for policymakers during times of global and regional turmoil.

Combining Bayesian VARs with survey density forecasts: does it pay off?

Presenter: Federica Brenna
Co-authors: Federica Brenna; Marta Banbura; Joan Paredes; Francesco Ravazzolo

This paper studies how to combine real-time forecasts from a broad range of Bayesian vector autoregression (BVAR) specifications and survey forecasts by optimally exploiting their properties. To do that, it compares the forecasting performance of optimal pooling and tilting techniques, including survey forecasts for predicting euro area inflation and GDP growth at medium-term forecast horizons using both univariate and multivariate forecasting metrics. Results show that the Survey of Professional Forecasters (SPF) provides good point forecast performance, but also that SPF forecasts perform poorly in terms of densities for all variables and horizons. Accordingly, when the model combination or the individual models are tilted to SPF's first moments, point accuracy and calibration improve, whereas they worsen when SPF's second moments are included. We conclude that judgement incorporated in survey forecasts can considerably increase model forecasts accuracy, however, the way and the extent to which it is incorporated matters.

Entropic tilting for macroeconomic variables: The role of asymmetrically distributed survey forecasts

Presenter: Anastasia Allayioti
Co-authors: Anastasia Allayioti;

There is a growing interest in incorporating external information extracted from a survey of professional forecasters into real-time macroeconomic predictions from vector autoregressive (VAR) specifications. The method of entropic tilting achieves this by modifying the baseline VAR distribution such that it matches certain moment conditions. Existing papers adopting this methodology focus on the first two moments of aggregate survey forecast distributions, thereby restricting their attention to symmetric environments. We propose a modification to the standard relative entropy approach which allows for asymmetry in the macroeconomic variables and explores the predictive content of higher-order moments. The proposed methodology involves tilting the VAR distribution towards aggregate survey projections that have been appropriately reshaped to match the non-Gaussian features of the sample data. We illustrate this methodology with an application examining real-time forecasts for four U.S. macroeconomic variables. We consider a variety of VAR models, ranging from time-varying volatility to non-Gaussian errors. Results across models indicate meaningful gains in terms of both point and density forecast accuracy relative to individual multivariate specifications and existing forecasting methods that blend model-based forecasts with external judgement.

Accelerate Open Source Forecasting with SAS (Part 1)

Presenter: Jessica Curtis
Co-authors: Jessica Curtis;

Forecasting is core to solving many different business challenges across virtually every industry. The impact of having an efficient and accurate forecasting process is far reaching and fundamental for better business decisions throughout any organization. Open source is widely used to initiate forecasting projects. Many organizations begin with an open source strategy, leveraging Python or R to build forecasts, and face challenges when trying to scale across many different forecasting use cases across the enterprise. Join us to learn the benefits of running open source forecast models with SAS Visual Forecasting, which

allows organizations to build upon existing open source strategies in an agile, efficient way. SAS offers a distributed, scalable, and resilient way to run open source forecasting projects in an open ecosystem. In every step of the analytics life cycle, from data preparation to model development to model deployment, SAS accelerates open source forecasts by automatically distributing the workload. You no longer have to choose between SAS and open source ”“ it is truly a complementary relationship.

Accelerate Open Source Forecasting with SAS (Part 2)

Presenter: Andrea Moore
Co-authors: Andrea Moore;

Join us to experience SAS open source integration in action. Get a hands-on view of the two main steps for integrating open source forecasting models into SAS Visual Forecasting, each with increasing benefits and value:1. Laying the Foundation: Accelerate Open Source Forecasts Programmatically2. Sharing the Wealth: Open Source Forecasts as a Custom Modeling NodeWhether you take the programmatic approach or build a custom modeling node within the Visual Forecasting UI, SAS enables scalable, resilient forecasting for transitioning from small project to enterprise solution.After this session, your organization will be able to:Â · Automate manual processesÂ · Accelerate open source run timeÂ · Augment with built-in best practicesÂ · Scale to enterprise use casesÂ · Enable an interactive, collaborative workflow

Deep Learning for Retail Sales Forecasting

Presenter: Szymon Haponiuk
Co-authors: Szymon Haponiuk; Michal Kurcewicz

This empirical paper discusses the use of deep learning techniques of retail sales forecasting. Compared to classical time series methods, deep learning models can fully utilize rich data sets available to modern retailers. Using data from a major Polish retailer we have developed short-term demand forecasting models for four product categories exhibiting different demand characteristics. We present a complete forecasting pipeline implemented in a mixed SAS and open-source environments. We discuss the network architectures and compare the forecasting performance of deep networks with time series and standard machine learning techniques. Preliminary results indicate that deep learning models perform significantly better than competing model types. The improvement in forecast quality is especially evident for slow moving items. We also discuss model stability, hyperparameter tuning and steps taken to avoid over-fitting.

Major Paradigm Shifts in Modern Forecasting Methodology

Presenter: Russ Wolfinger
Co-authors: Russ Wolfinger;

Driven by strong and extensive contributions from the open source community, classic time series and econometric forecasting methods have been supplanted by gradient boosted trees and neural networks in terms of performance and flexibility in handling large, complex, and messy data sets. We highlight two examples from Kaggle competitions: COVID Global Forecasting and M5 Uncertainty and provide intuition and tips on why the new approaches are so compelling.

Temporal Disaggregation of U.S. State Natural Gas Data

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

Current daily natural gas demand forecasting is focused on the local distribution company (LDC) level. But problems with the natural gas infrastructure are not confined by LDC boundaries. Outages occurring in one part of the country may influence supplies for LDCs in other parts pf the country as seen by the outages in Texas during the extreme cold event in February 2021. Predicting the demand across LDCs requires state level data, which is available from the U.S. Energy Information Administration (EIA). The EIA publishes monthly data for all U.S. states. However, daily data is needed for accurate and timely forecasts. In this paper, we focus on disaggregating state data from monthly to daily. This is done by using daily weather data and monthly natural gas consumption data. Our disaggregation method uses multi-parameter regression with prior day weather adjustments. This method achieves a 17.8% MAPE in disaggregating monthly to daily data. This state-level approach to forecasting natural gas demand helps address the uncertainties of extreme events.

Progress in the Quantitative Analysis of the Field of Forecasting (1985-2021)

Presenter: Claudio Antonini
Co-authors: Claudio Antonini;

The 3,023 files of the International Journal of Forecasting in terms of papers, letters and replies, reviews of books, editorial comments and announcements, list of contributors and other types of articles, provide an unsurpassed repository of information that allow to determine trends in the field of forecasting. A quantitative analysis of such a database will allow to enlighten qualitative discussions about the past, the present, and the future of the field.We will discuss progress in the analysis of the database and preliminary conclusions, such as (a) topics that virtually disappeared or are in fashion, (b) bubbles, and (c) a recent tendency to concentrate on technical aspects (as an end in itself) with a progressive decrease of the use of terms that indicated that the original purpose of the discipline was to support managerial decisions. This last aspect can be understood in various ways, one of which is that the concept of forecasting in industry is already mature and accepted, and, thus, it is not necessary to explain its purpose anymore.

Time Series Disaggregation

Presenter: Thomas Willemain
Co-authors: Thomas Willemain; Adam Petrie

Imagine a discrete event simulation that requires daily time series data as an input. If you only have data bucketed monthly, how do you proceed? Converting monthly data to daily data is an example of the time series disaggregation problem. We discuss current (inadequate) solutions to this problem and present a (less inadequate) solution that suffices for some practical instances.

Identifying the Global Transmission of the COVID-19 Pandemic on G-20 Economies

Presenter: Brian Sloboda
Co-authors: Brian Sloboda; Rolando Santos

We use a Global Vector Auto-Regression (GVAR) model featuring the G-20 economies to investigate the factors behind the dynamics of the effects of the COVID-19 pandemic in these economies. The GVAR approach enables us to make two key contributions: first, to model international linkages among many countries, which is a key asset given the diversity of countries and regions, and second, to analyze the impacts of the COVID-19 pandemic. The latter proves to be very important due to the internationalization of production chains. The model can be used to gauge the effect on the macroeconomics of various scenarios, such as an output shock in the United States, output shocks to foreign nations that are members of the

G-20. After estimating the GVAR model, we generate 12 one-quarter-ahead forecasts for the next quarter including real GDP, inflation, short-term interest rates, and other macroeconomic variables over the period 1995M1"“2021M.

The impact of Covid-19 on vector autoregressions:

Presenter: Florens Odendahl
Co-authors: Florens Odendahl; Luis Alvarez

The Covid-19 pandemic has led to large outliers in macroeconomic data that distort the estimation of macroeconometric models. We propose a method to deal with these outliers in the estimation of a Bayesian VAR (BVAR). Our method leaves the contemporaneous correlation of the reduced form error term unchanged, allows for different outlier magnitudes across variables, and known or unknown outlier dates. In an application with euro area data, we find that our specification leads to posterior distributions that are robust to the Covid-19 outliers. Morevoer, we find that while the BVAR produces reasonable forecasts, they are far from being competitive to survey or institutional predictions during the pandemic.

Cancelled: Comparing Global VAR with alternative macro models for forecasting and scenario analysis

Presenter: Jeremy Kwok
Co-authors: Jeremy Kwok;

Macroeconometric models such as Global Vector Autoregresive (GVAR), Factor-Augmented VAR (FAVAR) and Dynamic stochastic general equilibrium (DSGE) models are often constructed for analysing monetary policy shocks. However, the rationale behind the modelling is completely different. This paper aims to investigate how GVAR fares against other macro models. Particular interest is in forecasting and scenario analysis. This paper compares the forecasting ability of GVAR and also shock response from impulse response functions (IRFs) by FAVAR and DSGE. For the forecasting exercise, the ability is compared between a generic AR model with GVAR ex-ante and GVAR-ex post forecasts. For the scenario analysis, IRFs were constructed from GVAR, FAVAR and DSGE models with various shocks. It is easy to see that certain properties are similar among the models such as the long run appears to be unaffected by a monetary shock or that the GDP is negatively affected by it. However, there are also a lot of discrepancies in the short run, particularly in the first 4 quarters. From this, we can conclude that the GVAR model fares best in forecasting that it explicitly allows error correction mechanisms among country models, this is reflected by the dynamic responses from each economy. On the other hand, the FAVAR results look more uniform in their values and shape. The comparison was also made with DSGE IRFs and shows that there is a certain consensus among the theory-driven versus the data-driven models. In contrast with forecasting, the scenario analysis provided by IRFs cannot be evaluated against real-world events. There is no ‘true’ model to speak of compared to the true values in the forecasting application. Consequently, the IRFs inform us more about the underlying methodology and assumption of the models themselves than can be used to evaluate their accuracies. The paper concludes that the GVAR model is quite adaptable in terms of allowing the data to dictate the short run but also relying on more theory-led identification for the long run.

Tracking economic growth during the Covid-19: a weekly indicator for Italy

Presenter: Simone Emiliozzi
Co-authors: Simone Emiliozzi; Davide Delle Monache; Andrea Nobili

Following the breakout of the Covid-19 pandemic, economic forecasting has become more complex.

One way to address these new challenges is to exploit the information content of high frequency variables to construct a synthetic and timely indicator of the business cycle. Using data reduction techniques in a mixed-frequency framework, we develop an Italian Weekly Economic Index (ITWEI), which proves to be particularly useful for forecasting and policy analysis during the pandemic period. We are working on a Bayesian extension to fully take into account filter and parameter uncertainty.

Nowcasting the economy with news during the pandemic

Presenter: Luca Barbaglia
Co-authors: Luca Barbaglia; Sebastiano Manzan; Sergio Consoli

We evaluate the informational content of news-based sentiment indicators at forecastingthe Gross Domestic Product of the major European economies. We build aspect-basedsentiment indicators about various aspects of the economic activity and policy, and assesstheir added value at forecasting next to survey-based sentiment indicators. Our data setincludes 26 large newspapers, for a total of over 27 million articles. Our

findings outlinethe complementarity of news-based sentiment indicators to surveys at forecasting horizons-both in an in-sample and out-of-sample setting. We employ the proposed indicators tonowcast during the Covid-19 pandemic and show their added values to forecast the recoverypatterns.

Now- and Backcasting Initial Claims with High-Dimensional Daily Internet Search-Volume Data

Presenter: Erik Christian Montes Schütte
Co-authors: Erik Christian Montes Schütte; Daniel Borup; David E.Rapach

We generate a sequence of now- and backcasts of weekly unemployment insurance initial claims (UI) based on a rich trove of daily Google Trends (GT) search-volume data for terms related to unemployment. To harness the information in a high-dimensional set of daily GT terms, we estimate predictive models using machine-learning techniques in a mixed-frequency framework. In a simulated out-of-sample exercise, now- and backcasts of weekly UI that incorporate the information in the daily GT terms substantially outperform models that ignore the information. The relevance of GT terms for predicting UI is strongly linked to the COVID-19 crisis.

Textual Data for Time Series Forecasting

Presenter: David Obst
Co-authors: David Obst; Badih Ghattas; Sandra Claudel; Jairo Cugliari; Yannig Goude; Georges Oppenheim

While ubiquitous, textual sources of information such as company reports, social media posts, etc. are hardly included in prediction algorithms for time series, despite the relevant information they may contain. In our work, openly accessible daily written weather reports from France and the United Kingdom as well as tweets are leveraged for electricity consumption prediction. Using exclusively the weather reports, we are able to predict the load time series with sufficient accuracy to be used to replace missing data. Furthermore, the Twitter data pertaining to specific keywords helped to adapt to the new patterns of electricity consumption observed during the COVID-19 lockdown period in France, significantly reducing forecast errors of operational models.

Academicians and Practitioners Will Thrive Under our New ML Master

Presenter: Lawrence Vanston
Co-authors: Lawrence Vanston;

The sweep of the M5 competition by machine learning (ML) has verified that we have a classic substitution on our hands where a superior and rapidly improving new technology, ML, is displacing an older technology, statistical forecasting. It's a classic story: folks in the old technology are forced to adapt to survive. I believe we will not only survive, but thrive.Practitioners will find the most successful ML tools embedded in off-the-shelf systems. Better tools mean better business. Applied academicians will thrive because of ML's ability to model systems and capture relationships among variables in ways unimaginable before. Predictive power becomes prescriptive power. Theoretical academicians will find plenty of interesting ML problems to apply their genius to. We also note some of the constraints on ML adoption and suggest that those who help overcome them may thrive the most.

Using machine learning in building a sentiment analysis tool to forecast risk appetite

Presenter: Petrus Potgieter
Co-authors: Petrus Potgieter; Stuart Royden-Turner

This study is an analysis of the techniques used in quantifying sentiment (valency) for use in forecasting variables of interest in asset management. The neoclassical growth model is an important tenet, where variants, such as stochastic extensions, are used in public finance, contemporary finance and explaining business cycle theory. Evidence points to consumer expectations and sentiment in financial market that are extrapolative, showing a change in attitude or risk appetite in the run up to a financial crisis. Consumer sentiment measures have been augmented by the recent surge of accessible sources of data in financial services, termed *alternative data sources*. Data streams include free popular news sources (in our review this includes the Guardian in the UK and the New York Times), social media platforms (including Twitter) and internet site sentiment measures generated by Google's GDELT. We make use of alternative data with powerful unstructured data processing techniques, coupled with sentiment valence techniques ranging from simple text-based frequency based counting methods, to lexical dictionary-based methods and finally expanding to machine learning techniques to create sentiment indices.In a second step, we review the power of the advanced techniques against different target variables that are more specifically used in asset management. In this study we are largely focussed on the credit spread and the respective underlying variables driving this spread, including default rates, interest rates, inflation and cash. We find that the increased frequency of data, when used in conjunction with advanced modelling techniques provide promising results and overcome challenges in traditional statistical techniques and those covered by high levels of expert logic.

A picture is worth a thousand data points: An image-based time series forecasting approach

Presenter: Artemios-Anargyros Semenoglou
Co-authors: Artemios-Anargyros Semenoglou; Evangelos Spiliotis; Vassilios Assimakopoulos

Deep Learning (DL) has become standard practice in several scientific fields, including computer vision and natural language processing, among others. Nevertheless, DL has been adopted at a much slower pace in the area of forecasting, despite the strong evidence available in favour of its use. Moreover, existing DL forecasting approaches handle time series data as numeric vectors, thus not leveraging directly and to their full extent recent advances in DL. Motivated by these insights, we introduce a novel DL forecasting approach that transforms the traditional task of time series forecasting into a computer vision

one. First, instead of using numeric representations as input to our forecasting models, we consider visual representations of them in the form of images. Second, deep convolutional neural networks, inspired by popular image recognition neural network architectures (e.g. ResNet) are tasked with producing point forecasts from the aforementioned visual representations. We evaluate the forecasting accuracy of the proposed approach using data from the M4 forecasting competition. Our results indicate that image-based time series forecasting methods outperform traditional ones, both of statistical and machine learning nature.

Size does matter: Time series augmentation for enhanced cross-learning

Presenter: Evangelos Spiliotis
Co-authors: Evangelos Spiliotis; Artemios-Anargyros Semenoglou; Vassilios Assimakopoulos

Cross-learning, i.e., training models using data of multiple time series instead of the single series being forecast, has been proven to be an effective strategy for improving the forecasting accuracy of machine learning methods and especially neural networks that are "data-hungry" in nature. This was also one of the major findings of the M4 and M5 forecasting competitions where all top performing methods employed a form of cross-learning. However, in many cases the series originally available for training may be insufficient in number, limiting the potential benefits of cross-learning. To deal with this issue, augmentation techniques can be used to artificially increase the size of the train set available and allow time series patterns to be effectively learned. In this study we examine several time series augmentation techniques, both well-established and new ones, and evaluate their contribution to forecasting accuracy considering a feed-forward neural network as a baseline model. In order to link the impact of augmentation techniques with the size of the train set originally available, we also consider different pools of series of small, moderate, and large size. Our results suggest that time series augmentation can help us enhance the performance of cross-learning methods, especially when the size of the original train set is relatively small. However, we also find that accuracy improvements depend strongly on the augmentation technique used.

Location Network Analysis and Supervised Machine Learning Models to Identify Virus Spread Trends in the COVID-19 Pandemic

Presenter: Carlos Pinheiro
Co-authors: Carlos Pinheiro;

The recent outbreak of coronavirus disease 2019 (COVID-19) has infected and killed several thousands of people globally, resulting in a pandemic with enormous global impact. This disease affects the respiratory system, and the virus that causes it, SARS-CoV-2, spreads through droplets of saliva, coughs, and sneezes. As a disease caused by an extremely transmissible viral infection, COVID-19 is inflicting significant damage on the economies of both developed and lower- and middle-income countries (LMIC) because of its direct impact on the health of citizens and because the measures taken to contain the infection can lead to economic damage. Methods to reduce or control the spread of the virus and protect the global population are needed in order to avoid further deaths, to prevent long-term health issues and prolonged economic harm, and to provide time until vaccines are widely available and treatments are developed. The most effective approach to reduce viral spread and avoid a substantial collapse of the health care system is nonpharmaceutical interventions (NPI), such as enforcing social containment policies, monitoring overall population mobility, implementing widespread viral testing, and increasing hygiene measures. Our approach consists of combining network analytics with machine learning models by using a combination of health data and anonymized telecommunications data to better understand the correlation between population movements and virus spread. This approach, known as location network analysis, allows for accurate prediction of possible new outbreaks. It gives governments and health authorities a crucial tool that can help define more accurate public health metrics and can be used either to ease social containment policies to reopen the economy or to intensify them to avoid further spread.

Representing and Forecasting the COVID-19 Pandemic Using Differential Equation Models

Presenter: Marc Kessler
Co-authors: Marc Kessler;

The COVID-19 pandemic has put unprecedented demands on hospital resources such as personal protective equipment, intensive care unit (ICU) beds, and ventilators. The scarcity of these health resources together with their often slow supply chains has created a need for accurate, easily understood models that can forecast COVID-19 infection rates. In this paper, a susceptible, exposed, infectious, and recovered (SEIR) differential equation model is developed. This SEIR model includes a flexible parameterization of the reproduction number, R0, to account for interventions such as social distancing and mask wearing that vary in their participation rates over time. Additionally, this SEIR model incorporates vaccine data to improve forecasts of how long the pandemic will last. The use of intuitive intervention and vaccine efficacy parameters in this model in conjunction with estimates of these parameters using observed infection rates has allowed health services organizations like the Cleveland Clinic to explore a range of likely and hypothetical scenarios for health resource use.

Visualization by pattern similarity for Covid-19 data set

Presenter: Youngjin Park
Co-authors: Youngjin Park; Mahesh Joshi

COVID-19 data are available for 183 countries at the country level in a daily format. Among these countries, eight countries have state-level data. We use a pattern similarity measure to create groups of similar countries and to create groups of similar states within a few countries. To illustrate our ideas, we focus on daily confirmed cases as our target. We present some characteristics of the data and methods of smoothing the data. Because the data set is constantly growing larger, we use some similarity measures to identify several representative series to focus on in a subsequent modeling step. Given two ordered numeric sequences, similarity measures can be computed between an input sequence and a target sequence that "slide" the target sequence with respect to the input sequence. The "slides" can be measured by an observation index (sliding-sequence similarity measures) or a seasonal index (seasonal-sliding-sequence similarity measures). After clustering, vector time series plots and series plots show whether clustering by pattern similarity works well. Another clustering outcome is that we can define new characteristics in each group of the country-level cluster. From state-level clustering, we can determine which state pattern is most like its country pattern. We also explore how the groups change as the number of clusters increases by using a Sankey diagram, which visualizes the change in group membership for different numbers of clusters.

Evolution of Statistical Models for Producing Weekly COVID-19 Forecast

Presenter: Ran Bi
Co-authors: Ran Bi; Youngjin Park; Mahesh Joshi

We have witnessed a SARS-CoV-2 pandemic that has spread globally in the last 15 months and has significantly affected people's lives. Many insightful research projects contributed to COVID-19 forecasting at last year's ISF. The advantages and disadvantages of susceptible, exposed, infectious, and recovered (SEIR) models and statistical models were evaluated by Rahul Pathak (2020). Mechanistic models such as SEIR can outperform statistical models in the early stage of a pandemic when no information on transmission or past trends is available. By contrast, statistical models need more data and could make better predictions in the short term. The Centers for Disease Control (CDC) and Johns Hopkins University collect and publish data in detail every day, enabling researchers to create their statistical model forecasts in real-time and investigate the evolution of models over time. This paper examines the evolution process

of statistical models, including ARIMAX models, exponential smoothing models (ESMs), and unobserved component models (UCMs), over the past several months. We build an automated forecasting system that can choose the champion model for each time series for a given history region. We predict confirmed cases and fatalities for each model and evaluate the evolution every week. We compare the forecasts to the actual confirmed cases and fatalities and calculate a range of error measures, such as pinball loss, MAPE, MASE, and RMSE, in order to choose the best model for time series over different time periods. All models are wrong, but some are useful (Box 1976). Our work evaluates the value of the forecasting process that monitors how the models evolve over time and whether they capture the effects of external factors in time.

Surrogate Monte Carlo

Presenter: A. Christian Silva
Co-authors: A. Christian Silva; Fernando Ferreira

This article proposes an artificial data generating algorithm that is simple and easy to customize. The fundamental concept is to perform random permutation of Monte Carlo generated random numbers which conform to the unconditional probability distribution of the original real time series. Similar to constraint surrogate methods, random permutations are only accepted if a given objective function is minimized. The objective function is selected in order to describe the most important features of the stochastic process. The algorithm is demonstrated by producing simulated log-returns of the S&P 500 stock index.

MODELLING FINANCIAL MARKET TIME SERIES VIA MACHINE LEARNING AND STATISTICAL BASED MODELS; A COMPARISON STUDY

Presenter: Andressa Contarato
Co-authors: Andressa Contarato; Pedro Ivo Rodrigues; Marco Sanfins; Reinaldo Castro Souza

Among financial institutions it is common to analyze resource values aiming to obtainfuture results in prediction analysis. There are several methodologies from descriptive analysis and moving averages implemented in spreadsheets to powerful models.These models are fundamental when it comes to creating diverse and complex systems which help to simulate a range of investment scenarios and thereby choose the best among them. With Big Data and investments in technological improvement, the application of models from Machine Learning techniques became more feasible. These techniques have a high computational power to obtain the most accurate analysis and forecasts of financial market behavior. The objective of this work is to compare traditional based models using Time Series, such as Auto-regressive and Moving Averages (ARIMA) and GARCH, to Machine Learning models, namely: Support Vector Regression (SVR) and ArtificialNeural Network (ANN). These models have been applied to some financial assets and tokens. The results showed that for both types of resources, machine learning models had the best performances in financial series close values, but with different highlights for SVR and ANN respectively. However, the ARIMA + GARCH model resulted in good models also when a GHskew was included in the waste modeling. It was observed that the precision level of the prediction of the SVR and ANNem models in comparison with the ARIMA model are quite proven.

Mapping ABM of Financial Markets: A bibliometric and citation network analysis

Presenter: Conceicao Xavier
Co-authors: Conceicao de Castro;

An increasing number of financial investors base their investment decisions on computational analysis. Agent-based modeling (ABM) is an approach to model systems comprised of individual and how behaviors affect others in ways that have not been available before. Therefore, a new research method that allows the complexity and nonlinearity of many social processes to be managed in a simple way. The complexity of the environment in the financial markets in general has encouraged the use of modeling by multi-agent platforms and particularly in the case of the stock market. This paper provides a brief review of the recent advances in agent-based modeling research for stock markets. It uses a bibliometric analysis approach to summarize the status quo and development trends of agent-based modeling for stock markets. It uses the academic research and trends by conducting bibliometric research based on systematic key search-string-based analysis of the Clarivate Analytics Web of Science. The study's findings assist establishing a solid conceptual base and suggest promising avenues for further research.

Forecasting uncertainty - the quest for quantification

Presenter: Steve Morlidge
Co-authors: Steve Morlidge;

Back in 1921 Frank Knight drew a distinction between uncertainty - where we don't have the knowledge to predict the outcome of an event - and risk where there is sufficient knowledge about the class of events to be able to make a probabilistic prediction. This is a challenge faced by every forecaster because forecasting when you have perhaps years of demand history is completely different from forecasting demand for a new product, for example, because in the latter case we have no prior knowledge upon which to base a forecast. But does this mean that it is not possible to submit uncertain situations to quantification and devise a strategy we can apply in business to reduce it? The tentative answer to this is "˜no' as will be described in this talk which will outline a way to make uncertainty mathematically tractable in a manner that is within the compass of non expert and which is capable of being understood by most numerically competent managers.

The Forecaster's Predicament: Issues with Communicating Uncertainty

Presenter: Michael Gilliland
Co-authors: Michael Gilliland;

From a technical standpoint there are many ways to express the uncertainty in a forecast (including prediction intervals, fan charts, and full predictive densities). But when uncertainty is high, the forecaster faces a special challenge: How to communicate this high degree of uncertainty without making yourself look incompetent? Studies suggest that a moderate level of uncertainty is tolerated by decision-makers, but forecasts with extreme uncertainty may be disregarded as not useful. This presentation explores the challenge faced by forecasters to maintain credibility when providing an honest assessment of forecast uncertainty.

Cancelled: Recalibrating probabilistic forecasts to improve their accuracies

Presenter: Ying Han
Co-authors: Ying Han; David Budescu

There are several compelling examples in the forecasting literature (e.g., Baron, et al,. 2014; Turner at al., 2014) illustrating the benefits of recalibration of individual forecasts, as well as aggregates of multiple forecasts, of the target events. All these examples focused on binary events and, as such, amount to recalibrating "" extremizing or de-extremizing"" a single probability. In this paper we proposed, to our knowledge, the first extension of this approach that allows one to recalibrate a cumulative probability

function based on C of its quantiles in a consistent and coherent way that is captured by its singe recalibration parameter. The recalibration function is defined relative to the uniform distribution and its impact is defined in relation to the invariant "anchor" , Prob = 1/C, in the sense that probabilities below or above this anchor are transformed in different directions. The recalibration function generalizes Karmarkar's transformation that was used often in the special case C=2. We discussed some of the properties of the proposed function and illustrated its use by re-analyzing a large body of forecasts from the quarterly Survey of Professional Forecasters (SPF) conducted by the European Central Bank (ECB), for three economic indicators made by almost 100 experts and spanning 72 quarters. This analysis confirmed that recalibration can be highly beneficial and we found that its effects are not uniform, in the sense that not all indicators benefit equally. It also clearly showed that, on average, longer term forecasts require more aggressive recalibration. Finally, we have illustrated obvious practical applications of our approach by showing how one can use recalibration parameters estimated in previous periods to significantly increase the accuracy of future short-term forecasts.

Constrained Maximum Likelihood Estimation for Forecast Reconciliation in Hierarchical Time Series

Presenter: Zhilong Zhu
Co-authors: Zhilong Zhu;

What makes hierarchical time series (HTS) special are the summation constraints between child nodes and their parent node. While observations naturally satisfy the constraint, almost all forecast using non-linear model need to undergo a reconciliation process to have the summation constraints satisfied across the hierarchy. The reconciliation of HTS forecast typically involves two steps: (1) fitting a time series model independent of other nodes to get the base forecast, (2) an algorithm that determines weights of the base forecasts for each of the node based on the base forecast's historical performance. In this talk, the reconciliation of HTS forecast is posed as a constrained optimization problem during model fitting, where model parameters for all nodes are estimated simultaneously using maximum likelihood estimation with equality constraints between child and parent nodes' forecast. The major advantage of this approach lies in the utilization of model parameter uncertainties and thus model reliability in the reconciliation. Just like observation, the model parameter estimated has uncertainties, and the magnitude of the uncertainties in the parameter space determines the reliability of the model. As a case study, this novel reconciliation approach was tested with the Australian Tourism data that is commonly referenced in literature and showed very promising results.

A Follow-up on Robust Reconciliation

Presenter: Mauricio Lila
Co-authors: Mauricio Lila; Fernando Cyrino; Erick Meira

The demand for information to support an effective decision-making process is an increasing issue. The society and politicians want to understand the effects of policies on certain groups of interest, especially for hierarchical or grouped domains of interest.Time series sometimes offer special features, such as the natural disaggregation of their components according to a hierarchical structure. Hierarchical forecasting methods can take advantage over such structure by considering the base forecasts reconciliation, producing results which are usually unbiased and more accurate than the ones provided by standard methods. In this work, we provide a follow-up on the idea of robust estimation for hierarchical forecast reconciliation methods. We formalize different robust approaches applied to data from a Labor Force Survey in Brazil. To demonstrate the potential and validity of the proposed approaches, we compare their performance with those from traditional and state-of-the-art methods across a range of forecasting experiments. These are conducted by considering different combinations of base forecasting methods. Based on our findings we were able to validate the contribution of robust estimators in the context of Hierarchical Time Series.

A Machine Learning Factor-Based Interpretation for the Bond Risk Premia in the U.S.

Presenter: Caio Vigo-Pereira
Co-authors: Caio Vigo-Pereira;

In this paper, we study the time variation of the risk premia in U.S. Treasuries bonds. We propose a novel approach for deriving a single spanning state factor consistent with a dynamic term-structure with unspanned risks theoretically motivated model. Using deep neural networks to uncover relationships in the full set of information from the yield curve, we derive a single state variable factor that provides a better approximation to the spanned space of all the information from the term-structure. We also introduce a way to obtain unspanned risks from the yield curve that is used to complete our state space. We show that this parsimonious number of state variables have predictive power for excess returns of bonds over 1-month holding period. Additionally, we provide an intuitive interpretation of derived factors and show what information from macroeconomic variables and sentiment-based measures they can capture.Link: https://caiovigo.com/publication/A_ML_Factor-Based_Interpretation_Bond_Risk_Premia_US/ML_Factor_Based_Interpretation_US_Bond_Premia.pdf/subsection%7BEstimation of Short-run Predictive Factor for US Growth using State Employment Data}Presenter: Arabinda Basistha
Co-authors: Arabinda Basistha;

We estimate a predictive single factor model targeted to unobserved common growth in GDP and GDI using a state space framework with state employment data. We use likelihood based comparison to select the states to estimate the dynamic factor. The results show improved in-sample and out of sample performance than threshold principal component factors and financial spreads. Out of sample evaluations indicate larger gains for GDI growth with 14 to 20 percent lower mean squared forecast errors than other alternatives. These estimates compare favorably to using sectoral employment based predictive factors. An expanded model using both sectoral and state employment data show that their common component is the primary predictive factor.

ROC Approach to Forecasting Recessions using Daily Yield Spreads

Presenter: Kajal Lahiri
Co-authors: Kajal Lahiri; Cheng Yang

Even though many studies have established the existence of structural breaks and declining predictability in the relationship between GDP growth and yield spreads, business analysts continue to watch for the inversion of the spread as a leading indicator for recessions. Using two alternative definitions of the target variable, we reevaluate the enduring power of spread to forecast recessions, notwithstanding the temporal instabilities. We identify the threshold value of the spread that produces the highest discriminatory power as measured by the ROC curve, and its functionals such as the hit rate, false alarm rate, and the Youden's index. Based on data from January 2, 1962, we find that the threshold has drifted upwards from zero since the recession of 1980. Once the threshold is adjusted to its optimal value recursively, the type I and II errors associated with the forecasts for the four recent recessions remained unchanged around 90% hit rate accompanied by a false alarm rate of 25%. To account for the sampling variability, we use a block bootstrap procedure to construct the confidence intervals for these statistics.

Structural Breaks in Seemingly Unrelated Regression Models

Presenter: Shahnaz Parsaeian
Co-authors: Shahnaz Parsaeian;

This paper develops an efficient Stein-like shrinkage estimator for estimating the slopeparameters under structural breaks in seemingly unrelated regression models, which then isused for forecasting. The proposed method is a weighted average of two estimators: a restricted estimator which estimates the parameters under the restriction of no break in the coefficients, and an unrestricted estimator which considers break points and estimates the parameters using the observations within each regime. It is established that the asymptotic risk of the Stein-like shrinkage estimator is smaller than that of the unrestricted estimator which is the common method for estimating the slope coefficients under structural breaks. Furthermore, this paper proposes an averaging minimal mean squared error estimator where the averaging weight is derived by minimizing its asymptotic risk. The superiority of the two proposed estimators over the unrestricted estimator in terms of the mean squared forecast errors are derived. Besides, analytical comparison between the asymptotic risks of the proposed estimators is provided. Insights from the theoretical analysis are demonstrated in Monte Carlo simulations, and on two empirical examples of forecasting U.S. industry-level inflation rates, and forecasting output growth rates of G7 countries.

Applications of sequential change detection in continuous-time self-exciting point process forecasting

Presenter: Moinak Bhaduri
Co-authors: Moinak Bhaduri;

Recent decades have witnessed a surge in a specific type of forecasting interest: fitting different models to specific sections of the relevant history with an intention to improve forecast quality, among others. Crucial to such analysis is an estimation of the boundary that defines the separation between neighbouring periods. This work is devoted to such estimation problems. Continuous-time point processes are the tools through which the modelling is done. The intensities are taken to be data-dictated, forcing a dependence structure over non-overlapping compact time supports. To identify structural breaks in such intensities, we deploy a sequence of tests implemented through novel statistics inspired by time reversal. In comparison to competing options, these detections are shown to be more accurate and asymptotically consistent without sounding too many false alarms. Bootstrapped intervals are offered; change-point-based clustering tools are also proposed. Examples are sampled from economic announcements, rare events, natural hazards, etc.

Cancelled: Dynamic Model Averaging in the Presence of Structural Change

Presenter: Alisa Yusupova
Co-authors: Alisa Yusupova; Nicos Pavlidis

Model combination (averaging) methods have gained substantial attention in the time series forecasting literature due to their success in numerous and diverse applications. In this work we focus on methods capable of accommodating changes over time in both model parameters as well as in the optimal model combination. This problem is particularly relevant as a growing empirical literature provides strong evidence in favour of structural instability in numerous application areas, including macroeconomics and finance. Ignoring structural changes can have detrimental consequences for inference and forecasting. The approach we discuss relies on Dynamic Model Averaging (DMA). We explore the extent to which the specification of DMA, which is designed for a gradually changing (drifting) data-generating process, can handle abrupt structural change, and discuss methods to overcome this limitation. Our objective is to design an alternative formulation that can accommodate different types of dynamics, and provide reliable estimates of the timing of structural changes. We apply our proposed approach on time-series from macroeconomics and finance, and compare its performance to state-of-the-art DMA formulations.

Scalable Cloud-Based Automatic Time Series Imputation

Presenter: Thiago Quirino
Co-authors: Thiago Quirino; Michael Leonard

Many organizations need to process large numbers of time series for data analysis, decomposition, forecasting, monitoring, data mining, and risk analysis. Often, these time series contain missing values that must be imputed. The TSMODEL procedure provides a resilient, distributed, optimized, generic time series analysis scripting environment for cloud computing. It comes equipped with capabilities such as automatic time series model generation, automatic variable and event selection, and automatic model selection. It also provides advanced model-based time series imputation. This paper describes the scripting language that supports cloud-based automatic time series imputation. Examples that use SASÂ® Visual Forecasting software demonstrate the use of this scripting language.

Forecasting Software Trends for the Next Decade

Presenter: Michele Trovero
Co-authors: Michele Trovero;

A solid forecasting support system, in the form of a commercial package or based on open source, is essential to enable analysts for generating reliable and robust forecasts in a production system. In recent years, the availability of options of forecasting support systems has been growing and expanding steadily. There are two main drivers of this change: on one side, the evolution of technology that is common to much of the software industry; on the other side, the evolution in algorithms and methodologies specific to forecasting. This talk will look at these two aspects of the transformation that forecasting software is undergoing and will try to extrapolate some common trends that will drive the development in the next several years.

Using Open Source Machine Learning Algorithms in SAS Visual Forecasting

Presenter: Javier Delgado
Co-authors: Javier Delgado; Taiyeong Lee; Michael Leonard; Thiago Quirino

Open source software has become an integral part of the forecasting workflow of many organizations. Using open source software often requires specialized knowledge of the languages that the software is implemented in. For example, you must have a working knowledge of the Python language, including how to load data, set parameters, and construct models, in order to effectively use TensorFlow's Python programming interface. The External Languages package included in SASÂ® Visual Forecasting software enables you to utilize these open source packages without requiring much knowledge of the language itself. It also facilitates the process of assigning computing resources in multinode/multicore environments. There are many ways this package can benefit you. For example, if you are an existing SASÂ® user, it enables you to explore different forecasting algorithms without having a thorough understanding of the language that the model is implemented in. You can then use any SAS packages you are familiar with to perform data analysis and postprocessing. SAS Visual Forecasting in particular provides myriad tools to optimize forecasting workflows. You can even use the SAS environment to "glue" code written in the Python, R, and SAS languages.We show how to use SAS Visual Forecasting software to apply a TensorFlow neural networkâˆ'based forecasting algorithm to SAS time series data sets, visualize the results, and compare the results to those of native SAS algorithms. We describe how this workflow can be seamlessly scaled for data sets consisting of thousands of time series. We discuss the benefits of using SAS Visual Forecasting versus performing the analysis directly in Python. This discussion includes the evaluation of processing time, the amount of manual tuning required, and the analysis and postprocessing tools available, including tools that are packaged with SAS Visual Forecasting, such as automatic forecasting.

A tail risk penalty based combination approach to commodity price forecasting

Presenter: Yifei Zhang
Co-authors: Yifei Zhang; Jue Wang; Sheng Cheng; Xiao Zhang

Forecast combination is a well-established technique for improving forecasting accuracy. It investigates the integration of competing forecasts to produce a composite forecast which are superior to individual forecasts. Various combination methods have been presented over the past decades, like simple average, expert aggregation, Bayesian methods.Normally, simple averaging method and optimal weight method minimizing the combined in-sample mean square error are the most popular methods. The forecasting performance measure for the combination models is the error on future data outside the training set, also known as generalization error. This error may be undesirably large when direct minimization of the training error like mean square error for a fixed in-sample dataset is conducted. In other words, minimization of the training error does not necessarily imply a corresponding minimization of the generalization error. This phenomenon is usually referred to as "overfitting" . Upon choosing an appropriate loss function of the forecast combination process, a novel forecast combination method is presented which is expected to combat overfitting and improve the generalization ability. Based on the idea of portfolio theory, we defined the tail risk measurement in forecast combination using VaR (Value at Risk) and ES (Expected Shortfall). It is expected to capture and measure the tail risks of in-sample ensemble error. Furthermore, a novel target loss function with a penalty term is proposed which can make the trade-off between the global and tail risk of forecast combination models. Specifically, an ant bee colony algorithm based optimization method is introduced to achieving the optimal combination weights.The experimental results on gold price and oil price data demonstrate that the proposed forecast combination approach can outperform not only individual competitive benchmark but also combination approaches like the simple averaging, optimal weight method and other benchmark models. For instance, the MAPE achieved by the presented method could decrease by 11.6% and 6.6% in gold and oil price forecasting, compared to the best benchmark model.

Social Costs of the New Energy Policy in Mexico, 2005Q1-2020Q4. A Forecast.

Presenter: Eduardo Loría
Co-authors: Eduardo Loría; Raul Cossio

Despite the 2020 economic depression, inflation has continued to grow in Mexico. This has been the result of the government's new energy policy, which increased significantly fuel prices, and so it will be in the long run. We find that there are asymmetric effects of unemployment (U) and the Rate of Critical Labor Conditions (RCLC) on inflation. Through an Augmented New Keynesian Phillips Curve, estimated with a Non-Linear Autoregressive Distributed Lag (NARDL) model for 2005Q1-2020Q4, we calculate the asymmetric sacrifice rate on unemployment and RCLC to keep in line the Central Bank inflation goal. Based on a forecast of the price of hydrocarbons, we estimate how much it will cost in terms of U and RCLC.

Efficiency gains from the deregulations of refined oil prices in China

Presenter: Xun Zhang
Co-authors: Xun Zhang; Lin Zhao

This paper examines the efficiency gains from the deregulations of refined oil prices in China. Regressions based on a panel data set of 31 provinces show that, the efficiency improvements of oil products from deregulations on refined oil prices are significant, while the efficiency gains of crude oil consumption are insignificant. Specifically, the increase of labor productivity has positive effect on energy efficiency gains brought by deregulation, while the increases of share of transportation sector in economy and deficits in crude oil consumption have negative effects. Then, a Dynamic Stochastic General Equilibrium model with

Chinese characteristics is developed to provide theoretical verifications, and simulate these three channels to further explain how macroeconomic structures affect energy efficiency gains brought by deregulations. The model also generates forecasts of future energy efficiency.

Nowcasting Economic Activity with Mobility Data

Presenter: Koji Takahashi
Co-authors: Koji Takahashi; Oh Yusuke; Tomohiro Sugo; Kohei Matsumura

In this paper, we develop high frequency indexes to measure sales in service industries and production activity in the manufacturing industry by using GPS mobility data from mobile applications. First, focusing on the possibility that the number of customers in service industries can be estimated using mobility data, we develop indicators to capture economic activity in amusement parks, shopping centers, and food services. We show that using GPS mobility data, it is possible to nowcast economic activity in the service industries, in real time, with a high level of precision-something which conventional statistics are largely unable to assist. In addition, by using statistical methods such as clustering, we can construct an indicator with even better nowcasting performance. For the above analysis on the three sectors in the service industry, some may point out that after the pandemic, the number of visitors to commercial facilities in service industry has drastically declined, which increased the correlation between sales and the population. Therefore, some may claim that a high correlation in turbulent times does not guarantee high nowcasting performance in normal times. However, we only use data up to March 2020 when the effect of the pandemic on consumption was limited and show that even before the serious deterioration of the pandemic, the indicator demonstrates a high correlation with other statistics. Second, in the manufacturing sector we identify the locations of relatively large factories using panel data from the Economic Census for Business Activity and by utilizing hourly and daily mobility patterns such as a daytime ratio. We then construct indicators for nowcasting production based on the population in the specified areas. We show that we can nowcast production with a high level of precision for some labor-intensive industries including the transportation equipment and production machinery industries, despite the fact that we abstract important factors such as capital stock or materials that are thought to determine output. These results suggest that mobility data are a useful tool for nowcasting macroeconomic activity.

Nowcasting Thailand Economic Activity Using the Google Mobility Data

Presenter: Chaleampong Kongcharoen
Co-authors: Chaleampong Kongcharoen;

Forecasting economic activity in the covid-19 pandemic era is challenging. At the same time, policymakers and business leaders require timely evaluation of the health of the economy. Even in the period before the pandemic, most of Thailand's economic indicators lag in an announcement. This paper explores the benefit of adding publicly available data, i.e., the Google Mobility data, for assessing the economic situation. We compare the forecasting performance of the Autoregressive Moving Average (ARMA) model, ARMA with explanatory variable (ARIMAX), and Mixed-data sampling (MIDAS) model using the Google mobility index. We consider the monthly service production index and manufacture production index. We find that the Google Mobility data help to improve the forecasting performance of various service sector. While the models with the Google Mobility index perform worse than the pure time series model, i.e. ARIMA in finance, public administration and manufacturing sector.

Probabilistic ensemble forecasting of Australian COVID-19 cases

Presenter: Rob J Hyndman
Co-authors: Rob Hyndman;

In March 2020, I joined a team responsible for providing probabilistic forecasts of COVID-19 cases to all Australian state & territory Chief Health Officers. We use case-level data of all Australian positive COVID cases, along with nationwide surveys and mobility data from Google, Facebook and Apple. Three separate models have been built: (1) a stochastic susceptible-exposed-infectious-recovered (SEEIIR) compartmental model; (2) a stochastic epidemic model; and (3) a global autoregressive model based on public case data from 31 countries. These are then combined into a mixture ensemble to generate probabilistic forecasts of daily cases which are provided to the Australian governments each week. I will discuss the ensemble forecasting aspects of this work and how we evaluate the results.

Feature-based ETS model components selection

Presenter: Lingzhi Qi
Co-authors: Lingzhi Qi; Xixi Li; Yanfei Kang

Hyndman et.al (2002) proposed a well-developed modelling framework named ETS (ExponenTial Smoothing or Error, Trend, Seasonality) for automatic forecasting, which incorporates state-space models, parameter estimation, point prediction and interval prediction. The ETS framework provides an automatic way of selecting the best model components (E, T, S) with information criteria such as AIC and BIC. While the information criteria focuses more on the model complexity, leaving the forecasting performance of model over training data under-emphasized. To tackle this issue, this work adopts a meta-learning technology for optimal ETS model components selection. Specifically, we train LightGBM to obtain the optimal components for the ETS models based on the time series features. We evaluate our approach on the widely-used forecasting competition datasets (M1, M3, and M4), in terms of both point forecasts and prediction intervals. To demonstrate the practical value of our framework, we showcase the performance improvements from our approach on load data.

Improving intermittent demand forecasting : an empirical study on forecast combination methods

Presenter: Li Li
Co-authors: Li Li; Yanfei Kang; Feng Li

In the field of intermittent demand forecasting, limited attention has been given to forecasting combination methods, which have achieved good performance in forecasting fast-moving time series in recent years. The current study examines the empirical outcomes of some traditional forecast combination methods and proposes a diversity-based forecast combination model to improve the intermittent demand forecast combination. We analyze the point and quantile forecasting performance based on simulation and real datasets.

The uncertainty estimation of feature-based forecast combinations

Presenter: Xiaoqian Wang
Co-authors: Xiaoqian Wang; Yanfei Kang; Fotios Petropoulos; Feng Li

Forecasting is an indispensable element of operational research (OR) and an important aid to planning. The accurate estimation of the forecast uncertainty facilitates several operations management activities, predominantly in supporting decisions in inventory and supply chain management and effectively setting safety stocks. In this paper, we introduce a feature-based framework, which links the relationship between time series features and the interval forecasting performance into providing reliable interval forecasts. We propose an optimal threshold ratio searching algorithm and a new weight determination

mechanism for selecting an appropriate subset of models and assigning combination weights for each time series tailored to the observed features. We evaluate our approach using a large set of time series from the M4 competition. Our experiments show that our approach significantly outperforms a wide range of benchmark models, both in terms of point forecasts as well as prediction intervals.

Distributed Forecasting with Large Bayesian Vector Auto Regressions

Presenter: Feng Li
Co-authors: Feng Li; Yanfei Kang; Sune Karlsson; Xiaowen Man

This paper proposes a Bayesian distributed vector autoregressive (DVAR) model for the distributed system with the least square approximation method. The DVAR model properly handles the large p problem by introducing additional virtual observations. Our algorithm improves the computational efficiency for large T with the support of the Apache Spark platform. Furthermore, we consider the data streaming scenario and propose an efficient Bayesian updating scheme based on the DVAR model. The DVAR model is applied to the prediction of the ultra-long electricity load data and the stock market indexes data. Our empirical study shows the DVAR model performs well on both applications.

Online Hotel Product Click-Through Forecast Using Contents of Promotional Photos

Presenter: Chenyu Cao
Co-authors: Chenyu Cao; Doris Wu; Ji Wu

The development of deep learning technology allows us to process online hotel photos at a considerably larger scale. Image recognition tools based on deep learning are capable of processing pictures automatically and return semantic concepts of entities captured in them, with which we can quantify the "contents" of photos quickly and objectively. Since the show photo on the accommodation service vendor listings is the first image customers see when they are searching for accommodation products, this study aims to identify how the "contents" of these hotel photos affect customers' click decisions. Empirically, 386 hotel photos with 675,175 hotels and customers' product click data from an online tourism service platform in China are collected for examination. Based on graph theory, we adopt community detection algorithms to construct the relations of entities in photos and discover subcategories for different entities. Then regression is adopted to investigate how the "contents" of hotels' photos affect and predict customers' click decisions. We also include such controls as prices, time intervals of display and types of hotels in our regression model. Our results show that photos showing outdoor scene and building appearance of hotels result in more clicks, while photos of a guest room are not appealing. With deep learning and a well-tuned forecasting model, our study provides an "AI judge" for accommodation vendors to evaluate the attractiveness of their show photos and helps them to improve their promotional performance by designing more effective promotional photos.

Tourism demand forecasting using time-varying parameter global vector autoregressive model

Presenter: Long Wen
Co-authors: Long Wen; Gang Li; Han Liu

The advantages of time-varying-parameter (TVP) models in forecasting tourism demand have been well discussed, and recently several studies have applied global vector autoregressive (GVAR) models to examine tourism demand for a group of destinations in the same region. GVAR models are able to take account of the interrelationships between different countries as well as global shocks by combining

country-specific models into a global framework. It therefore provides an important tool to better model tourism demand from a global perspective as tourism demand of different countries are often interconnected, especially within the same region. However, the tourism demand system is also likely to change over time and a combination of TVP and GVAR techniques can provide further improvement in capturing the dynamics in a tourism demand system, which has not been investigated in the tourism forecasting literature. This study presents the first attempt to use the TVP-GVAR model to forecast tourism demand in a regional setting.

Towards tourism demand forecasting with multisource big data in cloud computing

Presenter: Xinyan Zhang
Co-authors: Xinyan Zhang; Richard Tianran Qiu; Doris Chenguang Wu

With the development of the tourism industry and the information and communication technologies, more and more tourism-related data that could be analyzed by both public- and private- sectors have become available and reached the terabyte (TB) or even petabyte (PB) level. This big data phenomenon has brought great opportunities for tourism demand forecasting research as the collection of up-to-date information and the key factors could help to enhance the forecast accuracy. However, most of the traditional forecasting systems in the tourism industry use the old construction method and have difficulties in tackling big data. As opposed to the traditional single machine computing, cloud computing is an Internet based computing in which the service provider takes care of the base infrastructure including computers, software, and data storage. Customers pay for the use of this infrastructure on a monthly pay-as-you-use basis. Therefore, the cloud could offer an optimum or even unlimited computing and storage capacity for effective big data storage and analysis. In addition, cloud computing also has the advantages of high speed, reliability, access to the latest applications, and collaboration efficiency.This study will explore the use of multisource big data for tourism demand forecasting in a cloud platform. The platform will adopt the B/S (Browser/Server) architecture to enhance the usability. Development of the back-end layer will be based on the Java environment using the SSH (Spring, SpringMVC and Hibernate) framework to process data with Apache Spark, which is an open-source computing architecture for real-time big data processing. Spark supports an extensive range of languages such as R, SQL, Python and Java. The forecasting system will interact with end-users through a dashboard and user interfaces at the front end. This front-end layer will be developed using HTML, JavaScript and CSS. Bootstrap will be integrated to enable the user interfaces and dashboard to adjust automatically. All the data will be stored, staged and processed in the cloud. A hotel occupancy forecast using sentiment analysis with online reviews will be utilized to demonstrate the superiority of the proposed system.

Probabilistic Forecasting of Solar Power Generation with Autoregressive Recurrent Networks

Presenter: Yao Zhang
Co-authors: Yao Zhang; Fan Lin

The need of solar power uncertainty quantification in power system has inspired probabilistic solar power forecasting. Recently, the DeepAR method has been proposed to train an autoregressive recurrent network for providing time series probabilistic forecasting. To the best of our knowledge, our work is the first effort to use the DeepAR model for probabilistic solar power forecasting. To effectively apply this model in solar power forecasting, it is necessary to make some customizations according to solar power data features. In this work, we make customizations on the DeepAR prototype from three aspects and propose multiple DeepAR-based solar power probabilistic forecasting (SDAR) models. First, solar power time series is preprocessed through normalization and Fourier decomposition considering its boundedness and seasonality. Second, a novel loss function form based on continuous ranked probability score (CRPS) is

proposed for model training. Third, several parametric and nonparametric distribution forms with bounded or unbounded support are utilized to depict probability density of solar power time series. Through customizing different components in DeepAR framework, 15 SDAR models are proposed in this paper.

Numerical results on public real-world data show that 1) the proposed CRPS-based loss function works effectively in training SDAR models and performs evidently better than the widely-used log likelihood function in the case of complex distribution forms. 2) The generalized parametric forms of nowcast distribution outperform clearly the original ones, indicating that the generalized parametric distribution could improve the forecasting performance. 3) The effectiveness of parametric distributions is somewhat associated with their "tailedness" , and long-tailed distributions (i.e., Laplace and two-side power distributions) provide better forecasting for solar data studied in this paper. 4) Compared with state-of-the-art method (e.g., quantile regression neutral network), our proposed best SDAR model significantly improves the quality of very short-term probabilistic solar power forecasting.

Enhanced SVR model for solar PV power forecasting

Presenter: Punam Pawar
Co-authors: Punam Pawar;

High penetration of renewable energy resources like solar PV provides a great challenge for system operators. Intermittency of solar PV instigates difficulties for power system stability. However, intermittency issue can be addressed using accurate solar PV power forecasting. In recent years, numerous machine learning methods are introduced and among them, Support Vector Regression (SVR) is one of the most common techniques used in solar PV power forecasting. SVR provides better forecasting accuracy, is computationally less complex and does not need large data for training. Nevertheless, providing initial guess values in training phase for SVR model is not only challenging but it also affects complexity and accuracy of the model. Additionally, SVR uses Lagrange multipliers for converting non-linear data into linear data and needs up to 100 steps for solution. For efficient solar PV power forecasting, quick learning, computationally simple and robust model is desired. Therefore, in this research, the traditional SVR model was modified using non-linear least square method called Gauss-Newton method. In this enhanced model, kernel function of Gaussian Radial Basis in SVR was replaced by more simple function, a sine function. Moreover, optimization problem of nonlinear data was solved using Jacobian matrices. This enhanced SVR model was tested for its forecasting accuracy in different weather conditions like sunny, partly cloudy and overcast. Robustness of the enhanced SVR model was firstly verified in bivariate form using only two variables of solar insolation and PV power output data. Secondly, the model was tested in multivariate form using temperature and humidity data along with solar insolation and PV power output. Finally, to explore the model strength, kernel function used in this enhanced model was changed to radial basis function. For performance analysis, this enhanced model was compared with traditional SVR model using error parameters like root mean square error (RMSE), mean absolute percentage error (MAPE), mean relative error (MRE), mean absolute error (MAE) and mean bias error (MBE). Modification in traditional SVR model allowed to solve the nonlinear problem in less than 12 steps. Efficient performance of this enhanced SVR model was confirmed for achieving better accuracy and less computational complexity.

Solar power forecasting method based on smooth trend estimation in three directions of time, date, and solar radiation

Presenter: Takuji Matsumoto
Co-authors: Takuji Matsumoto; Yuji Yamada

With the introduction of a large amount of solar photovoltaic (PV) power generation on a global scale, the needs for forecasting the PV generation are increasing year by year. To date, many methods such as machine learning have been proposed for their forecast, but most studies do not necessarily model the relation between inherent smooth seasonal trends and the PV output. Also, it is important to express

the dependence of other exogenous variables on the PV output using a suitable multivariate function. In this context, we adopt an approach using a generalized additive model (GAM) to estimate smooth trends, which is easy to handle for practitioners. In the past methods using GAM, some approaches have been proposed such as estimating seasonal trends separated by hourly zone and weather dummy variables and estimating a two-dimensional tensor product spline function that considers smoothing conditions in the date and hour directions. This study further develops these methods by constructing a model using a three-dimensional tensor product spline function that simultaneously incorporates smoothing conditions in three directions of date, hour, and solar radiation. By comparing with multiple previous models, we reveal the advantages of the proposed method in ensuring robustness and reducing prediction errors.In particular, the empirical results demonstrate that the proposed method based on the three-dimensional tensor product spline function can well express the non-linearity of the PV output with respect to the solar radiation, which changes with each date and time. We specifically visualize that the non-linearity of PV output with respect to the solar radiation becomes remarkably strong resulted from the decrease in PV generation efficiency due to the temperature rise around spring to early summer. It also be clarified that the proposed model has the highest forecast accuracy for both in-sample and out-of-sample periods among the compared models. Furthermore, by measuring the prediction error by each month, we demonstrate that the categorical variable-based approaches such as monthly or hourly zone are prone to over-learning and provide a clear evidence that the proposed method imposing the smoothing conditions in the three directions has a superiority in terms of securing robustness.

Second-hand housing price forecast with purchase purpose analysis: an empirical study

Presenter: Wei Shang
Co-authors: Wei Shang; Ruizhen Bai

Because of the rapid development of China's economy and the continuous development of urbanization, people pay more and more attention to real estate information. However, with less and less land available for development in the city, the second-hand housing market has gradually become an important role in real estate transactions, and the demand for fast and accurate second-hand housing valuation is becoming more and more large. The core content of this paper is to establish machine learning models for the prediction and analysis of second-hand housing prices, collect the listed housing prices of Xicheng district in Beijing between 2017 and 2018 as research data, using recursive feature elimination method to select suitable variables, and then build multi-linear regression model and random forest model to predict and evaluate second-hand house prices, comparing the result of model forecasts. In addition, this paper selects 7 different housing purposes with obvious characteristics. Purchase purposes have an impact on the house prices and decision-making of residents when choosing a home. Matching these purchase purposes with the characteristic variables, and incorporating the corresponding purchase purpose of each small area as the characteristic variable into the second-hand housing price prediction model, can greatly improve the prediction effect.

Nowcasting GDP with targeted predictors: A model averaging for dynamic factor models

Presenter: Yuying Sun
Co-authors: Yuying Sun; Yongmiao Hong; Yizhou Kuang

With the advent of the complex information system to collect data, real-time nowcasting faces various challenges, including a large number of predictors, the number of lags, unbalanced data structure, model uncertainty and how to bridge high-frequency information contained. To address these issues, this paper proposes a new real-time nowcasting forecast combination with dynamic factor regressions, which deletes redundant predictors and selects optimal weights for candidate models, simultaneously. It is shown

the selected weight achieves the asymptotic optimality and consistency, even all candidate models are misspecified. The proposed estimator is consistent and asymptotic normality, if a true model is included in candidate models. Simulation results show that the proposed method yields lower mean square forecast errors than alternative nowcasting methods, including MIDAS in Ghysels et al. (2004), GARS in Giannone et al. (2008), and FADL-MIDAS in Andreou et al. (2013). The proposed method is applied to forecast quarterly GDP with a set of 118 macroeconomic monthly data series, which compares favorably to other competing methods.

Nowcasting China's PPI inflation using low frequencyand mixed frequency dynamic factor models

Presenter: Jiawen Xu
Co-authors: Jiawen Xu;

We construct nowcasts and forecasts of China's PPI inflation using a large panel of data series with different frequencies (monthly, ten-day, weekly, daily). Mixed frequency data are incorporated in the dynamic factor model in two approaches: one is to convert high frequency data to low frequency, which is monthly in our example, and apply the expectation maximization (EM) algorithm in estimation; the other is to treat low frequency data as high frequency (daily) data with missing observations in a specific pattern and apply Banbura and Modugno (2014) to estimate the dynamic factor model with missing observations. We compare the forecast accuracy of these two approaches with some other alternative forecasting models, such as random walk, univariate autoregressive model and dynamic factor model with only monthly data. Our empirical results show that the first approach outperforms other models in most cases and horizons. Models utilizing high frequency data generally perform better than those do not. As high frequency information flows in, the forecasting accuracy improves substantially.

Optimal reconciliation with immutable forecasts from specific hierarchical levels

Presenter: Bohan Zhang
Co-authors: Bohan Zhang; Feng Li; Yanfei Kang

The practical importance of coherent forecasts in hierarchical forecasting has inspired many studies on forecast reconciliation, which generally improves the forecasting accuracy via the process of adjusting forecasts. A series of forecast reconciliation methodologies have been developed based on least squares, such as OLS, WLS and MinT. These methodologies are unconstrained, which means there are no additional constraints except the aggregation constraints. However, in practice, it might be necessary to keep forecasts of particular levels unchanged after forecast reconciliation (e.g., when there are fixed budgets for a company or the base models can capture the information in some specific levels very well). In this paper, we formulate this constrained forecast reconciliation problem as least squares with equality constraints, and propose the constrained OLS, WLS and MinT methods. Monte Carlo simulation demonstrates that the proposed methods can keep the forecasts of some levels unchanged after forecast reconciliation with substantial accuracy improvements. We also perform empirical experiments to assess the impacts of bringing in constraints. Note that the proposed method is different from Wickramasuriya et al. (2020), which deals with non-negative constraints, while our study aims to keep specific levels fixed after reconciliation. We have developed a python module pyhts to implement the proposed method (and the general forecast reconciliation methods), which is available at https://angelpone.github.io/pyhts/./subsection%7BExplainability versus privacy for hierarchical time series forecasting}Presenter: Surya Shravan Kumar Sajja
Co-authors: Surya Shravan Kumar Sajja; Sumanta Mukherjee; Dhinakaran Vinayagamurthy; Pankaj Dayama; Vikas Raykar

Sales and inventory data in supply chains has spatio-temporal characteristics, leading to a hierarchical time series framework. Accurate forecasting and decision making in a hierarchical framework depends on

the data shared amongst participants from the same hierarchical level and also data sharing across different levels of hierarchy. Although conceptually attractive, many participants in the supply chain are unwilling to share detailed information due to the fear of unfair exploitation by other competing parties.Â There would also be regulatory requirements that limit the data sharing outside the data owner in plain form. Hence, privacy preserving protocols are the need of the hour.Â Apart from privacy, explainability is another essential requirement imposed on the design of most AI based systems. Supply chain participants who rely on hierarchical forecasts, expect to understand why are the forecasts high or low?Â Which adjacent node or a faraway node in the hierarchy is responsible for a change in the forecast? These requirements are in direct opposition to the privacy requirements of the participants, thus posing several engineering and research challenges. In this work, we focus on developing forecasting systems that exploit hierarchical data structures while optimizing for a trade-off between privacy and explainability.

Forecasting hierarchical time series with multi-output ML models

Presenter: Mahdi Abolghasemi
Co-authors: Mahdi Abolghasemi; Rob J HyndmanRecently there has been a great interest in reconciling hierarchical series with machine learning models but the value of explanatory variables in forecasting hierarchical time series has not been explored enough. There are some advanced statistical techniques for reconciliation but these techniques often do not directly consider the explanatory variables that might be available, making them less effective to model sudden changes in various nodes of hierarchical time series. We conduct an experiment in a supply chain setting where we use multi-output top-down machine learning models with explanatory variables to generate forecasts for lower levels nodes from their parent nodes. We empirically evaluate our model on various hierarchical time series and show that using explanatory variables can significantly improve the accuracy of reconciled forecasts across different levels.

NA

Measuring Unobserved Judgment

Presenter: Emilio Zanetti Chini
Co-authors: Emilio Zanetti Chini;

This paper investigates the dynamics of the bias due to strategic judgment (SJB) in macroeconomic forecasts. The SJB arises as output of a game among three players: forecast producer, forecast user and reality. The interaction among these players is formalized by a mathematical object named Scoring Structure (SS) and summarized by a forecasting protocol. This last aim to test the coherence of professional forecast with the announcement by the forecast user via likelihood principle: if the forecast user does not introduce any judgment in his/her announcement in the quotation by forecast producer and viceversa, the two estimates corresponds to the likelihood of the forecasting model (assumed being known), which, in turn, is equivalent to the Logarithmic Score of the estimated predictive density. In a static system, the SJB is measured as the Lq-transform of the likelihood function (or "deformed likelihood"), where the tuning parameter q is the quote of judgment for unit of information. The deformed likelihood coincides with Log-likelihood estimator if q=1 and can be represented as weighted likelihood, where weights can be used to isolate the quote of the judgment from the quote of sample-based information. We introduce an estimator of the unobserved SJB for a dynamic system, which is named "Deformed Kalman Filter" (DKF, or "Judgmental Filter" , JF). This allows a recursive estimation of the weights of the Deformed Likelihood so that the q-parameter is detected without using complicated inference of resampling methods, as required by the Lq-Likelihood estimation literature. A Monte Carlo simulation experiment proves the feasibility and the good capabilities of the proposed method, also in samples characterized by additive outliers. The DKF is applied on U.S. data on GDP forecasts by Survey of Professional Forecaster (the forecast producer) and FRED (the forecast user). Our results reveal that (i) SJB varies significantly with the forecasting horizon; (ii) it is high during recessions and low during expansions; (iii) the increases in the quote of judgment are more frequent in the first half of the sample, while in the second half is characterized by lower quote.

Band-Pass Filtering in the Time Domain

Presenter: Alessandro Giovannelli
Co-authors: Alessandro Giovannelli; Marco Lippi; Tommaso Proietti

In this paper we show that the end-of-sample problem in band-pass filtering has a unique optimal solution, which can be determined either by frequency-domain or time-domain techniques. We discuss the method known as Christiano and Fitzgerald's (CF henceforth) based on the frequency domain approach and propose an alternative which is based on the time-domain representation of the series of interest, we call it the time-domain projection method. The projection method consists in band-pass filtering the available series, augmented with the forecast of the data outside the sample. The above method is then used to construct alternative indicators of the current state of the economy, which is a smoothing of the growth rate of GDP. The competing indicators are based on methods that make use of large datasets like New Eurocoin. The projection method compared to New Eurocoin has several differences. Two, in particular, are essential: New Eurocoin is calculated exclusively in the frequency domain and employs smooth factors, which are estimated using only frequencies above one year. For the projection method, instead, the estimation of the factors is performed through the Stock and Watson (2002) or Forni et al. (2005) procedures. In comparing the performance of these indicators on real data, we run a pseudo-real-time exercise based on data relative to the US economy (FRED dataset). The empirical analysis leads to the following results: the projection method performs substantially better than the Christiano-Fitzgerald filter (univariate) and is very competitive compared to New Eurocoin.

Forecasting GDP Growth Rates Using Google Trends in the United States and Brazil

Presenter: Evripidis Bantis
Co-authors: Evripidis Bantis; Michael Clements; Andrew Urquhart

Many studies explore the usefulness of "Big Data" in forecasting specific economic variables such as unemployment or inflation, but only a few focus on the overall economic activity. Thus, the purpose of this paper is to nowcast GDP growth rates by using a dynamic factor model based on traditional economic indicators as well as on Google search data. Our analysis covers Brazil and the United States from 2005 to 2019. Moreover, we employ several variable selection methods to investigate whether factor models with targeted predictors provide forecast gains when utilizing high-dimensional datasets. Empirical results show that factor models based on economic indicators and Google search data provide forecast gains compared to factor models based only on economic indicators. When we isolate the source of forecast improvements, we find that pre-selecting predictors indeed matters, but gains appear mostly in forecast horizons and tend to vanish as we move to nowcasting and backcasting horizons. Finally, only the main Google Trends categories can provide forecast benefits while performing better for the United States than in Brazil.

Time Series Segmentation Using Two-Stage Clustering Approach

Presenter: Sagar Mainkar
Co-authors: Sagar Mainkar;

Time series data have a unique data structure, each point in series is often is considered a variable and each time series is considered an observation. As the time dimension increases, the number of variables also increases, in proportion to the time dimension. The existing segmentation techniques have high complexityand large processing time because of distance-based clustering mechanisms as every time series variable compares with other time series variable. The objective of this paper is to utilize statistical properties of time series and reduce them to lower dimension using dimension reduction techniques and then cluster them into segments which have similar patterns.In this paper we use a two stage clustering

approach for segmentation.The time series data can be characterized using time domain statistics such as autocorrelations, partial autocorrelations and white noise test statistics. These statistics are used for model identification in the ARIMA model, so they represent the feature of time series well. If you use the ACF and PACF together, the resulting features represent an ARIMA model well. In the first stage we utilize statistical properties of time series viz. PACF. The correlation captures the seasonality and trend within the time series and this data can be utilized to cluster time series.At the end of first stage we obtain clusters of time series , however, to obtain a finer classification we reduce each time series within a cluster to a single number using either of the following technique Discrete Fourier Transform (DFT), Discrete Wavelet Transform(DWT), Singular Value Decomposition (SVD), Line Segment Approximation with Sum (LSAS) and Line Segment Approximation with Mean (LSAM). The single point time series are then clustered again. K-means clustering algorithm is used in both stages.Since this paper discusses unsupervised technique it is difficult to evaluate the final results but we would compare the results obtained with results obtained using similarity approach with on some public data sets. However, the end results of having similar segments of time series is the main objective of the paper that can be judged by visualizing the segments of time series.

Model Agnostic Feature Based Explainability For Timeseries Forecasting

Presenter: Sumanta Mukherjee
Co-authors: Sumanta Mukherjee; Vikas Raykar; Bhanukiran Vinzamuri; Giridhar Ganapavarapu

Explainability is the degree to which a human can understand the cause of a decision (or prediction) made by a prediction model. Various notions of explainability have been well studied in standard supervised learning paradigms, like classification and regression. In this work, we first formalize the notions of local and global explanations in the context of time series forecasting. A local explanation explains the forecast made by a forecaster at a certain point in time, whereas a global explanation provides insights into the forecasting model. We propose a feature-based explainability algorithm to explain the forecast of any forecasting model. The proposed method is model agnostic, needs access to only the predict (forecast) methods of the forecaster. For any given univariate time series we first generate an in-sample prediction time series based on a sequence of expanding window backtests. We then construct a surrogate time series forecasting task by reducing it to a standard supervised regression problem. For each time point, we generate a feature vector (lag features, seasonal features, date-time encodings, etc) based on which we need to predict the in-sample prediction time series values. The surrogate model is fitted using tree-based regressors like XGBoost, CatBoost, LightGBM, etc. We produce model explanations in terms of features that encode the time series. We rely on the Tree SHAP (SHapley Additive exPlanations) algorithm to explain the output of ensemble tree models. We evaluate the quality of the explanations in terms of (low) sensitivity, (high) faithfulness, and (low) complexity. To improve the sensitivity we extend the above approach by aggregating multiple explanations from bootstrapped versions of the time series. We also explore different strategies for multi-step forecasting, extensions to include external regressors, notions of semi-local/aggregate explanations, and parallelization strategies. We further point some shortcomings of the above approach in terms of not explicitly handling correlations in the time series.

Forecasting Model Territories

Presenter: Thiyanga Talagala
Co-authors: Thiyanga S. Talagala;

The field of time series forecasting has been evolving rapidly withadvances in techniques for modelling and forecasting. However, choosing the right technique for a given series is at the heart of forecasting research. This process is challenging because certain forecasting techniques will perform best on some series while different alternative will perform best on other series. Certainly, each forecasting technique has its own territory and dominance. Discovering the conditions under which a forecasting technique will function well and under which not is useful in identifying model territory and dominance. The forecast submissions of

the top 25 participants of the M4-competition are used for the analysis. Evaluating forecasting submission only using global measures such as forecasting error measure collapse all local information and does not allow to identify local differences of those methodologies. We explore the relationships between features of the methodologies used to generate forecasts, features of the resulting forecasts and features of the time series to identify model territories and their characteristics. Taking these local information into account can have benefit in developing new methods and shed some light for further development in the field of forecasting.

Modelling and Forecasting the Stock Market Volatility of China's Listed Tourism Companies in Unstable Environments

Presenter: Peng Yang
Co-authors: Peng Ynag; Han Liu

Tourism is an industry that is highly comprehensive, correlated, and sensitive to external influences, especially for the performance of its stock market. This paper uses the double asymmetric GARCH-MIDAS model to forecast the stock market volatility of China's listed tourism companies and investigate the impact of external shocks, such as the oil shock and the uncertainty of infectious diseases, on the tourism stock market volatility forecast. The empirical results show that: (1) The impact of external shocks on the volatility of stock market of Chinese listed companies has significantly asymmetric characteristics. (2) Frequent oil shock will have a significant impact on the long-term volatility of tourism stocks' performance, while sudden epidemics infectious diseases have a greater impact on the short-term volatility. (3) Forecasting performance of tourism stocks volatility is closely related to external shocks and the influence is asymmetry, and the forecasting performance tends to be worse when the volatility of external shocks is large. The research on the asymmetric relationship between the forecasting performance of listed tourism companies' stock volatility and external shocks can provide long-term and short-term guidance and suggestions for tourism practitioners and investors in different unstable stages.

Hotel demand forecasting with travel reviews: A topical sentiment analysis

Presenter: Shiteng Zhong
Co-authors: Shiteng Zhong; Chenguang Wu; Haiyan Song

With the rise of social networking, amounts of user-generated textual content data are generated exponentially. Tourist opinions in travel reviews play an important role in tourist decision-making process, which is crucial for hoteliers to anticipate the potential needs of customers. Travel reviews have already been applied in the field of tourism forecasting via sentiment analysis, but how to extract the sentiment of different topics of the reviews and make prediction has not been examined. As the development of research on fine-grained sentiment analysis, more and more studies pay attention to topical sentiment analysis. This study therefore aims to examine if conducting topical sentiment analysis can help to extract more effective sentiment information to improve forecasting accuracy. Particularly, the incorporation of Sentence Segment LDA and Bi-directional Long Short-term Memory model is proposed to conduct topical sentiment analysis and construct fine-grained sentiment indices. Both deep learning method and mixed data sampling (MIDAS) method are employed to evaluate the effectiveness of topical sentiment analysis-based forecasting. Demand for a number of high scale hotels in Hong Kong, Macao and Mainland China is used to compare the forecasting performance between traditional sentiment-based method and topical sentiment-based method. The present study is among the first to use textual reviews for hotel demand forecasting by the means of constructing topical deep learning sentiment analysis and the findings will highlight the importance of travel reviews in hotel revenue management and strategy formulation for hoteliers.Keywords: hotel forecasting, online review, topical sentiment analysis, deep learning, MIDAS

Product Picture Color and Consumers' Click Behavior

Presenter: Ming Cai
Co-authors: Ming Cai; Ji Wu; Chenguang Wu

The pictures in accommodation service vendor listings can affect consumers' decisions to click a vendor, which in turn influence their final purchase. Although a few studies have explored how the object or content of a picture affects a customer's perception and click intention, little is known about how the color of pictures determines customers' click decisions empirically in hospitality industry. We collected a unique dataset, which include 387 hotels' pictures and related customer clickstream data over 4 months, from an online travel service agent in China. Based on color theory, we propose a machine learning approach to extract color features from pictures, and use a regression model to empirically investigate the relationships between picture color and customers' click decisions in accommodation service. The results show that the degree of warmth and coldness of the vendor pictures significantly affect consumers' click behavior. Neutral colors and orange in warm colors increase consumers' product click, while blue and cyan in cool colors have a negative impact on consumers' click. Interestingly, we find that the number of colors presented in product images also influences consumers' click, with positive impact when the variance of the hue is low, and with negative impact when the variance of the hue is high. The results of our study offer managerial insight for merchants to display product images.

Some theoretical ways in which forecast value added analysis can be misleading and how this can be remedied.

Presenter: Paul Goodwin
Co-authors: Paul Goodwin;

Forecast value added (FVA) analysis has recently received a lot of attention in the literature. FVA is potentially useful when assessing whether judgmental adjustments made to algorithm-based forecasts are improving the value of these forecasts and whether these improvements are justified after taking into account any increased costs. This talk will consider a number of theoretical ways in which the results of an FVA analysis can be misleading. It will examine the conditions where adjustments that are associated with a negative FVA still have the potential to add value to forecasts and how this can be detected. It will link these conditions to particular causes of judgmental errors and the use of metrics that are based on inappropriate loss or cost functions.

Forecasting using expert causal models

Presenter: Fergus Bolger
Co-authors: Fergus Bolger;

Real-world forecasting is often messy. Key data may be scarce or non-existent and, where available, noisy. Further, considerable uncertainty regarding aspects of the forecast model, such as, the relevance of variables, and the nature of their interrelations, is commonplace. Also, there may be causal loops such that future actions and events "" including those resulting from policies based on our own forecasts"" affect the things we are trying to forecast. All these factors can make forecasting with traditional statistical methods problematic.A possible solution is to elicit expert causal models of the problem domain in the form of causal Bayesian networks (BNs); these provide powerful tools for representing and manipulating models of the world relevant to real-world forecasting and "" most importantly"" they permit the integration of "˜hard' data with judgment within a normative framework. Specific advantages of BNs include the ability to represent: rationales and confidence; causality; and probabilistic connections. Further, BNs permit sensitivity analysis, allow conclusions to be updated in the light of new evidence, and are adaptable to changing circumstances, such as new policy initiatives.Unfortunately, BNs are difficult to construct for

those without specialist modelling skills. For this reason, we have developed protocols based on Delphi"“ a structured-group technique ““ that allow elicitation of the elements needed to build a BN. Using these protocols, several domain experts, interacting through a facilitator, can collectively build a BN in a series of steps with feedback between each Delphi"round" . This can be completed with very little training of the participants in Bayesian modelling.In this paper I report the results of a study where our protocols are used to elicit BNs from a small group of participants for the causes of violent crime in a city. These models are then used to forecast changes in levels of crime in the city resulting from proposed policy changes (e.g. in policing, zoning, drug legislation). Forecast performance of the elicited judgmental models is evaluated in terms of accuracy and calibration against the results of simulations using a model derived from data from a real city.

Stylised facts of forecast value added, a meta-analysis' where do judgmental adjustments improve accuracy?

Presenter: Robert Fildes
Co-authors: Robert Fildes; Paul Goodwin

This paper examines five previously published data sets capturing product sales where judgmental adjustments have been applied to statistical forecasting models. We aim to understand where such adjustments prove beneficial and "˜Forecast Value Added: FVA' is positive. Various hypotheses have been advanced in the literature which include the differential effects of adjustment direction, the effects of adjustment size, and over-adjustment: these will be examined across data sets to establish robust conclusions. In addition, new hypotheses concerning the differences between organizations which achieve positive FVA and SKUs which prove amenable to judgmental adjustment will be explored. Despite the ubiquity of judgmental adjustments to demand data in practice and the growth of increasingly sophisticated forecasting methods aimed at offering a more effective route to improved accuracy, I will argue that there is still a lot of questions still need answering if demand forecasting is to effectively combine judgment with statistical models.

Risk Management in Wholesale Electricity Markets: A Signal Processing Approach

Presenter: Ritvana Rrukaj
Co-authors: Ritvana Rrukaj; Leif Sandal; Benjamin Fram

Because wholesale electricity markets exhibit high price volatility, market participantsroutinely hedge price risks through the use of financial derivatives, such as by acquiringfutures contracts that allow them to buy or sell electricity for a pre-determined price over aset time period. Taking the perspective of a market participant, this paper applies techniquesfrom the field of signal processing to map medium-term relationships between load and pricesin the New York State wholesale electricity market (NYISO). We first sort and filter rawmarket data from the years 2006 through 2018 into annual, seasonal (90-day), and weeklymoving averages. We then develop three nonlinear waveform models that relate the movingaverages of load and prices while accounting for seasonality in the data. As a next step,we extract the waveform from the filtered data and employ the Box-Jenkins methodologyfor time series analysis to model the resulting noisy residuals. We conclude by performingout-of-sample forecasts of the seasonal moving average model using data from 2019. Theforecast results indicate that our modelling approach is useful in predicting the relationshipbetween seasonal load and prices in the NYISO and may assist market participants to betterassess market risk and adjust their hedging positions.

SEE CROSS-BORDER TRADE ANALYSIS BASED ON ECONOMIC ELECTRICITY EXCHANGE

Presenter: Valbona Karapici
Co-authors: Valbona Karapici; Arsena Gjipali

AbstractA key development in the southeast European energy sector is the agreement between Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Greece, Macedonia, Kosovo, Romania, Serbia, and Montenegro to develop a regional electricity market. There are benefits that would arise from competition and co-ordination in a regional electricity market considering the diverse resources of the countries involved, difference in demand shapes and the possibility for sharing capacity reserves. Benefits would be displayed in the form of lower end-user prices for a given level of system security. In this context, one key issue is the regional electricity balance and potential for cross-border trading between the countries in the region. Under this framework, the concept of a Balkan Benelux has been developed regarding regional energy co-operation and views on the western Balkan six (WB6) countries initiative: Albania, Bosnia and Herzegovina, Kosovo, Macedonia, Montenegro and Serbia.The purpose of this paper is to provide an analysis of cross-border trade in SEE based on economic electricity exchange. This is achieved through the following objectives: first, an overview is made of the available energy and economic data in the region; second, a model is developed for regional least cost expansion planning when allowing for cross-border trade. The former aims to assess electricity supply and demand in the region with the purpose of making a comparative analysis regarding energy resource endowments. The second aim is achieved using the Solver software package in an optimisation model to estimate the least cost generation expansion plan required to meet the regional electricity requirement and evaluate the impacts of trade and co-operation in the region. The model system is based on maximising the benefits of cross-border trading to ensure cost reductions and improved system security, based on projections made by the regional power utilities. Key words: cross-border trade in SEE, regional least cost expansion planning, forecast.

Forecasting of system imbalance (volume and direction) in a real-time balancing market

Presenter: Konstantinos Plakas
Co-authors: Konstantinos Plakas; Athanasios Bachoumis; Nikos Andriopoulos; Panayiotis Alefragis; Alexios Birbas; Alex Papalexopoulos

The rapid change in the energy mix of power systems across Europe, has created several technical challenges in the system. One of the main challenges for the system operators remains the task of maintaining balance between production and consumption (minimize system imbalance) and thus to ensure the stability of the grid. The above challenge is at risk due to the constant increasing penetration of renewables into the grid, potentially leading to high system imbalances. The volume and direction of system imbalance are crucial parameters, both for system operators and market participants. In this work we are focusing on the market's participant side. An accurate forecast of the imbalance volume (regression problem) and the direction (classification problem) would allow the market player to contribute more efficiently to the system imbalance and thus be rewarded for that. The selected case study, is the real time balancing market of Belgium, operated by ELIA. Single pricing system is implemented whilst the gate closure time of the market is set 15 minutes prior to the delivery. The main principle in this market is that the imbalances should be cleared by Balancing Responsible Parties (BRPs) so as to limit the overall system imbalance to be cleared out by ELIA. To achieve that, two different Random Forest models are implemented both for regression and classification problem. For regression, the target variable is the system imbalance volume (MW) whilst for classification there are 2 possible classes (either positive or negative). Preliminary results have shown that interesting conclusions can be drawn by the correlation of system imbalance with different type of predictors. We will present the implemented models and illustrate the benefits compared to classical time series forecasting methods.

A new approach for change-point detection and reconstruction in time series forecasting

Presenter: Nachiketas Waychal
Co-authors: Nachiketas Waychal; Nachiketas Waychal; Arnab Laha; Ankur Sinha

In this paper, we extend the Adaptive Ensemble Generator (AEG[m,n]) algorithm for forecasting time series in the presence of multiple evaluation criteria and user preferences for situations when change-point(s) may be present in the series. Using a simple scanning heuristic, new sequences of dependent observations are derived and the problem is reduced to the detection of outliers in these newly derived sequences. An extensive simulation-based study is performed to evaluate the performance of the proposed method in the presence of multiple user-provided forecast evaluation criteria both when change-point is present and absent. Further, an empirical illustration of the performance of the method on a database of real-life time series data sets is provided.

Loss Function-based Change Point Detection in Risk Measures

Presenter: Xiaohan Xue
Co-authors: Xiaohan Xue; Xiaohan Xue; Emese Lazar; Shixuan Wang

We propose a new test to detect change points in risk measures, based on the cumulative sum (CUSUM) procedure applied to the Wilcoxon statistic of the FZ loss function class of Fissler and Ziegel (2016). The proposed test efficiently captures change points jointly in two risk measure series: Value-at-Risk (VaR) and Expected Shortfall (ES). In particular, we derive the asymptotic distribution of the proposed statistic. We also adopt a stationary bootstrapping technique to obtain the p-values of the test statistic. Monte Carlo simulation results show that our proposed test has better size control and higher power than the alternative tests under various change-point scenarios. The alternatives considered include change point detection methods based on self-normalized CUSUM statistics for the VaR series (Hoga, 2017) and the ES series (Fan et al., 2018) taken individually and a modi

fication of our proposed test using a statistic based on Renyi-type formulation. An empirical study of risk measures based on the S&P 500 index illustrates that our proposed test is able to detect change points which are consistent with well-known market events.

Forecasting Time Series using Trend Indicator Saturation

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

Trend indicator saturation is based on the recent developments of indicator saturation estimators, whereby significant trend indicators are selected from a saturating set to capture broken trends. Trends are essential to pin down accurately when forecasting. Extrapolating linear trends that are poorly estimated at the forecast origin, or that do not persist over the forecast horizon, lead to forecasts that deteriorate as the forecast horizon increases. There is a large literature that dampens trends to avoid this problem. In this paper we explore two aspects, (i) the importance of modelling trend breaks in-sample to ensure the 'correct' trend is extrapolated over the forecast horizon, and (ii) dampening the estimated trend by use of smoothed robust forecasts to address trend breaks or a lack of trend persistence over the forecast horizon. We analyse top income share data from a range of developed countries for data spanning almost a century. We identify many trend breaks and level shifts using saturation methods, and produce multi-step forecasts for top income shares using saturation estimation models, smoothed robust forecasting models, and Cardt, our univariate forecast method based on autoregressive models with damped trends, developed for the M4 forecasting competition. Our empirical results highlight the need to model in-sample trend breaks, allowing

for the full sample to be used rather than restricting the sample to post-break observations, and careful treatment of the trend for multistep forecasts when trend persistence can change over the forecast horizon.

Algorithm aversion or algorithm appreciation ?

Presenter: Shari De Baets
Co-authors: Shari De Baets; Nigel Harvey

Forecasting algorithms provide a significant opportunity for forecasters to improve accuracy. However, once people have seen an algorithm err, they are quick to abandon it. This phenomenon is known as "˜algorithm aversion' (Dietvorst, Simmons, & Massey, 2015). However, recently, Logg, Minson, & Moore (2019) have reported the opposite effect: results from six experiments showed that people preferred the advice of an algorithm over that of a person. The authors termed this"˜algorithm appreciation'. How can these findings be reconciled? One possibility is that algorithm appreciation occurs before people see the algorithm perform but, after seeing it perform (and err), they display algorithm aversion. Thus, we focus feedback. We set up experiments wherein AR-based forecasting models and human judgment models with similar performance metrics are overlaid on a historical sales time series in phase 1. Participants make a choice between both types of forecasting models (computer versus judgment), indicating their preference. The latter is measured two ways: with a binary choice, and with a bipolar slider indicating the degree of preference for algorithms (algorithm appreciation or judgment (algorithm aversion). In the first experiment, it will be limited to this process (no feedback condition). In a second version of the experiment, participants get immediate performance feedback on both methods, after indicating their preferences (feedback condition). Reconciling the findings of Logg et al. (2019) and Dietvorst et al. (2015, 2018), we hypothesize that the participant's preference changes from algorithm appreciation to algorithm aversion as trials progress. They will see both forecast types err, but should be more unforgiving towards the algorithmic one (Dietvorst et al., 2015). Subsequently, to investigate whether any preference was based on a pre-existing belief (e.g., that judgment is better) or based on the characteristics of the judgmental and statistical forecast series, we use the same experimental set-up as the first one described above (no feedback condition), but swap the labels of the two types of forecast. This setup serves to answer the question if people stay with the same label or with the same type of series, indicating another possible cause (next to feedback) of algorithm aversion versus appreciation

"Using judgmental forecasting and scenario thinking for anticipating the future: what are the differences, the similarities, and the advantages of each?".

Presenter: George Wright
Co-authors: George Wright;

This paper focuses on the areas of judgmental forecasting and scenario thinking - alternative approaches for anticipating the future. Topics discussed are: what are the differences, the similarities, and the advantages of each? Both approaches have, to a large extent, developed independently of one another. The first is the domain of psychologists interested in forecasting and the second, until recently, the domain of practitioners interested in helping organisations make better decisions in the face of uncertainty. The paper discusses the role of subjective probability and outcome verification, and the focus on single point estimates as opposed to creating multiple broad-brush futures. Until recently, very little academic scenario research has used experimental techniques to evaluate the quality of developed scenarios, but this position is changing. Additionally, recent work within the judgmental forecasting tradition has combined judgmental prediction within scenario story-lines , a focus that has also become part of practice - for example, within the UK National Grid energy scenarios that are used for national policy-making. Clearly, scenario thinking is now becoming strongly established in practice - perhaps as a response to the World's lack of preparedness for Covid-19 pandemic - and is the focus of quite intense social science-based research.

Judgmental interventions: model tuning and forecast adjustments in a retailing case study

Presenter: Anna Sroginis
Co-authors: Anna Sroginis; Nikolaos Kourentzes

With the increase in data sources and frequency of decision-making, organisations see an increase in the volume of forecasts that needs to be generated. Nonetheless, several studies verify that human intervention in forecasting remains a common practice. There are several ways that experts can augment statistical forecasts with judgment: (i) adjusting forecasts individually for a single item; (ii) batch-adjusting: correcting several time-series or categories at the same time; (iii) model tuning, indicating a location of corrections rather than size and feeding it to a statistical model, for example, by introducing indicator variables in a regression model. The literature has explored the first category extensively, but much less the other two. Yet, these are easily scalable for many products at once, making it easier and faster for forecasters to implement changes. There is limited research on the effectiveness and performance of these approaches. Furthermore, due to their ease of use, both batch-adjustments and model tuning might be overused and, as a result, potentially lose their effectiveness. For instance, in model tuning, introducing indicators for spurious events may result in overfitting rather than augmenting the statistical models, which increasingly employ more sophisticated algorithms. Using a case study from a UK retailer, which exhibits all three behaviours of interventions, we provide empirical evidence of the efficacy of these alternatives, as well as exploring the conditions where each alternative may be preferable.

The power of text-based indicators in forecasting the Italian economic activity

Presenter: Juri Marcucci
Co-authors: Juri Marcucci; Valentina Aprigliano; Simone Emiliozzi; Gabriele Guaitoli; Andrea Luciani; Libero Monteforte

Can we use newspaper articles to forecast economic activity? Our answer is yes and, to this aim, we propose a brand new economic dictionary in Italian with valence shifters, and we apply it on a corpus of about two million articles from four popular newspapers. We produce a set of high-frequency text-based sentiment and policy uncertainty indicators (TESI and TEPU, respectively), which are timely, not revised and computed both for the whole economy and for specific sectors or economic topics. To test the predictive power of our text-based indicators, we propose two forecasting exercises. First, using Bayesian Model Averaging (BMA) techniques, we show that our monthly text-based indicators greatly shrink the uncertainty surrounding the short-term forecasts of the main macroeconomic aggregates, especially during recessions. Secondly, we employ these indexes in a weekly GDP growth tracker, delivering sizeable gains in forecasting accuracy in both normal and turbulent times.

High frequency indicators. Why? When? and How? A users' guide

Presenter: Simona Delle Chiaie
Co-authors: Simona Delle Chiaie; Gabriel Perez-Quiros

This paper studies the role of high-frequency data to estimate the current state of the economy. Our research question is motivated by the evidence that since the early weeks of the coronavirus crisis, several high-frequency indicators provided meaningful signals of mounting economic damages well before the release of standard monthly indicators. High-frequency indicators provide timely but noisy signals of economic activity. However, when economic conditions do not change rapidly, they can be seen of a second-order importance compared to the common signal provided by standard monthly indicators. The covid-19 crisis opens the question of whether, in specific circumstances, the informative content of these indicators suddenly increases. While the answer to this question is merely empirical related to the

performance of these models in the recent recession, the unprecedented nature of the Covid-19 crisis poses a challenge for any empirical analysis based on past data. In this paper, we study the role of high-frequency indicators in nowcasting economic activity, relying on a simulation exercise. A controlled experiment that considers many possible situations allows us to derive some general lessons on the role and use of high-frequency data that are not conditioned by the quality of the data or the sample and country chosen. Moreover, the paper also illustrates our findings with an empirical analysis applied to US data. Our results show that high-frequency data are occasionally useful, with their importance clearly emerging in crisis times. We find that when economic conditions change abruptly, models augmented with high-frequency indicators can capture the shift in economic conditions more rapidly than models based on slower variables only. The empirical exercise based on actual data validates the simulation results. Our results show that a nowcasting model augmented with the signal from high-frequency data would have provided a timelier and accurate prediction of the depth of the Covid-19 recession.

Seismonomics: Listening to the Heartbeat of the Economy

Presenter: Luca Tiozzo
Co-authors: Luca Tiozzo; Elisa Tosetti

Seismic sensors continuously record a wide range of ground vibrations that are not necessarily related to earthquake activity, but are rather caused by human activity such as industrial processes, urban and air traffic. In this paper we show that human-generated seismic noise provides valuable information about the economic developments of a particular area, thus offering policymakers a useful tool for monitoring the heartbeat of that economy. We adopt a set of techniques developed within the seismic literature to disentangle sources of ground motion and propose a novel, daily indicator measuring vibrations caused by human activity. To demonstrate the usefulness of our procedure, we collect a huge data set made of nearly 20 years of continuously recorded seismic data in Beijing, China, and use our vibration indicator to forecast daily variations in regional industrial production.Our findings suggest that seismic data closely tracks business cycle fluctuations, with significant enhancements in the forecasting performance during economic crises. Our results support the usefulness of seismology as a nowcasting and forecasting tool in the area of economics and business, particularly for monitoring regional economies, for which timely and up-to-date indicators of economic activity are often not available, or for tracking national economies in periods of disruption of many key statistics.

Forecasting Low Frequency Macroeconomic Events with High Frequency Data

Presenter: Ana Galvao
Co-authors: Ana Galvao; Michael Owyang

High-frequency financial and economic activity indicators are usually time aggregated before forecasts of low-frequency macroeconomic events, such as recessions, are computed. We propose a mixed-frequency modelling alternative that delivers high-frequency probability forecasts (including their confidence bands) for these low-frequency events. The new approach is compared with single-frequency alternatives using loss functions adequate to rare event forecasting. We provide evidence that: (i) weekly-sampled spread improves over monthly-sampled to predict NBER recessions, (ii) the predictive content of the spread and the Chicago Fed Financial Condition Index (NFCI) is supplementary to economic activity for one-year-ahead forecasts of contractions, and (iii) a weekly activity index can date the 2020 business cycle peak two months in advance using a mixed-frequency filtering.

When Two Become One: Integrated Forecasting and Optimisation in the Newsvendor Problem

Presenter: Congzheng Liu
Co-authors: Congzheng Liu; Adam Letchford; Ivan Svetunkov

Newsvendor problems form a classical and important family of stochastic optimisation problems. The standard approach decomposes the problem into two steps: forecasting to get demand distribution, then determination of the optimal production quantity (or quantities) for the given distribution. We propose a new, integrated approach, which estimates the optimal production quantity directly from the data. Extensive experiments, on both simulated and real data, show that the proposed approach performs at least as well as conventional ones when applied to linear newsvendor problem, and performs better when applied to the nonlinear one. We also provide evidence that our approach is more robust with respect to model misspecification than the existing approaches.

Evaluating the impact of business practices on inventory performance

Presenter: Evangelos Theodorou
Co-authors: Evangelos Theodorou; Evangelos Spiliotis; Vassilios Asimakopoulos

Achieving high levels of product availability is crucial for retail firms that aim to improve the experience of their customers, enhance their profitability, and strengthen their position in the market. However, product availability is typically subject to numerous business decisions, agreements, and restrictions that directly or indirectly affect inventory levels. As a result, identifying cost-efficient practices that ensure adequate levels of product availability becomes a challenging task.To deal with this challenge, many firms focus on improving the accuracy of the methods used for forecasting demand and estimating sales uncertainty. Yet, more often than not, significantly more gains can be achieved through changes applied to the inventory practices considered. More importantly, these gains may vary for firms that sell products of different characteristics, each one requiring different review periods or target service levels. It becomes evident that, in order a company to be able to effectively revise its existing inventory practices and minimize risk, it has to quantify the expected impact of the changes examined, both at product level and in total.In this study, we introduce a novel methodological approach that assists retail firms revise their inventory practices in a direct, accurate, and computationally efficient way. Our approach consists of an off-line and an on-line phase. In the off-line phase we first generate numerous time series of various demand patterns to obtain a representative reference set of demand data. Then, assuming an order-up-to inventory control system, we use the reference set to conduct a large-scale simulation and assess the impact of different business practices on inventory performance, including different review periods, lead times, and target service levels, among others. Finally, we build a decision-tree-based model that relates inventory performance changes with inventory practices revisions. In the on-line phase, we use the constructed model to evaluate the effect of various inventory practices on the profitability of a major company using a large set of real series. Our results suggest that the proposed approach can be effectively used for supporting decisions as well as for conducting detailed optimizations.

Demand forecasting under lost-sales stock policies.

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

Demand forecasting is a crucial task within supply chain management. Stock control policies are directly affected by the precision of the demand probabilistic forecasts. For instance, safety stocks and reorder points are based on those forecasts. However, does the inventory assumption influence the forecasting model? In other words, if the inventory policy relies on the lost-sales assumption, should the forecasting

policy be modified accordingly? In this work, we investigate the different forecasting techniques as the Tobit Kalman Filter, which consider censored demand, to model the hypothesis of lost-sales situations. To the best of authors' knowledge, it is the first time that the Tobit Kalman Filter is applied to supply chain demand forecasting. Additionally, non-parametric methods as the Kernel density estimators will be also evaluated. Simulated and real data will be used to show the main results of this ongoing research.

Forecasting the all-time demand of spare parts with Bayesian hierarchical diffusion models

Presenter: Julius Mehringer
Co-authors: Julius Mehringer;

Predicting the all-time demand of spare parts for the time period after production ceased is a central issue in supply chain management in order to reduce costs in terms of storage, transportation, disposal, finance and to maintain a desired level of customer satisfaction. The main problem with modeling the future demand after this "End of Production" is that predictions about the whole product life cycle have to be made in the early stages of this life cycle. Thus, only few observations can be used to base the forecast on, which results in volatile and thus unstable predictions for single spare parts. In this paper, we propose a hierarchical Bayesian formulation for diffusion models in order to obtain valid predictions for the spare part demand during the whole life cycle. This model formulation allows us to 1) learn latent reference classes in order to base the forecasts on similar spare part demand patterns from the past; and 2) incorporate the few demand information already available to adequately scale the diffusion model. We validate the feasibility of our approach with simulation-based experiments. Finally, we evaluate our findings on a real world data set consisting of master and consumption data from 1985-2020 stemming from a big manufacturer of household goods and show that it yields significantly better forecasting results than a standard industry baseline model.

New product life-cycle forecasting with temporal hierarchies

Presenter: Oliver Schaer
Co-authors: Oliver Schaer; Nikolaos Kourentzes; Doug Thomas

Predicting new product sales with life-cycle curves have traditionally been applied to data of low-frequency nature, e.g. annually or quarterly observations. However, with big data, companies now often have access to sales data at higher frequencies, e.g. weekly or daily. While this naturally provides more data points, it can introduce seasonality, higher signal to noise ratio, and various irregularities. Although one can extend life-cycle curves, such as diffusion models, to capture seasonality, this can substantially increase model complexity and complicate the estimation of model parameters further. Furthermore, these additional high-frequency details can harm the long term predictive performance of the life cycle curves.To address these issues, we suggest using temporal hierarchies that use optimal suited time-series models at each aggregation level to extract model structure and subsequently combine it to increase predictive accuracy. For example, fitting a diffusion model at the quarterly level, with a long term focus, and a seasonal exponential smoothing model at the weekly level, with a short term focus. Combining these hierarchically results in a prediction that retains both aspects. Another benefit of using temporal hierarchies is that it allows continuing obtaining valuable model parameters, e.g. innovator and imitator coefficients of the Bass model, refined by the information in the more short-term and detailed models. We demonstrate the usefulness of the approach from a large computer manufacturer's dataset and provide insights on how to put these type of models into practice.

Cancelled: Challenges In Large Scale Hierarchical Demand Forecasting

Presenter: Nupur Aggarwal
Co-authors: Nupur Aggarwal; Brian Quanz; Sumanta Mukherjee; Vikas Raykar

Time series forecasting is a central problem encountered in many domains including retail, manufacturing, supply chain, workforce, finance, and health sciences. In many domains we are interested in analysing and forecasting for multiple loosely correlated time series. For example, the daily sales for category t-shirts in different stores, weekly country-specific commodity-prices, daily electricity consumption for different appliances across different locations etc. Very often the multiple time series are tied together by different levels of hierarchy. For instance, the products could be grouped into product hierarchies such as (department, category, and product) and stores there could be grouped as (country, region, and store). For planning, forecasts are needed at different levels of the hierarchy and incorporating hierarchy into the forecasting models generally improves the accuracy.Â For large datasets, hierarchical forecasting suffers from scalability issues. It is very common to have millions of time series (corresponding to thousands of products and stores). Due to the scalability issues of hierarchical time series, it is not employed in production systems even though they are generally proven to be more effective in dealing with multiple time series. In this paper, we present a cloud based architecture to scale hierarchical forecasting methods to large datasets. The proposed system can be deployed on any container application platforms such as Kubernetes or OpenShift. The system deploys containerised microservices using platforms such as Docker.Â It uses highly scalable databases such as InfluxDB or Elasticsearch for fast retrieval, and querying of time series datasets. It also employs distributed local storage such as MinIO to efficiently share data between different nodes of a cluster. Finally, we use state of the art tools like Ray to distribute computation of hierarchical time series forecasting across a distributed cluster.Â The predictions from the system are served using serving tools like Seldon or Falcon. The system is agnostic to the cloud provider and can work on on-premises, and hybrid cloud infrastructure also.Â

Predictive regressions under heteroskedasticity

Presenter: Robinson Kruse-Becher
Co-authors: Robinson Kruse-Becher; Matei Demetrescu; Christoph Hanck; Robert Taylor

Typical financial predictive regressions are characterized by time-varying volatility, high persistence and endogeneity leading to biased estimators and inefficiency. A feasible solution for the latter two problems is the so-called IVX procedure which builds on an instrument which is decisively less persistent than the predictor. These instruments are self-generated and are used to construct a test statistic whose critical values are taken from a standard limiting distribution. In this work, we tackle the important problem of time-varying volatility. While IVX predictive regressions can be robustified against time-varying volatility, we focus here on improving estimation and inference by using weighted IVX methods. The newly proposed procedure builds on a local non-parametric volatility estimator. Observations in the predictive regression are weighted according to the volatility estimator. It can be shown that under a set of weak assumptions regarding e.g. smoothness of the volatility function, the limiting distribution of weighted estimators and statistics remain the same. In addition, the behavior under local predictability alternatives is investigated. In our extensive Monte Carlo study, we first consider the estimation accuracy of standard and weighted versions under a set of different volatility patterns. It turns our that MSE ratios can be reduced up to thirty percent under heteroskedasticity, while there is almost no loss (up to one percent only) under homoskedasticity. Second, we focus on the problem of testing hypotheses about the slope coefficients. We study the empirical size and power and find that noticeable power gains are achievable under time-varying volatility, while the newly proposed tests perform well in terms of size. Furthermore, feasible and infeasible versions are nearly indistinguishable from each other. In an empirical application, we consider CRSP data for the equity premium and the logarithmic book-to-market ratio. We study both in-sample and out-of-sample predictability. While standard IVX tests do not indicate in-sample predictability from 1926 to 2018, the newly proposed heteroskedasticity-weighted tests clearly reject the null. For the out-of-sample

exercise from 2004 to 2018, we find that the weighted IVX estimator for the predictive regression provides the largest pseudo-R^2 measure in comparison to other approaches.

Multiplicative Non-Stationary Volatility Models with Exogenous Information

Presenter: Cristina Amado
Co-authors: Cristina Amado;

In this paper, we propose a multiplicative non-stationary volatility model allowing for nonlinear behaviour driven by exogenous information. The new model extends the time-varying GARCH model of Amado and TerÃ¤svirta (2013, 2017) by including an additional stochastic variable to allow the conditional variance to change smoothly between regimes. Modelling strategies for the proposed model are developed, and they rely on Lagrange multiplier tests. The estimation of the model is simpli

fied by employing maximisation by parts and the asymptotic properties of the proposed estimators are also studied. Finite-sample properties of these procedures and statistical tests are examined by simulation. An empirical application to agricultural commodity returns illustrates the functioning of the model in practice.

A time-varying graphical lasso approach to high dimensional portfolio selection

Presenter: Laura Reh
Co-authors: Laura Reh;

Gaussian graphical models are a powerful tool to estimate a high-dimensional inversecovariance (precision) matrix, which are particularly useful for portfolio selection in fi-nancial applications with »100 cross sectional units. Sparsity in the precision matrixtranslates to conditional independence among several pairs of returns which is a plausibleassumption particularly the context of financial systems that include large blocks of highlycorrelated assets inside industry sectors. In order to decode the underlying patterns infinancial asset returns, in this paper we adopt a time-varying graphical lasso (TVGL)approach which extends the graphical lasso to a dynamic context, such that it can takeinto account evolutionary patterns in the correlation structure over time. With the Alter-nating Direction Method of Multipliers (ADMM) we utilize a flexible framework in whichwe can solve the problem in an efficient way. Furthermore, it enables us to incorporateseveral constraints on the structure and the dynamic evolution of the precision matricesaugmenting the L1-type sparsity induced by the Lasso. Exploiting that the precisionmatrix, in contrast to the variance-covariance, is directly connected to several popularportfolio selection strategies like the Markowitz and the Global Minimum variance alloca-tion via a linear link function, we develop several problem-oriented penalty functions thatare tailored to specific portfolio selection strategies and which can incorporate specificeconomic requirements like limited gross exposure directly within the estimation processof the dynamic precision matrix. Finally, we show that our augmented time-varyinggraphical lasso (ATVGL) approach leads to improved forecasting performance in empir-ical applications to assets of the S&P500, in comparisons to the standard (time-varying)graphical lasso as well as compared to popular benchmark models from the literature.

Bayesian Optimization for Neural Networks with Small Sample Size Data: A Combined Approach for Feature Selection and Hyperparameter Tuning

Presenter: Nicki Lena Kämpf
Co-authors: Nicki Lena Kämpf; Sandra Spiegelberg; Jonas Krembsler; Nicola Winter; Thomas Winter; Robert Knappe

With the rise of Artificial Intelligence, the research in time series forecasting with Neural Networks is growing as well. However, there is a large gap between scientific research and the algorithms used in the

private sector. This gap is mainly caused by two problems: 1) the scientific studies mostly rely on large sample size data sets which are not available in practice and 2) the scientific papers seldomly focus on the importance of the feature selection.To narrow the gap between scientific research and the application of Neural Networks, this paper proposes a combined approach for feature selection and hyperparameter tuning with Bayesian Optimization. This combined approach allows for a generalizable, end-to-end training and optimization of Deep Neural Networks. Using a combined approach reduces the problems of overfitting and high-variance gradients for high dimensional, small sample size data sets. The advantages are demonstrated for a real-world revenue data set with a small sample size.The results will be used in a research project in the field of public transportation. The goal is to automate the revenue controlling and implement data-driven decision-making in the existing controlling processes.

Automatic time series feature selection for Neural Networks in Forecasting

Presenter: Dr. Sven F. Crone
Co-authors: Sven F. Crone; Tobias Kempcke; Heiko Kausch

Fully automatic specification of artificial neural network architectures is imperative for forecasting of large-scale industry data with hundreds or thousands of time series. Traditionally, neural network architectures apply only autoregressive lags as input features for univariate modelling. At the same time, network architectures of deep convolutional (CNN) and long-short-term memory models (LSTM) have shown promise in application domains of image, speech, and face recognition, by utilising architectural representations that implicitly extract input features from the input data. However, initial results in time series forecasting show only mixed results, questioning that the feature extraction in deep architectures may not be directly suited for time series data features. As an alternative, authors have suggested to explicitly generate time series features including seasonality, trends, level shifts and outlier with multiple transformations as inputs to shallow neural networks. In the case of explicitly generated time series features, the question arises if all features are needed, or automatic feature selection can positively affect forecast accuracy (and also runtime). We propose to use feature selection for multilayer perceptrons using linear statistics, checking their efficacy given data conditions of autocorrelated and multicollinear features. We assess the empirical accuracy on a representative dataset of a large sample of monthly industry data, using a valid experimental design of fixed-horizon multi-origin design and robust error measures. We compare the accuracy of shallow ANNs with innovations in statistical modelling of Exponential Smoothing variants, including automatic packages of es and ets, alternative modelling approaches of THIEF and TBATS, and advanced statistical methodologies of BSTS, Prophet, and XGBoost, all available via R. Our results show that shallow Multilayer Perceptrons show increased accuracy and also efficiency, outperforming advanced methodologies of BSTS, Prophet, XGBoost, and TBATS both in forecast error and runtime, making them also applicable in industry.

Cancelled: Optimize time series forecasting model using Artificial Neural Networks

Presenter: Marwan Ashour
Co-authors: Marwan Ashour;

Artificial neural networks (ANN) appear a powerful and effective tool in solving time series problems. The first aim of this paper is to diagnose better and more efficient networks of Back Propagation, Radial Basis Function Neural Networks (RBF), and Recurrent Neural Networks in solving the linear and nonlinear time-series behavior to get the desired results. The second aim disposes of the local minimum estimation problem; it is one of the very important problems of nonlinear models estimate. The goal of this paper is to discuss the performance efficiency of ANN and compare them to non-linear and linear time series models, used relative efficiency criteria and measuring accuracy of forecast Skill (SS), to diagnose any artificial neural networks more efficient and better accuracy in solving linear, nonlinear time series.

The second goal is testing the robustness of the forecasting accuracy. The very important finding reached in this paper was that the optimal neural network was the Back Propagation (BP) and Recurrent Neural Networks (RNN) to solve the time series, whether linear, semi linear, or non-linear. Besides, the result proved the inefficiency and inaccuracy (failure) of RBF to solve nonlinear time series, but their efficiency to solve linear or semi-linear time series only, this network's inability to is the disposal of local minimum problem estimation.

Forecasting the New Trends about the Consumer Behavior in the Cruise Industry post the COVID-19.

Presenter: Ana Silva
Co-authors: Ana Lucia Silva; Reinaldo Souza

The entire cruise industry has stopped due to the COVID-19 and a prohibition of resuming this industry was implanted worldwide. As a consequence of the cruise industry's importance, there are collective initiatives, such as the Healthy Sail Panel, the WHO's guidance, and the CDC framework among others that examined every aspect of the cruise journey and recommended new patterns to increase protection for all people involved. The routine on board for cruise ships is different in the cruise lines that are sailing post COVID-19. These protocols include screening travelers' temperature, filling health questionnaire, the COVID-19 swab test, and the medical review. Passengers should be wearing a mask, hand hygiene required everywhere, buffet restaurants should have crew members serving food, guests should be splited in sports and entertainment activities, and for most of the cruise lines, passengers should go onshore only if they book cruise lines' excursions. Some cruise lines are offering passengers wristbands to open doors and make the payment on board without physical contact. The ships should be better equipped with labs and health care. The itineraries available should be shorter and have fewer ports of call than usual, there is a cruise for nowhere in order to minimize the risks. In this beginning phase, the cruise lines are avoiding having different nationalities of passengers together and many ports of call. Also, they should be sailing with less than ship's capacity during the COVID-19. Until now, these new protocols, in addition to the mandatory COVID vaccine, have been bringing hope in this cruise industry resuming. However, it is crucial to reevaluate the trends before and after COVID-19 and how these trends could affect the customer journey and the experience of going on a cruise. Despite the fact that these new trends should be evaluated worldwide, in this paper it is considered only the forecast of expected number of passengers of Brazil. It will be used the available time series, as well as self-administered questionnaires with frequent Brazilians passengers, aiming information about the new desires about going on a cruise in the COVID-19 era.

Airbnb travel behaviour during the Covid 19 pandemic. The case of Switzerland.

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

PURPOSEBresciani et al. (2021) demonstrate that the pandemic creates a new need for physical distance between the accommodation host and the guest as well as among all guests, to avoid infection. This could increase the demand for specific kinds of Airbnb accommodation such as cottages, holiday flats and camping sites where people can keep to themselves. These kinds of accommodation are more common outside densely populated areas. This study aims to investigate the importance of property characteristics, population density and land use characteristics for the demand of Airbnb properties in Switzerland during the summer of 2020. RESEARCH METHODSEstimations are performed by use of panel spatial econometric models which allows to control for possible spatial effects of population density and characteristics of Airbnb properties nearby. The specification relates the change in the number of Airbnb bookings and revenues at the property level to characteristics of the Airbnb listing in the 2020 summer months to land use characteristics (surface covered by unusable land), population density as well as the spatially lagged variables. The data is a based on the full universe of Airbnb bookings for Switzerland comparing the

summer months 2020 with those of 2019 (with about 20000 observations for each month) provided by AirDNA.co.RESULTS AND DISCUSSIONResults based on the spatial Durbin error and spatially lagged X models show a significant link to the population density within the destination municipality as well as to the weighted average of the density of its neighbours in the 2020 summer months. This means that Airbnb revenues and bookings are lower is densely populated areas. This contrast demand for Airbnb listings those of the summers prior to 2020. Other characteristics such as the area covered by unusable land are also significant as well as the spatially weighted variables. Characteristics of the flat are also significant. IMPLICATIONSBefore the pandemic, cities exhibit a high growth rate of domestic and international overnight stays as well as demand for Airbnb properties. The Covid-19 pandemic has, at least temporarily, reversed this trend in urban tourism. During the crisis, areas with high population density experience the greatest decline in bookings and revenues for Airbnb accommodations. While part of the decline in Airbnb demand in high-density areas is due to the lack of international events and business meetings, it cannot be ruled

Forecasting and early warning of the holiday subway passenger flow in metropolis using attention-based LSTM model

Presenter: Xuerong Li
Co-authors: Xuerong Li; Fuxin Jiang; Shouyang Wang

Holiday travel has gradually become a worldwide popular lifestyle, which brings great pressure to the transportation system, especially in the case of nationwide holidays. Subway provides fast and cheap trips across urban tourist spots, serving as an effective transportation during holidays. The ridership explosion during holidays has triggered a series of issues, such as crowdedness in trains and the insufficient capacity of subway facilities. Therefore, forecasting subway passenger flows and building early warning systems under special events of Golden Week holiday benefits for both tourists' satisfaction and urban governance.This paper proposes an integrated framework for forecasting and early warning of holiday subway passenger flow based on deep learning models. An effective forecasting model of passenger flow considering external factors such as other stations and time of the day is constructed based on LSTM model. Furthermore, we propose a practical early warning framework integrating deep learning model results, which can be applied for the early warning of holiday metropolis subway stations around the world.Empirical results show that our forecasting model enables to forecast passenger flow from 15min to 1 hour ahead with favorable accuracy around 80%. Our early warning system show that the travel peak hour of Gold Week holiday is around 10 a.m. After that, we propose several emergent plans according to the risk level of our early warning system.Our study makes contribution in the aspects of: 1) Forecasting subway passenger flow within 1 hour based on attention-based deep learning model, with better performances compared to benchmark models; 2) Our early warning system provides three risk levels of subway passenger traffic, which helps subway stations to deal with peak passenger flows in advance. Our early warning system for nationwide holidays is hopefully applicable to almost all the metropolis around the world, i.e. New York, Tokyo, Hong Kong etc. 3) The integrated framework of forecasting and early warning for holiday subway passenger flow is contributing to emergencies responses and effective urban transportation management.

Comparison of different prediction methods for fare revenues in public transportation in Berlin

Presenter: Jonas Krembsler
Co-authors: Jonas Krembsler; Sandra Spiegelberg; Nicki Lena Kampf; Thomas Winter; Nicola Winter; Robert Knappe

We present first results from a case study of fare revenue forecast in public transportation in Berlin. The data is based on monthly fare revenues for different product segments. The results will be used in a research project in public transport with the goal of automating revenue controlling and implementing

data-driven decision-making in the existing controlling processes. The focus of this study is to obtain suitable and reliable predictions: on the one hand with autoregressive methods such as ARIMA, SARIMA as well as Holt-Winters Exponential Smoothing and on the other hand with methods that include exogenous variables such as SARIMAX, MLR, LASSO, Ridge, Random Forests, Gradient Boosting, and Neural Networks. The data concerning exogenous variables are freely available and cover a wide range from tourism data to labor market development and weather data. We discuss the different methods and compare the prediction results with common accuracy measures. The goal is to evaluate a wide range of different methods in order to decide in which situations they out- or underperform other methods.Besides simple prediction accuracy, another part of the study is the feature selection and interpretation of their impact. We address automatic feature selection using traditional approaches such as AIC optimization, a rolling window cross-validation approach optimizing the cv-error, and algorithmic approaches such as LASSO or Bayesian optimization. We discuss the interpretability of the results and the advantages and disadvantages of different approaches.

Intermittent demand forecasting in the Enterprise: Empirical verification

Presenter: Mariusz Doszyn
Co-authors: Mariusz Doszyn;

Forecasting methods are often valued by means of simulation studies. Forintermittent demand items there are often very few non“zero observations,so it is hard to check any assumptions, because statistical information is oftentoo weak to determine, for example, distribution of a variable. Therefore, itseems important to verify the forecasting methods on the basis of real data.The main aim of the article is an empirical verification of several forecastingmethods applicable in case of intermittent demand. Some items are sold onlyin specific subperiods (in given month in each year, for example), but mostforecasting methods (such as Croston’s method) give non”“zero forecasts forall periods. For example, summer work clothes should have non”“zero forecastsonly for summer months and many methods will usually provide non”“zeroforecasts for all months under consideration. This was the motivation forproposing and testing a new forecasting technique which can be applicableto seasonal items. In the article six methods were applied to construct separateforecasting systems: Croston’s, SBA (Syntetos”“Boylan Approximation), TSB(Teunter, Syntetos, Babai), MA (Moving Average), SES (Simple ExponentialSmoothing) and SESAP (Simple Exponential Smoothing for AnalogoussubPeriods). The latter method (SESAP) is an author’s proposal dedicated forcompanies facing the problem of seasonal items. By analogous subperiods thesame subperiods in each year are understood, for example, the same monthsin each year. A data set from the real company was used to apply all the aboveforecasting procedures. That data set contained monthly time series for aboutnine thousand products. The forecasts accuracy was tested by means of bothparametric and non”“parametric measures. The scaled mean and the scaled rootmean squared error were used to check biasedness and efficiency. Also, themean absolute scaled error and the shares of best forecasts were estimated.The general conclusion is that in the analyzed company a forecasting systemshould be based on two forecasting methods: TSB and SESAP, but the lattermethod should be applied only to seasonal items (products sold only in specificsubperiods). It also turned out that Croston’s and SBA methods work worsethan much simpler methods, such as SES or MA. The presented analysis mightbe helpful for enterprises facing the problem of forecasting intermittent items(and seasonal intermittent items as well).

Enhancing Short-Term Demand Sensing using Machine Learning.

Presenter: Charles Chase
Co-authors: Charles Chase;

Implementing a short-term (one to eight week) forecast is critical to understanding and predicting changing consumer demand patterns associated with sales promotions, events, weather conditions, natural disasters and unexpected shifts (anomalies) in consumer demand. Short-term demand sensing allows retailers and consumer goods companies to predict and adapt to those changing consumer demand patterns.This

session will demonstrate how machine learning combined with traditional time series techniques can be used effectively to enhance weekly and daily product demand forecasts. This new weekly forecasting methodology uses a combination of traditional time-series models and machine learning methods (Stacked NN + TS) to automatically choose the best model for each {Product "“ Ship to Location"“ Customer Location} combination. This new approach using machine learning establishes the efficacy of these methods by improving short-term forecasts for a large consumer products company. At the weekly and daily level, there was a significant improvement in forecast accuracy over existing forecasting procedures across multiple lag periods.The results also demonstrate that point-of-sale and customer inventory data further improves the daily and weekly forecast accuracy. We believe that our methods provide a flexible, transparent and scalable solution for effective supply chain management for CPG companies, which can also be utilized by retailers.

All You Need is Consistent Promotions

Presenter: Anneya Golob
Co-authors: Kanchana Padmanabhan

In retail demand planning, promotion planning is the task of deciding what products to promote, when to promote, and what promotion parameters (e.g., mechanic, discount, amount that the vendor/supplier will fund, etc.) to apply. Typically, the retailer manually estimates demand (# of units the product will sell) for many products under different conditions and picks the subset of products that provides the largest business benefit (e.g., revenue, market share) to the retailer. Optimize is an AI driven product, built by Kinaxis, that automates this process. The product allows for several “what if” promotional planning scenarios to be executed along with metrics that will help understand the benefit of each scenario. The retailers can use this product to complete their promotional planning tasks.An AI model built for this use-case needs to satisfy two important conditions of demand planners:1. Monotonicity: We want the estimated demand to be non-decreasing with the increase in value of certain promotion parameters. e.g., promotional discount. The demand predicted for a discount of 50% should be less than the demand predicted for a discount of 55%.2. Smoothness: A minor change in promotion parameters should correspond to a minor change in estimated demand. e.g., changing the discount value from 50% to 52% should not cause a significant change in the predicted demand output.While the above stated conditions seem simple enough, the underlying complexity of advanced machine learning models with highly complex retail data makes it hard to guarantee these conditions are met. In this talk, we will describe the problem in more detail, the issues with underlying data that makes this problem hard, and a model-agnostic methodology that can ensure the monotonicity and smoothness of the estimated demand.

Estimation of extreme conditional quantiles of wind speed: An application using South African data

Presenter: Caston Sigauke
Co-authors: Caston Sigauke;

Amongst several renewable energies, wind energy is known to be a pollution-free, cheap and clean energy resource. Due to its irregular, random and mercurial nature, wind energy offers challenges to system operators and decision-makers in power utilities. It is therefore important to know the highest possible energy which can be generated from wind. This study focuses on the prediction of extremely high conditional quantiles of wind speed, which is known to be the main driver of wind energy. The study applies a variety of methods in modelling and prediction of quantiles on the upper tail of the empirical distribution of wind speed data from Wind Atlas South Africa. The methods used in the study are extremal mixture, additive quantile regression, locally linear quantile regression, generalised additive extreme value and FKML GLD quantile regression models.

Long-term solar impact on groundwater in Long Island, New York

Presenter: Antonios Marsellos
Co-authors: Antonios Marsellos; Katerina Tsakiri

The Earth is considered an open system; however, if energy radiating mainly from the sun is constant, Earth may behave like a close system, while the reverse phenomenon may show otherwise. Here, we present a numerical model that shows a strong positive correlation between the daily change of sunspot number and groundwater. A time series model has been designed to analyze groundwater's long-term relationship with the daily sunspot number in several USGS groundwater monitoring stations in Long Island, New York. For the analysis, we use daily groundwater and sunspot number data between 2002 and 2021. The raw data show a weak correlation due to a low signal-to-noise ratio. A methodology is used to decompose the time series of all the variables into different time scales (long-, seasonal-, and short-term components). We use the Kolmogorov-Zurbenko (KZ) filter to decompose the time series, which separates the long-term variations from the short-term variations in a time series. The KZ filter provides a simple design and the smallest level of interferences between the time series scales. The long-term association between the groundwater and the sunspots reveals a negative correlation between the variables, which is consistent with the climate change over the last 200 years. However, the sunspots' change positively correlates with the groundwater's long-term component (r=0.85). We interpret this correlation as that sudden sunspot number changes may reflect stronger groundwater fluctuations not necessarily related to Earth's system mechanisms.

A novel forecasting system for medium term solar irradiance

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

Forecasts of Global Horizontal Irradiation are needed to obtain an efficient utilization of fluctuating energy output from photovoltaic plants. Most of the literature concentrates on the short term forecasting problem despite the need for forecasting horizons longer than 48 hours for tasks like unit commitment, transmission management, trading, hedging, planning, asset optimization, maintenance scheduling and spinning of power units. This paper aims at developing a medium-term forecasting method based on the adjusted combination of moving averages, which includes the yearly cycle although keeping hourly data. The proposed model, Adjusted Combination of Moving Averages (or ACMA), combines the information of short term and long term moving averages to form an annual profile that is finally re-scaled according to the last year of data available in order to produce the forecasts. The optimal number of days in the short term moving average is estimated by cross validation. ACMA is compared with several simple benchmarks (daily and yearly persistence), other traditional forecasting models (ARIMA and Dynamic Harmonic Regressions or DHR) and with deep learning models (LSTM and GRU), that are considered the indisputable option for medium-term forecasting in previous references. The results show that DHR and ARIMA models are the best for forecasting horizons shorter than two days, although GRU is the absolute winner for just a few hours ahead. However, the model proposed obtains the most accurate forecasts for horizons greater than 48 hours, overcoming the rest, with computing times that are considerably smaller. In addition, we have found out that the inclusion of the annual cycle in any model enhances the forecasting accuracy for forecasting horizons greater than 48 hours. The performance of all these methods are tested empirically on a dataset of 18 years (hourly sampled) obtained from a weather station located in Ciudad Real, Spain.

Optimizing Information and De-confounding Forecast Timing in Forecaster Assessment

Presenter: Mark Himmelstein
Co-authors: Mark Himmelstein; Pavel Atanasov; David Budescu

Several studies have demonstrated that forecasting is a unique and persistent skill. Forecasters who have been accurate in the past will tend to maintain their accuracy edge over time. It has also been demonstrated that psychological traits, such as intelligence and open mindedness, tend to correlate with individual forecasters' accuracy. The current study builds on this work and seeks to identify skilled forecasters in different information environments. Specifically, we study how to best leverage the information available as we learn more about forecasters' individual traits and history of forecasting performance over time. The issue of longitudinal assessment raises a further methodological complication: how do we compare the accuracy of forecasts made at different times about the same problem? Forecasters who make predictions about events far in the future are at a disadvantage compared to those who wait until they have more relevant information to rely on. As a result, the timing of forecasts can confound assessment of forecaster skill.Using methods based on hierarchical linear modeling and item response theory, we analyze results from the Hybrid Forecasting Competition (HFC) that lasted over two years and focused on forecasting geopolitical events. We demonstrate new ways of assessing the performance of forecasters, as information about them accrues over time, while accounting for differences in the timing of their forecasts. Results indicate that when information on forecasting performance is limited, psychometric traits are useful predictors of future accuracy, but that as performance information accrues, it becomes the stronger predictor of future accuracy. We also demonstrate that methods which account for forecast timing predict future accuracy better than those that do not. Accounting for forecasting timing and combining different sources of information optimizes performance weights in wisdom-of-crowds aggregation methods.

Using selected peers to improve forecasting accuracy

Presenter: Ye Feng
Co-authors: Ye Feng; David Budescu

Crowdsourcing approaches in which human forecasts are aggregated using optimal algorithms are becoming increasingly popular and are used in many contexts. In situations in which a large number of judges are offered the opportunity to predict multiple events, one often encounters large numbers of "missing" forecasts. We propose an approach that imputes the missing responses, based on the answers of other "similar" forecasters. Using each judge's recorded forecasts, we identify a group of "peer judges" , and we impute the missing forecasts based on the median forecast of these peers. For every forecaster, we measure two similarities with each of the other forecasters - Activity and Selection. Activity similarity measures the degree to which the two forecasters choose to answer same items, and Selection similarity measures how likely the two forecasters were to assign the highest probability to the same outcomes of these selected items. Various inclusion thresholds and different combination rules (Their union, intersection, etc.) can be applied for peer selection. We used data collected during a recent large-scale geopolitical forecasting tournament to illustrate the approach, test its feasibility, and quantify its benefits. The accuracy is measured by the Brier score. Results of preliminary analyses show that our method improves the collective accuracy under different inclusion thresholds and combination rules. We further used a subset of the data to test a dynamic version of the proposed approach that revises peer membership periodically. We found that methods that involve the peer forecasters are more accurate than imputation with the global median, and they are all beneficial compared to no imputation. Finally, we classified the forecasters into High, Medium and Low influence groups based on the total number of times their forecasts were included in the imputation process and compared the performance of the three groups. Analysis of the selected peers suggests that the proposed method is successful because it overweighs and propagates the responses of the most engaged and accurate forecasters. Influential peers also tended to score higher on various measures of intelligence and were better calibrated.

Models, Markets, and the Forecasting of Elections

Presenter: Rajiv Sethi
Co-authors: Rajiv Sethi; Julie Seager; Emily Cai; Daniel Benjamin; Fred Morstatter

The forecasting of elections is of broad and significant interest, but accurate prediction is notoriously challenging, and each approach to forecasting has significant limitations.We compare two different approaches to election forecasting, demonstrate the value of integrating them, and propose a method for doing so. One is a model-based approach that uses inputs such as polls and fundamentals to generate a probability distribution over outcomes. Daily state-level forecasts based on this model were published by The Economist for several months leading up to the 2020 presidential election in the United States. Over the same period, a market-based approach to prediction was implemented on a peer-to-peer exchange, PredictIt. Daily closing prices on this exchange, suitably adjusted, can be interpreted as probabilistic forecasts. Models and markets respond to emerging information in different ways. Most models are backward-looking by construction, calibrated and back-tested based on earlier cycles, based on variables selected prior to the forecasting exercise. New information is absorbed only when it starts to affect input variables. Meanwhile, markets are forward-looking and can rapidly incorporate novel information from essentially arbitrary sources. Neither the Economist model nor the PredictIt market showed clear superiority predicting battleground state outcomes for the seven months leading up to the election. Based on average Brier scores, models and markets have nearly identical performance. The market performs significantly better during the early part of the period, while the model performs better close to the election. The two approaches generate very different forecasts for individual states, and make errors of different kinds. As a result, a simple average of the methods outperforms both component forecasts when aggregating across dates and states, suggesting that more sophisticated synthetic approaches could be promising. We show how a hybrid prediction market'incorporating automated traders endowed with budgets, risk preferences, and model-derived beliefs'can be developed and explored experimentally. We also conduct a performance test, by examining the profitability of a bot that held and updated beliefs based on the model forecasts. This test can be used as a novel and dynamic criterion for a comparative evaluation of model performance more generally.

Quantile Forecast Optimality Testing

Presenter: Marc-Oliver Pohle
Co-authors: Marc-Oliver Pohle; Jack Fosten; Daniel Gutknecht

Quantile forecasts issued across multiple horizons have become an important prediction tool of many institutions such as central banks or international organizations. This paper proposes several misspecification tests for (method and model based) quantile forecasts. More specifically, we propose two tests that assess optimality of quantile forecasts over multiple forecast horizons and multiple quantiles. Our first test inspects the monotonicity of expected loss over horizons via a set of moment inequalities, while the second test builds on a multi-horizon version of Mincer-Zarnowitz quantile regressions cast into a moment equality or a Wald type test framework. Importantly, our tests do not just inform about general violations of optimality, but may also provide useful insights into specific forms of non-optimality. A simulation study investigates the finite sample performance of our tests, and empirical applications to the Bank of England's fan charts and to model-based Growth-at-Risk forecasts for Euro Area industrial production illustrate that our tests can yield interesting insights into forecast sub-optimality and its causes.

Quantile Regression and Predictive Distributions: With an Application Forecasting GDP Growth-at-Risk

Presenter: James Mitchell
Co-authors: James Mitchell; Aubrey Poon; Dan Zhu

Quantile regression methods are increasingly used to estimate and forecast tail risks and uncertainties in macroeconomic outcomes. But questions have been raised about their ability to do so. This paper proposes a one-step extension of a commonly adopted two-step quantile regression approach specifically

designed to deliver density forecasts. The proposed estimator lets us both accommodate parameter estimation uncertainties when forecasting and avoids fitting a specific parametric density to the quantile forecasts first obtained from the quantile regressions. We compare and evaluate alternative approaches, from both a frequentist and Bayesian perspective, to the production of density and tail risk forecasts from quantile regressions via simulations and an application forecasting US GDP growth.

Score-based calibration testing for multivariate forecast distributions

Presenter: Malte Knüppel
Co-authors: Malte Knüppel; Fabian Krueger; Marc-Oliver Pohle

Calibration denotes the "statistical consistency between the distributional forecasts and the observations" (Gneiting et al., 2007), and is often cast in terms of the probability integral trans-form (PIT). While PIT-based calibration testing is common in the univariate case, multivariate extensions face various challenges. We therefore propose two new approaches to calibration testing that cover univariate and multivariate predictive distributions alike. Both use proper scoring rules and are simple to implement even in high dimensions. One approach is based on comparing the expected performance of the forecast distribution (i.e. the expected score) to its actual performance based on realized observations (i.e. the realized score). The other employs the PIT of the score. We discuss the precise notion of calibration that is considered by the tests (auto-calibration) and propose a principled implementation. The tests have desir-able size and power properties in simulations. We illustrate the new approach using Bayesian (vector) autoregressive forecast distributions for economic time series data.

Calibrating and combining probability forecasts

Presenter: Benedikt Schulz
Co-authors: Benedikt Schulz; Tilmann Gneiting; Sebastian Lerch

Raw probability forecasts of binary events - including but not limited to predictions based on machine learning techniques - tend to be unreliable, and thus require calibration. Furthermore, there is overwhelming empirical evidence that the combination of individual forecasts improves predictive performance. We review and study parametric and non-parametric methods for the calibration and aggregation of probability forecasts, such as the traditional linear pool, the beta-transformed linear pool, beta and logit-normal calibration, isotonic regression via the pool-adjacent-violators (PAV) algorithm, and variants thereof. We assess the performance of the calibration and combination methods in simulation settings, and also in real data studies, based on classifier output from commonly used machine learning techniques.

Predictive properties and minimaxity of Bayesian predictive synthesis

Presenter: Ken McAlinn
Co-authors: Kenichiro McAlinn; Kosaku Takanashi

We examine and compare the predictive properties of classes of ensemble methods, including the recently developed Bayesian predictive synthesis (BPS). We develop a novel strategy based on stochastic processes, where the predictive processes are expressed as stochastic differential equations, evaluated using ItÃ´'s lemma. Using this strategy, we identify two main classes of ensemble methods: linear combination and non-linear synthesis, and show that a subclass of BPS is the latter. With regard to expected squared forecast error, we identify the conditions and mechanism for which non-linear synthesis improves over linear combinations; conditions that are commonly met in real world applications. We further show that a specific form of non-linear BPS (as in McAlinn and West, 2019) produces exact minimax predictive distributions for Kullback-Leibler risk and, under certain conditions, quadratic risk. A finite sample simulation study is presented to illustrate our results.

Quantile density combination: An application to US GDP forecasts

Presenter: Giulia Mantoan
Co-authors: Giulia Mantoan;

Often policy makers and practitioners wish to impose a desirable feature to predictive distributions (such as moments constraint, tails behaviour, shifts in support, etc…). Although constraining distributions is well discussed in literature (i.e. by exponential tilting), little study has been done on constraining through combination. Recently, Pauwels et al. (2020) investigates the advantages of high-moment constraints on optimal combination to forecast the SP&500 index. Compared to Pauwels et al. (2020), this paper allows the constraint to be any characteristic of the distribution and not just its moments. This forecast combination scheme assigns weights to individual predictive densities based on evaluation criterion and constrains them using Bayesian Importance sampling. This approach is applied to forecast US GDP growth under the Covid-19 pandemic: a combination of statical models is constrained to the higher probability of a negative growth event, recorded by survey of professional forecasters (SPF).

Automatic Forecasting with Gaussian Processes

Presenter: Giorgio Corani
Co-authors: Giorgio Corani; Alessio Benavoli; Marco Zaffalon

We propose an approach for automatic forecasting based on Gaussian Processes (GPs). GPs so far have not been successfully used for automatic forecasting, because of the difficulty of automatically identifying the optimal kernel for a given time series and because of the long training times. We avoid kernel search by considering a fixed additive kernel, which contains the components needed for modeling time series; it is the sum of a linear kernel (representing the linear trend): a periodic kernel (representing the seasonal pattern): an rbf and a spectral mixture kernel (representing the long-term trend). During training the unnecessary components are made irrelevant by automatic relevance determination (ARD). We assign priors to each hyperparameter in order sensibly estimate of the parameters: the priors allow the optimizer to discard unlikely region of the hyperparameter space. We define such priors through an empirical Bayes approach, by analyzing a separate set of time series. Thanks to the priors, the estimate of the hyperparameter is quite stable across different restarts. This removes the need for multiple restart; training the GP with a single restart on a monthly time series requires roughly the same time of auto.arima.We assess the model by measuring the mean absolute error (mae), the continuous ranked probability score (crps) and the log-likelihood (LL). Our model outperforms both ets and auto.arima on the quarterly and on the monthly M1 and M3 time series; it is instead outperformed on the monthly and quarterly M4 time series.The periodic kernel yields function which correspond to an infinite sum of Fourier terms. Thus our model can be applied to time series with non-integer seasonality (e.g., weekly time series). Moreover it can be applied to time series with multiple seasonality; to this end, we only need to add a second periodic kernel. We show results with time series characterized by non-integer seasonality and by double seasonality; the accuracy of our GP model is generally better than state-of-the-art methods such as TBATS and Prophet, both as as for the point forecast and the quantification of the uncertainty.

Monitoring Forecast Model Fitness Using Control Charts

Presenter: Joseph Katz
Co-authors: Joseph Katz;

Automatic forecasting software has been available for many years. These software tools will develop candidate models, incorporate events and independent variables, select the “best” model with optimized parameters, and generate forecasts’all without the need for human intervention. Yet the monitoring of model performance’determining when a model must be adjusted or discarded’has not reached a similar

level of sophistication.So when should a forecasting model be adjusted or replaced? This is an important consideration, because an inappropriate model results in avoidable forecast bias and error as well as consumes significant management resources making manual forecast adjustments.Until now, there has not been an automated capability for monitoring forecast model fitness for each time series to determine when models need to be discarded, adjusted/refit, or left unchanged. Automating this process becomes even more critical when the number of time series is very large. The two main questions that need to be answered are:1) Is it possible to evaluate forecast model suitability and viability for each individual time series as time passes, and as new data points are added to the historical data? 2) Can this forecast monitoring methodology be automated? The answer to both questions is yes. The approach described in this session demonstrates that it is possible to evaluate and automate the monitoring of forecast model fitness using control charts. Please note that this methodology is patented and owned by SAS Institute.

Fast and frugal time series forecasting

Presenter: Fotios Petropoulos
Co-authors: Fotios Petropoulos; Yael Grushka-Cockayne; Enno Siemsen

Over the years, families of forecasting models, such as the exponential smoothing family and Autoregressive Integrated Moving Average, have expanded to contain multiple possible forms and forecasting profiles. In this paper, we question the need to consider such large families of models. We argue that parsimoniously identifying suitable subsets of models will not decrease the forecasting accuracy nor will it reduce the ability to estimate the forecast uncertainty. We propose a framework that balances forecasting performance versus computational cost, resulting in a set of reduced families of models and empirically demonstrate this trade-offs. We translate computational benefits to monetary cost savings and discuss the implications of our results in the context of large retailers.

Forecasting Value-at-Risk and Expected Shortfall in Large Portfolios: a General Dynamic Factor Model Approach

Presenter: Carlos Trucios
Co-authors: Carlos Trucios; Marc Hallin

Beyond their importance from the regulatory policy point of view, Value-at-Risk (VaR) and Expected Shortfall (ES) play an important role in risk management, portfolio allocation, capital level requirements, trading systems, and hedging strategies. Unfortunately, due to the curse of dimensionality, their accurate estimation and forecast in large portfolios is quite a challenge. To tackle this problem, two procedures are proposed, being one of them based on a filtered historical simulation method in which high-dimensional conditional covariance matrices are estimated via a general dynamic factor model with infinite-dimensional factor space and conditionally heteroscedastic factors, and the other one based on a residual-based bootstrap scheme. The procedures are applied to a panel with concentration ratio close to one. Backtesting and scoring results indicate that both VaR and ES are accurately estimated under our methods, which outperforms alternative approaches available in the literature.

Comparison of macroeconomic, GARCH and neural networks approach on forecasting exchange rates' densities

Presenter: Gaëtan Le Floch
Co-authors: Gaëtan Le Floch; Miryan Constanza Bello Baracaldo; Qingyi Hu

This report aims to compare the predictive ability of USD/GBP and USD/JPY exchange rate macroeconomic fundamentals models, a GARCH process and a neural network. We compare one-step ahead density

forecasts from those models between them and with respect to a random walk. To evaluate performance, we use the Amisano and Giacomini (2007) weighted-likelihood test.

The reliability of equilibrium exchange rate models: A forecasting perspective

Presenter: Andrej Mijakovic
Co-authors: Andrej Mijakovic; Michele Ca'Zorzi; Adam Cap

In this paper we evaluate the predictive power of the three most popular equilibrium exchange rate concepts: Purchasing Power Parity (PPP), Behavioral Equilibrium Exchange Rate (BEER) and the Macroeconomic Balance (MB) approach. We show that there is a clear trade-off between storytelling and forecast accuracy. The PPP model offers little economic insight, but has good predictive power. The BEER framework, which links exchange rates to fundamentals, does not deliver forecasts of better quality than PPP. The MB approach has the most appealing economic interpretation, but performs poorly in forecasting terms. Sensitivity analysis confirms that changing the composition of fundamentals in the BEER model or modifying key underlying assumptions in the MB model does not generally enhance their predictive power.

Spatiotemporal Modeling with General and Geographical Covariates: Insights on Crime in Philadelphia

Presenter: Nicolò Bertani
Co-authors: Nicolò Bertani; Ville Satopaa; Shane Jensen

We explore whether crime depends on the urban geography (i.e. the individual components of the city, such as parks, hospitals, or churches). To do this, we extend Gaussian Process modeling to high-dimensional count data and devise a general procedure to relate any spatiotemporal phenomenon with important characteristics of its environment. The procedure encodes raw spatial objects (e.g the point location of a restaurant or the perimeter of a park) into geographical covariates. Geographical and general covariates (e.g. socioeconomic indicators such as income or population density) are then used to estimate how, how far, and how strongly they relate to crime. We consider the city of Philadelphia as a case study. Our results show that, when controlling for the urban geography, socioeconomic indicators are weak predictors. In particular, demographic information emerges as irrelevant to predict crime. In addition, the results indicate possible interventions to improve safety in the roughest parts of Philadelphia. Finally, we apply the model to make policy evaluation and assess the predicted effect on crime of the ongoing urban redevelopment in Sharswood-Blumberg. This study only relies on freely accessible non-commercial data, making it replicable for other municipalities.

Wildfire forecasting with deep neural networks

Presenter: Felipe Martins Choi
Co-authors: Felipe Martins Choi; Deive Ciro de Oliveira; Gabriel Rodrigo Gomes Pessanha; Reinaldo Antonio Gomes Marques

Wildfires have a noticeable impact on agriculture and public safety. Forecasting such events would enable preventive measures and improved insight for affected communities as well as private and public industries. The motivations behind the experiment include demonstrating the problem-solving potential of machine learning algorithms, exploring the possibility of forecasting fires for commercial use and providing an additional tool for tasks that involve monitoring ecosystems. Thus it was implemented as a system considering the tasks of fires forecasting. The data used for training and validating the model originates from fire event reports by INPE-queimadas and gridded weather forecasts by NOAA-GFS. The (balanced) sample for training and validation consists of 580 normalized samples of soil humidity, relative humidity of

air at 2m above ground, gust intensity and temperature, along with labels defining whether a fire event was detected (spatial and temporal resolution of 0.5Â° rectanguloid up to 6 hours later). A deep neural network model was used with 5 hidden layers of 100 neurons each, all ReLU-activated except for the final fire/no-fire output, which applies sigmoid. Training was then performed in a 5-fold cross-validation scheme over a shuffled copy of the original data. For comparison a similar scheme was applied to Random Forest, 3-Nearest-Neighbors and Support Vector Machine classifiers. The neural network achieved 86.38% accuracy. Compared to 78.8% for Support Vector Machines, 91.8% for Random forest and 87.7% for 3-Nearest-Neighbors, its accuracy is similar for most cases. The results of the experiment suggest that the weather strongly correlates to fires and encourage further investigation on the relationship between climatic conditions and probability of fire.An important issue of the work is the small sample size in order to train neural networks. Besides that the model does not take into account patterns across time, constraining maximum attainable accuracy. Future works may consider increasing sample size (temporal and spatial dimensions), adding time-series capabilities to the model and performing interpolation of weather grid points for improved performance.

Weather Patching: How good is good enough?

Presenter: Zachary Nordgren
Co-authors: Zachary Nordgren; Richard Povinelli

Historical weather data from the National Center for Environmental Information (NCEI) is a valuable dataset that is useful for energy forecasting and many fields of research. The data consists of thousands of weather stations geographically located around the globe that report hourly sensor measurements of many weather signals. The raw time series contain a high rate of missing data which make the data difficult to use for energy forecasting. The data covers a span of several decades and gaps range from a single hour to several months. There are many common ways of patching and error checking the historical records so they can be used to train other models. The methods compared in this paper are: imputed with mean, spline interpolation, multiple imputation by chained equation, and a recurrent neural network. Each of these methods is more complex than the previous method. This paper will answer the question of how well the various methods patch varying lengths of gaps, how the quality of the patching affects downstream forecasting models, and what data preparation methods are "good enough" when dealing with weather data.

Fan support in Major League Baseball: Can forecasting attendance improve game outcome predictions?

Presenter: Steffen Mueller
Co-authors: Steffen Mueller;

While it is widely acknowledged that fan support contributes to the home advantage effect in various sports, existing studies forecasting game outcomes only incorporate previous home game attendance as an outcome predictor, but they do not consider attendance forecasts. In contrast, this study explores the extent to which predicting crowd effects may improve game outcome forecasts in Major League Baseball. To this end, I use regular season data from 2010 to 2018 to make predictions for the 2019 season. Evaluating different statistical and machine learning methods, my preliminary result indicate that predicting fan attendance can improve game outcome forecasting accuracy; however, the improvement in forecasting accuracy largely varies across teams and appears to be mainly driven by positive attendance shocks.

Comparing penalized dynamic regression with time series methods for modeling and forecasting retail product sales

Presenter: Patrícia Ramos
Co-authors: Patrícia Ramos; Jose Manuel Oliveira; Nikolaos Kourentzes; Robert Fildes

Retailers depend strongly on accurate sales forecasts to manage their supply chains and make decisions concerning purchasing, logistics, marketing, finance, human resources, etc. Inaccurate forecasts of product sales can lead to stock-outs which are indubitably very negative for the business. If the product is not available in shelf its potential sales are lost and there is the chance of customers to look for it in the competitors, making loyalty difficult to maintain. Ordering excess inventory, to reduce the risk of stock-outs and to improve customer's satisfaction, increases costs significantly (e.g. labor and storage) reducing the profit margin. Additionally, there is an increased awareness that food waste should be reduced. In 2012 the European Parliament called for urgent measures to halve food waste by 2025 and to improve access to food for needy EU citizens. Efficient inventory management can be achieved with accurate forecasts of SKU (Stock-Keeping Unit) sales at the store level, which enable the retailer to replenish in time and meet the customers' expectations. However, how to develop forecasting models that consider the main drivers that affect demand still constitutes a major challenge.In this work we propose efficient and effective models to forecast retailer product sales at the SKU level. Working with a large dataset from a leading Portuguese retailer, we develop and test three different modeling approaches to forecast the demand during a promotion including routine against non-routine product categories. Integrated models that perform joint estimation of autoregressive components and promotional information outperform hybrid models that carry out this estimation on a two-step approach. On the other hand, this two-step approach performs best than univariate models which reveals the critical importance of external information. Models tested include ARIMA, TBATS, LASSO and Exponential Smoothing. This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within project UIDB/50014/2020.

Fitting three-parameter Weibull distributions with high negative skewness to generate monthly wind scenarios correlated with inflows to the Brazilian hydropower reservoirs

Presenter: Albert Melo
Co-authors: Albert Melo; Maria Elvira Maceira; Jose Francisco Pessanha

The Brazilian Electric Power System has been applying stochastic dual dynamic programming (SDDP) algorithms to solve the generation scheduling problem in the long-term operation planning. In the SDDP algorithm, system states are the storage of the reservoirs and the hydrological trend (past inflows to the hydropower plants). A Monte Carlo simulation scheme is used to iteratively construct multivariate expected cost-to-go functions representing the optimal operation policy. In the current SDDP implementation, a periodic auto-regressive model - PAR(p) - is used to generate the energy/water inflows scenarios that are used in the forward and backward passes of the SDDP algorithm and in simulating the system operation with the calculated operation policy. A large sample of Normal uncorrelated spatial residuals is generated and then a 3-parameter Lognormal distribution is fitted to them. Thus, the generated inflow scenarios are multivariate and present a skewed distribution similar to the original inflow process. A new development is underway to represent the uncertainties of wind production in the SDDP algorithm. To avoid increasing the number of state variables of SDDP algorithm, the wind speed will be modelled as a regression of the energy/water inflows of the current or previous periods plus a residual also comprising the cross correlation between hydropower plants and wind farms.Unlike the inflow residuals, the monthly wind speed residuals present negative skewness in several months, which prevents the use of the Lognormal distribution. Additionally, the monthly wind residuals present negative values. Therefore, one alternative is to use 3-parameter Weibull distributions.Several methods are available to estimate the shape, scale and location parameters of Weibull distributions. However, we found that the quality of the estimates provided varies

widely according to wind farms and the month of the year, with the worst performances associated with the months with high negative skewness.This work presents an approach to the modeling of monthly wind speed residuals through 3-parameter Weibull distributions, which present good performance when applied to Brazilian wind farms, especially in situations of high negative asymmetries. Case studies with the proposed approach are presented and discussed.

Forecasting Electricity Daily Load Profile Consumption in Brazil

Presenter: Daiane Oliveira
Co-authors: Daiane Oliveira; Soraida Aguilar; Reinaldo Souza

One of the most important aspects for electricity suppliers is the purchase and sale of energy for future demand. It is possible to affirm that the forecast of energy demand is essential for the efficient operation of any distribution utility, offering support for carrying out the planning and scheduling of energy production. The brazilian electrical matrix is predominantly renewable and this fact occurs because most of the electrical energy generated in Brazil originates from hydroelectric plants. According to the National Energy Balance (BEN) of 2020 (base year 2019), renewable sources account for 83% of the internal electricity supply in Brazil (considering the sum of the amounts related to national production and imports), 64.9% is represented by the water source. Consequently, electricity generation from non-renewable sources represented only 17.7% of the nation total in 2019 and 17.6% in 2018.Electricity load forecasting is known as one of the most important challenges in the management of electrical supply and demand and has been widely studied. This forecast is an important work to increase the energy efficiency and the reliability of the operation of the power system. In Brazil, the electricity market short-term prices are obtained through the Settlement Price for Differences, commonly called PLD and which represents the foundation for the free energy market. Nonetheless, since January 2021, the PLD started to be estimated on an hourly basis, according to the proposal made by the Permanent Commission for the Analysis of Methodologies and Computational Programs of the Electric Sector (CPAMP). Due to this fact, the present work aims to predict, through statistical models (Holt Winters Double Exponential Smoothing and Box & Jenkins Models), the new load profiles of consumers considering the effect of solar photovoltaic micro and/or distributed generation for the recent pricing dynamics. Thus, this work has the aim to forecast the day ahead hourly load/generation profile, since the introduction of the hourly rate requires that the daily load curves be accurately predicted.

Development of future biomass energy generation scenarios with intra-hourly frequency

Presenter: Gabriel Calvo
Co-authors: Gabriel Calvo; Paula Macaira; Reinaldo Castro

In the past few years, biomass power's share of the energy mix has grown significantly in Brazil. The biomass is considered a controllable renewable energy source such as dammed hydroelectricity, however, a particular characteristic of this source is the alternation between periods of zero generation and full generation. In addition, throughout the generation period, it is possible to divide it into at least three phases (initial, full, final). Based on this behavior, this work proposes a methodology, based on Markov Chain Monte Carlo and K-means clustering, to build future biomass energy generation scenarios with intra-hourly frequency. Preliminary results show that the division of the full generation period into three parts is promising, as well as the use of K-means to define the Markov states. Finally, the validation of the first results points to an adherence of the scenarios generated to the available historical data.

Considering Wind Uncertainties in the Long-Term Operation Planning of the Brazilian Power System

Presenter: Maria Elvira Maceira
Co-authors: Maria Elvira Maceira; Albert Melo; Jose Francisco Pessanha; Cristiane Cruz; Victor Andrade; Thatiana Justino

The Brazilian interconnected system presents a high share of renewable sources, accounting for more than 80% of the country's electricity consumption. Intermittent sources, especially wind, have experienced accelerated growth. For example, in 2019 the system had more than 15,000 MW of installed capacity in more than 600 wind farms and is envisaged to reach 39,500 MW by 2029. Despite the advantages, the intermittency of hourly wind generation, constitutes a challenge in terms of its integration in the electricity systems. Thus, it is essential to develop and improve methodologies to represent the uncertainties of intermittent renewable sources in long, medium and short term operation planning models.In Brazil, expansion and long term operation planning studies have been carried out since 1998 with the support of the NEWAVE model, which has been used in the routine and official activities of sector entities: generation dispatch by the National System Operator; calculation of the spot prices by the Whole Sale Energy Market Entity; expansion planning by the Ministry of Mines and Energy and the Energy Research Company; parameters of public auctions for the purchase of electricity by the Electricity Regulatory Agency; as well as by utilities of the power industry to develop corporate strategies.This paper introduces an approach to be used by the Brazilian power industry to represent the uncertainties of wind power production in the stochastic dual dynamic programming (SDDP) strategy used by NEWAVE model to solve the multi stage stochastic linear programming problem, The proposed methodology consists of four main stages: (i) statistical clustering of wind regimes; (ii) evaluation of monthly transfer functions (MTFs) between wind speed and wind production; (iii) generation of synthetic monthly wind speed scenarios considering a spatial correlation structure that also comprises, in addition to the spatial correlation between hydropower plants, the correlations between wind farms and the cross correlation between hydropower plants and wind farms; and (iv) obtaining monthly wind power production through MTFs, to be used in the SDDP algorithm.The proposed approach is applied in a case study with a real configuration of the Brazilian hydrothermal-wind system and the results are discussed.

Transfer learning for hierarchical forecasting: Reducing computational efforts of M5 winning methods

Presenter: Arnoud Wellens
Co-authors: Arnoud Wellens; Maxi Udenio; Robert Boute

The winning machine learning methods of the M5 Accuracy competition demonstrated high levels of forecast accuracy compared to the top-performing benchmarks in the history of the M-competitions. Yet, large-scale adoption is hampered due to the significant computational requirements to model, tune, and train these state-of-the-art algorithms.To overcome this major issue, we discuss the potential of transfer learning (TL) to reduce the computational effort in hierarchical forecasting and provide proof of concept that TL can be applied on M5 top-performing methods. We demonstrate our easy-to-use TL framework on the recursive store level LightGBM models of the M5 winning method and attain similar levels of forecast accuracy with roughly 25% less training time.In conclusion, we provide evidence for a novel application of TL to facilitate practical applicability of the M5 winning methods in large-scale settings with hierarchically structured data.

Forecasting with Deep Temporal Hierarchies

Presenter: Filotas Theodosiou
Co-authors: Filotas Theodosiou; Nikolaos Kourentzes

Temporal hierarchies are being increasingly used for forecasting purposes over the past years. They have shown to produce accurate and coherent forecasts which are beneficial for enterprises. Reconciling forecasts of different aggregation levels to achieve coherence, supports aligned decisions between different organizational levels. Current research focuses on analytical reconciliation methods which have shown to be more beneficial than conventional Bottom-Up and Top-Down approaches. However, such methods rely on a number of assumptions, primarily due to estimation requirements. This work proposes a novel approach for forecasting with temporal hierarchies. It results in a non-linear reconciliation method inspired by the architecture of an encoder - decoder deep neural network. A trainable encoder combines base forecasts into the reconciled bottom level predictions, while a decoder with fixed weights reconstructs the forecasts at all levels. The decoder ensures coherency and allows for a global loss minimization across all levels of the hierarchy. We suggest two alternative usages for the reconciler. One, to replace analytical expressions and reconcile base forecasts produced by models such as Exponential Smoothing. Second, as a part of a deep neural architecture which mimics the general framework for forecasting with temporal hierarchies. The proposed models are empirically evaluated using real data from different sources. Furthermore, we discuss the general effect of coherence on forecast accuracy and whether having a perfect reconciliation offers accuracy improvements.

A Flexible Factor Model and the Implications for Financial Modelling

Presenter: Necati Tekatli
Co-authors: Necati Tekatli;

This paper has two major contributions. First, we develop and implement a (latent) factor model, which we call the “flexible factor model”, that relaxes the orthogonality assumption in factor analysis, and allows for idiosyncratic correlations and correlation decompositions without any restrictions. Second, we explore the statistical and empirical implications of the flexible factor structure for asset pricing models, particularly, when there are omitted factors in the model. Running a comparative analysis between standard and flexible model structures, we fi

nd that the factor loadings - thus, the exposure to common (systematic) risk- are overestimated and unexplained variations -thus, the exposure to idiosyncratic (unsystematic) risk- are underestimated in a standard factor model. The

findings reveal that asset pricing models derived from a standard factor structure can give overly pessimistic or optimistic results for investors and analysts. The proposed methodology is applied to both the simulated data and the foreign exchange rate data.

Factor analysis and forecasting of data with heterogenous blocks of variables.

Presenter: Richard Haarburger
Co-authors: Richard Haarburger; Tatyana Krivobokova

It has been often empirically observed that including more variables into factor-based forecasting models may worsen the prediction considerably. We examine this issue assuming a factor model, which consists of heterogeneous and possibly dependent blocks of variables. We identify settings that cause the poor forecasting performance of such factor models estimated by principle component analysis (PCA) and suggest a simple modification of the standard PCA, called blocked PCA (bPCA), that leads to the proper identification of factors and prediction. Blocked PCA is a multi-step procedure, in which first a clustering algorithm is used to divide variables into blocks based on their statistical characteristics. Subsequently, from each block, a set of factors is extracted, which are used as a basis for factor-based now- and forecasting in the final step. Using Monte Carlo simulations designed to resemble macroeconomic now- and forecasting exercises, we demonstrate the greater robustness of bPCA over conventional PCA leading to improved fore- and nowcasting accuracy in the context of high dimensional data.

Forecasting with factor-augmented time-varying parameter models

Presenter: Douglas Eduardo Turatti
Co-authors: Douglas Eduardo Turatti;

In this paper we investigate the forecasting performance of time-varying parameter models (TVP) augmented with dynamic factors extracted from a large number of predictors. Hence, forecasts generated by the model are able to incorporate information from many predictors while retaining flexibility provided by time-varying parameter models. We consider TVP models with time variation in the following: stochastic volatility, temporal dependence, and cross-correlations with the dynamic factors. The model is estimated in two steps. In the first one, principal components are extracted from a large set of predictors and used as factor estimates. In the second step, the factors and the forecasting target dynamics are estimated in parallel. The factors' law of motion can be easily estimated by ordinary least squares. The forecasting target equation is a non-linear state-space model, and is estimated via simulated maximum likelihood using a novel technique that combines importance sampling and the Kalman Filter. Our modeling strategy thus marginalizes the computationally demanding estimation procedure only to the forecasting target equation and minimizes the curse of dimensionality. We apply the model to forecast several price indexes and real macroeconomic variables using factor estimates extracted from a large dataset of economic variables. Results indicate that stochastic volatility reduces forecasting errors for most real and nominal variables. Moreover, time-varying dependence and cross-correlations are most important for price indexes. For real variables, constant effects of the factors and stochastic volatility provide best forecasts. Our modeling specifications were able to outperform benchmark models for most nominal and real macroeconomic variables.

Automation of Causal Impact Measurement using BSTS Model at LinkedIn

Presenter: Xiaofeng Wang
Co-authors: Xiaofeng Wang; Ming Wu

Bayesian Structural Time-Series (BSTS) model is a widely used forecasting approach to measure the causal impact of a designed market intervention by forecasting the counterfactual KPIs of the treatment markets with the carefully chosen control markets. At LinkedIn, we have applied the BSTS model extensively for marketing campaign measurements when the member level targeting or tracking is not applicable. This approach is well received by not only data scientists, but also business partners. To ensure the quality of the causal studies and to improve the efficiency, we automated the solution by creating a web application to guide users in applying the approach to measure geo experiments and consume the results.In this talk, we will share our experience and learnings from applying the BSTS method, including common setup, popular use cases, challenges and limitations. In addition, we will discuss how to build a scalable platform for automating this forecasting method to drive the adoption.

Time-Series Forecasting with Random Curve Modeling Techniques

Presenter: Qiyao Wang
Co-authors: Qiyao Wang; Haiyan Wang; Chetan Gupta

In the era of big data, practitioners are frequently faced with the problem of forecasting a large number of related time-series with high resolutions. For instance, in day-ahead net-load forecasting problems in the energy industry, it's required to predict the electricity load for millions of residential customers using data with a resolution ranging from 1-second to 1-minute. Existing models have limitations in solving large-scale and high-resolution forecasting problems. The newly developed deep forecasting networks (e.g., deepAR) are designed to leverage the related patterns among multiple time-series. However, the recurrent architectures make them struggle in handling long-term dependencies, hindering their applications in high-resolution forecasting. To solve this issue, transformer models have been developed. Nevertheless, they

work only when all the time-series are equally spaced and share a homogeneous granularity. Motivated by functional data analysis in statistics, we consider time-series as a sequence of continuous random processes with a common support period. For instance, electricity load data with a 1-second to 1-minute resolution over 300 days is viewed as a sequence of 300 daily random curves, instead of a sequence of scalar-valued observations. The benefits of adopting curve techniques to tackle forecasting tasks are as follows. First, unlike other sequence models being cursed by the high-resolution, curve models enjoy granular information regarding the underlying continuous phenomenon. Second, working with the underlying curves can handle time-series with arbitrary granularity. Third, curve models effectively handle multiple layers of seasonality. Lastly, the curve modeling perspective allows access to derivatives and enables forecasts at any timestamps. We have applied the curve modeling techniques to forecasting problems across domains. Initial experiments have demonstrated the promising performance of our models.

Mid-flight Forecasting for CPA Lines in Online Advertising

Presenter: Hao He
Co-authors: Hao He; Tian Zhou; Lihua Ren; Niklass Karlsson; Aaron Flores

Verizon Media Demand Side Platform (DSP) is a system that helps advertisers to purchase online opportunities to show their ads, including display ads, video ads, and search ads, in an automated way. One major way is to purchase through real-time bidding in ad exchange, and the winning bidder gets the opportunity to serve their ad to users. For DSP, forecasting of ad campaign performance not only feeds key information to the optimization server to allow the system to operate on a high-performance mode but also produces actionable insights to the advertisers. The ad campaign performance, including the total number of impressions, total spend, effective Cost Per Action (eCPA), etc., highly depends on the quality and quantity of in-target inventory. However, without taking exact campaign-level bidding strategies and logic (such as max bid cap) into account, the forecasting results can be hundreds or even thousands of times off. In this paper, we propose a novel forecasting methodology that takes both inventories and bidding into account for conversion-based campaigns. First, we learn the joint distribution of two major factors of the campaign inventory, including bid price and bid values which are captured by the conversion rate. Then, forecasts are generated based on the joint distribution, combining with the bidding strategy. The proposed methodology generates relationships between various key performance metrics and optimization signals. It can also be used to estimate the sensitivity of ad campaign performance metrics to the adjustments of optimization signals, which is important to the design of campaign management systems. The relationship between advertiser spend and campaign eCPA is also characterized, which serves as a guidance for campaign adjustment in the middle of the flight. Several practical issues in implementation, such as downsampling of the dataset, are also discussed in the paper. At last, the forecasting results are validated against actual deliveries and demonstrates promising accuracy.

Forecasting FOMC Forecasts

Presenter: Jaime Marquez
Co-authors: Jaime Marquez; S. Yanki Kalfa

Hendry's (1980, p. 403) three golden rules of econometrics are "test, test, and test". The current paper applies that approach to model the forecasts of the Federal Open Market Committee over 1992—2019 and to forecast those forecasts themselves. Monetary policy is forward-looking and, as part of the FOMC's effort for transparency, the FOMC publishes its (forward-looking) economic projections. The overall views on the economy of the FOMC participants–as characterized by the median of their projections for inflation, unemployment, and the Fed's policy rate–are themselves predictable on information publicly available at the time of the FOMC's meeting. Their projections also communicate systematic behavior on the part of the FOMC's participants.

Forecasting US inflation in real time

Presenter: Chad Fulton
Co-authors: Chad Fulton; Kirstin Hubrich

We perform a real-time forecasting exercise for US inflation, investigating whether and how additional information – additional macroeconomic variables, expert judgment, or forecast combination – can improve forecast accuracy and robustness. In our analysis we consider the pre-pandemic period including the Global Financial Crisis and the following expansion – the longest on record – featuring unemployment that fell to a rate not seen for nearly sixty years. Distinguishing features of our study include the use of published Federal Reserve Board staff forecasts contained in Tealbooks and a focus on forecasting performance before, during, and after the Global Financial Crisis, with relevance also for the current crisis and beyond. We find that while simple models remain hard to beat, the additional information that we consider can improve forecasts, especially in the post-crisis period. Our results show that (1) forecast combination approaches improve forecast accuracy over simpler models and robustify against bad forecasts, a particularly relevant feature in the current environment; (2) aggregating forecasts of inflation components can improve performance compared to forecasting the aggregate directly; (3) judgmental forecasts, which likely incorporate larger and more timely datasets, provide improved forecasts at short horizons.

Extracting Information from Different Expectations

Presenter: Andrew Martinez
Co-authors: Andrew Martinez;

Long-term expectations are believed to drive future inflation. However, expectations are not directly observed and available measures present a range of values. To understand what drives these differences, we examine the evolution of survey and market-based measures. We show that inflation forecasts improve when incorporating the differences between them. Next, we decompose and extract the differentials in rigidity and information and find the information differential drives these improvements. Finally, using machine learning methods, we find that around half of this differential is explained by changes in liquidity. This explains forecast improvements and predicts the divergence in expectations in 2020.

Evaluating the Federal Reserve's Tealbook Forecasts

Presenter: Neil Ericsson
Co-authors: Neil Ericsson;

This paper examines publicly available Federal Reserve Board Tealbook forecasts of GDP growth for the United States and several foreign countries, focusing on potential time-varying biases and evaluating the Tealbook forecasts relative to other institutions' forecasts. Tealbook forecasts perform relatively well at short horizons, but with significant heterogeneity across countries. Also, while standard Mincer-Zarnowitz tests typically fail to detect biases in the Tealbook forecasts, recently developed indicator saturation techniques that employ machine learning are able to detect economically sizable and highly significant time-varying biases. Estimated biases differ not only over time, but by country and across the forecast horizon. These biases point to directions for forecast improvement.

Media Tone in Commodity Markets

Presenter: Nan Zhao
Co-authors: Nan Zhao; Ana-Maria Fuertes

We construct media tone measures for a cross-section of commodities in the main sectors-energy, agriculture, livestock, and metals-based on the textual tones retrieved from the news articles. Our media tone measures show strong explanation and in-sample and out-of-sample forecasting power for commodity futures returns even after controlling for hedgers hedging pressure, momentum, and carry factors. In addition, we design a simple portfolio strategy that can exploit the media tone measures and can significantly improve the performance of portfolios based on other signals. The predictability of our media tone measures generates from the information of the term structure of futures contracts and other behavioural measures rather than a risk channel.

Relationship between country risk volatility and indices based on unstructured information. Evidence for Argentina

Presenter: Martin Llada
Co-authors: Martin Llada;

This work assesses whether certain indicators constructed from unstructured information published in newspapers contain useful information regarding dynamics of Argentina's country risk volatility, estimated from a GARCH(1,1) model. The analysis covers the period 1998-2019. According to estimated models, the indicators that capture different manifestations of subjective states (e.g., uncertainty, fear, pessimism) in articles regarding the country risk topic provide valuable information regarding contemporary and future levels of country risk volatility. One standard deviation increment in the indicator that captures manifestations of pessimism is followed by an increment of approximately 0.2% in expected volatility of the risk premium in the consecutive quarter. Out-of-sample exercises confirm that these non-traditional indicators allow for gains in forecast accuracy. Moreover, the information provided by non-traditional indicators based on unstructured information is different from that provided by traditional macroeconomic indicators. These findings are robust to changes in the set of predictors, the specification of the model and the incorporation of new media content.

Bayesian Time Varying Coefficient Model with Applications to Media Mix Modeling

Presenter: Zhishi Wang
Co-authors: Zhishi Wang; Edwin Ng

Varying coefficient models are very useful tools in practice as they can be used to model parameter heterogeneity in a general way. Motivated by the need of enhancing Media Mix Modeling (MMM) at Uber, we propose a Bayesian Time Varying Coefficient (BTVC) model, equipped with a Bayesian hierarchy structure. This model is different from other varying coefficient models in the sense that the time varying coefficients are weighted over a set of local latent variables that follow certain probabilistic distributions. Markov Chain Monte Carlo (MCMC) sampling is used to estimate the dynamic coefficients. The proposed model also helps address some challenges faced by traditional MMM approaches. We used simulations as well as real-world marketing datasets to demonstrate our model's strength in terms of both accuracy and interpretability.

The impact of decomposition on the forecasting performance of Bagging

Presenter: Xinyang Liu
Co-authors: Xinyang Liu; Anyu Liu; Gang Li

Previous studies have shown that bootstrap aggregation (i.e., bagging) is an effective way to improve the forecasting performance. For time series bagging forecasting, the bootstrap procedure relies on

time series decomposition methods. Time series decomposition is based on a parametric or non-parametric method to decompose seasonal time series data into several components and offers multiple perspectives to analyze the original data. In the practice of time series bagging, the bootstrapped series are usually the re-aggregate of the original trend, seasonality, and the bootstrapped residuals. Tourism demand data are often non-stationary and easily affected by holidays or one-off events, which cause complex seasonality and large volatility. In recent years, with the development of machine learning and signal processing methods, an increasing number of time series decomposition methods have been proposed to deal with these issues. Based on various theories and assumptions, decomposition methods can significantly influence the accuracy of bagged forecasts, but the decomposition method's impact on bagging has been overlooked in the literature. Previous studies only performed bagging based on the seasonal trend decomposition procedure based on loess (STL) or the empirical mode decomposition (EMD) method. This study introduces bagging based on several time series decomposition methods (such as X11, STL, Prophet and structural time series model) for seasonal time series forecasting. We will use quarterly data of inbound tourism in the UK from key source markets to evaluate the performance of different decomposition methods applied to bagging in tourism forecasting. This study will contribute to the further applications of bagging in complex contexts of time series forecasting.

A hybrid method of exponential smoothing and neural networks for hotel demand forecasting

Presenter: Apostolos Ampountolas
Co-authors: Apostolos Ampountolas;

Over the past several years, the rapid development of information technology has been instrumental in the growth of demand in the hospitality sector. In the hotel industry, demand forecasting is challenging due to numerous anomalous days related to many annual holidays, a variety of events, promotions, and environmental features. This leads to solid irregularity and seasonal patterns; thus, the overnight demand exhibits variant arrival behavior that may contain outliers. Recent literature has looked at various techniques to forecast time series using linear, non-linear, and hybrid models to examine hotel occupancy accuracy. For example, hybrid models by appropriately combining linear models like the autoregressive integrated moving average (ARIMA) with other non-linear models such as an artificial neural network (ANN) have displayed highly accurate results than individual models, see the M4-competition. Therefore, the paper examines a hybrid methodology of obtaining the various combinations of linear and/or non-linear patterns in an overnight demand time series dataset. We evaluate our model's performance with alternative prediction approaches to forecasting the daily demand, including ARIMA, ETS, multilayer perceptron (MLP), in addition to a hybrid ARIMA"“ANN model. The goal is to measure how successfully combining exponential smoothing models (ETS) with ANN performs comparably to simple alternatives. Therefore, an analysis of specific standard metrics has been applied to measure forecast accuracy.

Forecasting Tourist Arrivals in Europe: A Causal Spatiotemporal Econometric Model

Presenter: Xiaoying Jiao
Co-authors: Xiaoying Jiao; Jason Chen; Gang Li

With the continuous growth of tourism industry over the past few decades, tourism demand studies have attracted increasing attention from academics and researchers. Because of the perishable nature of tourism, accurate tourism demand is considered crucial, and advanced tourism demand forecasting methods are continuously developed to improve forecasting accuracies. Among the advanced methods, spatial econometric methods have become a promising trend in tourism demand forecasting, as well as the general forecasting field because of the ability to capture the spatial spillover effects among destinations. However, applications of spatial methods are still rare in tourism demand forecasting, with three exceptions only (Jiao

et al., 2020; Long et al., 2019; Yang & Zhang, 2019). Inspired by the previous literature which confirms the existence of spatial dependence among destinations, this study further develops a general nesting spatiotemporal (GNST) model in an effort to improve the accuracy of tourism demand forecasts. The proposed GNST model accounts for the spatial and temporal effects of endogenous and exogenous variables as well as unobserved factors. Based on a panel dataset containing quarterly inbound visitor arrivals to 26 European destinations, this empirical study demonstrates that the GNST model outperforms both its non-spatial counterparts and spatiotemporal benchmark models. This finding confirms that spatial and temporal exogenous interaction effects contribute to enhanced forecasting performance. Keywords: Tourism demand forecasting, Spatiotemporal model, SAC model, GNST model, Panel data“fKey referencesJiao, X., Li, G., & Chen, J. L. (2020). Forecasting international tourism demand: A local spatiotemporal model. Annals of Tourism Research, 83, 102937. https://doi.org/10.1016/j.annals.2020.102937Long, W., Liu, C., & Song, H. (2019). Pooling in Tourism Demand Forecasting. Journal of Travel Research, 58(7), 1161”“1174. https://doi.org/10.1177/0047287518800390Yang, Y., & Zhang, H. (2019). Spatial-temporal forecasting of tourism demand. Annals of Tourism Research, 75, 106”“119. https://doi.org/10.1016/j.annals.2018.12.024/subsection%7BImproving Daily, Weekly, Monthly, and Quarterly Hotel Room Demand Forecasts for Vienna across Hotel Classes: Evidence from Single Models and from Combination Techniques}Presenter: Ulrich Gunter
Co-authors: Ulrich Gunter;

The present study employs daily data made available by the STR SHARE Center over the period January 1, 2010 to January 31, 2020 for six Viennese hotel classes and their total. The forecast variable of interest is hotel room demand. As forecast models, (1) Seasonal NaÃ¯ve, (2) ETS, (3) SARIMA, (4) TBATS, (5), Seasonal NNAR, and (6) Seasonal NNAR with an external regressor (annual seasonal naÃ¯ve forecast of the inflation-adjusted ADR) are employed. Forecast evaluation is carried out for forecast horizons h = 1, 7, 30, and 90 days ahead based on rolling windows. As forecast combination techniques, (a) mean, (b) median, (c) regression-based weights, (d) Bates-Granger weights, and (e) Bates-Granger ranks are calculated. In the relative majority of cases (i.e., in 13 of 28), combined forecasts based on Bates-Granger weights and on Bates-Granger ranks provide the highest level of forecast accuracy in terms of typical forecast accuracy measures.

Measuring uncertainty from a large set of models' predictions

Presenter: David Ardia
Co-authors: David Ardia; Arnaud Dufays

We propose a novel uncertainty measure exploiting the heterogeneity of a large set of models' predictions. The approach is forward-looking, can be computed in real-time, and can be applied at any frequency. We illustrate the methodology with expected shortfall predictions of worldwide equity indices generated from 71 risk models. We demonstrate the usefulness of the new measure for risk forecasting and asset pricing.

Duration-dependent volatility models with value-weighted approach

Presenter: Fernando Mendes
Co-authors: Fernando Henrique Mendes; Douglas Eduardo Turatti; Jose Henrique Mazzeu; Jose Caldeira

This paper attempts to revisit the duration-dependent Markov-switching model of Maheuand McCurdy (2000b) through a model combination approach. The existing literature didnot fully explore the issues of using a single duration value to characterize the conditionalvariance of returns. Hence, the potential of the model to predict conditional volatilityhas not been well investigated from an out-of-sample forecasting context. We highlightthe foundations of adopting different duration values to conduct a one-step-ahead weighing volatility forecast. To evaluate our empirical study's contribution, we conducted astatistical and

risk prediction evaluation for the daily bitcoin returns. In general, ourresults outperform GARCH-type models using different volatility proxies and robust lossfunctions, both for point volatility forecasting and Value-at-Risk analysis.

Leveraging Latent Spaces for Cryptocurrency Return Forecasting with Deep Networks

Presenter: Matthew Harrington
Co-authors: Matthew Harrington; Ali Habibnia

Cryptocurrencies are a newly popular asset class among both institutional and retail traders. In contrast with established markets like equities, cryptocurrencies experience lower trading volume and exhibit a high degree of comovement. These factors, combined with the market's distributed nature, make the asset class promising for quantitative traders. Herein, we leverage the comovement between the top 30 cryptocurrencies by market share to develop a trading strategy reliant on Deep Learning networks. Specifically, we compare the predictive capacity of popular dimension reduction techniques (Linear, Non-Linear PCA, Auto-Encoders). These latent representations are then fed as input to predictive models (Multi-Layer Perceptron, LSTM). Our deep models are trained on price data from Binance.us, a popular cryptocurrency exchange. Through portfolio simulation on out-of-sample data, our approach is shown to significantly outperform basic strategies like " buy-and-hold" and latent factor models.

Prediction Intervals: Neglected Diagnostics?

Presenter: Keith Ord
Co-authors: Keith Ord;

Our inferences are made conditionally on the assumption that the model is a good enough approximation to the data generating process. Checks of this assumption are usually based upon a combination of goodness-of-fit of the chosen model and the performance of out-of-sample point forecasts. Although it is well-known that prediction intervals are sensitive to the underlying assumptions, this sensitivity is rarely used to assess model adequacy. The classic airline series, which gave rise to the Box-Jenkins airline model, is used to illustrate how prediction intervals can improve the model selection process.

Estimating Interval Forecasts using Pruned Ensembles

Presenter: Erick Meira
Co-authors: Erick Meira; Fernando Cyrino; Lilian De Menezes

Drawing on knowledge from Statistics, Machine Learning and Forecasting, the present study introduces an ensemble-based approach to generate prediction intervals of electricity supply. It combines Bootstrap Aggregation (Bagging) algorithms, time series methods and a novel pruning routine capable of feature selection before aggregation. A comparative, out-of-sample analysis, is conducted using monthly data from 20 countries, and covers multiple time horizons (including the COVID-19 pandemic). Forecasting performance is compared with a range of methods, including traditional benchmarks and state-of-the-art ensemble forecasting approaches. The results demonstrate the capacity of the proposed approach in delivering accurate prediction intervals under different settings and even during periods of severe economic distress. Moreover, this is a flexible methodology, which can be used to forecast a wide range of univariate time series.

Orbit: Probabilistic Forecast with Exponential Smoothing

Presenter: Edwin Ng
Co-authors: Edwin Ng; Zhishi Wang

Time series forecasting is an active research topic in academia as well as industry. This paper introduces a package Orbit, where it implements a series of Bayesian time series exponential smoothing models with the help of probabilistic programming languages such as Stan and Pyro. Our model refinements include additional global trend, transformation for multiplicative form, noise distribution and flexible choice of priors. More importantly, Orbit provides a unified interface for time series inference and forecasting tasks, which can be readily extended to new model structures. A benchmark study is conducted on a rich set of time-series data sets for our proposed models along with other well-known time series models.

Forecasting in Big Data Environments: an Adaptable and Automated Shrinkage Estimation of Neural Networks (AAShNet)

Presenter: Ali Habibnia
Co-authors: Ali Habibnia; Esfandiar Maasoumi

This paper considers improved forecasting in possibly nonlinear dynamic settings, with high-dimension predictors ("big data" environments). To overcome the curse of dimensionality and manage data and model complexity, we examine shrinkage estimation of a back-propagation algorithm of a deep neural net with skip-layer connections. We expressly include both linear and nonlinear components. This is a high-dimensional learning approach including sparsity L1 penalties, allowing high-dimensionality and nonlinearity to be accommodated in one step. This approach selects significant predictors as well as the topology of the neural network. We estimate optimal values of shrinkage hyperparameters by incorporating a gradient-based optimization technique resulting in robust predictions with improved reproducibility. The latter has been an issue in some approaches. This is statistically interpretable and unravels some network structure, commonly left to a black box. An additional advantage is that the nonlinear part tends to get pruned if the underlying process is linear. In an application to forecasting equity returns, the proposed approach captures nonlinear dynamics between equities to enhance forecast performance. It offers an appreciable improvement over current univariate and multivariate models by RMSE and actual portfolio performance.

General NN Forecaster

Presenter: Slawek Smyl
Co-authors: Slawek Smyl;

The talk will describe a project to develop a NN forecasting system that offers competitive accuracy and superior forecasting speed, a useful characteristics for many scenarios, like anomaly detection.It follows a recent realization in forecasting community that powerful ML models trained on large collection of series can generalize very well.While training is a time consuming process, the trained model (partially in binary format) is able to forecast never-seen-before series very quickly, because during serving no optimization takes place, just a number of highly optimized matrix operations.

Time Series Forecasting with Time Series Plot and Computer Vision

Presenter: Taiyeong Lee
Co-authors: Taiyeong Lee; Michael Leonard

Time series forecasting is to predict future values with historical time series data. If the data is not available, but the plot image is the only available source, a human being can do a forecast without the underlined data source. For example, if a sine wave plot is given, you can guess the plot shape after the end time point, which is another sine wave. If you see the stock market chart, you may forecast the stock price in the next time points, whether your forecast is right or not. You would guess its upward or downward direction in the price move. Now the question is that computer vision can do a similar forecast with a time series plot as you do, which means whether computer vision can conjecture the future behavior of the given image of series. In this paper, we show how to use the image pixels to predict the future movement or shape of the given time series plot.

A Multiple Investigation with Time Series, Structural Models and Model Selection

Presenter: Camila Vasconcelos
Co-authors: Camila Vasconcelos; Eli Hadad Junior

The objective of this applied work is to investigate the predictive power of exchange rate projection models in the context of time series analysis. The relevance of this theme is associated with the possibility of arbitrage for investors who hold currency assets in their portfolios. For the study, 7 exchange rates were selected against the US dollar, referring to the countries: Japan, United Kingdom, Australia, Chile, Argentina, Mexico and Brazil. The choice was based on the currency function criterion, as defined by BIS (Bank of International Settlements). The study horizon will cover the period from 1995 to 2021, with data on a quarterly basis sourced in IMF and OECD databases. For the euro currency, high frequency data will be investigated, with a sampling interval of one minute, making possible to discuss arbitrariness in a speculative context. The contribution of this research is made in three aspects: i) the 26-year horizon of the investigation includes the entire period available after euro introduction and post-adoption of the floating exchange rate regime, in the case of Brazil; ii) structural models will be adopted with the application of VAR, VEC, combination of projections and use of filters for GDP econometric techniques, and iii) high frequency data will be analyzed using ARFIMA and neural network models. The innovation of this study is due to the inclusion of a country-risk proxy variable in the adopted empirical models. Additionally, this research assesses the predictive power of projection models by multiple criteria (loss function and direction of change), as well as addressing the uncertainty associated with projections based on the density function. The approach and techniques covered by the research will allow the comparability of the empirical results obtained with those pointed out by both classic and the most recent studies in the literature.

Combination of theoretical models for exchange rate forecasting

Presenter: Maria Paula Bonel
Co-authors: María Paula Bonel;

The main objective of this paper is to study whether there are exchange rate forecasting gains, with respect to a reference model, when combining fitted data from different empirical models based on economic theory. Data combination is performed using Bayesian Model Averaging (BMA) models. Working with both developed and emerging countries is a contribution to the recent literature, which usually focuses on developed countries. BMA models tend to provide more accurate forecasts than random walk. This difference is statistically significant for an important number of country and horizon combinations (3, 6 and 12 months ahead). Using pooled data by group of countries generates accuracy gains in about 60% of the cases, with respect to forecasts that use individual information. This percentage is maintained throughout the different forecast horizons. By country group, the gain is largest in panel A of developed countries. On the other hand, the gains in accuracy decrease as the forecast horizon is extended. BMA models for developed countries tend to be more “sparse” than emerging countries models. Also, BMA models, for

both country groups, become more dense as the forecast horizon is extended. Forecasting exercises on the direction of change also yield positive results.

Forecasting Fuel Prices with the Chilean Exchange Rate

Presenter: Pablo Pincheira
Co-authors: Pablo Pincheira; Nabil Jarsun

In this paper we show that the Chilean exchange rate has the ability to predict the returns of oil prices and of three additional oil-related products: gasoline, propane and heating oil. The theoretical underpinnings of our empirical findings rely on the present-value theory for exchange rate determination and on the strong co-movement displayed by some commodity prices. The Chilean economy is heavily influenced by one particular commodity: copper, which represents nearly 50% of total national exports and attracts a similar share in terms of Foreign Direct Investment. As a consequence, the floating Chilean exchange rate is importantly affected by fluctuations in the copper price. As oil-related products display an important co-movement with base metal prices, it is reasonable to expect evidence of Granger causality from the Chilean peso to these oil-related products. We find substantial evidence of predictability both in-sample and out-of-sample. Our paper is part of a growing literature that in the recent years has explored the linkages between commodity prices and commodity currencies.

Multivariate Business Forecasting with Causal Validity

Presenter: Sameer Manek
Co-authors: Sameer Manek; Alex Chin; Heng Kuang; DJ Rich; Sean Taylor

Forecasting under various scenarios plays a key role in Lyft's decision making and planning processes. In order to efficiently: design marketplace policies, allocate marketing spend, and plan for exogenous events, we need to construct forecasts incorporating the impacts of our decisions across core business metrics. In this paper, we describe the implementation of our forecasting platform as a deep multi-task neural network with observable intermediate concepts; we create a network of sub-models'themselves (deep) neural networks'to capture the causal relationships between metrics and compose them into a unified model of the business.The proposed approach has advantages over alternative multivariate time series forecasts: 1) we can forecast counterfactual scenarios by adjusting policy variables, 2) subject matter experts within Lyft are able to develop and improve sub-models in isolation'without any knowledge of any of the other models, 3) we can dynamically generate compositions of the sub-models to focus on the forecasting problem at hand, and 4) the differentiable architecture allows straightforward optimization of plans.Our architecture is an extension of a deeply-supervised neural network; the network encodes the directed acyclic graph of our business and some "hidden" layers are supervised directly. We've implemented our approach within PyTorch, and the model can be trained either in components or jointly using a global loss function..The accuracy of this approach is on par with our prior forecasting system based on time series methods, while being able to explicitly incorporate plans, decisions, and exogenous factors. This approach also reduces the time required to develop scenarios from days to hours, allowing us to plan at a higher cadence. This approach is highly flexible and well suited for other multivariate forecasting needs when they follow a similar causal structure.

Forecasting models in production systems

Presenter: Reza Hosseini
Co-authors: Reza Hosseini; Albert Chen; Kaixu Yang; Sayan Patra; Parvez Ahammad

In this presentation, we discuss the desirable properties of forecasting models in production systems: ability to accommodate complex patterns; being easily customizable; can support flexible objectives; can

produce robust forecasts; and can run fast (scalable). Then we present an algorithm we have developed at Linkedin to satisfy these properties. Our algorithm (Silverkite) which now supports many use cases company-wide is a two component model where the mean and volatility components are modeled separately to achieve high speed and robustness. Silverkite can easily customize very complex patterns, it supports flexible objective functions and has automatic changepoint detection and flexible holiday components built in.

Forecasting Daily Business Metrics On Short-Term Horizons

Presenter: Jerry Shan
Co-authors: Jerry Shan;

In this paper, we present a new modeling method for forecasting daily business metrics. The solution first uses daily data in multiple historical weeks in a training month to quantify the DoW (Day of Week) effects, and then uses weekly aggregates over a month to quantify weekly trend effects. In the process, we introduced a new concept, called "week to month association" , or "associative week to a month" , which uniquely determines the membership for a week to be associated with a month. Based on the quantified DoW patterns and the weekly patterns within a month or across two months, we build a non-parametric model for the DoW daily percentages in a week and a parametric model and non-parametric model for the weekly aggregates, and use the models coupled with any incremental daily data to forecast the metrics for the subsequent days and weeks. Finally, we use the forecasted values and the actual values as they arrive to quickly detect anomaly at varying levels of severity to serve for early warning detection purpose.

Electricity Demand Threshold Forecasting for Triggering Cost Saving Peak Demand Shaving Actions

Presenter: Omar Aponte
Co-authors: Omar Aponte; Katie McConky

Utilities around the world have increased their infrastructure investments to ensure the reliability and safety of their electric grids as new smart grid and renewable energy technologies become available. The financial burden of these investments has been passed on to the consumers in the form of dynamic pricing schemes and demand charges. These schemes and additional charges are designed to influence the energy consumption behavior of the consumers in an attempt to prevent grid failures. Researchers have placed significant emphasis on methodologies that focus on consumers reacting to time-based rate differentiation and price signals coming from the utility. However, little attention has been paid to methodologies that would allow consumers under demand charges to proactively determine an appropriate demand threshold value in order to trigger demand peak shaving actions. Peak shaving actions are designed to minimize demand charges, which can amount to up to 70% of an electric bill. On the other hand, these same actions often generate undesired inconvenience to users, such as sub optimal thermostat settings. A recently published paper detailed how both the demand charge and the need to perform demand peak shaving actions, were minimized by performing these actions only when demand reached a forecasted threshold. However, the accuracy of this demand threshold forecasting methodology has not been contrasted with other machine learning based alternatives in order to ensure that the consumers are indeed getting the appropriate threshold value to achieve the best possible results. Our research provides a comparison between the performance of the published methodology and three additional machine learning based methodologies at forecasting the demand threshold using real consumer electricity data. The comparison was made for two common scenarios: a consumer without self-generation of electricity and a consumer with behind the meter solar generation. This research contributes a superior methodology to proactively establish a demand threshold that allows electricity consumers to trigger demand response actions that minimize both the financial impact of demand charges, and the user inconveniences generated by performing such actions.

Long-term probabilistic forecasting for intra-hour reserve requirements

Presenter: Allison Campbell
Co-authors: Allison Campbell; Nader Samaan; Malini Ghosal; Marcelo Elizondo; Quan Nguyen; Tony Nguyen

This talk introduces the ten year ahead probabilistic reserve requirement calculations completed for the Western Electricity Coordination Council (WECC) by the Pacific Northwest National Laboratory using the 2030 Anchor Data Set (ADS). Long-term planning for ancillary service reserve requirements, such as load following and regulation, hinges on intra-hour variability. Increased targets for the penetration of solar and wind resources require extra consideration to accurately model the minute-to-minute uncertainty introduced by load, wind, and solar between the day ahead, hour ahead, and real time markets to account for these imbalances. Observed historical variability in load, wind, and solar forecasts are modeled for each Balancing Authority in the WECC and incorporated into the 2030 ADS in order to preserve actual imbalance characteristics. Discussion includes the methodology for ten year ahead forecasting of load following and regulation reserve requirements and the role of accurately recreating the observed uncertainty distribution for each resource. Results of the WECC 2030 ADS forecast for reserve requirements will be presented for BAs with anticipated high and low penetration of renewable resources, along with disaggregation by resource type. The talk will motivate the role of probabilistic forecasting to enable the adoption of variable resources with existing dispatchable resources.

Machine Learning Application in Identifying Representative Hours in Capacity Expansion Planning Model

Presenter: Zohreh Parvini
Co-authors: Zohreh Parvini; Laleh Behjat; Bill Rosehart

Capacity expansion planning (CEP) of power systems is a long-term optimization problem with embedded short-term operation constraints. A successful planning models all system conditions, including different generation levels of renewable generations and load. Given the long-term horizon of CEP models, large and interconnected power systems, and ever-increasing injection of renewable resources, it is not possible to consider every load and generation condition in the models. Conventionally, operation constraints are accounted for by considering a limited number of representative operating periods (hours) for each planning year, balancing between fidelity of the models and computational tractability. These "representative hours" are identified by dividing the load into blocks and choosing one load condition for each block. However, the intermittent and uncertain nature of renewable resources has introduced new challenges to these traditional approaches.While considering every operating period in the whole planning horizon is unrealistic, using a subset of operating periods by only considering load characteristics also fails to model dynamics of renewable energy resources accurately. Therefore, it is vital to obtain an appropriate temporal representation of the operating constraints to preserve its the chronological characteristics. In this research, we aim to propose new time-series clustering techniques to find the most suitable representation of operating periods for the planning problem, capturing key statistical characteristics of parameters considering load and renewable resources. We simulate CEP models using the representative hours obtained by these methods to find the optimal expansion plan. To compare the clustering techniques, we run production cost modeling simulations to identify the number of infeasibilities, overloads, or curtailments. The comparison enables us to identify the most efficient algorithm for selection of representative operation periods.

Tools for forecast reconciliation: the R package FoReco

Presenter: Daniele Girolimetto
Co-authors: Daniele Girolimetto;

FoReco is an R package designed for point forecast reconciliation of a multiple linearly constrained (e.g. hierarchical/grouped) time series. The present release (0.1.1, https://CRAN.R-project.org/package=FoReco) deals with cross-sectional (Hyndman et al., 2011), temporal (Athanasopoulos et al., 2017), and cross-temporal (Kourentzes and Athanasopoulos, 2019, Di Fonzo and Girolimetto, 2020) forecast reconciliation procedures. Projection and structural approaches have been considered to better exploit the linear relationships linking the data. Almost all of the state-of-the-art reconciliation procedures were considered, and a powerful tool to guarantee non-negativity of the reconciled forecasts was made available.In the new FoReco 0.2.0, some significant updates have been considered: first, building upon and extending a recent proposal by Hollyman et al. (2021), a new forecast combination based forecast reconciliation procedure is considered, with either exogenous or endogenous intermediate level constraints. Level Conditional Coherent (LCC) forecast reconciliation for elementary hierarchies, and the Combined Conditional Coherent forecast reconciliation approach are now available in the new command lccrec(). In addition, the new release, (i) besides the non-negative option, now permits to impose linear inequalities bounds on all the reconciled forecasts, which may be very useful in many practical situations, (ii) offers a wider freedom in choosing the covariance matrices (also different along the forecast horizon), and (iii) lets the user define a customized subset of the temporal aggregation orders to be used in the reconciliation.ReferencesAthanasopoulos, G., Hyndman, R.J., Kourentzes, N., Petropoulos, F. (2017), Forecasting with Temporal Hierarchies, European Journal of Operational Research, 262, 1, 60-74.Di Fonzo, T., Girolimetto, D. (2020), Cross-Temporal Forecast Reconciliation: Optimal Combination Method and Heuristic Alternatives, https://arxiv.org/abs/2006.08570.Hollyman, R., Petropoulos, F., Tipping, M.E. (2021), Understanding Forecast Reconciliation, European Journal of Operational Research (in press).Kourentzes, N., Athanasopoulos, G. (2019), Cross-temporal coherent forecasts for Australian tourism, Annals of Tourism Research, 75, 393-409.Hyndman, R.J., Ahmed, R.A., Athanasopoulos, G., Shang, H.L. (2011), Optimal combination forecasts for hierarchical time series, Computational Statistics & Data Analysis, 55, 9, 2579-2589

Understanding forecast reconciliation: further insights and extensions

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

In a recent paper, while elucidating the links between forecast combination and cross-sectional forecast reconciliation, Hollyman et al. (2021) have proposed a forecast combination-based approach to the reconciliation of a simple hierarchy. A new Level Conditional Coherent (LCC) point forecast reconciliation procedure was developed, and it was shown that the simple average of a set of LCC reconciled forecasts (called Combined Conditional Coherent, CCC) results in good performance as compared to those obtained through the state-of-the-art cross-sectional reconciliation procedures. We build upon and extend these results along some new directions. We shed light on the nature and the mathematical derivation of the LCC reconciliation formula, showing that this procedure can be viewed as the result of an exogenously linearly constrained minimization of a quadratic loss function with a diagonal associated matrix in the differences between the target and the base forecasts. We consider endogenous constraints as well, which results in level conditional reconciled forecasts of all the involved series, where both the upper and the bottom time series are coherently revised. We show that even in this framework it is still valid the interesting interpretation given by Hollyman et al. (2021) of the reconciliation formula as the combination of direct (base) and indirect forecasts, the latter ones depending on the accounting relationships linking upper and bottom series. The extension to the cases where a full metric matrix is considered in the definition of the loss function is straightforward, which might be useful when (if) suitable error forecast covariance matrices may be estimated. In addition, since the LCC procedure (i.e., with exogenous constraints, but the result holds in the endogenous case as well) does not guarantee the non-negativity of the reconciled forecasts, we argue that its interpretation as an 'unbiased top-down reconciliation procedure' leaves room for some doubts. Finally, in order to exploit the inherent relationships valid for a hierarchy/group of time series to achieve better forecast accuracy, we extend the LCC (and CCC as well) forecast reconciliation procedures in both temporal and cross-temporal frameworks.ReferencesHollyman, R., Petropoulos, F., Tipping, M.E. (2021), Understanding Forecast Reconciliation, European Journal of Operational Research (in press).

Forecast combinations, pooling, and hierarchies: how do they ”combine”?

Presenter: Nikos Kourentzes
Co-authors: Nikolaos Kourentzes;

Hierarchical forecasting has received increasing interest over the last years, motivated by a need to address cross-sectional hierarchical forecasting questions (e.g. forecasts of sales of different products, aggregating to product groups, and so on). Nonetheless, the same techniques can be used in more types of forecasting problems, particularly with the use of temporal hierarchies that can be applied to single time series. Most current hierarchical forecasting methods rely on different forms of forecast combinations, with additional restrictions to satisfy the hierarchical structure of the problem. In this work we explore the connection between hierarchical forecasting and forecast combinations further. In particular, we consider the recent advances in forecast pooling, where the modeler attempts to pre-select the set of forecasts to be combined, prior to constructing combination weights, with the objective of reducing the weight estimation errors, and by extension improving the accuracy of the combined forecasts. With the combination of both, we propose sparse hierarchies, attempting to improve both the accuracy, but also the flexibility of hierarchical methods. Complimentary, we explore the benefits of hierarchical thinking for forecast combinations more generally. The various approaches are compared empirically against established hierarchical and forecast combination benchmarks.

Forecasting the U.S. Unemployment Rate: Another Look

Presenter: John Guerard
Co-authors: John Guerard; Rong Chen; Han Xiao; Dimitrios Thomakos

Abstract. Twenty years ago, an eminent group of economic forecasters and statisticians published a seminal work of forecasting the U.S. unemployment rate, certainly one of the most important economic measures of the U.S. economy. Montgomery, Zarnowitz, Tiao, and Tsay (MZTT, 1998) reported that linear and nonlinear time series models are useful in predicting unemployment rates relatively accurately, in both short term (1 month) and medium term (5 months) prediction. One interesting and important ï¬ nding was that the weekly unemployment claims was a statistically signiï¬ cant input in forecasting the U.S. unemployment rate over the 1959 ”“ 1993 time period. The weekly unemployment claims time series is a component of the U.S. Leading Economic Indicators (LEI). In this paper we replicated and extended the MZTT analysis for the 1959 to 2019 time period. We report out-of-sample one-step to twelve-step ahead monthly prediction performance of various models for the 1990-2019 period, using a no-change (random walk) model as a forecasting benchmark. Results obtained from this study include: (1) weekly unemployment claims are indeed a useful and statistically signiï¬ cant input in a transfer function model to forecast the unemployment rate; (2) the leading economic indicators time series is a statistically signiï¬ cant input in a trans-fer function model to forecast the unemployment rate; (3) a seasonal ARIMA (SARIMA) model outperforms the no-change benchmark for all forecasting horizons; (4) the SARIMA and transfer function models are statistically signiï¬ cantly better forecasting models than a null, or no-change, forecast, particularly in the Global Financial Crisis (GFC), 2008 -2019 time period.Keywords: Time series forecasting; transfer function modeling; the unemployment rate

Modelling higher moments and density forecasting: a comprehensive look

Presenter: Xiaochun Liu
Co-authors: Xiaochun Liu; Richard Luger

Many GARCH-type models have been proposed in the literature for the higher moments of financial returns and their conditional distributions. In this paper, we examine comprehensively whether these models yield better out-of-sample density forecasts. Among a wide range of specifications for autoregressive

conditional volatility, skewness, and kurtosis, we find that the most promising approach rests on a decomposition of returns into their signs and absolute values. This approach specifies the joint distribution of the return components by combining a dynamic binary choice model for the signs, a multiplicative error model for the absolute values, and a dynamic copula function for their interaction. This flexible specification captures well the time-varying conditional skewness process and provides more accurate density forecasts than competing models, especially for the left tail of financial returns.

Optimal Prediction Under Multivariate Asymmetric Loss; Comparison of Multivariate Garch vs. Multivariate Realized Garch Models

Presenter: Yasemin Ulu
Co-authors: Yasemin Ulu;

We compare the performance of the optimal predictor for daily returns using Multivariate BEKK-GARCH model with normal innovations to that of Realized Multivariate GARCH model developed by Hansen et. al (2014), under the assumption that agents have asymmetric multivariate loss function. Our results indicate that under multivariate realized GARCH specification a further 5-7% more loss reduction is attainable for moderate degrees of asymmetry and can reach to about 12% as the degree of asymmetry increases.

Inflation Expectations and Uncertainty from the Perspective of Firms

Presenter: Xuguang Simon Sheng
Co-authors: Xuguang Simon Sheng; Brent Meyer; Nicholas Parker

We draw on the Atlanta Fed's Business Inflation Expectations Survey to draw inference about firm's inflation perceptions, expectations, and uncertainty. Using methods grounded in the survey literature, we find evidence that the concept of ''aggregate inflation" as measured through price statistics like the Consumer Price Index (CPI) hold very little relevance for business decision makers. This lack of relevance manifests itself through experiments (including randomized controlled trials) that show varying question wording researchers use to elicit inflation expectations and perceptions significantly changes firm's responses. Our results suggest firms have become rationally ignorant of a low inflation environment. Instead, we find that unit costs are the relevant lens with which to capture firms' views on the nominal side of the economy. We then investigate both firm-level (micro) and aggregated (macro) probabilistic unit cost expectations. On a firm-level, unit costs are an important determinant of firms' price-setting behavior. Aggregating across firms' beliefs, firms' unit cost perceptions strongly co-move with official aggregate price statistics and, importantly, firms' inflation expectations bear little in common with the ''prices in general" expectations of households. Rather, firms' aggregated beliefs strongly covary with the inflation expectations of professional forecasters and market participants.

How far ahead can we forecast US Data?

Presenter: Constantin Burgi
Co-authors: Constantin Burgi;

Several novel tests are applied to GDP and inflation rate forecasts from the SPF, and the Greenbook in order to assess, how far into the future professional forecasters can add value relative to a naive benchmark and their previous prediction. It is shown that the predictive ability is rather limited beyond three quarters ahead for GDP and CPI inflation and two quarters ahead for the GDP deflator. In a regression where longer horizon forecasts are regressed on shorter horizon forecasts, there is a limited difference between the two forecast sources, while the Greenbook outperforms the SPF when the forecasts are regressed on the actual, controlling for the prediction made for the previous period.

All Forecasters Are Not the Same: Time-Varying Predictive Ability across Forecast Environments

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

This paper examines data from the European Central Bank's Survey of Professional Forecasters to investigate whether participants display equal predictive performance. We use panel data models to evaluate point- and density-based forecasts of real GDP growth, inflation, and unemployment. The results document systematic differences in participants' forecast accuracy that are not time invariant, but instead vary with the difficulty of the forecasting environment. Specifically, we find that some participants display higher relative accuracy in tranquil environments, while others display higher relative accuracy in volatile environments. We also find that predictive performance is positively correlated across target variables and horizons, with density forecasts generating stronger correlation patterns. Taken together, the results support the development of expectations models featuring persistent heterogeneity.

Feature selection via a modern optimisation lens: Helping insurers ask the right questions

Presenter: Jessica Leung
Co-authors: Wai Yin Jessica Leung; Dmytro Matsypura

Insurtech, a confluence of insurance and technology, is transforming the insurance industry. Insurance policies are now mostly purchased via the internet. Along the way, immense amounts of individual-level consumer data are continuously generated and stored. Unfortunately, merely collecting vast amounts of data is insufficient. To translate a data-intensive environment into data-enabled competitive advantages, one must also carefully select the relevant attributes on which to focus. An appropriate subset of features informs not only which potential customer is of high risk but also the premium pricing decisions and guides the formulation of loss reserving, marketing, and customer management strategies. Enhanced understand- ing of the relevant features also simplifies the questionnaire design and improves the consumer experience during a machine-oriented information extraction process. Naturally, insurance data contains both continuous and categorical features and thus a feature selection procedure accounting for the grouped information is required. To identifying profitable potential customers, we introduce a novel classifier that selects the relevant features among the vast amount of Insurtech data available. Since insurance data contains both continuous and categorical variables, we formu- late the problem as a mixed-integer optimization problem and incorporate grouping information of categorical variables in the classifier. The enhanced understanding of the number of relevant attributes simplifies the questionnaire design and improves the consumer experience during a machine-oriented information process.

Estimation of Tail Risk Measures for Heteroskedastic Financial Time Series: A Extreme Value Approach With Covariates

Presenter: Robert James
Co-authors: Robert James; Artem Prokhorov; Wai Yin Jessica Leung

In this study, we explore how exogenous economic and financial information can be incorporated into the two-stage GARCH-EVT risk model of [**?**], and whether or not this additional information improves out-of-sample Value-at-Risk and expected shortfall forecasts. We hypothesize that the optimal model is sparse and that the statistical importance of covariates is time-varying. That is, the econometrician faces a variable selection problem. Accordingly, we incorporate Lasso regularization into the Maximum Likelihood objective function of the Generalized Pareto distribution in the second ''EVT" stage of the risk model. Using an extensive set of performance criteria we demonstrate that the Lasso regularized GARCH-EVT

risk model dramatically improves financial risk forecasting during periods of elevated volatility and financial stress when accurate forecasts are needed the most. We provide a detailed insight into how economic and financial covariates are related to financial risk.

Mixed Integer Optimization for time series change points detection

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

Identifying structural breaks in time series data is a major area of interest within econometrics and applied statistics, going back at least to the 1950s. Traditionally, researchers relied on statistical approaches e.g. Bai and Perron (1998). Recently, a growing literature is showing attractive properties of Mixed Integer Optimization (MIO) methods, as means of obtaining efficient solutions in a wide range of statistical problems. Specifically, it has long been thought that MIO methods are not suitable for machine learning problems. However the recent drastic improvements in the performance of MIO methods have lead many to reconsider this premise. In this paper we develop a MIO method for the identification and estimation of structural breaks in time series data. We design a formulation that transforms the classical structural break detection problem into a Mixed Integer Quadratic Programming problem. Then, MIO is capable of finding provably optimal solutions to the problem. Our framework allows to determine the unknown number of structural breaks. In addition to that, we demonstrate how to accommodate the required number of structural breaks, or a minimal required number of breaks. To demonstrate the effectiveness of our approach, we conduct extensive numerical experiments on synthetic and real-world data. The proposed MIO representation is solved using a well-known optimization solver. We examine optimal and sub-optimal solutions of the problem, and the effect of tuning the parameters. We show how to choose the tuning parameters and compare our results with established methods such as those proposed by Bai and Perron (1998).

BQML Time Series Forecasting and Anomaly Detection

Presenter: Xi Cheng
Co-authors: Xi Cheng;

In this talk, we introduce BQML time series, a scalable and serverless time series platform offered from BigQuery of Google Cloud. As part of the BQML platform, it lets the user accurately forecast hundreds of thousands of time series in parallel via ~10 lines of SQL code, without the need to worry anything about resource configuration, data pipeline, feature engineering, and hyper-parameter tuning. Modeling-wise, it automatically handles spikes and dips, level changes, multiple seasonalities, holiday effects and so on. It also performs the auto.ARIMA algorithm to find the best ARIMA model for the trend component. Platform-wise, it parallels the time series computations using the ambient resources in the Cloud, which however is subject to the amount that the user reserves. The forecasting results are self-explainable by the exposed time series decomposition results. Additionally, it lets the user perform anomaly detection on the time series.

ML Forecasts on Small Datasets Using Transfer Learning and Data Augmentation

Presenter: Haoyun Wu
Co-authors: Haoyun Wu;

In computer vision and natural language processing, transfer learning has been extensively studied to save training time and solve limited training data issues. In time series forecasting, we also face the

problem of not enough historical data to train or select the best model. Our model STARRY-N as a machine learning framework of globally learning statistical models has the potential ability of transfer learning, which could solve the problem of limited data. We explore the transfer learning ability of STARRY-N on multiple public datasets including M4, M3, electricity and etc. We also study data augmentation methods for further performance improvement.

Robust Ensembles of Quantile Forecasts

Presenter: Kashif Yousuf
Co-authors: Kashif Yousuf;

In this talk we discuss ways to produce robust quantile forecasts by ensembling many component quantile forecasts. The resulting ensemble should be robust in the sense that it can easily be applied to a new dataset with minimal tuning thus reducing the need for expert involvement. We discuss various quantile forecasting methodologies primarily based on empirical prediction errors, as well as different ensembling methodologies. Our end result is to propose a set of component quantile forecasts as well as an ensembling strategy. We show the effectiveness of our methods via simulation studies and an application to a public dataset.

Constructing Stationary Errors for Empirical Prediction Intervals

Presenter: Kenneth Lichtendahl
Co-authors: Casey Lichtendahl; Chris Fry

When the sample variance of either additive or multiplicative forecasting errors from a non-stationary model are used to construct an empirical prediction interval, the intervals are often too narrow or too wide. The problem mainly stems from heteroskedasticity in the standard additive and multiplicative error series: additive and multiplicative errors often have increasing and decreasing variance over time, respectively. When this heteroskedasticity arises, the corresponding time series themselves are often growing or declining over time. A lesser, but still significant issue is the positive serial correlation we find in most error series at longer horizons. To address the heteroskedasticity in an error series, we construct a stationary error series using a new error type that is in between an additive and multiplicative error form. The new h-step-ahead forecasting error is the difference between a Yeo-Johnson (YJ) transformation of a series' actual value and the same transformation of its h-step-ahead point forecast. To address the positive serial correlation in an error series, we develop a new estimator for the standard deviation of these YJ errors. We then use this YJ-error standard deviation estimator to set the widths of our h-step-ahead empirical prediction intervals. We demonstrate that these empirical prediction intervals offer improvements over the non-stationary ETS(Z,Z,Z) model's prediction intervals on two datasets, the U.S. EIA's electricity generation series by source and the M4 Competition's Monthly series.

Improving data-driven forecasting through constrained optimization

Presenter: Pablo Montero-Manso
Co-authors: Pablo Montero-Manso;

Time series forecasting models are usually fit by minimizing a functional on the one or multiple step-ahead in-sample loss. However, many time series contain information that can be exploited by adding terms in the objective function that isÂ being minimized, taking the form ofÂ constraints or penalizations.

For example, some time series exhibit multi-scale dynamics thatÂ can be captured by imposing that forecasts must be accurate at several levels of temporal aggregation.

In multivariate time series, cross-sectionalÂ hierarchies can also convey information.

In many forecasting contexts there is also external information available, business forecastsÂ might have to follow budget or logistic constraints, and physical systems might follow some conservation law.

We will present and analyse a unifying viewÂ of adding penalizationÂ terms of this kind to forecasting models. Summarizing the contributions of the talk:Â Â

Practical:

Â - Showcase several types of constraints and how they improve accuracy in many datasets, including the M5.

Â - The constraints can be added to existing models with minimal effortÂ (especially in modern implementations of neural networks and decision trees).

Â - They are sensitive to parameter tuning.

Methodological:

Â - Some constraints are related to data augmentation "tricks".

- In univariate models, cross-sectional constraints move local methods towards global or cross-learning.
- In modern machine learning, where data is fit to zero loss, the effect of some of these constraintsÂ disappears,Â Â the resulting model is the same as if no constraints were added, with profound implications.

Bayesian analysis of the mean-volatility dynamic factor model

Presenter: Mengheng Li
Co-authors: Mengheng Li; Dick van Dijk

The factor stochastic volatility (FSV) model of Chib et al. (2006) and Kastner et al. (2017) (among many others) extends the standard factor model by allowing for stochastic volatility (SV) in both factors and idiosyncratic components to account for empirically observed co-movement in the second-moment (i.e. volatility) of multivariate time series, such as financial returns. We observe, however that when we fit the FSV model is to the FRED MD dataset, there is a surprisingly strong co-movement in the variable-specific or idiosyncratic logvariance, regardless of the number of factors with SV we choose. Also, this common idiosyncratic volatility looks very similar to the factor volatility. In fact, it does not matter how many factors we use or if the factors are observable; as long as the DGP really features the first moment and second moment "commonalities" as coming from the same source, i.e. factors, projection onto the space spanned by factors should give residuals that are uncorrelated both in the first moment and in the second moment. However, our empirical results using the FRED MD data suggest otherwise.We introduce the mean-volatility dynamic factor model which explicitly separates the modelling of commonality in the first and the second moment of a vector time series. We motivate this model by illustrating the failure of the class of factor stochastic volatility models to explain the commonality in volatility left in residuals whatever the number of factors is chosen to be. We provide an order-invariant Bayesian estimation procedure that extends the result in Chan et al. (2018) via state space representation. In the empirical application, we show that sectoral industrial production and sectoral PCE inflation are better described by the proposed model rather than the factor stochastic volatility model.

Loss-Based Variational Bayes Prediction

Presenter: Gael Martin
Co-authors: Gael Martin; David Frazier; Ruben Loaiza-Maya; Bonsoo Koo

We propose a new method for Bayesian prediction that caters for models with a large number of parameters and is robust to model misspecification. Given a class of high-dimensional (but

parametric) predictive models, this new approach constructs a posterior predictive using a variational approximation to a loss-based, or Gibbs, posterior that is directly focused on predictive accuracy. The theoretical behavior of the new prediction approach is analyzed and a form of optimality demonstrated. Applications to Bayesian neural network models, autoregressive mixture models, and to the M4 forecasting competition, demonstrate that the approach provides more accurate results than various alternatives, including misspecified likelihood-based predictions.

Forecast combination puzzle in the HAR model

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

The Heterogeneous Autoregressive (HAR) model of Corsi (2009) has become the benchmark model for predicting realized volatility given its simplicity and consistent empirical performance. Many modifications and extensions to the original model have been proposed that often only provide incremental forecast improvements. In this paper, we take a step back and view the HAR model as a forecast combination that combines three predictors: previous day realization (or random walk forecast), previous week average, and previous month average. When apply- ing the Ordinary Least Squares (OLS) to combine the predictors, the HAR model uses optimal weights that are known to be problematic in the forecast combination literature. In fact, the simple average forecast often outperforms the optimal combination in many empirical applications. We investigate the performance of the simple average forecast for the realized volatility of the Dow Jones Industrial Average equity index. We find dramatic improvements in forecast accuracy across all horizons and different time periods. This is the first time the forecast combination puzzle is identified in this context.

Multi-population Mortality modelling: Extensions to Cairns-Blake-Dowd Model

Presenter: Pengjie Wang
Co-authors: Pengjie Wang; Athanasios Pantelous

Modelling mortality dependence for multiple populations is critically important in improving forecasting accuracy and examining population basis risk. We propose five multi-population extensions of the Cairns-Blake-Dowd (CBD) model to capture the mortality co-movements: the populations dependent on a common vector of time-varying factors, three variants of the common factor model, an augmented common factor model in which population-specific time effects are considered. We implement efficient and flexible Bayesian Markov Chain Monte Carlo (MCMC) algorithms for the models. We use the male mortality data in Group of Seven (G7) countries to show the efficacy of our models. Various model comparison and validation to demonstrate the models' adequacy, consisting of out-of-sample forecasting performance and posterior predictive check. A formal Bayesian model comparison using the integrated likelihood to approximate the marginal likelihood is also undertaken to compare the proposed models quantitatively. The empirical results show that the augmented common factor model yields satisfactory in-sample fit and out-of-sample forecast, and outperforms the other models in terms of marginal likelihood computation.

How Effective Is Social Distancing?

Presenter: Difang Huang
Co-authors: Difang Huang;

We identify the dynamic effects of social distancing policy on reducing the transmission of the COVID-19 spread. We build a model that measures the relative frequency and geographic distribution

of the virus growth rate and provides hypothetical infection distribution in the states that enacted the social distancing policy, where we control all time-varying, observed and unobserved, state-level heterogeneities. We apply our model to a panel of weekly COVID-19 infection cases and deaths of all states in the United States from February 20 to April 20, 2020, and find that during our sample period, social distancing intervention is effective in reducing the weekly growth rate in cases by 9.8% and in deaths by 7.0%. We show that the effects are time-varying that range from the weakest at the beginning of policy intervention to the strongest by the end of our sample period. We further demonstrate that the effects are cross-sectional heterogeneous as the states with higher income, higher education, more White people, more democratic voters, and higher CNN viewership have a more considerable reduction in the infection growth rate.

Assessing mortality inequality in the U.S.: What can be said about the future?

Presenter: Han Li
Co-authors: Han Li; Rob Hyndman

This paper investigates mortality inequality across U.S. states by modelling and forecasting mortality rates via a forecast reconciliation approach. Understanding the heterogeneity in state-level mortality experience is of fundamental importance, as it can assist decision making for policymakers, health authorities, as well as local communities who are seeking to reduce inequalities and disparities in life expectancy. A key challenge of multi-population mortality modeling is high dimensionality, and the resulting complex dependence structures across sub-populations. Moreover, when projecting future mortality rates, it is important to ensure that the state-level forecasts are coherent with the national-level forecasts. We address these issues by first obtaining independent state-level forecasts based on classical stochastic mortality models, and then incorporating the dependence structure in the forecast reconciliation process. Both traditional bottom-up reconciliation and the cutting-edge trace minimization reconciliation methods are considered. Based on the U.S. total mortality data for the period 1969"“2017, we project the 10-year-ahead mortality rates at both national-level and state-level up to 2027. We find that the geographical inequality in the longevity levels is likely to continue in the future, and the mortality improvement rates will tend to slow down in the coming decades.

A new hybrid ensemble framework for exchange rate forecasting using Taylor rule and Temporal convolutional network

Presenter: Yun Bai
Co-authors: Yun Bai; Yunjie Wei; Shouyang Wang

Exchange rate forecasting is a challenging problem for it is influenced by numerous factors, such as macroeconomics and the capital market. Existing research literature on exchange rate forecasting mainly focuses on a simple hierarchical structure and single models, which may have a large variance and instability. Our study proposes an ensemble framework for USD/GBP exchange rate forecasting, which considers different variables based on economic theory to measure the long trend and combines deep learning methods to capture short-term volatility. Wilcoxon rank test is used to classify decomposed series, which are then recomposed into different frequency components and predicted using a relatively appropriate method. Exogenous variables in our proposed model incorporate macro-economy, trading system, and short-term shocks. Low-frequency component is estimated based on the Taylor rule and uncovered interest rate parity. High-frequency component is modeled using technical indicators and temporal convolutional neural network. Empirical results indicate our framework presents superior forecasting capability and improve accuracy which is measured by four evaluation indicators.

Do sentiment indices always improve the prediction accuracy of exchange rates?

Presenter: Fumiko Takeda
Co-authors: Fumiko Takeda; Takumi Ito

This study aims to improve the prediction accuracy of the exchange rate model by changing how indices that capture market sentiment are constructed. Specifically, we construct the sentiment indices (SI) for the Japanese and American markets using the Google search volume index (SVI) for financial terms listed in the Japanese dictionary. For these SVIs, we select keywords based on the correlation between weekly changes in the yen-dollar rate and the SVI. We use 30, 20, and 10 keywords that are replaced at three different frequencies: three months, six weeks, and weekly. Weekly SVIs are obtained for the period from January 2013 to December 2017, which covers the tenure of the Bank of Japan' Governor, Haruhiko Kuroda. The training period is from January 2013 to June 2015, and the forecast period is from July 2015 to December 2017. To predict the exchange rate, we use a rolling regression, which keeps the length of the reference period constant at two and a half years. We use the interest rate parity and autoregressive models for the predictions. We compare the prediction accuracy using the mean squared prediction error, Clark and West's tests of equal predictive accuracy, and the direction of change test. When the SIs are updated every three months and six weeks, neither the interest rate parity model nor the autoregressive model shows improved prediction accuracy, even if the SI is added. However, when the SIs are updated weekly, prediction accuracy improves in both the interest rate parity and the autoregressive models as the number of words used to construct the SI increases. We conclude that frequently updated SIs can improve the short-term prediction accuracy, while SIs updated less frequently may not. Rather, when SIs are updated every three months or six weeks, the smaller the number of words used for SI, the better the prediction accuracy. Moreover, the best prediction accuracy is for the model without SI.

Exchange Rates Forecasting and Trading with Evolutionary Relaxed Support Vector Regression

Presenter: Shaolong Sun
Co-authors: Shaolong Sun; Erlong Zhao; Shouyang Wang; Yunjie Wei

A novel evolutionary learning approach, namely CS-RSVR, is proposed for exchange rate forecasting and trading. The proposed CS-RSVR approach can dynamically optimize the values of all SVRs' parameters through the CS evolutionary algorithm, and use acquired parameters to construct optimized RSVR in order for proceeded forecasting foreign exchange rates. Previous literature on exchange rate forecasting ignores how forecasting methods guide professional practice while focusing on forecasting performance. Many researchers have discussed exchange rate forecasting with the majority focusing on forecasting performance. However, accuracy is only one part of exchange rate forecasting. More important is how integrated approaches such as this can guide professional practice. To fill that gap, we extend our forecasting to test trading performance of exchange rates between the USD and four other major currencies (i.e. EUR, GBP, CNY and JPY). The experimental results demonstrate the CS-based optimized SVRs models significantly improve efficient in trading terms compared with other optimized SVRs models. Generally speaking, our proposed CS-RSVR model can be considered as a promising solution for exchange rates forecasting and trading.

Time series feature embedding for forecasting with deep learning

Presenter: James Nguyen
Co-authors: James Nguyen; Klaus Ackermann

The task of modelling time series data is difficult due to latent temporal characteristics present

within time series. This poses a greater issue with more granular time series data, which is becoming increasingly common in this digital age as a larger number of periods gives rise to more temporal characteristics. Deep learning methods struggle to capture temporal behaviour and require external features, to assist model development. These external features, however, are often unavailable, or if they are available, do not match the granularity of the original time series data, and hence do not provide a guaranteed method of capturing temporal behaviour.This paper proposes time2vec, a learned representation of time series features from fixed size time series windows. Time series features are useful in exploring different properties of the time series data from which they can be extracted, however, the process of extraction is generally complex and computationally inefficient. The representation time2vec presents a more efficient method to extract these features and can be implemented alongside deep learning methods to capture latent properties normally not captured within model where times series data is only available. We conduct a case study of load demand prediction in 24 different European countries and evaluate performance of multiple applications of feed forward neural networks with a time2vec implementation. We find that time2vec enables the model to capture finer patterns within time series data, and in some cases, provides greater prediction performance than other common time series methods.

A Look at the Evaluation Setup of the M5 Forecasting Competition

Presenter: Hansika Hewamalage
Co-authors: Hansika Hewamalage; Pablo Montero-Manso; Christoph Bergmeir; Rob Hyndman

Forecast evaluation plays a key role in how empirical evidence shapes the development of the discipline, guiding the concentration of research efforts to the most promising methods. Domain experts are interested in error measures that are relevant for their decision making needs, but these error measures might make the experiment unreliable, their results unlikely to hold in a future time or similar data. The reliability properties of dozens of metrics have already been discussed, along with the qualities of the experimental datasets, but reliability has hardly been quantified in an objective way. We propose a way to capture the reliability of an experiment using Rank Stability, which evaluates how much the rankings of an experiment differ in between similar datasets, when the models and the error measures are kept constant. We use rank stability to study the evaluation setup of the M5 and compare it to other error measures, at different levels of the hierarchy as well as across time. We find that the evaluation setup of the M5 is less reliable than other measures, its ranking is likely to change in future applications, both on other datasets and on the same dataset at a different time. The main drivers of instability are hierarchical aggregation and scaling. Price-weighting reduces the stability of all tested error measures. The scale normalization of the M5 error measure produces comparatively less stability than other scale-free errors. Hierarchical levels taken separately are less stable the more we aggregate, and their combination is even less stable than individual levels. Aggregation and stability can be linked to the influence of the much debated magic numbers. Though no measure is able to retain both stability and meaning, some do better than others in that spectrum. We show positive tradeoffs where hierarchical upper levels of aggregation can retain part of its importance without affecting stability. In addition to the analysis of the M5, many of these consequences can be applied to general hierarchical forecast benchmarking. Rank stability can be used in further research for the study and design of experimental setups.

Dependency Learning Graph Neural Networks for Multivariate Forecasting

Presenter: Abishek Sriramulu
Co-authors: Abishek Sriramulu; Christoph Bergmeir; Nicolas Fourrier

Graph Neural Networks (GNN) are models that can learn from graphs with nodes representing variables of interest and edges defining the relationships between variables. GNNs have recently gained popularity in the forecasting domain due to their ability to model complex spatial and temporal

patterns in tasks such as traffic forecasting and region-based demand forecasting. Most of these methods require a predefined graph as input, wherein real-life multivariate time series problems a well-predefined dependency graph very rarely exists. This requirement makes it harder for GNNs to be utilised widely for multivariate forecasting in other domains such as retail or energy. In this work, we review methods for constructing a dependency graph using neural network models and statistical structure learning models on multivariate time series ultimately aiming to enable the use of GNNs for multivariate forecasting even when a well-defined graph does not exist. Furthermore, we propose a hybrid approach combining neural networks and statistical structure learning models to self-learn the dependencies for multivariate forecasting. The statistical structure modelling in conjunction with neural networks provides a well principled and efficient approach by bringing in causal semantics to determine dependencies among the series. Finally, we demonstrate significantly improved performance using our proposed approach over real-world benchmark datasets without a pre-defined dependency graph.

Probabilistic forecasts using expert judgement: the road to recovery from COVID-19

Presenter: George Athanasopoulos
Co-authors: George Athanasopoulos; Rob J. Hyndman; Nikolaos Kourentzes; Mitchell O'Hara-Wild;

The COVID-19 pandemic has had a devastating effect on many industries around the world including tourism, and policy makers are interested in mapping out what the recovery path will look like. In this paper we propose a novel statistical methodology for generating scenario-based probabilistic forecasts based on a large survey of 443 tourism experts. The probabilistic scenarios map out pessimistic, most-likely and optimistic paths to recovery. Taking advantage of the natural aggregation structure of tourism data due to geographical locations and purposes of travel, we propose combining forecast reconciliation and forecast combinations implemented to historical data to generate robust COVID-free counterfactual forecasts, to contrast against. Our empirical application focuses on Australian tourism, analysing international arrivals and domestic flows. Both sectors have been severely affected by travel restrictions in the form of international and interstate border closures and regional lockdowns. The two sets of forecasts, allow policy makers to map out the road to recovery and also estimate the expected effect of the pandemic.

Daily hotel demand forecasting using a hybrid deep learning approach

Presenter: Yuan Qin
Co-authors: Yuan Qin; Shanshan (Vera) Lin

The artificial intelligence (AI) models have been widely applied in forecasting tourism demand given their strong ability to explain nonlinear patterns and relationships between input and output variables, as well as their potential to improve forecasting accuracy. The artificial neural network (ANN) and its variations are most frequently used AI-based techniques by tourism scholars in the past two decades due to their strong capability in processing imperfect data and handling nonlinearity. Long short-term memory (LSTM) networks, an extension of RNN, can automatically learn the time lags of time series data as well as the changes to the time lags. By having such feature, LSTM has been widely used to solve time series forecasting problems and also been demonstrated a powerful tool for tourism demand forecasting. This study aims to forecast daily hotel demand in Hangzhou, China adopting the LSTM networks that is capable of utilizing the historical data of multivariate time series including visitor volume data, search engine data, weather data, online consumer's ratings, confirmed cases of COVID-19 pandemic, and hotel room rates. A composite search index based on searching findings of Baidu index is constructed using a novel approach by combing machine-learning based feature selection methods and generalized dynamic factor model (GDFM). A universal thermal climate index (UTCI) is introduced as a proxy to assess the city's weather conditions and air quality.

Daily data is collected from Hangzhou's Tourism Economics Lab over a period of November 2019 and March 2021. To capture both linear and nonlinear patterns in the input and output variables, a hybrid modeling strategy that combines time series models and LSTM networks is adopted. A group of error measures and statistical tests including the mean absolute percentage error (MAPE), root mean squared percentage error (RMSPE) and Diebold-Mariano (DM) test are applied to evaluate the forecasting accuracy of all testing models. The findings of the study are expected to enrich empirical evidence of demonstrating LSTM networks' good forecasting ability.Key words: Hotel forecasting, LSTM, hybrid model, forecast combination, COVID-19

Forecasting Visitor Arrivals amid COVID-19

Presenter: Haiyan Song
Co-authors: Haiyan Song;

The study looks at the challenges faced by tourism forecasters during the COVID-19 pandemic. The discussion will be focused on the ways in which both qualitative and the quantitative approaches could be used jointly to develop scenario forecasts of tourism demand during a major public health crisis. The effectiveness and reliability of the forecasts generated by this hybrid method is tested based on the data on a number of tourist destinations within Asia Pacific where the demand for tourism were badly affected.

A Scalable Ensemble of Global and Local Models for Long-term Energy Demand Forecasting

Presenter: Kasun Bandara
Co-authors: Kasun Bandara; Hansika Hewamalage; Rakshitha Godahewa

In this work, we present key insights on the model development strategies used in our ensemble based energy demand forecast framework. The proposed framework outperforms state-of-the-art univariate models in the time series forecasting literature and has achieved overall 4th position in the 2nd Technical Challenge, organised by the IEEE Computational Intelligence Society. The proposed solution is a four-layered energy prediction framework that addresses multiple challenges posed by the competition dataset, including the presence of missing values, the limited availability of energy consumption data, the availability of the various types of exogenous variables, etc. The first layer of our framework consists of data aggregation and data imputation modules that handle the data completeness of the dataset. The second layer, the primary forecast engine of our framework, employs a combination of univariate and global forecasting models. In this competition, the employment of global models becomes particularly useful in generating accurate forecasts for households with limited energy consumption history. This is because of the ability of global models to learn common energy consumption patterns across multiple households. On the other hand, univariate forecast models are utilised to capture the energy dynamics within a single household. To bring model diversity to our ensemble framework, we use a host of linear and non-linear forecasting models, including Auto-Regressive Integrated Moving Average model (ARIMA), Pooled-Regression (PR) model, Lasso-Regression model (LR), Multilayer Perceptron Neural Network (MLP), Random Forest (RF) and Gradient Boosting Trees (GBT). The third layer of the proposed framework aggregates the forecasts of each base model to monthly level. Finally, the fourth layer, the ensemble layer combines the individual predictions of multiple forecast models using two ensemble schemes: the geometric mean ensemble scheme and the median ensemble scheme. Our results show that the proposed ensemble forecast methodology is able to improve the base accuracy of individual models.

Causal Inference Using Global Forecasting Models for Counterfactual Prediction

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

Causal inference determines causal relationships of interventions and effects, and measures the impact of interventions. It is important in situations where fully randomised control trials (A/B testing) are too costly, ethically questionable, or otherwise not possible. Under the "potential outcome" approach introduced by the Rubin Causal model, the estimation of causal effects is usually calculated by the difference of counterfactual predictions and the true values of the time series affected by some treatment. Nonetheless, the underlying forecasting methods used in the current counterfactual prediction frameworks are mostly univariate models. In contrast, the state-of-the-art in time series forecasting has moved from such local, per-series univariate modelling to global forecasting models (GFM) that learn across many time series.Our study proposes a global forecasting and inference method based on recurrent neural networks (RNN) to predict policy interventions' causal effects on an outcome over time through this counterfactual approach. This allows us to add to the training phase of the counterfactual modelling all the time series of a dataset. Once we train treated and control time series simultaneously over the pre-treatment period, we input more information to the modelling. Also, we do not need to make equivalence assumptions between distributions of the control and treated outcomes in the pre-treatment period. This allows us to achieve better accuracy and precisely isolate the effect of an intervention. We compare our novel approach with local univariate approaches on simulated and real-world datasets. Our results show that our novel method can outperform the accuracy of state-of-the-art predictions, thereby estimating the size of a causal effect more accurately. The experimental results are statistically significant, indicating our framework generates better counterfactual predictions. The contributions of this research are particularly useful to show that by incorporating the global forecasting approach we can implement more complex and realistic modelling to predict reliable counterfactual outcomes.

A Strong Baseline for Weekly Time Series Forecasting

Presenter: Rakshitha Godahewa
Co-authors: Rakshitha Godahewa; Christoph Bergmeir; Geoffrey Webb; Pablo Montero-Manso

Many businesses and industries require accurate forecasts for weekly time series nowadays. The forecasting literature however, does not currently provide easy-to-use, automatic, reproducible and accurate approaches dedicated to this task. We propose a forecasting method that can be used as a strong baseline in this domain, leveraging state-of-the-art forecasting techniques, forecast combination, and global modelling. Our approach uses four base forecasting models specifically suitable for forecasting weekly data: a global Recurrent Neural Network model, Theta, Trigonometric Box-Cox ARMA Trend Seasonal (TBATS), and Dynamic Harmonic Regression ARIMA (DHR-ARIMA). Those are then optimally combined using a lasso regression stacking approach. We evaluate the performance of our method against a set of state-of-the-art weekly forecasting models on six datasets. Across four evaluation metrics, we show that our method consistently outperforms the benchmark methods by a considerable margin with statistical significance. In particular, our model can produce the most accurate forecasts, in terms of mean sMAPE, for the M4 weekly dataset.

Forecast evaluation of electricity peak load in South Korea: from traditional to hybrid models

Presenter: Juyong Lee
Co-authors: Juyong Lee; Youngsang Cho

In order to improve the accuracy of the peak load forecasting, a comprehensive and comparative study is conducted. We compare the performance of seasonal autoregressive integrated moving average with exogenous variables (SARIMAX) model, artificial neural network (ANN), support vector regression (SVR), and long short-term memory (LSTM) for the single models, SARIMAX-ANN, SARIMAX-SVR, and SARIMAX-LSTM for the hybrid models, and extreme gradient boosting (XGboost) model are selected. One step further, this study proposes a multi hybrid boosting (MHboost) model that conducts iterative training with LSTM model and combines with SARIMAX model. Predictive power of MHboost (mean absolute percentage error (MAPE) 2.8%) is superior to all the other models of this study, and also more accurate than the current peak load forecasting model of Korea (MAPE 4.4%). In addition, we conducted an additional analysis to compare the predictive power of our models with the Korea Short-term Load Forecastor (KSLF) model, which Korea currently utilizes for peak load forecasting. The RMSE of the KSLF model is 3531.0519 and the MAPE is 4.37%; thus, the LSTM, SARIMAX-SVR, SARIMAX-LSTM, XGboost, and MHboost models outperform the KSLF model in both the overall performance and highest peak load prediction. Therefore, the performance of the current peak load forecasting model used in Korea, which consists of panel and time series models, can be improved by grafting hybridand boosting models.

Application of Deep Neural Network and Recurrent Neural Network for Short-term Electricity Load Forecasting

Presenter: Pyae Pyae Phyo
Co-authors: Pyae Phyo;

Short-term load forecasting (STLF) is an essential role in planning and operating electricity. It is hard to meet optimal units for electrical manufacturers. The objective of this research is to minimize the forecasting error using deep learning models for electric industries. The historical sequential data required is from the Electricity Generating Authority of Thailand (EGAT). Deep neural network (DNN) model is used to handle back propagation problems and recurrent neural network (RNN) model is used to handle sequential data. Moreover, long short-term memory (LSTM) is combined with RNN to overcome long dependency problems of sequential data. The key challenge of this forecast is the variation of load on each day and the seasonality of the load. Therefore, the proposed models are trained and tested using the cleaned data during the period of December 2013 to July 2017 by replacing the general holidays load, bridging holidays load and outliers. Mean absolute percentage error (MAPE) is used as the measurement for forecasting accuracy and tracking signal is applied to determine under forecast or over forecast. The results of DNN model are compared with the results of RNN with LSTM model. The results show that DNN model provide better accuracies than RNN with LSTM model.

Tidy Time Series Anomaly Detection for Load Forecasting

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

Accurate load forecasting is vital for effective energy management, as forecast underestimation can result in blackouts, whereas overestimation may result in energy wastage. However, the quality of historical data for load forecasting can be affected in several ways, such as data integrity attacks, missing values, incorrect readings, technical aberrations. These issues make data unreliable and untrustworthy and can have a direct impact on forecast accuracy and subsequent decision making. This work develops a framework for detecting anomalies in tidy time series data. An anomaly is defined as an observation that is predicted as very unlikely given the robust time series forecast models. The algorithm works with tidy temporal data provided by the 'tsibble' package and produces an 'outstable', a tsibble with flagged anomalies and their degree of outlierness. An approach based on extreme value theory is applied to residual series in order to calculate a data-driven anomalous

threshold. The proposed framework can also provide a cleansed tsibble that closely integrates with the tidy forecasting workflow used in the 'fable' package. A number of different approaches are available for the data cleansing process. The wide applicability and usefulness of this proposed framework in load forecasting will be demonstrated using various synthetic, real-world, and publicly available benchmark datasets including data from Global Energy Forecasting Competitions. This framework is implemented in the open-source R package 'outstable'.

Application of VAR Models for Semiconductor Manufacturing Equipment Market: Impulse Responsive Analysis to Chip Product Market and User's Financial Performance

Presenter: Takashi Ogawa
Co-authors: NA

In recent years, semiconductor chip has become a strategic key component to achieve emerging digital transformation in many industries. With an anxiety for the logistic constraint due to a global spread of COVID-19 and political frictions for trading of high technology products, leading countries and economic zone are forced to make sure the stable supply chains for key components including semiconductors.

Having these backdrop, the timely procurement and installation of semiconductor manufacturing equipment (SME) into the wafer fabs will be essential for semiconductor companies and investors. To achieve the strategic acquisition of manufacturing capacity, the market monitoring and evaluation for SME becoming critical.

In this study, we applied VAR, VEC, and BVAR models for SME market data and performed the impulse responsive analysis (IRA), aiming to evaluate how SME market will be influenced by changes of the end-product market and user's financial performance. Concretely, we constructed the models with monthly global shipment data of SME and semiconductor end-products "" Memory devices (DRAM & Flash memory), Logic device and Foundry products and performed the IRA after the determination the optimum model by statistical procedure.

As a result, the analysis successfully made it possible to evaluate impact of change in each end-product market on SME market at the first time. Further, as the recent capital spending in semiconductor industry has been dominated only by 5~10 major companies, we tried to evaluate how the quarterly financial performance of the single major company will impact on SME market in the mid-term by applying Mixed-Frequency VAR model and the IRA. In a case study, we analyzed the impact of Taiwan Semiconductor Company (TSMC) as well-known as the top foundry vendor and the results implied the mid-term impact of each financial performance such as gross-profit and PPE on SME market.

Short-term Facebook Network Traffic Forecasts for Data-center Planning

Presenter: Yilun Chen
Co-authors: Slawek Smyl;

Our work outlines the end-to-end machine learning modeling for Facebook Infrastructure's short-term network planning at data-centers level. We leverage causal signals from different Facebook Infrastructure systems, including hardware usage, power supply and service placement, to predict Fabric Aggregator network traffic.Our presentation focuses on two parts: machine learning model deployment and the models' interpretability and explainability. The whole model process includes data pipeline (data cleaning, pre-processing, feature engineering), modeling and training (model selection, hyper-parameter tuning and model interpretation), model serving (testing, deploying and maintaining). The input signals include power, hardware usage metrics, rack types and service types at hourly time granularity. The regressors are fitted into Gradient Boosting Tree based models with Pinball loss, not considering temporal dimension, for different data-centers. The initial feature space

is more than 100 dimensions, which might include unnecessary signals introducing noises to models. Therefore, we leverage RFE (recursive feature elimination) into the pipeline and trade off between models' complexity and explainability. After iterative experiments, we arrive a relevant stable feature space especially for top 5 to 10 important features evaluated with impurity score and permutation importance. The representability and predicability of features for traffic forecasts are validated with SHAP (SHapley Additive exPlanations). The 95% quantile loss function is used in feature selection, hyper-parameter tuning (with Hyperopt: Distributed Asynchronous Hyper-parameter Optimization), model selection and training. Specifically, we iteratively dropping least significant features or searching through hyper-parameter space to lower 95% Pinball loss on both training set and validation set. We report 95% Pinball loss and the percentage of exceedance on testing set and leverage operational decisions on power supply, hardware and services placements projections in inference period.

A new times series forecasting method and its application to the supercomputer power consumption prediction

Presenter: Jirí TOMCALA
Co-authors: NA

In general, accurate prediction methods are computationally very demanding and therefore their calculation takes a long time. Most of these are machine learning methods, where a lot of time is consumed to create a mathematical model. On the contrary, fast prediction methods are not very accurate. These are methods that are based on some simple principle such as zeroth algorithm, exponential smoothing, or moving average. In this paper, a new prediction method is introduced, which includes the advantages of both approaches mentioned above. This method does not create any mathematical model, but uses the procedures of some machine learning methods to refine originally inaccurate simple prediction methods. Although these fast simple methods lose a bit of speed, they gain a lot of accuracy. The level of the sacrifice of speed for accuracy can then be tuned using several parameters of this new method. The comparison with several currently most used prediction methods was performed on complex time series of the supercomputer infrastructure power consumption.

Can we measure inflation expectations using Twitter?

Presenter: Cristina Angelico
Co-authors: Cristina Angelico; Filippo Quarta; Marcello Miccoli; Juri Marcucci

Drawing on Italian tweets, we employ textual data and machine learning techniques tobuild new real-time measures of consumers' inflation expectations. First, we select somekeywords to identify tweets related to prices and expectations thereof. Second, we build a set of daily measures of inflation expectations on the selected tweets, combining the LatentDirichlet Allocation (LDA) with a dictionary-based approach, using manually labelled bigrams and tri-grams. Finally, we show that Twitter-based indicators are highly correlated withboth monthly survey-based and daily market-based inflation expectations. Our new indicators provide additional information beyond market-based expectations, professional forecasts, and realized inflation. Moreover, they anticipate consumers' expectations, proving to be a good real-time proxy. The results suggest that Twitter can be a new timely source for devising a method to elicit beliefs.

Nowcasting food inflation with a massive amount of online prices

Presenter: Pawel Macias
Co-authors: Pawel Macias; Damian Stelmasiak; Karol Szafranek

The consensus in the literature on providing accurate inflation forecasts underlines the importance of precise nowcasts. In this paper, we focus on this issue by employing a unique, extensive

dataset of online food and non-alcoholic beverages prices gathered automatically from the web since 2009. Using a highly disaggregated framework, we perform a real-time nowcasting experiment among popular, simple univariate approaches. We show that accounting for the information on online food prices in a simple, recursively optimized model already leads to a substantial reduction in errors in the nowcast. This framework outperforms a variety of other approaches, including the judgmental methods and various model combinations. In turn, pure estimates of online price changes are particularly effective in nowcasting inflation when working on short samples. During the COVID-19 the nowcasting quality of the model with online prices has been comparable with judgmental approaches, but still superior with respect to other frameworks. We also demonstrate that meticulous product selection and classification is essential for providing accurate inflation nowcasts. We conclude that the use of online prices can markedly aid the decision process at central banks by providing timely information on changes in the headline inflation.

Nowcasting Turkish Food Inflation Using Daily Online Prices

Presenter: Baris Soybilgen
Co-authors: Baris Soybilgen; Ege Yazgan; Huseyin Kaya

For this study, we have been scraping daily food prices from retail chains' websites since July 2018. Overall, we have collected over 5.9 million data points. Using these food prices, we formed 132 food price subindexes compatible with those of the Turkish Statistical Institute (Turkstat). Then we calculate the primary food inflation rate using Turkstat's consumer price index weights and our food price subindexes. When we compare our food inflation rate with that of Turkstat, it seems that both food price indexes move very closely until February 2020. However, the first Coronavirus restrictions caused our food prices to jump, whereas Turkstat's food prices didn't experience the same spike. As we observe food prices daily, we can also construct a daily food price index and use it to nowcast Turkstat's food inflation before Turkstat announces its own index. Our results show that we can nowcast Turkstat's food inflation rate successfully using a regular bridge equation framework even before the end of the current month. We also show that trimming each day's highest and lowest prices at the subindex level improves our daily food price index's nowcasting performance. Finally, we show that we can successfully nowcast headline Turkish inflation using our daily food price and daily exchange rate.

How local is the local inflation factor? Evidence from Emerging European Countries

Presenter: Oguzhan Cepni
Co-authors: Oguzhan Cepni; Michael P. Clements

We consider whether inflation is a 'global phenomenon' for European emerging market economies, as has been claimed for advanced or high-income countries. We find that a global inflation factor accounts for more than a half of the variance in the national inflation rates, and show that forecasting models of national headline inflation rates that include global inflation factors generally produce more accurate path forecasts than Phillips Curve-type models, and models with local inflation factors. Our results are qualitatively unaffected by allowing for sparsity and non-linearity in the factor forecasting models.

A new test for common breaks in heterogeneous panel data models

Presenter: Eiji Kurozumi
Co-authors: Eiji Kurozumi; Peiyun Jiang

In this paper, we develop a new test to detect whether the break points are common in heterogeneous panel data models where the time series dimension T could be large relative to cross section dimension N. The error process is assumed to be cross-sectionally independent. The test is based on the cumulative sum (CUSUM) of the ordinary least squares (OLS) residuals. We derive the asymptotic distribution of the detecting statistic under the null hypothesis, while proving the consistency of the test under the alternative. Monte Carlo simulations show good performance of the test.

Robust Covariance Matrix Estimation in Time-Series Econometrics: A Review

Presenter: Masayuki Hirukawa
Co-authors: Masayuki Hirukawa;

In time-series econometrics, long-run variance estimators play an important role in estimating model parameters more efficiently and drawing more accurate statistical inference on the parameters. In this paper, a non-technical review of the long-run variance estimation is provided. The review covers both parametric and nonparametric estimators. Considering that kernel methods are dominant among all estimation procedures, we also present recent developments in kernel-smoothed estimators and related inference. The information collected in this paper can help practitioners decide on a suitable long-run variance estimator.

Dynamically Time Warped Cointegration

Presenter: Anton Skrobotov
Co-authors: Anton Skrobotov; Artem Prokhorov; Giuseppe Cavaliere

Dynamic Time Warping is a pattern recognition tool from machine learning initially developed to handle distorted sound data recorded at unequal frequencies. In the simplest form, it is a way of aligning signals y_i and x_j which come at different frequencies ”“ a situation often encountered in forecasting. Instead of using time proximity, it matches x_j to y_i to maximize similarity of the two time series. We develop the theory of cointegration for dynamically time warped data. We show that DTW cointegration is more general than traditional cointegration analysis and hence offers a degree of robustness to the standard time series tools. We derive a form of the functional central limit theorem that can be applied to DTW and illustrate using simulations how to detect cointegration in the underlying same frequency data when all we observe is the DTW signal. Our approach works when standard cointegration tools cannot be applied directly. We construct, and use for forecasting, an error correction model based on a detected DTW cointegration relationship.

Inference in the Nonparametric Stochastic Frontier Model

Presenter: Valentin Zelenyuk
Co-authors: Valentin Zelenyuk; Christopher F. Parmeter; Leopold Simar; Ingrid Van Keilegom

This paper discusses conducting various types of inference in the stochastic frontier model when it is estimated using nonparametric methods. We discuss a general and versatile inferential technique that allows for a range of practical hypotheses of interest to be tested. We also discuss several challenges that currently exist in this framework to alert researchers to potential pitfalls. Namely, it appears that when one wishes to estimate a stochastic frontier in a fully nonparametric framework, separability between inputs and determinants of inefficiency is an essential ingredient for correct empirical size of a test. We showcase the performance of the test with a variety of Monte Carlo simulations.

Exponential Smoothing with Regularisation

Presenter: Kandrika Pritularga
Co-authors: Kandrika Pritularga; Ivan Svetunkov; Nikolaos Kourentzes

Exponential smoothing is widely used in practice and has shown its efficacy and reliability in many real cases. The state of the art approach to exponential smoothing is selecting the most appropriate model from a pool based on an information criterion, such as Akaike Information Criterion. However, there is evidence in the literature that information criteria do not consistently choose the most accurate models. The selection becomes particularly challenging on small samples, where parameter uncertainty makes the distinction between models more difficult. Motivated by the challenges in parameter estimation, building on the LASSO literature, we explore l1-regularisation for exponential smoothing that is potentially able to simultaneously estimate parameters and shrinking unneeded components to zero.We investigate regularisation for pure additive exponential smoothing models. Our objective is to obtain a continuous spectrum of model selection, rather than the potentially abrupt switching from one model to another that is currently the norm with information criteria. Even when the model form is adequately identified by both approaches, regularisation can help with parameter estimation, mitigating sampling-induced uncertainties. We propose a series of modifications to the implementation of state-space exponential smoothing and conduct a simulation study to assess the performance of the exponential smoothing with regularisation, tracking both the model selection efficiency, and the resulting forecasting performance.

A new taxonomy for vector exponential smoothing and its application to seasonal time series

Presenter: Huijing Chen
Co-authors: Huijing Chen; Ivan Svetunkov; John Boylan

In the context of short-term demand forecasting, businesses are often required to forecast the demand of seasonal products, based on very few complete seasonal cycles of data. This makes the forecasting task difficult, but one possible solution is to make use of the readily available cross-sectional information from a homogeneous group. For example, a product family may contain essentially the same product with many variations such as size or colour. It is reasonable to assume that seasonality is common within the family and can thus be estimated using cross-sectional information. In this research, we propose a framework based on vector ETS (VETS), and provide a conceptual contribution by devising a taxonomy of pure additive and multiplicative ETS models, bringing together the factors of components, parameters and initial values. This framework is general in the sense that it is not limited to, although inspired from, the modelling of seasonal components. It can be applied to level, trend (including damped trend) and any possible combination of these features. Experiments with artificially generated data will provide an example of how this taxonomy works. This is also evaluated empirically on the performance of the methods derived from the taxonomy. Model selection is addressed in this part to provide guidance on implementing the framework. Insights drawn from this research on the impact from improved efficiency and flexibility of estimation on multivariate forecasting will be discussed.

How to make multiplicative ETS work for you

Presenter: Ivan Svetunkov
Co-authors: Ivan Svetunkov; John Boylan

Exponential smoothing in state space form (ETS) is a popular forecasting technique, widely used in research and practice. While the additive error ETS models have been well studied, the multiplicative error ones have received much less attention in forecasting literature. Still, these models

can be useful in cases, when one deals with positive data, because they are supposed to work in this situation. Unfortunately, the classical assumption of normality for the error term breaks this property and might lead to non-positive forecasts on positive data. In order to address this issue we propose using Log Normal, Gamma and Inverse Gaussian distributions, which are defined for positive values only. We demonstrate what happens with ETS(M,,) models in this case, discuss conditional moments of ETS with these distribution and show that they are more natural for the models than the Normal one. We conduct the simulation experiment in order to study the bias introduced by point forecasts in these models and then compare the models with different distributions on an example of real data.

Forecasting bounded time series with time-varying bounds

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

Many forecasting applications involve response variables which are both continuous and known to be bounded. This is especially the case for all variables representing rates, percentages and proportions which are frequently encountered in statistical practice and lie in the unit interval. In particular one may be interested in forecasting the spread of an epidemic in terms of cases per a number N of persons, the unemployment rates of a given country or the proportion of time spent by animals in a certain activity. In all those applications the bounds of the interval the observations lie in are assumed to be fixed throughout the time series. If this assumption makes sense in some cases, it can be misleading and negatively impacts the quality of the forecasts when the bounds actually vary over time whilst not being observed. In particular it is highly relevant for energy applications, such as wind power forecasting. Indeed wind power generation is a stochastic process that is theoretically double-bounded by zero and the nominal power of the turbine. However in practice the upper bound may also change in time, while being unknown, for example in case of curtailment actions.First we perform a simulation study to show the impact of a wrong fixed bound assumption on synthetic time series scaled by a bound which varies over time. We focus on both point and probabilistic forecasting and evaluate the accuracy of the forecasts in the worst and best cases. In the former it is assumed that the bounds are fixed, to 0 and 1 in the case of the unit interval, while in the latter it is assumed that one has perfect knowledge of the actual time-varying bounds. Then we define a new modelling framework in which the bounds can be adaptively estimated from the data by different algorithms. In particular we consider Expectation-Maximization algorithms and sequential Bayesian filtering.Finally the proposed algorithms are both applied to the synthetic data generated during our simulation study and to the real test case of the Anholt offshore wind farm in Denmark, with emphasis placed on 10-minute-ahead point and probabilistic forecasts.

Forecasting Hourly Wind Power Production In Sweden With Time Series Models

Presenter: Xiangli Meng
Co-authors: Xiangli Meng; Dong Jin

This paper focuses on forecasting hourly wind power production data in Sweden with time series approaches. The main characteristics of hourly production data are periodicity and volatility jump. The spectral analysis is used for analyzing the seasonal patterns in hourly production. The HEGYtype seasonal unit root test for hourly data is used to test for stationarity and the data reject the presence of unit roots at most of the seasonal frequencies. To incorporate volatility jumps of wind power production, GARCH-type models are employed. Four methods are employed to forecast the conditional mean of hourly production: Seasonal ARIMA model, ARAR algorithm, Holt-Winter seasonal algorithm and Periodic Autoregressive models. The forecasts are made on differents horizons for intra-day market and day-ahead market. Evaluating forecasts results with root mean square error and mean absolute error.

Day-Ahead Forecasting of Instantaneous Power at Wind Farms

Presenter: Leo May
Co-authors: Leo May; Jethro Browell

A challenge in operating low-carbon electricity networks is the procurement of backup power for fast, corrective rebalancing of electricity generation and consumption immediately following the failure of large generators, transmission lines or other network components. Wind farms are well placed to secure revenue by providing this backup power, however they must be capable of contracting ahead of time to provide a fixed amount of power headroom at a high reliability and long lead times, so called 'firm power'. As stochastic generators, the capability to provide firm power is defined in the tail of a probabilistic forecast of minimum instantaneous power within time intervals specified by the power contracts. This is a deviation from the current practice of probabilistic energy (average power) forecasting at wind farms. A methodology for creating and scoring day-ahead probabilistic forecasts of minimum instantaneous power at wind farms is presented. The benefits of forecasting minimum power directly rather than repurposing energy forecasts are shown to be large. This suggests that specific forecasts of instantaneous power should be used to manage ancillary services provision. The competition winning approach of gradient boosted decision trees is utilized to handle multiple input features and create non-parametric quantile forecasts.This methodology, and a benchmark based on current practice, are demonstrated in a case study comprising three wind farms in Great Britain (GB). In GB, system security requirements necessitate firm power provision with a response time of 1 second to be contracted day-ahead, meaning probabilistic forecasts of minimum instantaneous power are required on lead-times of 11 to 35 hours ahead. The forecasting methodology is extended to create probabilistic forecasts of the complete cumulative distribution function of instantaneous power in one-hour periods. This process allows the integration of time limited constraints to firm power allocation strategies from wind farms including the allocation of storage capacity day-ahead. The resulting probabilistic forecasts are reliable (calibrated) in contrast to re-purposing conventional energy (average power) forecasts, which are not. The forecasting of aggregated firm power from multiple wind farms is also investigated to demonstrate spatio-temporal effects on the variability of instantaneous power and the resulting forecast skill.

Forecasting port container throughput with multisource big data and optimized machine learning

Presenter: Gang Xie
Co-authors: Gang Xie;

Accurately predicting port container throughput has become increasingly difficult due to international transfers in the manufacturing industry, trade policy changes, and frequent irregular events in the global trade. To improve predictive accuracy, we develop a least squares support vector regression model with a novel hybrid adaptive particle swarm optimization and gravitational search algorithm (LSSVR-APSOGSA) to forecast the throughput using multisource big data. These data include search query data from Baidu and Google, economic and trade indexes, and correlated container throughput at other ports. Moreover, a kernel principal component analysis is used to extract principal components having nonlinear information from the big data, to avoid multicollinearity or overfitting problems. Using data from twelve ports in China and in the United States, this study is the first to investigate methods for predicting port container throughput using different independent variables. The empirical results show that the proposed model demonstrates significant average improvement rates in both point and interval forecasting, ranging from 6.63-33.51% better than benchmark models. This indicates the approach can effectively predict port container throughput in increasingly complex trade situations. With these more accurate forecasts, the government and shipping companies can develop more appropriate maritime policy and plans to provide better customer service, and ports can improve the efficiency of container operations.

Predicting operating speed in transshipment processes

Presenter: Benedikt Sonnleitner
Co-authors: Benedikt Sonnleitner; Claudia Ehrig; Simon Rauch

In Process Management we need to predict how fast orders are processed among various workstations. In our study we examine trans-shipment in transportation. The headworker needs to anticipate how fast the shipments are processed, given the number of workers in and the overall load.In our project we compare three different formulations of this problem: (1) as a time series problem with a strong seasonality. Here the available workforce, the overall load that needs to be processed in this shift and the processing speed (the target) form a multivariate time series. (2) as a regression problem with a multidimensional target vector. The target vector describes the speed at each hour of the specific shift. (3) we use the setup of (2), but instead of predicting the volume of processed units, we predict the share of processed unit per period at a certain station. This is valid, since the units that will be processed in the predicted shift are known in advance, but not how early in the shift all of the orders are processed. We use a multilayer perceptron, an ARIMA(X) model, and a linear model for the first problem formulation, while for formulations 2 and 3 we rely on multilayer perceptrons. We present the various problem formulations, the various models linked to the formulations, and provide empirical evidence of their performance, supporting decisions in process management.

On direct horizon-specific forecasting with state space models

Presenter: Benedikt Sommer
Co-authors: Benedikt Sommer; Pierre Pinson; Klaus Holst

Many applications require forecasting of a dependent variable multiple-steps ahead into the future. Say, a forecaster is tasked by a client to forecast the weekly number of products to be delivered to customers in the upcoming weeks. Already closed sales with future delivery dates are naturally a strong predictor. However, the predictive power decreases with the forecast horizon since customers do not procure products with delivery dates far in the future. The forecaster's choice is thus to directly model the predictions of each forecast horizon with a separate model due to the natural ability to handle horizon-dependent features. Furthermore, caused by market changes and the client's varying product portfolio, the forecaster expects the model coefficients to be time-varying.Given the convenience to estimate time-varying coefficients in state space models (SSMs), we explore their capabilities for direct multi-step ahead forecasting. We first highlight that the residuals of a direct h-step ahead forecasting model are serially correlated for $h > 2$, even when the model is correctly specified. Then, we show that using parametrizations of linear Gaussian SSMs that ignore the residual serial correlations can have severe negative consequences when the standard Kalman filter is used for inference. We propose the following two methods. To allow for the application of the standard Kalman filter, we show that one can augment the SSM to whiten the observation noise of the model. Since this approach increases the size of the latent space, our second approach is to use an extension of the standard Kalman filter that avoids state augmentation. We compare both methods in a simulation study before we apply them to forecast the weekly empty container pickups and returns to the network of a large container shipping company.

What should be taken into consideration when forecasting oil implied volatility index?

Presenter: Panagiotis Delis
Co-authors: Panagiotis Delis; Stavros Degiannakis; Konstantinos Giannopoulos

Crude oil plays a crucial role in the international economy since it is considered a key commodity for all the international economies. In this study, we focus on forecasting oil volatility index (OVX), which is the market's expectation of future oil volatility, by incorporating information from other asset classes. Due to the fact that a number of papers have applied the heterogeneous autoregressive (HAR) modeling framework used for forecasting realized volatility in forecasting implied volatility (IV) indices, we consider crucial to investigate whether this implementation is sufficient in the case of OVX or it fails to offer predictive gains as opposed to the realized volatility case. First of all, we employ a long memory test by estimating the Hurst exponent, which shows strong long memory in the OVX time series. This test has been implemented in order to justify the use of the HAR structure. Apart from the HAR model specifications, we also use the dynamic model averaging (DMA) approach in order to allow for different sets of IV indices from other asset classes to be applicable at different time periods. Moreover, we applied the DMA approach by including only the main IV indices of stock market for investigating their impact on OVX. The evaluation framework consists of statistical loss functions and an options straddle trading strategy as an additional economic criterion to evaluate the volatility forecasts. The main results show that the forecasters should focus more on the major stock market's IV indices, and more specifically the DJIA Volatility Index (VXD), which provides strong evidence for high interconnectdness of oil and stock market. Finally, we draw the conclusion that the impact of the main stock market's IV indices is short-term because the crude oil is financialized and the fact that there is no longer-term predictive ability indicates evidence for market efficiency hypothesis.

Forecasting crude oil prices with DSGE models

Presenter: Michal Rubaszek
Co-authors: Michal Rubaszek;

We run an oil prices forecasting competition among a set of structural models, includingvector autoregressions and dynamic stochastic general equilibrium models. Our resultshighlights two principles. First, forecasts should exploit the mean reversion of the real oilprice over long horizons. Second, models should not replicate the high volatility of oil pricesobserved in sample. Abiding by these principles, we show that a small scale DSGE modelperforms much better in real oil price forecasting than the random walk as well as vectorautoregressions.

Crude oil and green energy finance: an asymmetric connection?

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

The Paris Agreement on climate change mitigation was negotiated in the midst of a period of crude oil glut which had sent the price of oil sharply downward. An oversupply of shale oil production, the deceleration of the Chinese economy, and environmental policy campaigns to get away from fossil fuels have all been blamed for a weak oil demand.In which way does the green-energy financial market interact with crude oil prices?We contrast a selection of green-energy stock indices with the benchmark crude West Texas Intermediate. In order to model the dynamics of co-movements and potential asymmetry in volatility spillovers between returns, we use a combination of VARMA and a special bivariate GARCH model, in the spirit of the GJR-GARCH. This approach allows for asymmetry in the analysis of the bi-directional pattern of volatility spillovers.Our findings give rise to interpretations in terms of the characteristics of news impact on volatilty. For instance, the steepest ascent of volatility predominantely occurred when both the oil price and the green-erergy index decreased. During the Covid-19 pandemic, however, increases in oil prices emerged as driving forces behind the fluctuations.

Effects of outcome feedback on inflation judgments and confidence in them

Presenter: Xiaoxiao Niu
Co-authors: Xiaoxiao Niu; Nigel Harvey

People's inflation judgments have been found to be right-skewed and heterogeneous. There is limited work on how their accuracy can be increased. Feedback provides one potential approach. Outcome feedback is the most common type that forecasters receive; it simply provides the forecaster with the realized values of the variable for which forecasts were generated. Feedback can have two effects. First, it may incentivise people; it has this effect only while it is present. Second, it can enable people to learn how to perform tasks better; this effect continues after feedback has been removed. To distinguish these effects experimentally, feedback is either given or not given in a first (manipulation) session. Better performance in the former than in the latter case in a second (test) session without feedback indicates that feedback produces a learning effect. In our experiment, 120 participants were recruited via the web platform Prolific. They were randomly accolated into either feedback condition or no-feedback condition and asked to make inflation judgments for 2019 for 20 different countries, 10 in the manipulation session and 10 in the test session. Confidence that each estimation would be within 20% of the outcome was indicated by moving a horizontal slider. Results showed that the overall error in judgments was not different between the two conditions. However, constant error (bias) was significantly lower and variable error (random noise or scatter) was significantly higher in the feedback condition. Confidence judgments showed lower bias and better calibration in the feedback condition. Moreover, the quality of these judgments was constant over the two sessions in the feedback condition but gradually deteriorated in the no feedback condition.

These results show that feedback has mixed effects. It helps people reduce the degree to which they overestimate inflation but increases the scatter in their judgments. Furthermore, without feedback, people erroneously think their performance improves with practice. Providing them with feedback enables them to appreciate that this is not so.

Ensembles of Judgmental Forecasts in Dynamic Environments containing Structural Breaks

Presenter: Matthias Seifert
Co-authors: Matthias Seifert; Yun Shin Lee

We study the effectiveness of judgmental forecasts in time series environments involving structural breaks. Inspired by Grushka-Cockayne et al. (2017), we propose two ensembling approaches, which aggregate trimmed individual judgments based on forecasters' empirically elicited or inferred belief that the underlying data series has been subject to structural change. We test the predictive performance of our models using judgments obtained from two online experiments as well as by studying analysts' forecasts in several real world contexts. Our findings indicate the superior accuracy of ensemble methods that utilize inferred beliefs about structural breaks as their aggregation criterion, which outperform other commonly known aggregation rules including symmetric & asymmetric trimming as well as simple averages. Our findings also suggest that empirically elicited break predictions suffer from systematic biases, which severely reduce ensembling effectiveness. We discuss implications for managerial practice.

Forecasting in Supply Chain Management

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

The first step of the supply chain in a company starts with PROCUREMENT. Here obviously the

price forecast of the raw products plays an important role. In two ways this this a difficult task: to estimate the best time for the procurement we have to forecast the whole price development along a planning horizon (and not only one time point in the future). Second, raw products (e.g. metals, chemicals, energy) are often world market prices, which implies that we have to forecast world models.The PRODUCTION itself includes forecasting models for quality indicators and predictive maintanance.At the end of the supply chain we have to forecast SALES. This is an important information for the control of the production lines and the planning of storage capacity. It is the forecast uncertainty which is an essential parameter for the planning of storage capacities to balance intermediate over and under forecasts.In the talk we will see different feedforward and recurrent neural networks for the 3 above tasks.

Bayesian Estimation of the Italian Business Cycle through SVSS through

Presenter: Lazzaretto Enrico
Co-authors: Enrico Lazzaretto;

In this paper we propose a measure for the Italian Business Cycle (or Output Gap) obtained by performing Bayesian Model Averaging after the implementation of the Stochastic Variable Specification Search procedure. We present a simultaneous trivariate unobserved component model of inflation, unemployment and output which are interrelated through statistical stochastic processes and theoretical economic models, namely the Phillips Curve and Okun's Law. The SVSS methodology is applied to include a wide variety of model specifications through the use of the binary indicators which allow us to make model comparison beyond the traditional test for nested models. Specifically, we apply thestochastic search to different trend specifications and distribution assumptions(Normal and Student-t) of innovation terms for both the inflation ad the output measurement equations. The model is estimated by a Monte Carlo Markov Chain routine and Bayesian model averaging is applied to estimate the model chosen after running the SVSS algorithm. The main findings are that output gap is an important determinant for the inflation gap, Okun's is a well established relationship even for the Italian case, potential output is best described by a time-varying trend and that innovation terms are generated from Student-tdistributions rather than Normal ones.

Decoupling Shrinkage and Selection for the Bayesian Quantile Regression

Presenter: Tibor Szendrei
Co-authors: Tibor Szendrei; David Kohns

In this paper we extend the idea of Hahn and Carvalho (2015) by decoupling shrinkage and sparsity for the Bayesian Quantile Regression (BQR). The procedure follows two steps: In the first step, we shrink the quantile regression posterior through state of the art global-local priors and in the second step we sparsify the posterior through an efficient variant of the adaptive lasso, the signal adaptive variable selection (SAVS) algorithm. We propose a new variant of the SAVS which automates the choice of penalisation through quantile specific loss-functions that are valid in high dimensions. We show in large scale simulations that our selection procedure not only decreases bias irrespective of the true underlying degree of sparsity in the data, compared to the un-sparsified regression posterior, but also attains model selection consistency. We apply our two-step approach to a high dimensional growth-at-risk (GaR) exercise using the McCracken & Ng (2020) database. The prediction accuracy of the un-sparsified posterior is retained while yielding interpretable quantile specific variable selection results. Our procedure can be used to communicate to policymakers which variables drive downside risk to the macro economy.

Global-Local Priors and the Illusion of Sparsity

Presenter: David Kohns
Co-authors: David Kohns;

Recent studies have shown that in high dimensions, the choice of model can have a substantial impact not only on forecasting performance, but also on the data informed degree of sparsity, leading often to a false certainty, or "illusion of sparsity" (Giannone et al., 2020). In this paper, we investigate whether the recently popularised global-local (GL) priors are able to uncover the true model complexity and propose a new variant of the horseshoe prior of Carvalho et al. (2010) that stays explicitly agnostic about the underlying degree of sparsity. The simulations as well as empirical applications show that the proposed prior not only provides better forecasts than the spike-and-slab, but also yields better mixing properties. The resultant cleaner identification of the underlying degree of sparsity beats out the spike-and-slab as well as competing GL priors in simulations and leads to reduced uncertainty in the empirical applications. The results suggest that the proposed prior is more robust at detecting the underlying degree of sparsity in the data.

A holistic approach for improving the renewable energy forecasting model and value chain: The Smart4RES project

Presenter: Georges Kariniotakis
Co-authors: George Kariniotakis; Simon Camal

In this paper we present the research directions and innovative solutions developed in the European Horizon 2020 project Smart4RES (http://www.smart4res.eu) for advanced modelling and forecasting of weather variables necessary to optimise the integration of renewable energy (RES) production (i.e. wind, solar) into power systems and electricity markets. Smart4RES gathers experts from several disciplines, from meteorology and renewable generation to market- and grid-integration. It aims to contribute to reach very high RES penetrations in power grids of 2030 and beyond, through thematic objectives including: "¢ Improvement of weather and RES forecasting (+10-15% in performance);"¢ Streamlined extraction of optimal value through new forecasting products, data market places, and novel business models;"¢ New data-driven optimization and decision-aid tools for market and grid management applications;"¢ Validation of new models in living labs and assessment of forecasting value vs costly remedies to hedge uncertainties (i.e. storage). Smart4RES focuses both on improving forecasting models of weather (e.g. physical models, data assimilation, Large Eddy Simulation) and RES production (e.g. seamless models, highly resolved predictions), and on addressing applications in power grids. Developments in the project have been formalized in Use Cases that cover a large range of time frames, technologies and geographical scales. For example, use-cases on power grids refer to the provision of ancillary services to the upper-level grid (e.g. balancing power) and the local grid (e.g. voltage control and congestion management), where the accurate forecasts of variable generation are key for accurate decision-making. Hierarchical forecasting permits to have compatible forecasts at different levels of the grid. A grid state forecasting will quantify dynamically issues of RES integration in distribution grids. Collaborative forecasting investigates the improvement associated to local data sharing between distributed RES plants. This data sharing paves the way to a data market where agents exchange measurements, predictions or other types of valuable data. Lastly, data-driven approaches will streamline decision-making by simplifying the model chain of bidding RES production, storage dispatch or predictive management electricity grids. They will also provide interpretable hindsight to decision-makers by integrating the decisions of experts (human-in-the-loop)

Forecast of future technological developments in German energy companies and energy start-ups

Presenter: Theresa Fritz
Co-authors: Theresa Fritz;

A large number of new framework conditions and trends have led to the development of new technologies in the German energy industry in recent years. As a result, the number of patent applications

rose sharply. Numerous studies reveal that by analysing patent data, relevant technological changes can be determined and forecasts can be modelled. In the context of the analysis, patent applications from German energy companies and energy start-ups are collected via the German Patent and Trademark Office. With the help of a time series modelling and the modelling of a bass diffusion curve, future technological developments are forecast. The results show that in the areas of energy storage, IT / telecommunications and electromobility, a sharp increase in patent applications can be expected from 2022. Classic fields such as energy generation, energy supply and infrastructure / transport, on the other hand, will have reached a phase of saturation by 2030. This also applies to technological developments in the field of sustainability. Based on these findings, a strong technological change and rethinking in the industry can be demonstrated.

An integrated forecasting and optimization approach applied in trading renewable energy.

Presenter: Akylas Stratigakos
Co-authors: Akylas Stratigakos; Simon Camal; Andrea Michiorri; Georges Kariniotakis

Recent years have seen an advent in the development of energy analytics tools, aiming at leveraging available data to improve decision-making under uncertainty. The typical approach for data-driven decision-making comprises of two components, namely forecasting and optimization. First, uncertain parameters, such as renewable energy or market prices, are forecasted, and then are used as input in an optimization module to derive optimal decisions. In the context of statistical learning, forecasting models are trained to minimize prediction error without considering the effect of forecasts on the cost of the subsequent decision, i.e., the forecast value. In this work, we present a data-driven approach that integrates forecasting and optimization in a single model that learns a policy conditioned on explanatory data. We propose a decision tree algorithm trained to minimize decision costs and derive prescriptions via a weighted Sample Average Approximation of the original stochastic optimization problem. Further, we develop a framework to assess impact of explanatory data on the efficacy of prescribed decisions. Finally, we evaluate the proposed method on the case of trading renewable energy in a day-ahead electricity market subject to imbalance costs.

Binary Choice with Asymmetric Loss in a Data-Rich Environment: Theory and an Application to Racial Justice

Presenter: Andrii Babii
Co-authors: Andrii Babii;

We study the binary choice problem in a data-rich environment with asymmetric loss functions. In contrast to the asymmetric regression problems, the binary choice with general loss functions and high-dimensional datasets is challenging and not well understood. Econometricians have studied the nonparametric binary choice problems for a long time, but the literature does not offer computationally attractive solutions in data-rich environments. In contrast, the machine learning literature has many computationally attractive algorithmsthat form the basis for much of the automated procedures that are implemented in practice, but is focused mostly on loss functions that are independent of individual characteristics. One of the main contributions of our paper is to show that the theoretically valid predictions of binary outcomes with a generic loss function can be achieved via a very simple reweighting of the logistic regression or state-of-the-art machine learning techniques, such as boosting or (deep) neural networks. We apply our analysis to racial justice in pretrial detention.

Protecting Time Series for Data Privacy with Minimal Forecast Loss

Presenter: Matthew Schneider
Co-authors: Matthew Schneider; Jinwook Lee

Forecasting models may be negatively impacted due to anonymization requirements in General Data Protection Regulation (GDPR). To measure the severity of this problem, we derive theoretical bounds for the loss to forecasts from additive exponential smoothing models using protected data. Following the guidelines of anonymization from the GDPR and California Consumer Privacy Act (CCPA), we develop the k-nearest Time Series (k-nTS) Swapping and k-means Time Series (k-mTS) Shuffling methods to create protected time series data that minimizes the loss to forecasts while preventing a data intruder from detecting privacy issues. We formally model optimization problems with a two-party data privacy framework that includes the utilities of a data provider and data intruder. We apply our data protection method to thousands of time series and find that it maintains the forecasts and patterns (level, trend, and seasonality) of time series well compared to standard data protection methods suggested in legislation. Substantively, our paper addresses the challenge of protecting time series data when used for forecasting.

Forecasting for Social Good

Presenter: Bahman Rostami-Tabar
Co-authors: Bahman Rostami-Tabar; Mohammad M Ali; Tao Hong; Rob J Hyndman; Michael D Porter; Aris Syntetos

Forecasting plays a critical role in the development of organisational business strategies. Despite a considerable body of research in the area of forecasting, the focus has largely been on the financial and economic outcomes of the forecasting process as opposed to societal benefits. Our motivation in this study is to promote the latter, with a view to using the forecasting process to advance social and environmental objectives such as equality, social justice and sustainability. We refer to such forecasting practices as Forecasting forSocial Good (FSG) where the benefits to society and the environment take precedence over economic and financial outcomes. We conceptualise FSG and discuss its scope and boundaries in the context of the "Doughnut theory" . We present some key attributes that qualify a forecasting process as FSG: it is concerned with a real problem; it is focused on advancing social and environmental goals and prioritises these over conventional measures of economic success; and it has a broad societal impact. We also position FSG in the wider literature on forecasting and social good practices. We propose an FSG maturity framework as the means to engage academics and practitioners with research in this area.Finally, we highlight that FSG: (i) cannot be distilled to a prescriptive set of guidelines,(ii) is scalable, and (iii) has the potential to make significant contributions to advancing social objectives

Frequency-domain information for active portfolio management

Presenter: Gonçalo Faria
Co-authors: Gonçalo Faria; Fabio Verona

In this paper we assess the benefits of using frequency-domain information for active portfolio management. First, we forecast the bond risk premium and equity risk premium using a methodology that allows isolating frequencies (of the predictors) with the highest predictive power. The resulting forecasts are more accurate than those of traditional forecasting methods for both asset classes. Next, we use these forecasts in the context of active portfolio management and find that they lead to better portfolio performance by several measures.

Modelling the Relation between the US Real Economy and the Corporate Bond-Yield Spread in Bayesian VARs with non-Gaussian Disturbances

Presenter: Tamas Kiss
Co-authors: Tamas Kiss; Par Osterholm; Hoang Nguyen; Stepan Mazur

In this paper we analyze how asymmetry and fat tails affect the relationship between the real economy and the corporate yield spread, a popular predictor of real activity. We use quarterly real-time US data to estimate Bayesian VAR models with stochastic volatility and non-Gaussian disturbances. We find that, after controlling for stochastic volatility, innovations in GDP growth can be well-described by a Gaussian distribution. In contrast, both the unemployment rate and the yield spread appears to benefit from being modelled using non-Gaussian innovations. When it comes to real-time forecasting performance, we find that the yield spread is an important predictor of GDP growth, and accounting for stochastic volatility matters for – mainly density – forecasts. The incremental improvements of non-Gaussian, skewed innovations are limited to forecasts of the unemployment rate, though. Our results suggest that stochastic volatility is of first order importance modelling the relationship between yield spread and real variables, which actually captures most of the non-Gaussian behaviour of the variables.

MEM or/and logARMA: Investigation of Predictive Performance for Realized Volatility

Presenter: Stanislav Anatolyev
Co-authors: Stanislav Anatolyev;

There are two popular autoregressive conditional density model classes for series of positive financial variables such as realized volatilities, price ranges, durations, and the like. One is a class of multiplicative error models (MEM) where the conditional mean is modelled autoregressively, while the specified shape of the conditional distribution imposes evolution on higher order moments. The other class contains logARMA models ”“ that is, ARMA models for logarithms of original series, with a flexible possibly heteroskedastic conditional distribution imposed. For MEM models, generating predictions is straightforward, while for logARMA models, additional numerical integration is required.We compare the two model classes in terms of out-of-sample predictability, using real data on realized volatility. We generate pseudo-out-of-sample forecasts from simple and complex MEM and logARMA models, as well as from their smoothed-AIC combinations. The estimated forecast combination weights show that both model classes are able to generate competitive forecasts, more parsimonious models having an advantage, with a serious edge given to the MEM modelling.

Modeling and Forecasting Macroeconomic Downside Risk

Presenter: Andrea De Polis
Co-authors: Andrea De Polis; Davide Delle Monache; Ivan Petrella

We document a substantial increase in downside risk to US economic growth over the last 30 years. By modeling secular trends and cyclical changes of the predictive density of GDP growth, we recover an accelerating decline in the skewness of the conditional distributions, with significant, procyclical variations. Decreasing trend-skewness, turning negative in the aftermath of the Great Recession, is associated with the long-run growth slowdown stared in the early 2000s. Short-run skewness fluctuation imply negatively skewed predictive densities ahead, and during recessions, often anticipated by deteriorating financial conditions, while positively skewed distributions characterize expansions. The model delivers competitive out- of-sample (point, density and tail) forecasts, improving upon standard benchmarks, due to financial conditions providing strong signals of increasing downside risk.

Capturing GDP nowcast uncertainty in real time

Presenter: Paul Labonne
Co-authors: Paul Labonne;

Nowcasting methods rely on timely series related to economic growth for producing and updating estimates of GDP growth before publication of official figures. But the statistical uncertainty attached to these forecasts, which is critical to their interpretation, is only improved marginally when new data on related series become available. That is particularly problematic in times of high economic uncertainty. As a solution this paper proposes to model common factors in scale and shape parameters alongside the mixed-frequency dynamic factor model typically used for location parameters in nowcasting frameworks. Scale and shape parameters control the time-varying dispersion and asymmetry round point forecasts which are necessary to capture the increase in variance and negative skewness found in times of recessions. It is shown how cross-sectional dependencies in scale and shape parameters may be modelled in mixed-frequency settings, with a particularly convenient approximation for scale parameters in Gaussian models. The benefit of this methodology is explored using vintages of U.S. economic growth data with a focus on the economic depression resulting from the coronavirus pandemic. The results show that modelling common factors in scale and shape parameters improves nowcasting performance towards the end of the nowcasting window in recessionary episodes.

Forecasting real GDP growth for Africa

Presenter: Philip Hans Franses
Co-authors: Philip Hans Franses; Max Welz

We propose a simple and reproducible methodology to create single equation forecasting models for low frequency macroeconomic variables. Our methodology is illustrated for forecasting annual real GDP growth rates for 52 African countries, where data start in 1960. The models (potentially) included lagged growth rates of all countries, as well as a cointegration relationship to capture potential common stochastic trends. With a few selection steps, our methodology quickly arrives at a reasonably small forecasting model per country. Compared with benchmark models, the single equation forecasting models seem to perform quite well.

Hierarchical forecasts of Diabetes mortality in Mexico by marginalization and sex to establish resource allocation

Presenter: Eliud Silva
Co-authors: Eliud Silva; Corey Sparks

The mexican population has experimented an astounding rise in type II Diabetes mortality as well as a growing trend for the economic burden in the recent years. Thus, this paper's purpose is to propose an approach to establish a distribution of resource allocation objectively to face the future economic burden. To this purpose, hierarchical forecasts of Diabetes mortality to 2030 by sub-domains of the population are estimated. Hierarchical time series model is used based on marginalization and sex. The forecasts confirm that differences related to sub-domains will be significant. In fact, the rates will increase most notably both in low and high marginalized levels of the country. The hierarchical method just provide point forecast and until now prediction intervals cannot be estimated according to the specialized literature. However, it is assumed that for our objective this kind of previsions are enough. Likewise, it is considered that the marginalization level is the same for all the forecast horizon.There is not a similar application for mexican data that estimates objectively the resource allocation to cope with the Diabetes mortality by marginalization and sex. Thereby, the most recommendable budget distribution should be mainly addressed among the low and high levels. Implications of these estimates should support unpostponable health policy in general and for the mentioned sub-domains in particular.

Ensemble CNN Approach for Patient Classification Covid-19 using CT Scan

Presenter: Ariela Stefanini
Co-authors: Ariela M. Stefanini; Gabriel Rodrigo Gomes Pessanha; Reinaldo Antonio Gomes Marques; Ciro de Oliveira

In 2019 coronavirus pandemic was required to find ways to further early diagnosis of COVID-19. This research aim is to implement deep neural networks (CNN). This model will classify COVID-19 through a chest CT scan at last assist in the diagnosis of COVID-19. To tackle this challenge, we employ the RSNA International COVID-19 Open Radiology Database (RICORD). This database is a collection of COVID-related imaging datasets with expert annotations. It has 120 identified chest CT scans of negative patients and 120 COVID-positive chest CT scans with DICOM extension. Machine learning has to analyze all images at the same time to generate an overview of patient analysis. This way, we implement a CNN along with cross-validation and data augmentation techniques for reducing false negatives and treating the patient set images as one. Besides, we implemented image processing to remove the noise present in a CT scan. Thus, only the parts of lung infections going to CNN training. So that makes a significant increase in 5% accuracy. After identifying the steps needed to generate a model for this problem, we applied it to five different architectures like VGG 19, Resnet 50, Inception, Xception, SqueezeNet. Each architecture has its advantages and disadvantages, and to supply the needs that each architecture presents, we generated ensembles among them. In general, the segmentation has shown that it is possible to capture regions with COVID-19 and differentiate them from other diseases. This study has pointed out the precision of 85,6% and with the lower false-positive rate for COVID-19 CT scan image detection in the worse case. Future works can be done to improve our results, in particular, more databases may include detection of multiple disease cases such as pneumonia, bronchitis. Thus, it may be indicated more information to the classification result.

A Comparative Study of Predicting ICU Length of Stay

Presenter: Qixuan Hou
Co-authors: Qixuan Hou;

Predicting ICU length of stay (LOS) is critical in improving the quality of care, enabling efficient hospital plans, and meticulously managing hospital resources. With modern electronic healthcare records containing a large amount of patient data, this study analyzes the complete historical medical records, medications, produces, and lab data of more than 40,000 critical care patients. Statistically significant factors in determining LOS will be identified. Different models will be implemented and compared, including Random Forecast, Gradient Boosting, Support vector machine, Decision Tree, and Neural Networks. Additionally, the ensemble method will be explored to further improve the results. Accuracy, sensitivity, specificity, and AUC will be used to evaluate and compare the models.LOS is defined as the number of days that a patient is hospitalized. Shortening LOS can considerably reduce hospital costs. Predicting LOS is regarded as the significant first step towards effectively controlling costs. A good prediction of LOS enables efficient hospital management. Surgeries can be scheduled electively based on the upcoming availability of beds. Besides, hospitals can intervene early in the patients who are identified to likely have prolonged ICU LOS. Thus, better predicting ICU LOS is important in the healthcare system. The study will also identify important factors which determine LOS. Understanding the factors is crucial in predicting LOS. Additionally, the awareness of the factors among doctors and caregivers could lead to the improvement of current clinical pathways.

Forecasting with Bayesian stochastic volatility Fourier series Model: The case study of Covid-19 effect on Food futures and Marine stocks

Presenter: Alexandros Pasiouras
Co-authors: Alexandros Pasiouras; Theodoros Daglis

In this paper, we construct a hybrid model, comprised of Bayesian Vector Autoregressive model with Bayesian stochastic volatility (SVAR-SV), and Fourier Series (FS). We test the proposed model's performance in terms of forecasting ability, comparing it with simple Bayesian stochastic volatility (SV), and also with baseline econometric autoregressive models. By estimating the average prices of the major Food futures in the stock market, and also the average prices of the biggest Marine companies' stocks, we test the effect of Covid-19 on these stocks, through the hybrid model proposed. The rationale of the case study is that the Covid-19 pandemic (fear of spread, economic instability, and lockdown measures) hindered the performance of marine companies, and affected the food prices, with those two affecting one another. Based on the findings, a shock is apparent from Food futures to the Marine companies' stocks, and the hybrid model proposed is the best model, in terms of forecasting ability.

Factors of predictive power for mineral commodities

Presenter: Patric Papenfuß
Co-authors: Patric Papenfuß; Amelie Schischke; Andreas Rathgeber

In our study, we individually forecast 26 metal prices, which are regarded important for the possible transformation pathways of the German Energy System, one-month ahead and outperform the predefined benchmark model, a random-walk (with drift) in 17 (18) cases. In addition, we investigate the resulting models in regard to the predictive power of the variables included. To start, we analyze variables, which the literature previously identified to have predictive power on commodity markets. The relevant, commodity-specific factors are pre-selected through a correlation analysis, followed by a BIC-based regression selection. We analyze their out-of-sample, one-month-ahead predictive content through linear regression models, in comparison to a random-walk and a random-walk with drift benchmark. First and foremost, our results show a significant outperformance of our models for 17 (18) of the 26 commodities considered, especially those in the minor metals sector. The differences regarding predictability between precious and industrial to minor metals are remarkable, highlighting the importance to individually analyze and model prices. Focusing on the influential factors, the value factor has a highly significant, negative effect on the prediction and determination of prices. Additionally, the US dollar index, as an approximation for the exchange rate, the S&P500, and the long-term interest rate, as well as the commodity-specific demand show significant predictive ability on prices.

Forecasting aggregate household consumption and aggregate income: A simulation-based model selection approach

Presenter: Robert Kunst
Co-authors: Robert Kunst; Adusei Jumah

Household consumption and disposable income provide a role-model example for error correction. On given national-accounts data, we explore whether and to what degree the cointegration properties benefit forecasting. It evolves that statistical evidence on cointegration is not always equivalent to better forecasting properties by the implied cointegrating structure. The exercise is conducted in the framework of simulation-based forecast-model selection. The simulation-based method explicitly permits letting the forecast model choice depend on the intended time horizon of the forecast. The simulation-based approach permits the determination of the sample size, beyond which the more sophisticated model dominates with regard to its forecasting properties.

Investigating Climate Variability Impact on Sugarcane Using Satellite-Based NDVI Time Series

Presenter: Willard Zvarevashe
Co-authors: Willard Zvarevashe;

Climate variability has a direct impact on the agricultural production yield therefore, an understanding of the trend and periods is very important for future planning and sustainability, especially in the rain-fed agriculture. The Normalised Difference Vegetative Index (NDVI) is positively correlated to the yield in many crops. In recent times it has been used to gauge the expected yield for crops. In this study it used to investigate the climate variability of sugarcane in a selected region in KwaZulu-Natal province, South Africa. This region supplies a large percentage of sugarcane to South Africa and export market. The NDVI time series is decomposed to identify the seasonal, trend and random series. In order to find the periods and the impact of precipitation and temperature in the NDVI a data adaptive method, ensemble empirical mode decomposition and synchronisation is applied. A non-linear regression model to find the impact of climatic variables on the variability of sugarcane is proposed. The study is the first to use NDVI to detect climate variability for the mostly rain-fed sugarcane plantations in KwaZulu-Natal province South Africa. Furthermore, very few statistical studies do not look at the growth phase of sugarcane but use yield as a basis for analysis of climate variability impact.

Tailings dam slope stability analysis using the WARIMAX-GARCH forecasting model

Presenter: Jefferson Royer Chaurais
Co-authors: Jefferson Royer Chaurais; Alvaro Faria; Luiz Albino Teixeira Junior; Gionei Gaio

Slope stability is a very sensitive issue for geo-technicians in the mining sector. This is because a dam break may destroy cities with many negative impacts for the environment, society and the mining companies themselves. In Brazil, there were two recent major disasters. The collapse of Samarco's FundÃ£o dam in 2015 and the break of Vale's B1 dam in 2019 that resulted in more than 300 deaths and homeless families, besides the catastrophic environmental consequences including the extinction of many flora and fauna endemic species. Therefore, it is very important for mining companies to focus on the stability of those structures. A slope is considered stable when its resistive inner forces are larger than the motive ones. This quotient is named Factor of Safety (FS). To monitor it, it is common to install sensors to measure the surface displacement and the phreatic level. Hence, the forecasting of the measurements of those instruments may be used to predict the FS value and therefore when an accident may occur, which gives the company more time for a possible repairing action, or even for evacuation. In this paper we present how the forecasting of instruments readings may be used to predict future values of a slope's FS.

OpenDrift Model : a Search and Rescue forecast evaluation at Rio Janeiro coast.

Presenter: Thiago Zamith
Co-authors: Thiago Zamith Cunha;

The safeguarding of human life at sea is intrinsically related to the continuous monitoring of the Brazilian coast and, mainly, to the use of Operational Oceanography systems that provide observations and forecasts. In this paper, an evaluation will be made through a case study that took place recently on the Brazilian southeast coast where the drift of the SAR target will be simulated on OpenDrift using the Leeway package, a Search and Rescue (SAR) forecast model developed by the US Coast Guard. Then, the results will be compared with the search areas calculated and referenced to the International Maritime and Aeronautical Search and Rescue Manual (IAMSAR). Nevertheless, highlighting its importance in the

decision-making process and discuss the advances of the forecast drift of objects in the ocean at Brazilian Navy.