International Symposium of Forecasting 2019

June 16-19, Thessaloniki, Greece

General Chair: George Athanasopoulos, Program Chair: Aris Syntetos

Book of Abstracts

Looking into the Future with the Reverend! Bayesian Forecasting in the 21st Century

Presenter: Gael Martin

Bayesian forecasting quantifies uncertainty about the future using the language and rules of probability. A forecast distribution is produced, conditioned only on past observations; all uncertainty about the parameters of the forecasting model, and about the model itself, having been integrated, or 'averaged', out, via these simple rules. Inherent to this natural and coherent approach however, are two key requirements: 1) that the integral which defines the forecast distribution can be computed; 2) that the assumed model (or model set) which underpins that integral is a correct representation of (or contains) the true data generating process.

In this presentation, I explore these two requirements. First, I place Bayesian forecasting within the broad historical context of Bayesian computation per se. Beginning in 1763, with the intractable posterior integral of Reverend Thomas Bayes, I track the journey to the 21st century, in which Bayesian forecasts are produced using the computational techniques of Markov chain Monte Carlo (MCMC), pseudo-marginal MCMC, approximate Bayesian computation and variational Bayes. A startling discovery is that - conditional on the forecasting model being correctly speciÖed - when it comes to forecast accuracy, the way in which parameter uncertainty is quantified, and the forecast distribution thus computed, matters very little! Second, I tackle the more challenging issue: how to re-think the Bayesian forecasting paradigm - and the use of computation therein - when one acknowledges empirical reality: namely, that any assumed forecasting model, or set of such models, is a misspecified representation of the true data generating process.

Recent advances in neural forecasting methods

Presenter: Tim Januschowski

Deep learning has seen increasing adoption in practical forecasting applications over the last years. Their practical success is also reflected in recent forecasting competitions such as the M4 competition and a number of Kaggle competitions. In my talk, I will review a number of deep learning approaches for forecasting and draw parallels to classical forecasting methods. In particular, I will summarize research at AWS AI Labs such as work on pure deep learning methods like DeepAR, a probabilistic forecasting algorithm based on recurrent neural networks, and recent efforts into combinations of neural networks and probabilistic models.

The effect of serialization on forecasting for remanufacturing: an empirical assessment

Presenter: Thanos Goltsos

In manufacturing, accurate demand forecasts drive accurate orders for procurement, which in turn lead to efficient inventory control of SKUs. In that sense, the demand generating process is the predominant source of uncertainty in the system, whereas supply (orders for replenishment) is treated as a deterministic decision variable. In remanufacturing however, where items are recovered from customers to be brought back into a good-as-new state and re-sold, uncertainty on the returns process compounds on any uncertainties from the demand generating process, leading to a 'dual sourced uncertainty'. The focus then shifts to forecast net demand (demand – returns), to inform procurement that will serve any part of the demand that cannot be covered by remanufactured returns. Returns forecasting literature has focused on forecasting returns, with some forgiving assumptions on the demand generating process. Based on a unique dataset from a UK

remanufacturer, we assess the performance of these models, and extend their application to accommodate demand assumed unknown. We find that serialization accounts for considerable improvements by better characterising the relationship between returns and past sales. We find that demand forecasting performance is as important as that of return forecasts, and should not be brushed aside. Finally, we explore how this double-sourced uncertainty affects the inventory performance. We conclude with opportunities for further improvements for remanufacturers.

Demand Forecasting Using Sell-through Data

Presenter: Jente Van Belle
Operational forecasting in supply chain management supports a variety of short-term planning decisions among which production scheduling, inventory management and raw materials procurement. In this respect, improving short-term forecast accuracy is a way to build a more agile and responsive supply chain for manufacturing companies. Traditionally, demand forecasting relies on well-established univariate forecasting methods to extrapolate historical demand. Collaboration across the supply chain, including information sharing, is suggested in the literature as a means to improve upon the forecast accuracy of established traditional methods. In this work, we investigate different modeling approaches and forecasting methods to incorporate downstream data in demand forecasting. Where empirical findings on information sharing mainly focus on point-of-sale data in two-level supply chains, this research empirically investigates the added value of using sell-through data originating from intermediaries, next to historical demand figures, in a multi-echelon supply chain. In a case study concerning a US drug manufacturer, we evaluate different methods in order to incorporate this data. We consider both time series modeling and machine learning techniques to produce multi-step ahead weekly forecasts. The results show that the manufacturer can effectively improve its short-term forecast accuracy by integrating sell-through data into the forecasting process. This conclusion holds for all forecast horizons considered, though it is most pronounced for one-step ahead forecasts.Keywords: Bullwhip effect, Demand forecasting, Information sharing, Machine Learning, Sell-through data.

Data driven inventory control using causal demand forecasting

Presenter: Ruud Teunter
In theory, data-driven demand forecasting allows us to set inventory levels without specifying the type of demand process. However, to attain some target service level, the safety stock levels still need to incorporate all demand forecast uncertainties, and so also uncertainty about the type of demand process. In this presentation, we compare several ways of doing so, applying results from the field of stochastic optimization.

Forecasting Individual Electric Utility Customer Hourly Loads from AMI Data

Presenter: Bradley Lawson
Utilities across the world are rapidly adopting AMI (Automated Meter Infrastructure), often giving the company access to hourly load data for all customers for the first time. Prior to AMI, customer metering only captured a single consumption value each month. With staggered meter reading schedules, customer consumption data was difficult to use because it did not encompass a calendar month. With the massive amounts of data now produced by AMI, companies increasingly see value in forecasting individual customer hourly loads. Hourly customer load forecasts can be used to improve and maintain customer relationships with customized high bill alerts following severe weather. As Electric Vehicles (EVs) and Distributed Energy Resources (DER) continue to expand on the distribution system, customer level forecasting can provide valuable insights into load levels and load patterns on unmetered segments of the distribution feeders. The underlying problem in forecasting individual customer load is the continuing variability of the individual customer's schedule and the effect on their electrical consumption patterns. Simple examples are when the customer is on vacation/holiday in June last year but August this year, or when children leave for college. There are scores of smaller differences that add to create variability from one period to the next for an individual customer. Strategies for dealing with this underlying variability include forecasting aggregated customer loads where that is structurally appropriate, such as in forecasting total load on a distribution feeder. Another strategy is to forecast individual customers as a member of a customer segment. This paper

uses actual customer AMI data to investigate the relative accuracy of different forecast strategies; forecasting customers individually, aggregated or as a member of a customer segment. The temperature-based modelling approach used in each case follows the approach of Tao Hong's 2010 dissertation, "Short Term Electric Load Forecasting," at North Carolina State University.

Development of an end-use load forecasting model for peninsular Malaysia

Presenter: Mohd Azlan Uda Kanardin
Load forecast plays a central role in the planning and operation of the electric power system in Peninsular Malaysia. Previously, electricity demand is derived using traditional econometric techniques which are proven to be capable of capturing historical relationships between electricity demand and important drivers. However, in recent years, forecasts based on these methods alone have underperformed due to lacks of ability to capture new programmes or policies that are not reflected in historical data such as structural changes in how customers use electricity and implementation of energy efficiency (EE) initiative.

To overcome the above challenges, Single Buyer (SB) has embarked on a study to develop an end-use load forecasting model for Peninsular Malaysia. The LoadMAPTM modeling tool, which was developed by Applied Energy Group (US), is used to develop the end-use model. The process of developing the end-use model involves several steps. The first step involves developing an electricity usage profile for the main tariff sectors. For this reason, a commercial and industrial customer survey was conducted by Single Buyer in 2016. Using the surveys and TNB's billing data, the commercial sector was segmented into 11 building types, the industrial sector segmented into six industry segments and the domestic sector, the smallest of the three sectors, was modeled as a single segment. The second step is to develop comprehensive information on customer energy use in a recent historical year, called the "base year". For each customer segment included in the model, the model needs the following base year data: annual electricity use, market size, and energy intensity. The end-use model also requires forecast assumptions such as customer growth, trends in appliance saturations, energy-use patterns in new construction, assumptions about appliance efficiency and purchase decisions. The forecasted electricity reduction due to EE initiatives is modelled based on Government's National Energy Efficiency Plan (NEEAP), which specifies various EE initiatives such as the promotion of 5-star rated appliance, Minimum Energy Performance Standards (MEPS) and Energy Audit and Energy Management in Commercial Building and Industries. Base-year data development results in Retail segment being the largest electricity consumer in the commercial sector, followed by government and private offices. Cooling accounts for almost half of the electricity consumption in the commercial sector, followed by interior and exterior lighting at 26.0%. The top electricity consuming industries are iron and steel, petrochemical, electrical and electronics and cement. More than 40.0% of the electricity used by Malaysian manufacturers are for operating machinery (motors), followed by the process at 24.0%. In the domestic sector, electricity demand is driven by appliances at 26%, followed by cooking at 15% and cooling at 13%. As the population grows together with number of homes , demand for these services and appliances will also increase.

Daily Peak Load Forecasting with Mixed-frequency Input Data

Presenter: Tao Hong
Daily peak load forecasting is typically conducted using hourly load and temperature data. In this presentation, we will compare different methods that use load and temperature of different frequencies. The empirical study based on Global Energy Forecasting Competition 2012 data demonstrates that using daily load with weighted hourly temperatures performs reasonably well in ex post forecasting and turns out to be more robust than the models based on hourly load and temperature for ex ante forecasting.

Forecasting the Volatility of Asset Returns: The Informational Gains from Option Prices

Presenter: Wenying Yao
The Realized GARCH class of models is extended to include option prices to forecast the volatility of asset returns. As analytical expressions are not available to evaluate option prices in the presence of GARCH volatility specifications, the VIX is used to approximate option prices. This is formally achieved by deriving

an expression relating the integrated measure of volatility and the VIX where the parameters of the equation are subject to a set of cross-equation restrictions. The full model is characterized by a nonlinear system of three equations containing asset returns, RV and the VIX, which is estimated by maximum likelihood methods. The forecasting properties of the joint model are investigated by forecasting daily volatility on the S&P500 using daily and intra-day data from July 2001 to November 2017. For comparison a number of special cases are also considered including the GARCH and RV-GARCH models as well as the GARCH model augmented by the VIX but not RV. The forecasting results show that augmenting the GARCH model with the VIX generates superior forecasts for a broad range of forecast horizons. The inclusion of RV as well as the VIX, or RV without the VIX does not improve the forecasts.

Country Characteristics and Currency Return Predictability

Presenter: Ilias Filippou
In this paper, we evaluate the out-of-sample performance of a large number of country characteristics when both time-series and cross-sectional dimensions are considered. In order to achieve this goal, we deviate from ordinary or weighted least squares which are susceptible to model overfitting and employ machine learning techniques that allow for high-dimensional models. We find that the latter methods offer an improved description of the behaviour of exchange rate returns in contrast to traditional methods. The improvements in accuracy of measuring currency risk premia is indicated by the highly positive out-of-sample R-squares as well as the strong economic value offered to a U.S. investor with mean-variance preferences.

Sparse Macro Factors

Presenter: David Rapach
We use machine learning to estimate sparse principal components (PCs) for 120 monthly macro variables spanning 1960:02 to 2018:06 from the FRED-MD database. In contrast to conventional PCs, each of the sparse PCs is a sparse linear combination of the underlying macro variables, whose active weights allow for intuitive economic interpretation of the sparse PCs. Innovations to the conventional (sparse) PCs constitute a set of conventional (sparse) macro factors. Robust tests indicate that only one of ten conventional macro factors earns a significant risk premium, while three of the sparse macro factors (yields, housing, and optimism) do. A four-factor model comprised of the market factor and mimicking portfolio returns for the yields, housing, and optimism factors performs on par with or better than leading multifactor models from the literature in accounting for numerous anomalies in cross-sectional stock re-turns. The sparse macro factors also provide insight into the economic underpinnings of well-known factors from the literature, including size, value, momentum, investment, and profitability factors. Finally, we provide out-of-sample evidence supporting the relevance of the sparse macro factors.

How and when does Judgment improve the Accuracy for Macroeconomic Forecasts?

Presenter: Ana Galvao
This paper evaluates the impact of judgment on the accuracy of macroeconomic density forecasts of GDP growth and inflation. To isolate the role of judgment we exponentially tilt a range of statistical models' density forecasts towards the predictive moments provided by a set of professional forecasters in the UK, where judgmental adjustments are believed to play some role. In addition to point (mean) forecasts, we evaluate the role of judgment about the forecasted variance and skewness (or balance of risks). Our set of forecasters include the Bank of England, the National Institute of Economic and Social Research and two surveys of professional forecasters from Consensus Economics and the Bank of England's Survey of External Forecasters. We consider a range of statistical models. This is to acknowledge that conclusions about the role of judgment inevitably depend on relative to 'what'. Since, in the UK, we neither know what statistical or macroeconometric model(s) is (are) used by the professional forecasters nor what judgment they may then apply to these, we consider the density forecasts from a range of models. We consider workhorse models of the type likely used by professional forecasters, namely autoregressive models, Bayesian VAR models and a density forecast combination of many statistical models (as published by the Warwick Business School Forecasting

System). Our empirical results indicate that judgment does tend to improve upon the point forecasts from statistical models, especially at short horizon forecasts and at times of hightened macroeconomic uncertainty. But, in contrast, judgment about higher moments and future macroeconomic uncertainty systematically takes value away from statistical density forecasts.

Tracking inattention: A new expectations-based measure

Presenter: Nathan Goldstein
Information heterogeneity and inattention play a key-role in the expectations formation process, according to a prominent strand of literature. This study proposes a new measure of inattention to information, which exploits data of survey forecasts at the individual level. This measure provides direct estimates of the inattention parameter according to leading models of expectations formation, which highly improve the precision of available estimates that are based on aggregate expectations. The proposed methodology could also be used to differentiate between the informational approach and alternative approaches of modelling the expectations formation. In an application to the US Survey of Professional Forecasters, the new measure sheds light on variations in inattention over time and across different macroeconomic variables.

Why can't professional macroeconomic forecasters predict recessions?

Presenter: Víctor López-Pérez
The professional forecasters' inability to anticipate macroeconomic recessions is well documented. The literature has found that aggregate or consensus forecasts are too optimistic before downturns and too pessimistic before recoveries. This paper explores whether this result also holds at the individual level or is the result of an aggregation bias. Using a Spanish survey of professional forecasters conducted by Funcas, I find that individual forecasters are indeed too optimistic before recessions. The reason is not that they become inattentive when the economy is in good shape. Instead, they put too much weight on the most recent developments when producing their forecasts. The analysis of their forecast revisions reveals that better-than-expected data makes some forecasters to revise their forecasts upwards too much. These revisions raise the consensus forecast and trigger herding by other forecasters, who also revise up their forecasts. Both factors lead to subsequent negative forecast errors, especially when a recession occurs. Consequently, professional forecasters could improve their forecasting performance by reacting less aggressively to the latest data releases and by avoiding inefficient herding.

Co-err with Thy Neighbour: Understanding the Covariance Matrix in Hierarchical Forecasting

Presenter: Kandrika Pritularga
Keyword: hierarchical forecasting, reconciliation, covariance matrix, generalised least squares

Regularized regression for hierarchical forecasting without unbiasdness conditions

Presenter: Souhaib Ben Taieb
Forecasting a large set of time series with hierarchical aggregation constraints is a central problem for many organizations. However, it is particularly challenging to forecast these hierarchical structures. In fact, it requires not only good forecast accuracy at each level of the hierarchy, but also the coherency between different levels, i.e. the forecasts should satisfy the hierarchical aggregation constraints. Given some incoherent base forecasts, the state-of-the-art methods compute revised forecasts based on forecast combination so that the aggregation constraints are satisfied. However, these methods assume the base forecasts are unbiased and constrain the revised forecasts to also be unbiased. We propose a new forecasting method which relaxes these unbiasdness conditions, and seeks the revised forecasts with the best tradeoff between bias and forecast variance. We also present a regularization method which allows us to deal with high-dimensional hierarchies, and provide its theoretical justification. Finally, we compare our proposed method with the state-of-the-art

methods both theoretically and empirically. The results on both simulated and real-world data indicate that our methods provide significant improvement in forecast accuracy.

Probabilistic forecast reconciliation

Presenter: Anastasios Panagiotelis
Forecast reconciliation involves adjusting forecasts to ensure coherence with aggregation constraints. We extend this concept from point forecasts to probabilistic forecasts using a geometric interpretation. A new theorem is proved that the popular log score is an improper scoring rule with respect to the class of incoherent forecasts when the true predictive distribution coheres with aggregation constraints. Probabilistic forecast reconciliation is then discussed in practic including methods that exploit ideas from cross-validation.

Forecasting scores in games with multiple scoring modes

Presenter: Rose Baker
In some sports there are several ways of scoring, which accrue differing numbers of points. Such sports include National Football League (NFL) games, rugby, rugby league, basketball, Australian football, Gaelic football, and hurling.Forecasting exact scores cannot then be done using the usual Bradley-Terry or Poisson models, and to obtain the probability of the final scorefrom a model,one has to sum probabilities over the vast number of ways in which that particular final score could be arrived at.An approach is described that generalizes that of Baker and McHale, in 'forecasting exact scores in NFL games', International Journal of Forecasting 29 (2013), 122-130.In that work there is a hazard function for each type of score by either team. The hazard at a time t in the game can be a function of covariates such as past team performance and bookies' odds,and of events such as total points accrued by either side that have occurred in the current game to time t. The state of the game is assumed to evolve according to a Markov process in continuous time, and is a bivariate birth process.It is desired to compute the likelihood function and hence to estimate unknown model parameters, and then use the model to forecast.The starting point is the Chapman-Kolmogorov forward differential equation.From this, the likelihood can easily be written down as a formal solution involving the exponential of the generator matrix.The substantial difficulty with this approach is in doing the extensive computations required.Analytic solutions, if they could be found using an algebra package, are difficult to evaluate because of numerator and denominator sometimes both going to zero, and simulation is easy to program but very inefficient when there is a vast range of possible final scores, so thatany particular final score obtained occurs with very low probability.The computations for the Markov method can be greatly speeded up using three 'tricks'. One is the stochastic-process technique of uniformization, that greatly speeds up the summation of a series expansion of the matrix exponential, and another isexploiting the sparsity of the generator matrix, so that many multiplications need not be carried out. Finally, terms that cannot lead to the known final state need not be computed.The results from the 2013 NFL analysis are briefly reviewed and explained with some examples, and some generalizations of the methodology discussed.

Informational efficiency and price reactions in exchange betting markets

Presenter: Luca De Angelis
We investigate the degree of efficiency of exchange betting markets. In particular, we test whether prices on the exchange market are set efficiently both before the beginning of a football match and in the aftermath of an event occurred during the match. First, we propose a forecast-based approach for formally testing the efficiency of pre-match prices on the exchange betting markets. Then, using event study analysis on high-frequency data, we examine the reaction of prices to events and the arrival of major news. In particular, we analyse the post-event evolution of in-play odds and their dynamic behaviour at different horizons. We test for informational efficiency in football exchange betting markets in three different directions: (i) we investigate whether there is evidence of mispricing in in-play odds as reaction to major news events (i.e. goals, red cards), (ii) we test whether the arrival of new information on the market creates systematic bias which can be exploited to set a profitable betting strategy, (iii) we focus on possible over/under-reaction by investigating the main drivers which may create deviations from efficiency. To do so, we consider a large dataset comprising prices collected every ten seconds from Betfair Exchange for all the English Premiership matches played in

the five seasons from 2009/10 to 2013/14. Our main findings show the presence of a reverse favourite-longshot bias in both pre-match and in-play exchange market odds and that this bias is not fully absorbed over time, creating market inefficiency. In particular, as our event study analysis shows that anticipated events (e.g. goal scored by favourites) are underpriced while surprise events (e.g. goal scored by underdogs) are overpriced, we find evidence that overall exchange market participants appear to be less risk averse than fixed-odds bookmakers.

Forecasting netball scores

Presenter: Phil Scarf
We model the scores in Netball matches. Our model borrows some parts from the Poisson match—the notion of attacking and defensive strengths—and some parts are new—three sources of randomness: the total score; and the conversion of possession into goals by each team. In this way, the structure of netball leads to what we think is an interesting model. We estimate model parameters using maximum likelihood and for a UK tournament, the Vitality Superleague, which has been in-play for about a decade. Match data are somewhat sparse in parts, as new teams have entered the league and others have left. Therefore, our estimation method has an element of shrinkage. Strength varies over time so we use a discounted likelihood. We assess forecast performance for the 2019 season, which is currently in progress. Other uses of the model—assessing competitiveness, tactical questions—are briefly mentioned.

Application of data compression techniques to time series forecasting

Presenter: Boris Ryabko
It is known in Information Theory that the problems of data compression and time series forecasting are closely related. More precisely, in the case of stationary processes, a data compressor which gives an asymptotically minimal length of a compressed sequence can be used as a time-series predictor whose prediction error (per symbol) is asymptotically minimal, see details in the book (Ryabko, Astola, Malyutov, Compression-Based Methods of Statistical Analysis and Prediction of Time Series. Springer, 2016). In this study we show that standard well-known file compression programs (zlib, bzip2, etc.) are able to forecast real-world time series data well. The strength of our approach is its ability to use a set of data compression algorithms and "automatically" select the best one of them during the process of forecasting. Besides, modern data-compressors are able to find many kinds of latent regularities using some methods of artificial intelligence (for example, some data-compressors are based on finding the smallest formal grammar that describes the time series). Thus, our approach makes it possible to apply some particular methods of artificial intelligence to time-series forecasting. As examples of the application of the proposed method, we made forecasts for the monthly T-index and the Kp-index time series using standard compressors. In both cases, we used the Mean Absolute Error (MAE) as an accuracy measure. For the monthly T-index time series, we made 18 forecasts beyond the available data for each month since January 2011 to July 2017. We show that, in comparison with the forecasts made by the Australian Bureau of Meteorology, our method more accurately predicts one value ahead. The Kp-index time series consists of 3-hour values ranging from 0 to 9. For each day from February 4, 2018 to March 28, 2018, we made forecasts for 24 values ahead. We compared our forecasts with the forecasts made by the Space Weather Prediction Center (SWPC). The results showed that the accuracy of our method is similar to the accuracy of the SWPC's method. As in the previous case, we also obtained more accurate one-step forecasts.

R Rlgt package - Bayesian Extensions of Exponential Smoothing models

Presenter: Slawek Smyl
Rlgt is an R package that implements a number of time series forecasting models that are Bayesian generalizations and extensions of some Exponential Smoothing models. The main differences/additions include 1) nonlinear global trend, 2) Student-t error distribution3) a function for the error size, so heteroscedasticity. The methods are particularly useful for short time series. When tested on the well-known M3 dataset, they are able to outperform all classical time series algorithms. The models are fitted with MCMC using the 'rstan' package.

Component Forecasting Using Singular Spectrum Analysis

Presenter: Don Poskitt

It is well known from Wold's decomposition that the spectrum of a stochastic process can be separated into discrete and continuous components that correspond to so called singular and regular components of an observed series that are, respectively, deterministic (ultimately perfectly predictable) and purely non-deterministic (ultimately unpredictable). Under very general conditions concerning the structure of an observed series, Singular Spectrum Analysis (SSA) constructed using re-scaled trajectory matrices (RT-SSA) can be shown to reproduce the spectral features of an underlying data generating process. This motivates a forecasting methodology based on an application of step-wise RT-SSA in which the unobserved singular and regular components of Wold's decomposition of a series are forecast separately and then combined to forecast the series itself. The operation of this methodology and associated theoretical results are illustrated via numerical examples involving trend stationary and difference stationary processes, a random walk with drift, and an empirical example.

One model or many? Exchange rates determinants and their predicitve capabilities.

Presenter: Piotr Dybka

In this article, we focus on the identification of potential determinants of exchange rate movements and their application to exchange rate forecasting. Economic literature offers several different theories regarding the exchange rate movements, that include, among others, Purchasing Power Parity (PPP), Taylor rule fundamentals or the Behavioral equilibrium exchange rate (BEER) models. First, based on the extensive literature review, we create a list of potential exchange rate determinants. Second, we apply the Bayesian Model Averaging (BMA) techniques to evaluate the posterior inclusion probabilities of each of the analysed variables. This way we can establish an optimal variable composition of the exchange rate model and verify its predictive capabilities. We compare the predictive performance of a combined model based on the BMA results with standard models used in the literature and two additional benchmarks: the half-life Purchasing Power Parity model and a random walk. We examine model performance at various forecast horizons using differing metrics (mean squared error, the direction of change). BMA-combined model outperformed the standard models, however its forecasts provide better accuracy than the benchmarks only in the longer term. Interestingly, the BMA-combined model performs better than the model re-estimated only for the variables with the highest Posterior Inclusion Probability.

Application of Google Trends Data in Exchange Rate Prediction

Presenter: Fumiko Takeda

Exchange rate forecasting has always been considered a difficult task, due to the complexity of the factors affecting FX rates. Meese and Rogoff (1983) claim that the random walk model performs just as well as any structural model, implying that any effort to predict exchange rates is useless. Since then many studies have tried to prove this claim wrong (Mark 1995; Rossi 2005; Hashimoto 2011), with recent studies using online data such as search frequencies from Google Trends (Chojnowski and Dybka 2017; Bulut 2017) to try and create a useful model. In this paper, we explore the possibilities of applying Google Trends to exchange rate forecasting. Specifically, we use Google Trends to try and capture market sentiment in Japan and the United States, and construct a sentiment index similar to the FEARS index of Da et al. (2014). We produce forecasts of one-month ahead USD/JPY rates using three structural (PPP, Interest Rate Parity, and Monetary) and two auto-regressive models (AR[1] and AR[2]), and see if our sentiment index has the ability to improve the predictive power of these models. We evaluate the forecasts with the MSPE (mean squared prediction error), Clark and Wests' (2007) test for equal predictive accuracy, and a DOC (direction of change) test. The data we use is from Jan 2004 to Aug 2018, treating Jan 2004-Feb 2011 as a training sample, and March 2011-Aug 2018 as a forecast sample. We contribute to existing literature in two ways. First, we create a unique sentiment index for Japan using a Japanese dictionary, to capture sentiment in Japan as accurately as possible. Second, we show that the Google Trends sentiment index we propose indeed has the potential to improve one-month ahead forecasts of log USD/JPY returns. Overall, our best results are seen with the AR(1) and AR(2). These base models show a decrease in MSPE of 6.09% and 6.99% compared to the random

walk, and the addition of the sentiment index improves these numbers to 6.38% and 8.65%, respectively. Furthermore, in the former model, its ability to predict directional changes of exchange rate returns increases by a statistically significant amount. We also test sentiment indices of different word numbers (10, 20, 25, 30, 35, 40) and find that the 25 and 30-word index perform best; especially, the 30-word index improves all models tested in this study to a certain degree.

The impact of expert advice on forecasts based on presumed credibility, identification, and the media source of conveying the advice.

Presenter: Sun Rui
On the wide literature in expert knowledge elicitation for forecasting, presumed credibility of experts has been consider as a factor that could possible influence decision making of novices. This study attempts to contribute further by examining the impact of presumed credibility on forecasts when different levels of information sharing relevant to presumed credibility are becoming available to the advisees. Comparing the impact of plain status of an expert, the actual identity of the expert or the media source that the advice of the expert is conveyed may bring a different weighting by the novice's side. The main goal of the paper is to examine degrees of presumed credibility under different degrees of information sharing and media sources.

Demand Forecasting for Retail Pricing

Presenter: Brian Seaman
Building trust with customers is key to the long-term success of a company and price is an important driver of that trust. The philosophy of Every Day Low Prices (EDLP) is how Walmart maintains that fairness around price that customers have grown to trust. In order to execute this strategy, there are several areas where the importance of accurate forecasts are highlighted including future demand and sales, the impact of past and future prices, and market responses. This presentation will focus on the challenges and opportunities of these forecasting requirements.

Design, implementation and operation of a large-scale, mission-critical sales forecasting system

Presenter: Phillip M. Yelland
In light of Parkinson's famous dictum, it's perhaps unsurprising that as the computing power available to organizations increases, their ambitions expand correspondingly. Commercial forecasting systems are no exception to this trend: As a matter of course, a modern retail sales forecasting system is expected to produce forecasts for hundreds of thousands of SKUs, often sold in hundreds of physical locations and through multiple on-line channels. Furthermore, these forecasts are often updated daily (if not more frequently), and must be produced reliably, with a minimum of manual intervention. Finally, since sales forecasts constitute a vital input to a wide array of business processes, provision must be most for automated monitoring of forecast performance, with alerts dispatched as appropriate, and explanations of forecasts should be available on demand. In this presentation, we will discuss the design, implementation and operation of such a sales forecasting system at the major US retailer Target. We will highlight the ways in which we've sought to address the requirements outlined above, the challenges we've faced in doing so, and steps we've taken to overcome those challenges.

From models to data: Forecasting through similarity

Presenter: Evangelos Spiliotis
The ability to make predictions for the future is of utmost importance in all areas of science. Especially in areas such as industry, finance and medicine, it is vital that data are modeled and transformed into forecasts for the future. There are many methods to make predictions for the future. One of the most popular and simple methods used in predicting the future is the simple exponential smoothing method. However, despite its success and widespread use in many areas, simple exponential smoothing method has some shortcomings that negatively affect the accuracy of forecasts made using this method. Therefore, ATA-simple has emerged

as a new method of statistical estimation alternative to simple exponential smoothing. ATA-simple can be used to model times series that do not show any trend or seasonality behavior. ATA-simple model is very similar in mathematical form to simple exponential smoothing but the smoothing constant of ATA-simple is parameterized so that it depends both on the sample size at hand and a smoothing parameter that is optimized on a discrete space. This makes the weighting scheme of ATA-simple dynamic so that the weights are distributed amongst the observed values depending on where in time the smoothing is carried out. Also because the optimization is limited to discrete values, the process is much faster and easier. Another positive difference that ATA-simple brings to the table is that it does not require an initial value. The two models are applied to the M3 and M4 competition data sets simultaneously and their forecasting accuracies are compared. The method performed better than simple exponential smoothing for both data sets. The detailed results will be discussed in this talk in addition to the model's properties and the advantages that the model brings to the forecasting literature.

From models to pictures: Forecasting through automatically generated image features

Presenter: Artemios-Anargyros Semenoglou
The development of accurate generalized methods and frameworks suitable for time series forecasting has always been the aim for researchers and practitioners in the field. This is particularly true when dealing with large and diverse datasets in terms of domain, frequency, length and characteristics. Over the last years, many novel approaches to this problem have been suggested, the most promising ones involving concepts such as feature extraction, cross and meta learning. This trend in the research can be attributed mainly to two factors, data availability and advances in the field of Machine Learning (ML) algorithms, both of them enabling the identification of useful features and common, possibly non-linear, patterns found in the series. The value of such cross learning approaches became evident in the M4 Competition in which the two top performing methods exploited the large dataset provided through advanced ML algorithms. Based on the aforementioned observations, the purpose of this paper is to propose an innovative alternative to the matter. The suggested framework, inspired by the success of Artificial Neural Networks (ANNs) in the area of digital image processing, transforms the series into graphic objects which can be then classified more effectively through feature extraction techniques without making any assumptions on the characteristics of the series and the way these are measured. A simple ANN can then be used for extrapolating the series. The main novelty of this approach lies at the forecast generation procedure: each series is forecasted by observing the patterns of similar ones without hypothesizing its underlying data generation process. The performance of the proposed method is evaluated using the data of the M3 and M4 Competition datasets, for testing and training the framework, respectively.

Deep Learning (DL) for time series forecasting: An evaluation and a comparison with other ML/NN algorithms.

Presenter: Spyros Makridakis
Deep learning is used widely to solve problems in games, fraud detection, computer vision, natural language processing, automatic translation, robotics, autonomous vehicles and a host of other applications. Additionally, it has been proposed as the means to revolutionize time series forecasting. It is the purpose of this paper to empirically test such an assertion by evaluating its performance in comparison to other ML and NN methods. This will be done by first using various forms of DL to forecast the 1045 monthly time series of the M3 Competition. Consequently, such results will be compared with those available to the authors in a previous published paper in PLOS ONE. In addition, to supplement the study and avoid the criticism that DL and ML/NN methods required much longer series, the 1045 lengthier monthly series from the M4 Competition will be selected to determine how longer number of observations will affect the accuracy of DL and ML/NN methods on their own as well as in comparison to traditional statistical methods.

Forecasting AI in Forecasting

Presenter: Lawrence Vanston
Artificial Intelligence, especially machine learning, has been applied in forecasting for many years. Recent progress has been significant and raises the question of AI's role in the future of forecasting. This paper provides a platform for addressing that question. We take a future-oriented view, examining the drivers and constraints for the continued adoption of AI. We also discuss how further progress in AI might or might not change the balance. Finally, we discuss the question of whether AI could substantially substitute for traditional forecasting methods, and, if so, what that means for the forecasting profession.

Forecasting algorithm assignment to distribution grid service points in the context of demand response

Presenter: Jonathan Farland
Demand side management programs such as demand response and energy efficiency have been evolving across US utility companies for decades. Demand response (DR) in particular has been widely used not only to manage loads during times of system peaks but has also been shown to provide locational benefits to distribution grid operators. Providing accurate ex-ante forecasts and ex-post measurements of load reductions from DR program participation is extremely challenging especially at more granular levels of aggregation. This is mostly due to the stochastic nature of meter-level historical time series, and also the heterogeneity across program participants. This presentation describes one approach used to assign algorithms to distribution grid service points based on historical forecasting performance, customer level characteristics and time series attributes. This approach is currently available in TROVE's Demand Response solver within the TROVE data science platform.

Modelling Uncertainty: Probabilistic Load Forecasting Using Weather Ensemble Predictions

Presenter: Nicole Ludwig
Predicting electricity demand is challenging. Although demand patterns are regular to some extent, they also exhibit uncertainty and depend on the weather, which itself is a highly stochastic system. A vast majority of the recent probabilistic forecasting literature addresses how to incorporate uncertainty into the output of a load forecasting model. However, there are still only a few papers that consider accommodating uncertainty in the input variables, especially uncertainty in the weather forecasts. In this paper, we generate probabilistic forecasts for the national demand of electricity in Great Britain using weather ensemble predictions to introduce uncertainty in the input variables. For this purpose, we adopt two different modelling approaches. The first is a linear regression model that serves as a benchmark, while the second is a quantile random forest that can also capture nonlinear relationships between demand and weather ensembles in a nonparametric framework. The load data was obtained from the National Grid, whereas the weather data comes from the European Centre for Medium-Range Weather Forecasts. More specifically, the weather data we use to capture the uncertainty, are perturbed (ensemble) forecasts of temperature, wind speed, cloud cover and precipitation. We forecast the demand at different times during each day. For each of these times, models are estimated separately, with features included to capture special day effects along with seasonality. We evaluate the out-of-sample accuracy of point forecasts, quantile forecasts and distributional forecasts using the MAPE, quantile score, and CRPS, respectively, on a forecasting horizon from one day to seven days ahead.

Understanding the Impacts of Distributed PV Resources on Short-Term Load Forecasting – A Comparative Study on Solar Data Availability

Presenter: Adam Wigington
Distributed PV is expected to play an increasingly significant role in the energy production of many regions over the coming decade, due to federal or state regulations and/or incentives, lowering installation cost and other factors. Recent experience and projections of future scenarios have shown that behind-the-meter distributed PV can result in increased uncertainty as to the load that operators will observe and need to

dispatch generation against. While tools to forecast hour-ahead to days-ahead loads have been commercialized for decades and models are fitted well for traditional demand, adding PV can result in significant uncertainties. Understanding how load forecasting should consider distributed PV will require new PV-focused data, such as irradiance and insolation measurements, and potentially new techniques. Challenges arise when system operators and those who produce forecasts do not have 1) knowledge of PV production in real time or after the fact, 2) information about irradiance or other meteorological influences on PV production at various points on the network, and 3) the design specifications of the PV installations. The relative significance of these challenges will vary across geographical regions due to their different PV deployment trends and penetrations, the relative size of the system operator's territory, how the system operator is organized, and the methodologies used to generate their load forecasts. This paper will present on a current Electric Power Research Institute (EPRI) project[1] that is working to study these impacts across multiple utilities and ISOs, and to provide insights to what data – specific to distributed PV resources – provide the most benefit to short-term load forecasting. In the current phase of the project, we are studying geospatial distributions of irradiance and weather data across multiple utility and ISO territories, as well as Numerical Weather Prediction (NWP) models, to help provide guidance on sensor deployment and data infrastructure investments for utilities. We will present interim results for the project that benchmark performance on typical forecast models used in industry – multiple linear regression and Artificial Neural Networks (ANNs). The results will include comparisons of bottom-up vs top-down approaches for generating PV and load forecasts, and analyze different methodologies for incorporating the PV into the load forecasts.
[1] https://www.epri.com/#/pages/product/000000003002008444/

Access forecasting for safety-critical crew transfers in offshore environments

Presenter: Jethro Browell
Cost reduction in the construction and operational phase of offshore assets has the potential to significantly reduce OPEX in both the oil & gas and offshore wind industries. Improved access in tandem with improved prognostic capability from condition monitoring has the potential to realise significant savings. However, the pressure to achieve increased access to offshore structures implies operating on a greater number of marginal-weather days, which potentially carry a greater safety risk. This raises a number of challenges to both the forecaster, who must ensure safety-critical information is communicated effectively, and the forecast user who must interpret a range of complex information before making a decision.Here a new access forecasting and decision-support tool and underpinning analytical methods developed by the ORACLES project (Offshore Renewables Access, Loss Estimation & Safety) are presented with the aim of achieving low cost operation while maintaining very high levels of safety. By utilising statistical methods for probabilistic forecasting and leveraging crew transfer vessel telemetry data, the ORACLES system predicts characteristics vessel motion during crew transfer from vessel to structure in order to provide marine planners with an assessment of weather windows in terms of "transfer quality" for day-ahead and intraday planning. This information is presented alongside longer term forecasts of accessibility and energy yield in order to assess the value of potential weather windows. This system has been developed in close collaboration with vessel and wind farm operators, and a selection of visualisations is proposed for different user preferences.

Long-Term Economic Forecasting with Structured Analogies and Interaction Groups

Presenter: Konstantia Litsiou
In this study, we employ judgmental forecasting techniques, Structured Analogies and Interaction groups for long-term forecasting. The aim of the paper is not to evaluate forecasting accuracy per se but to highlight the potential of such techniques in this so complex and challenging task. The case study is about Saudi Arabia and its aim to adopt a diversification strategy to reduce its dependency on the oil sector, where oil revenue consists 90% of its budget currently. The study has four phases: Unaided Judgment, Structured Analogies, and Interaction Groups with Structured Analogies - all three using disguised data – before finally working on the undisguised case study under review over a significant amount of time. Adopting judgmental methods are attributed to three main reasons: in an attempt to derive long-term economic forecasts about Saudi Arabia's ability to diversify its investments, to discover the impact of different factors on financial and

economic outlooks, and to explore the main reasons for deviating the accuracy of financial and economic forecasts.

Contrast effects in judgmental forecasting when assessing the implications of worst- and best-case scenarios

Presenter: Paul Goodwin
Two experiments investigated whether individuals' forecasts of the demand for products and a stock market index assuming a best or worst-case scenario depend on whether they have seen a single scenario in isolation or whether they have also seen a second scenario presenting an opposing view of the future. Normatively, scenarios should be regarded as belonging to different plausible future worlds so that the judged implications of one scenario should not be affected when other scenarios are available. However, the results provided evidence of contrast effects in that the presentation of a second 'opposite' scenario led to more extreme forecasts consistent with the polarity of the original scenario. In addition, people were more confident about their forecasts based on a given scenario when two opposing scenarios were available. We examine the implications of our findings for the elicitation of point forecasts and judgmental prediction intervals and the biases that are often associated with them.

The effects of scenarios on judgmental demand forecasts and the subsequent production decisions

Presenter: Sinan Gonul
In production planning and inventory management activities, demand forecasts play a crucial role by shaping the decisions of how many to produce from different products subject to a fixed total capacity restriction. Such demand predictions can be supported by the presence of optimistic and pessimistic scenarios that provide a vibrant narrative to envisage the future. In this paper, through an experimental study, we are investigating i) the effects of providing only time-series information on the generation of judgmental demand forecasts and the subsequent production order decisions and ii) the interaction of presenting optimistic and pessimistic scenarios with this forecast generation and ordering process. The experimental design involved three between subject groups: i) only time-series information, no scenarios ii) time series information accompanied with weak optimistic and pessimistic scenarios iii) time series information accompanied with strong optimistic and pessimistic scenarios. Findings are disclosed and directions for future extensions are suggested.

Decomposition of contextual information for forecast adjustments

Presenter: Anna Sroginis
When making judgmental adjustments to statistical forecasts, forecasters rely on their knowledge based on additional information (e.g. promotional/marketing activities, weather, holidays, strikes) that is unavailable to the statistical model. Most Forecasting Support Systems (FSS) do not record or structure any qualitative statements, and the whole process of final forecast elicitation is usually difficult to track and evaluate. Hence, we aim to find a more structured method to use domain knowledge in order to improve accuracy. One possible way is to decompose this information into relevant categories and subtasks, which are suggested to be easier to process. However, there has been limited research on the value of decomposition in judgmental forecasting, specifically when applied to contextual information. This research investigates how the decomposition of contextual information can influence judgmental adjustments. Previous research has found that the resulting adjustment from the same contextual information differs depending on the factors included in the statistical model. We build on this finding, exploring how decomposition of contextual information affects adjustments of forecasts originating from simple baseline and statistical promotional models in a retailing setting. We conduct laboratory-based experiments, simulating an FSS interface. We provide participants with a graphical representation of past sales, statistical forecasts and markers for promotional periods. In addition, various qualitative details with or without any predictive power are presented. The latter is provided either in a step-by-step fashion, decomposing the information by source, or simultaneously, allowing us to contrast the effects of different forms of decomposition and leading to useful conclusions for forecasting practice and the design of FFS.

Forecasting mortality rates with time series models

Presenter: Farshid Vahid

With populations ageing in almost all countries in the world, the need for reliable mortality forecasts has become more important than ever. In the voluminous literature on mortality modelling and forecasting, the use of data driven age specific univariate time series models is often summarily dismissed on the basis that long-term point forecasts from such models may imply an unreasonable age structure of mortality rates. As a result, the literature has focused on models that postulate a single (or a small number of) unobserved common factor(s) driving the dynamics of mortality rates for all ages, and various adjustments have been suggested to make these models more compatible with the data. In this paper we argue that there is a vast array of possibilities between age specific univariate time series models and a restricted multivariate model that assumes a single common unobserved factor for all ages. Specifically, we study multivariate time series models with error correcting mechanisms that ensure that long term forecasts of age-specific mortality rates follow a locally common trend. We show that such models perform at least as well as the existing models. The real advantage of our models is that they provide prediction intervals quite readily, whereas providing such intervals for the existing single index models that are often estimated in two steps and/or involve ad-hoc modifications such as jump-off corrections is not straightforward.

Verification and prediction of the effect of healthcare control policy.

Presenter: SangKyun Jo

The Korean medical authorities have recently introduced new healthcare policy that imposes extra burden on patients who visit a hospital due to mild diseases. Although the government intends to enhance the efficiency of the overall healthcare system by strengthening the gatekeeping role of the primary medical care, the effectiveness of the policy has not been clearly explained in detail. In this paper, we verify the effect of the policy through statistical techniques such as difference in difference and regression discontinuity design. In most cases, we find statistically significant policy effect that suppress the patients' choice of secondary medical care. Examining different demographic subgroup, we also find that the effect of policy tends to be highest in the middle-income group rather than the lower income group, which can be interpreted as information gap among income brackets. Based on the statistical models, we predict the effect of the newly expanded health care control policy. It shows that the economically vulnerable classes could be highly affected by the policy.

The effects of health shocks on hospital visit rates and healthcare spending

Presenter: Jinhwan Park

Middle East respiratory syndrome coronavirus (MERS-CoV) is a highly lethal respiratory disease that requires hospital admission. Although most cases of MERS have occurred in Saudi Arabia and the United Arab Emirates, cases have been reported in South Korea in people who traveled from the Middle East or their contacts. In this paper, we analyze the change in the number of hospital visits and healthcare spending of patients with mild disease using difference in difference (DID) method. We also compare the effects of MERS and H1N1 influenza outbreaks on hospital visits and healthcare spending. We find that the number of hospital visits and healthcare spending decreased when MERS occurred. Possible reason is that patients may be reluctant in visiting hospitals due to incidence of infection. By contrast, the number of hospital visits and healthcare spending increased when H1N1 influenza occurred. Patients tend to receive treatment as soon as they find early symptoms of H1N1 to prevent from getting worse. Our results highlight the need for forecasting the behavior of patients under unexpected health shocks, and provide evidence that people behave differently depending on the characteristics of infectious disease.

Changes in Stock Returns Predictability using Artificial Neural Networks

Presenter: Adam Chudziak

Scientific predictions in financial markets are commonly based on theoretical finance foundation. Nonetheless, many techniques used by practitioners do not come from theoretical considerations. Successful trading strategies based on Artificial Neural Networks (ANN) have been reported and are used by leading hedge funds. Although many ANN methods still used today, such as the multi-layer perceptron, have origins in

1950s and 1960s, the interest in them was rather small for almost half a century. Recently more results on stock predictability using the ANN are published. However, they are usually constrained to specific time periods. This study investigates the changes in the US stock prices predictability using Artificial Neural Networks since the 1960s. We show that over this time there were periods when using past market data, even simple Artificial Neural Network methods could, save for computational hardware limitations, exploit market inefficiencies.

Maximum likelihood estimation of the Hull-White interest rate model for pricing and forecasting

Presenter: Kamil Kladivko

We develop a maximum likelihood approach to rigorous parameter estimationof the Hull-White interest rate model from a time series of yield curves.The Hull-White model, although originally designed for pricing interest rate derivatives, has gained substantial popularity in portfolio and risk management.The common practice is to calibrate the model parameters to yield curve data of a single date. However, this parameter identification only allows for a risk-neutral evolution of the yield curve, which is inappropriate for computations of risk-related measures, especially over longer time horizons. Our estimation strategy identifies the model parameters under the both risk-neutral and statistical probability measure, and therefore facilitates the use of the model not only for valuation of interest rate contingent claims, but also for yield curve forecasting, and thus portfolio and risk management. We formulate the likelihood function so that the one-step ahead forecasting error of the yield curve is minimised while the perfect fit of the current yield curve is preserved. The short-rate, which is the sole yield curve factor, is modelled as a latent process, and it is filtered out from yield curve data. In addition to the short-rate, we only estimate three parameters, namely the mean-reversion speed, volatility, and market price of interest rate risk. Our estimation methodology provides robust parameter estimates and encouraging out-of-sample forecasting results across yield curve data sets from various time periods and currencies. Indeed, the out-of-sample forecasting performance of our Hull-White model implementation is comparable to the performance of the popular Diebold and Li model [Diebold, F.X. and Li, C. (2006). Forecasting the term structure of government bond yields. Journal of econometrics, 130(2), 337–364.] In contrast to the Diebold and Li model, the Hull-White model is arbitrage-free and fits the yield curve perfectly. Our approach unifies pricing and forecasting of interest rates in a parsimonious, robust and highly functional way.

Economic uncertainty and stock market volatility prediction

Presenter: Vasiliki Skintzi

We investigate the question of whether economic uncertainty measures contain information about future stock volatility beyond that contained in common predictors such as past volatility, measures based on option-implied information or high-frequency data, macroeconomic fundamentals and sentiment indicators. There is ample theoretical and empirical evidence that economic uncertainty affects stock returns and volatility. Veronesi (1999), Bollerslev et al (2009) and Pastor and Veronesi (2012) provide a theoretical framework for the link between economic uncertainty and stock market volatility.Moreover, Pastor and Veronesi (2012) find that individual US stock returns are more volatile and that pairwise US stock returns correlations are higher when economic uncertainty is higher. Liu and Zhang (2015) show that incorporating economic uncertainty into volatility prediction models significantly improves forecasting performance. Asgharian et al (2018) use an economic uncertainty index to forecast stock market correlations and conclude that incorporating economic uncertainty into forecasting US-UK stock market correlations improves out-of-sample forecasting performance.Following the importance of the impact of economic uncertainty on various economic and financial variables, an active area of research is proposing alternative measures of economic uncertainty. A group of studies (e.g. Baker et al, 2016) have attempted to measure uncertainty by quantifying the newspaper coverage of uncertainty. The proposed uncertainty indices are based on the frequency of newspaper articles that mention specific word or phrases that relate to various concepts of economic or financial uncertainty. Another strand of the literature propose uncertainty measures based on econometric analyses of macroeconomic time-series data and professional forecasts. The econometric based uncertainty measures aggregate the variation of objective or subjective forecast errors over a number of macroeconomic or financial variables.

Based on previous research, we expect that economic uncertainty measures are of importance for stock market volatility. We draw upon several measures of economic uncertainty, and examine the impact of economic uncertainty shocks on long run stock market volatility. We apply the class of GARCH-MIDAS models that has been proven useful for analyzing the impact of the macroeconomic environment on financial volatility. Our study makes a number of important contributions. First, we link various measures of economic uncertainty to US stock market volatility based on a MIDAS-GARCH model and examine both the in-sample estimation fit and the out-of-sample forecasting performance. Second, the forecasting performance of pooled economic uncertainty measures based on a number of combination methods is also examined. Third, we combine economic uncertainty measures with traditional predictors of financial market volatility in the MIDAS-GARCH models to examine whether economic uncertainty incorporates any additional information in forecasting future volatility. Forth, as a robustness check we study the predictability of financial volatility by economic uncertainty measures using standard predictive regressions for the future realized volatility.

Estimating high dimensional stochastic volatility models

Presenter: Matteo Pelagatti

This paper proposes tree main results that enable the estimation of high dimensional multivariate stochastic volatility models. The first of these results is the closed-form steady-state Kalman filter for the multivariate AR(1) plus noise model. The second result is an EM algorithm that uses the steady-state Kalman filter to quickly estimate the parameters of the same model. The third result is an estimator of the correlation of two elliptical random variables with time-varying variances that is consistent and asymptotically normal regardless of the evolving scheme of the variances.We combine these results to estimate high dimensional multivariate SV models cast in linear state-space form. Let us call d the number of time series and n their (common) length. The problem with the standard approach to SV model estimation cast in linear state-space form and estimated using numerical Gaussian quasi maximum-likelihood (QML) is that each pass of the Kalman filter implies, for every time point t in $\{1, 2, \ldots, n\}$, sums, multiplications and an inversion of $d \times d$ matrices. Thus, for large d the computational burden becomes too expensive and the approach infeasible. Furthermore, the typical quasi-Newton algorithms used to maximize the log-likelihood function become very unstable when the number of parameters is very large (for a portfolio with $d = 100$ assets the parameters to estimate are more than 10,000!).Our approximate QML estimation, instead, is based on the closed-form steady state Kalman filter combined with the EM algorithm and a robust estimator of the returns' correlation matrix. The speed and precision of our methodology is assessed through a simulation experiment. Finally, we implement our method for the estimation of two different types of stochastic volatility models and compare its performance with that of the GARCH-DCC approach in a minimum variance portfolio composed by the constituents of the S&P 100 index.

Modelling Count Time Series with Zero Inflation and Overdispersion

Presenter: Siuli Mukhopadhyay

In this talk a new approach for modelling count time series with excess zeros and high dispersion is proposed. Such time series count data with excess zeros is frequently measured in biomedical, public health and environmental studies. For example, one may consider a respiratory disorder related daily counts of emergency room visits or the counts of a rare disease over a specified time line. The conditional distribution of the dependent variable, given the past history of the process and covariates is assumed to follow a two level zero inflated Poisson distribution, where the overdispersion is modeled using a Gamma distribution. The parameters of the two level zero inflated distribution depend on the past process and covariates through an autoregressive and moving average model. An EM algorithm in the partial likelihood framework is used to compute the maximum likelihood estimates. Convergence results for the estimation procedure is discussed. The proposed modelling approach is illustrated using several simulated and real data examples.

The Effect of Hybridization on Time Series Models for Forecasting Baltic Dry Indices

Presenter: Christos Katris

This study considers whether hybridization of time series methods can improve forecasting accuracy of the Baltic Dry Index (BDI), being a barometer of freight rates in the ocean dry-cargo carrying shipping industry. The long memory property of the BDI series - detected through Hurst's exponent –is modelled through FARIMA models. Feed-Forward Artificial Neural Networks (ANNs) being a machine learning alternative forecasting method, displays comparable, to FARIMA models, forecasting accuracy. The ultimate goal is to improve further the forecasting ability from individual FARIMA and ANN models. To that effect, a hybridization scheme, which incorporates different hybridization approaches, is considered. On the other hand, 'forecast combinations' constitute an alternative way of improving forecasting accuracy of individual forecasting models. The best hybridization alternative is compared with mean and Bates/Granger combination methods. Forecasting accuracy comparisons are made through the well-known RMSE and MAE criteria. Moreover, forecasting methods are compared over different cycles, the latter being identified through the detection of breakpoints in the series with the use of Bai-Perron test. Finally, the robustness of the methods is investigated over the indices that constitute the BDI: that is, for Capesize, Panamax and Supramax indices, representing freight rates of different size dry bulk vessels. In this way, conclusions are drawn of whether the phase of the business cycle and/or the ship size plays a role in the forecasting accuracy of the methods.

Box-Cox transformation in forecasting sales. Evidence of the Greek Market.

Presenter: Maria Voulgaraki

We investigate whether transforming sales leads to an improvement in forecasting accuracy in a trubulent economic environment and a critical time period. We treat sales as a time series and we consider the class of Box-Cox power transformation. We compare several time series models for estimating and forecasting, using Box-Cox transformation parameters on historical monthly data sales of new car retail sector in Greece. Simple models like the mean, seasonal naïve, linear models with seasonal dummies and more complicated ones, like exponential state space smoothing (ETS), seasonal autoregressive integrated moving average (SARIMA), and generalized autoregressive conditional heteroskedastic (SARIMA-GARCH) models, conduct an extensive recursive forecast experiment on a set of seasonal monthly microeconomic time series related to domestic imports and retail turnover in the local market. Empirical evidences consider that Box-Cox transformation produces significantly better forecasts than the untransformed data. In most of the cases the logarithmic and the squared transformation is used as the best ones. Evidences in favour of a transformation become less strong and forecast accuracy becomes weaker as forecast horizon increases. Untransformed data minimize forecasting accuracy measures (like mean squared error-MSE, mean absolute error-MAE), when forecasts are made by simple naïve or the ETS models. When sales data are transformed then the best forecasts are made with linear models with seasonal dummies. During a period of economic crisis is quite difficult to predict sales levels. The use of Box Cox transformation in economic series and simple time series models help in forecasting new car sales levels more accurate in the Greek Market.

Threshold and Multivariate Adaptive Learning Forecasting

Presenter: Foteini Kyriazi

We expand our earlier work on univariate adaptive learning forecasting to account for threshold learning and multivariate settings. In particular, we provide new theoretical results on the use of univariate adaptive learning when the learning rate accounts for the sign of the forecast error on which learning is based and also on how to expand the univariate case to the multivariate one. Both extensions to the univariate case contain new insights and practical usability characteristics that extend the usefulness of the adaptive learning method. We illustrate these theoretical characteristics via simulations and several empirical exercises across a broad context of forecasting benchmarks. Our results are suggestive of the potential of these two extensions.[special request by authors: schedule on Mon or Tue]

Forecasting Related Time Series with Linear methods via Data-driven Regularization

Presenter: Pablo Montero-Manso
The top two most accurate methods in the recent M4 Forecasting Competition showed a way forward in forecasting by successfullyexploiting cross-series information, i.e., they produced better forecasts for individual series by considering other series in the dataset in some form. However, both approaches rely on complex nonlinear methods, which limits our understanding of how cross-series information affects the fitting process and its resulting model. Because of their interpretability and computational complexity, it is interesting to develop linear models that effectively leverage the dataset-level information. We will present a method for related-series forecasting with low order autoregressive linear models, where the cross-series information is introduced via regularization. In our scheme, per-series insample loss of a given set of the autoregressive coefficients is penalized by the loss that these coefficients produce across all time series in the database. We develop a closed-form solution that is extremely computationally efficient, able to forecast hundreds of time series per second while achieving superior forecasting accuracy than the popular automated versions of Theta, Arima or Exponential Smoothing, on both Yearly and Daily frequency series in the M4 dataset (the nonseasonal series). This simple yet effective method avoids costly model selection of Arima and its solutions are always memoryless. These results show the potential of forecasting with linear models when introducing machine learning techniques to exploit the availability of related data, while the insights gained along the way can be applied to more complex models. This approach has connections to Bayesian and Multitask learning methods.

Comparison of ATA trended models with other trended benchmark models on both M3 and M4-Competition Data Sets

Presenter: Tugce Ekiz Yilmaz
ATA method which is a new forecasting method that has a similar form to exponential smoothing (ES) has surprisingly gained important success compared to ES models on both M3 and M4-competition data sets. ATA method is a purely statistical and simple method that can produce accurate forecasts quite fast. The proposed method can be adapted to model time series with various types of trend such as additive, multiplicative and damped. These trended ATA models' predictive performances will be compared to their counter exponential smoothing models. Therefore, in this talk the reasons behind ATA method's success on giving more accurate forecasts in both M3 and M4-competitions compared to some other trended benchmark models will be discussed.

Comparison of The Forecasting Performances of Ata Simple and Simple Exponential Smoothing on M3 and M4 Competition Data Sets

Presenter: Beyza Çetin
The ability to make predictions for the future is of utmost importance in all areas of science. Especially in areas such as industry, finance and medicine, it is vital that data are modeled and transformed into forecasts for the future. There are many methods to make predictions for the future. One of the most popular and simple methods used in predicting the future is the simple exponential smoothing method. However, despite its success and widespread use in many areas, simple exponential smoothing method has some shortcomings that negatively affect the accuracy of forecasts made using this method. Therefore, ATA-simple has emerged as a new method of statistical estimation alternative to simple exponential smoothing. ATA-simple can be used to model times series that do not show any trend or seasonality behavior. ATA-simple model is very similar in mathematical form to simple exponential smoothing but the smoothing constant of ATA-simple is parameterized so that it depends both on the sample size at hand and a smoothing parameter that is optimized on a discrete space. This makes the weighting scheme of ATA-simple dynamic so that the weights are distributed amongst the observed values depending on where in time the smoothing is carried out. Also because the optimization is limited to discrete values, the process is much faster and easier. Another positive difference that ATA-simple brings to the table is that it does not require an initial value. The two models are applied to the M3 and M4 competition data sets simultaneously and their forecasting accuracies are

compared. The method performed better than simple exponential smoothing for both data sets. The detailed results will be discussed in this talk in addition to the model's properties and the advantages that the model brings to the forecasting literature.

Wavenet + Dropout – An Efficient Setup for Competitive Forecasts at Scale

Presenter: Tobias Bischoff

Neural networks have traditionally been regarded as inferior in their ability to generate forecasts compared to traditional statistical frameworks (e.g., ARIMA and variants or exponential smoothing) [1]. However, recent developments have shown that machine learning models based on neural networks are competitive or, on some benchmarks, superior to traditional methods [2]. Unfortunately, in those instances when neural networks achieve equal or better results than traditional methods, the exact architecture and setup that was used to train seems to matter more than the fact that neural networks were involved (e.g., on M4). Here, we present a single deep neural network setup that without extensive hyperparameter tuning achieves state-of-the-art results on some recent and relevant benchmarks including the M4 Forecasting Competition dataset. Neural networks can contain millions of parameters that need to be fitted during training. This can be a major advantage when large amounts of diverse training examples are available. Time series datasets can appear large, but the diversity of individual training examples can be low, and overfitting becomes a problem [3]. Appropriate regularization schemes (e.g., dropout or batch normalization) can reduce overfitting but may require extensive hyperparameter tuning, usually via cross-validation, on validation datasets. Depending on the dataset, this can be a problem when, for example, data-hungry techniques like temporal cross-validation are used. We use concrete dropout [4] at each neural network layer, a technique that allows dropout rates to be learned at training time, to regularize the model. This allows the model to learn the appropriate amount of regularization for each dataset similar to some nonparametric Bayesian methods. A side benefit of concrete dropout is that when applied at prediction time, forecast samples that account for model parameter uncertainty can be generated [4]. Although many neural network architectures are conceivable, we show that, similar to what has been reported in the machine learning literature, a Wavenet-like architecture [6] based on causal temporal convolutions is fast to train and achieves very good results irrespective of which dataset is used [5]. We found that these architectures often train within an hour to very respectable results on the benchmark datasets chosen here. Our efforts demonstrate that in order to achieve state-of-the-art performance similar to, or better than, many statistical or hierarchical methods on recent benchmarks, one need not look further than simple convolutional neural networks and efficient regularization schemes. [1] Makridakis, Spyros, et al. "The M4 Competition: Results, findings, conclusion and way forward." International Journal of Forecasting 34.4 (2018): 802-808. [2] Rangapuram, Syama Sundar, et al. "Deep state space models for time series forecasting." Advances in Neural Information Processing Systems. 2018. [3] Goodfellow, Ian, et al. Deep learning. MIT press, 2016. [4] Gal, Yarin, et al. "Concrete dropout." Advances in Neural Information Processing Systems. 2017. [5] van den Oord, Aaron, et al. "Wavenet: A generative model for raw audio." arXiv preprint arXiv:1609.03499 (2016). [6] Borovykh, et al. "Conditional time series forecasting with convolutional neural networks." arXiv preprint arXiv:1703.04691 (2017).

Bayesian functional predictor for survival data

Presenter: Tanya Garcia

Neuroimaging data now plays an imperative role in understanding the progression of neurodegenerative diseases such as Huntington's, Parkinson's and Alzheimer's disease. A primary focus with such data is discovering and evaluating those neural regions most predictive of clinical outcomes such as age of disease onset. Identifying these neural regions is of high public health relevance as it aids to determine when and how therapeutic treatments should intervene. But handling the complex structure of high-dimensional neuroimaging data, extracting relevant regions and overcoming large censoring rates in event responses complicates image analysis. Overcoming these challenges requires developing advanced statistical tools for neuroimaging analysis which is the focus of this talk.

We propose a new Bayesian framework for identifying brain regions of interest associated with disease onset. Our methods are shown to be consistent and leads to new scientific discoveries in multiple studies of neurodegenerative diseases.\subsection{Building a forecasting model where a 1% error is worth \$5M}Presenter:

Doron Bartov
What do you do when you are tasked with building a yearly revenue model, when the prediction isalmost $1B dollars? At Wix we've been able to successfully create a highly accurate sales forecast model, which approaches this accuracy. We faced some significant challenges along the way:<U+F0B7> Ongoing business changes in the prediction period<U+F0B7> Model should be optimized for many metrics (multiple time horizons)<U+F0B7> Costumers are CXOs who love to get into the algorithm<U+F0B7> Demand for high level precisionIn the talk I will present our journey from a one-model system to using an ensemble of models with different nature. Our current system includes classical statistical models as Arima, Prophet, Rnn based modeland our own home-grown model tailored specifically for Wix. I will show how the ensemble approach is able to help us improve our accuracy by more than 25%

Forecasting Demand at the Mix-Level using Deep Learning and Hierarchical Time Series

Presenter: Sarah Boufelja
Demand Forecast is of paramount importance in Sales and Operations Planning (S&OP) process, intended at improving Supply Chain management through a rigorous supply and demand balancing exercise, rapid-response methods to parts shortage risk and an effective inventory management. However, conventional forecasting methods used in the industry today focus primarily on predicting global sales volumes. While this could make sense for the long-run strategic forecasts, the short-run operational forecasts rely upon a more detailed view, at the Mix Level, that is, at the finest level of granularity for items defining a vehicle (a Lexicon in a more technical jargon): Engine, Energy, Gears, Rooftop, Color ... This paper presents a recent project conducted at Renault for Short termSales Forecast at the Mix level in different geographical areas, using twodifferent Machine Learning methodologies: Deep Learning and LSTM net-works on one hand and Hierarchical Time Series on the other one. Thisproblem presents a twofold challenge: first, Renault Products diversity is rel-atively complex, leading to numerous, sometimes noisy, disaggregated timeseries. Second, Sales data are inherently non-self-comprehensive, therefore itis necessary to integrate explanatory (and exogeneous) data into the model,while avoiding overfitting. Although different exogenous explanatory vari-ables have been identified as suitable candidate for the model (GDP, Infla-tion, Data from Renault configuration system...), this first study relies solelyon historical Demand Data and Inventory Data. The two selected methodologies are complementary, in the sense that they capture different sets of characteristics in the time series: while LSTM networks are extremely efficient at modeling the long-term dependencies in multi-variate time series problems, recent work on hierarchical and grouped time series are naturally suited for disaggregated forecasts. We first demonstrate the performance of these methodologies compared toVanilla models used traditionally for demand forecast and then discuss somefuture improvements to be added to the framework.

Estimating success of own and competitor's new products with pre-release buzz

Presenter: Oliver Schaer
Pre-release forecasting is a vital task for organisations to adjust advertising strategies and operational decisions. Past studies have demonstrated the predictive value of pre-release buzz for forecasting the adoption of new products. Since the pre-release buzz is publicly available and it is well understood that homogenous products follow similar sales pattern we want to investigate how useful such data is to estimate competitors' new products success. For this purpose, we propose to construct profiles of pre-release buzz patterns and associate them with product success. The resulting model allows forecasting the success of a new product by observing its relatively easy to measure pre-release buzz. This approach will not only provide marketers with useful information about their products but will also allow insights to be gained about the competition.We test our approach on sales data of 240 individual video games where we use cluster profiles of pre-release Google Trends information to investigate how those can be effective in predicting the success of own and competitor new products.

Evaluation of multi-horizon strategies for electricity load forecasting

Presenter: Akylas Stratigakos

Forecasting electricity load is an important task for the safe and reliable operation of the power system and has been a prominent area of research for years. Whether the forecasting tasks concern daily or hourly values the relevant literature is mostly focused on comparing the performance of different models. However, when considering multi-step ahead forecasting, such as the day-ahead hourly load, strategy selection is an important factor to be considered that has not been yet fully examined. The selection of the strategy for h-step ahead usually focuses between the two prevalent single-horizon strategies: either the recursive, in which a one-step model is iterated h times, or the direct, in which different h models are trained, with each one producing a single-horizon forecast. Another important strategy is the multiple input multiple output (MIMO), capable of forecasting the whole horizon in one-shot. The MIMO strategy preserves the stochastic dependencies between future values and can outperform the previous strategies in multistep-ahead forecast tasks. In this frame, we focus our attention in evaluating different MIMO strategies for the task of hourly load forecasting, implemented with artificial neural networks (NN), which are gaining lately a lot of attention. In particular, first we present the various multistep-ahead strategies evaluated, namely the MIMO strategy and three variations of it: the direct-MIMO, the recursive-MIMO and a combination of all three of them, in all of which NNs are trained to predict a fraction of the horizon H. Subsequently, we evaluate the performance of the strategies for different h-step ahead horizons, spanning from one day to one week ahead, and compare them to the basic recursive strategy. Our study showcases that variations of MIMO strategy, in which the model is trained to predict a portion of the horizon H, consistently outperform other strategies in terms of error prediction in all time horizons examined, with the results being robust to different configurations of the NNs. In addition, they do so without increasing the computational cost prohibitively, such as the case with the direct strategy. Therefore, our results suggest that splitting the forecasting horizon in smaller fractions can be beneficiary for hourly load forecasting and underline the importance of strategy selection.

Zero Initialization of modified Gated Recurrent Encoder-Decoder Network for Short Term Load Forecasting

Presenter: Madan Mohan Tripathi

Single layer Feedforward Neural Network(FNN) is used many a time as a last layer in models such as seq2seq or a simple RNN network. The importance of such layer is to transform the output to our required dimensions. When it comes to weights and biases initialization, there is no such specific technique that could speed up the learning process. We could depend on deep net-work initialization techniques such as Xavier or He initialization. But such initialization fails to show much improvement in learning speed or accuracy. In this paper we propose Zero Initialization (ZI) for weights of a single layer network. We first test this technique with on a simple RNN network and compare the results against Xavier, He and Identity initialization. As a final test we implement it on a seq2seq network. It was found that ZI considerably reduces the number of epochs used and improve the accuracy. The developed model has been applied for short-term load forecasting using the load data of Australian Energy Market.

Impact of meteorological variables in short-term electric load forecasting

Presenter: Eduardo Caro

This work analyses the impact of meteorological variables (such as temperature, solar radiation, cloudiness, wind speed and wind direction) over the Spanish electric load forecasting. 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.

Unpacking value from a new granularly resolved global dataset on internet activity

Presenter: Klaus Ackermann
Sometime during 2018 an astonishing human technological milestone was reached: over half of humanity became connected to a single open communications platform, the internet. Here we show how the internet can be re-conceptualised as a data-driven insights platform at global scale. We constructed a functional data-set covering over 1,600 cities during 7 years, from over 1.5 trillion observations. To analyse this data-set, we developed a new machine learning technique that provides accurate predictions of temporal patterns of sleep and work. We apply the same method to US city wide electricity demand data, demonstrating improved accuracy compared to current state of the art functional data with scalar response methods. We used our trained model to predict out of sample sleeping and working time for cities around the globe. We find strong agreement between our predictions to data obtained from wrist band monitors, aggregated by city.

Forecasting house prices using online search activity

Presenter: Erik Christian Montes Schütte
We show that Google search activity is a strong out-of-sample predictor of futuregrowth in U.S. house prices and that it strongly outperforms standard predictive models based on macroeconomic variables as well as autoregressive models.We extract the most important information from a large set of search termsrelated to different phases of the home search process into a single Google-basedfactor and then use it to predict movements in future house prices. At the one-month forecast horizon, the Google factor delivers an out-of-sample R2-statisticof about 50% for the aggregate U.S. market over the period 2009-2018. Weshow that the strong predictive power of Google search activity holds for longerforecast horizons, for various house price indices, for seasonally unadjusted andadjusted data, and across individual U.S. states.

Estimation of Weak Factor Models and Its Application to Forecasting

Presenter: Yoshimasa Uematsu
In this presentation, we propose a novel consistent estimation method for the approximate factor model of Chamberlain and Rothschild (1983), with large cross-sectional and time-series dimensions (N and T, respectively). Their model assumes that the r (<< N) largest eigenvalues of data covariance matrix grow as N rises without specifying each diverging rate. This is weaker than the typical assumption on the recent factor models, in which all the r largest eigenvalues diverge proportionally to N, and is frequently referred to as the weak factor models. We extend the sparse orthogonal factor regression (SOFAR) proposed by Uematsu et al. (2018) to consider consistent estimation of the weak factors structure, where the k-th largest eigenvalue grows proportionally to N^{alpha_k} with some unknown 0 < alpha_k <= 1 for k = 1, ..., r. Importantly, our method enables us to consistently estimate alpha_k as well. The finite sample evidence suggests that our method works well. In our experiment, the performance of the new estimator dominates that of the principal component estimators in terms of mean absolute loss, and its superiority gets larger as the factor components become weaker. We apply our method to analyze S&P 500 firm security monthly returns, and the results show that the largest eigenvalue grows proportional to N, whilst the second and third much less slowly diverge.

Improved spare part inventory management using installed base information

Presenter: Sarah Van der Auweraer
Time series methods are widely used in spare parts demand forecasting, to improve the predictability of future demand. However, the lumpiness in spare part demand and limited availability of historical data render time series approaches hard to apply. Alternatively, installed base information can be used to improve spare parts demand forecasting and inventory management. Installed base information comprises information with regard to the whole set of systems or products that an OEM has sold and that are still in use. It includes, for example, machine monitoring tools, information technology, or detailed maintenance data from customers. The rationale behind the use of installed base information to control spare parts inventories is that this information captures the factors that generate the demand for spare parts. As the demand of spare parts

originates from the maintenance activities that require their use, it is related with the number of machines in the field that make use of this part, known as the active installed base, in combination with the part's failure behaviour and the maintenance plan. Our method makes a dynamic inventory decision that minimizes the expected costs, making use of information on the size of the installed base at a particular time, as well as the part and machine's lifetime. We benchmark our method against a series of other approaches and show the added value of different types of information under different scenario settings.

What is the value of POS data in supply chain? The empirical analysis of a two-echelon supply chain.

Presenter: Mahdi Abolghasemi
Demand forecasting is an important input for a lot of managerial decisions in the supply chain such as production planning, material resource planning, and inventory control. Despite advances in technology that eases information sharing, e.g., point of sales (POS) data, and companies' ability to use these data, there are compelling results about their value. It is not clear how much POS sharing can improve the accuracy of demand forecasting, and consequently, reduces the inventory level. We will investigate a two echelon SC where the manufacturer does not have access to the retailers' inventory, and demand indicates a high level of variations due to sales promotions. In this case, the manufacturer should rely either on demand history or POS data to forecast retailers' demand. Our results show that sharing POS can improve the accuracy of demand forecasting under certain conditions. The empirical results are drawn through real data of a food manufacturing company in Australia.

Forecasting Repairing Time of Automotive Parts: An Ordinal Logit Model using LASSO and Elastic Net

Presenter: Shixuan Wang
In this paper, we propose to employ an ordinal logit model with the least absolute shrinkage and selection operator (LASSO) and Elastic Net to forecast the repairing time of automotive parts. The motivation to use the method is two-fold. Firstly, the purpose of forecasting repairing time of automotive parts is to support job scheduling in the garage. Because the repairing jobs are scheduled based on the block of time (such as half an hour), it is satisfactory to have the forecasting repairing time within the block of time as the actual repairing time, even though there is a small gap. It is unnecessary to pursue the precise forecasting in this context. Thus, we convert the repairing hours into ordinal numbers and further employ the logit model to conduct forecasting. Second, the available explanatory variables are mainly nominal variables, such as the vehicle's make, the model, and model year, which needs to be transformed into a large number of dummy variables. The standard estimation techniques are inefficient due to the "curse of dimensionality". To tackle the dimensionality problem, we propose to employ the LASSO and Elastic Net to regularise the estimation results and automatically choose the best combination of explanatory variables and their interactions. The forecasting performance has been examined by an empirical dataset of repairing time of automobile parts, provided by a sizeable fleet company in the United Kingdom. We compare our methods with three benchmarks, a naïve method, k-nearest neighbours, and deep learning. Our methods consistently outperform the three benchmarks in the out-of-sample, and the superior predictive power is confirmed by robustness checks. Our developed method can support garage managers to make more efficient job scheduling and facilitate the operational management of the garage.

Forecasting International Market Segmentation

Presenter: Nigel Meade
Subtitle: The (in)stability of international market segments over four generations of technology Purpose: We examine the stability of segment membership over four generations of mobile phone technology using the interval between a technology's international introduction and its introduction to a specific country as the criterion for international market segmentation. To quantify stability, we develop a metric of the difference between segmentations of countries according to different criteria. In our review, we quantify the lack of consensus between published studies. Methodology: Using data for 172 countries, economic, social,

political and geographical covariates are assembled to examine how well segment membership is explained by a country's observable characteristics. Preliminary modelling using proportional hazards indicates the main drivers of change in the introduction interval over four generations of technology. K means clustering identifies, typically, three segments: early, intermediate and late introducers. Using multinomial logistic regression to explain segment membership, we compare the explanatory power of different subsets of covariates. Findings: Differences in sector membership between generations are shown to be greater than those between most published studies. A logistic regression model using all socio-economic and geographical variables reveals the changing emphasis on drivers of segmentation over the generations and provides the most accurate predictions. Across the digital generations, this model identifies segment membership accurately for 79% of countries. Research Implications: This retrospective analysis demonstrates the difficulty of explaining segment membership. Further research is needed to identify the idiosyncratic factors determining the decision-maker's choice of segment for the remaining 21%. Originality: Counterintuitively, we demonstrate that segment membership in one generation of a technology is a poor predictor of membership in the next.

Do the individual personal characteristics of customers impact the change in importance weights of features?

Presenter: Sheik Meeran
Choice Based Conjoint (CBC) models have been considered as an effective tool to express the trade-off customers make in the form of a preference structure giving the preference weights (importance weights) of the features. Traditionally these weights have been assumed to be fixed and constant over time. However in a recent study involving multiple rounds in extracting such a preference structure for laptops, it was found the weights significantly changed between two rounds which are six months apart. This is observed for the whole cohort of 161 participants collectively. In this study we have attempted to see whether individual's characteristics had any influence on such lability of weights over a six months period. In addition to usual demographic variables such as gender, age, education and employment additional attributes of the customers in using laptops such as perceived competency level with technology, the upgrade and change duration of laptops, the importance to consumers of the technological specifications of laptops, the daily usage by consumers of laptops, the importance to consumers of laptops, were considered in the experiment. The results of this study indicate that the choice and preferences of the participants' were affected by the length of time that elapsed before they changed or upgraded their laptops. The more often participants changed or upgraded their laptops, the more unstable their choices were over time. None of the other characteristics considered was found to be associated significantly with the lability of feature weights.

Analysis of the Key Adoption Predictors in the Market for Mobile Internet

Presenter: Mohsen Hamoudia
This paper focuses on the analysis of the key adoption predictors in the market for Mobile Internet. This is done through a two-stage strategy, separately estimating the effects of the direct and of the indirect network externalities on the diffusion process of Mobile across some selected countries. Positive direct network externalities in the diffusion of Mobile internet, highlighting the relevance of herding and of imitation in adoption behaviour, are captured by estimating alternative specifications of nonlinear, S-shaped, aggregate diffusion curves. A key consequence of these nonlinearities is the qualitative changes they allow to identify in adopters' behaviour, before and after the peak of adoptions has been reached. This peak has long been identified as discriminating between early and late adopters (or majorities Rogers (2003)) and represents an obvious partition between users' attitudes towards the adoption of a new product or technology. This paper's main research objective is to assess how these differences in adoption behaviour, between early and late adopters, interact with other exogenous predictors of Mobile Internet adoption. In more detail, the focus of the paper is on how indirect network externalities may play a different role in the adoption decisions, depending on whether they affect early or late adopters or, in other words, how these indirect externalities' impact, on Mobile Internet diffusion, may change before and after a country's diffusion peak has been reached. The separation of the effects of the indirect network externalities between early and late adopters is captured by estimating country-specific interaction effects. These estimates are then used to assess the role played by marketing strategies on penetration prices, smartphones' diffusion and usage time for Mobile diffusion and

whether these influences are changing depending on their timing along the Mobile diffusion process.

Temporal aggregation level and forecast horizon: investigating the connection

Presenter: Devon Barrow

A key modelling challenge in time series modelling, and particularly for univariate models, is the identification of the time series structure. Over the years, there have been substantial advances in automatic specification of established models, such as the exponential smoothing and ARIMA. Recently, it was demonstrated that by using multiple temporal aggregation it is possible to facilitate the identification further, improving the accuracy and reliability of forecasts. Aggregating temporally a time series filters high-frequency components, making the identification of any underlying long-term dynamics easier and potentially helping model specification. Current research has demonstrated that considering multiple aggregation levels instead of a single "optimal" one, is preferable, due to the difficulty of identifying the optimal aggregation in a general context and because considering multiple levels makes it easier to have a holistic view of low- and high-frequency components of a time series. This has motivated the development of various multiple temporal aggregation approaches, namely the Multiple Aggregation Prediction Algorithm (MAPA) and Temporal Hierarchies (THieF). Both have demonstrated promising performance over conventional forecasts that use only the original time series, particularly for the longer horizon forecasts. This motivates our question: how does the range of aggregation levels considered impact the accuracy of different forecast horizons? For instance, research has shown that these approaches are inferior to conventional modelling for very short horizons, but it is unclear at what point they become beneficial. We argue that although multiple temporal aggregation is overall beneficial, blindly applying for all forecast horizons the same range of temporal aggregation levels, as the literature prescribes, is not ideal and propose an algorithm to dynamically change the temporal aggregation range, depending on the forecast horizon. Our work provides insights in how the framework of multiple temporal aggregation connects with conventional time series modelling.

Can bagging improve the trading performance of exchange rates? The LSTM touch at work

Presenter: Yunjie Wei

A new LSTM-B ensemble learning approach is proposed for exchange rate forecasting and trading by integrating bagging predictor and long-short term memory (LSTM). Many literatures have drawn attention to exchange rate forecasting, but most of them are only focus on the forecasting performance. The accuracy is only one part of exchange rate forecasting, and how to guide practice also is more important part. We extend our forecasting to test the trading performance of exchange rates between the USD and four other major currencies, EUR, GBP, CNY and JPY. The experimental results demonstrate the effectiveness of Bagging predictors, which can significantly improve forecasting accuracy. And the proposed LSTM-B ensemble learning approach outperforms other benchmarks with or without bagging in terms of both forecasting accuracy and trading performance.

What to do with negative weights when combining forecasts?

Presenter: Andrey Vasnev

This paper gives the first thorough investigation of the negative weights that can emerge when combining forecasts. The usual practice in the literature is to ignore or trim the negative weights, i.e., set them to zero. This default strategy has its merits, but it is not optimal. We study the theoretical conditions for the negative weights to emerge in the unconditional framework of Bates and Granger (1969) and in the conditional framework of Gibbs and Vasnev (2018). In the unconditional framework the negative weights are driven by high positive correlations, in the conditional framework, the same effect can also be observed if several forecasts conditionally under(over)estimate the true value. Another important observation is that the region where negative values are theoretically optimal is unstable, i.e., a small estimation error of the underlying parameters can result in significant changes in the weights and the forecast. We then investigate the effect of estimation and trimming using the framework of Claeskens et al. (2016). The positive effect of trimming comes from the variance reduction of the estimated weights, i.e., stabilization, but the threshold of zero is

arbitrary and can be improved. Our proposal is to use an optimal trimming threshold, i.e., an additional tuning parameter to deliver better forecasting performance. We illustrate our theoretical discussions using the European Central Bank Survey of Professional Forecasters where we find that the new strategy performs exceptionally well for the 1-year horizon and can deliver more than 23% improvement upon the trimming with zero threshold (and up to 30% improvement upon the simple average benchmark that uses equal weights).

Designating the Endogeneity in the Cryptocurrency Market Using an Intensity-based Hawkes Process

Presenter: Alev Atak
We study the self-excitability and price clustering properties of the cryptocurrency market using an intensity based Hawkes process. The branching ratio, which is the average ratio of the number of price moves caused by endogenous interactions to the total number of all price changes, is used as a proxy for market "reflexivity" or the endogeneity in the cryptocurrency market. We find that the price process for Bitcoin has the highest endogeneity among the processes for other cryptocurrencies and other financial markets including SP500, Gold, and VIX and finalize our analysis with a forecast application.

Econometric analysis of the Bitcoin market: a successful prediction, and looking forward

Presenter: Spencer Wheatley
We demonstrate the limited explanatory and predictive use of "social signals" – in this case google search activity – to drive value of Bitcoin, as it is rather found to follow price changes. The Metcalfe law fundamental valuation has been supported through conventional bivariate econometric analyses, which features a long run causality from user growth to value, as well as diminishing short run effects from price onto user growth. The JLS model (also known as LPPLS) – a highly non-linear model, which stands in contrast to standard time series models, or unit root tests – was shown to be a useful model for speculative bubbles in Bitcoin. We revisit our successful prediction for Bitcoin made a year ago, and updating our analysis, find that Bitcoin may be approaching the current Metcalfe-based fundamental value. A 1 year out forecast is made.

Predictability, Spillover and Disagreement in Signed Financial Networks

Presenter: Michele Costola
Given the threat to financial stability and the real economy, quantifying systemic risk is now investigated by scholars as well as policy makers. More recently, graph theoretic measures and in particular convergence of autonomous agents on the network to a consensus have been involved in the systemic risk measurement such as early warning indicator for banking crises. In this paper, we extend the study of rate of convergence to consensus of autonomous agents on an interaction network. In particular, we introduce antagonistic interactions and thus a signed network. This will allow to include the, previously discarded, sign information, in the analysis of disagreement on financial networks.

When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage

Presenter: Laurent Ferrara
Nowcasting GDP growth is extremely useful for policy-makers to assess macroeconomic conditions in real-time. In this paper, we aim at nowcasting euro area GDP with a large database of Google search data. Our objective is to check whether this specific type of information can be useful to increase GDP nowcasting accuracy, and when, once we control for official variables. In this respect, we estimate shrunk bridge regressions that integrate Google data optimally screened through a targeting method, and we empirically show that this approach provides some gain in pseudo-real-time nowcasting of euro area GDP quarterly growth. Especially, we get that Google data bring useful information for GDP nowcasting for the four first weeks of the quarter when macroeconomic information is lacking. However, as soon as official data become available, their relative

nowcasting power vanishes. In addition, a true real-time analysis confirms that Google data constitute a reliable alternative when official data are lacking.

Weighted-Covariance Factor Reduction of VARMA Models: Illustrated with Forecasting Quarterly GDP at Monthly Intervals Using Monthly Indicator Variables

Presenter: Baoline Chen
We develop and apply a method, calledweighted-covariance factor reduction (WCFR), for reducing an estimated VARMA"data" model of observed variables being considered to a smallerVARMA "factor" model of a subset of observed variables of primaryinterest. Although WCFR is conceptually and computationally closely related toprincipal components analysis (PCA), it has 3 notably different features: (1)whereas PCA strictly applies only to stationary data, WCFR applies to anymixture of stationary and nonstationary data and models; (2) whereas PCAimplicitly takes a long-term perspective, in WCFR the user sets the perspectiveof any finite duration; (3) like PCA, WCFR can reduce data to factors, but,unlike PCA, WCFR can also reduce data models to factor models. We illustrateWCFR with U.S. monthly indicators (4 coincident, 8 leading) and quarterly realGDP. We estimate 2 monthly models of 5 and 13 variables in differenced-logform, apply WCFR to the estimated models; determine the number of significantfactors for each model; accordingly reduce each data model to a univariate"implicit-factor" GDP model; and, compare the accuracies ofout-of-sample GDP forecasts of the estimated data and implicit-factor models.The application's main conclusion is that WCFR can reduce moderately-largemonthly models of quarterly GDP and up to 12 monthly indicators to univariatemonthly GDP models that forecast out-of-sample GDP at monthly intervals aboutas accurately as their larger antecedent models.

SETAR forecasts with weighted observations

Presenter: Marcella Niglio
In time series analysis the generation of forecasts, based on weighted averages of past observations, has been largely examined.The interest risen by this forecasting approach has been often related, not only to the relative ease of application and interpretation but mainly to the accuracy of the generated forecasts, even when they are compared to more complex forecasting approaches.Starting from this theoretical idea, we propose a new predictor for the nonlinear Self Exciting Threshold Autoregressive (SETAR) obtained as weighted mean of past observations. In particular we provide the theoretical results that allow to derive a predictor whose weights are obtained from the minimization of the Mean Square Forecast Error.The proposed predictor has a number of advantages: it is built to take into account all the observed data and not only the last values of the time series (as usually done in the parametric domain); the weights decrease to zero as the distance between the observed values and the forecast horizon increases; the accuracy of the forecasts is higher than other predictors largely used in the SETAR and in the linear time series literature.The forecast accuracy has been evaluated through a large Monte Carlo study and the application of the proposed predictor to empirical datasets.

Workforce forecasting

Presenter: Carla Di Cairano-Gilfedder
BT is continually investing in making the Digital Britain vision a reality with ever faster and more flexible communication services. To support the associated large scale network infrastructure programmes, BT requires an efficient workforce field service operation able to serve geographically dispersed customers with diverse needs and expectations. Key to a successful engineer field operation is the end to end service chain planning process. This requires optimising resource deployment so as to meet demand while maximising resources' utilisation and leading to lower operational costs and maximising profitability. For such processes BT integrates forecasting and planning at both strategic and operational time-scales, with various AI and optimisation methods being applied to improve forecasting and planning at each timescale.

Delivering forecast models using Plumber APIs in R

Presenter: Alex Hallam
In businesses with many analysts and developers using diverse tools — R,Python, Java, etc — delivering forecast models to client analysts which are accessible and explorable across all tooling may be difficult. We show how forecasters can use the Plumber library in R to serve models via anAPI to users with various tooling; allowing users to access models built in R while using their programming language of preference. We give an example of building these APIs using the open Corporación Favorita Grocery Salesdataset from Kaggle. We close with best practices for the hosting and security ofthese APIs.

PypeCast: A framework for time series forecasting with artificial neural networks and uncertainty modelling

Presenter: Guilherme Cano Lopes
In this paper, we present PypeCast: an open source Python package for time series analysis and prototyping of models for forecasting with artificial neural networks. PypeCast consists of four main modules: Data, Descriptor, Features and Models. The data module currently provides resources to process High Frequency Trading (HFT) data from Brazilian stock market. Other functionalities and common data visualization functions for forecasting are presented in the descriptor module, which aims to provide early insights about the series. In the feature module, a set of transformations, such as the series log return (a feature commonly used in financial forecasting tasks), can be easily obtained by setting a few parameters in PypeCast. The package also brings some predefined artificial neural network forecasters, such as LSTM based models, which are known to perform well in time series forecasting. It also contains deep regression networks forecasters with uncertainty estimation, by using Mixture Density Networks (MDN). As demonstrated in the literature, the later model can be applied in situations when making a wrong forecast is specially costly. At last, PypeCast provides several error metrics along some tools for evaluating a forecasting model, such as comparison to naive forecasts, that helps to avoid some common traps when dealing with time series forecasting as a regression problem. In this work, we also showcase a practical example of analysis and forecasting with PypeCast, demonstrating the ease of use of this package in an example dataset. In this demonstration, we evaluate the accuracy gain of a forecast in a real time series, which was achieved through the discard of uncertain forecasts.

Predicting software development skill from effort estimates

Presenter: Magne Jørgensen
If software developers' estimates of how much effort they would need to complete tasks were strongly correlated with their actual use of efforts, we could safely select among the developers with the lowest effort estimates and predict that they will be among the most skilled ones. Unfortunately, this is not necessarily the case. As documented in studies in many domains, those with lower skill tend to know less about how little they know and, for this reason, give over-optimistic estimates of their own performance. This effect is termed the Dunning-Kruger effect. Does this mean that effort estimates are useless as predictors of software development skill?To find out more on this we requested 104 software developers to estimate four larger and five smaller software development tasks and measured their programming skills. We assessed the connection between lower estimates and higher skill through rank correlations and hit rates, where the hit rate measures the frequency of selecting the most skilled out of two developers when selecting the one with the lower effort estimate on the same task. The results were as predicted by the Dunning-Kruger effect for the larger tasks. The developers in the lowest skill quartile had on average lower estimates than those in the highest skill quartile. To predict relative programming skill based on these estimates would be very inaccurate. The correlations between task estimates and skill were between -0.15 and 0.01 and the hit rates between 44% and 50%.The developers' effort estimates on the smaller tasks were, however, much better connected with their measured programming skills. The rank correlations were between 0.29 and 0.50 and the hit rates between 62% and 68%. Clustering the estimates of all the smaller tasks into the same factor gave a rank correlation of 0.45 between this factor and the measured skill and a hit rate of 70%. While not very strong, we show that the estimates on these tasks is better connected with measured skill than those in use in software development contexts, i.e., company-assesses skill category, length of experience, self-assessed skill

and confidence in knowing how to solve the tasks. These indicators had rank correlations between 0.14 and 0.29, and hit rates between 57% and 65%.A possible explanation of why lower skill was connected with higher effort estimates for the smaller, but not for the larger tasks, is that lower skill in solving a task can both lead to lower skill in identifying the complexity and the simplicity of a task. In the situation with the simpler tasks, we propose, those with better skill were also better in identifying how easily the tasks could be solved at the time of estimation. This explanation suggests that we may be able to design estimation tasks even better at separating those with high and low task completion skill than in our study, when focusing on hidden simplicity, i.e., simplicity that the more skilled are more likely to identify than the less skilled ones.

Determining the Demand Elasticity in a Wholesale Electricity Market

Presenter: Sergei Kulakov
The main focus of researchers in energy markets is typically placed on the analysis of the supply side. The demand side, despite its critical importance, is a subject which still deserves a more profound academic investigation. In particular, the number of studies on the demand elasticity in a wholesale market is limited to merely several pieces. In this paper we extend this field of study and propose a new method for determining the demand elasticity. More specifically, we decompose the data observed in the wholesale market into individual supply and demand schedules of the market participants. These schedules are then used to construct a fundamental model of the market. Our approach allows us to better understand the bidding behavior of the market participants and thus make more precise inferences about the functioning of the electricity market or produce more accurate forecasts.

Horse and Cart: A Scalable Electricity Load and Price Forecast Model

Presenter: Michael O'Leary
A new method for forecasting electricity load and day-ahead price is developed and backtested for the electricity market managed by the Electric Reliability Council of Texas (ERCOT), one of the largest and most heavily traded electricity markets in North America. The Horse and Cart (HOmochronic Recessed Smoothing Ensemble with Corrective Auto Regression Transfer function) method defines the sample space of the forecast based on the region, weather station data within that region, season, holiday, day of the week, and hour of the day to estimate the load. The residuals of the estimate are tracked for the month prior to the 7-day forecast and a vector autoregression (VAR) model of the errors, temperature, and load is generated to produce an error forecast. The Horse model then estimates the 7-day forecast and the error forecast from the Cart model is added to calculate the Horse and Cart load forecast.The load forecast is fed into the day-ahead Price forecast model, the methodology of which is similar to the load model. The day-ahead price of the previous month is calculated based on the Horse method (with day-ahead price as the dependent variable) and the error is tracked to produce a VAR model for the error forecast. The Horse method is used to estimate the day-ahead price for a 7-day forecast and the error forecast is added to calculate the Horse and Cart day-ahead price forecast. The Horse and Cart load model is backtested against the actual load for 18 months and benchmarked against the ERCOT load forecast. The Horse and Cart day-ahead price model is backtested against the actual day-ahead price for 18 months. Parameter tuning is discussed relative to the initial results. The advantages of the model, including the quality of the forecast and limited model training required, are outlined in terms of scalability.

Temporal hierarchies with autocorrelation for load forecasting

Presenter: Peter Nystrup
We propose three different estimators that take into account the autocorrelation structure when reconciling forecasts in a temporal hierarchy. Combining forecasts from multiple temporal aggregation levels exploits information differences and mitigates model uncertainty, while reconciliation ensures a unified prediction that supports aligned decisions at different horizons. In previous studies, weights assigned to the forecasts were given by the structure of the hierarchy or the forecast error variances without considering potential autocorrelation in the forecast errors. Our first estimator considers the autocovariance matrix within each aggregation level. Since this can be difficult to estimate, we propose a second estimator that blends

autocorrelation and variance information, but only requires estimation of the first-order autocorrelation coefficient at each aggregation level. Our third estimator facilitates information sharing between aggregation levels using a sparse representation of the inverse autocorrelation matrix. We demonstrate the usefulness of the proposed estimators through an application to short-term electricity load forecasting in different price areas in Sweden. We find that by taking account of the autocovariance when reconciling forecasts, accuracy can be significantly improved uniformly across all frequencies and areas.

Forecasting International Tourism Demand Using Global and Local Spatiotemporal Autoregressive Models

Presenter: Xiaoying Jiao

Given the rapid development of the tourism industry and the perishable nature of tourism related goods and services, accurate tourism forecasting becomes increasingly important and attracts much attention from both academics and practitioners. A main focus has been put on developing new and innovative models to increase the forecasting accuracy. Spatial econometrics remains to be a popular research area in the general field of econometrics, and has been applied in many different fields of research including tourism. However, applications of spatial econometrics in tourism demand forecasting are still rare, with the only two exceptions (Yang and Zhang, 2019; Wen and Song, 2018). These studies only used global spatial models and local estimation of spatial models has never been used in any tourism demand forecasting study. This study fills the gap by using both global and local forecasting models, the latter of which allows for different parameters and specifications for different regions to reflect both spatial heterogeneity and spatial spillover in forecasting international tourism demand. A simple local spatio-temporal autoregressive model can be expressed as follows (Li et al., 2016): U(i)Yt = U(i)tYt-1 + U(i)<U+03C1>WYt-1 + U(i)μ + U(i)eHere U(i) denotes an N × N diagonal spatial weight matrix assigning one to region i's neighbors within the sample and zero to other regions; U(i) extracts the whole sample into a sub-sample for region i; Yt denotes an N × 1 vector of tourist arrivals at time t; W is a spatial weight matrix; μ denotes the spatial fixed effect; t and <U+03C1> are parameters to be estimated and e is the error term.This study uses annual data of international tourism arrivals to 42 European countries for the period of 1995-2016. The data are collected from the United Nations World Tourism Organization. The forecasting performance of the proposed global and local models will be evaluated against three benchmark models including ARIMA, naïve, and exponential smoothing models. The forecasting accuracy is measured by the mean absolute percentage error (MAPE) and the root mean squared error (RMSE). Besides, the global and local forecasting results are also compared between each other to see if unique parameterizations and specifications for different countries improve the forecasting performance. Furthermore, for each country, direct, indirect, and total spatial effects are calculated and their role in improving the forecasting accuracy is investigated as well. Keywords: Spatial temporal autoregressive model, tourism demand, spatial spillover, forecasting, local estimation Key references LeSage, JP (2008) Introduction to Spatial Econometrics. Boca Raton, FL: CRC Press. Li H, Chen JL, Li G and Goh C (2016) Tourism and regional income inequality: Evidence from China. Annals of Tourism Research, 58: 81–99. Long, W., Liu, C. and Song, H. (2018) 'Pooling in Tourism Demand Forecasting',<U+202F>Journal of Travel Research. doi:<U+202F>10.1177/0047287518800390. Yang Y and Wong KKF (2012) A spatial econometric approach to model spillover effects in tourism flows. Journal of Travel Research 51(6): 768–778. Yang, Y. & Zhang, H. 2019, "Spatial-temporal forecasting of tourism demand",<U+202F>Annals of Tourism Research,<U+202F>vol. 75, pp. 106-119.

Forecasting asymmetric tourism demand over the business cycle

Presenter: Andrea Saayman

Recent research by among other Smeral (2017) showed that income elasticities of tourism demand varies across the business cycle. The effect that the business cycle has on tourism demand has received increasing attention in recent research efforts, with Croes and Ridderstraat (2018) showing that business cycles also explain tourism demand cycles and that these effects are asymmetric. This confirm Smeral's (2017) results of asymmetric income elasticities in tourism demand. This paper extends the work of Croes and Ridderstraat (2018), as well as Smeral (2017) and Mayers and Jackman (2011), by modelling and forecasting tourism demand taking into account asymmetric behaviour of tourism over the business cycle for 5 of the most-travelled nations,

namely Germany, UK, France, USA and Italy to two destinations, namely South Africa and Canada. Three different rolling 1-step and 4-step ahead forecasts generated using the Markov-Switching model, including both state-dependent and state-independent variables, are compared to traditional symmetric models, such as the seasonal ARIMA and seasonal naïve forecasts.

Tourism demand forecasting using a nonlinear ARDL model

Presenter: Doris Chenguang Wu
This study aims to examine the performance of a nonlinear autoregressive distributed lag (ARDL) model in forecasting tourism demand. Given its importance for strategy formulation and policy-making, tourism forecasting has obtained continuous attention from academia. Majority of previous studies adopt linear forecasting techniques which assume that independent variables such as tourist income and destination prices have constant impacts over time on tourism demand (Song and Li 2008; Wu, Song and Shen 2017). However, these kinds of impacts may be asymmetric due to consumers' irrational behaviors. A nonlinear modeling technique, the nonlinear ARDL model, first advanced by Shin et al. (2014), is an effective way to capture this kind of asymmetry and has been applied for asymmetric analysis in different fields (e.g., Raza et al 2016; Nusair 2016). Its application in tourism demand analysis however is very rare with only one exception (Fareed et al 2018), and its forecasting performance has not been assessed in the tourism context. According to Smeral (2019), taking account of asymmetric income effects across business cycles may lead to significant improvement of tourism demand forecasting accuracy. This study therefore aims to first introduce this nonlinear ARDL technique to tourism forecasting and evaluate its performance using Hong Kong tourism demand as an empirical case. The traditional linear ARDL model, naive model and ARIMA model are used as benchmarks for forecasting performance assessment.Keywords: nonlinear ARDL model; asymmetric impact; tourism forecastingReferencesShin, Y., Yu, B. and Greenwood-Nimmo, M.J. (2014) Modelling Asymmetric Cointegration and Dynamic Multipliers in a Nonlinear ARDL Framework. In: Horrace, William C., Sickles, Robin C. (Eds.), Festschrift in Honor of Peter Schmidt. Springer Science & Business Media, New York.Smeral, E. (2019) Seasonal forecasting performance considering varying income elasticities in tourism demand. Tourism Economics. Online.Raza, N., Shahzad, S .J. H., Tiwari, A. K., & Shahbaz, M. (2016) Asymmetric impact of gold, oil prices and their volatilities on stock prices of emerging markets. Resources Policy, 49, 290-301.Fareed, Z., Meo, M. S., Zulfiqar, B., Shahzad, F. & Wang, N. (2018) Nexus of tourism, terrorism, and economic growth in Thailand: new evidence from asymmetric ARDL cointegration approach. Asia Pacific Journal of Tourism Research, 23(12) 1129-1141.Nusair, S. A. (2016) The effects of oil price shocks on the economies of the Gulf Co-operation Council countries: Nonlinear analysis. Energy Policy, 91, 256-267.Wu, D. C., Song, H., & Shen, S. (2017). New developments in tourism and hotel demand modeling and forecasting. International Journal of Contemporary Hospitality Management, 29(1), 507-529.Song, H., & Li, G. (2008). Tourism demand modelling and forecasting – A review of recent research. Tourism Management, 29(2), 203-220.

Evaluating heterogeneous forecasts for vintages of macroeconomic variables

Presenter: Philip Hans Franses
There are various reasons why professional forecasters may disagree in their quotes for macroeconomic variables. One reason is that they target at different vintages of the data. We propose a novel method to test forecast bias in case of such heterogeneity. The method is based on Symbolic Regression, where the variables of interest become interval variables. We associate the interval containing the vintages of data with the intervals of the forecasts. An illustration to 18 years of forecasts for annual USA real GDP growth, given by the Consensus Economics forecasters, shows the relevance of the method.

Endogenous Time-Variation in Vector Autoregressions

Presenter: Danilo Leiva-Leon
This paper proposes a new econometric framework to provide robust inference on the origins of instabilities in the relationship between key macroeconomic variables. We introduce a new class of Time-Varying Parameter Vector Autoregression (TVP-VAR) models where the set of underlying structural shocks are allowed to

potentially influence the dynamics of the autoregressive coefficients. The proposed Endogenous TVP-VAR framework is applied to study the sources of instabilities in the relationship between the unemployment, inflation and interest rates of the U.S. economy. We also show how the proposed framework can be use to deal with issues related to the curse of dimensionality in large TVP-VAR, when employed for forecasting purposes.

Multivariate business cycle by Circulant SSA

Presenter: Eva Senra
Circulant Singular Spectrum Analysis (CSSA) allows to extract the unobserved components associated to any ex-ante desired frequency in a time series in an automatic way. In this paper we first generalize the technique to a multiple setup and automatize it in the same way by the use of circulant matrices applied to the new multivariate trajectory matrix. Second, we extend our proposed methodology to perform factor analysis by frequency. In this way we can jointly extract common factors associated to the trend, to the cycle or to the seasonal component. Finally, we show the application of MCSSA to find common cycles in economic data by sectors orcountries.

Focused Bayesian Prediction

Presenter: Ruben Loaiza-Maya
Bayesian predictive distributions quantify uncertainty about out-of-sample values of a random process conditional only on observed data, where uncertainty about the model-specific parameters is integrated out via the usual probability calculus. Uncertainty about the assumed model itself can, in turn, be accommodated via model-averaging, with the implicit assumption being that the true data generating process (DGP) is contained in the set over which the averaging occurs. We move away from this so-called M-closed world, in which the true DGP is assumed to be either known with certainty, or known to lie in a finite set of models. Herein, we propose a novel method for constructing Bayesian predictive distributions that explicitly acknowledges that practitioners operate in an M -open world. This new approach is not based upon the construction of a conventional Bayesian predictive for a given model, or set of models, but is instead driven by a user-supplied concept of predictive performance loss. To develop such machinery in the Bayesianparadigm, we rely on the principles of approximate Bayesian computation (ABC). A posterior distribution defined over predictive functions is constructed by selecting draws from a prior class of predictives that minimize the loss over a pre-specified evaluation period. We illustrate the potential of our method in a simulation exercise based on misspecified models for a stochastic volatility process, where we find substantial improvements over exact Bayesian methods. Improved forecast accuracy is also a feature of a series of empirical analyses, auguring well for the potential of the new paradigm to reap benefits more broadly.

Automatic Bayesian Aggregation of Predictions

Presenter: Ville Satopää
Many researchers consider the Bayesian paradigm to be the most appropriate for aggregating probability predictions about a future event (see, e.g., French 1983; Genest and Zidek 1986; Clemen andWinkler 1999). This requires a likelihood for the predictions and a prior probability, i.e., a base rate for the future event. The idea is then to use the Bayes theorem to update the base rate in the light of the predictions.The resulting aggregator is likely to work well only if the likelihood describes the heterogeneity in the observed predictions reasonably well. Therefore it should capture important aspects such as bias, accuracy, and dependence. Past literature has found this to be a source of frustration (Clemen andWinkler, 1999). Previous authors have proposed likelihoods with many parameters that must be chosen by the decision-maker (DM) (see, e.g., Lindley 1985). Therefore, before using the aggregator, the DM must specify the base rate but also conduct a complicated elicitation process about human behavior. Furthermore, the aggregator can be very sensitive to the chosen parameter values, and consequently the DM may not feel comfortable using it (Roback and Givens, 2001). We believe this to be one of the main reasons why the existing Bayesian aggregators have not found a strong position in the practitioner's toolkit (see, e.g., O'Hagan et al. 2006; Budescu and Chen 2014).Our main contribution is an Automatic Bayesian Aggregator (ABA) that requires no tuning or

choosing of parameters, does not slow down as the number of predictions grows large, and hence can be applied "out of the shelf" to any set of one-off predictions. First, we explain different ways to choose the base rate and motivate a simple "default" choice based on the predictions. Second, we construct a likelihood that captures accuracy, bias, and dependence, and describes idiosyncratic differences in terms of both information asymmetry and noise. Instead of asking the DM to specify the parameters in the likelihood, we estimate them from the predictions. The procedure is Bayesian and requires a prior distribution for the parameters. We derive a prior that has minimal impact on the aggregate prediction and yields a proper posterior distribution. The final aggregate is then estimated with Markov Chain Monte Carlo and runs very quickly. For instance, 500 predictions aggregate in less than 1 second.The method is evaluated on synthetic and real-world data. Our real-world dataset is from the Good Judgment Project (GJP) (Ungar et al., 2012; Mellers et al., 2014) and consists of millions of expert predictions about around 500 international political future events deemed important by the Intelligence Advanced Research Projects Activity (IARPA). This dataset offers an ideal testing platform because it contains many different topics (finance, politics, etc) and forecasting horizons (1-250 days). On average, ABA dominates over all time horizons. More specifically, on expert predictions the mean squared error of ABA is 24% lower than that of the average prediction and 14-22% lower than that of other state-of-the-art aggregators.

Approximate Bayesian Computing for asymmetric loss functions

Presenter: Georgios Tsiotas

Loss function is an important component in all optimization problems. Optimal forecasts depend on the specification of the loss function. Apart from the quadratic loss which implies the conditional mean forecasts, other loss function have been used in Economics and Finance literature. These involve: the asymmetric linex loss, utility-based losses and loss function bases on the sign of the forecasted value compared such as the MFTR and the MCFD. However, the use of these loss function often create intractability problems. Approximate Bayesian Computing (ABC) is a likelihood-free Bayesian method that intends to by-pass problems such as the unavailability or the intractability of the likelihood function. Thus, instead of deriving the posterior distribution of the parameters' under estimation using the typical Bayesian way, one can derive posterior inference based on those draws that guarantee a relative low distance between simulated and observed sufficient summary statistics expressed by a bandwidth. Efficient ABC algorithms have been implemented using Markov Chain Monte Carlo (MCMC) algorithms. Here, we propose an efficient ABC-MCMC algorithm based on asymmetric kernels. The method is illustrated using several examples in Economics and Finance such as Stochastic Volatility models, Regression models in real estate price predictions, etc.

Deep State Space Models for Time Series Forecasting

Presenter: Syama Sundar Rangapuram

We discuss a novel approach to probabilistic time series forecasting that combines state space models with deep learning. By parametrizing a per-time-series linear state space model with a jointly-learned recurrent neural network, our method retains desired properties of state space models such as data efficiency and interpretability, while making use of the ability to learn complex patterns from raw data offered by deep learning approaches. Our method scales gracefully from regimes where little training data is available to regimes where data from large collection of time series can be leveraged to learn accurate models. We also discuss how our method can be applied to arbitrary likelihood functions. We provide qualitative as well as quantitative results with the proposed method, showing that it compares favorably to the state-of-the-art.

Combinations of Machine Learning and Traditional Approaches to increase accuracy in Forecasting

Presenter: Marinela Profi

This discussion reflects results on the impact of Machine Learning (ML) methods on forecasting process in various industries and contexts, such as retail, manufacturing, energy companies, etc. Massive amount of sparse and noisy data, long- and short-term time series' high variability together with the proliferation

of Machine Learning algorithms, have been challenging traditional statistical methods over the past years. However, results achieved in different industries proved that the key is the combination of the two approaches and, most importantly, figuring out how combinations should be performed. Accuracy of individual statistical or ML methods proves to be lower under conditions of uncertainty and large variances of time series. On the contrary, hybrid methods, based on combination of statistical models and ML components, have a great potential and are next big step. The ability to exploit the benefits of this new "Forecasting Era" will be the differentiating point for organizations' long-term and sustainable succes

Time Series Forecasting via Spline Quantile Function RNNs

Presenter: Konstantinos Benidis
One of the difficulties of applying classical probabilistic forecasting methods as well as recent neural network-based techniques to large scale, real world data sets is the tension between the parametric assumptions made on the output distribution (or the forecast error distribution) and the observed data. The vast majority of techniques assumes a Gaussian distribution, a choice often based on mathematical convenience rather than evidence, which is often not adequate.We propose a flexible method for probabilistic modeling that combines the classical quantile regression, using monotonic regression splines, with recurrent neural networks that parametrize the shape of the splines. The parameters are learned by minimizing the continuous ranked probability score (CRPS). Within this framework, we introduce a method for probabilistic time series forecasting, which leverages the modeling capacity of recurrent neural networks and the flexibility of a spline-based representation of the output distribution.

The relationship between traffic accidents and real economic activity revisited: old targets and new policy implications

Presenter: Antonio García-Ferrer
This work analyzes the relationship between road traffic accidents and real economic activity in Spain, using data on accidents, fatalities and injuries from January 1975 to December 2016. Our results show the historical asymmetriccyclical behaviour of traffic accidents variables. This relationship is more evident for accidents and injuries, while fatalities have shown a different pattern since 2002. Besides using aggregate data, we have analyzed urban and nonurban accidents separately. We analyze the effect of economic variables, public policy interventions and other potential factors affecting traffic series. Regarding policy interventions, we confirm a permanent reduction in all accident rates associated with the mandatory use of seatbelts on car passengers since 1992, as well as withmethodological changes in the data collection from January 2006. However, the penalty points system introduced in July 2006 has only had temporary effects. We have also shown the effect of economic variables such as Industrial Production Index, gasoline and diesel consumption and registration of new vehicles and, as a novelty, the benefits ofusing the composite coincident and leading indicators of the Spanish economy.

Bayesian estimation of Threshold Autoregressions for short-term forecasting of traffic occupancies

Presenter: Yiannis Kamarianakis
Traffic management centers in metropolitan areas use real-time systems that analyze high-frequency measurements from fixed sensors, to perform short-term forecasting and incident detection. The literature focuses primarily on forecasting models for traffic speeds and volumes. Traffic occupancy approximates vehicular density and is frequently used to characterize parts of the network as "congested". This work presents a Bayesian three-step procedure for parsimonious estimation of Threshold-Autoregressions (TAR), designed for location- day- and horizon-specific forecasting of traffic occupancy. In the first step, multiple regime TAR models, reformulated as high-dimensional linear regressions, are estimated using Bayesian horseshoe priors. Next, significant regimes are identified through a forward selection algorithm based on Kullback-Leibler distances between the posterior predictive distribution of the full reference model and TAR specifications with fewer regimes. Finally, the same procedure is applied to identify significant autoregressive terms per regime. Empirical results applied to data from TRB's traffic forecasting competition, illustrate the efficacy of

the proposed scheme in producing interpretable models, which lead to satisfactory point and density forecasts at multiple horizons.

On the validation of the performance of micro-scale vehicle emission models in a metropolitan city network

Presenter: Christos Keramydas

Microscale vehicle emissions modelling is an advanced approach employed by leading local authorities worldwide to acquire information on the emission behaviour of vehicles and calculate emission inventories. The selection of the appropriate model is of critical importance as this information supports decisions in estimating, assessing, controlling and regulating the environmental impacts of road transport in metropolitan areas.Although there are several studies on the validation of this kind of models, the existing literature does not include a generic framework of metrics or any guidelines on the overall validation process. Very little information on the sufficiency of the metrics used is available. There are limited studies based on Portable Emissions Measurement Systems (PEMS).This paper assesses four microscale vehicle emission models using a set of real-world second-by-second vehicle emissions data measured with PEMS. The models, namely Modal and Peak Model (MAPM), P<U+0394>P, Passenger Car and Heavy-Duty Emission Model (PHEM) and PHEM Light, were developed using a modelling dataset provided by the Environmental Protection Department of Hong Kong. The dataset included Franchised Double-Decker Buses (FBDD), Non-Franchised Buses (NFB), Liquified Petroleum Gas (LPG) Taxis and Heavy Goods Vehicles (HGV). Predictions were made for 7 pollutants, including CO2, NOx, and PM10. The models were validated using the vehicles of the modelling dataset and also other "unknown" vehicles.The paper proposes a set of widely accepted measures aiming to assess different aspects of the performance of the models, e.g. level, trend and peaks. The strengths and the weaknesses of the metrics are discussed. An overall evaluation of microscale modelling and a comparative assessment of the models is provided. Some ideas on the respective decision-making process, i.e. to select the most appropriate model, are presented.The results showed that the "best" predictive model varied with the measure used to evaluate performance. Some metrics exhibited extreme values, e.g. mean absolute percentage error. The performance of the models was improved by 15%-20% when emissions were aggregated over time, i.e. 10 seconds vs. 1 second. The performance also improved when predicting the total emissions of a trip.The ratio of the mean absolute error to the respective observed mean value of the time-series ranged from 65% to 77%, and the associated R values (observed-predicted) ranged from 0.55 to 0.72, depending on the vehicle class and the pollutant. No large differences were observed between models. There was not a dominant model within a vehicle class or a pollutant. CO2 was the most predictable pollutant and NO2 the less predictable one. NFBs and FBDDs were better predicted than HGVs and Taxis. The predictions were often found to be more accurate for the unknown vehicles than for the known ones. The metrics revealed that the models could be further improved.The prediction of vehicle emission time-series is a challenging task due to their high variance and peaks. A generic framework for the evaluation of microscale emission models is necessary. The performance of microscale models could trigger an interesting debate on the role of this type of modelling in estimating real-world emission inventories.

Bayesian Risk Forecasting for Long Horizons

Presenter: Agnieszka Borowska

We present an accurate and efficient method for Bayesian forecasting of two financial risk measures, Value-at-Risk and Expected Shortfall, for a given volatility model. We obtain precise forecasts of the tail of the distribution of returns not only for the 10-days-ahead horizon required by the Basel Committee but even for long horizons, like one-month or one-year-ahead. The latter has recently attracted considerable attention due to the different properties of short term risk and long run risk. The key insight behind our importance sampling based approach is the sequential construction of marginal and conditional importance densities for consecutive periods. We report substantial accuracy gains for all the considered horizons in empirical studies on two datasets of daily financial returns, including a highly volatile period of the recent financial crisis. To illustrate the flexibility of the proposed construction method, we present how it can be adjusted to the frequentist case, for which we provide counterparts of both Bayesian applications.

Forecasting economic decisions under risk: The predictive importance of choice-process data

Presenter: Steffen Mueller
We investigate various statistical methods for forecasting risky choices and identify important decision predictors. Subjects are presented a series of 50/50 gambles that each involves a potential gain and a potential loss, and subjects can choose to either accept or reject a displayed lottery. From this data, we use information on 8800 individual lottery gambles and specify four predictor-sets that include different combinations of input categories: lottery design, socioeconomic characteristics, past gambling behavior, eye-movements, and various psychophysiological measures that are recorded during the first three seconds of lottery-information processing. The results of our forecasting experiment show that choice-process data can effectively be used to forecast risky gambling decisions; however, we find large differences among models' forecasting capabilities with respect to subjects, predictor-sets, and lottery payoff structures.

Bootstrap confidence intervals for multi-period value at risk

Presenter: Esther Ruiz
The Basel Accords require financial institutions to report daily their Value at Risk (VaR) computed using 10-day returns. Together with point predictions of the VaR, the Basel Accords require measures of the associated uncertainty. Due to its importance, there is a large literature dealing with VaR prediction. However, most authors deal with VaR point predictions based on one-day returns. Our first contribution is to compare several popular VaR point prediction procedures when implemented to predict the more realistic VaR based on 10-day returns. Using simulated data, we compare procedures based on direct approaches and procedures based on iterative approaches when the conditional variance is assumed to be specified by a given GARCH-type model. Second, we consider available procedures to compute the uncertainty associated with VaR predictions based on one-day returns and, when possible, extend them to compute the uncertainty of VaR for 10-day returns. The performance of these procedures is analyzed by means of Monte Carlo experiments. Finally, we propose a bootstrap procedure for obtaining confidence intervals of future Value at Risk (VaR) estimated based on multi-period returns and parametric specifications of the conditional variance. These intervals incorporate the parameter uncertainty associated with the estimation of the conditional variance and/or of the conditional quantile without assuming any particular error distribution. The results are illustrated computing the VaR of three different series of returns, namely, the obiquitous S&P500, the exchange rate Dollar/Euro and IBM.

Adapting the Supply Chain to match the Acceleration of Change

Presenter: Mike Wilson
We are living through a period of accelerated change, which is having a dramatic impact upon global supply chains. Consumer expectations, technological advances, changing demographics, environmental and socio-political trends all call for a new approach.

We have seen inventory levels rise to cope with demand for immediacy (for example, next-day delivery), but we are relying on supply chains that are over-stretched and not designed for the open-source, flexible and dynamic demand that we now experience.

We have spent many years trying to perfect a stable process with incremental improvements. Can we adapt our current tools, methods and models to the new world? Or should we be developing new models of supply chain, which supersede anything we have seen before?

The demand and supply equation is no longer balanced. Therefore, we need to be adapting forecasting and inventory optimization to a holistic view using a new model that is dynamic, circular, flexible and connected.

Inventory levels have plateaued and organisations must now position inventory as close to market as possible to remain competitive. Domestic transportation will grow as international transportation shrinks; there will be more value-add closer to market. We need new models and tools to facilitate rapid inventory optimisation and dispositioning. We need to re-forecast the entire supply chain.

The impact of time series characteristics on temporal aggregation performance

Presenter: Dejan Mircetic

Non-overlapping temporal aggregation is an intuitively appealing approach that transforms a time series from higher into lower frequencies strengthening or attenuating different elements. This is achieved by dividing the time series into consecutive non-overlapping buckets of time with length m, where m is the aggregation level. A new series is then generated through the summation (bucketing) of every m periods of the high frequency time series. Forecasters then have two options for forecasting purposes: using original series ii) non-overlapping aggregated series. Recent studies show that two approaches have their own merit.In this paper, we analyze the impact of time series characteristics on the performance of the non-overlapping temporal aggregation approach. We conduct an empirical study using M4 competition data set and considering three of the most popular forecasting methods Naïve, ARIMA and Exponential Smoothing. We extract measures from a comprehensive set of features such as trend, seasonality, periodicity, autocorrelation, skewness, kurtosis, nonlinearity, and randomness. We first examine the evolution of time series characteristics by applying temporal aggregation. We then determine the conditions under which non-overlapping temporal aggregation provides more accurate forecasts and provide recommendations rules for forecasters when using temporal aggregation.

Supply Chain Management: Bridging the gap between business and academia

Presenter: Evangelos Theodorou

This conference paper focuses on challenges and limitations related to the development and implementation of decision support systems for supply chain management. The utilization of such systems becomes mandatory as the volume of data increases and the operational constrains become more complex. Although lots of research has been conducted in the field, the academic approach usually differs from the one applied in practice. For instance, the literature focuses mostly on constructing accurate forecasting methods and inventory control algorithms without taking into consideration key elements, as the commercial agreements between the companies and their suppliers or the uncertainty related to the lead times. On the other hand, companies' decisions are significantly affected by such variables, putting little emphasis on forecasting accuracy and measures that monitor their actual performance or their true potential. Another issue for applying science in practice is the unfaithful environment that is observed and time required for trusting and adopting new systems and practices, which can be only achieved after communicating some tangible and intuitive results. In order to discuss these issues and demonstrate some suggestions, we present an interesting case study about a major Greek retail company involving more than 15,000 SKUs, 271 stores and 3 warehouses. We show that state-of-the-art forecasting methods and optimization techniques can be efficiently combined with business-as-usual practices through integrated supply chain management systems and representative reports that maximize service level and minimize storage cost.

Is new-age, virtual engagement for large cohorts, useful in predicting the overall performance of students?

Presenter: Gerrie Roberts

The goal of this study is to determine if virtual engagement plays a role in predicting student performance. Specifically, we will determine if two continuous forms of in-semester engagement are useful in predicting the overall performance of students enrolled in a first level business statistics unit at Monash University, Australia. The first of these is the new-age, virtual form of engagement captured by time spent in an online, statistical analysis platform that has built-in adaptive learning. The second form of engagement is attendance at interactive lectures wherein unique polling technology facilitates group work and response submission, in a large lecture theatre setting. Whilst there are a plethora of studies on the impact that the various forms of online engagement has on exam performance (e.g., Davies and Graff, 2005, Chen and Lin, 2008), the focus of this study is different in two aspects. Firstly, the impact of in-semester engagement is measured against overall student performance, which includes both in-semester and final examination assessments, not just an exam score. Secondly, our forms of online engagement, i.e., attendance at interactive lectures and time spent online are both direct measures of time spent by students in active learning platforms, rather than simply time logged in a Learning Management System such as Blackboard or Moodle. For our study, data is available

for large cohorts of students (600-900) enrolled in a first level business statistics unit over four semesters from Semester 1 2017 to Semester 2 2018. Based on the overall grades of High Distinction (80 and above), Distinction (70 to 79), Credit (60 to 69), Pass (50 to 59) and Fail (below 50), we draw stratified samples from the population of students enrolled in each semester. The in-semester engagement, viz., interactive lecture attendance and time spent in online study employing adaptive learning, were electronically recorded in each semester.Consistently, across all four different models fitted to the combined samples from the four semesters, i.e., least squares regression, quantile regression and both binary and ordinal logistic regression, we are able to show that virtual engagement captured by time spent in an online, statistical analysis platform and attendance at interactive lectures, both have significant positive impacts on the overall performance of students; hence, enabling us to use these models to predict future academic student performance in this new-age, virtual learning environment.ReferencesChen, J., & Lin, T. 2008. Class attendance and exam performance: A randomized experiment. Journal of Economic Education 39:213–27.Davies, J. & Graff, M. (2005). Performance in e-learning: online participation and student grades. British Journal of Educational Technology, 36(4), 657-663.

A Comparison of Seasonal Adjustment Approaches using State Space Representation

Presenter: Steve Matthews
Two approaches for seasonal adjustment are used predominately by statistical agencies throughout the world; namely X12-ARIMA which is uses moving averages to separate seasonal and irregular effects from the trend-cycle component, and SEATS which uses explicit parametric time series models to do so. We consider several state space models including one to proxy each of these methods, and empirically compare the models using real datasets from Statistics Canada to gain insight into different decompositions. The components of the models are contrasted to illustrate their similarities and differences. An overall comparison between the methods will be drawn, using parameters from their state space representations, and practical considerations for their use in official statistics will be discussed.

An unobserved component modelling approach to evaluate fixed-event forecasts

Presenter: Jing Tian
We propose an unobserved modeling framework to evaluate a set of forecasts that target the same variable but are updated along the forecast horizon. The approach decomposes forecast errors into three distinct horizon-specific processes, namely, bias, rational error and implicit error, and attributes forecast revisions to corrections for these forecast errors. By evaluating multi-horizon daily maximum temperature forecasts for Melbourne, Australia, we demonstrate how this modeling framework can be used to analyze the dynamics of the forecast revision structure across horizons. Understanding forecast revisions is critical for weather forecast users to determine the optimal timing for their planning decisions.

Online Bayesian Density Forecasts for Vehicle Trajectory with Heterogeneous Driver Behaviour

Presenter: Nathaniel Tomasetti
Self-driving vehicles operate by continuous application of three basic steps: Detect the position of surrounding vehicles, forecast their future trajectory, and plot a route that will avoid possible collisions. To increase safety these forecasts should be probabilistic so the self-driving vehicle may avoid routes that may cause a collision with low probabilities. In this paper we provide a two stage methodology to provide online density forecasts of vehicle trajectory. First we infer a wide range of differing human behaviours present in the NGSIM US Route 101 Highway dataset through a Bayesian hierarchical model. Later we leverage this distribution as prior information to construct individual vehicle models, and repeatedly forecast the trajectory of these vehicles as information is observed. individual vehicles The bivarariate auto-regressive models employed for vehicle acceleration and steering angle are significantly more accurate than the standard constant models commonly employed, while providing forecasts of the entire distribution of possible vehicle trajectories. An

online recursive posterior updating scheme allows forecasts to be computed within a short time frame suitable for driving.

Variational Inference for high dimensional structured factor copulas

Presenter: Hoang Nguyen
Factor copula models have been recently proposed for describing the joint distribution of a large number of variables in terms of a few common latent factors. In this paper, we employ a Bayesian procedure to make fast inferences for multi-factor and structured factor copulas. To deal with the high dimensional structure, we apply a variational inference (VI) algorithm to estimate different specifications of factor copula models. Compared to the Markov chain MonteCarlo (MCMC) approach, the variational approximation is much faster and could handle a sizeable problem in a few seconds. Another issue of factor copula models is that the bivariate copula functions connecting the variables are unknown in high dimensions. We derive an automatic procedure to recover the hidden dependence structure. By taking advantage of the posterior modes of the latent variables, we select the bivariate copula functions based on minimizing the Bayesian information criterion (BIC). The simulation studies in different contexts show that the procedure of bivariate copula selection could be very accurate in comparison to the true generated copula model. We illustrate our proposed procedure to model and predict the temperatures at different stations in Germany.

Error Metrics for Time Series Forecasting

Presenter: Lilian Wong
In this talk, we present the challenges we encountered when choosing the proper error metrics for training and validating time series forecasting models. Then we elaborate on the shortcomings of a few established error metrics such as MAPE to illustrate how textbook reliance on those metrics could lead to errors in judgment, cause us to miss some critical edge cases, and in turn impact customer confidence.

Next, we present several custom new metrics for time series forecasting use cases where spikes are of particular significanceto our customers. We demonstrate how we obtain a more wholistic description of model errors by combining these error metrics, each with their strengths and weaknesses.

On the parametrization of simple autoregressive models with neural networks

Presenter: Laurent Callot
This paper demonstrate how simple, interpretable univariate time-series models can be parametrized using feed-forward neural networks in the context of a panel of data.

Using neural networks to parametrize these models has several advantages. The estimation procedure can take advantage of the panel structure of the data to efficiently pool information while allowing for heterogeneity, without requiring the researcher to specify what should be pooled a priori. The neural network can be trained by likelihood maximization, the model is fully probabilistic with estimates of uncertainty. Standard residual-based tests can be applied. Constraints on the parameter values are easily imposed, including joint constraints on multiple parameters. The parameters can be made to be functions of covariates such has seasonal features or past observations. The loss function can be written to optimize forecasting accuracy one or several steps ahead, or jointly over multiple steps. The efficacy of the proposed method is illustrated with autoregressive models on synthetic and real data.

Forecasting for Social Good: Why it Matters

Presenter: Bahman Rostami-Tabar
Forecasting for Social Good is an initiative sponsored by the International Institute of Forecasters. The aim of this session is to discuss activities undertaken under this initiative, increase awareness in the area of forecasting for social good, facilitate a constructive exchange of ideas and enable IIF members to further contribute in this area. We summarise past actions and consider the way forward. The forecasting discipline's focus has largely been on forecasting for financial and economic reasons; much less has been contributed in the area of forecasting for social impact. We wish to explore and engage the forecasting community in the

following questions: i) What does "forecasting for social good" mean? ii) Why is it important? iii) How is it different from business forecasting? iv) What role can forecasting play to benefit society? and v) What ought to be the role of our forecasting community as part of this activity?

Making text count

Presenter: George Kapetanios
We provide several novel methodological and empirical contributions in forecasting using large datasets. The first contribution is a novel analysis of a large textual dataset using high dimensional regression methods. The second extends this high dimensional analysis by developing a time series model that accommodates unstructured datasets allowing for several attributes of granular observations to affect the time series properties of the forecasting outcomes. Finally, we present a comprehensive methodology for incorporating time variation in machine learning forecasting models while allowing for elements of interpretability and model uncertainty.

Googling Fashion: Forecasting Fashion Consumer Behaviour using Google Trends

Presenter: Emmanuel Silva
This paper aims to discuss the current state of Google Trends as a useful tool for fashion analytics, show the importance of being able to forecast fashion consumer trends and then presents a univariate forecast evaluation of fashion consumer Google Trends to motivate more academic research in this subject area. Using Burberry, a British luxury fashion house, as an example, we compare several parametric and nonparametric forecasting techniques to determine the best univariate forecasting model for "Burberry" Google Trends. In addition, we also introduce Singular Spectrum Analysis as a useful tool for denoising fashion consumer Google Trends and apply a recently developed hybrid Neural Networks model to generate forecasts. Our initial results indicate that there is no single univariate model (out of ARIMA, Exponential Smoothing, TBATS, and Neural Network Autoregression) which can provide the best forecast of fashion consumer Google Trends for Burberry across all horizons. In fact, we find Neural Networks Autoregression (NNAR) to be the worst contender. We then seek to improve the accuracy of NNAR forecasts for fashion consumer Google Trends via the introduction of Singular Spectrum Analysis for noise reduction in fashion data. The hybrid Neural Networks model (Denoised NNAR) succeeds in outperforming all competing models across all horizons, with a majority of statistically significant outcomes at providing the best forecast for Burberry's highly seasonal fashion consumer Google Trends. In an era of Big Data, we show the usefulness of Google Trends, denoising and forecasting consumer behaviour for the fashion industry.

Seeking Forecasting Accuracy Gains at Scale- Does Pre-model Data Clustering Help?

Presenter: Koel Ghosh
In retail, in context of servicing enterprise demand forecasting at scale, the approach taken sometimes is to fit a time series model to a group of products. The product 'groups' could be the result of the retails firm's internal organization of the products or operational handling that clubs' products together or grouping driven by consumer buying behavior. A good part of retail demand forecasting practice at scale is about the search of the one forecasting method that works for most products in a group (or the entire assortment), and many recent works compare the aggregate forecasting accuracy of parametric and non-parametric models in being that one generalizable forecasting model. Irrespective of the model chosen, forecast accuracy at an aggregate level is a function of how similar the individual underlying time series are in terms of time series characteristics like seasonality and trend. In real life applications, there is always that subset of outlier products in any grouping context that time series wise behave differently than most others in their group impacting aggregate forecasting accuracy for that group. Many retail and other organizations doing forecasting at scale handle this by putting domain informed analysts in the forecasting process loop to make needed modeling adjustment for those non-confirming time series in a group. But this manual insertion breaks the automation loop. In this presentation, an alternative automated way of improving on forecasting errors based on clustering of time series in a group prior to modeling is explored- the clusters force design of

meaningful sub groups of similar items allowing for pooling of information from different products to fit a single model. The presentation will include supporting discussion on implications of doing so at scale both for forecasting accuracy as well as computations considerations.The rationale and the efficiency of the clustering prior to modeling is demonstrated empirically using a dataset of anonymized products. The experimental set up anchors on a basic retail demand forecasting model (i.e. multivariate time series model of sales on price, promotions, retail holidays, seasonality and trend) being applied to each product in the group as the control. The forecasting accuracy at individual product level and in aggregate for the group for this base model is measured in terms of MAPE, for the validation period as well as a 'pseudo' out of sample period. In the approach outlined in this presentation, the products are clustered based on time stamp dependent and independent descriptive measures of model variables for each product, and the same base model is run for each cluster. The forecasting accuracy from this experiment is then measured against the control. The presentation will then address the challenges and benefits of automating the approach and applying it at scale.

Large-scale retail forecasting – a hybrid approach

Presenter: Michal Kurcewicz
This paper presents a large-scale forecasting system that implements a hybrid approach to retail demand forecasting, using both machine learning as well as time-series techniques to generate demand forecasts. While traditional time series techniques like ARIMA(X) or exponential smoothing have been widely and successfully applied to model retail demand, these techniques are not able to fully utilize the data available to modern retailers. These data include not only detailed sales and promotion information but also data from the company's loyalty program, product database, e-commerce system, information on local events as well as data supplied by expert business users and external sources. To get most out of these data, time series techniques can be augmented with machine learning methods. We present a comprehensive hybrid forecasting pipeline that combines time series methods with gradient boosting. We discuss data cleansing, creation of features (explanatory variables) and time-series segmentation based on the characteristics like seasonality or intermittency. Next, we discuss the recommended forecasting approach for each segment. Results indicate that for regular demand items gradient boosting offers better forecast performance while being easier to develop. Finally, we discuss how business experts can influence the forecasts. The discussed system has been implemented at a major Polish retailer.

Air Passenger Demand Forecasting using Deep Learning Methods

Presenter: Nahid Jafari
Demand for air travel is essentially driven by economic and tourism activities. The economic recessions decreases demand for leisure and business travel, consequently, demand for air transport which is the most important transportation mode in the United States. As the economy recovers from economic downturns, the aviation industry will continue to grow over the long run.Air transport demand forecasting is receiving increasing attention, especially because of intrinsic difficulties and practical applications. The forecast of the air passenger demand over time can be analyzed as time series forecasting in which it has a complex behavior due to their irregularity, high volatility and seasonality. Among quantitative forecasting approaches, econometric, statistical and artificial intelligence approaches are well-known. In this study, we examine various deep learning methods to predict the air passenger demand of the U.S. domestic airlines. Deep neural networks, a multiple layer perception (MLP) with multiple hidden layers, can learn more complicated function of the input. We frame the problem where, given the air passenger demand for prior years, we forecast the passenger demand (aggregate and disaggregate) for a short term. We implement the models with Keras, the Python deep learning library.

Air Quality Forecast with Recurrent Neural Networks

Presenter: Michel Tokic
Air quality forecast is of importance for the control of air polution in cities caused by cars and households. To allow a government control of the air pollution there should be a forecast of the essential pollution parameters

(NO2/PM2.5/PM10) for the next days. Obviously the pollution depends on traffic density and weather, but we get only measurements of a part of these variables. With recurrent neural networks we can build models for several days ahead forecasts even if there are missing variables or missing data. In the talk we will show different model types attacking these application problems. The recurrent neural networks are deviations of historical consistent neural networks (HCNN). The talk represents our state-of-the-art on this forecasting field.

Enhancing predictive classification of bank customers with machine learning and network information

Presenter: Viacheslav Lyubchich
Modern capabilities of customer analytics allow retailers to track individual clients and their peers and to study their behavioral patterns to enhance customer segmentation, optimize next best offer and customer retention strategies. While telecommunication providers have been using social peer network data to improve their customer analytics and business intelligence solutions for a number of years, the knowledge of peer-to-peer relational effects in retail banking is still very limited. We demonstrate a cost-effective way to collect peer information in retail banking and integrate these data for a higher predictive accuracy. Using an example of predicting activity churn of retail bank customers in Canada, we demonstrate utility of several machine learning methods and propose an efficient stacking procedure of multiple predictions, using a deep learning convolutional neural network.

Econometric modelling and forecasting of intraday electricity prices

Presenter: Michal Narajewski
We analyse the ID3-Price on German Intraday Continuous Electricity Market using an econometric time series model. A multivariate approach is conducted for hourly and quarter-hourly products separately. We estimate the model using lasso and elastic net techniques and perform an out-of-sample very short-term forecasting study. The model's performance is compared with benchmark models and is discussed in detail. Forecasting results provide new insights to the German Intraday Continuous Electricity Market regarding its efficiency and to the ID3-Price behaviour.

On the Importance of Cross-border Market Integration under XBID: Evidence from the German Intraday Market

Presenter: Christopher Kath
The introduction of XBID (the Cross-border intraday project) marked a fundamental step towards European market integration and the idea of one pan-European copper plate. It allows intraday market participants to trade based on a shared order book independent of countries or local energy exchanges. In theory, this leads to an efficient allocation of cross-border capacities and ensures a maximum of market liquidity across European intraday markets. We review this postulation with a particular focus on the German EPEX intraday market. Does the introduction of XBID influence prices to an extent that is measurable by means of statistical methods such as regression analysis or dedicated tests for structural breakpoints? However, we do not only focus on the impact of XBID on the German ID3 index and the volume weighted average price (VWAP) but elaborate if we can transfer such insights to the world of forecasting. XBID enhances the level of market integration which makes it crucial to consider electricity markets as an integrated system of several exchanges. We demonstrate that an explicit consideration of neighboring countries by means of their price levels and its connected fundamental determinants increases forecasting accuracy. We present a forecasting model that exploits cross-border capacities as well as several European load forecasts and prices and compare its performance to other state-of-the-art models.

A Generative Model for Multivariate Probabilistic Scenario Forecasting

Presenter: Tim Janke
In recent years the paradigm in forecasting, especially in the weather and energy related forecasting literature,

has been shifting from point forecasts to probabilistic forecasts, i.e. forecasting a predictive distribution instead of only reporting a mean estimate. Having access to a predictive distribution allows an agent that bases his actions on a forecast to incorporate the associated risk in his decision making process, e.g. by using tools from stochastic programming. This usually requires to draw samples from the predictive distribution. In the univariate case this is straightforward. However, in the case of a multivariate forecasting setting this is more challenging as drawing independently from the marginal distributions does not preserve the dependence inherent in the data. Here, copulas can provide a way to construct functions that couple the known marginal distributions with the joint distribution.We introduce an alternative method that allows to bypass the specification of a copula by training a generative model that learns to generate consistent multivariate samples by minimizing the Energy Score. We draw on ideas from Implicit Generative Models (IGM), e.g. Generative Adversarial Networks. IGMs are likelihood free generative models, i.e. they are trained to approximate an arbitrary distribution via generated samples.We demonstrate our approach on a data set for the German-Austrian electricity spot market. We first train an ensemble of point forecasting models to predict the mean for each of the 24 hourly prices. We then train a linear generative meta model that maps samples drawn from a set of univariate latent distributions to 24-dimensional price vectors. We parametrize the latent distributions of the generative model by estimates for the historical noise and the ensemble dispersion. We find that our approach gives better results than a naïve benchmark and a Gaussian copula approach in terms of Energy Score.

Going With Your Gut: The (In)accuracy of Forecast Revisions in a Football Score Prediction Game

Presenter: Carl Singleton
We study individuals who each chose to predict the outcome of fixed events in an online competition, namely all football matches during the 2017/18 season of the English Premier League. We ask whether any forecast revisions the individuals chose to make (or not), before the matches began, improved their likelihood of predicting correct scorelines and results. Against what theory might expect, we show how revisions tended towards significantly worse forecasting performance, suggesting that individuals should have stuck with their initial judgements, or their 'gut instincts'. This result is robust to both differences in the average forecasting ability of individuals and the predictability of matches. We find evidence that this is because revisions to the forecast number of goals scored in football matches are generally excessive, especially when these forecasts were increased rather than decreased.(Link to preliminary working paper: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3260428)/subsection%7BForecasting the outcome of professional tennis matches: Evidence from Wimbledon 2018}Presenter: Leighton Vaughan Williams
Significant advances in the forecasting of tennis match outcomes can be traced to the publication of early mathematical and statistical models. These models can be broadly categorised as regression-based, point-based and paired comparison models. They have also been compared with odds-based predictions. Regression-based models have been based on official player rankings and ELO-based ratings (the difference in the ELO rating between two players serves as a measure of the probability that one or other player will win the match). Points-based models have been based on specifying probabilities of winning a point on both serve and return. Paired comparison models, such as the Bradley-Terry model, have also been used to predict match outcomes. Of all of these, perhaps the most well known is the FiveThirtyEight model, the approach of which is a variant of the ELO rating system, originally used to gauge the strength of chess players. In this paper, we use data from the Wimbledon 2018 Lawn Tennis Championships, to compare the forecasting efficiency of the official rankings of players, of the ELO ratings of each of the players, and the best odds offered by bookmakers, deflated to generate implied probabilities. We do this for both the men's singles tournament and the women's singles tournament. We derive conclusions as to which of the approaches generated the best forecasts, and draw some implications from our findings for forecasting more generally.

Evaluating Strange Forecasts: The Curious Case of Football Match Scorelines

Presenter: James Reade
This study analyses point forecasts for a common set of events. These forecasts were made for distinct competitions and originally judged differently. The event outcomes were low-probability but had more

predictable sub-outcomes upon which they were also judged. Hence, the forecasts were multi-dimensional, complicating any evaluation. The events were football matches in the English Premier League. The forecasts were of exact scoreline outcomes. We compare these with implied probability forecasts using bookmaker odds and a crowd of tipsters, as well as point and probabilistic forecasts generated from a statistical model suited to predicting football match scorelines. By evaluating these sources and types of forecast using various methods, we conclude that forecasts of this type are strange. We argue that regression encompassing is the most appropriate way to compare point and probabilistic forecasts, and find that both types of forecasts for football match scorelines generally add information to one another.

What is the optimal strategy of aggregation for forecasting sales? Time series forecast reconciliation by region, product category, and channel.

Presenter: Carla Freitas Silveira Netto

While some companies still struggle to gather, store and analyse data necessary to make better predictions, others are worried with increased requirements for data minimization and anonymization. This scenario raises questions about which variables are important to gather, and the resources necessary to do so. An important topic for academia and practice is how to improve forecasts, having limited access to data. This difficulty is specially perceived by small and medium-sized enterprises (SMEs) that do not have the resources to handle such data. In this paper we consider such difficulties and propose forecasting strategies based on sales data in different aggregations criteria and structures. We introduce to the marketing field the optimal reconciliation approach to forecast sales. We explain why it has advantages over other approaches and suggest which type of marketing variables should be used in such strategy. The paper compares aggregation criteria using both hierarchical and grouped time series, applying the most basic and prevalent type of variables existing in our field, that is, marketing mix variables. The proposed method is generalizable to all types of goods, not only the packaged goods, most commonly studied in marketing literature. Our paper indicates whether product category, channel type or region (geographic location) works best alone or combined when using the optimal reconciliation approach. Our research allows to run sales forecasting more efficiently, using open-source tools.

Analyzing Mortality Bond Indexes via Hierarchical Forecast Reconciliation

Presenter: Han Li

In recent decades, there has been significant growth in the capital market for mortality and longevity linked bonds. Therefore, modeling and forecasting the mortality indexes underlying these bonds have crucial implications for risk management in life insurance companies. In this paper, we propose a hierarchical reconciliation approach to constructing probabilistic forecasts for mortality bond indexes. We apply this approach to analyzing the Swiss Re Kortis bond, which is the first "longevity trend bond" introduced in the market. We express the longevity divergence index associated with the bond's principal reduction factor in a hierarchical setting. We first adopt time series models to obtain forecasts on each hierarchical level, and then apply a minimum trace reconciliation approach to ensure coherence of forecasts across all levels. Based on the reconciled probabilistic forecasts of the longevity divergence index, we estimate the probability distribution of the principal reduction factor of the Kortis bond, and compare our results with those stated in Standard and Poors' report on pre-sale information. We also illustrate the strong performance of the approach by comparing the reconciled forecasts with unreconciled forecasts as well as those from the bottom up approach and the optimal combination approach. Finally, we provide first insights on the interest spread of the Kortis bond throughout its risk period 2010-2016.

Bayesian Forecast Reconciliation

Presenter: Florian Eckert

This paper contributes to the literature on optimal hierarchical forecast combination by introducing a Bayesian reconciliation framework. It allows for an explicit estimation of the reconciliation biases, which leads to several innovations: It is possible to use prior judgment to assign weights to specific forecasts, the occurence of negative reconciled forecasts can be ruled out and prediction intervals can be easily computed. The method is compared to existing reconciliation techniques using a comprehensive hierarchical dataset of Swiss goods

exports, grouped by geographical and categorical characteristics.

A feature-based framework for detecting technical outliers in water-quality data from in situ sensors

Presenter: Priyanga Dilini Talagala
Outliers due to technical errors in water-quality data from in situ sensors can reduce data quality and have a direct impact on inference drawn from subsequent data analysis. However, outlier detection through manual monitoring is infeasible given the volume and velocity of data the sensors produce. Here, we propose an automated framework that provides early detection of outliers in water-quality data from in situ sensors caused by technical issues. We compare two approaches to this problem: (1) using forecasting models; and (2) using feature vectors with extreme value theory.In the forecasting models, observations are identified as outliers when they fall outside the bounds of an established prediction interval. For this comparison study we considered two strategies: anomaly detection (AD) and anomaly detection and mitigation (ADAM) for the detection process. With ADAM, the detected outliers are replaced with the forecast prior to the next prediction, whereas AD simply uses the previous measurements without making any alteration to the detected outliers.The feature-based framework first identifies the data features that differentiate outlying instances from typical behaviours. Then statistical transformations are applied to make the outlying instances stand out in transformed data space. Unsupervised outlier scoring techniques are then applied to the transformed data space. An approach based on extreme value theory is used to calculate a threshold for each potential outlier. This threshold calculation process starts by computing a boundary using a subset of data containing half of the observations with the smallest outlier scores and then tests for potential outliers in the remaining subset. This approach successfully identified outliers involving abrupt changes in turbidity, conductivity and river level, including sudden spikes, sudden isolated drops and level shifts, while maintaining very low false detection rates.The proposed framework was evaluated using two data sets obtained from in situ sensors in rivers flowing into the Great Barrier Reef lagoon. The proposed framework is implemented in the open source R package 'oddwater'.

Probabilistic forecasting models for NO2 concentrations

Presenter: Sebastien Perez Vasseur
Air quality is an increasingly alarming issue in cities around the world and thus there is an urgent need to reliably forecast high concentration episodes of certain pollutants in the air. In this sense, NO2 is one of the most worrisome pollutants, as it has been proven its direct implication in a variety of serious health affections. In some cities, high concentration episodes for NO2 are dealt with by authorities through traffic restriction measures, amongst others, which are activated when air quality deteriorates beyond certain thresholds. Probabilistic forecasting (as opposed to deterministic or point forecasting, which is designed to predict a single expected value), allows for the prediction of the future full distribution function for any magnitude. In the case of NO2 forecasting, this is a key feature which allows for the calculation of the future probabilities of exceeding the protective thresholds set by the authorities, and further cost-effectiveness analysis.In this study, we have extended previous research on forecasting future NO2 concentrations with the implementation of four different probabilistic predictive models: a probabilistic version of k-nearest neighbors, quantile random forests, linear quantile regression and quantile gradient boosting. We have tested those models against a wide set of forecasting horizons (up to 60 hours), and we have studied the quality of these forecasts.Furthermore, we have modified those models by applying semi-parametric approach to the output of the original models. We show how this proposal improves the results and speeds up training and prediction times. In our experiments, quantile gradient boosting is the best performing model as it provides the best results for both the expected value and the full distribution. We also provide visualizations of the output of the models that shed light into the shortcomings and advantages of each model.

Probabilistic forecasting of an air quality index

Presenter: Jooyoung Jeon
In many countries, air pollution has emerged as a major issue affecting human health, with the resulting

burden on health systems having economic and political implications. Urban air pollution is believed to be the cause of more than a million premature deaths worldwide each year. Respiratory illnesses are the main health risk factor, but the prevalence of heart disease, stroke and cancer is also increased. Air quality indices are widely-used to summarise the severity of the level of a set of pollutants, with a traffic light signal often used to provide a visual indicator. The index is a convenient measure used by policy makers, but is also used, on a day-by-day basis, by health professionals and the public, especially those with a history of respiratory conditions. Forecasts of the index are typically produced each day for lead times up to several days ahead. The predictions are usually provided by meteorologists using atmospheric models that have chemistry features incorporated. Probabilistic forecasting from such models is not straightforward, and hence is very rare. In this paper, we consider the use of time series models to produce density forecasts for the index. Our approach involves the fitting of a multivariate model to a set of six pollutants. To capture the dependencies between the pollutants in a practical way, we use an empirical copula. Our empirical work uses hourly data from South Korea, where more than half the population are considered to be exposed to dangerous levels of pollutants.

Does Money Growth Predict Inflation? Evidence from Vector Autoregressions Using Four Centuries of Data

Presenter: Pär Österholm

In this paper, we add new evidence to a long-debated macroeconomic question, namely whether monetary aggregates have predictive power for inflation or, put differently, whether monetary aggregates Granger cause inflation. We study this issue by employing vector autoregressive models to unique data. Using a historical dataset – consisting of annual Swedish data on money growth and inflation ranging from 1620 to 2018 – we conduct analysis both within- and out-of-sample. The models are estimated using different assumptions and methods: Models with constant parameters and covariance matrix are estimated using both maximum likelihood and Bayesian methods; using state-of-the-art Bayesian methods, we also estimate models with drifting parameters and/or stochastic volatility. Our results indicate that within-sample analysis – based on Wald tests and marginal likelihoods for the models estimated with maximum likelihood and Bayesian methods respectively – tends to provide strong evidence in favour of money growth Granger causing inflation. In our out-of-sample analysis, the forecasts from various univariate and bivariate models are assessed based on forecast precision (in terms of root mean square forecast errors and accompanying tests) and, for Bayesian models, predictive likelihoods. This evidence turns out to be more mixed.

Forecasting food-at-home inflation using weather data

Presenter: Julia Regina Scotti

Food inflation is the most unpredictable component of CPI and the one that mostly affects people's lives, especially the poor. One of the reasons for food inflation volatility is its dependence on weather variations. In developing countries, all these three facts are more pronounced: First, food prices are more volatile due to the higher prevalence of fresh over processed foods; Secondly, fresh foods are more susceptible to weather variations not only because they are fresh, but also because production is usually less technological than that of processed foods; Finally, in developing countries, food-at-home represents a higher share of the already limited household budget. This has two significant consequences, the weight of food-at-home inflation is larger in developing countries CPI, and it is more difficult for people to accommodate food price increases. Therefore, we research whether meteorological data can help forecast food-at-home inflation. We study the case of Brazil, a developing country with rich historical daily weather data and a reliable monthly inflation index published 24 times per year. We retrieved data for the last 18 years. As method, we use the lasso and Random Forest because they handle well high dimensional models. All our estimations are pseudo-out-of-sample, and we use, as benchmark, a direct estimated AR with lag order selected by BIC. Using a validation set, we choose to use climate normals averaged over the last 30 and 90 days. Our results show that the weather data improve forecasts when compared to the benchmark for each of the 1- to 7-month horizons considered. On average over the horizons, the ratio to benchmark of the Root Mean Squared Error (RMSE) in the holdout set was 0.70 for the lasso model and 0.73 for the Random Forest model. Additionally, we use our forecasts of food-at-home inflation to forecast headline inflation. We find that, again, we could

improve forecasts for all seven horizons with both models. In the holdout set, the average RMSE ratio to benchmark for the lasso was 0.87, and for the Random Forest, 0.91. Our findings suggest that weather data might substantially improve inflation forecasts in developing countries.

Forecasting Inflation for the Long Run

Presenter: Jamus Lim
Given the abysmal track record at short-run inflation forecasting, attempts to forecast longer-run inflation appear to be fools' errands. Yet monetary authorities across the globe appear to be willing to commit themselves to precisely this seemingly-impossible task; inflation-targeting central banks are routinely expected to do so, which suggests that independent central banks either do believe that they are able to anticipate and attain—at least on average—a given inflation rate through the adoption of a nominal anchor, or that such targets offer sufficiently stable focal points for private-sector expectations that monetary authorities are willing to risk their credibility to post ex ante targets. In this paper, we ask the question of whether monetary authorities do indeed induce more accurate long-run inflation forecasts when they commit to a specific inflation target, or if they should instead rely on projections derived from model or agent-based signals of inflationary pressures. More specifically, we consider a typical forecasting horse-race between models, markets, and surveys, but instead of benchmarking such forecasts only to the standard random walk or a simple autoregressive model, we examine whether medium-to-long-term deviations from the declared target rate of inflation-targeting central banks are greater or smaller than alternative projections of long-run inflation, including pseudo out-of-sample model forecasts, financial market-implied inflationary expectations, or forecasts drawn from survey measures. We then go on to ask whether the credibility of central banks may be compromised as a result of persistent failure to attain their inflation mandate.

Spread of Zika virus in the Americas

Presenter: Ana Pastore y Piontti
In February 2016 the World Health Organization declared the Zika virus (ZIKV) epidemic in the Americas a Public Health Emergency of International Concern. During that period, it was unknown the time and location of introduction of the virus in the region, and even less was known about its timeline, as symptoms, if any, are mild and very similar to other endemic diseases in the area such as dengue and chikungunya.In order to answer these questions, we use a data-driven global stochastic epidemic model to project past and future spread of the ZIKV in the Americas. We use the Global Epidemic and Mobility model (GLEAM), that has been shown to work very well on previous instances such as the H1N1 influenza pandemic and the 2014 West African Ebola outbreak, now extended to model vector-borne diseases for the first time. The model has high spatial and temporal resolution, and integrates real-world demographic, human mobility, socioeconomic, temperature, and vector density data. We estimated that the first introduction of ZIKV in Brazil likely occurred between August 2013 and April 2014. This results was later corroborated through phylogenetic analysis. We provided simulated epidemic profiles of incident ZIKV infections for several countries in the Americas through February 2017. The ZIKV epidemic was characterized by slow growth and high spatial and seasonal heterogeneity, attributable to the dynamics of the mosquito vector and to the characteristics and mobility of the human populations. We projected the expected timing and number of pregnancies infected with ZIKV during the first trimester, and provided estimates of microcephaly cases assuming different levels of risk as reported in empirical retrospective studies [1].[1] Zhang et al., Proceedings of the National Academy of Sciences May 2017, 114 (22) E4334-E4343; DOI: 10.1073/pnas.1620161114

Stacking Probabilistic Time Series Forecasts of Infectious Disease Outbreaks using Gradient Tree Boosting

Presenter: Nicholas Reich
Accurate and reliable forecasts of infectious disease dynamics can be valuable to public health organizations that plan interventions to decrease or prevent disease transmission. A great variety of models have been developed for this task, using different model structures, covariates, and targets for prediction. Experience has shown that the performance of these models varies; some tend to do better or worse in different seasons

or at different points within a season. We considered a range of ensemble methods that each form a predictive density for a target of interest by stacking predictive densities from component models. In the simplest case, equal weight is assigned to each component model. In the most complex case, we use gradient tree boosting, with appropriate regularization, to estimate weights that vary by region, prediction target, week of the season when the predictions are made, a measure of component model uncertainty, and recent observations of disease incidence. We applied these methods to predict measures of influenza season timing and severity in the United States, both at the national and regional levels, using three component models. We trained the models on retrospective predictions from 14 seasons (1997/1998–2010/2011) and evaluated each model's prospective, out-of-sample performance in the five subsequent influenza seasons using the log score. In this test phase, medium-complexity ensemble methods showed optimal performance that was on average similar to the best of the component models, and offered reduced variability in accuracy across seasons than the component models. Regularization of weight estimation improved ensemble forecast stability. Ensemble methods offer the potential to deliver more reliable predictions to public health decision makers, for whom variability in forecast accuracy is a large concern.

Large Scale Tourism Demand Forecasting

Presenter: Haiyan Song
The presentation focuses on an automated forecasting system that is developed to predict the demand for tourism within Asia Pacific. This system consists of more than 2000 econometric models specifically estimated for all source markets of the 40 destinations within the region. The main forecasting results over a period of 7 years are presented and evaluated. The presentation further discusses the challenges faced by the forecasting team and the lessons learned from the forecasting exercise. Policy implications and future research directions are also addressed.

Cross-temporal coherent forecasts for Australian tourism

Presenter: Nikolaos Kourentzes
Key to ensuring a successful tourism sector is timely policy making and detailed planning. National policy formulation and strategic planning requires long-term forecasts at an aggregate level, while regional operational decisions require short-term forecasts, relevant to local tourism operators. For aligned decisions at all levels, supporting forecasts must be 'coherent', that is they should add up appropriately, across relevant demarcations (e.g., geographical divisions or market segments) and also across time. We propose an approach for generating coherent forecasts across both cross-sections and planning horizons for Australia. This results in significant improvements in forecast accuracy with substantial decision making benefits. Coherent forecasts help break intra- and inter-organisational information and planning silos, in a data driven fashion, blending information from different sources.

Visualizing the consequences of forecasts, logistics and costs using R and Shiny

Presenter: Stephan Kolassa
Forecasts are not an end in themselves, but contribute to different decisions, such as retail shelf replenishment. Other input data to these processes are logistics, e.g., pack sizes or replenishment schedules, and costs for over- and understock. The interplay between these different drivers and levers is not always easy to understand. We will demonstrate an R/Shiny tool to visualize the implications of forecasts, logistical settings and cost assumptions in the context of fresh item forecasting and replenishment. This tool helps understand the comparative influence of the different ingredients, the sensitivity to inputs (e.g., of the costs of over-/understock, which are often hard to assess) and where to focus efforts for improvements.

Retail sales forecasting with meta-learning

Presenter: Shaohui Ma
Sales forecasting in retail often concernswith the tasks to generate forecasts for thousands of products over a number ofstores. It is hard to believe that there is a model which can do the best forall products in all stores and at all times, as product demand patterns inretail are varies among stores and products, and also are

ever-changing acrosstime. We present a meta-learning framework based on a deep learning model forretail sales forecasting, which first learns the knowledge on which model in alarge pool of models predicts well for what demand patterns before theforecasting and under what conditions during the forecasting periods, and thenwe use the knowledge to select appropriate model to generate forecasts for theproduct. We conduct a series of forecasting experiments with a large retaildataset, and report the forecasting comparisons between the meta-learning and anumber of benchmarks to explore the value of meta-learning in retail salesforecasting.

Forecasting at Scale

Presenter: Trevor Sidery
For any business, the challenge of creating reliable forecasts is fundamental to having a stable and efficient operation. Tesco is one of the largest grocery retailers in the world, with 3500+ stores in the UK alone; having a robust scalable automated forecasting system is a requirement for good decision making. We have been moving from a scenario where each team in Tesco was making their own forecasts, to having a single framework that allows expertise from one group to inform others. We will discuss the many challenges in producing such a system, as well as the progress we have made in setting ourselves up with a versatile system that can iterate quickly as new techniques are developed both internally and from academia.

Issues in Retail Forecasting: Research and Practice

Presenter: Robert Fildes
The retail industry is rapidly changing, in part throughdevelopments in on-line retail activity. This presentation first considers thedifferent decision problems that retailers must face; shorter life cycles,increased promotional intensity and on-line activity by both organizations and consumers make modelling and forecasting more challenging. Which methods workand under what circumstances? This presentation reviews the research literature,identifies successes and the many gaps; and in the light of current retailpractice, identifies paths for future research.

Forecasting oil and natural gas prices with futures and threshold models

Presenter: Marek Kwas
Crude oil and natural gas are key sources of energy in the globaleconomy, which explains why there is vast literature focusing on thedynamics of their spot and futures prices. We contribute to this literatureby evaluating the predictive power of futures for both commodities,within a regime dependent framework. Using weekly data over the period1999-2018, we show that the dynamics of the crude oil and natural gasmarkets are different in contango than in backwardation regimes. Next,we evaluate the accuracy of forecasts for the nominal prices over theperiod 2009-2018, for horizons ranging from one week to one year. Weconsider two sets of threshold models and their linearcounterparts. The first set of direct forecast models assumes that thespot price gradually adjusts to the spread between spot and futuresprices. The second set of threshold vector autoregressive modelsexplains the joint dynamics of both prices within a multivariateframework. Finally, we consider models with financial marketregressors, which take into account the evidence from the literaturethat energy commodity prices react to the dynamics of exchange rates,stock indices, interest rates, uncertainty indices, precious andindustrial metal prices.Our main finding is that futures are relatively good predictors ofspot prices. The accuracy of futures-based forecasts turns out to bebetter than of those from the random walk model, with the gains morepronounced for the crude oil than for the natural gas prices. Thissuggests that the common practice of financial institutions, to relyon futures prices as predictors for future spot prices, isjustified. Additionally, for the natural gas marketdirect forecasts or VAR models improve the accuracy of forecasts evenfurther. Moreover, we find that the forecast accuracy ofnonlinear threshold models is lower than that of the linear ones. Ourinterpretation is that the potentially better specified thresholdmodels, as evidenced by in-sample fit to the data, provide lessaccurate forecasts due to higher estimation error. In other words,their significantly better in-sample fit does not pay offout-of-sample. The last finding is that adding information on thedynamics of financial variables does not lead to better accuracy. Thissuggests that the information content of futures contractsalready incorporates the information on present and past dynamics ofother financial variables.

Ensemble-based approaches and regularization techniques to enhance natural gas consumption forecasts

Presenter: Erick de Oliveira
This paper puts forward a novel forecasting approach combining Bootstrap aggregating (Bagging) algorithms, time series methods and regularization techniques to obtain highly accurate forecasts of natural gas consumption across different countries. In addition, a new variant of Bagging, in which the set of classifiers is built by means of a Maximum Entropy Bootstrap, is proposed. The methodologies are used to forecast several steps ahead monthly gross inland consumption data from different major and minor gas-consuming countries in Europe. A comparative out-of-sample analysis is then conducted using different performance metrics. The obtained results attest that the suggested approaches can substantially improve forecasting accuracy, outperforming important benchmarks in the forecasting literature. Findings and policy implications are further discussed.

A multi-granularity heterogeneous combination approach to crude oil price

Presenter: Jue Wang
Crude oil price forecasting has attracted much attention due to the potential impact on commodities markets as well as for nonlinear complexity in forecasting model. Forecast combination is a well-established technique for improving forecasting accuracy and the primary benefit is the ability to provide constituent forecasts using diverse information. An increasing amount of research focuses on efficiently generating individual forecasts for combination, specifically for crude oil prices. Granular computing (GrC) is about processing complex information entities, called information particles, which appear in the process of data abstraction and knowledge acquisition from data. Referring to the idea of GrC, complex distribution data can be transformed into local data by feature selection methods and a multi-granularity approach is proposed in this paper. Combining the forecasts in different granular spaces, we propose a multi-granularity heterogeneous combination approach to improve the forecast accuracy. Firstly, we introduce various feature selection techniques, such as Randomized-Lasso and Lasso, to identify key factors that affect crude oil price and construct different granular spaces. Secondly, distinct feature subsets identified by different feature selection methods are introduced to generate individual forecasts with three popular forecasting models including Linear regression (LR), Artificial neural network (ANN) and Support vector machine (SVR). Finally, the final forecasts are obtained by combining the forecasted value from individual forecasting model in each granular space and the optimal weighting vector is achieved by artificial bee colony (ABC) techniques. The experimental results indicate that feature selection technique can identify key factors in stock and flow series, macroeconomic, financial series and price series that affect crude oil prices, reduce computational complexity and increase model interpretation. Simultaneously, it can lower LR, SVR, and ANN by up to 18.6%, 33.6%, and 4.2% in terms of MAPE, respectively. Comparing with best single model, the novel multi-granularity heterogeneous combination approach based on ABC can reduce RMSE, MAE and MPAE by 18.9%, 15.7% and 23.9%, respectively. Specifically, it is also superior to single-granularity heterogeneous and multi-granularity homogenous approaches based on ABC with a reduction of 12.6% and 8.3% of MAPE on average, respectively. The study helps in accurately capturing the trends of crude oil price trends, which can offer guidance for governments, industry and individuals for decision making and budgetary planning.

Predicting Natural Gas Pipeline Alarms

Presenter: Colin Quinn
This paper provides a method to forecast natural gas pipeline pressure alarms. Forecasting natural gas pipeline pressure alarms help control room operators maintain a functioning pipeline and avoid costly down time.Natural gas production companies use pipelines to transport natural gas from the extraction well to a distribution point. As gas enters the pipeline and travels to the distribution point, it is expected that the gas meets certain specifications set in place by either state law or the customer receiving the gas. If the gas meets these standards and is accepted at the distribution point, the pipeline is referred to as being in a steady-state. If the gas does not meet these standards, the production company runs the risk of being shut-in, or being unable to flow any more gas through the distribution point until the poor-quality gas is removed. Being shut-in is costly and time consuming for the production company. For the pipeline to become

functional again, the unexpected gas must either be diffused with gas further down the line or flared from the system entirely.The internal pressure of the pipeline measured at the distribution point is the variable forecasted to help controllers deliver acceptable gas in a safe and reliable way. Sensors are used to collect real-time pressure information from within the pipe, and alarms are used to alert the control operators when a threshold is exceeded. If operators fail to keep the pipeline's pressure within an acceptable range, the company risks being shut-in or rupturing the pipeline.Predicting pressure alarms enables operators to take appropriate action earlier to avoid being shut-in and is a form of predictive maintenance.We forecast alarms by using an autoregressive model (AR) in conjunction with alarm thresholds. The alarm thresholds are defined by the production company and are occasionally adjusted to meet current environment conditions. A regression-based approach to predicting alarms is used in favor of a classification-based approach because the alarm thresholds can be changed after the algorithm is deployed. The benefit of a regression-based model is in its output, since it can be used to diagnose the state of the pipeline rather than just classifying an alarm being imminent.The results of the alarm forecasting method show that we can accurately forecast the pressure time series up to a 30-minute time horizon. This translates into true positive rates that drop of linearly from around 100% at one minute to approximately 65% at a 30-minute forecast horizon. This means that at 30 minutes, we correctly forecast 65% of the alarms, but have no false positives.

Bayesian functional forecasting with locally-autoregressive dependent processes

Presenter: Matteo Ruggiero
Motivated by the problem of forecasting demand and offer curves, we introduce a class of nonparametric dynamic models with locally-autoregressive behaviour, and provide a full inferential strategy for forecasting time series of piecewise-constant non-decreasing functions over arbitrary time horizons. The model is induced by a non Markovian system of interacting particles whose evolution is governed by a resampling step and a drift mechanism. The former is based on a global interaction and accounts for the volatility of the functional time series, while the latter is determined by a neighbourhood-based interaction with the past curves and accounts for local trend behaviours, separating these from pure noise. We discuss the implementation of the model for functional forecasting by combining a population Monte Carlo and a semi-automatic learning approach to approximate Bayesian computation which require limited tuning. We validate the inference method with a simulation study, and carry out predictive inference on a real dataset on the Italian natural gas market.

Bayesian nonparametric methods for macroeconomic forecasting

Presenter: Maria Kalli
Vector autoregressive (VAR) models are the main work-horse model for macroeconomic forecasting, and provide a framework for the analysis of complex dynamics that are present between macroeconomic variables. Whether a classical or a Bayesian approach is adopted, most VAR models are linear with Gaussian innovations. This can limit the model's ability to explain the relationships in macroeconomic series. We propose a nonparametric VAR model that allows for nonlinearity in the conditional mean, heteroscedasticity in the conditional variance, and non-Gaussian innovations. Our approach differs to that of previous studies by modelling the stationary and transition densities using Bayesian nonparametric methods. Our Bayesian nonparametric VAR (BayesNP-VAR) model is applied to US macroeconomic time series, and compared to other Bayesian VAR models. We show that BayesNP-VAR is a flexible model that is able to account for nonlinear relationships as well as heteroscedasticity in the data. In terms of short-run out-of-sample forecasts, we show that BayesNP-VAR predictively outperforms competing models.

Variational Forecasts of Observation-driven Models

Presenter: Alex Cooper
We demonstrate that approximate Bayesian forecasts constructed using variational methods are both accurate and extremely fast to compute. We focus on the class of generalized linear autoregressions (GLAR), a simple but large class of observation-driven forecasting models with wide practical application. We present both theoretical and numerical evidence that variational forecasts can have the same predictive power as

'exact' (ie unapproximated) Bayesian forecasts constructed using sampling methods.Our theoretical results provide conditions for variational GLAR forecasts to be asymptotically equivalent to exact forecasts, in two respects. First, we show that variational approximate probabilistic forecast distributions 'agree' with their exact counterparts in the sense of statistical merging. Second, we show that in the large-data limit, variational forecasting performance is no worse than for exact forecasts, as measured by a proper scoring rule.We also present numerical experiments on variational forecasts for different data types, including continuous, binary, and count data. In each case, the experiments confirm our theoretical results and demonstrate that variational forecasts can be several orders of magnitude faster than forecasts based on sampling methods.

Bayesian forecasting for high-dimensional state-space models: A variational approach

Presenter: Matias Quiroz

The Bayesian approach to forecasting averages the data likelihood over the posterior density of the unknown parameters. In a state-space setting, this entails computing the joint posterior density of the state vector and the static model parameters, a task that is particularly difficult outside the linear Gaussian state-space model, especially when the state vector is high-dimensional. We propose to approximate the posterior density by a multivariate Gaussian variational approximation, in which the variational parameters to be optimized are a mean vector and a covariance matrix. The number of parameters in the covariance matrix grows as the square of the number of model parameters. Hence, in a high-dimensional setting, it is necessary to find simple yet effective parameterizations of the covariance structure. We do so by approximating the joint posterior distribution of the high-dimensional state vectors by a dynamic factor model, with Markovian time dependence and a factor covariance structure for the states. This gives a reduced dimension description of the dependence structure for the states, as well as a temporal conditional independence structure similar to that in the true posterior. We illustrate the usefulness of our approach for prediction in a multivariate stochastic volatility model for financial returns via a Wishart process.

On the empirical performance of Bayesian approaches for spare parts inventory forecasting

Presenter: M. Zied Babai

Spare parts demand forecasting and inventory control are increasingly challenging tasks for inventory managers. Different approaches have been proposed in the literature to address these issues, originating in both the frequentist and Bayesian paradigm. Under the frequentist approach, the lead-time demand distribution is assumed to be known and its parameters are constant and estimated using a forecasting method. The Bayesian approach assumes that the parameters of the lead-time demand distribution are random and characterised with a prior distribution. Although a considerable amount of work has been developed under the frequentist approach, there is very little research that is dedicated to the Bayesian one. Moreover, most of this latter research has been developed under the Poisson demand distributional assumption. However, it is well known that the Poisson distribution is associated with certain limitations to modelling spare parts demand. In particular, it cannot reflect lumpy demands and Compound Poisson distributions are rather more appropriate for such demand patterns. Addressing this gap in the literature constitutes the objective of our work. In this paper, we propose a new Bayesian method based on the Compound Poisson distribution to model spare parts demand. By means of an empirical investigation on demand data of more than 8,000 stock keeping units in the aerospace and automotive sectors, the performance of the proposed method is evaluated and compared to the Bayesian method with Poisson demand as well as some forecasting methods originating in the frequentist paradigm.

Forecasting intermittent demand of an airline's spare parts using incrementally trained neural networks

Presenter: Athanasios Tsadiras

In this paper, a new Neural Network method is proposed that incorporates incremental training to forecast intermittent demand. The use of incremental training is chosen because such a training is suitable for time

series especially in cases where the training data becomes available gradually over time. The advantages of such a training to forecast intermittent demand is thoroughly studied in this paper. Its performance is examined based on 10 years of demand data of 5,135 spare parts of the aircrafts' fleet of an airline company.The structure of the proposed Neural Network has three layers. In the first layer the inputs are based on: (1) the demands at different time steps before the target period, (2) the number of periods separating the last two nonzero demand transactions before the target period, and (3) the number of periods between the target period and first zero demand before the target period.The second layer is the hidden layer consisting of H hidden neurons, with typical value of H to be 3. The third layer is that of the output of the proposed Neural Network that represents the predicted value of the demand for the target period.To evaluate the proposed method, a comparison with the following well-established parametric methods is done: a) Single Exponential Smoothing, b) Croston's method, and c) Syntetos and Boylan Approximation. It is also compared with another Neural Network method, that of Gutierrez et al. (2008) to verify the advantages of incremental training. The evaluation of all the above methods is based on both a) their forecast accuracy and b) their inventory efficiency in leading to the trade-off between holding volumes and backordering volumes.The results of the evaluation show that Single Exponential Smoothing and the Incrementally Trained Neural Networks outperform all other examined methods, including the non-incrementally trained Neural network method of Gutierrez et al. (2008) in both a) forecast accuracy and b) inventory efficiency.

What about those sweet melons? Using mixture models for demand forecasting in retail

Presenter: Ivan Svetunkov

The area of retail forecasting has experienced the rapid growth of interest in academic society over the last several years, which was followed by an increase in the number of publications in this direction. One of the reasons for this is because retailers have a lot of data to work with, and there are many distinctive features in it, which differ from other areas of application. For example, the sales of product scan demonstrate multiplicative seasonality, spikes and level changes which are influenced by many factors, such as price changes, sales promotions and calendar events. At the same time, there can be periods with no demand, which usually correspond to zero or missing values in the data. So working with such data becomes a challenging task, and there are not many models that can be efficiently applied in this context. For instance, the conventional multiple regression does not perform well, as it does not handle the multiplicative effect together with zero sales. The dynamic forecasting approaches,such as exponential smoothing or intermittent demand methods, do not take the seasonality of the demand occurrence into account and thus fail to produce accurate forecasts. This means that there is a need in a statistical model that would be able to perform better given these data features. Yet, there already exists an instrument that is used in statistics and econometrics for similar problems. It is called a "mixture model", and it allows separating the demand into two parts, namely "the demand occurrence" and "the demand sizes". This gives the necessary flexibility, taking the retail data features into account. In this talk we will discuss several examples of such models and compare their forecasting performance on retail data.

Intermittent Demand Forecasting with Deep Temporal Point Processes

Presenter: Caner Turkmen

Intermittent demand is a common and challenging problem in many real-world forecasting scenarios. Often, such data arises as a result of discretizing an event sequence localized in continuous time, i.e. available with timestamps, to a preconceived grid in time. This process not only leads to discretization error but also introduces model misspecification problems often encountered in intermittent demand scenarios.However, such data sets can be addressed equally well with a different set of models, temporal point processes (TPP). In this presentation, we compare and contrast TPP and explore their application toward intermittent demand forecasting use-cases. We present a recent introduction to literature, deep TPP, which tackle computational concerns related to traditional conditional-intensity TPP while yielding favorable predictive accuracy. Finally, we present preliminary empirical findings on intermittent demand data; comparing point process models to their discrete-time counterparts.

A new metric of consensus for Likert-type scale questionnaires: An application to forecast the unemployment rate using the degree of agreement in consumer expectations

Presenter: Oscar Claveria

Inthis study we present a metric of consensus for Likert-type scales. The measuregives the level of agreement as the percentage of consensus among respondents.The proposed framework allows to design a positional indicator that gives thedegree of agreement for each item and for any given number of reply options. Inorder to assess the performance of the proposed metric of consensus, in aniterated one-period ahead forecasting experiment we test whether the inclusionof the degree of agreement in consumers' expectations regarding the evolutionof unemployment improves out-of-sample forecast accuracy in eight Europeancountries. We find evidence that the degree of agreement among consumerscontains useful information to predict unemployment rates in most countries.The obtained results show the usefulness of consensus-based metrics to trackthe evolution of economic variables.

Effects of information on the market performance - application in forecasting.

Presenter: Michal Chojnowski

With recent progress in communication technology and media, customers have been given an infinite source of information regarding the product of their interest.On the other hand, there is more and more news articles, which describes how 'fake news' are influencing society.It seems that nowadays information is becoming more influential. Question remains - by how much?Author in his paper presents an artificial neural network, which combines Internet activity data (here: Google Trends) with economic and market fundamentals. The network allows to include non-observable factors into forecasting models, both in a linear and non-linear manner (ARIMA and LSTAR). Factors are derived from Google Trends using Structural Equation Modelling. Method not only brings better performance of the forecasts (by 30%), but also provides an analysis of news' importance on the market.It enables both economists and business to better understand how given market is shaped by unobserved fundamentals (in economics known as sunspots).With further research the method can be used to predict upcoming up- and downturns on the market.

How does financial information impact judgment in forecasting company stocks?

Presenter: Ariel Cecchi

Investors, particularly those who trade every day, are constantly confronted with short-term forecasting decisions based on graphically presented time series. These decisions are often influenced by market news and the information investors have access to. The news valence (positive or negative) appears to have an impact on market tendencies (Hayo & Neuenkirch, 2012; Reeves & Sawicki, 2007), buying or selling behaviour (Barber & Odean, 2008; Sobolev, Chan & Harvey, 2017), or the understanding of graphical information (Tuckett, 2012). It seems that people tend to ignore good news but react to bad news (Galati & Ho, 2003; Harvey & Reimers, 2013). In addition, judgmental forecasts are often subject to trend damping (Bolger & Harvey, 1993; Eggleton, 1982; Lawrence & Makridakis, 1989); the tendency to underestimate future values for upward trends (under-forecasting) and overestimate them for downward trends (over-forecasting). We investigated the influence of news on judgmental forecasting in an experimental study where we manipulated the influence of market news presented before the forecast of stock prices. In this study, participants were presented with two sets of 55-day time series company stocks and asked to indicate their predictions of the price for the next 5 days. The first set of graphs was forecast based solely on the information in the graphs. Before the second set of time series, participants were provided with news regarding some of the companies they were going to forecast. These pieces of news resulted from a combination of either positive or negative and integral (related to the company's performance) or incidental (related to the companies' products) information. I will present our findings on the impact of market news on trend damping and discuss their implications.

Nudging for improved forecasts of future expenses and savings

Presenter: Shari De Baets

Many households do not have the necessary savings to deal with unexpected shocks, such as a car breakdown

or a household member falling into unemployment (Grinstein-Weiss, Russell, Gale, Key, & Ariely, 2017). Consequently, many people suffer from economic insecurity and are at risk for future economic problems (Weller & Logan, 2009). The issue is further complicated by people's behavioural tendencies: they are more oriented towards the present than the future (Frederick, Loewenstein, & O'Donoghue, 2002; Tam & Dholakia, 2011), prefer instant gratification over long-run benefits (Ainslie, 1992; Angeletos, Laibson, Repetto, Tobacman, & Weinberg, 2001; O'Donoghue & Rabin, 1999), underestimate future rise in expenses compared to rise in income and underestimate the risk of unexpected expenses in the near future compared to those they experienced in the near past (Howard, Hardisty, Sussman, & Knoll, 2016). Additionally, the broader literature on judgment and decision making teaches us that people suffer from a general optimism bias (Weinstein & Klein, 1996) and tend to ignore pessimistic scenarios and focus on the positive ones (Newby-Clark, Ross, Buehler, Koehler, & Griffin, 2000). Consequently, these factors all add up to an under-estimation when forecasting future expenses, increasing the risk of households without savings to fall into economic disarray. The objective of this study is therefore to test, with the use of scenarios, how to best aid people's judgment on predicting financial risks and nudge them towards the best decisions. The main research question of this project focusses on the effect of risk information on participant's forecasts of expenses and savings. We will investigate whether (1) participants adjust their forecast based on the information they receive, (2) whether this adjustment is different for expenses and savings and (3) what the effect is of different types of risks (low-risk, high-risk, control/no-risk). We hypothesize that individuals given high-risk scenarios will adjust their expense forecast and savings forecast, but insufficiently so (due to an anchoring effect and a general optimism bias; Tversky & Kahneman, 1974; Weinstein & Klein, 1996); while people given low-risk scenarios will decrease their expenses and increase their savings (more so than expenses will increase in the high risk scenario). The neutral scenario serves as a control.

Economic Sentiment in Europe: Disentangling Private Information from Public Knowledge

Presenter: Axel Lindner

This paper addresses a general problem with surveys asking agents for their assessment of the state of the economy: answers are highly dependent on information that is publicly available, while only information that is not already publicly known has the potential to improve a professional forecast. We propose a simple procedure to disentangle the private information of agents from knowledge that is already publicly known (that is common knowledge) for surveys that are structured like that for the European Commission's consumer sentiment indicator. Utilizing the fact that this survey asks agents how they assess their own as well as the general economic prospects we show that the weight rational agents give to their private information is higher for their assessment of their own prospects than for their assessment of the general economic prospects of the economy. As a consequence, the indicator representing private information is simply a linear function of the difference between the assessment of the general prospects and the assessment of the private household's own economic prospects. Private information extracted from the consumer survey in this way can be used as input for forecasts. We show that this method works quite well in practice for some EU economies such as Germany.

Tests of Conditional Predictive Ability: Some Simulation Evidence

Presenter: Michael McCracken

In this note we provide simulation evidence on the size and power of tests of predictive ability described in Giacomini and White (2006). Our goals are modest. First, we establish that there exist data generating processes that satisfy the null hypotheses of equal conditional predictive ability. We then consider various parameterizations of these DGPs as a means of evaluating the size and power properties of the proposed tests. While some of our results reinforce those in Giacomini and White (2006), others do not. We recommend against using the fixed scheme when conducting these tests and provide evidence that very large bandwidths are sometimes required when estimating long-run variances.

Analyzing Different Facets of Forecast Quality through Decompositions of Loss Functions

Presenter: Marc-Oliver Pohle

Sound decision-making in a wide variety of situations relies on high-quality forecasts of economic variables. However, what does quality mean with regard to economic forecasts? The econometric forecast evaluation literature mainly focuses on forecast accuracy, i.e. on measuring the distance between forecasts and realizations via loss functions. Besides that, a large literature on forecast optimality testing has emerged. I propose to augment this approach by using measures of forecast quality and the so-called Murphy decomposition originating from the meteorological literature, provide a proper theoretical framework, discuss the estimation of the decomposition via nonparametric regression and apply it for the first time to mean and quantile forecasts from economics. The Murphy decomposition, which I discuss for a general type of loss functions, namely consistent scoring functions, expresses the scoring function as the sum of an uncertainty component of the observations - which can be regarded as a generalized variance or entropy, the (mis-)calibration and the resolution. Miscalibration measures the statistical compatibility of forecasts and observations and is closely related to forecast optimality, while resolution measures the ability of the forecasts to discriminate between outcome values, in other words their informational content. The decomposition illuminates the relationship between the forecast-accuracy and the optimality-testing approach to forecast evaluation by adding a measure of the informational content to the equation. It shows that both the informational content and the optimal use of information determine the size of forecast errors and reveals which of the two is the driving force. Thus, it can be crucial in understanding the sources of forecast errors and consequently lead to improved forecasts. Firstly, I set up a proper theoretical framework with an explicit time series structure. This enables me also to show connections between formerly seemingly unrelated strands of the literature and to reassess the sharpness principle put forward by Tilmann Gneiting and his coauthors. Secondly, I discuss the estimation of the decomposition terms, employing kernel regressions to estimate the occurring conditional functionals. Thirdly, I apply the decomposition to evaluate mean forecasts of US inflation and GDP growth from the Survey of Professional Forecasters and quantile forecasts of inflation derived from the Bank of England fan charts. For the former I find that miscalibration plays a negligible role and that resolution drives the mean squared forecast error. Resolution drops to zero already for two-quarter ahead forecasts, putting into question the usefulness of the forecasts from then on, even though they are calibrated and thus do not to contain systematic mistakes. For the latter, a quite different picture arises with a gradual rise in miscalibration as well as a gradual fall in resolution. Here, even though the informational content naturally falls with rising forecast horizon, it would suffice to provide useful forecasts for several quarters ahead if the forecasts were calibrated. The applications show the potential of the proposed methodology in complementing the existing approaches to forecast evaluation in econometrics, providing valuable insights and at the same time being sufficiently simple to implement and interpret.

Evaluation of Government Budget Forecasts

Presenter: Neil Ericsson

This paper reviews the literature on the evaluation of government budget forecasts, outlines a generic framework for forecast evaluation, and illustrates forecast evaluation with an empirical analysis of different US government agencies' forecasts of the US federal debt. Techniques for forecast evaluation include comparisons of mean square forecast errors, tests of stability, and tests of bias and efficiency. These techniques are generally applicable, including to forecasts of components of the government budget, and to budgets from municipal, state, provincial, and national governments. Evaluation of forecasts is fundamental for assessing the forecasts' usefulness; and evaluation can indicate ways in which the forecasts may be improved.

A comparison of wind speed probabilistic forecast via quantile regression models

Presenter: Soraida Aguilar

Considering the current global climate changes that around the world are being experienced, the need arises to increase the use of renewable sources as energy resource that support electrical systems and the energy transition around the world. Thus, wind speed represents a strategic alternative to attend these requirements. Nonetheless, the high variability of this sort of resource affect the energy production planning for different

horizonts. To cope to this situation a probabilistic model based on local quantile regression is used and compared with Box & Jenkins models. The results are encouraging and show that the model is robust and efficient when consider a suitable hypothesis of the errors distribution.

Online Distributed Learning in Wind Power Forecasting

Presenter: Benedikt Sommer
Forecasting wind power generation up to a few hours ahead is of utmost importance for the efficient operation of power systems and for participation in electricity markets. Recent statistical learning approaches exploit spatio-temporal dependence patterns among neighbouring sites but their requirement of sharing confidential data with third parties may limit their use in practice. This explains the recent interest in distributed, privacy preserving algorithms to high-dimensional statistical learning, e.g., for auto-regressive models. The few approaches that have been proposed are based on batch learning, hence being potentially computationally expensive while not allowing to accommodate nonstationary characteristics of stochastic processes like wind power generation. This paper closes the gap between online and distributed optimization by presenting two novel approaches that recursively update model parameters while limiting information exchange between wind farm operators and other potential data providers. The first one is an online variant of the Alternating Direction Method of Multipliers (ADMM), while the other one is inspired by mirror descent approaches, being less complex and computationally lighter. A simulation study allows underlining the convergence and tracking ability of both approaches. In addition, a case study using a large dataset for 311 wind farms in Denmark confirms that online distributed approaches generally outperform existing batch approaches.

Probabilistic Solar Power Forecasting: Long Short-Term Memory Network vs. Simpler Approaches

Presenter: Vinayak Sharma
The high penetration of volatile renewable energy sources such as solar, make methods for coping with the uncertainty associated with them of paramount importance. Probabilistic forecasts are an example of these methods, as they assist energy planners in their decision-making process by providing them with information about the uncertainty of future power generation. Currently, there is a trend towards the use of deep learning probabilistic forecasting methods. However, the point at which the more complex deep learning methods should be preferred over more simple approaches is not yet clear. Therefore, the current article presents a simple comparison between a long short-term memory neural network and other more simple approaches. The paper consists of training and comparing models able to provide one-day-ahead probabilistic forecasts for a solar power system. The paper makes use of an open source dataset provided during the Global Energy Forecasting Competition of 2014 (GEFCom14).

A non-parametric approach to wind power forecast

Presenter: Paula Maçaira
The wind power generation depends on the wind speed, air density, turbine rotor area and many other technical features. In order to simplify the mathematical calculation, wind power curves are used to estimate how the power output from a wind turbine varies with a steady wind speed. However, this approximation is one of the uncertainty sources of the wind power forecasting models, since: (i) it assumes wind speed distributed on fixed intervals and (ii) a given speed will always generate the same power. Therefore, this work aims to develop a methodology to forecast wind power generation based on clustering techniques, to estimate the wind speed intervals, Kernel Density Estimation (KDE) approach, to define the probability density function (PDF) for each interval, and the Markov Chain Monte Carlo method to simulate and forecast from the PDF. The datasets, with real measures, from two wind farms with twelve turbines each in Brazil, are used to test the proposal.

Measuring the amount of disorder in financial networks

Presenter: Angi Roesch
There are multiple ways to characterize and quantify dynamic aspects of information flow in networks of

financial markets. Our network perspective assumes that information flow to one market triggers a cascade of perturbations changing the network's intrinsic properties with respect to informational balance. We show how certain notions from the theory of information entropy, applied to network dynamics running forward and backward in time, can contribute to diagnosing turbulences, thus detecting network anomalies and systemic risk. The empirical examples of our study include networks of the euro area as well as networks of international asset markets, using daily returns on stock indices as input.

An early-warning indicators' framework for Polish banking sector

Presenter: Marcin Lupinski
The Polish banking is perceived as one of the soundest Europe's banking sectors. However, during recent 30 years it experienced several serious pressures on its stability. In this paper we present set of leading macroprudential indicators (early warning indicator, EWIs) based on statistical data gathered directly from Polish banks. These composite measures are constructed with help of conceptual framework based on Markov switching approach. We assess their relative performance with standard aggregated indicators based on credit-to-GDP gap approach. While the traditional aggregated EWIs perform well over long horizons, the Markov switching indicators allow to catch signals of incoming shocks within one-year perspective. The paper addresses also the issue how the prepared EWIs can be used to support both micro- and macroprudential policy-makers.

Designing the early warning model that do not incorporate ex-post knowledge of crisis mechanisms

Presenter: Kamil Jonski
Effective operation of macroprudential tools like CCyB critically depends upon reliable early warning models. To get a clue how such models would work in real-life crisis prevention, the 'out-of-sample' evaluations are carried out. In such exercises, models are estimated using data available in the real-time. However, ex-post knowledge of crisis mechanisms is silently embedded in the model specification and during the variable selection. For example, the subprime crisis became fairly 'predictable', once housing sector variables – absent in pre-2008 works – entered the models. We argue that contrary to the academics, macroprudential practitioners always attempt to predict crisis whose causes and dynamics is inherently unknown. We refer to this as a 'knowledge gap'. To immune our early warning model to this problem, we employ only two variables utilized in the pre-crisis literature (Borio, Lowe, 2002): debt-to-GDP ratio and stock prices index. We demonstrate, that using rolling correlation between these two variables – instead of their levels, gaps or growth rates – enabled extracting warning signal within policy relevant horizon. The proposed model provides reasonably good forecasting performance (almost equal during in-sample and out-of-sample evaluation), signal stability and conceptual interpretability. Quantitative out-of-sample evaluation is complemented by the case studies of historic financial crisis driven by specific mechanisms: Nordic Crisis and Subprime Crisis (banking crisis ignited by housing debt) as well as Asian Crisis (banking crisis ignited by currency crisis). Although past performance is not indicative of future success, obtained results suggests that proposed approach could be useful in spotting the general build-up of imbalances in financial system – without looking at the specific variables that ex-post allowed to 'predict' these events in out-of-sample exercises (like house prices or exchange rate dynamics).

Forecasting Defaults of US Banks in Early Warning Systems

Presenter: Claudio Antonini
We have developed an Early Warning System tool in Shiny to analyze financial aspects of counterparties, industries, and countries. As part of that tool, the paper introduces a number of models to determine the probability of default of U.S. banks within 6 months. The dataset builds on previous work by Gogas, Papadimitriou and Agrapetidou (IJF, 2018) and consists of 555 banks that failed between 2000 and 2018, and a variable number of solvent banks (from 555 to 8,300) in the same period, using quarterly data. The models analyzed are logistic regression (ridge and LASSO), support vector machines, and decision trees. A preliminary conclusion is that different regimes can be seen before and after 2008. For that reason, the

coefficients of the models change to accommodate these regimes. Commonalities among the variables used in the different models will be discussed. As a new development, an immediate application of the dataset is to refit and simplify James Ohlson's probabilistic bankruptcy 1980 model–which was not applicable to non-financial entities–extending it to U.S. banks. In all cases, either with the machine learning models or Ohlson's, the accuracy of the fits is in the mid to upper 90s.

Judgment in forecasting

Presenter: NA
Forecasting may be based on unaided judgment (judgmental forecasting), on a formal procedure usually implemented by computer (computer-based forecasting), or on a combination of judgmental and formal approaches (computer-aided forecasting). Most research on judgment in forecasting has been concerned with how it is used in judgmental forecasting. However, as computer-aided forecasting has become more prevalent in business contexts, research has increasingly focussed on judgmental adjustment of formal forecasts. Recently, researchers have also come to appreciate that judgment may be involved even in computer-based forecasting. We know that different formal approaches typically produce forecasts (from the same data set) that vary in accuracy. Users of forecasting support systems may use their judgment to decide which formal approach to use. I'll discuss research in these three areas: judgmental forecasting, judgmental adjustment of formal forecasts, and judgmental selection of formal forecasts.

Peeking inside FFORMS: Feature-based FORecast Model Selection

Presenter: Thiyanga Talagala
Features of time series are useful in identifying suitable forecast models. Talagala, Hyndman & Athanasopoulos (2018) proposed a classification framework, called FFORMS (Feature-based FORecast Model Selection), which selects forecast models based on features calculated from the time series. The FFORMS frameworkbuilds a mapping that relates the features of a time series to the "best" forecast model using the random forest algorithm. In this paper we explore what is happening under the hood of the FFORMS framework and thereby gain an understanding of what features lead to the different choices of forecast models and how different features influence the predicted outcome. This is accomplished using model-agnostic machine learning interpretability approaches. Partial-dependency plots are used to visualize both main and interaction effects of features. The results of this study provide a valuable insight into how different features and their interactions affect the choice of forecast model selection. This gives a more refined picture of the relationship between features and the choice of forecast model which is particularly valuable for ongoing research in thefield of feature-based time series analysis.ReferencesTalagala, TS, RJ Hyndman & G Athanasopoulos (2018). Meta-learning how to forecast time series. Working paper 6/18. Monash University, Department of Econometrics and Business Statistics

Time series forecasting based on automatic feature extraction

Presenter: Feng Li
Feature-based time series representation has attracted substantial attention in a wide range of time series analysis methods. Recently, the use of time series features for forecast model selection and model averaging has been an emerging research focus in the forecasting community. Nonetheless, most of the existing approaches depend on the manual choice of an appropriate set of features. Exploiting machine learning methods to automatically extract features from time series becomes crucially important in the state-of-the-art time series analysis. In this paper, we introduce an automated approach to extract time series features based on images. Time series are first transformed into recurrence images, from which local features can be extracted using computer vision algorithms. The extracted features are used for forecast model selection and model averaging. Our experiments show that forecasting based on automatically extracted features, with less human intervention and a more comprehensive view of the raw time series data, yields comparable performances with the top best methods proposed in the largest forecasting competition M4.

GRATIS: GeneRAting TIme Series with diverse and controllable characteristics

Presenter: Yanfei Kang

The explosion of time series data in recent years has brought a flourish of new time series analysis methods, for forecasting, clustering, classification and other tasks. The evaluation of these new methods requires a diverse collection of time series benchmarking data to enable reliable com- parisons against alternative approaches. We propose GeneRAting TIme Series with diverse and controllable characteristics, named GRATIS, with the use of mixture autoregressive (MAR) models. We generate sets of time series using MAR models and investigate the diversity and coverage of the generated time series in a time series feature space. By tuning the parameters of the MAR models, GRATIS is also able to efficiently generate new time series with controllable features. In general, as a costless surrogate to the traditional data collection approach, GRATIS can be used as an evaluation tool for tasks such as time series forecasting and classification. We illustrate the usefulness of our time series generation process through a time series forecasting application.

Averaging probabilistic forecasts of day-ahead electricity prices across calibration windows

Presenter: Tomasz Serafin

Most day-ahead electricity price forecasting (EPF) studies focus on developing models that better represent the inter-variable dependencies or implementing faster and more efficient estimation algorithms in order to increase the accuracy of the forecasts. However, somewhat surprisingly, authors have almost completely ignored the problem of finding the optimal length of the calibration window for their models. The typical approach is to select ad-hoc, a 'long enough' window. The problem is that there is no consensus in the literature as to the length of this window – training samples as short as two weeks and as long as six years have been used.We build on a recently introduced concept for point forecasts, that averaging day-ahead electricity price predictions over calibration windows of various lengths produces better results than selecting, even ex-post, only one 'optimal' window length. We introduce a new, well-performing weighting scheme for averaging forecasts, which assigns larger weights to windows that have performed well in the past. Furthermore, we extend the approach of averaging predictions over different calibration windows to probabilistic forecasts. In order to test our methodology, we consider four autoregressive structures as forecasting models and utilise datasets from three major power markets that differ in the geographic location and the generation mix: Nord Pool (Norway, Denmark, Finland, Sweden), EPEX (Germany, Austria) and PJM (the United States). Based on point forecasts, we construct prediction intervals (PI) using Quantile Regression Averaging (QRA) on multiple types of inputs. We consider averaging point forecasts and using the results as a single-vector input for quantile regression as well as using QRA on a multi-vector input of point forecasts. We also focus on the length of the calibration window for probabilistic forecasts – ranging from 7 to 364 days. Moreover, we look if gains in the accuracy of interval forecasts can be achieved by averaging quantiles obtained from different calibration windows.

Regularization for quantile regression averaging. A new approach to constructing probabilistic forecasts

Presenter: Bartosz Uniejewski

Over the last two decades electricity price forecasting has gradually become a fundamental decision making process for energy companies. The literature concerning point forecasts is abundant. However, the recent introduction of smart grids and renewable integration requirements has had the effect of increasing the uncertainty of future prices. This has created a need not only for point but also for probabilistic electricity price forecasts (EPFs). Nowotarski and Weron (2018) claim that now probabilistic forecasts are more important for energy systems planning and operations than ever before. To address this issue we examine possible accuracy gains from forecast averaging in the context of probabilistic EPFs. We propose a new method for constructing prediction intervals (PIs), which utilizes quantile regression (QR), a pool of point forecasts of individual models and LASSO regularization. We show that our approach can significantly outperform the standard Quantile Regression Averaging (QRA) technique of Nowotarski and Weron (2015) when the set of individual point forecasts is large.

Revisiting the jackknife method for construction of prediction intervals – application to electricity market

Presenter: Katarzyna Maciejowska
The probabilistic forecasting in electricity markets has gained a lot of attention over the last few years (see Nowotarski, Weron (2018) for a comprehensive review). Various methods have been applied in the field: bootstrap method, quantile regression, conformal predictions and forecast averaging. This research revisits the jackknife approach and extends it to allow for multiple data removal. It is shown that this procedure is robust to a model overspecification. A simple Monte Carlo experiment is conducted to evaluate the behavior of the method for problems with a high number of parameters to the sample size ratio. The results indicate that the procedure gives prediction intervals with the coverage closer to the nominal one than traditional bootstrap or jackknife. The approach is finally applied for probabilistic forecasting of electricity prices. The results confirm the initial results and show that it is superior to other sampling methods in terms of robustness.

Forecasting performance of seasonal tourism market share models in the EU-15

Presenter: Egon Smeral
Export market shares are key figures for evaluating the ex post competitiveness of tourism destinations (countries). The performance of tourism exports provide important information for determining tourism policy regarding the international competitiveness of the country specific tourism industries. However, studies analyzing tourism market shares are very rare, especially in cases where the explanation of the short term seasonal dynamic and forecasting issues are the main focus. To narrow the research gap, we developed tourism export market share models for each of the EU-15 countries, focusing only the summer season and short term fluctuations. The main explaining factors are the export prices of the country and of the EU-15. We analyzed the seasonal tourism market share performance not only with the usual longitudinal approach using data capturing all quarters, but also with a model showing distinct relationships only for the summer season (here the third quarter). In other words, for the specific countries, the third quarter is understood as its own time series. The estimation results showed that, for most of the EU-countries, the export prices of the specific countries and the competitive price set of the EU-15 are the main factors to explain the short term market share fluctuations. However, the price elasticities of the two approaches were significantly different. In the case of the forecasting performances, neither approach showed a clear superiority as each model type performed only in fifty percent of the cases better than the other type. These results indicate that a combination of the two approaches might optimize the forecasting performance. Key words: export market shares, short term fluctuations, export prices, forecasting performance, combined forecasts.

Forecasting big tourism flows at attractions and destinations using mixed-frequency data

Presenter: Shanshan Lin
This study aims to explore the potential determinants of tourism demand at the attraction and city levels from the perspectives of using multiple big data sources. The number of visits to Hangzhou (China), and its two major tourist attractions (the Qiandao Lake and West Lake) will be forecast by considering a list of indicators, such as, OTA bookings, online hotel room rates, search engine query volumes, traffic flows on the airport, expressways and subway, and weather conditions. Due to different frequencies of arrivals data and its regressors, a mixed-data sampling (MIDAS) modelling approach will be applied and compared to the traditional time series models. The contributions of each indicator in predicting the visits to the city and its attractions will be evaluated and compared.

Tourism Forecasting Combining Time-varying Parameter and Mixed Frequency Techniques

Presenter: Long Wen
This study aims to apply a hybrid model combining both the time-varying parameter (TVP) and mixed-frequency modeling techniques for tourism demand forecasting. This idea is derived from two aspects. First,

as the internet development, the online tourist behavior data such as search engine data and online review data have demonstrated ability of helping with tourism forecasting (such as Bangwayo-Skeete and Skeete 2015; Pan, Wu and Song 2012; Yang, Pan, Evans and Lv 2015). Since this kind of data are often of high frequency, mixed frequency modeling technique has been introduced to tourism forecasting by which the variables with different frequencies are allowed to be estimated in one model (such as Bangwayo-Skeete and Skeete 2015). This is an effective way to avoid the loss of information in the high frequency data (Wu, Song and Shen 2017). Second, TVP technique has shown excellent performance in tourism forecasting due to its advantage of allowing the model parameters to vary instead of keeping constant over time, which is obviously more consistent with reality (Song, Li, Witt and Athanasopoulos 2011; Song, Witt and Jensen 2003). Though the advantages of the TVP and mixed frequency modelling techniques have been investigated separately, no study has combined the merits of the both for tourism forecasting. This study therefore aims to examine if this combination is an effective way to improve forecasting accuracy of tourism demand. Particularly, two mixed-frequency modelling approaches, namely mixed-data sampling (MIDAS) and mixed-frequency VAR are combined with the TVP approach respectively to generate the TVP-MIDAS and TVP-VAR models. These two new models are applied for Hong Kong inbound tourism forecasting exercise. Explanatory variables include tourist income, tourism price, substitute price and Google search data. All data are on quarterly frequency except Google search data which is with monthly frequency. The competition models include traditional TVP, MIDAS, VAR and AR models. Different forecasting horizons from one- to four-step-ahead will be examined. The results of this study will provide practitioners with useful information and guidance on adopting high frequency data for tourism forecasting more effectively.

Pharmaceutical forecasting with low-volume data: A case-study in Greece

Presenter: Nikolaos Athiniotis

In modern businesses, successful decision making is closely connected to accurate forecast. This is particularly true in processes related to budgeting and target setting, i.e., short and long-range planning. However, forecasting accuracy can be significantly affected from data availability and quality. For instance, when dealing with short-life-cycle or low-volume products, forecasting may become rather challenging. To cope with these issues, forecasters have to apply special pre-processing techniques and exploit information extracted from similar products or cases, a process which is not typically supported by off-the-shelf solutions. This paper demonstrates and discusses these problems by presenting an indicative case-study of a Greek pharmaceutical company involving 167 SKUs of 5 types, distributed across 70 regions. It is shown that the proposed extrapolation approaches enhance forecasting performance when compared to traditional forecasting methods, providing accurate forecasts even for cases where data are abnormal or scarce. The main advantage of the suggested framework is identified for short seasonal SKUs, for which capturing seasonality poses a challenge, and low sales volume areas, where uncertainty becomes dominant.

Automatic selection of predictors for ARIMA models

Presenter: Patrícia Ramos

Typically, organizations make their forecasts by combining simple univariate statistical modeling with managerial judgement to include relevant new information. Yet, as the volume of forecasts increases, and the promotions and marketing activities become more intensive, that practice is not efficient, and an automated and reliable multivariate forecasting system is required. This is particularly true in the retailing context, where forecasts are needed for thousands of Stock Keeping Units (SKUs), at a high frequency. The problem is further exacerbated by multiple promotional events being active simultaneously, introducing multicollinearity and a complex variable selection problem, even when halo and cannibalization effects are not considered.Although in principle regression models can capture the required effects, in particular when augmented with ARIMA errors, this is not practical, as the users need to manually specify models for the different SKUs to avoid the aforementioned challenges. To achieve an automated solution, that scales appropriately in the retailing context, we investigate the usefulness of models for high-frequency seasonality coupled with variants of lasso regression. The empirical evaluation is conducted using data from a Portuguese retailer. To evaluate the performance of the competitive models, a robust experimental design is implemented. This design is based on a rolling forecasting origin scheme that considers several forecast horizons and calculates suitable error

measures. Our findings provide important insights for the use of appropriate models in the retail business that can reflect improvements on decision-making and the importance of developing models that include key marketing features.

Demand forecasting model taxonomy for short seasonal time series using cross-sectional information

Presenter: John Boylan
In the context of short-term demand forecasting, it is often difficult to estimate seasonality accurately due to short data histories. However, companies usually have multiple products with similar seasonal demand patterns. A possible solution in this case is to use the components of several time series of products from a homogeneous family, thus estimating seasonal coefficients based on cross-sectional information. In this research we propose a framework based on vector exponential smoothing, that exploits different features of time series and improves the efficiency of the estimates of the components. We propose a taxonomy of vector exponential smoothing seasonal models based on this framework, taking into account the commonality of seasonal components, smoothing parameters and seed values. Empirical findings will be discussed to assess the performance of the models within this framework, relative to each other and a benchmark model.

A generalized threshold stochastic volatility model incorporating with realized measures

Presenter: Feng-Chi Liu
A generalized threshold stochastic volatility (THSV) model jointly modeled with realized measures (RMs) is considered in this study. Advancing from trading technology, high frequency data is used to construct RMs for accurate volatility forecasting. This study proposes a threshold-type RM equation to jointly model with a generalized THSV model. Based on the Bayesian approach, parameters of the proposed model are estimated by the designed Markov chain Monte Carlo method. In the real data analysis, we employ the risk measures of value-at-risk (VaR) and expected shortfall (ES) to evaluate the performance of volatility forecasting. The results show that the proposed model can produce more accurate volatility forecasting than the model with a simple RM equation.

Adaptive Dynamic Model Averaging for House Price Forecasting

Presenter: Alisa Yusupova
In this work, we propose a novel, adaptive dynamic model averaging methodology (ADMA) which extends the approach of dynamic model averaging (DMA) first proposed in Raftery et al. (2010), and modified in Koop and Korobilis (2012). DMA involves two critical parameters, called forgetting factors. The first forgetting factor determines the rate of adaptation of the estimated coefficients of each dynamic linear model (Kalman filter) in the pool of considered models. The second forgetting factor determines the rate at which past performance is discounted in the determination of the weights assigned to each model in the pool. Lower values of these forgetting factors correspond to more aggressive discounting of past data and thus enable a more rapid adaptation to changes in the data generating process. On the other hand, if the data generating process is static, higher values of these forgetting factors are appropriate since these enable more stable estimation and hence better performance. In Raftery et al. (2010) and Koop and Korobilis (2012), both forgetting factors are set by the user and are static throughout the execution of the DMA algorithm. However, as the results in these papers show, the performance of DMA is heavily affected by the choice of these parameters. Moreover, in real world applications, the assumption of a constant rate of adaptation may be too strict since changes in the data generating process may occur smoothly over some time periods and abruptly over others. To address these issues, we develop a stochastic gradient descent scheme to adapt the forgetting factor involved in the update of the coefficients of each dynamic linear model. This allows the forgetting factor of each model to adapt to changes in the evolution of the data generating process. We further replace the weighting scheme employed in the traditional DMA with a parameter-free model aggregation algorithm that has finite time performance guarantees with respect to the optimal sequence of forecasting models (V'yugin and Trunov, 2019).As an empirical application, we examine the performance of ADMA in forecasting UK

house prices at a regional and national level. The latest boom and bust in real estate markets and its role in the Great Recession has generated considerable interest in the dynamics of house prices. International organizations and central banks are increasingly concerned about monitoring the developments in real estate markets across the world (e.g., the IMF established the Global Housing Watch in 2014), and policy makers have attached a larger weight on the importance of these markets in financial stability and the real economy (see, e.g., the 2018 BoE Financial Stability Report and Stress Test Results). Our results document that the proposed ADMA algorithm produces more accurate house price forecasts than the benchmark methods in this area, as well as competing DMA specifications. These results are consistent with recent evidence that suggests that the relationship between real estate valuations and conditioning macro and financial variables displayed a complex of time-varying patterns over the last decades (Aizenman and Jinjarak, 2014).

Density Forecast Combination using Bayesian Global VAR Models and Bayesian Neural Networks

Presenter: Roberto Morales-Arsenal

Forecasting with large sets of data is one of the most important fields in Macroeconomics. In general, if the information is good and relevant and if the analyst is able to build a useful model, a greater amount of information should produce a greater accuracy in the predictions made. In order to enlarge the information set we use two models: 1) Bayesian Global Vector Autorregresive Models (combining BVARs and Global VARs models) and 2) Bayesian Neural Networks (BNNs). In both models, we use different prior specifications to obtain the density forecast, providing a whole description of the uncertainty associated with a point forecast, applied to the harmonised index of consumer prices (HICP) in the euro area. In a second step we combine the two obtained density forecasts from the two models used in order to obtain the final density forecast. For this task, we use bayesian techniques in a similar way to the one used in Hall and Mitchell (2004). For the evaluation and calibration of density forecasts we use: 1) the probability integral transform (PITs) and 2) scoring rules. The final result shows strong evidence in favor of the combined approach instead of the individual ones.

Forecasting US recessions using large scale data sets and machine learning algorithms

Presenter: Baris Soybilgen

Identifying the start of a US recession as early as possible is crucial for both market participants and policy makers. However, it is a well-known fact that it is hard to forecast US recessions accurately in a timely fashion. Usually the best models to predict US recessions are simple probit models including yield spread and stock market returns. In this study, we analyze whether using large scale data sets with state art of machine learning algorithms improve over basic models. We use three main types of machine learning algorithms: tree based algorithms, neural network based algorithms and support vector machine variants. As the large scale data set, we use the FRED-MD data set enriched with financial variables. We further use dimension reduction techniques such as factor models, autoencoders and lasso models before feeding large scale data sets into the machine learning algorithms. In our results, we can not identify a single dominant model for all forecasting horizons and all sample period. Our results also show that in various cases simple models can compete with highly sophisticated models.

Forecasting UK GDP growth with large survey panels

Presenter: Nikoleta Anesti

Extensive research undertaken by academics, central bankers and financial market participants has been dedicated on finding the best set of predictors to be used in forecasting key macroeconomic variables. In this paper, we consider the task of short-term forecasting of GDP growth for the UK economy at monthly frequency using large panels of survey and macro data. We review standard linear approaches for regularisation and dimension reduction, like Lasso and Ridge Regressions, Principal Components (PC) and Partial Least Squares (PLS) and provide empirical evidence on the predictive ability of these methods using the newly released measure of monthly GDP growth as a target variable. We also use the same forecasting environment

as a laboratory to test techniques drawn from the machine learning literature. In particular, we apply a set of supervised algorithms which little if any attention have been given in the context of macroeconomic forecasting so far. Random forest is one example, which is a popular ensemble method that can be used to build predictive models for both classification and regression problems. Since its inception, random forest has gained substantial ground and is frequently used in many machine learning applications across various fields of the academic community. The main driver for its wide adaptation is the unexcelled accuracy among other machine learning algorithms. One of its main features is that it is supported by an efficient calculation algorithm offering a useful framework for analysis of big datasets. Based on the structures of this method it can handle a large number of input variables without any correlation restrictions. We also make use of Support Vector regressions (SV) which is a family of non-linear, large-margin models. Support Vector regressions estimate a separating hyperplane that achieves minimum separability between the input data and the target. We evaluate soft-margin support vector regressors using linear, radial basis function, polynomial, and sigmoid kernels, and retain the model configuration yielding optimal performance. To select the hyper parameters of the evaluated kernels as well as the cost hyper parameter (related to the adopted soft margin), we resort to a grid-search ten-fold cross-validation. In an out of sample evaluation exercise, we find that the PLS, the Lasso and the Ridge regressions have better predictive content compared to our benchmark specification, a simple AR(1) model, for up to 12 steps (months in our case) ahead. Among the machine learning techniques evaluated in this exercise, we find that the random forecast yields the most accurate predictions in terms of RMSFE over the same period. Finally, we examine the usefulness of 'soft' information from business and consumer surveys, as opposed to information from more general macroeconomic time series in the spirit of the so-called Stock and Watson dataset and we find that the richer information set only improves our predictions for longer horizons forecasts, whereas for the shorter horizons (1 and 3 steps ahead) the surveys-only models outperform.

Understanding Automatic Time Series Modeling and Forecasting: A Case Study of Real GDP, the Unemployment Rate, and Leading Economic Indicators

Presenter: John Guerard

Thetime series modeling approach has evolved from the Box and Jenkins (1970)approach for the identification, estimation, and forecasting of stationary (orseries transformed to stationarity) series, through the analysis of the seriesautocorrelation and partial autocorrelation functions, to the world of Clive Granger (2001) and causality testing, to current applications of automatic timeseries modeling and forecasting of Hendry and Doornik (2014). What paths should a practitioner seek to follow in analyzing time series models and using time series model forecasts?In this case study, we applied the Hendry and Doornik automatic time series PCGive (OxMetrics) methodology to several well-studied macroeconomics series,real GDP and the unemployment rate. We report that the OxMetrics andAutoMetrics system substantially reduce regression sum of squares measuresrelative to a traditional variation on the random walk with drift model. Themodeling process of including the Leading Economic Indicator in forecastingreal GDP has been addressed before, but our results are more statisticallysignificant. A similar conclusion isfound for the impact of the LEI and weekly unemployment claims series leadingthe unemployment rate series. Such results should excite practitioners!

Conditional forecast and scenario analysis in DSGE models: how to improve causal interpretation.

Presenter: Olga Croitorov

Policy makers are often interested in assessing the outcomes related to possible future events. However, usual conditional forecast techniques remain silent about the causal economic effects of constrained future paths. This paper proposes a two-steps algorithm allowing for scenario analysis in presence of technical forecast assumptions while preserving structural interpretation. First, we infer the path of the structural shock associated with the scenario variable. Second, we apply the Kalman filter to a conditioning set augmented by the shock from the first step. We show the algorithm at work employing the European Commission's Global Multi-country model that is regularly used to identify the key drivers of the Euro Area macroeconomic variables forecast. Given the unprecedentedly low levels of worldwide nominal interest rates, we set up several scenarios associated with a path reversal to historical average. In particular, we evaluate the sensitivity of

Euro Area variables forecast to the set of technical assumptions and alternative shocks triggering the increase in global interest rates. Namely, we consider an increase in productivity, a surge in risk premium and a slowdown of rest of the world growth.

Forecasting with DSGE-VAR Model: An Application for Brazilian Economy

Presenter: Fernanda Valente
The purpose of the current paper is to use a small open economy Dynamic Stochastic General Equilibrium (DSGE) model as a prior distribution in the estimation of a Vector Autoregressive (VAR) model for the Brazilian economy, following the approach proposed in Del Negro and Schorfheide (2004). The main goal is to assess the relative performance of the one-, three- and six-steps-ahead forecasts, generated by the DSGE-VAR model in relation to alternative models such as a Bayesian VAR (BVAR) with Minnesota prior and diffuse prior, and an estimated DSGE. Monthly data on output, inflation, exchange rate and interest rate of the Brazilian economy from January 2003 to December 2016 were used. A Model Confidence Set (MCS) procedure was adopted in order to compare the accuracy of the forecasts obtained by the different models. The results show that the inclusion of prior information from DSGE model is advantageous when the objective is to generate forecasts, especially in medium and long-term. According to the MCS procedure, the DSGE-VAR model is the only one contained in the set of the best models for all variables and horizons, showing the importance of incorporating information from economic theory. In addition, the BVAR model with Minnesota prior was on the set of better models in most cases, emphasizing the importance of prior information to reduce the imprecision of the forecasts of overparameterized models.

Evaluation of alternatives for modeling and forecasting non-stationarity DSGE models

Presenter: Norberto Rodriguez
This paper shows the benefits of using two approaches suggested in the literature to estimate DSGE models with observable variables that are non-stationary and do not undergo any transformation before estimation; this is achieved by decomposing variables into cyclical (driven by a DSGE model) and trend components (determined by an structure selected), whose parameters are estimated simultaneously. Using a simulation exercise with a small new-Keynesian model, the benefits of these alternative approaches are exemplified. In general, they provide estimations (parameters and IRF) closer to the real counterparts of the model. We suggest using a Bayesian forecast combination strategy which shows to have a better performance compared to traditional approaches.

Forecasting Births Using Google

Presenter: Francesco D'Amuri
Monitoring fertility change is particularly important for policy and planning purposes. New data may help us in this monitoring. We propose a new leading indicator based on Google web-searches. We then test its predictive power using US data. In a deep out-of-sample comparison we show that popular time series specifications augmented with web-search-related data improve their forecasting performance at forecast horizons of 6 to 24 months. The superior performance of these augmented models is confirmed by formal tests of equal forecast accuracy. Moreover, our results survive a falsification test and are confirmed also when a forecast horse race is conducted using different out-of-sample tests, and at the state rather than at the federal level. Conditioning on the same information set, the forecast error of our best model for predicting births is 35% lower than the Census bureau projections. Our findings indicate the potential use of Google web-searches in monitoring fertility change and in informing fertility forecasts.

The Bayesian Nested Lasso and its Application to Forecasting Problems with Mixed Frequency Data

Presenter: George Michailidis
For a number of forecasting problems, one is interested in modeling and forecasting an outcome variable

measured at a low frequency, whereas the predictors associated with it can be measured at both low and high frequencies. A typical example involving economic data is forecasting a quarterly time series such as gross domestic product (GDP), or aggregate consumption, by predictors measured at similar or higher frequencies (e.g. interest rates).To that end, mixed data sampling regression methodology has been developed, wherein low frequency data are modeled as a linear function of high frequency data with coefficients constrained to belong to a lower dimensional space. Further, more distant lags of the predictors exhibit diminishing magnitude coefficients in the model.In this work, we introduce a fully Bayesian methodology which will select both the appropriate lag for the predictors under consideration, but also estimate the model coefficients simultaneously. The lag selection procedure is based on a novel Bayesian variant of the nested lasso, where the specification of the variance of the prior distribution encourages decreasing magnitude coefficients over time. We develop an efficient sampling scheme to calculate the posterior distribution and illustrate the methodology on synthetic and real economic data.

Hybrid New Keynesian Phillips Curve for Mexico, 2002Q1-2018Q4

Presenter: Eduardo Loria Diaz
With GMM (Hansen 1982), we estimate the New Keynesian Hybrid version of the Phillips Curve (HPC) for the Mexican economy (2002Q1-2018Q4) in its closed and open economy versions. We find robust evidence that in both estimates inflation has a statistically significant and positive relationship with the indicator of the economic cycle (through the real marginal cost gap), and that this relationship is much stronger in the open economy specification due to the impact of exchange rate on labor costs and inflation.With an in-sample-forecast, we find that the open economy model is more robust, so is the out-of-sample-forecast.In both specifications, the forward-looking rule of inflationary expectations is greater than the backward-looking one, which suggests that a greater proportion of agents form their expectations making systematically correct forecasts rather than just following past inflation. However, this relative difference is much more pronounced in the estimate for a closed economy, perhaps because it is more difficult to forecast inflationary shocks that come from abroad (depreciation).We find structural change in trend and in intercept in the estimate of the open economy model in 2015Q4. Thus, through estimating by subsamples (2002Q1-2015Q3 and 2015Q4-2018Q4) we find that in recent years the forward-looking rule of inflation has gained strength while the backward-looking one has weakened, which can be accounted for by the inflationary outbreaks (explained by domestic factors and by depreciations) that have occurred since 2015Q4.In environments characterized by higher inflation, agents form their expectations taking into account, to a greater extent, the expected evolution of inflation determinants.Likewise, while in the first sub-sample the cyclical indicator of the economy is not statistically different from zero, in the second one it is, and its associated parameter is high, which indicates that before 2015Q4 inflation used to be mainly explained by expectations and, in the second period, the economic cycle has gained importance since its associated parameter is much higher than that of expectations.We find that both the relative weight of the forward-looking rule of inflation versus the backward-looking one and sensitivity of inflation to changes in the economic cycle have increased significantly in recent years.

Receiver operating characteristic (ROC) curves: What are they and what are they good for?

Presenter: Tilmann Gneiting
Receiver operating characteristic (ROC) curves are used in an astonishing range of settings where predictors for binary events are to be evaluated. I will examine and review their construction and usage from an academic perspective and move on to make explicit recommendations for practitioners. In particular, ROC curves ought to have a concave shape, and their key usage is in model development, rather than expost forecast evaluation.

The M4 Forecasting Competition: A Practitioner's View

Presenter: Chris Fry
The M4 Forecasting Competition was a huge success along many dimensions. It enabled a comparison of 60

different forecasting methods across 100,000 real-world time series. It also created the opportunity for both the participants and observers to learn from each other through open sharing of both the full set of test data as well as the solutions applied.

As practitioners who have worked on a wide range of forecasting problems, we were excited about the competition and its results, as well as the innovative ideas that were shared as a result of the competition. At the same time, we also saw some differences between the nature of the competition and the type of problems that we work on and have seen others working on. We share our suggestions for 5 themes which we observe in the forecasting community and which we would like to see reflected in future competitions to tie them more closely to today's real-world forecasting and prediction challenges. We also outline some attributes of forecasting approaches that we expect will be key success factors in prediction for those types of problems, and discuss how those resonate with the competition findings. Lastly, we compare attributes of the M4 competition data set with those of a recent Kaggle competition on web traffic time series forecasting which was hosted during the same time period.

Evaluating machine learning approaches for spare part demand forecasting

Presenter: Julius Mehringer

Predicting spare parts demand 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. This paper evaluates machine learning approaches to (i) find patterns in and (ii) forecast demands for a large dataset of master and consumption data from 1985-2017 stemming from a big manufacturer of household goods. More specifically, we apply different clustering methodologies with varying distance measures and parameter calibrations to determine products that share similarities in terms of (i) master data and (ii) consumption patterns. We use these clusters to construct predictions for "new" products where historical data is scarce using the data from similar (older) products where more data is already available. To this end, we use diffusion models and artificial neural networks. Our results indicate that this step-wise approach of combining clustering and forecasting methods yields significantly better forecasting results than a baseline model and improves the spare parts planning and controlling process.

Looking back to drive ahead: contrasting insights from meta-analyses of innovation launches with forecasting techniques and ideas for improved methods for new product demand forecast.

Presenter: Antonio Schuh

Forecasting future demand of new products and services is common problem with multiple important applications. Amongst the various techniques identified in the literature, the use of the Bass model with parameters selected based on analogies to past launches of similar innovations is one the most established. However, actual implementations raised issues on procedures and criteria for selecting analogies and on incorporating inputs from managerial judgement in order to increase confidence in the forecast results. This paper proposes principles for revised methods based on Bass and analogies. The proposal is based on insights from studies employing meta-analyses of the diffusion of new products and services and their success or lack of thereof. We analyze six comprehensive meta-analytics studies of the drivers of new product success (measured in different manners), and of empirical generalizations derived from them. These insights are compared to the process employed in two major detailed description of the analogy-based Bass, plus 8 other forecasting exercises focusing on the ICT sector. We identify that while the forecasting processes already incorporates some of the insights identified from the meta-analyses, there is room to incorporate additional ones aiming to improve the forecasting applications. In particular, there are implications for the analysis the factors to be used in the selection of p and q parameters from analogies, and also on the need of procedures for adapting them to specific circumstances. In addition, we identify a number of elements impacting substantially diffusion success that are usually not captured in innovation forecasting processes. These factors include both company-specific (and project-specific) factors and strategy characteristics such as order of entry, fit with company orientation or likelihood and intensity of competitive response. These factors that are difficult to adjust in the usual parameters, but could be incorporated as part of a structured

discussion stage with managerial decision makers aimed at improving both forecasting likelihood of success and management confidence in the forecast. The paper concludes with an outline of recommended actions to forecasting practitioners to augment models and processes with insights from the meta-analytics innovation success empirical findings.

Predicting demand for road haulage with external data: a comparison of methods

Presenter: Benedikt Sonnleitner
Road haulage companies have to plan about one week in advance, which transport capacities they need. Therefore, dispatchers manually forecast the required trucks and drivers with by means of expert knowledge. However, the human experts underestimate or overestimate the requirements that arise. This leads to costs since short-term booking of freight capacity is expensive - just like booked but unused capacity.We develop a forecasting model that predicts demand on trucks based on a variety of external factors, such as weather, public holidays, macroeconomic indicators or disturbances in other transport modes with regard to their influence on road transport demand of two road haulage companies. In particular, we use various filter and wrapper methods in order to select potential features. We then use those identified features to compare the performance of multivariate statistical methods and artificial neural networks. First results indicate that both methods improve the forecasts when comparing them to a baseline model.

Forecasting Time Series with Multiple Seasonal Patterns using a Long Short-Term Memory Neural Network Methodology

Presenter: Kasun Bandara
As sensors and data storage capabilities advance, time series with higher sampling rates (hourly, daily) are more and more common in many industries, e.g. in the utility demand industry (electricity and water usage). In such series, the presence of multiple seasonal patterns is a common observation. The accurate modelling of multiple seasonal cycles is then necessary to estimate the demand on various time horizons precisely, and may lead eventually to better demand planning and efficient resource utilization.The current state of the art to handle multiple seasonal patterns are mostly univariate forecasting methods that treat each time series as an independent sequence of observations, and forecast them in isolation. In contrast to global prediction models, where a single model is built across all the time series, the univariate models are unable to incorporate any key patterns and structures that may be shared by a group of time series. Although several unified models have been proposed to learn better under these circumstances, how to handle multiple-seasonal patterns in a set of time series has not yet been thoroughly addressed.In this preliminary research, we explore techniques that effectively account for the multiple seasonal periods present in a time series, using Long Short-Term Memory Networks (LSTM). Building on recent promising developments in the area, we propose a forecasting model that employs a global prediction model. We then aim to uncover common seasonality structures and behaviours available in a collection of time series and supplement the LSTM learning process with a series of state-of-the-art statistical time series decomposition techniques, including Mstl, stR, and TBATS.We evaluate the proposed methodology on a benchmark competition dataset and a real-world energy consumption dataset, where our method achieves competitive results.

Statistical and machine learning methods combination for improved energy consumption forecasting performance

Presenter: Guillaume Hochard, PhD
The recent M4 competition has demonstrated that the use of one forecasting method alone is not the most efficient approach in terms of forecasting accuracy. In this talk, I will review classical statistical forecasting methods and the advantage of combining them to machine learning algorithms, such as gradient boosting, for increased performance. In particular, I will focus the talk on an energy consumption forecasting use case integrating exogenous data such as weather conditions. Generalising the use of these methods can be a major help to address the challenge of electricity demand and production adjustment.

Probabilistic Forecasting of Electricity Demand using Markov Chain and Statistical Distribution

Presenter: Eralda Gjika

Electricity production is one of the key factor in the economic progress of a country. The climatic position of Albania makes the power sector heavily dependent on electrical energy produced mainly by hydropower plants. The electrical power system is divided into three main sectors: the manufacturing sector which is the Albanian Power Corporation (KESH), the Transmission System Operator (OST) and the distribution Sector by Electricity Power Distribution Operator (OSHEE). The manufacturing sector produces regarding the demand from the distribution sector. Analyzing the capacity of electricity production system in the recent years it is noticed an ability to cover the demand and exceed it which can be sold in the region. Electric energy consumption in most countries varies with time and is defined as a continuous random variable. This study presents two methodologies to estimate daily probabilities of electric demand by consumers in Albania. Data are observed daily from January 2011 to December 2018 in three basic HPP which produce the most amount of demand in the country. Since there is a noticeable difference between working and non-working days in terms of daily demand we have analyzed two subgroups. This distinction between these two classes significantly reduces periodicity. In the first approach we use a homogeneous Markov chain model to evaluate the n-step transition probability between classes. Given that our time series are characterized by periodic patterns we transform the demand time series through a normalization process and create the appropriate conditions to apply an approach based on homogeneous Markov chain (HMC). In our approach we use quartiles to classify the demand on each day in one of three classes: low, medium and high. Depending on the results and in order to increase the forecast accuracy we have taken into consideration the division of each class into subclass based on quartiles. The HMC method is used then to provide stochastic and deterministic forecasts of daily electricity demand. In the second approach, we have considered continuous probability distributions such as: normal, exponential, Weibull, log-normal, gamma, Pareto and log-logistic to fit daily data and use the "best" fit to achieve short-term forecast of demand. Maximum likelihood parameter estimation method is used to estimate the parameters of the models. Graphical tests, accuracy measures and goodness-of-fit statistics are considered to decide on the suitable statistical distribution. The results show that Gamma, Weibull and log-logistic statistical distributions are accurate to fit the level of daily energy produced. The two proposed models are important for the production analysis department, to plan an efficient scheduling by taking into consideration the possibility of a peak or exceeding the peak in the electric production system. This way, they may prevent failure probability to deliver the request and design a production strategy to optimize at the same time cost, need satisfaction and maximize profits. Key words: forecasting, probability, estimation, electrical energy, Markov chain, linear programming.

Early event classification in spatio-temporal data streams

Presenter: Sevvandi Kandanaarachchi

From the Internet of Things to social media, data streams are omnipresent in today's world. Typically, it is not the whole data stream that is of interest, but certain events that occur within the stream. For example, events of interest might be traffic congestion resulting from road accidents or a network intrusion in a high performance computing facility.We are interested in events that start, develop for some time, and stop at a certain time. Such events can be characterised by measurable properties or features, and involve a class label in most scenarios. It is a challenge to predict the class of these events while they are still developing because only partial information is available until the events are complete. For example, it is easier to differentiate a daffodil from a tulip when both are in full bloom, but more difficult to differentiate a daffodil bud from a tulip bud without resorting to other information such as characteristics of leaves.Our focus is on early event prediction in spatio-temporal data streams. Our framework encompasses an event extraction algorithm as well as two early event prediction algorithms called SAVEC and DAVEC, which predict the class of events that are still developing, using partial information. SAVEC is a static classifier suitable for stable environments while DAVEC is a dynamic classifier suited for changing environments. We test our framework on synthetic and real data and achieve better results compared to logistic regression.

Multi-way Relationships in Multivariate Time Series of Brain Activity Data

Presenter: Peter Bajorski

In the current research of human brain activity, correlations between brain regions is the most common and fundamental method used to perform this task. However, correlation describes only a two-way relationship. This work explores a new approach by analyzing multi-way relationships. Due to computational complexities, we concentrate on three-way relationships. In particular, we compare conventional two-way correlations and three-way regression models in order to assess significant component of a three-way relationship relative to already observed two-way relationship. Data transformed and processed from 3,280 MRI scans of the human brain are used in modeling and analysis. In order to measure how much advantage the three-way model has, compared to the corresponding two-way models (correlations), we use a measurement called "D value" calculated from subtracting maximum squared correlation of the two-way models from the R-square of the three-way model. We test all possible 253,460 variable selection combinations, and recursively model each scan over the time domain. We then evaluate consistency of the D-values over time in order to identify the most interesting cases of consistent three-way communications among the brain regions. This consistency is defined as the percentage of the length of time points with high D-value over the total length of a scan. We then also evaluate persistence of the three-way communications across multiple scan of the same person. This is done with variance analysis in a random effect model with "person" as a random factor. These methods provide an effective path to understand and evaluate the models as well as a new statistical perspective to analyze human brain activity. The results of this research show specific three-way relationships that have a significant advantage relative to their corresponding two-way relationships. The algorithm proposed in this paper can potentially outperform the conventional two-way correlations in exploring the activity of human brain regions.

Short-term forecasting for product prices with sporadic constant segments

Presenter: Benjamin Buchwitz

Forecasts for future price realizations are an essential component for consumer-oriented decision support systems. These systems aim to advice customers to determine and schedule buying decisions. Common application areas are online sales channels with homogeneous products such as electronic goods. In this area, the customer base is usually opportunistic about the place of purchase and highly price sensitive. Accurate price forecasts are especially valuable for technology products as they account for a major part of the e-commerce sector.Product prices of electronic consumer goods usually deteriorate over time, thus exhibit a time dependent level. Price adjustments occur irregularly and with varying magnitude leading to calmer and more active price changing phases as well as to entirely constant segments. Additionally, customers often decide on the basis of the lowest price over a selection of retailers making seller-specific pricing strategies less effective. Time series of such minimum prices exhibit similar characteristics than the ones from individual sellers but show more extreme values and shorter, more irregular segments with constant prices. In this study, we introduce a hybrid forecasting approach for historic daily minimum product prices using an exponentially-weighted markov-switching model with components for mean, variance and zero-inflation (of the first differences). Besides presenting estimation and modeling details, we compare different (global, local and component-wise) configurations of the model and discuss their contribution to forecast accuracy.Using a customized bootstrap approach, the competing forecasts are used to predict probabilities of price decline events over a pre-specified horizon. These price probabilities are used to generate recommendations to either "buy immediately" or "wait with the purchase". Applied repeatedly as part of a buying recommendation engine, the results can be used to evaluate the economic performance using real gains and losses caused by adhering to issued recommendations. Based on an empirical, asymmetric cost function and a rolling out-of-sample recommendation simulation, we quantify the economic impact of the excess modeling effort and comment on the scalability of the approach. Findings suggest that the proposed model has severe benefits over model-free benchmark strategies as well as forecasts generated by simple models that neglect some of the data properties. Thus, our developed method supports utilizing historic price information and helps generating precise and beneficial buying recommendations.

Linking forecasting to inventory for humanitarian operations: an empirical evaluation

Presenter: Marwa Hasni

The increasing frequency of humanitarian crises calls for efficient methods to improve the effectiveness of emergency response for medical items because an inventory shortage could mean a difference between life and death. On the other hand, holding more inventory may result in burning medical items which is very expensive for humanitarian organisations. Humanitarian operations are characterised by unpredictable demand, making it difficult to estimate demand for goods needed for disaster relief operations in conflict zones. Additionally, there is a lack of research linking forecasting to stock control and planning in humanitarian organisations. The current approach to deal with uncertainty in one of the major humanitarian organisations based on Geneva is to use moving average demand of the last six months as safety stock. We conduct an empirical study based on over 2000 Stock Keeping Units from three warehouses of a major humanitarian organisation located in the Middle-east and Africa. We examine the use of parametric, bootstrapping and probabilistic forecasting approaches when an order up to level inventory policy is considered. We then benchmark against the current approach used by the humanitarian organisation. The results demonstrate the gain obtained in terms of inventory cost by improving the forecast capabilities. Valuable insights are offered to demand planners and the developed approaches will be accessible through a package in R software for practitioners.

Anticipating special events in Emergency Department forecasting

Presenter: Bahman Rostami-Tabar

The Emergency Department (ED) allows patients to access a nonstop urgent medical care in the health service. Accurate forecast of daily ED attendance helps planners in the ED to better allocate available resources. It may reduce i) cost associated with calling external agencies, ii) staff sickness, iii) pressure on other health services iv) waiting time and finally v) improve the quality of service for major stakeholders. Among factors that may impact the ED attendance, special events such as public holidays and festive days play an important role. Since special events affect human behaviour, they may increase or decrease the demand for ED service. Most of studies that look at ED forecasting: i) are descriptive in nature, ii) fail reproducibility and lack methodological rigor, iii) do not consider modeling the type of special events, and iv) completely ignore the future uncertainty of attendance. We generate both point and probabilistic forecasts of daily ED attendance for the first time, which also explains the effect of the type of special events. We propose a forecasting model to generate both point and probabilistic daily forecast attendance for the first time in the context of ED. Moreover, we analyze the impact of special events on the ED attendance by considering real-life ED data from a hospital in Wales, UK. The model includes: i) Auto-regressive effect; ii) weekday effect; iii) long term trend effect; iv) special events effects and v) date effects. Using a daily ED attendance data of over 6 years from a hospital in the UK, we demonstrate the effectiveness of the proposed model on forecast accuracy. We benchmark against existing approaches: 1) Naive; 2) Auto-regressive order p; 3) exponential smoothing state space (ETS) model; and 4) Simple regression without special events effects. We analyze in-sample results to determine the effect of parameters. Moreover, we evaluate the forecast accuracy in out-of-sample using Median Absolute Error, Root Mean Squared Error for point forecasts and using Pinball score and energy score for probabilistic forecasts. We show that the proposed model outperforms benchmarks across all horizons for both point and probabilistic forecasts. We provide evidences that modeling special events will improve the forecast accuracy of ED attendance. Moreover, we describe in detail how might different type of events increase or decrease the ED attendance. Results also show that our model is more robust with increasing forecasting horizon. Our model can easily be adapted to be used not only by EDs but also by other health services. It could apply to any service such as Ambulance service or hospitals trying to generate accurate forecast by considering special events. Our model cloud also be generalized to include more type of special events if required.

Optimization approach to the low-level forecast reconciliation for two-dimensional hierarchical time series

Presenter: Igor Gusakov
When forecasting hierarchical time series using top-down (or middle-out) approach, there is always a problem of forecast distribution to the lower level. Common practice is to use lower level forecasts as a base for this distribution. For two-dimensional hierarchical time series, we propose to solve an optimization task of a special type (transportation problem) to minimally change obtained low-level forecasts and at the same time fit the restrictions: sum of the low-level forecasts should be equal to the forecast of the upper level. Although to find a solution to this problem generally means to solve a system of linear equations (which requires computational power), we found a direct analytical solution of a similar problem with fewer restrictions and then applied a simple and powerful iterative technique which keeps the solution near optimum, fit it into all restrictions and is not computational power consuming.

Non-negative forecast reconciliation for hierarchical and grouped time series

Presenter: Shanika Wickramasuriya
A problem with MinT and its variants proposed by Wickramasuriya, Athanasopoulos and Hyndman (2018) is that the coherent forecasts may result in negative values, even when all of the base forecasts are non-negative. This has become a serious issue in applications that are inherently non-negative such as with sales data. While overcoming this difficulty, we considered the analytical solution of MinT as a least squares minimization problem. The non-negativity constraints were then imposed on the minimization problem to ensure that the reconciled forecasts are strictly non-negative.Considering the dimension and sparsity of the matrices involved, and the alternative representation of MinT, this constrained quadratic programming problem was solved using three algorithms. They are the block principal pivoting algorithm, projected conjugate gradient algorithm, and scaled gradient projection algorithm. It was observed that when obtaining the coherent forecasts using the OLS approach, the scaled gradient projection algorithm simplifies to the standard gradient projection algorithm. Hence, use of this algorithm is problematic for very large collections of time series. A series of Monte Carlo simulations were performed to evaluate the computational performances of these algorithms. The results demonstrated that the block principal pivoting algorithm clearly outperforms the rest, and projected conjugate gradient is the second best. The superior performance of the block principal pivoting algorithm can be partially attributed to the alternative representation of the weight matrix in the MinT approach. An empirical investigation was carried out to assess the impact of imposing the non-negativity constraints on forecast reconciliation. It was observed that slight gains have occurred at the most disaggregated level. At the aggregated level slight losses were also observed. Although the gains or losses are negligible, the procedure plays an important role in decision and policy implementation processes. References Wickramasuriya, SL, Athanasopoulos, G, and Hyndman, RJ (2018). Optimal forecast reconciliation of hierarchical and grouped time series through trace minimization. Journal of the American Statistical Association. To appear.

Hierarchical Forecast Reconciliation at Scale

Presenter: Robert Davies
This paper discusses an exercise in performing hierarchical forecast reconciliation across tens of thousands of time-series at a large e-commerce company. In order to accomplish this reconciliation, we adapted the methods originally developed in Hyndman et al (2011) and follow up papers. In particular, we needed to address issues when the forecasts at different nodes varied greatly in scale, ensure that the reconciled forecasts did not violate non-negativity constraints, and address circumstances when the hierarchy was unbalanced. The first and second issues were addressed by introducing a new loss function based on percentage errors when performing the reconciliation. We found that using an L2 loss function as in the original paper would sometimes result in reconciled forecasts that differed quite implausibly from the original data. This would occur when reconciling forecasts that differed greatly in terms of scale. This new loss function also resulted in decent accuracy gains in our backests over using an L2 loss function. The third issue was addressed by careful attention to how the summation matrix was defined. Finally, we also showed that hierarchical reconciliation provided substantial accuracy improvements in our setting over a bottoms-up method across several backtests.

What do we know when? A real time investigation of short-term forecasting

Presenter: Christian Hepenstrick
Nowcasting methods to assess the current state of the economy are often evaluated in terms of their out-of-sample forecast performance for quarterly GDP. However, such an assessment heavily depends on the point in time, i.e. the information set, for which such an evaluation is conducted. Furthermore, it ignores that the target variable – the first release of quarterly GDP – is highly prone to revisions. Therefore, in this paper, we assess such methods in an alternative way and focus on the convergence of each models forecast weekly by week to its current quarter nowcast at the time when the first estimate of GDP is published. In other words, we answer the question of what can be known when. We use DFM based forecasts for four countries: US, Japan, Switzerland and Germany. For Switzerland, we find that within a period of two years after the respective quarter has ended, the real-time assessment of economic developments based on nowcasting approaches outperform first and subsequent GDP releases in terms of prediction ability for final GDP. For Japan and the US, we come quite close to the precision of the first GDP estimate. Additionally, we compare the relative convergence of DFM based forecasts for the four countries and find that convergence is faster for Switzerland and the US compared to JP and Germany.

Can inflation expectations in business or consumer surveys improve inflation forecasts?

Presenter: Raisa Basselier
** Please note that this is an updated version of what was presented last year at the ISF in Boulder **In this paper we develop a new model that incorporates inflation expectations and can be used for the structural analysis of inflation, as well as for forecasting. In this latter connection, we specifically look into the usefulness of real-time survey data for inflation projections. We contribute to the literature in two ways. First, our model extracts the inflation trend and its cycle, which is linked to real economic activity, by exploiting a much larger information set than typically seen in this class of models and without the need to resort to Bayesian techniques. The reason is that we use variables reflecting inflation expectations from consumers and firms under the assumption that they are consistent with the expectations derived from the model. Thus, our approach represents an alternative way to shrink the model parameters and to restrict the future evolution of the factors. Second, the inflation expectations that we use are derived from the qualitative questions on expected price developments in both the consumer and the business surveys. This latter source, in particular, is mostly neglected in the empirical literature. Our empirical results suggest that overall, inflation expectations in surveys provide useful information for inflation forecasts. In particular for the most recent period, models that include survey expectations on prices tend to outperform similar models that do not, both for Belgium and the euro area. Furthermore, we find that the business survey, i.e. the survey replies by the price-setters themselves, contributes most to these forecast improvements.

Perceptions and Expectations of Inflation by Indian Households

Presenter: Kajal Lahiri
Using a large household survey conducted by the Reserve Bank of India, we estimate the dynamics of aggregate inflation perceptions and expectations over a volatile inflation regime over last ten years. We find persistent upward bias and considerable heterogeneity across socio-economic groups and regions. The prices of necessities and certain frequently purchased commodities affect perceptions and expectations disproportionally. Third, the degree of inertia in the expectation formation process varies over inflation regimes. Our findings suggest that, rather than discounting household expectations due to their apparent biases, policy makers should take advantage of the survey results to monitor how inflations is affecting households differentially.

On jumps models and tampered innovations in volatility models and applications

Presenter: Oyebimpe Adeniji
Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models have been used to model non-constant variances in financial time series models. Previous works have assumed error innovations of GARCH models of order (p,q) as: Normal, Student-t and Generalised Error Distribution (GED), but these distributions

failed to capture conditional volatility series adequately, leading to low forecast performance. This study is therefore aimed at developing variants of GARCH(p,q), Asymmetric Power ARCH (APARCH(p,q)) models, Exponential GARCH EGARCH(p,q) model and comparison with Jumps GARCH models such as Generalized Autoregressive Score (GAS) , the Exponential GAS (EGAS) and the Asymmetric Exponential GAS (AEGAS)) with asymmetric error innovations for improved forecast estimates.The two error innovations considered were the Generalised Length Biased Scaled-t (GLBST) and Generalised Beta Skewed-t (GBST) distributions, obtained by remodifying Fisher Concept of Weighted Distribution and McDonald Generalised Beta Function, respectively, in the Student-t distribution. The properties of the proposed distributions were investigated. The proposed innovations were imposed on GARCH(1,1), EGARCH(1,1) APARCH(1,1) models to obtain GARCH-GLBST(1,1) and APARCH-GLBST(1,1), EGARCH-GLBST(1,1) models, respectively. Similarly, GARCH-GBST(1,1), EGARCH -GBST(1,1), APARCH-GBST(1,1) models were also obtained by incorporating proposed innovations into GARCH(1,1), EGARCH(1,1) APARCH(1,1) models. Data from the Central Bank of Nigeria All Share Index (ASL) were used to illustrate the models. The proposed models were compared with jumps and classical models. The performance of the proposed models over the existing ones were investigated using the Log-likelihood function, Root Mean Square Error (RMSE), Adjusted Mean Absolute Percentage Error (AMAPE) and Akaike Information Criterion (AIC).Out of the 18 models in consideration, EGARCH-GLBST(1,1) was the best, followed by APARCH-GLBST(1,1) and EGAS models, in terms of the AIC values (7.856,7.988 and 9.984).The forecast evaluation criteria (RMSE, AMAPE), EGARCH-GLBST(1,1) model also ranked best (RMSE =0.281, AMAPE = 0.280), followed by APARCH-GLBST(1,1) model (RMSE = 0.291, AMAPE = 0.290) and EGAS model (RMSE = 0.309, AMAPE = 0.301). The least performing in terms of forecasts was the GARCH(1,1)-Normal model. The proposed volatility models with error innovations outperformed existing models in terms of fitness of conditional volatility and forecasts. The proposed models will be good alternatives for volatility modelling of symmetric and asymmetric stock returns.

Can Solar Data Help to Forecast NASDAQ's Finance Sector Price Index? An ARIMA - GARCH Approach

Presenter: Theodoros Daglis

Although the impact of solar events on financial activity appears to be of great interest, very limited work has been done, so far, on the subject. In this work, we examine the impact of solar weather on Earth's financial activity and more precisely, on NASDAQ's finance sector price index, spanning the period 1998-2008. We find that the comparison of the MAPE and RMSFE for the ARIMA-GARCH model augmented with exogenous variables, capturing the solar activity, over the standard model without exogenous variables, shows that the augmented model is statistically significant in the solar variables and better in terms of its forecasting ability. This means that the solar variables provide information for the forecasting of NASDAQ's finance price sector index, in the 1998 –2018 time period. These results give credit to the impact of solar events on financial activity.

Finding a well performing Box-Jenkins forecasting model for annualised patent filings counts.

Presenter: Peter Hingley

At the European Patent Office (EPO), a forecasting exercise takes place at the beginning of each year. This predicts how many incoming patent filings there will be in the current year and also for each of the following five years. A scenario is proposed that is then converted into a business plan for the volumes of EPO's major products and activities. The plan is used as a basis for the EPO's budget and investment decisions. Here we apply an automatic data based model selection procedure for Box-Jenkins time series models for the annualised patent filings series. The method is applied separately to autoregressive integrated moving average (ARIMA) and auto-distributed lag (ADL) models (Dikta, 2006). For ADL, independent variables include terms that are based on GDP and R&D expenditures series in the source countries for the patent filings. A universe of such models is considered that specifies no more than four lags in any of the dependent or independent variables. Firstly, the optimal degrees of differencing for the series are established using the KPSS test. Then a subset of differenced models is considered that passes normality and autocorrelation

tests. From these, the model with the lowest Akaike's Information Criterion (AIC) is selected as the winner. Forecasts for the next five years' patent filings are generated from the winning model. But in some years the procedure finds no appropriate model. Or it can be that the model that is selected cannot be fitted or generates unrealistic forecasts. We take a multi-year approach to recreate the historical sequence of forecasting exercises that took place as the time series grew in length by one observation each year. We analyse the collection of allowed models over these years in order to identify a group of models that performs relatively well in terms of AIC from year to year, while not necessarily having the lowest AIC in any particular year. We argue that this is an appropriate model selection approach for the time series of patent filings (and similar business time series), where the data generating process either does not follow exactly the assumptions of a model or contains unspecified heterogeneities. The practical usefulness of the selected models is assessed retrospectively by using MAPE like statistics. Implications for the accuracy of the forecasts to be made in the future are discussed. Reference: Dikta, G., 2006. "Time series methods to forecast patent filings". In "Forecasting Innovations", P. Hingley and M. Nicolas (editors), Springer, Heidelberg, pages 95 – 124.

Equal Predictive Ability Tests for Panel Data with an Application to OECD and IMF Forecasts

Presenter: Oguzhan Akgun

This paper focuses on testing for equal predictive ability using panel data. Weextend the tests developed by Diebold and Mariano (1995) to a general panel data context.The first aim of the paper is to adapt such testing procedures to panel data takinginto account the complications arising while using data from different units, in particular cross-sectional dependence across units. Namely,the objective is to build tests to compare the predictive ability of two forecasters, based on n units, hence n pairs of time series of observed forecast errors, fromtheir forecasts on an economic variable while allowing for different strength andtypes of cross-sectional dependence. Various panel data tests of equal predictiveability are proposed, extending that of Diebold and Mariano (1995) which concerns a single time series.The second aim of the paper is to investigate the small sample properties of the testsproposed. For the treatment of spatial dependence in the errors, we use spatial heteroskedasticity and autocorrelation consistent(SHAC) estimators of the covariance matrix. In a time series framework the smallsample properties of HAC estimators are well known and comparison of the role of different kernel functions in the estimation performance is readily available. Here, the analysis is extended to allow forseveral additional kernel functions and their small sample performance is evaluated through Monte Carlo experiments.Finally, the paper contributes also to the empirical literature. These tests areapplied to compare the economic growth forecasts errors of the OECD and the IMF.We investigate the equality of accuracy for different time periods and country samples.

Effects of External Assumptions on Forecast Errors

Presenter: Katja Heinisch

This paper examines the extent to which forecast errors on the economic development are driven by external or technical assumptions that prove to be incorrect ex post. Therefore, we use a new data set comprising an unbalanced panel of annual forecasts from different institutions forecasting German GDP, exports and imports and the underlying assumptions. We test whether forecast errors decrease significantly with smaller horizons. A high fraction of the variation in forecast errors for GDP, exports and imports growth can be accounted for by the variation in the underlying assumption errors. We find that forecast errors in world trade drive most of the variation in GDP forecast errors.

Effects of External Assumptions on Forecast Errors - A Global Perspective

Presenter: Christoph Schult

This paper analyses the impact of ex-ante technical assumptions for forecast errors of GDP growth and inflation of international forecasters such as the IMF, OECD and European Commission. First, we construct a new data set covering forecasts and underlying assumptions on commodity prices, financial markets indicators and world economic growth from various international institutions and for different countries. Second, we calculate the forecast and assumption errors with respect to first releases and estimate to which extent

forecast errors could be explained by wrong assumptions and whether the results differ significantly with various horizons and for different countries (country groups). Preliminary findings indicate that the oil price assumption matter in particular for inflation forecast errors and world trade assumptions for forecast errors of GDP growth. However, the results vary for different country groups and horizons.

Practical forecaster construction through cloud-based model search

Presenter: Tomas Singliar

Modern forecasting makes extensive use of machine learning methods. The typical ML strategy is to convert the forecasting problem into a regression problem. The formulation creates both an opportunity to improve accuracy with "wigglier" models (with more degrees of freedom) and a challenge of how to search the larger space of models. Automated Machine Learning (AutoML) is a machine learning SDK distributed via PyPi, the open python package repository. AutoML makes statisticians and data scientists more productive by automating the sometimes-tedious process of selecting the best model – data transforms such as lagging, estimator type such as linear regression, and their hyper-parameters such as rolling window length – that the data will support. The package can run locally or leverage cloud computing to scale. Now, Microsoft is making forecasting a first-order task in AutoML. Proper rolling-origin cross-validation, horizon-aware confidence intervals from back-testing and missing value imputation that respects time ordering are some of the features the AutoML is adding to support forecasting model search. AutoML aims to be the easiest path from a time series dataset to an operationalized forecaster. We will discuss several practical considerations not often appreciated in the retrospective train/test framework and how AutoML gives the statistician/forecaster the tools to addresses them. For instance, the concept of "information set" is underemployed in modern data science. However, especially with time series forecasting, it becomes central to developing and deploying a correct model. In analyzing a time series which describes a sequence of historical intervals, we typically assume that "data describing a period p becomes fully known at end of p". This default assumption is a pitfall that may result in a developing a model which is not operable because it lacks input data at time of prediction, or inaccurate because it fails to consider available data that describes future periods. For example, GDP estimates are published several months after the end of the quarter, while the price of potatoes at the grocery store in two months is known to the management today, and becomes known to the customers one month in advance. Another aspect that the typical forecast in a business context must be done "in bulk", with many series in a single input dataset, with groupings of series for modeling to improve accuracy and enable "cold-start" forecasting. A short demonstration/tutorial of forecasting within AutoML will be included in the talk, with the requisite Azure resources provided for the time of the conference free of charge to participants. We will make permanently available all instructional material including working examples. We will share competitive performance results for AutoML on M3 and other datasets. Our M4 evaluation is also being planned and likely to complete in Q2 2019.

Demand forecasting beyond the time series vector: meeting the business requirements

Presenter: Akshey Gupta

So you've built the new best forecasting model in the world that solves "given (Xt, yt) for t=1,. . . ,T, what is yt for t=T+1,. . . ,T+h?". It beats everybody on M3 and M4. It should be easy to sell it to some large businesses where a small accuracy improvement creates a lot of value – or maybe not. What practical considerations does a business need? What lies between the abstraction of (Xt, yt) and software running the business every day? Turns out there is quite a bit. Following an actual implementation of ML-based forecasting, we will talk about how these business requirements are addressed. The actual implementation is the product demand forecasting solution in Dynamics AX, Microsoft's ERP ("an operating system for a company"). Most of these decisions come down to appropriately grouping the time series. This grouping makes the sets of time series more homogeneous from the business perspective and is separate from any "ML" grouping that is done for the benefit of model accuracy. There are several reasons for grouping. First, forecasts are used to drive decisions about how things should be, for instance, "how many red t-shirts should we stock next week"? However, historical data normally describes how things were, even if the overall system performed sub-optimally. For example, we fulfilled the order from a distant warehouse B, when nearby

warehouse A was out of stock. Using uncorrected data will perpetuate the problem of under-forecasting at A and over-forecasting at B. We will cover the process of constructing a "usage file", a picture of what the true demand should have been. Second, products time series will require a wide mix of forecast horizons, due to varying production and procurement lead times. Frequently, the system-configured lead times do not match the actual lead times, and thus the appropriate horizon is itself in need of being forecast. Third, while the model you just built is universally excellent, and does not require coarser aggregation for slow moving items to attain the necessary accuracy, the business still might need to manage these items differently due to considerations such as economic order quantity. Finally, we will comment on approaches that have worked well in Microsoft for forecasting demand for new products. Forecasting demand for completely new products is very challenging and nobody has a very good forecaster. However, we have had good results with modeling demand for new generations of existing products. The successful solution uses hedonic regression combined with the time-variable notions of "current" and "previous" generation. Other successful approaches involved demand sensing on social networks for computer game "buzz" before release.

Moving your custom forecaster to the Azure cloud

Presenter: Austin Gross

As the number of available potentially useful predictor data feeds multiplies, business decision makers expect from their forecasters more agility in developing fully operational custom forecasters that take advantage of the predictors. While forecasters are traditionally excellent statistical modelers, "operationalizing" a custom forecasting model is also a prerequisite for the model to become practically useful. However, not all forecasting personnel have the expertise to quickly turn forecasting models into operational models at business scale and IT departments can impede modeler's impact. Operationalizing means moving the forecaster's code, often written in R or python, from a statistician's environment where it was developed and evaluated, into an environment where it can connect to the business data sources and run regularly and independently of the forecasting expert's intervention. For a forecasting job to scale to millions of time series (think daily sales forecasts for 10,000s items across 100s of stores), the work has to be suitably divided into smaller jobs and executed on a scalable platform. Azure Machine Learning Services are a cloud platform to train machine learning models, including forecasters, either on your own computer at no cost or on cloud compute resources at a large scale. We will give an overview of the end-to-end process of how to develop your own machine learning model, starting with a historical dataset and ending with a fully deployed, operationalized forecaster in Azure cloud. But that is only the "inner loop". We will also demonstrate how to schedule regular (e.g. weekly) executions of the forecaster on data in regularly updated data warehouses, and present the forecasts in standard productivity software like Excel and PowerBI reports that are automatically updated. Python Jupyter notebooks with hands-on exercises will be made available for the session. The short duration of the session will not allow for interactive participation, but participants will be able to use the notebooks as a template to their own machine learning model with their own data later. The intent is to mimic an actual business forecasting scenario. Then we will walk through an example of building a long short-term memory (LSTM) networks to enerating stock market predictions. (The choice of an arguably unpredictable series underlines that this is an example intended for the users to adapt to their own more predictable data.) We will explain the mapping of the time series forecasting problem onto the neural architecture and show how it is reflected in code. The resulting model will be deployed as a web service in Azure Container Instance (ACI) or Azure Kubernetes Service (AKS). The web service runs a python script that encapsulates the scoring logic and the model itself. Then we will explain how to use AzureML Pipelines to create a scheduled run moving data from a data warehouse to the forecasting web service and back into the data warehouse. We will discuss how the data can be partitioned for consumption in AKS so that the automatic scaling properties of the platform enable you to generate many forecasts in a short time.

Compound Poisson demand parameter estimation for inventory control

Presenter: Dennis Prak

Most companies store demand data periodically (e.g. weekly) and make periodic demand forecasts, although inventories are controlled continuously. The same 'mismatch' can be observed in the literature; most forecasting methods are periodic, but inventory control theory often assumes the demand process to be continuous and

given. Guidance on estimating the parameters of a continuous demand process from period demand data is lacking, in particular for the popular and well-studied compound Poisson class of demand. Commercial software packages wrongly fit compound Poisson processes based on period Size-Interval forecasting methods, leading to dramatic overshoots of the target service level and therefore too high inventory costs. Also the standard method-of-moments (MM) parameter estimators suggested in textbooks and inventory control research papers have substantial biases in finite samples. We propose an intuitive, consistent, closed-form MM alternative that dominates in terms of estimation accuracy and on-target inventory performance.

Demand Forecasting Under the Fill Rate Constraint – The Case of Re-Order Points

Presenter: Joanna Bruzda

The aim of this paper is to discuss and compare theoretically and through simulations the performance of different methods of forecasting under the conditions concerning the type II service level (i.e., the fill rate). The exact formulation of the fill rate constraint is case sensitive, with the two most popular specifications being the one applying to the one-period model: FR(S) = 1- E(D-S)+/ED = t (1) and the one in the multiperiod re-order point based inventory model: FR(ROP) = 1-E(DL+1-ROP)+/Q = t (2) where D is the random demand, DL+1 is the demand during the lead time increased by 1, S is the production level, ROP denotes the re-order point and Q is the order quantity. The elicitability of the functional S defined in (1) and its certain variation have been noted in Bruzda (2018). Here we concentrate on the re-order point formula (2), for which we specify a family of loss functions uniquely minimized at the ROP. Focusing on the special case of the linear-quadratic loss of the prediction-error type of the form: L(e) = (e-C)21{e=0} + [-2Ce+C2]1{e<0} (3) where e is the forecast error and C = (1-t)Q, and assuming some regularity conditions, we prove consistency and asymptotic normality of our estimator of the ROP defined as: ROPn = argmin <U+03A8>n(ROP), (4) where <U+03A8>n(ROP) = (1/n)<U+2211>L(yt-ROP), for the case of iid data. The loss function (3) is treated here as an implied inventory cost resulting from the service level assumption of the decision maker (compare Brandimarte and Zotteri, 2007). Properties of the estimator (4) in comparison with those of two-step estimators obtained under different distributional assumptions and the maximum likelihood estimation of parameters are studied theoretically and through extensive simulations studies. The normal, Laplace and gamma distributions are covered. For both types of estimators, we also examine the small sample coverage of confidence intervals for the ROP constructed according to the appropriate asymptotic formulae. In the outlook part of the paper, we formulate a general consistency theorem for the estimator (4) in the case of dependent stationary data. Some classes of dynamic models for forecasting under the type II service level constraint are also discussed, such as generalized autoregressive score (GAS) models based on the loss function introduced here or its multivariate extension and different uni- and multivariate ROP smoothing equations corresponding to the quantile smoothing models discussed in Bruzda (2019). In particular, the simplest univariate GAS model is as follows: ROPt+1= <U+03C9> + ßROPt + a(yt-ROPt)1{yt=ROPt}. (5) The model (5) assumes that the re-order point is substantially increased each time the demand during the lead time is above the available stock, while otherwise it converges towards a steady state. References Brandimarte P., Zotteri G. (2007), Introduction to Distribution Logistics, Wiley, Hoboken. Bruzda J. (2018), Multistep Quantile Forecasts for Supply Chain and Logistics Operations: Bootstrapping, the GARCH Model and Quantile Regression Based Approaches, Central European Journal of Operations Research, to appear. Bruzda J. (2019), Quantile Smoothing in Supply Chain and Logistics Forecasting, International Journal of Production Economics, 208.

Quantifying Economic Impact of Classical Forecast Errors

Presenter: Safi Elegbede

In recent years, among forecasting practitioners, managers and academics, there has been increasing awareness and growing interest in the economic impact of classical forecast errors on inventory control parameters and business bottom-line. In this paper, we consider the accuracy of common and most often used forecasting methods and their economic impact in terms of relevant stockout cost, stockover cost and inventory total cost for replenishment decisions characterised by the order up to (OUT) level inventory policy. The demand

for this replenishment system follows either a first order autoregressive, AR (1) process or a mixed first order autoregressive and moving average, ARMA (1,1) process. The study assesses the contemporary methods of forecast evaluation and the impact on decisions at the levels of operations and financial functions. In particular, we evaluate and quantify the traditional error metrics such as the mean squared error (MSE),the mean absolute error (MAE) and the mean absolute percentage error (MAPE) interms of inventory variables such as relevant inventory costs and revenues with reference to predefined service level. Using less restrictive assumptions, we show that the typical traditional safety stock model is neither adequate nor robust and can thus be enhanced. We also fill some gaps in the literature by extending results on utilisation of forecast methods to Naïve, ExponenTial Smoothing(ETS) and Auto ARIMA in the economic quantification of statistical forecast errors such as MSE, MAE and MAPE. The results of this study should enhance our understanding of how to better implement inventory control and to better mitigate inventory costs and revenue risks in practice.Keywords:demand forecasting, forecasting errors, inventory cost risks, inventorycontrol, inventory decisions

Impact of demand volatility forecasting on inventory systems

Presenter: Juan R. Trapero

Most of the research carried out in supply chain forecasting has focused mainly on point forecasts, where point forecasts are estimations of the central tendency of the density function of the demand. Point forecasts are crucial, but they are just half of the problem, where forecasting the variability is equally important. Such variability is of paramount importance since it is a measure of risk and then, it is directly related to the determination of the safety stock size. Those safety stocks are also employed to set the reorder points in many replenishment policies. In this presentation, we will talk about different methodologies to forecast volatility, which are present in other disciplines as finance and which are not entirely utilized in a supply chain context. Forecast errors will serve as a measure of future demand variability and inputs to the volatility forecasting models. Note that, if the underlying demand generating process is correctly identified by the forecasting model (ARIMA, State Space, etc...), such errors are typically assumed to be independent and identically distributed (iid), following a Gaussian distribution with zero mean and constant variance. Nevertheless, given the market complexities, it is unrealistic to assume that your forecasting models have captured the wide variety of underlying demand patterns, and thus, the fulfillment of iid assumptions should be, at least, checked. For instance, deviations from iid can be found in products subject to promotional campaigns, where demand density functions tend to be asymmetric. Moreover, homoscedastic assumption has also been questioned in previous works, where the presence of heteroskedasticity in in industrial time series has been shown. Essentially, this work explores empirical parametric (GARCH and single exponential smoothing) and non-parametric (Kernel and percentiles) methods to forecast the volatility of several SKUs real demand, among others. The performance of the forecasting methods, considering lead times of different length, will be validated in terms of cycle service level, inventory investment and number of backorders, i.e., trade-off curves. The results show that for shorter lead times, non-normality is the main issue and the non-parametric kernel method produced better results. In turn, for bigger lead times, conditional heteroscedasticity became dominant and parametric GARCH models outperformed the rest of benchmarks.

Forecasting power generation for Small Hydropower Plants using inflow data from neighboring basins

Presenter: Margarete Afonso de Sousa

Electric power generation forecasting has an important role for Utilities when planning their operations. The difficulty increases ha target are Small Hydropower Plants, in general, run-of-river plants with a strong dependence of hydrological regimes. Not only that, inflow or precipitation data are, not only unavailable but also not long enough to be used in the modelling process. Obtaining good prediction closer to reality depends on the choice of models and methods. Considering the possibility of not having enough inflow observation data to more efficiently forecast the generation of Small Hydropower Plants, this study aims to present two approaches for the use of neighboring inflow series via causal models. The first one considers the use of neighboring inflow that is highest correlated with the generation directly in the model and the second one by backforecasting inflow series from small hydropower plant using neighboring inflow using a straight forward

linear regression. In this study, SARIMA, Transfer Function and Linear Regression models were used.

Probabilistic short-term Water Demand Forecasting

Presenter: Jens Kley-Holsteg
The water demand is a highly important variable for operational control and decision making. Hence, the development of accurate forecasts is a valuable field of research to further improve the efficiency of water suppliers. In this paper, a time series model is introduced to capture the typical periodicities, to account for a time-varying variance and to quantify uncertainty and time dependences of the water demand process. As estimation procedure an automatically shrinkage and selection operator (lasso) is applied. It allows to obtain a parsimonious, simple interpretable and fast computable forecasting model, which is well suited for the application in real time operational control systems. The methodology is applied to hourly water demand data of a German water supplier. Moreover, the importance for more advanced evaluation criteria is outlined and the energy score (ES) is introduced. The ES is an appropriate score to account apart from the accurateness of the mean also for the marginal properties as well as time dependences in an issued forecast.

When is water consumption extreme?

Presenter: Clara Cordeiro
We propose new models for analysing and forecasting hourly water consumption data. Such data can include extreme values due to meter malfunction, burst water pipes, etc. Therefore, special care must be given to these types of events in the series, and specific statistical procedures based on the behaviour of extremes are required to handle them. Unlike most traditional statistical theory, which typically examines the usual (or average) behaviour of a process, extreme value theory deals with models for describing unusual behaviour or rare events. Our aim is to model the statistical characteristics of such time series in order to understand the probabilities of extreme events and allow better management of water utilities. We illustrate the key ideas using data from a water company in Portugal, Infraquinta.Acknowledgements: Clara Cordeiro was funded by Fundação para a Ciência e a Tecnologia, IP (FCT) through the Grant SFRH/BSAB/142999/2018, and through the project FCT Portugal UID/MAT/00006/2019.

Forecasting Water Usage Demand in Sydney

Presenter: Vasilis Sarafidis
There is currently a high level of interest among policy agencies and international organisations in using water usage prices to balance the supply and demand for water (scarcity pricing), as well as in forecasting future water usage demand. This study applies dynamic panel data analysis to household level data in order to forecast water usage demand in Sydney in the short- and long-term. Estimation is implemented using the GMM approach, due to its appealing properties and generality in regression models with endogenous regressors. The approach to forecasting water use involved econometric analysis in 31 distinct segments, formulated depending on dwelling type (single or multi-residential properties), whether the property was built before or after the introduction of the BASIX regulation, tenancy status (owner occupied or tenanted), and whether or not households have access to reticulated recycled water. The models quantified the impact of identified variables (most notably, weather conditions and seasonal variables and price) on water use patterns. Forecasts of the explanatory variables were then used to develop forecasts of water use.

Optimal Density Forecast Combinations

Presenter: Gergely Ganics
How should researchers combine predictive densities to improve their forecasts? I propose consistent weight estimators building on the uniformity property of the Probability Integral Transform. The estimators provide density forecast combinations approximating the true predictive density, conditional on the researcher's information set. Monte Carlo simulations confirm that the proposed methods work well for sample sizes of practical interest. In an empirical application of forecasting monthly US industrial production, the estimator outperforms well-known benchmarks, such as the equal weights scheme. I show that housing permits had valuable predictive power before and after the Great Recession. Furthermore, stock returns and corporate

bond spreads proved to be useful predictors during the recent crisis, suggesting that financial variables help with density forecasting in a highly leveraged economy.

Forecaster Efficiency, Accuracy and Disagreement: Evidence using Individual-Level Survey Data

Presenter: Michael Clements
Recent theories of expectations which stress the role of informationrigidities suppose that agents make efficient forecasts given theirinformation sets. These theories have generally been tested on aggregatequantities, such as (cross-sectional) mean forecasts and mean forecasterrors. We use individual-level data to consider whether there aresystematic differences between forecasters in terms of their degrees ofcontrarianism, and the accuracy of their forecasts, and whether these areexplicable by inefficiencies in the use of information. We find thatforecaster inefficiency cannot the explain persistence in levels of disagreementacross forecasters, but there is evidence that the inefficient use of informationis responsible for persistent differences in accuracy across forecasters.

Robust Optimization of Forecast Combinations

Presenter: Stelios Arvanitis
A methodology is developed for constructing robust forecast combinations which improve upon a given benchmark specification for all symmetric and convex loss functions. The optimal forecast combination asymptotically almost surely dominates the benchmark and, in addition, minimizes the expected loss function, under standard regularity conditions. The optimum in a given sample can be found by solving a large Convex Optimization problem. An application to forecasting of changes of the S&P 500 Volatility Index shows that robust optimized combinations improve significantly upon the out-of-sample forecasting accuracy of simple averaging and unrestricted optimization.

Nonparametric Tests for Superior Predictive Ability

Presenter: Valerio Poti
A nonparametric method for comparing multiple forecast models is developedand implemented. The hypothesis of Nonparametric Forecast Optimality generalizes the Superior Predictive Ability hypothesis from a single given loss function to an entire class of loss functions. Distinction is drawn between General Loss functions, Convex Loss functions and Symmetric Convex Loss functions. The hypothesis is formulated in terms of moment inequality conditions. The empiricalmoment conditions are reduced to an exact and finite system of linear inequalitiesbased on piecewise-linear loss functions. The hypothesis can be tested using ablockwise Empirical Likelihood Ratio test. An empirical application to inflationforecasting reveals that a very large majority of thousands of forecast models areredundant, leaving predominantly Phillips Curve type models, when convexityand symmetry are assumed.

Forecasting excess return correlation with short selling information

Presenter: Marco Valerio Geraci
We use public short selling disclosures made to the Financial Conduct Authority (FCA) of the United Kingdom to connect stocks based on their common short sellers. We find that the number of common short sellers can predict abnormal stock return correlation one month ahead, controlling for similarities in size, book-to-market, momentum, and several other common characteristics. In our most flexible specification, a standard deviation increase in the number of common short sellers is associated with a future rise of 2.13% of the average four-factor excess return correlation for a given stock pair. We show that when predicting correlation our variable capturing the number of common short sellers can lead to significant out-of-sample forecast gains. Moreover, we show that the forecasted correlation can be used to establish a trading strategy that can yield positive cumulative abnormal returns of over 9% after 12 months, gross of transaction costs.Several possible mechanisms can explain the positive association between short selling and future correlation because, fundamentally, several reasons may lead a short seller to initiate a short position. We investigate three

mechanisms which we believe are most important and determine whether they have support in our data.Short selling could induce higher correlation by applying negative price pressure on several stocks. This price impact effect should be stronger for illiquid stocks (Brunnermeier and Pedersen, 2005, Brunnermeier and Oehmke, 2014). We use this prediction to verify whether there is evidence of the price impact effect in our data. Specifically, we test whether the uncovered association between the number of common short sellers and future correlation is stronger for less liquid stocks. Our empirical results do not find support for this mechanism. At least at the frequency and periodicity of our study, the positive association between short selling and future excess correlation does not appear to corroborate the price impact of short selling.As an alternative, our results could be a by-product of informative trading strategies of short sellers. Ultimately, by initiating a short position, a short seller predicts that stock prices will decline in the future to gain a positive return from the trade. Previous studies have shown that short sellers are sophisticated market agents, who trade on the basis of superior information. If common short sellers trade according to superior information that stock prices should decline in the future, this will coincide with higher future correlation across the shorted stocks. Thus, if this mechanism is occurring, the positive relationship between common short selling and future excess correlation can be explained by informative trading. To verify this mechanism, we isolate the number of common short sellers that can be associated with informative trading using several measures of value obtained from financial statement analyses. We then show that the informed number of common short sellers is a strong predictor of future returns, hence corroborating the informative trading mechanism.Overall, our results indicate that the informative trading mechanism is a likely explanation for the uncovered relationship between short selling and future correlation.

Spurious Dividend Smoothing

Presenter: Dae Keun Park

It is necessary to explain whether and why firms make dividend smoothing in connection with persistent earnings in order to understand their dividend policies. In many studies carried out so far, the focus was only on identifying the persistent earnings and modeling the dividend and persistent earnings as cointegration relation. However, there is no analysis of spurious dividend smoothing in which the cointegration relation is not established or temporary earnings affect dividends even though the cointegration relation is established. Our study aims to decompose the earnings into persistent and temporary earnings through trend-cycle decomposition method and to analyze the impact of these earnings on dividends. With the effects of the persistent and temporary earnings, it is examined whether and why firms make dividend smoothing by reflecting the capital procurement characteristics of firms in the error correction model. As a result of analyzing Korean firms, it is found that firms that finance from direct financial markets or have growth opportunity exhibit dividend smoothing whereas other firms show spurious dividend smoothing. This study can suggest a solution to understand the dividend policy of a firm by providing a systematic method to judge whether the firm makes spurious dividend smoothing or not.

Credit Risk Calculation: an application in the Brazilian market using the CreditRisk+ model with uncertainties

Presenter: Beatriz Jardim Pina Rodrigues

Due to an increasing economic instability worldwide, financial institutions are demanding more robust and powerful methodologies of credit risk modeling in order to ensure their financial health. The statistical model CreditRisk+, developed by Credit Suisse Financial Products (CSFP), is widely spread in the insurance market since it is not necessary to make assumptions. This is because the model is based on the default risk, that is, non-payment risk. The main goal of the above-mentioned model is to measure expected and non-expected losses in a credit portfolio. In order to measure default events, the model suggests grouping the debtors in exposure ranges so that the loss distribution can be approached to a Poisson. In the basic model, the default rates are fixed. To portray reality, we propose a new modeling in which the uncertainties and volatilities of default rates are incorporated. In this case, a new model which assumes a Gama distribution in association with these uncertainties. From the obtained distribution, not only is it possible to calculate the credit VaR (Value-at-Risk) but also the loss distribution and some point estimates, such as the expected loss in a certain period of time and the allocated economic capital. The main goal of this article is the CreditRisk+ model

application with uncertainties in a segment of Brazilian industry. The allocated economic capital, that is, the difference between VaR and the expected deprival value is always higher, depending on the proposed modeling (with the incorporation of uncertainties, volatilities and the default rates). Our result is important, since financial institutions can be underestimating their losses in stressful moments.

Stalking the Lizard: Validation of a Technical Analyst's Track record

Presenter: Roy Batchelor
Commentators on stock markets are more than willing to share their trading philosophies. Very few share their track records. This paper takes a careful look at the public record of trades by one analyst, an avowed follower of the cultish Elliott Wave (EW) Theory. We find that (a) there are many disjoints between prices published by the trader, and prices available on the trade dates, (b) claims of apparent outperformance are not supported by simple benchmarking, and by more formal bootstrap tests, and (c) it is unclear how, or how consistently, the recommended trades align with EW principles. The study raises issues over the way forecast performance is communicated to clients, and the use of sophisticated-looking theories as window dressing for more eclectic approaches to forecasting in the stock market.

A feast of time series tools

Presenter: Rob Hyndman
Modern time series are often high-dimensional and observed at high frequency, but most existing R packages for time series are designed to handle low-dimensional and low frequency data such as annual, monthly and quarterly data. The feasts package is part of new collection of tidyverts packages designed for modern time series analysis using the tidyverse framework and structures. It uses the tsibble package to provide the basic data class and data manipulation tools.feasts provides Feature Extraction And Statistics for Time Series, and includes tools for exploratory data analysis, data visualization, and data summary. For example, it includes autocorrelation plots, seasonality plots, time series decomposition, tests for units roots and autocorrelations, etc.I will demonstrate the design and use of the feasts package using a variety of real data, highlighting its power for handling large collections of related time series in an efficient and user-friendly manner.

Flexible futures for fable functionality

Presenter: Mitchell O'Hara-Wild
The fable ecosystem provides a tidy interface for time series modelling and forecasting, leveraging the data structure from the tsibble package to support a more natural analysis of modern time series. fable is designed to forecast collections of related (possibly multivariate) time series, and to provide tools for working with multiple models. It emphasises density forecasting, whilst continuing to provide a simple user-interface for point forecasting. Existing implementations of time series models work well in isolation, however it has long-been known that ensembles of forecasts improve forecast accuracy. Hybrid forecasting (separately forecasting components of a time series) is another useful forecasting method. Both ensemble and hybrid forecasts can be expressed as forecast combinations. Recent enhancements to the fable framework now provide a flexible approach to easily combine and evaluate the forecasts from multiple models.The fable package is designed for extensibility, allowing for easier creation of new forecasting models and tools. Without any further implementation, extension models can leverage essential functionality including plotting, accuracy evaluation, model combination and diagnostics. This talk will feature recent developments to the fable framework for combining forecasts, and the performance gain will be evaluated using a set of related time series.

Forecasting performance of Bayesian and Mixed Frequency VAR Models in EViews 11

Presenter: Gareth Thomas
Vector autoregression models (VARs) are a staple of modern macroeconomic analysis and prediction. However, the accuracy of their out-of-sample forecasts in macroeconomic studies is hampered by two issues; a large

parameter set and the inherent differing frequencies of the underlying time series data. To alleviate the first of these issues researchers have turned to Bayesian VARs (BVAR) which are able to model relationships with many parameters. BVARs require selecting a choice of prior alongside any hyper-parameters or initial conditions specified with the prior. Traditionally, the second issue has recently been addressed by using aggregation to convert all the underlying time-series to the lowest frequency. This method involves loss of fidelity as multiple high-frequency observations can be reduced to a single low-frequency observation. In recent years the introduction of Mixed Frequency VARs (MFVAR), which allow for the underlying data to be sampled at differing frequencies, has provided an alternative solution. In this presentation we demonstrate the estimation and forecasting of BVARs with Litterman (1979), Normal-Wishart (Koop and Korobilis 2010), Independent Normal-Wishart (Koop and Korobilis 2010), Sims-Zha (1998) or Giannone, Lenza and Primiceri (2012) priors in EViews 11. We also demonstrate the use of the MFVAR of Ghysels (2016) in EViews. We compare the forecasting power of each of the BVAR priors, utilizing different initial covariance assumptions, and the MFVAR using a standard US macro-economic dataset (Stock and Watson 2008).

Automatic Forecasting of Unobserved Components Models with the UComp Toolbox for MATLAB

Presenter: Diego J. Pedregal

The era of big data has triggered a revolution in many research areas. Indeed, it can be said that this effect is particularly dramatic in the area of time series forecasting. In a short period of time, forecasting needs in private companies and public entities have gone from a crafted analysis of each individual problem to a tsunami of information that must be processed efficiently, online and in record time. This change made the use of automatic identification algorithms compulsory to meet the users' needs. Some of the classical methods have adapted for these changes, like ARIMA and exponential smoothing models, but there is a promising class of models that has been systematically out of the main stream, namely the Unobserved Components Models (UC). There are many reasons for this. Firstly, UC have been developed in academic environments, with no strategy for their dissemination among practitioners for their everyday use in business and industry. Secondly, the widely-held feeling that UC models do not really have anything relevant to add to exponential smoothing methods has deterred its use in practice. Third, UC models are usually identified by hand, with automatic identification being very rare. Finally, software is rather scarce compared with other methods.This paper introduces a new toolbox (UComp) developed in MATLAB for the automatic identification and forecasting of UCs with some enhancing features. The toolbox searchers among many combinations of different specifications for each of the components following a step wise algorithm to reduce the universe of possible models and selects the one with the best metrics. The forecasting results suggest that UC models are powerful potential forecasting competitors to other well-known methods. Though there are several pieces of software available for UC modeling, this is the first implementation of an automatic algorithm.

Integrating theta method and multiple temporal aggregation: optimising aggregation levels

Presenter: Bonan Wang

Combining forecasts is beneficial as it leads to improved accuracy and reduced uncertainty. In this paper, we focus on the combined performance of two forecasting methods – theta and multiple temporal aggregation. Theta method produces forecasts through modelling lines with different curvatures, aiming to capture separately short- and long-term components. Multiple temporal aggregation (MTA) handles series of different frequencies so that specific patterns are amplified. The merits of the two methods imply potential gains in terms of forecasting performance from considering a hybrid approach. The aim of this paper is to explore and evaluate the structural integration of the theta method with multiple temporal aggregation through an empirical investigation. We compare the performance of applying the theta method on a single (optimal) aggregation level, as well as multiple temporal aggregation levels with equal or unequal weights derived by validation evaluation and clustering analysis. We also examine the connection between the selection of aggregation levels and time series features including length of time series, trend, and seasonality, etc. We create a feature-based sampling pool for categorising time series in order to identify the best combination of

aggregation levels. Our results suggest that clustering and feature-based weighting of aggregation levels can offer improvements over a simple integration of Theta-MTA.

Simplex combination and selection of forecasters

Presenter: Aranzazu de Juan
This paper considers the Split-Then-Combine (STC) approach (Arroyoand de Juan, 2014) to combine forecasts or models in the simplex: thesample space of positive weights adding up to one. Using statistical mea-sures in the simplex, we construct combinations of forecasts by varyingweights and compare them with the fixed weight approach (the average offorecasts). We also get rid of redundant forecasts, and make out-of-sampleone-step ahead forecasts through our Combine-After-Selection (CAS) ap-proach. We apply STC and CAS to construct combination of forecasts forseveral stock and flow economic variables, even when the sample size ismuch smaller than the number of forecasts

Trimming of forecasts in simple average combinations

Presenter: Mathias Käso
For a given set of competing forecasts, it is well known that forecast combinations are a proven procedure to achieve a higher forecast quality. In this context, it has been found that the simple average (weights equal to 1/N, where N is the number of forecasts) is often hard to beat. Based on this insights, there are some attempts to further improve the forecast performance of the simple average. The most common approaches are based on two different kinds of "trimming" (Timmermann, 2006). Firstly, we have the neglect of one or more complete forecasts in the combination. Secondly, there are approaches in which some of the smallest and largest forecast values are removed in each combination step ("as a form of outlier reduction rule", c.f. Timmermann, 2006). In empirical investigations these strategies has shown the potential for further improvements, although there are many open questions. Granger and Joen formulates this circumstance as follows "However, there is also a belief that trimming the forecasts, by throwing away the k% worst performers is a very good idea, before combining, when k = 5 or 10, say. Again, this is more of a pragmatic folk-view than anything based on a clear theory."(Granger & Joen, 2004, p. 335).In this article, we develop a theoretical framework describing the trimmed combination for N different individual forecasts. This approach allows to investigate both cases of trimming described above. For the first trimming approach, we derived a criterion which allows us to understand why and when it makes sense to remove a certain forecast from the combination. It turns out that this criterion is not only determined by the statistical properties of the forecast itself. Therefore, all performance criteria that only consider the individual forecast itself (e.g. the MSE) are unsatisfactory. It may be beneficial to use even the worst individual forecast in a combination, as this, for example, shows advantageous covariance properties with other forecast errors. For the second type of trimming, our analysis shows that it is a similar but less restrictive form of the first trimming approach. In this case, however, the variances and covariances of the forecast errors are only partially included in the MSE of the combination. Especially terms with a high contribution to the MSE of the untrimmed combination can now tendentially obtain a smaller weight, resulting in a smaller MSE. Whether this form of trimming enhances the combined forecast depends on certain forecast properties and finally follows a similar discussion as the neglect of complete forecasts. Under consideration of typical forecast properties (e.g. high correlations between the forecast errors) our analyzation delivers a theoretical explanation for the findings from empirical investigations. We illustrate our theoretical results with a Monte Carlo simulation.Timmermann, A. Forecast combinations, In Handbook of Economic Forecasting, G. Elliot, C. W. J. Granger, and A. Timmermann Ed., North-Holland, 2006.Granger, C. W. J. & Jeon, Y. (2004) Thick modeling, Economic Modelling, vol. 21, pp. 323-343.

Principals of Combination Forecasts

Presenter: Simon Spavound
Forecast combinations have long been known to have the potential to improve forecast performance over any individual model. The methods used for effectively combining these disparate forecasts are still under some debate. Typically, parsimonious combination techniques are found to be superior to any more complex

methodologies. Our motivation is straightforward; can the most salient features from a variety of forecasts be extracted, and sensibly combined, to improve out of sample forecast performance? We propose utilising Principal Component Analysis on the forecast set and combining the generated features to forecast out-of-sample. PCA reduces the dimensionality of the entire forecast set (which can be arbitrarily large) and extracts orthogonal components which may be representative of features of the underlying time series of interest. The final model attempts to determine the most useful features of the basket of forecasts and combine them to assess forecast performance.We test our approach on a variety of different series to examine under which (if any) conditions PCA forecast combination is effective and provide some further guidance for practitioners; comparing this performance against other well-known combination techniques.

Forecasting Industrial Production Using the State Dependent Models

Presenter: Bo Guan

This study evaluates the forecasting performance of a general class of non-linear models called "State Dependent Models" SDM, which is developed by Priestley in (1980). Three different forecasting methods of state dependent models - the traditional recursive method, and two new forecasting methods involving a smoothed and an unsmoothed coefficient grid search methods were considered and employed in forecasting industrial production indices for the UK. Their forecast performances were then compared with ARIMA, Error Trend Seasonality (ETS) model and neural network autoregressive models in forecasting the eight industrial production series of seasonally unadjusted industrial production indices for the United Kingdom. The results show the dominance of the three SDM models compared to other three methods. Improvement in Root Mean Squared Error (RMSE) is more pronounced in the long-run forecasting.

Predicting Group Membership for Interval Time Series

Presenter: Ann Maharaj

An interval time series (ITS), [Xt], is a sequence of intervals observed at successive instants in time, where each interval is represented by its lower and upper bounds; [X]t = [Xt,L; Xt,U], where -8 < Xt,L < Xt,U < 8. Each interval can also be represented by its centre and radius; [X]t = < Xt,C ; Xt,R >, where Xt,C = (Xt,L + Xt,L)/2 and Xt,R = (Xt,U – Xt,L)/2.In many practical situations, it may be more useful to analyse an ITS which provides information about the variability between the upper and lower values at each time point rather than analyse a time series with a single value at each point.We may wish to identify similar patterns of two or more ITS. For example, how similar or how different are sea levels at locations around the coastline of a particular country, given daily maximum and minimum levels? This information could be useful in addressing rising sea levels more efficiently at two or more locations at which the sea levels are similar. Another example could be to determine how similar or how different daily peak and off-peak electricity demand are between two or more locations in a particular city; this information could be useful in addressing electricity demand that is similar in two or more locations, more efficiently.Having identified similar patterns amongst groups of ITS (Maharaj et al., 2019), the aim of this study is to predict group membership of other related ITS. We investigate this by applying linear and quadratic discriminant functions, and the K-nearest neighbour classifier to two specific sets of discriminating features of the given groups of ITS. For the first set, we use wavelet variances at the relevant number of scales of the radius and centre of each ITS. For the second case, we use the Euclidean distance between the upper and lower bounds of each ITS. For all classifiers, we use the hold-out-one cross-validation technique to evaluate the performance of the methods.Simulation studies using ITS generated from space-time autoregressive models (Teles and Brito, 2015), show very good classification rates, and applications to sets of real ITS reveal the usefulness of this approach for ITS classification.ReferencesMaharaj E.A., Teles, P., Brito, P. (2019). Clustering of interval time series. Statistics and Computing, https://doi.org/10.1007/s11222-018-09851-z.Teles, P., Brito, P. (2015). Modelling interval time series with space-time processes. Communications in Statistics: Theory and Methods, 44, 17, 3599-3627.

Correlation functions of complex-valued variables in economic forecasting

Presenter: Sergey Svetunkov

The economic processes are complicated in their nature and rely on a variety of non-linear relations, which are difficult to capture using the conventional models. Forecasting such processes is one of those challenging tasks that can be improved by the usage of more sophisticated mathematical instruments.One of the options in this case is the use of the complex variables theory, which seems to be a promising direction. The complex variable z is a pair of real numbers y and x, connected with each other by an imaginary unit: z=y+ix.In economic forecasting, various functions of complex-valued variables can be used, among which the more interesting for the purposes of extrapolation is the complex autoregression, when a complex variable at a moment of time depends on its past values. In order to efficiently analyse any time series processes, variances, covariances and correlation moments are used. Similarly, the complex-valued processes can be analysed by the respective complex-valued statistics.We propose to consider these characteristics in the case of random variables as complex-valued and develop a new theory based on that. This gives new results that allow better analyse the existing processes. For example, the research of the complex correlation coefficient has demonstrated that its real part characterizes the closeness of the relation between two complex random variables to the linear. As for the imaginary part, it contains the information about the degree of variation of variables around a linear function. Knowing these properties of the complex pair correlation coefficient, it is possible to carry out precise data analysis and construct more sophisticated complex-valued models.This is especially useful in case of time series, where a set of complex-valued random economic variables, z[t], z[t-1], z[t-2], z[t-3], etc. needs to be analysed in order to determine and order of a complex autoregression model. Using autocorrelation function in this context is one of the possible solutions to the problem, allowing making conclusions about the structure of the time series and the existing patterns.Since we deal with complex-valued correlation coefficient, the autocorrelation function will have an unusual appearance. The correlograms in this case will be three-dimensional, where the x- and y- axes correspond to the real and the imaginary parts of the complex correlation coefficient and the z-axis depicts the lag t.Analysis of the autocorrelation functions and their correlograms allows determining the structure of the complex-valued processes in economics in order to identify the cyclical nature of their development produce more adequate economic forecasts.

Definitions and a Forecasting Method for Single-Valued Neutrosophic Time Series

Presenter: Ali Zafer Dalar

In the literature, fuzzy sets and systems have been used for obtaining forecasts. Fuzzy time series is one of the common classes of fuzzy methods. Fuzzy time series forecasting methods do not need strict assumptions unlike classical time series. Fuzzy time series was firstly defined by Song and Chissom (1993). In this study, single-valued neutrosophic time series is firstly defined. Neutrosophic sets are introduced by Smarandache (1998). Single-valued neutrosophic sets are special form of neutrosophic sets and their membership values are real numbers. Single-valued neutrosophic sets were proposed by Wang et al. (2010). In this study, single-valued neutrosophic time series definition and high order single-valued neutrosophic time series model are proposed. Moreover, a forecasting algorithm is proposed to solve the proposed model. The forecasting algorithm uses single-valued neutrosophic fuzzy clustering algorithm to create membership values for each membership functions. Single multiplicative neuron model artificial neural network is used to estimate G function. The training of single multiplicative neuron model artificial neural network is done by particle swarm optimization. The proposed algorithm is applied for some real-world time series data sets. The proposed approach's forecasting performance is compared with well-known forecasting methods. References

Song, Q. and Chissom, B.S. (1993). Fuzzy time series and its models. Fuzzy Sets and Systems, 54, 269-277.

Smarandache F. (1998). Neutrosophy. Neutrosophic Probability, Set, and Logic. Amer. Res. Press, Rehoboth, USA, 105 p.

Wang, H., Smarandache F., Zhang Y.Q., Sunderraman R., (2010), Single valued neutrosophic sets, Multispace Multistruct, 4, 410-413.

Li, Q., Ma, Y., Smarandache, F., Zu, S., (2018) Single valued neutrosophic clustering algorithm based on

Tsallis entropy maximization, Axioms, 7, 57, 1-12.

Single-Valued Neutrosophic Inference System for Time Series Forecasting based on Ridge Regression and Particle Swarm Optimization

Presenter: Ufuk Yolcu

Neutrosophic sets are firstly introduced by Smarandache (1998). Single-valued neutrosophic sets proposed by Wang et al. (2010) are special form of neutrosophic sets. Single-valued neutrosophic sets are extension both fuzzy sets and intuitionistic fuzzy sets. These sets can be used in an inference system and these sets provide more additional information to real observations compared with fuzzy sets and intuitionistic fuzzy sets. In the literature, Mamdani fuzzy inference system, Takagi-Sugeno-Kang fuzzy inference system and ANFIS have been used for forecasting time series. While these systems used fuzzy sets, there are other fuzzy systems based on intuitionistic sets. In the literature, neutrosophic sets have not been considered in an inference system for forecasting. With this study, a new inference system based on single-valued neutrosophic sets is proposed. The new inference system uses more membership values than fuzzy or intuitionistic fuzzy inference systems. The new inference system uses ridge regression to obtain neutrosophic functions. Each membership values are obtained from single-valued neutrosophic clustering algorithm proposed by Li et al. (2018). Final predictions of the system are created from combination of predictions for membership functions. Combination weights are obtained from particle swarm optimization. The proposed system is applied for real-world data sets and compared with some other forecasting methods. Key words: Forecasting, single valued neutrosophic sets, inference system, ridge regression, single valued neutrosophic clustering. References1- Wang, H., Smarandache F., Zhang Y.Q., Sunderraman R., (2010), Single valued neutrosophic sets, Multispace Multistruct, 4, 410-413. 2- Smarandache F. (1998). Neutrosophy. Neutrosophic Probability, Set, and Logic. Amer. Res. Press, Rehoboth, USA, 105 p. 3- Li, Q., Ma, Y., Smarandache, F., Zu, S., (2018) Single valued neutrosophic clustering algorithm based on Tsallis entropy maximization, Axioms, 7, 57, 1-12.

Picture Fuzzy Inference System for Forecasting Time Series based on Ridge Regression and Genetic Algorithm

Presenter: Eren Bas

Recent years, fuzzy inference systems are efficient tool for solving forecasting problems. Fuzzy inference systems are based on fuzzy sets proposed by Zadeh (1965). Neutrosophic sets are generalized form of fuzzy sets. The neutrosophic sets are introduced by Smarandache (1998). Picture fuzzy sets are special form of neutrosophic sets. Fuzzy inference systems use membership values besides original data so a data augmentation mechanism is employed in the fuzzy inference. Picture fuzzy sets provides additional information to original data via positive degree membership, negative degree membership, neutral degree membership and refusal degree membership apart from fuzzy sets. The data augmentation with these additional information will be provide to build a better inference system than fuzzy inference systems. In this study, picture fuzzy inference system is proposed for forecasting purpose by using multiple linear regression and genetic algorithm. Ridge regression method is used to obtain picture fuzzy functions and genetic algorithm is used to emerge different information coming from systems which are designed for positive degree membership, negative degree membership and neutral degree membership. In the proposed method, picture fuzzification is provided by picture fuzzy clustering proposed by Thong and Son (2016). The proposed inference system is tested by using various stock exchange data sets. The proposed method is compared with well-known forecasting methods. The obtained results are evaluated according to different error measures such as root of mean square error and mean of absolute percentage error. Key words: Forecasting, picture fuzzy sets, inference system, ridge regression, picture fuzzy clustering. References1- Zadeh L.A. (1965). Fuzzy Sets. Inform and Control, 8,338-353.2- Smarandache F. (1998). Neutrosophy. Neutrosophic Probability, Set, and Logic. Amer. Res. Press, Rehoboth, USA, 105 p. 3- Thong P.H., Son L.H. (2016). Picture fuzzy clustering: A new computational intelligence method, Soft Computing, 20: 3549-3562.

Picture Fuzzy Time Series Definitions and a Forecasting Method to Forecast to Picture Fuzzy Time Series

Presenter: Erol Egrioglu

Forecasting problems have been solved by using fuzzy sets and their applications in the literature. Fuzzy time series was defined by Song and Chissom (1993). Fuzzy time series is a kind of time series and its observations are fuzzy sets or fuzzy numbers. In the literature, there are different fuzzy time series methods and many of them do not use classical Song Chissom's fuzzy time series definitions. Many of fuzzy time series methods need to use new fuzzy time series definitions. To overcome these problems, Egrioglu et al. (2018) proposed new fuzzy time series definitions, intuitionistic fuzzy time series definitions, fuzzy time series forecasting models and intuitionistic fuzzy time series forecasting models. Neutrosophic sets are generalized form of fuzzy and intuitionistic fuzzy sets. The neutrosophic sets are introduced by Smarandache (1998). Picture fuzzy sets are special form of neutrosophic sets and it is also called as standard neutrosophic sets. In this study, picture fuzzy time series and a picture fuzzy time series forecasting model are defined. Moreover, a new picture fuzzy time series forecasting method is proposed. The proposed method has basic three steps. These steps are picture fuzzification, model construction and forecasting. In the proposed method, picture fuzzification is provided by picture fuzzy clustering proposed by Thong and Son (2016). Positive, neutral and negative memberships are obtained by picture fuzzy clustering. Model construction step contains estimating G function. In this study, G function is preferred to estimate via Pi-Sigma artificial neural network. References

Song, Q. and Chissom, B.S. (1993a). Fuzzy time series and its models. Fuzzy Sets and Systems, 54, 269-277.

Smarandache F. (1998). Neutrosophy. Neutrosophic Probability, Set, and Logic. Amer. Res. Press, Rehoboth, USA, 105 p.

Thong P.H., Son L.H. (2016). Picture fuzzy clustering: A new computational intelligence method, Soft Computing, 20: 3549-3562.

Egrioglu, E., Yolcu, U. & Bas, E. Granul. Comput. (2018). https://doi.org/10.1007/s41066-018-00143-5.

M4 Competition Winning Method

Presenter: NA

M4 Forecasting Competition in 2018 was won by a hybrid method that extended computational graph of a Recurrent Network stack with some Exponential Smoothinginspired formulas. The talk will introduce the method, retrace its development steps, and put it in a broader context of Differentiable Programming.

The Delphi Technique: Does (Group) Size Matter?

Presenter: Fergus Bolger

A key question for anyone planning to use the Delphi technique – a commonly used method for judgmental forecasting with groups of experts – is how many panel members do they need to recruit? To date, there is little empirical research exploring the relationship between Delphi group size and performance, and what past findings there are have been inconclusive. While perceived wisdom suggests "bigger is better", larger panels are more costly, time-consuming and logistically challenging. In addition, the Delphi process involves panellists reviewing other panellists' responses, which often include written rationales in support of those responses. This can be a cognitively demanding process. Research based on Cognitive Load Theory suggests that increasing the volume of information a person needs to process beyond a certain point can create information overload, leading to a rapid decline in task performance. It is also generally proposed that Delphi exercises should work best with a diverse group of individuals, who can provide a wide range of perspectives, opinions, experience and so forth. However, as with information quantity, there is evidence that exposure to greater information diversity can increase the risk of information overload, which could in turn impact on the effectiveness of the Delphi process. In the present study, participants answered 10 geopolitical forecasting questions and then reviewed 6, 9, 12, 15 or 18 answers (i.e. Delphi groups of 7, 10, 13, 16 and 19) – supported by either diverse or uniform rationales – and then had the opportunity to change their original answers. Participants' judgment accuracy was measured along with their confidence, likelihood of changing answer post-Delphi, perceived cognitive load, and other self-report measures. The study used

a novel online protocol to simulate 'real' Delphi groups by providing participants with pre-prepared sets of responses collected in advance from other participants. Group size was found to have no effect on the accuracy of forecasting judgments or confidence. Participants in larger groups were, however, significantly more likely to change their answer – most likely in groups of 13 and 16 – and were slightly more likely to change a correct answer for an incorrect one (but this did not translate into a reduction in accuracy overall). Giving participants a diverse set of rationales was found to improve judgment accuracy and reduce confidence, though (both statistically significant, though small-sized, effects). There was no effect of group size or diversity on participants' perceived cognitive load, except for groups of 19 given diverse rationales, who found the task somewhat more cognitively demanding. Overall, this study should provide some reassurance for those struggling to recruit panel members for a Delphi exercise: group size per se may not be important, provided the group represents a sufficiently diverse range of viewpoints on the issues under investigation. Further research should include groups larger than 19 and smaller than 7 and explore the effects of opinion diversity in more detail.

BVAR Forecasts, Survey Information and Structural Change in the Euro Area

Presenter: Florens Odendahl
We incorporate external information extracted from the European Central Bank's Survey of Professional Forecasters into the predictions of a Bayesian VAR, using entropic tilting and soft conditioning. Both methods significantly improve the plain BVAR point and density forecasts. Importantly, we do not restrict the forecasts at a specific quarterly horizon, but rather their possible paths over several horizons jointly, as the survey information comes in the form of one-year-ahead expectations. Besides improving the accuracy of the variable that we target, the spill-over effects to "other-than-targeted" variables are relevant in size and statistically significant. We interpret the results as evidence that survey forecasts can help mitigate the effects of structural breaks such as the recent crisis on the forecasting performance of a popular macroeconometric model.

Producing "effective" scenarios: an evaluation of enhancements to the Intuitive Logics scenario development method.

Presenter: George Wright
We consider recent augmentations of the Intuitive Logics scenario development method and evaluate whether these augmentations enhance the basic method's capabilities. We find that there is a strong case for arguing that these scenario methods are designed to address two of the three objectives that we identified from the literature, namely: (i) enhancing understanding: of the causal processes, connections and logical sequences underlying events — thus uncovering how a future state of the world may unfold, and (ii) challenging conventional thinking in order to reframe perceptions and change the mind-sets of those within organizations. However, other than in the augmentation in which multi-attribute value analysis is applied, none of the recent developments that we detail address directly the third objective that we distil from the literature: (iii) improving decision making: to inform strategy development. We conclude by considering the new methodological developments of (i)"Scenario Improvisation" (Cairns, Wright and Fairbrother, 2015), (ii) "Branching Scenarios" (Cairns, Wright, Fairbrother and Phillips, 2017), and (iii) Horizon Scanning based on the Backwards Logic scenario method - and evaluate their separate usefulness.

Forecasting Italian spot electricity prices using random forests and intra-daily market information

Presenter: Luigi Grossi
In this paper we estimate random forest (RF) models in the time period January 2014 - April 2018 in order to forecast electricity prices on the Italian spot market. Moreover, we introduce information about intra-daily markets within a set of explanatory variables which enable to improve the forecasting performance of the models. Random forest (Breiman, 2001) belongs to the family of ensemble learning models, with the decision tree, which is applied also in the time series context (Ahmed et al., 2010), as base learner. Ensemble learning models are methods developed for reducing model instability and improving the accuracy of a predictor

through the aggregation of several base learners. RF combines k decision trees based on bagging and random selection of input features. Petropoulos et al. (2018) have explored the applicability of bagging for time series forecasting. The variables involved are a response variable and a set of covariates. In addition to the predictor, it is possible to compute, for each covariate, a variable importance (VI) measurement, which is defined as total heterogeneity reduction produced by a given covariate on the response variable when the sample space is recursively partitioned. It was observed that VI measures tend to favor covariates having more values, so the method introduced by Sandri and Zuccolotto (2008) was used in order to eliminate this kind of bias. In the present paper, RF have been estimated for all phisical zones through which the Italian market is organized: North, Centre-North, Centre-South, South, Sicily, Sardinia. For each zone we have collected hourly spot prices, prices in each of the 5 intra-daily markets (M1, M2, M3, M4, M5), one-day-ahead forecasted electricity demand and one-day-ahead forecasted wind generation. The four months of 2018 have been left out for forecasting purposes. Observed spot prices have been compared with one-day-ahead forecasts obtained through a rolling window procedure. Results have been compared with those obtained from the estimation of a linear AutoRegressive (AR) model with the same regressors, which is considered as a benchmark. Taking advantage of the variable importance index, which is intrinsically computed by the RF, we have been able to rank regressors according to their relevance in different zones. The nonlinear nature of RF looks to help better capturing the link between spot prices and other regressors in the northern macro-zones, while the AR model can be considered equivalent to RF in the southern macro-regions where wind generation plays a crucial role in forecasting spot prices. References Ahmed N.K., Atiya A.F., El Gayar N., El-Shishiny H. (2010) An Empirical Comparison of Machine Learning Models for Time Series Forecasting. Econometric Reviews, 29:5-6, 594-621. Breiman L. (2001) Random Forests. Machine Learning, 45, 5–32. Petropoulos F., Rob J. Hyndman R.J., Bergmeir C. (2018) Exploring the sources of uncertainty: Why does bagging for time series forecasting work? European Journal of Operational Research, 268, 545–554. Sandri M., Zuccolotto P. (2008) A bias correction algorithm for the Gini variable importance measure in classification trees. Journal of Computational and Graphical Statistics, 17(3), 1–18.

Forecasting Northern Italian Electricity Prices

Presenter: Anna Gloria Billé

This paper predicts day-ahead hourly regional electricity prices. Focusing on the Northern Italian market, models based on econometric properties and fundamentals are implemented. Among regressors, forecasted demand, traditional fossil productions, renewable energy sources (RES), flows among Italian zones and neighbouring countries are included. Moreover, linear and nonlinear specifications, such as GARCH models, are included. Results indicate a strong predictive power from forecasted demand at any hour, from RES mainly at peak hours, and a non-diminishing role from fossil productions such as gas prices. Adding GARCH residuals does not improve forecast accuracy because the previous documented time-varying volatility is now captured by the intermittent behaviour of the RES.

Application of a SVM-based model for day-ahead electricity price prediction for the single electricity market in Ireland

Presenter: Conor Lynch

This paper espouses an innovative approach to predict the €/MWh day-ahead electricity prices for the single electricity market (SEM) in Ireland. An upsurge in demand response and the proliferation of distributed energy resources continue to drive the requirement for more accurate and computationally efficient models for forecasting the system marginal price (SMP). This paper presents such a model, a k-means, Support Vector Machine (SVM) and Support Vector Regression (SVM) model (k-SVM-SVR model). While obtaining prediction accuracy comparable with the best known models in the literature, the k-SVM-SVR model requires limited computational effort. The computational efficiency is achieved by eliminating the use of a price feature selection process, which is commonly used by existing models in the literature. The developed model achieved approximately 20% improvement and reduced error variances over the existing predictions available to market participants in Ireland.The k-SVM-SVR model is tested using SMP electricity market data from the periods 2010–11, 2015-16 and 2016–17 respectively for the Irish energy market. Future work will investigate the application of such a model to other energy markets including the now Integrated-Single Electricity Market

(I-SEM) in Ireland and the N2EX market in the UK.Keywords: Energy Price Forecasting, Single Electricity Market (SEM) and Single Marginal Price (SMP)

Big Data and the Macro Economy: A Bayesian Mixed Frequency Estimator

Presenter: David Kohns

In light of recent advances in computing and data availability, nowcasting low interval economic aggregates such as GDP, inflation or industrial production has increased in importance for policymakers and businesses alike. Google Trends in particular have received a lot attention in the nowcasting literature which is mainly attributed to their granularity (30 top-level and 250 second-level categories) and ability to measure latent consumer intent and expectations (Scott and Varian, 2014). When dealing with Big Data such as Google Trends for nowcasting, often encountered issues are that of mixed frequency, high dimensionality and the resulting computational complexity. A further issue, which has mostly been neglected in the nowcasting literature, is that of the economic structure or interpretability of the Big Data included in the nowcasting model.This paper develops a new estimator for nowcasting macroeconomic aggregates with high frequency Big Data that is able to handle any kind of frequency mismatch, high dimensionality, while remaining computationally simple and allowing for economic interpretability of the Big Data included. This is achieved by leveraging Harvey's (2006) state space framework such that in a computationally inexpensive way shrinkage priors, model uncertainty and unrestricted high frequency data are included in the sampling algorithm. By allowing the high frequency parameters to be unrestricted, one can gauge the intra-period dynamics of the Big Data included. The paper further shows how to use marginal inclusion probabilities of the individual high frequency regressors to yield deeper economic interpretation.The estimator is showcased by using Google Trends to augment nowcasts of real U.S GDP growth and is further compared to competing mixed frequency nowcasting models in terms of forecasting fit. We find that the proposed estimator outperforms all other estimators and that Google Trends increase forecasting fit. The tractability of the model shows that Google Trends act as an early warning signal of financial distress.

Nowcasting Monthly GDP: a Model Averaging Approach

Presenter: Alessandro Giovannelli

Gross domestic product (GDP) is the most comprehensive and authoritative measure of economic activity. The macroeconomic literature has focused on nowcasting and forecasting this measure at the monthly frequency, using related high frequency indicators. The paper addresses the issue of estimating monthly gross domestic product using a large dimensional set of monthly indicators, by pooling the disaggregate estimates arising from simple and feasible bivariate models that consider one indicator at a time, in conjunction to GDP or a component of GDP. The weights used for the combination reflect the ability to nowcast the original quarterly GDP component. Our base model handles mixed frequency data and ragged-edge data structure with any pattern of missingness.Our methodology allows to assess the contribution of the monthly indicators to the estimation of monthly GDP, thereby providing essential information on their relevance. This evaluation leads to several interesting discoveries.

Nowcasting macroeconomic variables with a new tool: TF-MIDAS

Presenter: Nicolas Bonino-Gayoso

Economic policy-makers, entrepreneurs and investors need to have access to real-time assessments of the state of the economy. The sooner they have access to information about the real economic situation, the better prepared they will be to make decisions. Unfortunately, data offered by the System of National Accounts is delivered with considerable delay. This delay, as Castle and Hendry (2013) point out, is the result of several difficulties: (i) not all disaggregated data are available when is needed to compute an aggregate; (ii) many disaggregated time series are only preliminary estimates. On the other hand, there is a considerable number of short-term economic indicators available at an earlier stage (e.g., monthly data from consumer surveys, daily data from financial markets). Several classes of models have been proposed to work explicitly with mixed-frequency datasets, most widely used being MIDAS (MIxed DAta Sampling) family of models. MIDAS models (Ghysels et al., 2002, 2003, 2006) are defined in terms of a Distributed Lag (DL) polynomial. This

kind of model has been applied to nowcasting GDP and corporate bond spreads, among other variables. A specific variation of the standard model, known as Unrestricted MIDAS (U-MIDAS), is introduced by Foroni et al. (2015). Based on a series of simulation exercises, these authors state that U-MIDAS' nowcasting accuracy outperforms that one of standard MIDAS, specially when the difference in sampling frequencies is not high, as it is usually in the case of macroeconomic nowcasting. In this paper we apply a transfer function representation to the Distributed Lag (DL) polynomial, deriving a new type of mixed-frequency model, named Transfer Function-MIDAS (TF-MIDAS). We present evidence from Monte Carlo simulation exercises confirming that TF-MIDAS beats currently available MIDAS models in terms of out-of-sample nowcasting performance. Working with simulated data allow us to consider different variants of DGPs and so to identify for which specific processes the advantage of TF-MIDAS is significantly greater. Besides, we consider separately the cases of stock and flow variables, thus taking into account different aggregation rules. Following Foroni et al. (2015), we initially consider a HF-VAR(1) process, finding that there is no considerable difference in nowcasting performance between U-MIDAS and TF-MIDAS in the case of a stock variable, while for a flow variable TF-MIDAS presents a slightly better relative performance. We also consider the cases of a HF-VMA(1) and HF-VMA(3) DGPs. For a stock variable following a HF-VMA(1) process we do not find any significant difference between U-MIDAS and TF-MIDAS performances, but we do find a considerable gain for TF-MIDAS in the case of a flow variable. Finally, in the case of a HF-VMA(3) process, we find that TF-MIDAS shows a considerable better performance than U-MIDAS, whether if the variable is a stock or a flow. We conclude that TF-MIDAS presents an overall much better performance than U-MIDAS. The advantage of TF-MIDAS over U-MIDAS is specially remarkable for DGPs that include a MA component, no matter if the LF variable y is a stock or a flow. These results are robust to different parameter specifications.

Probabilistic reconciliation of hierarchies: a closed-form Bayesian approach.

Presenter: Giorgio Corani

Often time series are organized into hierarchies (hierarchical time series). For example, the visitors of a country can be disaggregated according to region. In this case, the forecast for the whole country should match the sum of the forecasts for the different regions. Independently forecasting each time series yields base forecast that generally do not satisfy the summing constraints; for this reason, they are called incoherent. Algorithms which adjusts the base forecast in order to obtain a set of coherent forecasts are called reconciliation algorithms. The approach originally proposed by Hyndman et al. (2011) and further pursued in later works consists of generating base forecasts for each time series; updating the forecast of the bottom series on the basis oft he forecasts available for the whole hierarchy; obtaining forecast for the whole hierarchy by summing the updated bottom forecasts. Yet such methods only return point forecasts, without probabilistic information. This prevents quantifying the uncertainty of the reconciled forecasts. The approach by Taieb et al. (2017) yields probabilistic forecasts but does not have a closed-form solution. Our novel algorithm yields probabilistic reconciliation in closed-form. We treat the bottom time series as the random variable about which we want to make inference. The base forecasts act as prior information about them. We then note that the upper time series are linear combination of the bottom ones,and that the forecasts for the upper time series can be seen as noisy observation about linear combinations of the bottom series. We adopt Bayes' rule in order to compute the posterior distribution of the bottom time series, given the forecasts for the upper time series. By assuming the forecast to be multi-variate normally distributed, we analytically derive the reconciled forecasts. Our algorithm yields reconciled bottom forecasts that are a weighted linear combination of the forecasts available for the whole hierarchy; the weights depend on the variances of the forecasts. We prove that our approach minimizes the expected value of the mean square error overall the time series of the hierarchy. Experimentally, our approach compares favorably to the state-of-art minT algorithm by Wickramasuriya et al. (2018), simplifying moreover the definition of the covariance matrix. We apply our approach also to temporal hierarchies (Athanasopoulos et al., 2017), in which the same variable is forecasted at different scales. Extensive experiments on the M3 time series show that our approach compares favorably to the state-of-the-art thief algorithm (Athanasopoulos et al., 2017). References Athanasopoulos G,Hyndman RJ, Kourentzes N, Petropoulos F (2017) Forecasting with temporal hierarchies. European Journal of Operational Research 262(1):60–74 Hyndman RJ, Ahmed RA, Athanasopoulos G, Shang HL (2011) Optimal combination forecasts for hierarchical time series. Computational Statistics & Data Analysis, 55(9):2579 – 2589 Taieb SB, Taylor JW,Hyndman RJ (2017) Coherent probabilistic forecasts for hierarchical time series. In: Precup D,

Teh YW (eds) Proceedings of the 34th International Conference on Machine Learning, vol 70, pp 3348–3357
Wickramasuriya SL,Athanasopoulos G, Hyndman RJ (2018) Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. Journal of the American Statistical Association.

Revisiting Bottom-Up Forecasting for Hierarchical Time-Series Data

Presenter: Nicolò Bertani

Collections of time-series can naturally form hierarchies due to geographical or organizational structures that tie them together. For instance, a production facility can serve multiple distribution centers, which in turn serve multiple stores. Clearly, demand aggregates from the stores according to this organizational structure. Accurate forecasts at any level of aggregation are required by separate decision makers.Hierarchical forecasting methods aim at proving accurate forecasts that are coherent with the hierarchical structure, i.e. where forecasts respect aggregation constraints. It is crucial to realize that these structures are merely summing constraint, and that data generation does not take place at every level of the hierarchy. In most applications, data generation only occurs at the most disaggregated level (the bottom level or leaves). Unfortunately, this realization about the data generation process is not shared by all hierarchical forecasting methods.Consistently with the nature of the data generating process, we show that the hierarchy carries no informational value for the purpose of achieving the best estimates for the whole hierarchy, in the least squares sense. Minimizing squared residuals for the hierarchy coincides with minimizing squared residuals for the leaves, where all the randomness is generated. This means that the Bottom-Up (BU) method is best in the least squares sense. However, it must be noticed that the contemporaneous covariance amongst the leaves plays a pivotal role in squared residuals minimization. This role has been ignored in previous literature, and it has led to incomplete comparisons between the performance of BU and its alternatives. We showcase the improvement in the performance of BU when revisited with the inclusion of the contemporaneous covariance, and compare it to the best available alternative, i.e. the reconciliation method (Hyndman et al. 2011, Athanasopoulos et al. 2009, Wickramasuriya et al. 2018). This is operationalized by partially replicating the comparisons of Wickramasuriya et al. (2018).Additionally, we discuss an unaddressed rationality requirement that hierarchical forecasting methods ought to respect: conditional on observations, forecasts for a given node should only change if the number of leaf-nodes that it aggregates changes. For instance, the prediction at the production facility should not depend on whether the model incorporates and hence also predicts the intermediate level of distribution centers. Amongst the hierarchical forecasting methods that have been popularized in previous literature, only BU is both unbiased and hierarchically invariant. In consideration of its performance, its simplest implementation, and its aptitude to produce probabilistic forecasts, we suggest that the revisited version of BU be considered as the reference in hierarchical forecasting, unless idiosyncratic reasons justify the use of an alternative.

Probabilistic Hierarchical Reconciliation via a Non-parametric Bootstrap approach

Presenter: Puwasala Gamakumara

Forecast reconciliation for hierarchical time series involves adjusting forecasts to ensure coherence with aggregation constraints. This idea is now widely used in point forecast reconciliation, but there are only a few papers on probabilistic forecast reconciliation. We introduce a novel non-parametric reconciliation approach to produce hierarchical probabilistic forecasts. This method involves simulating future paths of the whole hierarchy using bootstrapped training errors followed by the reconciliation of these sample paths to assure the coherency of aggregation constraints. We use an extensive Monte-Carlo simulation to find an optimal reconciliation with respect to a proper multivariate scoring rule and show that it is similar to reconciling each sample path via the MinT approach (Wickramasuriya, Athanasopoulos, and Hyndman, 2019) used in point forecast reconciliation. An empirical application on forecasting Australian domestic tourist numbers shows that the proposed non-parametric bootstrap approach improves the accuracy of both point forecasts and probabilistic forecasts.

Personalized Protection of User-Generated Text with Shapley Values

Presenter: Shawn Mankad

Companies seek to enhance the value of their customer data by linking it at the customer level to publicly available user-generated content (UGC) on platforms such as Amazon, Facebook, Yelp, Twitter, Kaggle, Four-Square, Kickstarter, and so on. The effects of such data integration helps companies develop profiles about their customers; however, there are also a number of data privacy implications which need to be studied. Notwithstanding regulatory attempts, such as the Federal Trade Commission's Consumer Reviews Fairness Act that aims "to protect people's ability to share in any forum their honest opinions about a business' products, services, or conduct", we believe a technological solution is most effective to prevent the personal identification of reviewers. In this paper, we study rigorously the ability of a data intruder to reverse engineer the true author of a piece of UGC text subject to practical constraints (e.g., data availability, effort and complexity of the model, and so on.), and use the results to inform text protection policies when anonymity is desired. Specifically, we solve the significant issue of knowing why a particular authorship prediction is made by the de-anonymization method, which the platform could in turn use to implement a personalized protection policy. We solve this issue by introducing the Shapley Value to the privacy forecasting literature, which, for a given test document, provides a positive/negative contribution of each variable towards the final predicted probability. Thus, the Shapley values can be used to develop "reason codes" that direct the platform in its text protection efforts.

Coordinating tax revenue forecasting and Regulatory Impact Assessment (RIA) of the tax laws – evidence form Polish efforts to curb tax frauds and avoidance during 2015 - 2019

Presenter: Kamil Jonski

Budget revenue forecasting involves both (i) business-as-usual forecasting of macroeconomic aggregates and corresponding tax revenues as well as (ii) estimation of the consequences of the systemic changes in the tax system. The former is achieved using quantitative macroeconomic models – and quite well understood, tested and described. The latter is far less straightforward.Typically, changes in the tax system took form of primary legislation – and as such became subject of the impact assessment process (IA). In most OECD economies, it involves so called Regulatory Impact Assessment (RIA) - a brief statement of expected social and economic (including budget) consequences of the adoption of given law. In this context RIA can be viewed as conditional forecasts (data and theory driven statements about future budget revenues given implementation of the law) and should complement ordinary revenue forecasting.In the organizational (bureaucratic) level, that implies some degree of coordination between the two distinct processes – budgetary planning and legislation management (including RIA). The goal of this paper is to examine quantitative performance and bureaucratic coordination of these processes in the context of Polish revenue forecasting during the reform aimed at curbing tax frauds and avoidance. Specifically, it draws on business-as-usual forecasts developed for FY 2005 – 2019 for VAT, PIT and CIT revenues, and RIAs accompanying tax laws prepared during 2015 – 2018. The choice of Poland as a testing ground for the interplay of ordinary budgeting and RIA processes is motivated by three reasons. The first is the scale of tax fraud/tax avoidance problems, and thus significance of plausible gains. For example despite extraordinary growth performance (the only EU economy that haven't experienced recession after 2008 crisis), the peak value of the Polish 'VAT Gap' almost equated the Greek one, according to the CASE (2017) estimates for the European Commission. According to the same source, over 2015-2018 period the gap shrunk by 70 percent. The second is the political and policy relevance, as after 2015 elections new government had embarked substantial social spending program, that was supposed to be financed by curbing tax frauds/avoidance (thus, reliable estimation of the likely gains from implemented tax laws became crucial for appropriate budgeting process). Third, although RIA process is already well rooted in Poland (and was substantially reformed before 2014), the illiberal flavor of the new government cast serious doubts on its integrity (that became perceived – at least by some scholars and commentators – as resembling PR activities rather than rigorous analytical process). Obtained results could inform public policy decisions in the area of structuring tax revenue forecasting process and integrating it with legislative planning and analysis.

The Effects of Data Protection on Forecasting Models

Presenter: Matthew Schneider
Data protection requirements such as access control will have profound effects on the quality and use of data worldwide. Access control restricts the use and sharing of data after it has been gathered. As a consequence, many analytics use cases to include forecasting will be disrupted or altered due to the unavailability of data. To address this problem, we focus on data protection that does not rely on access control and quantify the reductions in accuracy using a variety of forecasting models. We transform the original data to protected data using random noise, rounding, top coding, swapping, aggregating, and synthetic data and illustrate the trade-off between forecast accuracy and privacy. We show that it is possible to limit the loss in forecast accuracy using data protection methods that do not rely on traditional rules found in legislation. Unlike access control, a key advantage of our approach is that a data breach only exposes the protected data.

Forecasting Multivariate Distributions of Exchange Rates, Interest Rates and Commodity Futures: Factor Quantile Model and Multivariate Scoring Rules

Presenter: Yang Han
This article introduces a flexible semi-parametric model for multivariate distribution forecasting where conditional marginals have a common factor structure, their distributions are interpolated from conditional quantiles and the dependence structure is derived from a conditional copula. A comprehensive comparison of a latent factor version of our model with GARCH and copula models for forecasting different multivariate distributions (exchange rates, interest rates and commodity futures) is the first application of proper multivariate scoring rules to evaluate the accuracy of forecasts for financial asset returns. Overall, the Model Confidence Set approach indicates favourable forecasting performance, matching or exceeding the accuracy of more complicated GARCH models.

Denoising the Equity Premium

Presenter: Antonios Alexandridis
NA

Survey forecasts of macroeconomic activity in the Euro area: Country specific versus area-wide information

Presenter: Malte Knüppel
The accuracy of economic activity forecasts for the entire Euro area is compared to the accuracy from aggregated country specific forecasts. In contrast to a well-known earlier study (Marcellino, Stock & Watson, 2003. Macroeconomic forecasting in the Euro area: Country specific versus area-wide information, European Economic Review, vol. 47(1), pp. 1-18), the investigation relies on survey forecasts instead of model-based forecasts. Survey forecasts tend to have important judgemental elements, and such elements might play a larger role in forecasts based on country specific information. The analysis uses survey forecasts from Consensus Economics for quarterly growth rates of real GDP and industrial production in France, Germany, Italy, the Netherlands, and Spain. These 5 countries account for about 80% of the Euro area's GDP. Despite of ignoring all other countries in the Euro area, it turns out that the aggregated country specific survey forecasts appear to predict macroeconomic activity in the Euro area better than survey forecasts for the entire Euro area. However, this superiority is essentially due to a couple of quarters during the Great Recession only. Moreover, the statistical significance of this superiority is only observed with encompassing tests for industrial production, relying on the usual heteroskedasticity- and autocorrelation-consistent standard errors and the corresponding normally distributed test statistic. The significance vanishes if inference is instead based on a recently advocated weighted-periodogram long-run variance estimator and the corresponding t-distributed test statistic.

Evaluating the Joint Efficiency of German Trade Forecasts - A nonparametric multivariate approach

Presenter: Christoph Behrens

In order to make policy or investment decisions, economic agents rely on professional forecasts (Carroll, 2003). The efficiency of professional forecasts is therefore of crucial importance, especially when it comes to a field receiving increased public attention, such as German trade policy, which has been in the center of political debates in recent years. I contribute to research on multivariate forecast evaluation by analyzing the joint efficiency of annual German export and import growth forecasts by leading economic research institutes for the years 1970 to 2017 in a multivariate setting.Despite Germany's role as one of the largest exporters in the world, the literature on the evaluation of German trade forecasts is very scarce. I focus on the evaluation of the joint efficiency of trade forecasts for Germany in a multivariate setting, as both export and import volumes make up a country's net exports. Since net exports are often in the focus of political debates on the introduction of protectionist trade policies, it is crucial to consider both trade aggregates in a multivariate approach. Furthermore, macroeconomic aggregates, such as the exchange rate, tend to influence export volumes as well as import volumes at the same time (Engelke et al., 2019).I adapt an approach brought forward by Behrens et al. (2018) who analyze the joint efficiency of German GDP growth and inflation forecasts using multivariate random forests. To this end, I compute, in a first step, multivariate random forests in order to model links between forecast errors and a forecaster's information set, consisting of several trade and other macroeconomic predictor variables (on multivariate random forests, see Segal & Xiao, 2011). The reason I use a nonparametric tree-based approach is that linear forecasting models or evaluation techniques run into problems with the data at hand, as they exhibit relatively few and irregularly spaced observations, as well as possible nonlinearities among the predictor variables or between predictor and response variables. A nonparametric, data-driven approach overcomes resulting problems, such as a lack of degrees of freedom or model misspecification issues. As performance criterion, I use the Mahalanobis distance as it is a superior performance metric, in a multivariate setting, compared to the Euclidean distance, since it captures correlations between the response variables (on the Mahalanobis distance, see McLachlan, 1999). In a second step, I compute permutation tests to check whether the Mahalanobis distance between the predicted forecast errors for the trade forecasts and actual forecast errors is significantly smaller than under the null hypothesis of forecast efficiency.I find evidence against joint forecast efficiency for two forecasters, however, for one forecaster I cannot reject joint forecast efficiency. For the other forecasters, joint forecast efficiency depends on the examined forecast horizon. By means of variable importance plots, I find evidence that real macroeconomic variables as opposed to trade variables are inefficiently included in the analyzed trade forecasts. Finally, I compile a joint efficiency ranking of the forecasters.

German Forecaster‘s Narratives: How Informative are Forecast Reports for German Business Cycle Forecasts?

Presenter: Karsten Müller

The paper investigate and evaluate Forecast reports for the German business cycle via textual analysis methods. The main research question is how useful and informative are German forecaster's narratives for forecast consumers as well as for forecaster itself. Clements and Read (2016) use the Inflation reports of the Bank of England, Sharpe et al. (2017) use FED's Greenbook to analyse whether and how informative are forecaster's narratives for England, respectively the US. Both studies find that the narratives contain useful information and that changes in sentiment (expressed in the narratives) can predict subsequent changes in the point forecasts. With reference to cited studies, the paper analyse German forecaster‘s narratives and the question, how informative are forecast reports for German Business Cycle Forecasts? I use a longitudinal and cross section sample of German forecast reports which we collected for the development of the German forecast database founded by the German Research Foundation (DFG)-Priority Programme 1859: “Experience and Expectation. Historical Foundations of Economic Behaviour”. The sample includes 11 forecasters with one to four reports (projections) per year from 1993 until 2017 and covers nearly 700 documents. Based on linguistic pre-processed text corpus, I apply textual analysis tools to analyse the informational content of German forecast reports. In a first step, frequency analysis is used to determine term frequencies and document structure for comparison with the economic development in Germany. In a second step, I evolve sentiment

and uncertainty indices by mapping textual data into numerical indices. Following Di Fatta et al. (2015), who argue that sentiment indices have to be adapted to the content which they are applied, I use a dictionary that take the special meaning of words in an economic forecasting context into account. After that, I use econometric forecast evaluation techniques to test for forecast efficiency and rationality of (numerical) GDP growth and Inflation point forecasts and test whether forecaster's narratives can be used to enhance the accuracy of point forecasts. Furthermore, I examine whether the sentiment and uncertainty indices are able to predict future changes in numerical forecasts or if they have incremental power for predicting the economy, specifically GDP growth or inflation.

Forecasting the impact of recruitment decisions in soccer

Presenter: Ian McHale
In the first part of the talk, I will consider recent developments in the forecasting in sport literature including the impact of, amongst other things, 'big data' and machine learning. The second part of the talk will look at some new work on estimating the impact on team results of changes player line-ups in soccer. At the heart of the problem is a forecasting model which uses player ratings to estimate the scoring rates of the two teams playing in a single match.

Recurrent Neural Networks for Time Series Forecasting: An Overview and Empirical Evaluations

Presenter: Hansika Hewamalage
Accurate time series forecasting affects many domains such as transportation, tourism and healthcare. Popular statistical models such as Exponential Smoothing (ETS) and ARIMA have traditionally supported forecasting in a univariate context. Owing to their relative simplicity and efficiency, these statistical models have dominated forecasting for a long time. However, in the current context of Big Data, the availability of many related time series opens up the opportunity to develop global models which simultaneously learn from many homogeneous time series. Thus, Neural Networks (NN) become the ideal candidate for forecasting in this era.Although the research regarding NNs has presented many complex novel NN architectures, empirical evaluations against statistical methods have challenged their theoretical capabilities. This brings the question whether NNs are really such powerful models for forecasting as anticipated. The lack of established guidelines to tune NNs for forecasting leaves most novel NN architectures unused and ignored in practical forecasting tasks. Newly introduced NN architectures are often not rigorously tested for their performance against state-of-the-art statistical techniques. Furthermore, such research publications are often not accompanied with reproducible code samples. To bridge these gaps, this research aims to carry out an extensive empirical study of existing Recurrent Neural Network (RNN) architectures for forecasting across a number of datasets. In particular, we consider univariate forecasting while leveraging cross-series information on datasets with single seasonality, without using any exogenous variables. The performance is compared to results of ETS, applied on the same datasets. All the implementations are integrated into a single open-source software framework. A guideline is introduced detailing how different RNN architectures work on datasets with different characteristics.

Deep Learning for Forecasting Model Selection

Presenter: Sasan Barak
Deep Learning has achieved breakthrough accuracy in classification tasks of image, speech and general pattern recognition. As a result, the underlying algorithms of deep neural networks (DNN) have seen a resurgence of interest across disciplines, including time series forecasting. However, these applications of DNNs see them applied as forecasting algorithms, similar to conventional neural network algorithms, using autoregressive input vectors, and thus far removed from their original domains of classifying image data. However, in forecasting model selection such applications of image recognition exist. Traditionally, expert-based forecasting model specification utilises time series graphs, seasonal (year-on-year) plots, autocorrelation functions and spectral analysis charts in order to identify the existence and type of seasonality, trend, outliers and structural breaks, serving as model selection filters to narrow down the choice of potentially useful models. While visual data

exploration allows accurate forecasting model selection, it does not facilitate large-scale automation of model selection over many individual time series. In this paper, we propose a novel use of deep learning in time series image recognition for model selection. We train deep neural networks on an image of a time series for a multi-class classification of its patterns of the level, trend, seasonality, and combination of these components, thus select between Exponential Smoothing (ETS) base learners. We assess the efficacy of the Convolutional (CNN) neural networks on a synthetic time series created with different patterns and noise distributions. Results are compared to benchmarks of statistical tests for seasonality and trend, aggregate model selection, and wrapper-based model selection using information criteria and forecast errors. Our results show the capability of DNNs to identify time series patterns directly from graphs, and improve accuracy on statistical tests and wrappers whilst being more efficient in computing resources.

Predicting Monetary Policy Using Artificial Neural Networks

Presenter: Natascha Hinterlang
This paper analyses the forecasting performance of monetary policy reaction functions using U.S. Federal Reserve's Greenbook real-time data. The results indicate that artificial neural networks are able to predict the nominal interest rate better than linear and nonlinear Taylor rule models as well as univariate processes. While in-sample measures usually imply a forward-looking behaviour of the central bank, using nowcasts of the explanatory variables seems to be better suited for forecasting purposes. Overall, evidence suggests that U.S. monetary policy behaviour between 1987-2012 is nonlinear.

Day-Ahead vs. Intraday — Forecasting the Price Spread to Maximize Economic Benefits

Presenter: Weronika Nitka
Recently, a dynamic development of intermittent renewable energy sources (RES) has been observed. In order to allow for the adoption of trading contracts for unplanned events and changing weather conditions, the day-ahead markets have been complemented by intraday markets; in some countries, such as Poland, balancing markets are used for this purpose. This research focuses on a small RES generator, which has no market power and sells electricity through a larger trading company. The generator needs to decide, in advance, how much electricity is sold in the day-ahead market. The optimal decision of the generator on where to sell the production depends on the relation between prices in different markets. Unfortunately, when making the decision, the generator is not sure which market will offer a higher price. This article investigates the possible gains from utilizing forecasts of the price spread between the intraday/balancing and day-ahead markets in the decision process. It shows that the sign of the price spread can be successfully predicted with econometric models, such as ARX and probit. Moreover, our research demonstrates that the statistical measures of forecast accuracy, such as the percentage of correct sign classifications, do not necessarily coincide with economic benefits.

Enhancing wind and solar generation forecasts to yield better short-term electricity price predictions

Presenter: Tomasz Weron
Recently, a rapid development of renewable energy sources is observed, among which wind and solar play a central role. Although it seems highly beneficial, resulting in the reduction of the greenhouse gases and the decrease of wholesale electricity prices, expanding renewables creates new challenges for the market and its participants. Total dependence on weather conditions makes energy generation volatile and difficult to forecast. In order to compensate for this, prediction of electricity prices may be supported by weather forecasts. However, more can be done. This research shows, that wind and solar energy production forecasts can be further improved by utilizing the information contained in past production values. Moreover, it examines the impact of this correction on predicting electricity prices.

Prediction Intervals in High-Dimensional Regression

Presenter: Sayar Karmakar
We construct prediction intervals for a time-aggregated univariate response time series in a high-dimensional regression regime. Consistency of our approach is shown for cases when the number of observations is much smaller than the number of covariates, particularly for the popular LASSO estimator. We allow for general heavy-tailed, long-memory, non-linear stationary process, fixed and stochastic design matrix and thus generalize the existing literature for prediction consistency. After validating our method against some of the existing forecasting methods in the literature through extensive simulations, we construct prediction intervals for hourly electricity prices over horizons spanning 17 weeks and compare them to selected Bayesian and bootstrap interval forecasts.

Forecast Value Added (FVA) Analysis

Presenter: Michael Gilliland
Does your forecasting process improve the accuracy of your forecast? Or are your efforts – such as elaborate models or judgmental overrides – just making the forecast worse? How would you ever know? Given the considerable time and resources that most organizations spend on forecasting, these are important question. Yet traditional performance metrics, like MAPE, don't provide the answer. This session shows how to get the answer using Forecast Value Added (FVA) analysis.FVA is defined as the change in a forecasting performance metric that can be attributed to a particular step or participant in the forecasting process. FVA analysis is a way to evaluate the performance of each activity in the process, comparing all the way back to a naïve forecast. It is now used by many organizations worldwide to identify those process activities that are failing to make the forecast better – or may even be making it worse. This session begins with an overview of the FVA approach, showing how to map your forecasting process, collect the data, and interpret results. It shows how to create two useful reports (the "stairstep" and "comet chart"), and shares case studies from organizations that have applied FVA with often surprising results.FVA analysis helps you identify and eliminate the waste, inefficiency, and worst practices in your forecasting process. The result can be better forecasts, with fewer resources and less management time spent on forecasting.

How a large manufacturer automates their planning process using Forecast Value Added (FVA) analysis

Presenter: Christoph Hartmann
A manual forecasting process, low forecast accuracy, a high Excel dependency and a weak scalability of existing systems and processes have been bothering this manufacturer's supply chain planning process. The manufacturer decided to use SAS Analytics software to enhance scientific forecasting capabilities and to enable demand sensing and shaping linking historical data with external data sources to make forecast more accurate.Forecast Value Added is important for this manufacturer because it helps to eliminate the non-value adding activities in the planning business and helps to streamline their supply chain planning process. This session begins with a summary of the manufacturer's business challenge and goals as well as a summary of the scientific forecasting and FVA project. In the second part of the session there will be a demo of how this company is analyzing and interpreting FVA information in order to make their supply chain planning more efficient.

Forecasting-focused Climate: Development and Validation of a Measurement Scale

Presenter: Anthonelli White
Recently, the results of three consecutive annual Gartner surveys found that managers ranked forecast accuracy/demand planning related challenges as their number one barrier to achieving goals and objectives in their supply chains (Klappich & Eschinger, 2013). Thomas et al. (2011) reported similar findings from their study wherein managers were asked to identify one issue they would choose to solve if given one wish and a magic wand. Forecasting and demand planning related challenges were among the top 5 issues identified by managers in their study. As has been captured in multiple surveys of organizational forecasting practices, and

has been repeatedly observed in case studies of business organizations, much of the forecasting process takes place at a disaggregate level where individual forecasters review and adjust statistical forecasts for hundreds, potentially even thousands of stock-keeping units (SKUs) (Davydenko & Fildes, 2013; Fildes & Goodwin, 2007; McCarthy-Byrne et al., 2011). According to Morelidge (2016), when it comes to supply planning, it is the quality of forecasts such as these, typically for lower levels of product hierarchies, which are most impactful to forecasting performance. Forecasting literature concerned with the individual forecaster has sought to explain forecaster performance as a function of a variety of factors. However, the explanatory power of this literature has been limited by a lack of focus on theory testing and development (Fildes et al. (2003). In addition to the need for further theory building, there is also a need for research oriented toward issues which inform managerial practice. Syntetos et al. (2016) highlighted this disparity in their assessment of forecasting literature, candidly attributing the gap between forecasting theory and practice to the fact that "... the issues being tackled are simply of no practical importance at all" (p. 16). Similarly, commenting on the lack of forecasting research in logistics and supply chain literature, Thomas et al. (2011) concluded: "there is a clear disconnect between what managers view as relevant and what academics provide through research activities" (p. 664). Given these assessments of both forecasting and logistics and supply chain management research, it is apparent that efforts to help close the gap require both theoretical grounding and consideration of issues of practical relevance to managers. In line with these assessments of both the forecasting and logistics/supply chain management research, this research effort contributes to forecasting management theory by building upon the efforts of Davis and Mentzer (2007). Specifically, in recognition of the importance of the organizational context within which forecasting takes place, Davis and Mentzer (2007) proposed the construct of organizational forecasting climate as a key factor affecting forecasting performance. Through a mixed methods research design, the organizational-level construct was adapted to an individual-level forecasting-focused climate construct. Further, qualitative and quantitative methods contributed to the development of a corresponding measurement scale based on the survey responses of 155 spare parts forecasters. Results of the scale development process are presented and the applicability to future forecasting theory testing and development efforts are discussed.

Urban Residential Real Estate Appraisal Using Transportation Data

Presenter: Wei Shang

Location has been proved to be one of themost important features in real estate appraisal models. However, there is no directvariables to measure the location value of a real estate. Most existingresearches used the value of nearby housings to give an estimation of thelocation value of given housings. However, if the sample trading prices ofnearby houses are insufficient, the location value is hard to estimate withouton site investigations or prior knowledge about local housing market. For thetask of automatic valuation of mass real estates, human involved investigationcould be costly and infeasible. Subway is considered as the most efficientmeans in urban transportation. The subway station network reveals the potentialliving and working demand and thus may serve as an important proxy of the valueof location.This paper aims at using subway transportationflow data to construct transportation attraction variables to forecast residentialreal estate prices in urban area. One week data of more than 300 subway stationsin Beijing is collected. Daily in and out flows as well as working day peakhour in and out flows are calculated, to represent the attraction of a givensubway station. Two housing price data set is collected. One data set includes thetrading prices of 98 communities in one district of Beijing. The other data setis real estate website listing prices collected from the Internet. The second housingprice data set includes three years listing prices in more than 1600communities in Beijing. We use the first data set to construct subway transportationattraction model and use the second data set to test the prediction power ofsuch transportation data in housing price forecast. Both hedonic real estate appraisalmodels and machine learning based models have been employed in our research. Empiricalresults showed the subway attraction variables weighted by distance between thesubway station and the community location may increase the prediction accuracyof the real estate valuation model. Therefore, the proposed approach is usefulin quantifying the location value of residential communities. The subwayattraction variable may hopefully serve as a new predictor in urban real estateappraisal models. Future research will explore other open transportation data sources,such as the satellite maps and government transportation division websites.

Automatic Mass Valuation for Non-Homogeneous Housing Markets

Presenter: Dimitrios Karlis

In recent years big financial institutions are interested in creating and maintaining property valuation models. The main objective is to use reliable historical data in order to be able to forecast the price of a new property in a comprehensive manner and provide some indication for the uncertainty around this forecast. The need for unbiased, objective, systematic assessment of real property has always been important. This need is urgent now as banks need assurance that they have appraised a property on a fair value before issuing a loan and also as the government needs to know the fair market value of a property in order to determine accordingly the annual property tax. In this study we compare various linear, nonlinear and machine learning approaches. We apply a large set of variables, supported by the literature, describing the characteristics of the real estate properties as well as transformation of these variables. The final set consists of 60 variables. We answer the question of variables selection by extracting all available information with the use of several shrinkage methods, machine learning techniques, dimensionality reduction techniques and combination forecasts. The forecasting ability of each method is evaluated out-of-sample is a set of over 30,000 real estate properties from the Greek housing market which is both inefficient and non-homogeneous. Special care is given on measuring the success of the forecasts but also on identifying the property characteristics that lead to large forecasting errors. Finally, by examining the strengths and the performance of each method we apply a combined forecasting rule to improve forecasting accuracy.

Forecasting Commercial Real Estate Yields

Presenter: Alimou Bah

Property yield (or cap rate) forecasts are broadly generated by combining econometric predictions with a subjective market overlay process. The use of subjective market knowledge sometimes leads to forecasts substantially deviating from the econometric models' predictions. This makes the forecast process of yields sometimes too skewed towards a qualitative analysis, and less reliant on technical analysis. How to improve the econometric modelling and forecasting of real estate yields? How to forecast turning points of property yields? In this analysis, we estimate a classic OLS model of yields for the London City office market. Yields are modelled on 10-year government bond as the benchmark pricing asset. On top of that, rent growth expectation is added to the model, as well as a risk premium to reflect the additional risks when investing in property relative to government bonds. Investor's demand for real estate needs is captured in the model by using two variables, one reflecting the general economic environment, and an investor's risk appetite. The former is proxied by city or country GDP growth while the investor's general appetite for risky assets is measured by the corporate bond spread. The model's estimation presents a good fit with high adjusted R-squared, but the model performs weakly for identifying turning points. However, the model appears to be more reliable in forecasting the cumulative yield change over a five-year period.

Forecasting of Sugarcane Production Index in Four Main Regions of Thailand

Presenter: Ekachidd Chungcharoen

Sugar industry generates close to US$1 Billion dollars per year and has put Thailand ranks second in the world in term of export after Brazil. Sugarcane which is the main raw material in sugar production is one of the most important agricultural products of the country. Effective forecasting of sugarcane production index comprising of sugarcane yield per rai (approximately 0.4 acre) and commercial cane sugar (CCS) will provide more efficiency in raw material management of sugar industry. This study aims to compare different types of models; Time-Series Decomposition, ARIMA, three MLR models; MLR with annual rainfall; MLR with periodically rainfall; MLR with LASSO, and combined forecasts in order to forecast production index of sugarcane plantation in four main regions: north, northeast, middle, and east of Thailand. The data on sugarcane area, yields, efficiencies of plantation, and monthly rainfall from 2005 - 2016 has been used for present study. The performances of models are validated by running regression between forecast values and actual Index values. The forecast accuracy shows that ARIMA(4,0,5) model provides R-Square and RMSE of 88.49% and 0.65, respectively. Decomposition model provides significantly lower R-Square and RMSE at 69.40% and 0.95, respectively. Among the three MLR models, the best MLR model which uses average periodic rainfall as predictor has the lowest R-Square and RMSE 49.43% and 1.35, respectively. However,

the combined forecast between MLR and ARIMA with optimal weights provides very satisfactory result with R-square of 91.07% and RMSE of 0.65. This result is much better than the sugarcane yield forecasting model that uses a non-linear function of Normalized Difference Vegetation Index (NDVI) currently practiced by the Office of Cane and Sugar Board, Ministry of Industry of Thailand nowadays.

Trends and shock persistence in precious metal prices

Presenter: Harald Schmidbauer
How fast, if ever, does the price of precious metals recover from a shock? Considering the industrial importance of precious metals, as well as investor requirements, measures of shock persistence can provide essential benchmarks. This study investigates the shock persistence in four major precious metal prices (gold, silver, palladium, and platinum) and two equity indices (Nasdaq and DJIA) reflecting technological and industrial development. After accounting for structural breaks, the persistence of a shock in prices can be assessed through the magnitude of autoregressive parameters and expressed in terms of cumulative impulse response (CIR) and half-life of a shock (HLS). We use weekly data from 2003 through 2019, and state-of-the-art methods known to yield unbiased estimates. Our findings suggest a breakpoint in mid-2008 indicating a regime change. Gold prices are found to recover quickly from shocks. Our findings thus contribute to the discussion whether gold is a "safe haven" for investors.

Forecasting agricultural product and energy prices: A simulation-based model selection approach

Presenter: Robert Kunst
The aim of our contribution is twofold. First, we study whether and to what degree the dynamic interaction between commodity prices and energy prices can be exploited for forecasting. Second, we present informative examples for the simulation-based forecast-model selection procedure. Apart from prediction by competing specifications to be selected from a small choice set, we also explore forecast combinations constructed from a continuum in the same framework. The simulation-based method explicitly permits letting the forecast model choice depend on the intended time horizon of the forecast.With regard to classical Granger causality, the evidence supports a causal direction from food prices to fuel prices, without feedback and somewhat in contrast to our expectations. This causal link, however, only benefits forecasting accuracy at relatively large sample sizes. Similarly, clear evidence on considerable seasonal patterns cannot be fused to a seasonal time-series model that outperforms non-seasonal rivals. Finally, the simulation experiments generally favor the handling of all price series in first differences.

Nowcasting and Short-Term Forecasting of Tourism Demand with a Lasso-MIDAS Model

Presenter: Han Liu
Background: Predicting tourism demand is important for managing destinations, allocating resources, and formulating pricing strategies. However, in the era of big data, an urgent problem needs to be solved in tourism demand forecasting research: how to choose the most predictive variables to improve prediction accuracy.Objective: The perishable nature of the tourism industry makes the accurate forecasting of tourism demand extremely important for tourism decision makers. This study aims to provide accurate nowcasting and short-term forecasting for the demand from mainland Chinese tourists to visit Hong Kong. Methods: This study automatically selected the optimal indicators from more than 100 Baidu search indexes according to 6 aspects that affect tourism demand: "dining," "shopping," "transportation," "tour," "attraction," and "lodging." Then, the mixed-data sampling model augmented by the least absolute shrinkage and selection operator method is used to conduct nowcasting and short-term forecasting of tourism demand. The Lasso method, introduced by Tibshirani (1996), is a dimension reduction technique that filters keywords by introducing a penalized variable into the objective function. Indicator sets with the strongest predictive ability among different prediction lengths can be selected based on recent forecasting performances. To be more specific, this study used the cross-validation approach to comparatively analyze the accuracy of the Lasso-MIDAS model relative to those of the benchmark spatial auto-regressive and moving average (SARMA)

model, the generalized dynamic factor model (GDFM) (Li et al., 2017), the MIDAS model (Bangwayo-Skeete & Skeete, 2015), and the GDFM–MIDAS model in the nowcasting and short-term forecasting of tourism demand.Results: Our experimental results suggest that (1) online data is a useful source of information that significantly improves nowcasting and forecasting performances at different stages of tourism development, whichever data processing techniques are used to select for dimension reduction or explanatory variables. (2) The Lasso variable selection method can identify the most influential information according to the forecasting performances under the given prediction period. Thus, the research can determine a set of optimal predictors, which significantly improves the accuracy and interpretability of the prediction. (3) Compared with the benchmark SARIMA model and the GDFM-MIDAS model, the Lasso variable selection method with broad applicability and objectivity can determine the optimal variable sets for prediction more scientifically and accurately. Moreover, the combination with the MIDAS model deals with the problem of the temporal aggregation of different frequency data to some extent. The nowcasting and short-term forecasting results of this model are more accurate than other models.Conclusion: This paper used Lasso, a penalized variable selection method, successfully selected online search indicators that can improve tourism demand forecasting, and the results of objective variable selection with broad applicability are obtained. Then, combined with the MIDAS model, nowcasting and short-term forecasting for the tourism demand from mainland Chinese tourists to Hong Kong were conducted. The experimental results suggest that the Lasso-MIDAS model is superior to the benchmark SARIMA and the GDFM-MIDAS model in both nowcasting and short-term forecasting significantly. It is a scientific and effective variable screening technique that can obtain more accurate prediction results.

Forecasting tourism demand with web search indexes based on kernel principal component analysis

Presenter: Gang Xie

Search query data (SQD) can be helpful in predicting tourism demand by generating web search indexes. Valuable nonlinear information in SQD may be neglected by researchers. To effectively capture the nonlinear information, we used kernel principal component analysis (KPCA) to extract web search indexes from SQD. As there are generally both linearly and nonlinearly correlated data in the time series of tourist arrivals, we propose hybrid autoregressive integrated moving average (ARIMA) and machine learning models, together with extracted web search indexes, for tourism demand forecasting. An empirical study was conducted with collected SQD and real data of tourist arrivals at Hong Kong. The results suggest that our proposed hybrid models are more accurate than other models because of the nonlinear data processing ability of both the KPCA and machine learning models, which demonstrate the validity of the proposed methodology for forecasting tourism demand with nonlinearity.

Diffusion innovation: evolution of AirBnb listings in Switzerland

Presenter: Miriam Scaglione

The rapid emergence of Airbnb in Switzerland is at the basis for this study. The main objective of this work is to describe the development of Airbnb's offer in Switzerland in terms of temporal and geographical evolution from the perspective of diffusion methods. A study by Larpin, B., Mabillard, J., Scaglione, M., Favre, P., & Schegg, R. (2019). An Analysis of Regional Developments of Airbnb in Switzerland: Insights into Growth Patterns of a P2P Platform, Cham shows a link between the relative weight of AirBnB listings and geographical concentration of traditional accommodation such as hotels. This study also shows an increasing level of supply professionalization, meaning an increasing number of listings managed by the same entity. The aim of the present study is to analyse the effects of these latter phenomena on the dynamics of adoption of the AirBnb platform using Bass models. Are there any differences in the diffusion patterns between regions depending on their tourism intensity? Is the professionalization of the sector accelerating the diffusion process? The data under study are provided by airdna.co and consist of all listings of proprieties for the whole country but aggregated at different geographical levels (town, canton, tourism destinations, etc.) from January 2016 to May 2019 on the platform Airbnb.ch, the URL for Swiss object supply. Strictly speaking, the series under study is the number of objects listed per month for the first time. Keywords: sharing economy, Switzerland, Airbnb professionalization, regional differences, P2P platform, Bass model

Panel discussion: Advancing forecasting research and practice: the contribution of the Institute and its journals

Presenter: Robert Fildes; Spyros Makridakis; Gael Martin; Esther Ruiz; Rob Hyndman
Chair: Robert Fildes, Founding Editor, International Journal of Forecasting (IJF) and Lancaster University Centre for Marketing Analytics and Forecasting. Spyros Makridakis, Founding Editor in Chief, International Journal of Forecasting (IJF) and Director, University of Nicosia Institute for the Future Gael Martin, Professor of Econometrics, Monash University and Associate Editors, IJF Esther Ruiz, Editor in Chief, IJF, from 2019, Universidad Carlos III de Madrid, Spain. Rob Hyndman, Editor in Chief, IJF, 2005-2018, Monash University.

The Institute and the two journals dedicated to forecasting (International Journal of Forecasting, Journal of Forecasting) were set up starting in 1981 to act as a central focus for research into all the wide variety of methods and forecasting practice. The proposed audience for the journals aimed to include practitioners, users and researchers in the social, behavioural, management and engineering sciences. The underlying philosophy was to be truly Interdisciplinary and to bridge the gap between theory and practice. Areas identified as important included: - Evaluation of different methodologies - New Methods in judgement, econometrics and time series - Technological forecasting - Organisational aspects - Impact of uncertainty on decision making Papers would be selected depending on their Importance, competence, replicability and their use of the method of multiple hypotheses. Further exploration and dissemination of research ideas were to be through the Forecasting Symposia, the first being in Quebec City in 1981. Since then the Institute has added Foresight, its practitioner journal as well as focussed workshops and most recently an early career network and summer school to its portfolio.

In this 39th Symposium on Forecasting the aim of this panel is to review these founding objectives and critically evaluate where the journals and the Institute have succeeded and where there have been failures. Finally, panellists will consider the future of forecasting – where gains should be made and where they are likely to occur.

IIF Fellows presentation

Presenter: Mohsen Hamoudia
NA

Remembering an IIF legend

Presenter: Robert Fildes
NA

Awards: best IJF paper and best student presentation

Presenter: Esther Ruiz; George Athanasopoulos
NA

ISF2020 - Brazil

Presenter: Reinaldo Castro Souza
NA

Closing

Presenter: Aris Syntetos
NA