‘Side-taking’ effect on unaided judgement for forecasting decisions in labour conflicts

Vasileios Bougioukos
Wenzhou-Kean University
China

Christos Ioannou
THE GREEK OMBUDSMAN
Greece

There are various concerns reported in the literature of forecasting about the ability of experts to provide accurate unaided estimation in situation of conflict. Due to the complexity of the task, the impact of subjectivity, as ‘side-taking’ effect, from experts (Decision Makers) on unaided forecasts should be investigated further as a factor hindering forecasting accuracy. In order to examine the aforementioned phenomenon, real-life labour conflicts have been distributed to both leaders-representatives of Employers’ Associations and of Labour Unions asking them for their forecasts. The participants were negotiators, who took part in real life labour bargaining processes, as found by the list of the Organisation for Mediation and Arbitration of a European country.

The main goal of this research is to show if there is any ‘side-taking’ effect, due to the position one holds and not due to the expertise, which in turn, might lead to biased forecasts. Matters of using different fairness criteria on forecasting in conflicts are addressed.

A Bayesian Approach to Oil Price and Real Economy Relationship

Qiao Yang
ShanghaiTech University
China

John Maheu
McMaster University
Canada

Yong Song
University of Melbourne
Australia

This paper revisits the predictability of U.S. economic growth by oil price. Existing works focus on point forecasts through the lens of linear models. We apply nonlinear models to investigate the impact of oil price on higher moments of the growth. In addition to point forecasts, we evaluate density forecasts as a measure of the overall predictive power.

A Bayesian Model for Forecasting Time Series in a Hierarchically Structured Organization

Beatriz Etchegaray Garcia
IBM Research
United States

Juie Novak
IBM Research
United States

This talk will focus on the problem of statistical forecasting of hierarchically structured data within a Bayesian framework. We develop a novel approach to hierarchical forecasting that provides an organization with optimal forecasts that reflect their preferred levels of accuracy while maintaining the proper additive structure of the business. We start out by assuming that there is a method in place to obtain forecasts for each node of the hierarchy. We treat the forecasts initially provided as observed data. In order to obtain the desired relative accuracy prescribed by the user, we use the past accuracy of the forecasts at all levels to determine which ones to trust depending on their historical ability to predict. We use a context specific heterogeneous loss function to penalize errors differently depending on the context and which levels we care about most in terms of accuracy. To illustrate our method's ability to improve forecasts, we compare our method to competitor methods. We include two sets of studies: one with simulated data and one with real data.

A Bayesian shrinkage approach to short-term seasonal demand forecasting

Huijing Chen
University of Portsmouth
United Kingdom

John Boylan
Lancaster University
United Kingdom

This paper proposes a Bayesian shrinkage approach to improve the estimation of seasonal patterns that are commonly exhibited in the sales of many products across different industries. Improved forecasting performances can contribute to significant cost reductions in inventory and replenishment planning. Traditionally seasonal indices are calculated from the individual SKU's history. It often does not produce satisfactory results because of short data history and the high level of noise in the data. There have been suggestions in the literature to use additional information. One way of doing it is the Bayesian approach by incorporating prior knowledge in the form of a prior distribution and then updating it when data are observed. Limited research indicates that it is promising for seasonal demand forecasting, but there is a lack of theoretical base. In this paper, we propose a Bayesian shrinkage estimator for a simple additive seasonal model and compare it to the traditional individual approach. We examine the different ways that the prior distribution can be specified, subjectively or empirically, and investigate the effect of a mis-specified prior on the performance of this approach. A James-Stein type of Empirical Bayes estimator is derived when the parameters of the prior distribution need to be estimated from the data. Forecasting performances will be tested on simulated and real data to gain insights as well as providing guidance on implementation.

A Hierarchical model for China's foreign trade prediction

Xun Zhang
Academy of Mathematics and Systems Science, Chinese Academy of Sciences
China

Yuying Sun
Academy of Mathematics and Systems Science, Chinese Academy of Sciences
China

Shouyang Wang
Academy of Mathematics and Systems Science, Chinese Academy of Sciences
China

The importance of understanding the underlying characteristics of foreign trade movements has attracted much attention. However, few works produce consistently good forecasting results because of the intrinsic complexity of exports and imports. In this paper, we propose a novel hierarchical model to forecast China's foreign trade. The foreign trade data are disaggregated from perspectives including trading products, trading partners and product processing forms, which forms three independent hierarchy models with same target variable. Then, a bottom-up strategy is applied. All bottom time series are modelled by corresponding control variables, e.g. China's export to major trading partners are largely affected by each partner's demand. Forecasts for bottom time series are then integrated to generate forecasts for total exports/imports. The final forecasts are obtained by combining three hierarchal models. Empirical experiments show that this hierarchal model significantly outperforms bench mark models and produce consistent forecasts for both total imports/exports or detailed items, which helps a lot for analyzing future trading structure change and making foreign trade policies.

A Joint Volatility Model for Overnight and Trading Day Returns

Vinu CT
IIM Tiruchirappalli
India

Malay Bhattacharyya
IIM Bangalore
India

We propose a class of Periodic AR-GARCH models for overnight, trading day, and daily returns series using opening and closing trading prices. Apart from the timing differences in the overnight (close to open) and trading day (open to close) returns series, in most cases, trading mechanisms are also different in 'close to open' and 'open to close' periods. The proposed class of models captures the relationship between the two series through mean and variance equations. Properties of the models and estimation of the model parameters are discussed. We have considered normal, Pearson Type IV, and Johnson S U distributions for the innovations. The proposed models are tested on five different stock indices. It is found that the interrelationship of the two series is significant in all indices.

A New Model of Inflation, Trend Inflation, and Long-Run Inflation Expectations

Joshua Chan
University of Technology Sydney
Australia

Todd Clark
Federal Reserve Bank of Cleveland
United States

Gary Koop
University of Strathclyde
United Kingdom

A knowledge of the level of trend inflation is key to many current policy decisions and several methods of estimating trend inflation exist. This paper adds to the growing literature which uses survey-based long-run forecasts of inflation to estimate trend inflation. We develop a bivariate model of inflation and long-run forecasts of inflation which allows for the estimation of the link between trend inflation and the long-run forecast. Thus, our model allows for the possibilities that long-run forecasts taken from surveys can be equated with trend inflation, that the two are completely unrelated, or anything in between. By including stochastic volatility and time-variation in coefficients, it extends existing methods in empirically important ways. We use our model with a variety of inflation measures and survey-based forecasts for several countries. We find that long-run forecasts can provide substantial help in refining estimates and fitting and forecasting inflation. The same evidence indicates it is less helpful to simply equate trend inflation with the long-run forecasts.

A Step Towards Demand Sensing: Employing EDI 852 Product Activity Data in Demand Forecasting

Jente Van Belle
Vrije Universiteit Brussel
Belgium

Wouter Verbeke
Vrije Universiteit Brussel
Belgium

This paper deals with short-term demand forecasting of medicines in a US drugs factory based on historical sales and EDI 852 product activity data. Traditionally, demand forecasting relies on statistical methods such as ARIMA, and smoothing methods such as exponential smoothing, to extrapolate the series of historical sales. Although these methods produce reasonably accurate forecasts in many cases, companies are looking for an increasingly higher level of forecast accuracy to further enhance the efficiency of their supply chains. With more and more diverse data becoming available at a higher velocity, it becomes possible to employ non-traditional data sources and generate forecasts for increasingly shorter time periods. From this point of view, in recent years, the concept of demand sensing emerged. However, in the scientific literature and empirical studies, the concept has gained only very little attention to date. Essentially, demand sensing comes down to leveraging a combination of near real-time (i.a. orders), downstream (ideally point-of-sale) and external data to improve the accuracy of short-term forecasts. In their purest form, however, the traditional models do not allow for the inclusion of covariates. In this paper a particular type of downstream data, EDI 852 product activity data, is employed in producing short-term weekly demand forecasts obtained from an ARIMAX, ETS-X, artificial neural network and support vector regression model. The EDI 852 is a data standard used for exchanging product related information (e.g. sales over the previous reporting period and ending inventory levels) between the supplier and the (end)customer. In order to produce the weekly demand forecasts, three years of historical weekly factory sales data and weekly wholesaler sales, ending inventory and receipts for the same time period are used. As a benchmark, also forecasts from an artificial neural network autoregression (NNAR), ARIMA and ETS model are produced. Forecast accuracy is measured out-of-sample by MAE, MAPE and MASE, so that both comparison between the methods as well as comparison across different series is facilitated.

Aggregation of judgmental forecasts: A structured literature review on adopted methodologies

Violina Sarma
Cardiff Business School, Cardiff University
United Kingdom

Aris Syntetos
Cardiff Business School, Cardiff University
United Kingdom

Mohamed Naim
Cardiff Business School, Cardiff University
United Kingdom

Geoffrey Haddock
School of Psychology, Cardiff University
United Kingdom

A structured literature review is conducted to investigate the different methodologies adopted in aggregation of judgmental forecasts. The study follows a content analysis process of material collection, descriptive analysis, category selection and material evaluation. For material collection, English journal articles considering aggregation of judgmental forecasts in the area of manufacturing or production are reviewed. A descriptive analysis of these articles is carried out providing the background for subsequent analysis. The descriptive findings provide insightful knowledge on how the methodological approaches have developed with time. The articles are categorised according to various methodological approaches: research designs, research methods and analysis methods. The selection of categories is based on the existing literature in a deductive fashion. However, a few other categories are added inductively after reading these articles. Finally, the articles are analysed based on the previously mentioned categories. The findings from material evaluation indicate the existence of methodological bias towards quantitative methods in judgmental forecasting research. It is shown that there is an increased use of analytical and simulation methods as compared with experimental and qualitative methods. In addition to the methodological approaches, recent research trends are discussed which help in identifying the present gaps in the literature. To compliment the literature review, exploratory interviews are conducted with managers from different organisations. These interviews help in understanding the aggregation process from an empirical perspective. Based on the structured literature review and exploratory interviews, theoretical propositions are suggested to address the research gap of judgmental aggregation of hierarchical forecasts. By following a structured order, this study allows for a transparent and systematic review with less bias. It aims to advance the practice of judgmentally reconciling forecasts by suggesting theoretical propositions for consensus decision making when such exercises take place in a hierarchical context. It concludes with implications and suggestions for further research.

An Intelligent Multivariate Ensemble Framework for PV Output Power Forecasting

Muhammad Qamar Raza
The University of Queensland, Australia
Australia

Mithulananthan Nadarajah
The University of Queensland, Australia

Australia

Intermittent and partly predictable nature of generated electricity from renewable energy resources especially solar PV create different challenges in electricity market, power system stability and operation. The possible benefits of solar PV integration can be fully received by accurately forecast the solar output power and used in energy planning decisions. In addition, PV output power forecast become more and more crucial for islanded grid operation and energy management system of PV integrated smart buildings. Therefore, multivariate ensemble forecast framework is proposed by combining the multiple predictors by taking the advantage of all of them. A weighted average technique is applied to aggregate the output of the predictors in proposed framework. The historical lagged PV output power, solar irradiation and weather variables such temperature, humidity, wind speed are applied to as forecast framework input. The performance of proposed framework is analyzed by applying training data with different resolution, length and quality based on clearness index. The framework is evaluated using a range of different performance matrix that quantify the point and quantile forecast accuracy. The proposed forecast framework is validated using real time recorded data of The University of Queensland, St. Lucia and Gatton campus.

An Overview of Retail Forecasting

Robert Fildes
Lancster Centre for Forecasting
United Kingdom

The retail industry in developed countries forms a major component of economic activity. It is rapidly changing, in part through developments in on-line selling. Based on a literature survey this introduction will review the limited research on different aspects of retail forecasting, in particular the different problems that must be faced: these include the volume of category, brand and SKU-level sales at distribution centre and store as well as volume of store sales by location. Such different types of important forecasting problem have seen surprisingly little research. This presentation aims to identify the gaps, present the limited evidence on effectiveness and act as an introduction to the remaining papers in this invited session.

Analysis and Short-term Predictions of non-Technical Loss of Electric Power Based on Mixed Effects Models

Paulo Fernando Mahaz Simões
Pontifícia Universidade Católica do Rio de Janeiro
Brazil

Reinaldo Castro Souza
Pontifícia Universidade Católica do Rio de Janeiro
Brazil

Rodrigo Flora Calili
Pontifícia Universidade Católica do Rio de Janeiro
Brazil

José Francisco Moreira Pessanha
Universidade do Estado do Rio de Janeiro
Brazil

In this paper we analyse and predict short-term non-technical loss (NTL) of electric power of Brazilian energy service companies based on different assumptions for the covariance structure of the errors and controlling for socio-economic confounding variables. Although the correlation among repeated responses is not usually of intrinsic interest, it is an important aspect of the data that must properly be accounted for to produce valid inferences in longitudinal or panel data analysis. In the extended linear mixed effects model, the covariance matrix of the response vector is comprised by two sub-components, a random effect component that can represent between group variation and a intraclass or within group component. So as to adequately treat the longitudinal character of NTL data, we use the decomposition of these variance components to evaluate different architectures to the within group errors. Using data of 60 Brazilian distributing utilities from 2004 to 2012, we fit a conditionally independent errors model and three other models with autoregressive-moving average parametrization to the intraclass disturbances. Finally, we compare models using the MAD metric in the prediction of NTL for the year of 2013. The findings suggest that the model with MA(1) conditional disturbances can be used as a benchmark in future statistical analysis of NTL.

Anomaly Detection in Streaming Time Series Data

Priyanga Dilini Talagala
Monash University
Australia

Rob J Hyndman
Monash University
Australia

Kate Smith-Miles
Monash University
Australia

Sevvandi Kandanaarachchi
Monash University
Australia

Mario A. Muñoz
Monash University
Australia

Rapid advances in hardware technology have enabled a wide range of physical objects, living beings and environments to be monitored using sensors attached to them. Over time these sensors generate streams of time series data. Finding anomalous events in streaming time series data has become an interesting research topic due to its wide range of possible applications such as: intrusion detection, water contamination monitoring, machine health monitoring, etc. This paper proposes a framework that provides real time support for early detection of anomalous series within a large collection of streaming time series data. By definition, anomalies are rare in comparison to a system's typical behaviour. We define an anomaly as an observation that is very unlikely given the forecast distribution. The proposed framework first forecasts a boundary for the system's typical behaviour using a representative sample of the typical behaviour of the system. An approach based on extreme value theory is used for this boundary prediction process. Then a sliding window is used to test for anomalous series within the newly arrived collection of series. Feature based representation of time series is used as the input to the model. To cope with concept drift, the forecast boundary for the system's typical behaviour is updated periodically. Examples are provided using data related to an intrusion detection application.

Assessing the performance of prediction models: Development, internal validation, external validation and model updating

Mohammadreza Mohebbi
Deakin University
Australia

Masoumeh Sanagou
Australian Radiation Protection and Nuclear Safety Agency
Australia

Prediction models are increasingly used to clinical reasoning and decision making in general, and in the cardiovascular research, in particular. Developed models need to provide accurate and both internally and externally validated estimates of probabilities of specific health conditions in targeted patients. We review the necessary steps for developing a prediction model. We focuses on how to estimate a model's predictive performance using internal validation techniques, and how to quantify the incremental value of new predictors to existing predictors. In addition to the area under the receiver operating characteristic (ROC) curve (c statistic), and goodness-of-fit statistics for calibration several new measures including reclassification tables, net reclassification improvement (NRI), and integrated discrimination improvement (IDI) has been proposed. Moreover the consecutive steps for assessing model's predictive performance in new individuals, and how to adjust or update existing models to local circumstances or with new predictors has been explained. We conclude that model development should follow a rigorous methodology. The minimum requirements for validation of a prediction model must include an assessment of the agreement between predicted and observed event rates, and a quantification of the model's ability to distinguish between individuals who prognoses the health condition of interest and those who not. Updating an existing prediction model to new circumstances to improve its performance is preferable to developing new models. These steps are illustrated with an empirical example from the cardiovascular field.

Assessing the uncertainty in central banks' macroeconomic outlook

Malte Knüppel
Deutsche Bundesbank
Germany

Guido Schultefrankenfeld
Deutsche Bundesbank
Germany

We assess the accuracy and informativeness of forecast uncertainty issued by the Bank of England, the Banco Central do Brasil, the Magyar Nemzeti Bank and the Sveriges Riksbank. Common to the density forecasts of these four central banks is that they have a two-piece normal (TPN) distribution, and their calibrations are subject to judgemental adjustment by policy-makers. Descriptive statistics of ex-ante and ex-post forecast uncertainty suggest that, while the Riksbank assesses fairly well inflation uncertainty across forecast horizons, the Brasilian and Hungarian central banks are fairly overconfident across forecast horizons. The Bank of England is rather underconfident when assessing inflation and output growth uncertainty at the very short end, yet fairly overconfident at the larger forecast horizon. When approaching the policy horizon of eight quarters out, however, overconfidence on inflation uncertainty is declining. Simple Mincer-Zarnowitz regressions of squared forecast

errors on predicted variances reveal that only the Bank of England's uncertainty forecasts are informative in the sense that a significant positive link between squared forecast errors and predicted variances can be found. This finding, however, does not apply for smaller forecast horizons. The Banco Central Do Brasil's uncertainty forecasts are positively related with squared forecast errors, but the coefficients are not significant. For the Magyar Nemzeti Bank and Sveriges Riksbank, respectively, forecast variances and squared forecasts errors are negatively correlated.

Assisting judgmental adjustments in hierarchical forecasts through interactive Forecasting Support Systems

Ilektra Skepetari
National Technical University of Athens
Greece

Evangelos Spiliotis
National Technical University of Athens
Greece

Achilleas Raptis
National Technical University of Athens
Greece

Vassileios Assimakopoulos
National Technical University of Athens
Greece

A typical practice in business forecasting is to adjust statistical forecasts through judgment, a process which enables experts to incorporate the effect of special factors and other available information. Yet, it has been repeatedly reported that such adjustments may be inconsistent and reduce forecasting performance. The task of judgmental adjustment becomes even more complex when dealing with hierarchical time series, since the expert must perform valid adjustments at multiple time series and then reconcile them appropriately. To assist experts make better decisions, Forecasting Support Systems (FSS) can be exploited. According to the literature, there are two main ways of designing an effective FSS to support judgment: through restrictiveness and guidance. In the first case, the options of users are limited to specific processes, discouraging that way unnecessary adjustments and model modifications. Yet, excessive restrictiveness can discourage users and decrease their engagement in the system. On the other hand, guidance including tips, graphs and past performance measures can motivate user engagement by directly or indirectly supporting decisions throughout the forecasting process. In this respect we propose a hybrid solution which enables the user perform adjustments in hierarchical time series through an interactive interface that incorporates both elements of restrictive and guidance. The proposed FSS allows users to re-arrange the hierarchy easily, for example changing the hierarchical tree of factors (market, product category and features, etc.), to intuitively understand the effect of adjustment and make proper decisions based on tips and error metrics, while at the same time restricting accordingly the potential options. We demonstrate our interactive solution and show how even untrained users can benefit from it and engage to the system to improve their performance.

Bagged artificial neural networks in forecasting inflation: An extensive comparison with currentmodelling frameworks

Karol Szafranek
Warsaw School of Economics, Narodowy Bank Polski Poland

Accurate inflation forecasts lie at the heart of effective monetary policy. By utilizing a thick modelling approach, this paper investigates the out-of-sample quality of the short-term Polish headline inflation forecasts generated by a combination of thousands of bagged single hidden-layer feed-forward artificial neural networks in the period of systematically falling and persistently low inflation. Results indicate that the forecasts from this model outperform a battery of popular approaches, especially at longer horizons. During the excessive disinflation it has more accurately accounted for the slowly evolving local mean of inflation and remained only mildly biased. Moreover, combining several linear and non-linear approaches with diverse underlying model assumptions delivers further statistically significant gains in the predictive accuracy and statistically outperforms a panel of examined benchmarks at multiple horizons. The robustness analysis shows that resigning from data preprocessing and bootstrap aggregating severely compromises the forecasting ability of the model.

Bagging-Clustering methods to forecast time series

Tiago Dantas PUC-RIO Brazil

Fernando Oliveira PUC-RIO Brazil

Bagging Exponential Smoothing procedures have recently arisen as an innovative way to improve forecast accuracy. The idea is to use Bootstrap to generate multiple versions of the time series and, subsequently, apply an Exponential Smoothing (ETS) method to produce forecasts for each of them. The final result is obtained aggregating the forecasts. The main drawback of existing procedures is that Bagging itself does not avoid generating highly correlated ensembles that might affect the forecast error. In this paper we propose and evaluate procedures that try to enhance existing Bagging Exponential Smoothing methods by an addition of a clustering phase. The general idea is to generate Bootstrapped versions of the series and use clusters to select series that are less similar among each other. The expectation is that this would reduce the covariance and, consequently, the forecast error. Since there are several cluster algorithms and dissimilarity measures, we consider some of them in the study. The proposed procedures were evaluated on monthly, quarterly and yearly data from the M3-competition. The results were quite promising, indicating that the introduction of a cluster phase in the Bagging Exponential Smoothing procedures can reduce the forecast error.

Bayesian Argumentation via Delphi (BARD): Eliciting Bayes nets from domain experts for use and reuse in forecasting.

Fergus Bolger
University of Strathclyde
United Kingdom

Real-world forecasting is often a messy business. Key data may be scarce or non-existent and, where available, noisy. Further, there is generally considerable uncertainty regarding aspects of the forecast model, such as, the relevance of variables, and the nature of their interrelations. Often also there will be causal loops such that future actions and events, including those resulting from our own forecasts, affect the things we are trying to forecast. However, help is at hand. Augmented Causal Bayesian Networks (ACBNs) provide powerful tools for representing and manipulating models of the world relevant to real-world forecasting. Advantages of ACBNs include the ability to represent: rationales and confidence (the "Augmented" part); causality (the "Causal" part); and probabilistic connections (the "Bayesian" part). Further, ACBNs permit automatic generation of reports with supporting arguments, and also analysis of, for instance, sensitivity, and updated conclusions in the light of new evidence and changing circumstances. This potential for reuse is a focus of the current study.

Unfortunately, ACBNs are difficult to construct for those without specialist modelling skills. For this reason we are developing a system that allows for structured elicitation of the elements needed to build an ACBN, then assists in putting these elements together and refining the final model. This system ("BARD") is an extension of the well-known Delphi technique that is often used for long-term forecasting. With the help of new software utilizing a graphical user interface, several domain experts, interacting through a facilitator, both individually and collectively build an ACBN in a series of steps with feedback between each sub-step (Delphi "round"). This is completed with relatively little training of the participants in Bayesian modelling.

In this paper I report the results of a pilot study where small groups of participants are given a problem of forecasting changes in levels of crime in a city resulting from a proposed policy change (e.g. in policing, zoning, drug legislation). Once a model is constructed, new forecasts are made for other policy changes that were not previously considered.

The models produced in the study are considered in terms of the performance (e.g. in terms of accuracy and calibration assessed against empirical values for crime risk, and coherence) of the forecasts they can make, the quality of the reasoning in the reports generated from the models, and system usability relative to other approaches.

Bayesian Ensembles: When and Why to Extremize Binary-Event Forecasts

Kenneth Lichtendahl
University of Virginia
United States

Yael Grushka-Cockayne
University of Virginia
United States

Victor Richmond Jose
Georgetown University
United States

Robert Winkler
Duke University
United States

We examine challenges related to aggregating forecasts of binary events – e.g., whether a borrower will default on a loan or not or whether a customer will click a link or not. Because linear combinations of probability forecasts are known to be underconfident, we introduce a class of aggregation rules, called Bayesian ensembles, that are non-linear in the experts' probabilities. These ensembles are generalized additive models of experts'

probabilities with three key properties. They are coherent, i.e., consistent with the Bayesian view. They can aggregate calibrated or miscalibrated forecasts. And they tend to be more extreme, and therefore more confident, than the commonly used linear opinion pool. We use this framework to aggregate several forecasts of loan defaults from several leading machine-learning algorithms in the Fannie Mae single-family loan performance data. Our Bayesian ensemble is easy to implement and offers an improvement out-of-sample over the linear opinion pool and over any one of the individual machine learning algorithms considered, two of which -- the random forest and extreme gradient boosted trees -- are already ensembles themselves. We also show empirically that our results are robust in that the Bayesian Ensemble performs well in other binary event prediction challenges (sports and car sales classification data sets).

Bayesian Rank Selection in Vector Autoregression

Bin Jiang
Department of Econometrics and business Statistics, Monash University
Australia

Anastasios Panagiotelis
Department of Econometrics and business Statistics, Monash University
Australia

George Athanasopoulos
Department of Econometrics and business Statistics, Monash University
Australia

Rob J Hyndman
Department of Econometrics and business Statistics, Monash University
Australia

Farshid Vahid
Department of Econometrics and business Statistics, Monash University
Australia

Estimating the rank of the coefficient matrix is a major challenge in multivariate regression, including vector autoregression (VAR). In this paper, we develop a novel fully Bayesian approach that allows for rank estimation. The key to our approach is reparameterizing the coefficient matrix using its singular value decomposition and conducting Bayesian inference on the decomposed parameters. By implementing a stochastic search variable selection on the singular values of the coefficient matrix, the ultimate selected rank can be identified as the number of nonzero singular values. Our approach is appropriate for small multivariate regressions as well as for higher dimensional models with up to about 40 predictors. In macroeconomic forecasting using VARs, the advantages of shrinkage through proper Bayesian priors are well documented. Consequently, the shrinkage approach proposed here that selects or averages over low rank coefficient matrices is evaluated in a forecasting environment. We show in both simulations and empirical studies that our Bayesian approach provides forecasts that are highly competitive against two of the most promising benchmark methods, dynamic factor models and factor augmented VARs.

Bayesian semi-parametric realized-CARE models for tail risk forecasting incorporating realized measures

Richard Gerlach
The University of Sydney
Australia

Chao Wang
The University of Sydney
Australia

A new model framework called Realized Conditional Autoregressive Expectile (Realized-CARE) is proposed, through incorporating a measurement equation into the conventional CARE model, in a manner analogous to the Realized-GARCH model. Competing realized measures (e.g. Realized Variance and Realized Range) are employed as the dependent variable in the measurement equation and to drive expectile dynamics. The measurement equation here models the contemporaneous dependence between the realized measure and the latent conditional expectile. We also propose employing the quantile loss function as the target criterion, instead of the conventional violation rate, during the expectile level grid search. For the proposed model, the usual search procedure and asymmetric least squares (ALS) optimization to estimate the expectile level and CARE parameters proves challenging and often fails to converge. We incorporate a fast random walk Metropolis stochastic search method, combined with a more targeted grid search procedure, to allow reasonably fast and improved accuracy in estimation of this level and the associated model parameters. Given the convergence issue, Bayesian adaptive Markov Chain Monte Carlo methods are proposed for estimation, whilst their properties are assessed and compared with ALS via a simulation study. In a real forecasting study applied to 7 market indices and 2 individual asset returns, compared to the original CARE, the parametric GARCH and Realized-GARCH models, one-day-ahead Value-at-Risk and Expected Shortfall forecasting results favor the proposed Realized-CARE model, especially when incorporating the Realized Range and the sub-sampled Realized Range as the realized measure in the model.

Biased forecasts: The impact of service level and product shelf life

Behnam Fahimnia
The University of Sydney
Australia

Tarkan Tan
Eindhoven University of Technology
Netherlands

Meysam Arvan
The University of Sydney
Australia

Enno Siemsen
University of Wisconsin-Madison
United States

Most supply chain decisions such as sourcing, inventory, production planning and logistics decisions are demand-driven. Sales forecasts produced by the forecasting team are assumed to be unbiased and simply display the most likely value for product demand based on historical data and sales patterns. These forecasts are then used as input into supply chain decisions in other departments where a number of other factors such as service level, capacity constraints, supplier reliability, and safety inventory will be taken into consideration. Forecasts are thus not plans or budgets, they are made to simply support the managerial decisions in various departments.

This study aims to understand the extent to which this assumption holds in reality. We want to know if forecasts – in a real-world context – are treated as "the most likely values" or more like "decisions". We refer to the latter as "biased forecasts". In particular we are interested in the impact of retailer service level requirement and product shelf life in generating the point forecasts for highly-perishable products. It is a common practice in supply chain planning decision-making to consider the trade-off between service level requirements set by retailers and the amount of wastes due to oversupply of perishable products (overage versus underage costs, also referred to as 'asymmetric loss function'). Forecasts are assumed to disregard these factors when generating the point forecasts to avoid double counting.

With this introduction in mind, we design an experiment to explore answers to the following questions: (1) to what extent service level requirement is taken into account when developing the base forecasts?, (2) to what extent waste minimisation is taken into account when developing the base forecasts for perishable products?, and (3) is there a clear relationship between a forecaster's sustainability beliefs and under-forecasting?

Call centre forecasting using multiple temporal aggregation

Devon Barrow
Coventry University
United Kingdom

Nikolaos Kourentzes
Lancaster University
United Kingdom

Bahman Rostami-Tabar
Cardiff University
United Kingdom

With thousands of call centres worldwide employing millions and serving billions of customers as a first point of contact, accurate scheduling and capacity planning of resources is important. Forecasts are required as inputs for such scheduling and planning in the short medium and long-term. Current approaches involve forecasting weekly demand and subsequent disaggregation into half-hourly, hourly and daily time buckets as forecast are required to support multiple decisions and plans. Once the weekly call volume forecasts are prepared, accounting for any seasonal variations, they are broken down into high frequencies using appropriate proportions that mainly capture the intra-week and intra-day seasonality. Although this ensures reconciled forecasts across all levels, and therefore aligned decision making, it is potentially not optimal in terms of forecasting. On the other hand, producing forecasts at the highest available frequency, and aggregating to lower frequencies, may also not be ideal as very long lead-time forecasts may be required. A third option, which is more appropriate from a forecasting standpoint, is to produce forecasts at different levels using appropriate models for each. Although this has the potential to generate good forecasts, in terms of decision making the forecasts are not aligned, which may cause organisational problems. Recently, Kourentzes et al. (2014) proposed the Multiple Aggregation Prediction Algorithm (MAPA), where forecasting with multiple temporal aggregation (MTA) levels allows both accurate and reconciled forecasts. The main idea of MTA is to model a series at multiple aggregation levels separately, taking advantage of the information that is highlighted at each level, and subsequently combine the forecasts by using the implied temporal hierarchical structure. Athanasopoulos et al. (2017) proposed a more general MTA framework than MAPA, defining appropriate temporal hierarchies and reconciliation mechanisms, and thus providing a MTA forecasting framework that is very flexible and model independent, while retaining all the benefits of MAPA. Given the high frequency, multi-temporal nature of the forecast requirements and the subsequent planning associated with call centre arrival forecasting, MTA becomes a natural, but yet unexplored candidate for call centre forecasting. This work evaluates whether there are any benefits from temporal aggregation both at the level of decision making as well as at the level of aggregation in terms of forecast

accuracy and operational efficiency. In doing so, various methods of disaggregation are considered when the decision level and the forecasting level differ, including methods which results in reconciled and unreconciled forecasts. The findings of this study will contribute to call centre management practice by proposing best approaches for forecasting call centre data at the various decision levels taking into account accuracy and operational efficiency, but will also contribute to research on the use of temporal hierarchies in the area of high frequency time series data.

Can forecasters interpret salient information when making judgmental adjustments?

Anna Sroginis
Lancaster University
United Kingdom

Robert Fildes
Lancaster University
United Kingdom

Nikolaos Kourentzes
Lancaster University
United Kingdom

While many academic papers propose different statistical methods for forecasting, many of which have been implemented for business use, practitioners still rely on judgmental approaches for product forecasting. Since pure judgemental forecasts are highly biased and cannot scale up to the requirements of modern supply chains, judgmental adjustments to statistical forecasts are common practice and seem to be more beneficial when extra soft information, such as marketing and promotional input, is provided. However, we do not know how people interpret this domain knowledge and how they evaluate whether this information is relevant to making a demand forecasting. The answer to this question could help to identify better ways of presenting this contextual information in Forecasting Support Systems (FSS) in order to guide forecasters in their tasks and to increase the accuracy of their forecasts. We propose an experimental setting that will help to illuminate how experts perceive the usefulness of information and implement it in their adjustments, when such soft information is provided. Crucially, we are interested when this information is inherently useful (or not). In the proposed computer-based experiment, forecasters are provided with historical sales, statistical forecasts and some 'soft' information. In particular, this information includes reasons or explanations of success (or failure) of past promotions and some statements for the forecasted period. The objective is to identify how people react to these statements and how they value their usefulness. We present preliminary results of the experiment identifying the circumstances when such diagnostic information improves forecast accuracy.

Changing the Paradigm for Business Forecasting

Michael Gilliland
SAS
United States

For at least the last 60 years, since Robert G. Brown's monograph on Exponential Smoothing for Predicting Demand, business forecasting has been dominated by an "offensive" paradigm. This paradigm is characterized by a focus on models, methods, and organizational processes designed to extract every last fraction of a percent of accuracy in our forecasts. More is thought to be better – more data, bigger computers, more complex models, and more elaborate collaborative processes. While this paradigm has taken us a long way, and today's software can do automatic forecasting on an unfathomable scale, the offensive paradigm is reaching its limits. We seem to

be stuck in making progress toward more accurate forecasts.

To meaningfully improve organizational forecasting performance, the solution may be a shift in mindset to a "defensive" paradigm. The defensive paradigm focuses on identifying the things we do that fail to improve the forecast – the so-called worst practices. It is now recognized that a significant percentage of real-life business forecasts have accuracy below that of even a naïve no-change model. So worst practices are a serious, and unfortunately common business problem.

This presentation outlines the playbook for defensive business forecasting. It shows how simple methods like Forecast Value Added (FVA) analysis can identify the bad practices that plague our forecasting processes. Eliminating harmful practices can bring improvements in forecasting performance that dwarf the accuracy gains you get from the elaborate models and methods of an offensive approach.

Characterizing retail demand with promotional effects for model selection

Patrícia Ramos
INESC TEC and School of Accounting and Administration, Polytechnic Institute of Porto
Portugal

José Manuel Oliveira
INESC TEC and Faculty of Economics, University of Porto
Portugal

Robert Fildes
Lancaster Centre for Forecasting, Lancaster University
Portugal

Shaohui Ma
School of Economics and Management, Jiangsu University of Science and Technology
China

Dramatic changes in the retail environment have led to increasing product variety with decreasing life cycles, making sales at the SKU (Stock Keeping Unit) level in a particular store difficult to forecast, as times series for these items tend to be short and often intermittent. Moreover, retailers are increasing their marketing activities, such as price reductions and promotions. Products are typically on promotion for a limited period of time during which demand is usually substantially higher leading to potential stock-outs due to inaccurate forecasts. 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 intense, that practice is not efficient and an automated and reliable multivariate forecasting system is required. This study focuses on taking advantage of the characterization of demand for effective model selection and improved forecast accuracy. The investigation is supported by a large dataset of daily data from a Portuguese retailer that operates in food distribution and consumer products manufacturing. A range of univariate models (ETS, ARIMA, TBATS, STL and Naïve), multivariate regression that includes price and promotional information with variable selection using several techniques, and methods for intermittent demand forecasting, are compared including the effects of different seasonalities on a weekly and daily basis. To evaluate their performance, a robust experimental design was implemented. This design is based on a rolling forecasting origin scheme that considers several forecast horizons and calculates multiple error measures suitable for intermittent time series. Our finds provide important insights for the relationship between the characteristics of the demand and the most adequate models.

Choice model based on bounded rational utility

SangKyun Jo
KAIST
South Korea

DukBin Jun
KAIST
South Korea

Rational choice theory is based on an economic principle which states that individuals are aware of their preferences and select the best choice in given situation. However, due to persistence of bounded rationality and information overload, search cost occurs in order to obtain alternative information. Previously, several researchers have analytically modeled bounded rational situations and applied search cost into consumer utility. Applying search cost into the model gives implication that there exists a finite optimal number of alternatives that maximizes the probability of choice. In this paper, we empirically investigate the results due to bounded rational consumer utility applying choice model.

Cloud-Based Time Series Analysis and Forecasting

Michael Leonard
SAS
United States

Many organizations need to process large numbers of time series for analysis, decomposition, forecasting, monitoring, and data mining. The TSMODEL procedure provides a resilient, distributed, optimized generic time series analysis scripting environment for cloud computing. It provides automatic forecast model generation, automatic variable and event selection, and automatic model selection. It provides advanced support for time series analysis (time domain and frequency domain), time series decomposition, time series modeling, signal analysis and anomaly detection (for IoT), and temporal data mining. This paper describes the scripting language that supports cloud based time series analysis. Examples that use SAS® Visual Forecasting software demonstrate this technique.

Coherent mortality forecasting by the weighted multilevel functional principal component approach

Ruhao Wu
University of Leicester
United Kingdom

Bo Wang
University of Leicester
United Kingdom

In human mortality modelling, if a population consists of several subpopulations it can be desirable to model their mortality rates simultaneously while taking into account the heterogeneity among them. The mortality forecasting methods tend to result in divergent forecasts for subpopulations when independence is assumed. However, under closely related social, economic and biological backgrounds, mortality patterns of these subpopulations are expected to be non-divergent in the future. In this article, we propose a new framework for

coherent modelling and forecasting of mortality rates for multiple subpopulations, in the sense of nondivergence of life expectancies among the subpopulations. The mortality of subpopulations is treated as multilevel functional data and then a weighted multilevel functional principal component approach is proposed to model and forecast the mortality rates. The proposed model is applied to sex-specific data for nine developed countries, and the results show that, in terms of overall forecasting accuracy, the model outperforms the independent model and the Product-Ratio model.

Coherent probabilistic forecasts for hierarchical time series

Souhaib Ben Taieb
Monash University
Australia

James Taylor
University of Oxford
United Kingdom

Rob Hyndman
Monash University
Australia

Many applications require forecasts for a hierarchy comprising a set of time series along with aggregates of subsets of these series. Although forecasts can be produced independently for each series in the hierarchy, typically this does not lead to coherent forecasts — the property that forecasts add up appropriately across the hierarchy. State-of-the-art hierarchical forecasting methods usually reconcile these independently generated forecasts to satisfy the aggregation constraints. A fundamental limitation of prior research is that it has looked only at the problem of forecasting the mean of each time series. We consider the situation where probabilistic forecasts are needed for each series in the hierarchy. We define forecast coherency in this setting, and propose an algorithm to compute predictive distributions for each series in the hierarchy. Our algorithm has the advantage of synthesizing information from different levels in the hierarchy through a sparse forecast combination and a probabilistic hierarchical aggregation. We evaluate the accuracy of our forecasting algorithm on both simulated data and large-scale electricity smart meter data. The results show consistent performance gains compared to state-of-the art methods.

Collecting Information from Multiple Forecasters: Inefficiency of Measures of Central Tendency

Ville Satopaa
INSEAD
France

Even though aggregating multiple predictions typically outperforms the average individual prediction, there is no consensus about the right way to do this. Optimally, an aggregator would use all the information in the predictions. This is possible but requires a model for how the predictions differ from each other and the outcome. To this end, past literature has discussed two philosophically distinct sources: measurement error ascribes differences to forecasters' errors in measuring some ``ideal'' prediction, whereas information diversity supposes that the predictions differ because the forecasters use different information.

Largely because the individual contributions of each source cannot be separated from the predictions, existing aggregators often assume differences to stem purely from one source or the other. For instance, measures of central tendency, such as the average, median, and many others, represent measurement error and hence assume that the forecasters use the exact same information but make mistakes in interpreting it. This can be a reasonable model in many classical wisdom-of-the-crowd examples, such as guessing the weight of the ox at the country fare or the number of jelly beans in a jar, where the forecasters' information sets are almost identical.

A crowd predicting outcomes in politics, national security, strategy, and many other fields, however, tend to draw on very different sets of information. In such high impact cases the main driver of differences is likely to be information diversity – not measurement error. Unfortunately, aggregation continues to be done via some measure of central tendency, suggesting a potential discrepancy between the choice of the aggregator and the main source of forecast heterogeneity.

Motivated by this, our work examines whether measures of central tendency can be optimal even when the differences stem from information diversity. Our analysis does not rely on any specific probability distribution or make any assumptions about the structure of the forecasters' information. The main results show, among others, that all weighted arithmetic, harmonic, geometric, quadratic, and cubic means distort the forecasters' information and are never optimal. In fact, no aggregator that remains strictly within the convex hull of the forecasts is optimal as long as the individuals rely on finite information. Given that the physical universe only contains a finite amount of information, for all practical purposes, measures of central tendency are sub-optimal under information diversity. To provide resolution, we briefly explain how optimal information-diversity-based aggregators can be constructed in practice.

Company Performance Predictive Analysis - Application to Leveraged Buy-Out Participation

Maxime Burlot
ECE Paris
France

Maxime Vindolet
ECE Paris
France

Adrien Lalanne
ECE Paris
France

Jérémy Guilhem
ECE Paris
France

Paul Bossy
ECE Paris
France

After 2008-2009 economic turmoil, the Leveraged Buy-Out (LBO) market growth has been since restored and stabilized. In 2015, debt ratios of targeted companies have reached the 2006 historical highest level representing an average of 7.5 times the EBITDA. During last two decades, the LBO market became more structured and it is therefore of paramount importance for LBO stakeholders - in particular private equity funds - to accurately assess the profitability of their operations.

To that extent, present article aims at providing a (statistically based) approach to forecast financial profile of companies over LBO. In technical terms, it is intended to determine a transition matrix forecasting short-term useful financial data of companies under LBO. Based on a sample of 50 French LBO companies operating in Wholesale industry of which financial data have been collected, the following two-step statistical method approach is applied:

1) Principal Component Analysis (PCA) is used to allow dimensionality reduction (maintaining data variance): 8 key financial variables help identify different groups of companies.

2) Multiple linear regression models (Method of Least Squares) are applied to each group in order to predict the 8 key financial variables post LBO that help assess LBO operations profitability.

From new data analysis based on financial statement per sector of activity, financial standing of a targeted company can be forecast with this method. The result will help private equity funds to estimate future profitability of their LBO portfolio.

Comparing alternative approaches to the production of robust inflation forecasts using the Phillips Curve: an application to Europe

James Mitchell
Warwick Business School
United Kingdom

Gian Luigi Mazzi
Independent Expert
Luxembourg

This paper considers how to accommodate instabilities in the Phillips Curve relationship between inflation and the output gap. The paper compares two methodological approaches, namely time-varying parameter models (e.g. as used by Cogley and Sargent; and Stock and Watson) and quantile based regression methods. It is considered how the quantile approach, similarly to time-varying parameter models, can offer robustness to temporal instabilities and to changes in volatility. In an empirical application this paper establishes that in the Euro Area as whole, and in France, Germany, Italy and Spain - similarly to in the UK and US - it is important to accommodate temporal instabilities when density forecasting inflation via the Phillips Curve. But extending previous work, we show that these instabilities can be captured, in general across countries and forecast horizons, better via quantile autoregressive distributed lag models than via what has become the industry-standard, time-varying Bayesian VAR models with stochastic volatility. Importantly, this is so across alternative measures of density forecast evaluation, including the log score and the cumulative ranked probability score.

Comparing Forecasting Methods for Time Series with Calendar Time Effects with Application to Google Search Traffic Data

Amy Richardson
Google
United States

Weijie Shen
Google
United States

Funan Shi
Google
United States

Daily or hourly time series driven primarily by human behavior typically exhibit a number of calendar time effects. There often are effects for multiple levels of seasonality due to daily, weekly or monthly schedules and/or other seasonal changes throughout the year. Some of these seasonal effects may be irregular, as is the case with effects due to holidays from non-Gregorian calendars such as the Easter holiday, Chinese New Year, or Eid al-Fitr. Estimation of such effects can be further complicated by large amounts of noise, change points in the level or trend of the series, or in series with a short history. Prediction accuracy for such time series may have important implications for decisions regarding resource allocation of all kinds including machine hardware, engineering efforts, marketing and sales efforts and more. In this case study, we compare four time series forecasting methods on several hundred Google daily search volume time series both globally and in several countries exhibiting varying levels of noise and series length. Measures include forecast accuracy, stability, scaleability and prediction interval coverage rates. The four methods are (1) TBATS model (Exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components, using the forecast R package), (2) a bayesian structural time series model incorporating effects for major holidays (using the bsts R package), (3) the prophet model with effects for major holidays (using the prophet R package) and (4) an ensemble modelling approach employing separate decompositions for each seasonal effect.

Computationally Efficient Prediction Intervals for Machine Learning Forecasts

Jocelyn Barker
Microsoft
United States

Gagan Bansal
Microsoft
United States

Machine learning forecasting methods have shown improved forecast accuracy over traditional time series methods (e.g. ARIMA, ETS, STL) in some forecasting problems. Unlike many traditional methods, most machine learning methods do not have a theoretical basis for creating predictions intervals, which are highly valuable in many applications. In this paper, we propose a method for computing prediction intervals using out-of-sample errors. The method builds on a common cross-validation method for time series, rolling forecasting origin validation. Using forecasts created during cross-validation, back testing can be performed, complete with hyper-parameter selection, with minimal computational cost (~72ms). The out-of-sample errors from the back testing are then fit to a Gaussian distribution with a mean of zero to create a prediction interval for the forecasts. Since the framework is model agnostic, it can be used on any machine learning forecasting method. To test our method, forecasts and prediction intervals were computed using four diverse machine learning methods (Random Forest, Elastic Net, K-Nearest Neighbors, and Support Vector Machines) on two datasets; the M3 dataset, which is a commonly used benchmark and Microsoft's revenue data, which tests our method in a real-world application. Prediction intervals were evaluated by computing the quantile of the predicted error distribution for the actual value for each of the forecasts. For comparison, prediction intervals were created using the same Gaussian assumption, but using in-sample errors rather than out-of-sample errors. If the prediction intervals accurately reflect the actual error distribution, we would expect the distribution of error quantiles to be uniform on the interval of 0:1. The distribution of error quantiles were visualized by plotting a cumulative distribution

function of forecast error quantiles similar in concept to a Q-Q plot. The ideal distribution would be a line through the origin with a slope of one. The cumulative error distribution using out-of-sample errors showed high correspondence to the expected linear distribution particularly when compared to those derived from in-sample errors (R2 in-sample vs out-of-sample in M3 data; Random Forest: -0.05 vs 0.89; Elastic Net: 0.77 vs 0.97; K-Nearest Neighbors: -0.89 vs 0.88; Support Vector Machines: 0.63 vs 0.99), indicating this method's prediction intervals characterize the error distribution well. Our method's ability to quickly and accurately compute prediction intervals for machine learning methods has caused it to be a valuable tool in Microsoft's financial planning process.

Conditionally Optimal Weights and Forward-Looking Approaches to Combining Forecasts

Christopher G. Gibbs
UNSW
Australia

Andrey L. Vasnev
University of Sydney
Australia

In applied forecasting there is a trade-off between in-sample fit and out-of-sample forecast accuracy. Parsimonious model specifications usually outperform richer model specifications. As a consequence, there is often predictable information in forecast errors that is hard to exploit. However, we show how this predictable information can be exploited in forecast combinations. In this case, optimal combination weights should minimize conditional mean squared error, or a conditional loss function, rather than the unconditional variance as in the commonly used framework of Bates and Granger (1969). We prove that our conditionally optimal weights lead to better forecast performance. The conditionally optimal weights support other forward-looking approaches to combining forecasts, where forecast weights depend on expected model performance. We show that forward-looking approaches can robustly outperform the benchmark random walk and many commonly used forecast combination strategies including equal weights in real-time out-of-sample forecasting exercises of inflation.

Improving forecasting performance using covariate-dependent copula models

Yanfei Kang
Beihang University
China

Feng Li
Central University of Finance and Economics
China

Copulas provide an attractive approach for constructing multivariate densities with flexible marginal distributions and different forms of dependence. Of particular importance in many areas is the possibility of explicitly forecasting tail-dependence. Most of the available approaches are only able to estimate tail-dependence and correlations via nuisance parameters, but can neither be used for interpretation nor for forecasting. We propose a general Bayesian approach for dynamically modeling tail-dependence and correlations as explicit functions of covariates. The proposed method allows for variable selection among covariates in marginal models as well as in copula parameters. Posterior inference is carried out using a novel and efficient MCMC simulation

method. Our approach is applied to S&P 100 and S&P 600 stock indices prediction and forecasting performances of different modeling strategies are compared based on log predictive likelihood.

Customer Flow Forecasts on Koupei.com

Shaohui Ma
Jiangsu University of Science and Technology
China

Forecasting customer flow is one of the key factors for a successful retailer. Basing on the forecasts, retailer can optimize their operations, reduce cost and improve user experience. Traditionally, the customer flow forecasting is based on the seasonal sales data collected by the POS scanners. Today, with the prevalent mobile location-based service, consumers can easily obtain shopping information on opportunities nearby, get recommendation based on the reputation and promotional information of stores, and pay for their purchases with their phone. And in return, those browsing and paying histories are recorded by mobile app service providers. Therefore, we now have far more comprehensive data to generate customer flow forecasts. Koupei.com, one of subsidiary of Alibaba, provides an online platform for consumers sharing their experiences in the store, and for retailers to release their promotional information, also collect spending information if the customers making payment via Alipay (similar with PayPal) in the shop, and therefore it accumulates huge amount of user data on the platform every day. With these data, Koupei.com aims to provide merchants with a customized back-end business intelligence services, including offering a customer flow forecasting services for every merchant on the platform. In this research, we investigate the value of different sources of data on forecasting customer flow, which here is limited to the customers using Alipay in the store, including the customers' browsing history, seasonality, holidays, weather, and air quality: also the benefit of data pooling by stores in different category but in nearby location.

Data as a service: Providing new datasets to the forecasting community for time series analysis

Evangelos Spiliotis
National Technical University of Athens
Greece

Argyris Patikos
National Technical University of Athens
Greece

Vassilios Assimakopoulos
National Technical University of Athens
Greece

When analyzing or forecasting time series, it is common in practice to assess the effectiveness of the examined methods by testing their performance on datasets regularly studied in the literature. Yet, it has been repeatedly reported that these data collections are not representative as they include limited number of time series of specific domains and characteristics. Moreover, the time series are not uniformly distributed across the resulting subsets. Consequently, methods which perform well on the larger subsets will be favored over others. On the other hand, if a researcher wishes to evaluate a method using a data sample of specific discipline areas, frequencies and characteristics, it may be difficult for him/her to access a reliable dataset which fits his/her needs. To mitigate these problems in this paper we introduce a web-service which enables researchers form datasets of interest by collecting time series from a multidimensional pool. For cases when the data found are inadequate, another service which generates time series of specified characteristics can be exploited. By

comparing the data collection introduced over other well-studied datasets, we also provide some insights regarding the benefits arising for setting it as a new benchmark for method evaluation.

Data characteristic analysis and model selection for container throughput forecasting within a decomposition-ensemble methodology

Gang Xie
Academy of Mathematics and Systems Science
China

With the current development of economic globalization, ports have become more and more important in the operation of international trade activities. A port is a logistic and industrial node which accommodates seagoing vessels. A port is characterized by a functional and spatial clustering of cargo transport, storage and transformation processes, all of which are linked to global supply chains. The handling of maritime cargo at specialized terminals remains a core function of ports. Particularly, as many manufacturing industries have experienced rapid growth, particularly in emerging economies, many coastal cities in those economic centers have invested heavily in ports. The prediction of container throughput at ports performs the basic function of helping port managers make strategic decisions, as well as with developing the port in terms of scale, general layout and district division. Accurate predictions of container throughput mean that investments in port capacity and other transportation infrastructure can be made consistent with the needs generated by that traffic. If these throughput predictions are not accurate enough, policy bias will occur, which could cause huge financial losses. Therefore, developing an effective container throughput forecasting model has become a crucial and challenging task. In this study, a novel decomposition-ensemble methodology is proposed for container throughput forecasting. Firstly, the sample data of container throughput at ports are decomposed into several components. Secondly, the time series of the various components are thoroughly investigated to accurately capture the data characteristics. Then, an individual forecasting model is selected for each component, based on the data characteristic analysis (DCA) and prediction horizon. Finally, the forecasting results are combined as an aggregated output. An empirical analysis is implemented for illustration and verification purposes. Our results suggest that the proposed hybrid models can achieve better performance than other methods.

Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate models

Florian Ziel
University Duisburg-Essen
Germany

Rafal Weron
Wroclaw University of Science and Technology
Poland

Conducting an extensive empirical study on short-term electricity price forecasting (EPF), involving state-of-the-art parsimonious expert models as benchmarks, datasets from 12 power markets and 32 multi-parameter regression models estimated via the lasso, we show that using the latter shrinkage approach can bring statistically significant accuracy gains compared to commonly-used EPF models. We also address the long-standing question

on the optimal model structure for EPF. We provide evidence that despite a minor edge in predictive performance overall, the multivariate modeling approach does not uniformly outperform univariate models across all datasets, seasons of the year or hours of the day, and at times is outperformed by the latter. This may be an indication that combining advanced structures or the corresponding forecasts from both modeling classes may bring a further improvement in forecasting accuracy. Finally, we also analyze variable selection for the best performing multivariate and univariate high-dimensional lasso-type models, thus provide guidelines to structuring better performing forecasting model designs.

DCaf: Detecting Change-points using Average Forecast Errors

Funan Shi
Google
United States

Amy Richardson
Google
United States

Weijie Shen
Google
United States

A time series step change is a shift in mean after a change point T, often as a result structural changes such as product launches or shifts in user behavior. Most existing change-point detection algorithms assume the underlying time series to be piecewise stationary, meaning if a single change point occurs at some time T, then the time series from time 1 to T should be stationary, and so should the series from T+1 and on. However, time series with trend, as so many real world data are, violate this fundamental assumption. Detrending time series afflicted with steps changes, however, is not a viable option because it would end up adding bias. In this work we present an intuitive change point detection method called Detecting Change-points using Average Forecast errors (DCaf) based on m-step forecasting errors. The method has three notable features: (1) The method is nonparametric and does not assume any particular generating model for the underlying time series, thus sidestepping the aforementioned stationarity requirement. (2) The method outlines a precise window (start to end) of a step change. Knowing the precise window of impact is crucial for adjusting the step changes. (3) The method is extremely fast, and suitable to be deployed to preprocess thousands of time series. We will show numerically that the method achieves excellent precision and recall using both simulated and real world data. Its speed and precision makes DCaf a suitable add-on for large scale forecasting tasks.

Deep Neural Network Regression as a Component of a Forecast Ensemble

Richard Povinelli
Marquette University
United States

Gregory Merkel
Marquette University
United States

Ronald Brown
Marquette University
United States

Marquette University's GasDay Project specializes in short-term load forecasting of natural gas demand. Traditionally, this forecasting is done using artificial neural networks and linear regression. These two forecasts are used as components of an ensemble that adjusts its weights based on recent residuals of each component. For example, if the linear regression forecaster is consistently outperforming the artificial neural network, then the weight placed on the linear regression forecast is greater. Despite the quality of this ensemble forecaster, literature would suggest that including addition quality forecasters would improve this ensemble. In many fields, deep neural networks are being used to replace standard backpropagation neural networks and other machine learners in solving time series and classification problems and often have found good solutions. In a related paper, the authors have shown that deep neural networks (DNN) often outperform either the linear regression or standard artificial neural network components on the short-term energy load forecasting problem. This paper examines the viability of using DNNs as component models in the GasDay ensemble. The ensemble of interest is evaluated using weighted MAPE on 90 natural gas data sets and compared to the current GasDay ensemble as well as ensembles with other candidate component models. The DNN-enhanced ensemble performs better on this metric than the other ensembles evaluated. In addition, the weights given by the ensemble are evaluated to see how often each of the components receives the greatest weighting. Although the DNN component usually has the greatest weight, it rarely has a majority weight. This is to be expected based on the performance of each of the component models.

Deep Neural Network Regression for Short-Term Load Forecasting of Natural Gas

Gregory Merkel
Marquette University
United States

Richard Povinelli
Marquette University
United States

Ronald Brown
Marquette University
United States

Short-term load forecasting is important for the day-to-day operation of natural gas utilities. Traditionally, short-term load forecasting of natural gas is done using linear regression, autoregressive integrated moving average models, and artificial neural networks. Many purchasing and operating decisions are made using these forecasts, and there can be high cost to both natural gas utilities and their customers if the short-term load forecast is inaccurate, so energy analysts continue to explore new ways to make better forecasts. Recently, deep neural networks (DNN) have emerged as a powerful tool in machine learning problems. DNNs have been shown to greatly outperform traditional methods in many applications, and they have completely revolutionized some fields. Given their success in other machine learning problems, we evaluate DNNs in energy forecasting. This paper discusses good practices including architecture, input features, and transfer functions in using DNNs for short-term load forecasting of natural gas and forecasting in general. Further, it proposes a network architecture with five hidden layers and evaluates the performance of this architecture on 90 natural gas data sets. The performance of the model is evaluated using weighted MAPE and compared against several traditional forecast strategies, including artificial neural networks and linear regression short-term load forecasting strategies. The DNN forecaster offers an average 10% decrease in weighted MAPE compared to standard ANNs and a nearly

30% decrease in weighted MAPE compared to linear regression, suggesting that DNNs are a promising option for the short-term energy load forecasting problem.

Demand forecasting by temporal aggregation: The overlapping vs. non-overlapping case

Bahman Rostami-Tabar
Cardiff University
United Kingdom

Zied Babai
Kedge Business School
France

Aris Syntetos
Cardiff University
United Kingdom

Demand forecasting performance is subject to the uncertainty underlying the time series an organisation is dealing with. There are many approaches that may be used to reduce uncertainty and thus improve forecasting performance. One intuitively appealing such approach is temporal aggregation. There are two different types of temporal aggregation: non-overlapping and overlapping. In the former case the time series are divided into consecutive non-overlapping buckets of time where the length of the time bucket equals the aggregation level. As a consequence the numbers of time periods in the aggregate demand time series is less than that corresponding to the original demand time series. The overlapping case is somewhat different in that it resembles the mechanism of a moving window technique where the window's size equals the aggregation level. It is observed that the information loss in the latter case is negligible as compared to the former case. This is an important observation in terms of data availability with considerable implications for the cases where little history of data is available. Unlike the fast growing area of non-overlapping temporal aggregation in a supply chain context, the effects of overlapping temporal aggregation have been rather neglected in the (supply chain forecasting) literature. To the best of our knowledge, the effects of overlapping temporal aggregation of ARMA processes, the identification and properties of the resulting (aggregate) series, the relationship between original and aggregate demand parameters, and related forecasting performance have not been studied yet. The aim of this work is to address the issues discussed above. To do so, we focus on an ARMA(1,1) demand processes in conjunction with the Simple Exponential Smoothing (SES) forecasting method. Our contribution is three-fold: i) we first determine the structure of the aggregate demand processes and identify their properties; ii) we evaluate the impact of overlapping temporal aggregation on demand forecasting performance; and iii) we compare the effectiveness of overlapping against non-overlapping temporal aggregation by considering the effect of the length of series. The results enable insights to be gained into the impact of the control and process parameters on the comparative performance of the two approaches.

DEMAND FORECASTING IN MARKETING: METHODS, TYPES OF DATA, AND FUTURE RESEARCH

Carla Freitas Silveira Netto
Universidade Federal do Rio Grande do Sul - UFRGS
Brazil

Vinicius Andrade Brei
Universidade Federal do Rio Grande do Sul - UFRGS
Brazil

Demand forecasts are fundamental to plan and deliver products and services. Despite such relevance, marketers have difficulty to choose which forecast method is the best for their organizations. One possible explanation for this baffling task is that the literature is not clear about demand forecasting methods' classifications, approaches, complexity, requirements, and efficiency. This theoretical paper tries to improve this scenario, reviewing the state of the art about demand forecasting in marketing. More specifically, we focus on: (1) the most frequently used models by academics and practitioners; (2) different classifications and approaches of those models, especially the ones based on statistics/mathematics and big-data; (3) challenges of big data/computer based forecasting; (4) types of data used; and (5) research gaps and suggestions of future research on demand forecasting in marketing. The most important research gaps are related to the types of models applied in marketing literature (structural). Besides simpler, easier to implement models, further research is necessary to develop forecasting techniques that incorporate dynamic effects, primary data, and nonparametric approaches more efficiently. The literature also evidences some gaps concerning the optimal use of types of data and data sources. Of foremost importance are data sets about durable goods, location/geographical data, big data, and the combination of different data sets. Based on the state of the art about forecasting methods, types and use of data, and research gaps found, we present suggestions for future research. New studies about demand forecasting in marketing should focus on durables goods and other types of less frequently purchased products. They could also combine different sources of data, such as free public data, firm property data, commercially available market research, big data, and primary data (e.g., surveys and experiments). Future studies should also analyze how to improve the use of location/geographical data, incorporating their dynamic perspective, without creating barriers to the method implementation. We also discuss how marketing and computer science should be integrated to fulfill those gaps.

Demand Forecasting of Autonomous Vehicles Based on Consumer Preference Analysis

Kyuho Maeng
Department of Industrial Engineering, Yonsei University
South Korea

Seung Ryong Jeon
Department of Industrial Engineering, Yonsei University
South Korea

Yongsang Cho
Department of Industrial Engineering, Yonsei University
South Korea

Autonomous vehicles(AVs), one of advanced vehicular technologies, are expected to greatly change the social environment as well as automobile industry. It is expected that the commercialization of AVs will not only enhance the convenience of the driver but also improve the transportation efficiency and reduce the traffic accidents. However, there is very limited research on consumer preferences for this emerging technology, and most of previous researches approach AVs from the perspective of engineering and information technology. The purposes of this study are evaluating what attributes are important for potential customers when purchasing AVs, forecasting demand of AVs, and planning of market penetration scenarios of AVs. For this, we conducted conjoint survey with 1,008 Koreans, and derived utility function of potential customers for AVs using discrete choice model. In the design of conjoint survey, we considered technological factors of AVs for autonomous driving such as road condition, weather condition, time condition, fuel efficiency, autonomous driving level, and price. According to the results, "road condition" is the most important factor (39.58%) followed by "price" (26.26%) , "time condition" (10.22%), "autonomous driving level" (9.15%), "fuel efficiency" (8.32%), and

"weather condition" (6.48%). Moreover, we found that the marginal willingness to pay (MWTP) of potential customers for all road condition, all time condition, and all weather condition are 3.42 million KRW, 2.32 million KRW, and 1.47 million KRW, respectively. We also analyzed the market potential of AVs by changing the level of each attribute. Based on the consumer utility function and demand forecast, we derived some strategic and policy implications to promote the diffusion of AVs. This study provides an important information which can be used in developing technology strategy, marketing strategy and related government policy.

DENSITY FORECAST COMPARISON USING BAYESIAN AND FREQUENTIST MULTIVARIATE MODELS WITH DISAGGREGATED MACROECONOMIC VARIABLES

Roberto Morales-Arsenal
CUNEF
Spain

Miguel-Ángel Gómez-Villegas
Universidad Complutense de Madrid
Spain

Multivariate macroeconomic models require a set of variables to increase the available information set in order to improve the predictive capacity of the model. In many cases, this set of random variables shows an interdependence or feedback that must be taken into account. Macroeconomic models of simultaneous equations, the so-called Vector Autoregressive (VAR) (Sims, 1980), are able to capture this property of interdependence. The problem with these models is the high number of parameters to estimate, a number that increases when we consider time-varying parameters (TVP-VARS) and also allows that the error covariance matrix to change over time. This gives rise to a problem of overparameterization. This has led to the use of Bayesian procedures that, using prior information, impose restrictions (shrinkage) on the parameters to estimate in the model in a logical and consistent way. In this work, we compare a point forecast and the predictive density function obtained from a frequentist VAR model versus a set of multivariate bayesian models: Bayesian VARs (with different priors), TVP Bayesian VARs (Canova, 1993), and TVP-FAVARs (Kose, Otrok and Whiteman, 2003) and Neural Networks-BVAR. We apply these models to the Harmonized Consumer Price Index (HCPI) in the Euro Area. We also compare the aggregated approach (forecasting directly the variable of interest) vs the disaggregated approach (indirectly using the components of the HCPI). The obtained results show the superiority of Bayesian procedures, especially in the disaggregated approach.

Detecting and Correcting Bias in the Fed's Forecasts

Neil Ericsson
Federal Reserve Board
United States

Forecast bias is central to forecast evaluation, especially because forecast bias is inherently systematic, and because ignored forecast bias may have substantive adverse consequences for policy. Drawing on recent econometric developments in saturation techniques, this paper sets out a comprehensive framework for detecting, analyzing, and correcting forecast bias, including bias that is time-invariant (Mincer-Zarnowitz),

anchoring (Tversky-Kahneman), informationally inefficient (Holden-Peel), or forecast non-encompassing (Chong-Hendry and Diebold-Mariano). Saturation techniques define a generic procedure for examining forecast properties; they explain why standard tests may well fail to detect bias; and they provide a mechanism for improving forecasts.

Using that saturation framework, this paper assesses the Fed staff's quarterly Greenbook and Tealbook forecasts, which are provided to the Federal Open Market Committee when meeting to decide U.S. monetary policy. Standard tests typically fail to detect bias in current-quarter and one-quarter-ahead forecasts of U.S. and foreign economic activity. However, additional tests detect economically large and highly significant bias that depends on the variable being forecast, the forecast horizon, and the phase of the business cycle. For example, for post-1980 U.S. output growth, the estimated forecast bias is approximately +0.5% per annum during expansions and -1.5% per annum during contractions. These biases are economically substantial, noting that output growth averages only about +2.8% per annum over the entire sample. Forecasts of Chinese real GDP growth--by contrast--have a relatively time-invariant bias of about +2% per annum. Bias correction is straightforwardly implemented. For all countries examined, there is little observed predictability beyond two quarters ahead.

Detecting Causality between Oil Prices and Tourist Arrivals: Evidence from US and European Nations

Xu Huang
De Montfort University
United Kingdom

Hossein Hassani
Institute for International Energy Studies
Iran

Emmanuel Silva
University of the Arts London
United Kingdom

Mansi Ghodsi
Institute for International Energy Studies
Iran

This paper aims at investigating the causal relationship between oil price and tourist arrivals, thus to further explain the impacts of oil price volatility on economical activities closely related to tourism. The analyses are conducted by both empirical and novel methods with comprehensive considerations of time domain, frequency domain and information theory domain perspectives. The causality tests adopted in this paper includes time domain, frequency domain and Convergent Cross Mapping (CCM), which has no assumption of linearity, no limitation of nonlinearity and complex fluctuations. Furthermore, both US and nine European countries are evaluated and compared. This paper contributes by not only employing advanced methods, but also via a comprehensive comparison across countries on a global scale which results in uncovering consistent and significant causal links from oil price on tourist arrivals.

DIY forecasting: judgment, models and judgmental model selection

Fotios Petropoulos
University of Bath
United Kingdom

Nikolaos Kourentzes
Lancaster University
United Kingdom

Konstantinos Nikolopoulos
Bangor University
United Kingdom

In this paper we explore how judgment can be used to improve model selection for forecasting.We benchmark the performance of judgmental model selection against the statistical one, based on information criteria. Apart from the simple model choice approach, we also examine the efficacy of a judgmental model build approach, where experts are asked to decide on the existence of the structural components (trend and seasonality) of the time series. The sample consists of almost 700 participants that contributed in a custom-designed laboratory experiment. The results suggest that humans perform model selection differently than statistics. When forecasting performance is assessed, individual judgmental model selection performs equally if not better to statistical model selection. Simple combination of the statistical and judgmental selections and judgmental aggregation significantly outperform both statistical and judgmental selection.

Do inflation surveys hold information that improves short-term inflation forecasts?

Raïsa Basselier
National Bank of Belgium
Belgium

David de Antonio Liedo
National Bank of Belgium
Belgium

Geert Langenus
National Bank of Belgium
Belgium

In this paper, we examine whether forecasts of Belgian HICP inflation can be improved by incorporating information from monthly surveys that are conducted by the National Bank of Belgium, but harmonised at the European level. Both in the consumer and in the producer survey, respondents are asked about their expectations on the evolution of prices, respectively twelve and three months ahead. While HICP inflation is, to some extent, triggered by rather volatile components such as oil and food prices that are difficult to predict, it should at least partly be driven by firms' decisions as well, in particular as regards the pass-through of rising or falling costs. Hence, it seems likely that the survey responses of business owners can provide an indication of how the overall price level will evolve in the near term.

In order to assess the information content of survey data on inflation, we elaborate specific short-term models of inflation in Belgium and analyse whether survey data can be a relevant predictor. We verify, in particular, whether survey data offer some additional explanatory power to HICP inflation, in addition to the traditional determinants. We then use the model to produce out-of-sample inflation forecasts and check how the accuracy of these forecasts compares to that of regular inflation projections (made in the context of Eurosystem projection exercises), which also include expert judgment. Finally, we extend the (first part of the) analysis to the euro area.

Do the lecturers predict accurately the student evaluation of their teaching? The mechanism behind and the implication of the possible mismatch between them.

Sheik Meeran
University of Bath
United Kingdom

Isa Sheik

United Kingdom

Ian Crawford
University of Bath
United Kingdom

Zhiqi Wang
University of Btah Spa
United Kingdom

This paper investigates the extent to which lecturers in a University are able to accurately estimate teaching evaluation scores given by their students. Study of the accuracy of such estimate is paramount, given that the student evaluation of teaching is taken as the 'de facto' standard by many universities to measure teaching performance (although objectionable to many). The inaccuracy in this estimate translates into inaccuracy in lecturers' self-evaluation of their performance, which will have an impact on their self-development plan, their perception of training needs, career progression decisions and also their job-satisfaction and motivation. A number of unique results are found in this study. We conduct two experiments using a recently developed model and find evidence that there is a varying mismatch between lecturers' estimate and their students' actual teaching evaluation scores. The lecturers who are rated low over-estimate whereas the lecturers who are rated high underestimate their expected student evaluation score. In some cases, the latter group of lecturers also over-estimate. However, their overestimate has been less than that of the lecturers who are rated low. We investigate this phenomenon (which is also known in other domains) in the context of 'Lecturers in the University sector' being the subject. We question the suitability of some of the existing theories to explain this phenomenon such as 'unskilled-unaware' theory and 'natural regression' theory. We find the lecturers anchor their estimates to an externally imposed university norm (a target student evaluation score to be achieved by the lecturers) rather than an internal evaluation of their own teaching capability. This finding along with the pattern of over-estimates and under-estimates leads us to propose that 'anchoring and under adjustment' heuristics could be a plausible explanation.

We also find that change in the externally imposed norm had a noticeable impact on the mismatch. This finding tallies with the recent research on the impact of changing the external anchor on the subject of student self-estimation of their marks. We have explored possible reasons for the lecturers' estimate not matching with student evaluation score. Although it has been well established that the student evaluation of teaching performance is riddled with many biases, lack of a usable standard against which to measure the teaching performance leads to the differing perception of teaching standard by lecturers and students. This makes the mismatch between evaluations of lecturers and students more probable and in turn, the resulting repercussions also more probable.

Dynamic Term Structure Models with Score-Driven Time-Varying Parameters: Estimation and Forecasting

Siem Jan Koopman
Vrije Universiteit Amsterdam
Netherlands

Andre Lucas
Vrije Universiteit Amsterdam
Netherlands

Marcin Zamojski
University of Gothenburg
Sweden

We consider score-driven time-varying parameters in dynamic yield curve models and investigate their in-sample and out-of-sample performance for two data sets. In a univariate setting, score-driven models were shown to offer competitive performance to parameter-driven models in terms of in-sample fit and quality of out-of-sample forecasts but at a lower computational cost. We investigate whether this performance and the related advantages extend to more general and higher-dimensional models. Based on an extensive Monte Carlo study, we show that in multivariate settings the advantages of score-driven models can even be more pronounced than in the univariate setting. We also show how the score-driven approach can be implemented in dynamic yield curve models and extend them to allow for the fat-tailed distributions of the disturbances and the time-variation of variances (heteroskedasticity) and covariances. Finally, we propose a novel way of obtaining multi-period forecasts for both the dynamic factors and the variables of interest based on previous normalised score adjustments in score-driven models which may improve performance when the model is mis-specified. This procedure may be applied to other models estimated with score-driven methods.

Dynamically Ensembling Forecasting Models

Maral Fakoor
Marquette University GasDay Laboratory
United States

Sarah Graupman
Marquette University GasDay Laboratory
United States

David Kaftan
Marquette University GasDay Laboratory
United States

Jarrett Smalley
Marquette University GasDay Laboratory
United States

Richard Povinelli
Marquette University GasDay Laboratory
United States

ronald brown
Marquette University GasDay Laboratory
United States

george corliss
Marquette University GasDay Laboratory
United States

Demand forecasting is a necessary process in the supply-chain of natural gas. One of the largest challenges in demand forecasting is adapting to systematic changes in demand. While there are many types of mathematical models for forecasting, there is no perfect formula. It has been found that ensembling several models often results in a better forecast. A common method for ensembling component models is taking a weighted average of the model forecasts. Due to the challenge of adapting to changes in demand, it is important to track the weights associated with each component model in an ensemble. We have developed an ensembling method, called the Dynamic Post Processor (DPP). The DPP ensembles several forecasting models, while tuning the weights based on recent performance of the models. It also removes biases from the component models in order to track changing patterns in natural gas demand. The ensemble results in better forecasts than any of the individual component models. It also reduces the mean forecasting error caused by systematic changes.

Economic Exchange Rate Models: Do Their Forecasts Fit in Colombia?

Juan Julio
Banco de la Republica (Central Bank of Colombia)
Colombia

Daniel Arocha
Universidad de los Andes
Colombia

We evaluate the forecast performance of time series and fundamentals-based exchange rate models in Colombia. Conditional rolling window forecasts derived from co-integration and VAR models as well as trend/gap decompositions are compared to naive, time-varying drift naive and PPP forecasts. Unlike previous studies in Colombia, our sample consists exclusively of records belonging to the free floating regime currently undergoing in Colombia, and our forecasts are derived from the conditional distributions of the exchange rate given that the fundamentals are known along the forecasting horizon. Among other finding our results support that (i) fundamentals-based models predict better the exchange rate only when fundamentals are known along the forecast horizon, (ii) this improvement is far greater for the cycle component of trend/cycle decomposition models, (iii) the fast adaptability of real time forecasts of the smooth nominal exchange rate long-term trend with harmonic models leads to further improvement of the long term exchange rate forecast upon random walks forecasts.

Electricity consumption forecast for the Brazilian residential sector: a bottom-up approach

Fernando Cyrino
PUC-Rio
Brazil

Paula Maçaira
PUC-Rio
Brazil

Reinaldo Souza
PUC-Rio
Brazil

Vanessa Oliveira
ENEVA
Brazil

The importance of the residential class of clients in the electricity consumption in Brazil can be recognized by its share: in 2016 it was responsible for 29% of the total consumption and 86% of total consumers. In this paper, the annual residential electricity demand is modeled through a bottom-up type of approach and forecasts are produced up to year 2050, for three macroeconomic scenarios. The bottom-up is a vintage stock model allowing a detailed modelling of the stock turnover, taking into account the development and diffusion in energy efficiency of end-uses over the years. The annual electricity demand is calculated as the product of the specific consumption per end-use and the corresponding stock. For the residential electricity use, the model covers entertainment appliances (television), air-conditioning, water heating (electrical showers), refrigeration (refrigerator and freezer), lighting and small household services (washing machines and others not distinguished due to a limited data basis). The macroeconomic variables used to explain the residential consumption were: Gross Domestic Product (GDP), population, end consumer energy carrier prices and number of households. Three possible macroeconomic scenarios, using an econometric model, were drawn from the main variables of the bottom-up model in order to obtain different electricity consumption scenarios. The most likely scenario, called as “base”, assumes an annual growth of 6.73% for the GDP, 0.27% annual growth for the population and 1.02% annual growth for number of households up to 2050. The pessimist scenario expects a timid growth than the base, while the optimist scenario assumes a stronger growth. The results show a grow of the electricity demand from about 136 TW in 2016 to 283 TW in 2050 in the base scenario, to 282 TW in the pessimist and to 284 TW in the optimist. As these figures are generated by a bottom-up model, it is possible to analyze the result by end-uses, accordingly, the freezer and the electrical shower have a consumption expected reduction until 2050, due to the increase in the ownership rate of frost-free refrigerators and the increase of natural gas as an alternative source of heating water. When comparing with the official sources, the bottom-up approach showed an average growth of 2.18% a year while the government expects an average growth of 2.84% per year, totaling 352 TW in 2050. This difference can be explained by the adoption of more optimistic assumptions for economic variables.

Electricity End Use Forecasting Using Non-Intrusive Load Metering Technology

Jonathan Farland
DNV GL
United States

Curt Puckett
DNV GL
United States

Fred Coito
DNV GL
United States

Nations with developing economies are often characterized by an increasing population growth rate and consequently a larger increase in the demand for electricity. The construction of new generation facilities, as well as the better management of electricity uses can be used to meet this increase in demand. Unfortunately, collecting metered consumption data on electricity end uses is data intensive and often cost prohibitive. The advent of Non-Intrusive Load Metering (NILM) technology allows for the collection of an unprecedented amount of data and information related to how electricity is actually demanded across utility service territories. Of particular interest is the saturation and energy intensity of electricity end uses across building types and economic sectors. This paper provides a bottom up forecasting framework and results from recent application that incorporates end use level data into a long term energy and peak demand forecast.

Empirical Mode Decomposition and ANN Based Hybrid Approach to Forecast High Frequency Foreign Exchange Rates

Hasanthi Pathberiya
University of Sri Jayewardenepura
Sri Lanka

Liwan Liyanage
Western Sydney University
Australia

Chandima Tilakaratne
University of Colombo
Sri Lanka

Chaotic nature of high frequency exchange rates has made it a challenging task to develop forecasting models with higher accuracy. This unanticipated behavior is subjected to surprises in the economic, political and social environments. However, these altering environments are reflected by various kinds of indicators such as Gross Domestic Product (GDP), Trade Balance, Unemployment Rate, etc. to some extent. This fact is evident in improving capability of forecasting models while incorporating such reflective indicators in such models. On the other hand, as supported by the Efficient Market Hypothesis which states that the market prices reflect all the available information, historical rates are supposed to add much information to these models on top of the aforesaid external factors. Hence, historical prices become the most imperative input in the forecasting model. This study proposes EMD-ANN hybrid approach to develop a model to forecast high frequency exchange rates by combining Empirical Mode Decomposition (EMD) and Artificial Neural Networks (ANN). Instead of considering historical series at different lags as a whole, within this study we propose to decompose the original series to a set of Intrinsic Mode Functions (IMFs) defined by considering their level of fluctuations employing EMD algorithm. Ultimate goal of such decomposition is the incorporation of these IMFs into the forecasting model as a set of weighted functions. Contribution of each IMF leading to improve accuracy of forecasts is quantified by defining weights through a feed forward neural network. Results revealed an improvement in the forecasting accuracy in the model incorporates a set of decomposed functions compared to the model with raw historical data. This result implies the significance of evaluating different components in a time series with respect to their level of fluctuations in the process of deriving short term forecasts with high level of accuracy. The model concludes forecasting accuracy as 68% in terms of yielding at most three percentage in points of absolute residual. Moreover, as evident in the results, the proposed approach reduces the misclassification rate from 29% to 21% in comparison to the model fed with raw historical data.

Ensemble of Specialized Neural Networks for Time Series Forecasting

Slawek Smyl
Uber
United States

The idea of using a pool of diverse forecasting algorithms and applying one of them that is best for a particular situation is an attractive one, but very difficult to realize - numerous efforts failed to create a meta-algorithm that would be able to a priori recommend a forecasting algorithm (just by analyzing the input time series). However, creation of such recommender is easier if the performance of the algorithms on each series is considered. I describe a system that includes a recommender and a pool of specialized neural networks. The networks have the same architecture and inputs, but specialize automatically (to deal well with a particular type of time series) by competing with each other and being reinforced by exposure to successfully predicted examples. This is possible because the neural networks have ability to be trained in the on-line fashion, i.e. gradually, with new examples one by one updating the model.

Estimating the Probability of Occurrence of Parkinson's Disease via Time Series Analysis

Brian Sloboda
Center for Management and Entrepreneurship (CME) School of Advanced Studies (SAS) University of Phoenix
United States

Chandrasekhar Putcha
Department of Civil and Environmental Engineering California State University, Fullerton
United States

Sathyasheel Kondapalli
Department of Civil and Environmental Engineering California State University, Fullerton Fullerton, CA
United States

This paper uses time series analysis to forecast the probability of occurrence of Parkinson Disease in the United States. In brief, Parkinson's disease is caused by the progressive impairment or deterioration of neurons in the area of the brain known as the substantia nigra. Data for Parkinson's disease, collected over a period of time are taken from literature. After the collection of the data, OLS specifications are estimated via linear, quadratic, and cubic specifications. After the implementation of these, splines are also explored. However in some cases, the smooth curves can lead to erroneous results due to rounding-off error and overshooting. Once the number of Parkinson's patients is estimated via OLS, the probability of this critical number of patients being greater than the mean number of Parkinson disease patients as well as the median number of Parkinson disease will be calculated. Parkinson's disease is mainly caused by loss of dopamine (produced by the substantia nigra cells). Consequently, data for dopamine are also collected for these patients. The probability of an actual dopamine level being less than the acceptable dopamine level is calculated for various age groups in order to study in which age group(s) Parkinson disease is most prevalent. This study also concentrates on the probability of Parkinson disease due to various risk factors: age, gender, family history, race and ethnicity. From this information, the critical factors are identified and a prediction is made as to which of these risk factors are likely to lead to the development of Parkinson's disease. From the development of this model, forecasts can be made about the number of people being affected by the disease.

Exchange Rate Prediction from Twitter's Trending Topics

FULYA OZCAN
UNIVERSITY OF CALIFORNIA, IRVINE
United States

This paper investigates whether incorporating sentiment extracted from Twitter's trending topics would improve the intra-day exchange rate predictions. What makes this paper unique is that unlike previous similar studies which only consider tweets that contain the symbol or the name of the currency or stock, it looks at all trending topics irrespective of whether they contain the name or the symbol of the currency. This allows to capture the general sentiment among the users, which implicitly affects the exchange rates.

This paper contributes to two different strands of literature in two disciplines: exchange rate prediction and machine learning. This is the first paper that makes use of data from social media to predict exchange rates with an unsupervised topic clustering model. Studies available in this literature mostly make use of news headlines as their data, and use neural networks or ensembles as their models. The novelty of this paper in exchange rate prediction literature is using Twitter data and also employing a Dirichlet process mixture algorithm to cluster the sentiment-based topics to use them in a time series model to predict the short term exchange rates. This paper is also the first paper that makes use of unsupervised learning in topic extraction, in exchange rate prediction using natural language processing as sentiment from all trending topics are extracted regardless of whether they contain the symbol of the currency in question.

Intra-day foreign exchange rates are predicted by making use of the trending topics from Twitter, using a sentiment based topic clustering algorithm. Twitter trending topics data provide a good source of high frequency information, which would improve the short-term or intra-day exchange rate predictions. This project uses an online dataset, where trending topics in the world are fetched from Twitter every ten minutes since July 2013. First, using a sentiment lexicon, the trending topics are assigned a sentiment (negative, positive, or uncertain), and then using a continuous Dirichlet process mixture model, the trending topics are clustered regardless of whether they are explicitly related to the currency under consideration. This unique approach enables to capture the general sentiment among users, which implicitly affects the currencies. Finally, the exchange rates are estimated using a linear model which includes the topic based sentiment series and the lagged values of the currencies, and a VAR model on the topic based sentiment time series. The main variables of interest are Euro/USD, GBP/USD, Swiss Franc/USD and Japanese Yen/USD exchange rates. The linear model with the sentiments from the topics and the lagged values of the currencies is found to perform better than the benchmark AR(1) model. Incorporating sentiments from tweets also resulted in a better prediction of currency values after unexpected events.

One of the important results from this study is that although this approach cannot predict surprise events, since events need to be heard by the users before a user sentiment can be obtained, it performs better than the benchmark AR(1) model when predicting what will happen after such unexpected events. Moreover, when the topics are added as dummies to the lagged values of the currencies, the prediction errors are found to be lower than the ones from using only the lagged values AR(1). This shows the implicit effect of the sentiment among users on the currency values.

Extending the Logit-Mixed Logit model: Estimation of Demand for Alternative Fuel Vehicles in China

Prateek Bansal
Cornell University
United States

A semi-parametric logit-mixed logit (LML) model to specify the mixing distribution of the preference heterogeneity is the recent advancement in logit-type models. LML provides a general specification for many previous semi-parametric and non-parametric models. Additionally, a special form of the likelihood gradient makes LML computationally more efficient than other parametric models. However, the original LML formulation assumes all utility parameters to be random. This study extends LML to a combination of fixed and random parameters (LML-FR) and motivates such combination in random parameter choice models in general. We further show that the likelihood of LML-FR specification looses its special properties, leading to a much higher estimation time than that of the original LML specification. In an empirical application about preferences for alternative-fuel vehicles in China, estimation time increased by a factor of 15-20 when introducing fixed parameters to the LML model. The comparison of the original LML and LML-FR specifications with other parametric models indicate that LML-FR model outperforms these parametric models in terms of model-fit-criteria (e.g., AIC, BIC, and likelihood), but the same does not hold for the original LML, i.e. when all parameters are assumed random.

Extracting sentiments and their application in forecasting

Michał Chojnowski
Warsaw School of Economics
Poland

This article elaborates on the role of sentiments in the economy. It seeks answers whether the sentiments leave a trace in macroeconomic fundamentals, if they can be tracked back and therefore extracted from historical data. Moreover it discussed their role in the economy and suggests the way of observing them in real life.

To detect the sentiments SVAR model with short-term identification was proposed to extract the signals individuals might have received. Based on those signals sentiments can be detected using Hodrick-Prescott filter. To check if the extracted time-series can be regarded as the sentiment, they were implemented into the forecasting models (ARMA) as exogenous variable to observe improvement of a prediction power. Additionally, Google trends were used to investigate if sentiments leave traces in Internet traffic as well, henceforth they could be detectable in real-time data.

The results are promising – for three examined markets we succeeded to extracts time-series, which may be regarded as sentiments present in this market. Although only car market shows significant improvement, in other two the RMSE of forecast is still smaller than original ARMA model. The Google trends data was used to imitate the sentiments using Internet activity and although sentiments can be observed via browsing traffic, it does not bring as much improvement as sentiments extracted. To check how model behaves when there is no influence of sentiments implied, control scenario was considered.

Feature-based Model Selection for Time Series Forecasting

Thiyanga Shamini Talagala
Monash University
Australia

Rob J Hyndman
Monash University
Australia

George Athanasopoulos
Monash University
Australia

Many applications require a large number of time series to be forecast. Providing better forecasts for these time series is important in decision and policy making. However, large scale time series data present numerous challenges in modelling and implementation due to the high dimensionality. It is unlikely that a single method will consistently provides better forecasts across all time series. On the other hand, selecting individual forecast models when the number of series is very large can be extremely challenging. In this paper we propose a classification framework which selects forecast models based on features calculated from the time series. A Random Forest approach is used to develop the classifier. The proposed framework is tested using the M3 data and is compared against several benchmarks and other commonly used approaches of forecasting.

Fiscal Sustainability – Measurement Using Event Probability Forecasts in OECD Countries

Kevin Lee
University of Nottingham
United Kingdom

Kian Ong
University of Melbourne
Australia

Kalvinder Shields
University of Melbourne
Australia

This paper proposes a measure of public sector debt sustainability based on probabilistic statements of two future events relating to debt sustainability. The approach accommodates two alternative definitions of sustainability, which is new in the literature. We illustrate the approach employing a Global Cointegrated Vector Auto Regression model, modelling both cross-country interactions and long-run relations. Applying on ten developed countries over the period 1958-2013, shows that our measures offer a multi-faceted description of fiscal sustainability defined not only on the fiscal balances, but also on the movements of real interest rates and output growth rates.

FORECAST INFLATION FROM INDIVIDUAL COMPONENTS: THE COLOMBIAN CASE

Wilmer Martinez
Banco de la República de Colombia (Central Bank)
Colombia

Eliana Gonzalez
Banco de la República de Colombia (Central Bank)
Colombia

We develop an empirical methodology to obtain forecast for total inflation from aggregating individual disaggregate forecast. To do that we estimate a model for each component of the CPI basket at the most disaggregated level. Furthermore, we evaluate the relationship between each component with other potential explanatory variables such as components of the IPP index, exchange rate, some climate indicators, and economic activity indexes. Finally, using Bayesian moving averaging (BMA) and ARIMAX models we evaluate the forecast performance of each possible model for each item, and once we identify the best model, we aggregate the individual forecasts with its respective weight in the CPI basket. We found that for some items the best forecasting model is an ARIMA while for others the additional variables have forecasting power. The aggregated forecast performs better than the one obtained from a single model for the aggregate inflation.

Forecast revisions and Business cycles: Evidence from 12 years of ESP Forecast Survey experience in Japan

Nobuo Iizuka
Kanagawa University
Japan

The ESP Forecast Survey (ESPF) is the first monthly survey of macroeconomic forecasts conducted by professional forecasters in Japan. The survey was launched in 2004 and is conducted by the Economic Planning Association for Japan. In 2012, the Japan Center for Economic Research took over this survey. The introduction of ESPF has made it possible to observe monthly changes in macroeconomic forecasts. Consensus Economics, the world's leading international economic survey organization, has a longer history than ESPF, and its surveys were analyzed in earlier studies concerning Japan (Chen et al., 2016). The participants of the ESPF panel are twice as many as for the Consensus Economics panel for Japan.Moreover, the ESPF not only has annual multidimensional data but also quarterly multidimensional data. This paper analyzes the forecasts of GDP growth and inflation contained in ESPF for the period 2004- 2016.Annual forecasts are for the current and the following year. Quarterly forecasts are for up to four quarters ahead. The forecasts comprise an unbalanced three-dimensional panel with multiple individual forecasters, target years, and forecast horizons (Davies & Lahiri, 1995). We evaluate the accuracy of the forecasts, examine the revision process, and try to discriminate between forecast errors that arise from unforecastable macroeconomics shocks and forecasts errors that are idiosyncratic errors, as performed by Capistrán & López-Moctezuma (2014) for Mexican data.

Reference Capistrán, C. & López-Moctezuma, G. (2014). Forecast revisions of Mexican inflation and GDP growth. International Journal of Forecasting, 30, 177-191. Davis, A. & Lahiri, K. (1995). A new framework for analyzing survey forecasts using three-dimensional panel data, Journal of Econometrics, 68, 205-227. Chen, Q. & Constantini, M. & Dechamps, B. (2016) How accurate are professional forecasts in Asia? Evidence from ten countries. International Journal of Forecasting, 32, 154-167.

Forecasting annual incidence and mortality rate for prostate cancer in Australia until 2022 using Autoregressive Integrated Moving Average (ARIMA) models

Arul Earnest
Monash University

Australia

Sue Evans
Monash University
Australia

Fanny Sampurno
Monash University
Australia

Jeremy Millar
Monash University
Australia

Introduction: Prostate cancer is the second most common cause of cancer-related death in males after lung cancer, and thus poses a significant burden on the healthcare system. Current methods used to forecast prostate cancer in Australia such as linear and logarithmic trend models may not be adequate. Challenges include temporal correlation and the impact of population changes. We propose the use of ARIMA models in conjunction with population forecasts to provide for robust annual projections of prostate cancer incidence and mortality rate in Australia. Methods: Data on the incidence of prostate cancer was obtained from the Australian Institute of Health and Welfare (AIHW) which included data from the Australian Cancer Database. Mortality data was from the National Mortality Database. The study period included cases diagnosed from 1982 to 2012. We formulated several ARIMA models and chose one which provided for an accurate fit of the data based on the Mean Absolute Percentage Error (MAPE). We assessed the model for external validity by developing the model using data from 1982-1999 and then validating the predictions using data from 2000 till 2013. Based on the final model selected, we estimated the annual number of cases expected to be diagnosed in Australia from 2013 to 2020. Results: The annual number of prostate cancer cases diagnosed in Australia increased from 3606 in 1982 to 20065 in 2012. There were two peaks observed around 1995 and 2010 have been thought to be related to PSA testing being listed in the Medicare Benefits Schedule in 1989. Among the various models evaluated, we found that the model with an autoregressive term of 1 (coefficient=0.45, p=0.028) as well as differencing the series provided the best fit, with a MAPE of 5.2%. External validation showed a good MAPE of 5.8% as well. Projections for prostate cancer incidence cases from 2013 till 2022 will be presented. Conclusion: Our study has accurately characterised the trend of prostate cancer incidence in Australia, and we have used this information to prospectively forecast the annual number of new cases expected to be diagnosed up till 2022. This information will prove useful for resource planning and manpower allocation.

Forecasting by Iterative Semi-supervised Learning with Application to Customer Lifetime Values

Fan Yang
Uber
United States

Venkatesh Periaswamy
Uber
United States

Jimmy Feng
Uber
United States

Prakhar Mehrotra
Uber
United States

Traditional forecasting methods such as exponential smoothing and ARIMA are designed for, and have seen great success in solving macro-metrics forecasting problems. As we move into big data era and are blessed with high-volume micro level data, it poses new challenges: aggregation over granular data to enable traditional methods implies considerable loss of information. In this article we propose a novel semi-supervised learning-based modeling framework that tackles these challenges. It accommodates incorporation of these features and allows granular level forecasts. At the same time, it assumes similarity between temporally neighboring data points, just like traditional methods. We illustrate the strength of this framework by applying it to customer lifetime value (LTV) forecasting problem. Results show improved forecast accuracy while preserving high dimensional granularity.

Forecasting Challenges at Amazon

Tim Januschowski
Amazon
Germany

At Amazon, forecasting problems arise naturally across the different businesses such as retail, cloud computing, logistics, professional services, etc. In my talk, I will describe some of these forecasting problems in more detail and highlight the technical and modeling challenges. Examples include data volume, short time series and intermittent time series with low counts. I will sketch the systems, models and algorithms the Core Machine Learning team develops in order to handle these problems at scale.

Forecasting Compositional Time Series: A State Space Approach

Ralph Snyder
Monash University
Australia

Keith Ord
Georgetown University
United States

Anne Koehler
Miami university of Ohio
United States

Keith McLaren
Monash University
Australia

Adrian Beaumont
University of Melbourne
Australia

A framework for the forecasting of composite time series, such as market shares, is proposed. Based on Gaussian multi-series innovations state space models, it relies on the log-ratio function to transform the observed shares

(proportions) onto the real line. The models possess an unrestricted covariance matrix, but also have certain structural elements that are common to all series, which is proved to be both necessary and sufficient to ensure that the predictions of shares are invariant to the choice of base series. The framework includes a computationally efficient maximum likelihood approach to estimation, relying on exponential smoothing methods, which can be adapted to handle series that start late or finish early (new or withdrawn products). Simulated joint prediction distributions provide approximations to the required prediction distributions of individual shares and the associated quantities of interest. The approach is illustrated on US automobile market share data for the period 1961–2013.

Forecasting Demand in Supply Chain Using Markov-Switching Model

Mahdi Abolghasemi
University of Newcastle
Australia

Ali Eshragh
University of Newcastle
Australia

Mojtaba Heydar
University of Newcastle
Australia

Demand forecasting is an inevitable task in supply chain management. Due to the high volume of uncertainty imposed to the supply chains both endogenously and exogenously, the regime of the system may vary significantly, and consequently a single statistical model may not suffice to forecast demand with a desirable level of precision. Hence, incorporating a stochastic model may improve demand forecasts derived by an Autoregressive Integrated Moving Average(ARIMA) model. In this paper, we utilize Markov-switching model for demand forecasting in a two-stage supply chain.

Forecasting Electricity Spikes with Conditional Asymmetries in Return Innovations

Laleh Tafakori
The University of Melbourne
Australia

Armin Pourkhanali
Monash UNiversity
Australia

Farzad Alavi Fard
RMIT
Australia

In this paper we consider the problem of modelling and forecasting the distribution of a vector of returns from electricity spot markets. We use a class of observation-driven time series models referred to as asymmetric exponential generalised autoregressive score (AEGAS) models. The mechanism to update the parameters over time is provided by the scaled score of the likelihood function in the AEGAS model. Based on this new

approach, the results provide a unified and consistent framework for introducing time-varying parameters in a wide class of non-linear models. Since the AEGAS model is based on the score, it exploits the complete density structure rather than means and higher moments only. This differentiates the AEGAS model from other observation-driven models in the literature, such as the generalized autoregressive moving average models and the vector multiplicative error models.

From the financial risk management point of view, the central problem is the actual quantification of the risks. We evaluate the accuracy of several hundred one-day-ahead value at risk (VaR) forecasts in a comprehensive set of in-sample and out-of-sample tests. Our study is conducted in the Australian National Electricity Market (NEM), which is the most efficient power auction in the world. Unlike all others, electricity is traded in a constrained real time spot market. Further, electricity prices in NEM exhibit highly stylised features such as extreme price spikes and a strong seasonal patterns, which complicates the estimations process and hampers the forecasting accuracy. In contrast to the jumps observed in the financial markets, spot price spikes are normally quite short-lived, and as soon as the weather phenomenon or outage is over, prices fall back to a normal level.

We evidence that our model substantively provides accurate forecasts of the underlying distribution and outperforms a set of competing models in their abilities to forecast one-day-ahead VaRs. The set of competing models include GJR-GARCH, EGARCH, and EVT-EGARCH models.

Forecasting eSchool: from gamification to superforecasting

Nikoletta Zampeta Legaki
Forecasting & Strategy Unit, National Technical University of Athens
Greece

Konstantinos Nikolopoulos
Bangor Business School, Prifysgol Bangor University, Wales
United Kingdom

Fotios Petropoulos
School of Management, University of Bath
United Kingdom

Leonard Tashman
Foresight: The International Journal of Applied Forecasting, International Institute of Forecasters
United States

Vassilios Assimakopoulos
Forecasting & Strategy Unit, National Technical University of Athens
Greece

Dimitrios Thomakos
School of Economics, Management and Informatics, Department of Economics, University of Peloponnese
Greece

Forecasting eSchool (ForceS) aspires to advance the educational process of delivering forecasting courses: via introducing gamification and innovative assessment approaches. ForceS is a modular information system that currently consists of three different games aiming to innovatively deliver the learning outcomes in respect to three popular academic works: a) “Horses for Courses” based on recent work on forecasting model selection, b) “JudgeIT” based on the seminal work on judgment and forecasting under uncertainty of Tversky & Kahneman, and c) “Metrics to escape” based on contemporary ideas of how to measure forecast accuracy and uncertainty.

Preliminary results have illustrated the positive effects of ForceS on both students' engagement as well as the learning process per se. In order to build on the aforementioned amplified learning effects and assess them respectively, ForceS integrates a superforecasting application, aiming to identify the top performers from a set of trained forecasters - along the principles and respective outcomes of the "good judgment project" in USA.

Forecasting for Customer Service in Booking.com

Brammert Ottens
Booking.com
Netherlands

Booking.com is the world's number one website for booking hotels and other accommodation, processing on average over 1 million room nights every 24 hours. In order to support our customers and partners, Booking.com operates a worldwide 24/7-customer service center servicing 40+ different languages, employing thousands of customer service agents. Having enough agents available in the right language at the right time is a challenging problem. Forecasting the right amount of work is paramount in achieving this while keeping the costs in check. In this talk we will focus on three aspects of our forecasting operation. We start with the needs that our business has, and the peculiarities of forecasting for a multi lingual contact center. We will discuss the models we are using and how we have automated running these forecasts daily for the business to use.

Forecasting implied volatility in foreign exchange markets: Multivariate and multilevel functional time series methods

Han Lin Shang
Australian National University
Australia

Fearghal Kearney
Queen's University Belfast
Ireland

Multivariate and multilevel functional time series methods are adapted for forecasting daily implied volatility in foreign exchange markets. On one hand, the multivariate functional time series method first standardizes all series, then combine them into a long vector for each time period. Via a functional principal component analysis, the extracted functional principal components and their associated scores can be obtained from a joint covariance structure consisting of all series. On the other hand, multilevel functional time series method uses multilevel functional principal component analysis of aggregate and maturity-specific data to extract the common trend and maturity-specific residual trend among different maturities. Both multivariate and multilevel functional time series methods are able to capture correlation among maturities, and are shown to produce more accurate forecasts than the ones obtained from the univariate functional time series method. To demonstrate the effectiveness of the proposed methods, we examine the daily implied volatility curves of foreign exchange option, namely EUR-USD, EUR-GBP and EUR-JPY for a number of maturity periods.

Forecasting in the presence of unknown distributions with the Linear-Ellipsoidal-Bounded (LEB) Filter

Claudio Antonini
BNY Mellon
United States

When modeling and forecasting dynamic systems, on many occasions, the type of a distribution is not known and one has to resort to assumptions, dubious approximations, or enlarging the model with parameters corresponding to the unknown dynamics. We present a method that creates a filter that describes both the system and measurement noises only by their minimum and maximum temporal values, independently of the noise distributions. The resulting filter has exactly the same structure and follows the same two-step process (propagation of the state vector and measurement update) than a Kalman filter, which can be easily replaced for testing purposes. By describing the noises by their bounds, it is not necessary to add artificial states to the Kalman filter to estimate extra parameters, allowing a reduction in the number of states.

The process works by (a) describing the state and system noise by two ellipsoids, (b) propagating these ellipsoids by creating a new ellipsoid that contains them, (c) intersecting the result with the noisy measurement, (d) creating a new ellipsoid that circumscribes the intersection. This last ellipsoid becomes the new estimated state and the cycle repeats itself. Given that each step of creating a circumscribing ellipsoid needs one parameter, we need two new parameters at each step of the iteration: one for the propagation--beta--, and one for the estimate--rho--. We derive the optimal parameters (beta and rho) by minimizing the volumes of the circumscribing ellipsoids (in steps (b) and (d) above), which are proportional to the trace and the determinant of the matrices describing them. One of the parameters results in a closed-form formula and the other is solved by a few Newton iterations.

We apply this methodology to a tracking problem and to a model of the economy, discussing its benefits in comparison to the Kalman filter.

Forecasting Industrial Production: Gain from Disaggregating

Dennis Jansen
Texas A&M University
United States

Dandan Liu
Kent State University
United States

Industrial production is a key economic indicator. The business cycle is largely influenced by the dynamics of industrial sectors, and the industrial production (IP) index, available at a monthly frequency, is more timely than real GDP data and hence is closely watched by business, finance, and economic professionals. Timely and accurate forecasts of industrial production can help business planners, financial companies, and economic policy makers better anticipate future developments of the business cycle and the aggregate economy.

We investigate the idea of using information in subcomponents to improve forecasts of aggregates, as applied to the growth rate of industrial production (IP). Our work is related to the macroeconomic forecasting literature that looks at the usefulness of summarizing information in large datasets by estimating common factors, but in our case we summarize information contained in a dataset of subcomponents of the series to be forecast. Three disaggregating ways are considered: by industry, by markets and by broad industries. Several forecasting approaches are used to predict the aggregated IP: univariate AR model on the aggregate, AR models for IP components and then combination of forecasts with several weighting schemes, univariate factor-augmented AR model on the aggregate.

We conduct a simulated out-of-sample forecasting exercise to compare forecasts to those of a benchmark univariate AR model at forecast horizons ranging from 1 month to 24 months. In terms of point forecasts, we find that the best forecasting model is one that forecasts the subcomponents of IP and aggregates them using the actual weights of the subcomponents in IP from the date the forecasts are made. This result holds over all horizons, and it holds not only in our preferred increasing-window forecasts but also generally holds for our rolling-window forecasts as well. We find the similar results for the density forecasts.

Forecasting Multiple Time Series with one sided Dynamic Principal Components

Daniel Peña
Universidad Carlos III de Madrid
Spain

Ezequiel Smucler
Universidad Buenos Aires
Argentina

Victor yohai
Universidad de Buenos Aires
Argentina

We define one sided dynamic principal components (ODPC) for time series as linear combinations of the present and past values of the series that minimize the reconstruction mean squared error. Previous definitions of dynamic principal components depend on past and future values of the series. For this reason, they are not apt for forecasting purposes. On the contrary, it is shown that the ODPC introduced in this paper can be successfully used for forecasting large dimensional multiple time series. Monte Carlo results show that a forecasting procedure based on the OSDPC compares favourably with other forecasting methods based on dynamic factor models.

Forecasting Performance of Markov-Switching GARCH Models: A Large-Scale Empirical Study

David Ardia
University of Neuchatel
Switzerland

Keven Bluteau
University of Neuchatel
Switzerland

Kris Boudt
Vrije Universiteit Brussel
Belgium

Leopoldo Catania
University of Rome, 'Tor Vergata'
Italy

We perform a large-scale empirical study to evaluate the forecasting performance of Markov-switching GARCH (MSGARCH) models compared with standard single-regime specifications. We find that the need for a Markov-switching mechanism in GARCH models depends on the underlying asset class on which it is applied. For stock data, we find strong evidence for MSGARCH while this is not the case for stock indices and currencies. Moreover, Markov-switching GARCH models with a conditional (skew) Normal distribution are not able to jointly account for the switch in the parameters as well as for the excess of kurtosis exhibited from the data; hence a Markov-switching GARCH model with (skew) Student-t specification is usually required. Finally, accounting for the parameter uncertainty in predictions, via MCMC, is necessary for stock data.

Forecasting Realized Volatility and the Role of Time-Varying Dependence with Market Returns

Leopoldo Catania
University of Rome, Tor Vergata
Italy

Tommaso Proietti
University of Rome, Tor Vergata and CREATES
Italy

Forecasting realized volatility has become of primary importance for business applications in risk management, asset allocation and pricing of derivative instruments. This paper proposes new models specifically tailored for realized volatility predictions. Our proposals effciently account for several well-known stylised facts characterising realised volatility time-series, such as: (i) long memory of the volatility process, (ii) presence of outliers, and (iii) negative dependence with daily market returns. Furthermore, the influence of volatility in volatility, time-varying leverage and quarticity effects are also investigated in the forecasting performance. Our analysis also covers the heavily undertaken aspect of density prediction of the future level of market volatility and its relevance in real applications. Our results show that the new models provide very high outperformance rates with respect to the benchmark HAR-RV, both in terms of point and density forecast.

Forecasting Sets of Similar Time Series with Recurrent Neural Networks

Kasun Bandara
Monash University
Australia

Christoph Bergmeir
Monash University
Australia

Slawek Smyl
Uber Technologies Inc
United States

Nowadays, in many applications, large quantities of similar time series are available, such as server performance measures in computer centers, sales in retail of thousands of different products, measurements for predictive maintenance, smart meter data. Forecasting time series in these domains with traditional univariate forecasting procedures leaves great potentials for producing accurate forecasts untapped.

This research introduces a methodology to exploit the similarities between time series to improve the forecasting accuracy over the state of the art of univariate time series forecasting that forecast time series in isolation. With the recent developments in the area of time series forecasting, we propose a prediction model using Long Short Term Memory (LSTM) networks, a special type of recurrent neural networks (RNN) trained in a way to exploit similarities between time series, which is supplemented by traditional time series preprocessing techniques such as Seasonal and Trend Decomposition using Loess (STL) with Log Transformation.

Specifically, our proposed method initially discovers cluster of similar series from the overall set of time series as an augmentation step to exploit the similarity between time series. After applying a log transformation to the data, with the objective of stabilizing the variance of the time series,we then deseasonalize the time series using STL decomposition, abiding to the premise that neural networks often have difficulties to model seasonality directly. Then, we train and build predictive models for each of the preprocessed subgroups using LSTM neural networks.

We evaluate the proposed methodology on different benchmark competition datasets and are able to achieve very competitive results.

Forecasting spare part demand with Installed Base information

Sarah Van der Auweraer
KULeuven
Belgium

Robert Boute
KULeuven and Vlerick Business School
Belgium

Aris Syntetos
Cardiff Business School
United Kingdom

Service maintenance is commonly used to extend the lifetime of capital assets, such as manufacturing equipment or heavy infrastructure. When a service part, necessary to perform the maintenance action, is required but not immediately available, the incurred shortage costs may be substantial. For this reason, companies keep large stock buffers to deal with uncertain demand of these spare parts. Specialized service parts models should therefore focus on improving the availability of parts whilst limiting the investment in inventories. An important characteristic of most service parts is their intermittent demand pattern, for which specific forecasting techniques have been developed (see e.g., the review of Boylan and Syntetos, 2010). Many of these methods, however, rely on the time series of the historical demand and do not take into account the factors that generate the spare part demand: the failure behaviour of the components, the maintenance policy, etc. We refer to these factors as the Installed Base information. In our work we provide an overview of the papers which use installed base information for forecasting future service parts demand and we develop a new model which incorporates this information. Dekker et al. (2013) define installed base information as the information on the set of systems or products for which a company provides after sales services. It can include the number of installed and serviced machines (i.e. the size of the installed base), its evolution over time, the failure behaviour of the parts, part age information, and the part replacement probability. In addition to that, it is also possible to include information on the sudden and scheduled service needs of the products. Because the maintenance policy in use has an impact on the demand of spare parts, taking this information into account will improve the predictability of service parts demand. We aim to provide a new model which uses installed base information to predict future demand. This model combines information on the maintenance policy, the size of the installed base and its evolution over time,

the part failure behaviour, and the replacement probability, in order to capture the full picture of the demand generating process.

Time-Series Modeling with Neural Networks at Uber

Nikolay Laptev
Uber
United States

As Uber services such as UberEats and UberRush grow, we need to accurately forecast their demand. With accurate forecasting, we can optimally scale our infrastructure, provide estimated time of arrival (ETA) given demand and prepare our budget. Over-forecasting or under-forecasting can cost millions in over-spending, underutilized hardware or unhappy customers. Currently, service demand forecasting requires a manual process and is prone to errors and continuous readjustment. We present several models, dealing with the short term and long term forecasting challenge, ranging from the classic time-series models to recurrent neural nets in a univariate and multivariate setting. We demonstrate the performance of these models on real use-cases including compute resources, ETAs, and special event trip forecasting.

Forecasting the Term Structure of Interest Rates Using Macroeconomic Factors

Weiqi Xiong
Rutgers
United States

In this paper, an extensive set of forecast experiments is conducted to provide further evidence on the effect of adding macro-variables on improving out-of-sample forecast performance of interest rate models in multiple sub-samples and in periods of economic recessions. In particular, we assess the forecast accuracy of a number of models commonly used for monthly prediction of the term structure of interest rates. These models include: (i) diffusion index type models developed and discussed in Stock and Watson (2002b); Bai and Ng (2006); Ayi Armah and Swanson (2010); (ii) DNS type models of the variety recently examined by Diebold and Li (2006), and various benchmark models, including VAR and AR models. Common macroeconomic factors are extracted from 103 U.S macro-variables (McCracken and Ng, 2016). It is found that, adding macro factors extracted from a large macroeconomic data set could improve the out-of-sample forecast accuracy in 1992-2008 for the whole yield curve, regardless of the sample period in recession or in normal time. However, in the post crisis sample from 2007:Dec - 2009:Jun, no “data-rich” prediction models can beat an AR(1) benchmark.

Forecasting unemployment in Scandinavia: multivariate and hierarchical modelling

Nikolaos Kourentzes
Lancaster Centre for Forecasting
United Kingdom

Rickard Sandberg
Stockholm School of Economics
Sweden

The four major Scandinavian economies (Denmark, Finland, Sweden and Norway) have high workforce mobility and depending on market dynamics the unemployment in one country can be influenced by conditions in the neighbouring ones. We provide evidence that Vector Autoregressive modelling of unemployment between the four countries produces more accurate predictions than constructing independent forecasting models. However, given the dimensionality of the VAR model its specification and estimation can become challenging, particularly when modelling unemployment across multiple factors. To overcome this we consider the hierarchical structure of unemployment in Scandinavia, looking at three dimensions: age, country and gender. This allows us to construct multiple complimentary hierarchies, aggregating across each dimension. The resulting grouped hierarchy enforces a well-defined structure to the forecasting problem. By producing forecasts across the hierarchy, under the restriction that they are reconciled across the hierarchical structure, we provide an alternative way to establish connections between the time series that describe the four countries. We demonstrate that this approach is not only competitive with VAR modelling, but as each series is modelled independently, we can easily employ advanced forecasting models, in which case independent and VAR forecasts are substantially outperformed. Our results illustrate that there are three useful alternatives to model connections between series, directly through multivariate vector models, through the covariance of the prediction errors across a hierarchy of series, and through the implicit restrictions enforced by the hierarchical structure. We provide evidence of the performance of each, as well as their combination.

Time series forecasting: an eccentric way forward

Ralph Snyder
Monash University
Australia

A common approach nowadays to time series forecasting is model centric: a model is specified, a compatible filter is devised to generate in-sample one-step ahead prediction errors needed to evaluate fit, parameters are estimated, and finally out-of-sample predictions are made. It will be argued, however, that this general approach has largely run its course and that a more flexible alternative is needed. The case will be made for one which eschews model specification, replacing it instead with filter specification: a filter is no longer derived but specified.

Forecasting with Approximate Bayesian Computation

Gael Martin
Monash University
Australia

Brendan P.M. McCabe
University of Liverpool
United Kingdom

David Frazier
Monash University
Australia

Worapree Maneesoonthorn
University of Melbourne

Approximate Bayesian Computation (ABC) has become an increasingly prominent tool of choice in challenging inferential problems, most notably those characterized by an intractable likelihood function. ABC requires only that one can simulate pseudo data from the empirical model, for given draws of the parameters from the prior. Parameter draws that produce a `match' between the pseudo and observed data are retained and used to estimate the posterior distribution, with the accuracy of the resultant estimate of the exact (but inaccessible) posterior dependent on the informativeness of the statistics used in the matching. Along with the growth in applications of ABC, attention has recently been paid to the theoretical underpinnings of the method, in particular to the asymptotic behaviour of ABC posterior distributions and point estimates of parameters. In this paper we provide the first exploration of ABC in a forecasting context. That is, attention is switched from parameter inference to the production of Bayesian predictive distributions that condition on summary statistics. Summaries that are themselves linked with forecasting accuracy are investigated, as are different approaches to the use of auxiliary models to produce the summaries. The extent to which existing asymptotic results regarding parameter inference have relevance in the forecasting realm is of particular interest.

Full trajectory prediction: What will you do in the rest of the day?

Amin Sadri
RMIT University
Australia

Understanding and predicting human mobility is the key problem in a large number of applications. Existing works on human mobility prediction mainly focus on the next location prediction after a specified time (e.g. one hour). Some other works predict a set of location that would be visited by the user without taking into account the locations sequence and the timing of the departures. In this paper, given the historical data and the user's trajectory in the morning, we predict the full trajectory in the afternoon. The full trajectory includes the sequence of the locations, the staying times, and the departure times.

Future Scenario of Dairy Developing States of India in 2030: A Delphi Analysis

Rajiv Kale
ICAR- Agricultural Technology Application Reseacrh Institute, Jodhpur, Rajasthan, 342005
India

K Ponnusamy
ICAR-National Dairy Research Institute, Karnal, Haryana, 132001
India

R Sendhil
ICAR-Indian Institute of Wheat & Barley Research, Karnal, Haryana
India

M S Meena
ICAR- Agricultural Technology Application Reseacrh Institute, Jodhpur, Rajasthan, 342005
India

S K Singh
ICAR- Agricultural Technology Application Reseacrh Institute, Jodhpur, Rajasthan, 342005
India

Dairy development in India is phenomenal since the white revolution but with regional disparities. The dairy developing states are highly important for growth of milk production. An attempt has been made to analyze the future scenarios of dairy developing states of India using Delphi and scenario development methodology. Experts from different domains were used as Delphi panelists to identify the key uncertainties and important driving forces and scenarios development for 2030. Index values for indicators were used to priorities the uncertainties and important driving forces. Climate vulnerability and technological up-scaling were found to be key uncertainties. Four plausible dairy scenarios for 2030 were developed during the scenario workshop. Study provided strategies for achieving the desirable future scenario.

Gauging the Uncertainty of the Economic Outlook Using Historical Forecasting Errors: The Federal Reserve's Approach

David Reifschneider
Federal Reserve Board of Governors
United States

Peter Tulip
Reserve Bank of Australia
Australia

Since November 2007, the Federal Open Market Committee (FOMC) of the U.S. Federal Reserve has regularly published participants' qualitative assessments of the uncertainty attending their individual forecasts of real activity and inflation, expressed relative to that seen on average in the past. The benchmarks used for these historical comparisons are the average root mean squared forecast errors (RMSEs) made by various private and government forecasters over the past twenty years. This paper documents how these benchmarks are constructed and discusses some of their properties. We draw several conclusions. First, if past performance is a reasonable guide to future accuracy, considerable uncertainty surrounds all macroeconomic projections, including those of FOMC participants. Second, different forecasters have similar accuracy. Third, estimates of uncertainty about future real activity and interest rates are now considerably greater than prior to the financial crisis; in contrast, estimates of inflation accuracy have changed little. Finally, fan charts—constructed as plus-or-minus one RMSE intervals about the median FOMC forecast, under the expectation that future projection errors will be unbiased and symmetrically distributed, and that the intervals cover about 70 percent of possible outcomes—provide a reasonable approximation to future uncertainty, especially when viewed in conjunction with the FOMC's qualitative assessments. That said, an assumption of symmetry about the interest rate outlook is problematic if the expected path of the federal funds rate is expected to remain low.

Global temperatures and green house gases - A common features approach

Li Chen
Monash University
Australia

Jiti Gao
Monash University
Australia

Farshid Vahid
Monash University
Australia

Despite the difficulty in determining the exact nature of trend in global temperatures, we can establish that the trend in global temperatures and green house gases is common. We then estimate the long-run relationship between global temperatures and green house gases and provide theoretically justified measures of confidence in our estimates. We compare our results with the existing results in the literature.

Health state monitoring of engineered systems based on variational autoencoders and K-means clustering

Peng Jiang
School of Materials Science and Engineering, UNSW Australia
Australia

Serkan Saydam
School of Mining Engineering, UNSW Australia
Australia

Hamed Ramandi
School of Mining Engineering, UNSW Australia
Australia

Alan Crosky
School of Materials Science and Engineering, UNSW Australia
Australia

Mojtaba Maghrebi
Department of Civil Engineering, Ferdowsi University of Mashhad
Iran

Effective monitoring of engineered systems is critical for condition-based maintenance (CBM). In this study a hybrid clustering algorithm for health state monitoring of engineered systems is developed. This model utilizes Variational Autoencoders (VAE) for transforming original data into a low-dimensional latent space. Then K-means algorithm is used to automatically separate the latent representations into several clusters which represent different health states. The proposed model is tested on a dataset obtained from an aircraft engine run-to-failure simulation. Results show that the developed model is able to predict the health states of this system with a high classification performance. In addition, the visualizations produced by VAE assist in gaining valuable insights into the degradation mechanism. This hybrid model can be also applied to other engineered systems such as batteries, bearings and chips to effectively monitor system health states.

Hierarchical forecasting with linear boosted models for GEFCom 2017

Cameron Roach
Monash University
Australia

Time-series can exist in hierarchical configurations. Examples include sales of a product within towns and also by state; forecasting economic indicators within states and for an entire country; and in this case, forecasting demand in bottom level and aggregated zones of an electricity network. When forecasting hierarchical time-series it is often necessary to manually adjust base forecasts to ensure they reconcile sensibly.

GEFCom 2017 provided the motivation for developing this model. Quantile forecasts for eight bottom level zones and two aggregated zones were produced for every hour of a future month. A simulation based approach was used to generate demand forecasts for the month. Adjustments were then made to each of the demand simulations to ensure all zone forecasts reconciled appropriately. Several methods for adjusting the base forecasts were tested.

This paper outlines a methodology used to produce hierarchical quantile forecasts and provides a discussion of results and modelling performance. Our model is compared against a credible benchmark method and is found to perform better. A brief review of hierarchical time-series forecasting and gradient boosting is also included.

How do we forecast climate changes?

Steven Sherwood
University of New South Wales
Australia

Many forecasting problems benefit from a large number of past (training) examples, in which case forecasting techniques can be refined empirically. Climate changes however, on time scales from seasonal to long-term, are a prime example where data are too limited to support empirical predictions and in some cases no close past analogues exist. This talk discusses how with respect to climate one develops confidence in predictions of events for which there are no perfect past analogues, but where this is a rich set of indirect knowledge that can be brought to bear. A particular focus is the prediction of global-mean warming under a given future greenhouse-gas emission pathway. As posed, this is a pure boundary-value problem, unlike the situation in weather forecasting, which is an initial value problem—which renders climate prediction easier than weather forecasting in some respects. Climate prediction problems on shorter (e.g. seasonal or annual) time scales, however, present an interesting mix of initial- and boundary-value aspects due to the presence of multiple Earth system components with widely varying intrinsic time scales. Current methods and future prospects for both shorter-term and long-term climate prediction will be reviewed.

Individual vs group: Advice Taking in Judgmental Forecasting

Hyo Young Kim
KAIST College of Business
South Korea

Yun Shin Lee
KAIST College of Business
South Korea

Duk Bin Jun
KAIST College of Business

South Korea

In this study, we examine advice taking behavior when individuals or groups make judgmental forecasts given external forecast advice. Using a controlled laboratory experiment, we investigate the differences between individuals and groups in the degree of advice taking and the ability to discern advice quality. Participants in an individual condition make decisions individually without any interactions with others whereas participants in a group condition interactively discuss with other two participants and agree on single decisions. All participants are randomly assigned to one of the two advice quality conditions where reasonable or unreasonable forecast advice is given. We find that groups are less receptive to advice than individuals and this relationship is mediated by confidence level; groups feel more confident than individuals, which decreases reliance upon advice. We also find that groups, who feel more confident, are able to discriminate the quality of advice better than individuals. While groups are able to take reasonable advice more than unreasonable advice, their reliance upon reasonable advice falls short compared to individuals due to their higher tendency to discount advice in general.

Inheriting hierarchical information for retail sales forecasting at sku-store level

Daniel Waller
Lancaster University
United Kingdom

John Boylan
Lancaster University
United Kingdom

Nikolaos Kourentzes
Lancaster University
United Kingdom

In retail, sales time series of a large number of products or stock-keeping units (SKUs) are often collected, normally existing within a two-dimensional hierarchy. Firstly, sales can be summed across groups of SKUs into various categories of increasing size, and secondly, they can be summed across groups of stores. In choosing a level of aggregation to forecast sales at, we can consider two questions: at which level of aggregation can we get the most accurate forecasts, and at which level does the retailer actually require sales forecasts?

Forecasting sales of a group of SKUs, or of a single SKU over a group of stores, is normally an easier task than forecasting sales of a single SKU-store combination, due to the stability and increased signal to noise ratio of the more aggregate series. In practice, however, forecasts of each SKU-store combination are often those desired by the retailer, in order to aid inventory replenishment decisions. Disaggregation of predictions made at higher levels is one option to achieve these, but requires a choice of weights be made, which can frequently introduce significant additional error.

In this research, we investigate alternative methods which inherit information from higher aggregation levels to produce SKU-store level forecasts. Exploring the family of promotional sales models frequently used in retail forecasting, we look to utilise relationships between the theoretical model forms at the different levels. The goal is to find quantities that can be transferred easily between levels without the difficulties involved in disaggregation. Based on this exploration, we propose an approach which is then tested against other approaches on real world retail sales data. Our particular focus is on the resulting forecast accuracy, rather than the accuracy of our parameter estimates.

Integrated Business Planning (IBP) Drives Decision Making Across the Extended Business

Elton Brown
Logility
Australia

Tim Williamson
Logility
Australia

As more advanced and detailed forecasting techniques are available, in theory companies should be experiencing greater levels of improved accuracy and reduced bias. However, in practise, this is not happening. Why not? Achieving best-in-class results is dependent on getting the right mix of powerful statistical techniques combined with practical, business centric capabilities to review, analyse and amend the forecasts. Companies that are able to get this mix right are seeing tangible and sizable benefits. Elements to consider are: - Do the statistics make sense from a practical point of view? - Do you have a truly collaborative and effective forecasting process? - Are you managing the forecasts effectively? - Forecast error is guaranteed – how do you protect yourself?

Interval-valued Time Series Forecasting using A Novel Combination Method

Tao Xiong
College of Economics and Management, Huazhong Agricultural University
China

Yukun Bao
School of Management, Huazhong University of Science and Technology
China

Accurate forecasts of interval-valued time series over future horizons is challenging and of great interests in forecasting community, by providing a range of value rather than a point estimate. Traditional interval forecasting methods, i.e., vector autoregressive and vector error correction model, provide good forecasts only when interval-valued time series under study are linear and stationary. In addition, neural network and multi-output support vector regression have shown excellent nonlinear modeling capability for interval-valued time series. However, it is difficult in practice to determine when an interval-valued time series is generated by either a linear or nonlinear process, and thus which particular technique is more suitable than another for a specific interval-valued time series forecasting task. Following the well-established "linear and nonlinear" modeling framework in point-valued forecasting practices, to this end, this study extends it to forecast interval-valued time series with vector error correction model (VECM) and multi-output support vector regression (MSVR) (named as VECM-MSVR), which is capable of capturing the linear and nonlinear patterns exhibited in interval-valued time series abstracted from finance markets. Two interval-valued agricultural commodity futures prices series are constructed from Chinese futures market are used to justify the performance of the proposed VECM-MSVR method against selected competitors. The quantitative and comprehensive assessments are performed and the results indicate that the proposed VECM-MSVR method is a promising alternative for forecasting interval-valued time series with high volatility.

Inversion Copulas from Nonlinear State Space Models

Michael Smith
University of Melbourne
Australia

Worapree Maneesoonthorn
University of Melbourne
Australia

While copulas constructed from inverting latent elliptical, or skew-elliptical, distributions are popular, they can be inadequate models of serial dependence in time series. As an alternative, we propose an approach to construct copulas from the inversion of latent nonlinear state space models. This allows for new time series copula models that have the same serial dependence structure as a state space model, yet have an arbitrary marginal distribution-- something that is difficult to achieve using other time series models. We examine the time series properties of the copula models, outline measures of serial dependence, and show how to use likelihood-based methods to estimate the models. To illustrate the breadth of new copulas that can be constructed using our approach, we consider three example latent state space models: a stochastic volatility model with an unobserved component, a Markov-switching autoregression, and a Gaussian linear unobserved component model. We use all three inversion copulas to model and forecast quarterly U.S. inflation data. We show how combining the serial dependence structure of the state space models, with flexible asymmetric and heavy-tailed margins, improves the accuracy of the fit and density forecasts in every case.

Jointly Testing Multi-Step System Forecasts

Andrew Martinez
University of Oxford
United Kingdom

This paper develops a new approach for testing differences in multi-step system forecasts. Building on Diebold and Mariano (1995), it compares the performance of several tests and loss functions when jointly testing multi-step system forecast accuracy. In doing so, it extends the classic Diebold Mariano test to jointly compare multi-step system forecasts using the generalized forecast-error second-moment (GFESM). Simulations illustrate that GFESM based tests are well sized and are best able to detect alternative types of misspecification. Modifications to calculations improve performance when there are relatively few forecast-error observations. The tests are then used to assess the performance of the Federal Reserve's Greenbook forecasts against the Survey of Professional Forecasters.

Leverage, asymmetry and heavy tails in high-dimensional factor stochastic volatility models

Mengheng Li
VU University Amsterdam
Netherlands

Marcel Scharth
The University of Sydney Busines School

Australia

It is known that time-varying volatility and leverage effects are often observed in financial time series which is believed to be asymmetrically distributed with heavy tails. The rich literature studying various forms of univariate stochastic volatility model tends to confirm such empirical findings. Yet the literature focusing on high dimensional stochastic volatility models lacks a corresponding general modelling framework and efficient estimation method due to curse of dimensionality. The aim is to propose a flexible factor stochastic volatility model with leverage based on generalized hyperbolic skew Student's t-error to model asymmetry and heavy tails. With shrinkage, the model leads to different parsimonious forms, and thus is able to disengage leverage effects and skewness in idiosyncratic noise from those in factors. A highly efficient Markov chain Monte Carlo estimation procedure which uses efficient importance sampling to exploit the Gaussian mixture representation of the error distribution is proposed to analyze the univariate version of the model. Multivariate extension is achieved with marginalization of factors and boils down to many univariate series which can be estimated in parallel. We assess the performance of our proposed method via a Monte Carlo study with both univariate and multivariate simulated data. Finally, we apply our model to a equally weighted portfolio consisting of stocks from the S&P100.

Long Term Forecasted Marketing Optimization

John Sears
Uber
United States

Venkatesh Periaswamy
Uber
United States

Prakhar Mehrotra
Uber
United States

Uber spends millions of dollars a year on paid marketing campaigns to attract new riders and drivers to the platform. Investment of this scale is planned months in advance, using forecasts of advertising channel efficacy. Given thousands of forecasted investment channels and a several month planning cycle, it is impossible for human decision-makers to set a budget that is coherent, well-informed by historical data and strategically-aligned. This paper outlines the approach we have taken using convex optimization with regularization to allocate across thousands of forecasting investment channels, taking into account both operational and strategic constraints. We address concerns around multiple comparisons, model risk, performance requirements, varying regional strategies and human intervention.

Long Term Forecasting for Dynamic Businesses

Neha Gupta
Uber
United States

Radhika Anand
Uber
United States

Prakhar Mehrotra
Uber

United States

In this talk, we would like to discuss the problem of long term forecasting of financial metrics at Uber which include forecasting trips, up to 12 months ahead, for over 450 cities around the world impacting millions of riders and drivers. The main challenges of this type of forecast are dynamic nature of the business, impact of exogenous variables, exponential growth leading to sudden changes, impact of weather and various events, forecasting multiple cities and regions and forecast of young cities and regions with very less or no past data. The main contributions of our research are 1) We performed a thorough empirical analysis of various forecasting techniques including - ARIMA, ARIMAX, TBATS and Bayesian Structural Time Series to find the model that fits a particular time series the best. 2) We implemented times series models with exogenous variables to understand the impact of external levers such as spend, weather, events, etc. on our trips. Spurious regressions and confounding variables were a few interesting issues that came up during the process. 3) We have developed an automated machine to pick the best model amongst various time series models and their ensembles, based on out of sample performance and cross validation, for future predictions. The key here is to generalize well out of sample, without overfitting. 4) We studied how the forecast of new cities with less or no data can be made more informative by utilizing data from other similar, big cities. We studied this in a Bayesian setting where we used posteriors of big cities to act as informative priors for younger cities. 5) Lastly, we'll discuss ongoing work on our bottoms-up approach to study forecasting, by modeling it as a big data regression problem. We use past behavior of the rider/driver such as the past number of trips, average wait times to come up with an estimate of his/her future trips. This algorithm helps us understand, at the lower level, the factors affecting trips and provides interesting insights into our user behaviour. This automated forecast machine can be used as a black box to forecast any business metrics. Our research shows promising results in terms of accuracy, with significant improvements in out of sample MAPE for trips forecast, compared to the baseline algorithms already used in production. Furthermore, we have achieved more than 50% accuracy improvements in younger cities.

Long-Term Projections using Extreme Value Analysis

Ann Maharaj
Monash University
Australia

Pierpaolo D'Urso
Sapienza - Universitá di Roma
Italy

Andres Alonso
Universidad Carlos III de Madrid
Spain

In this study, we explore long-term projections of time series with similar seasonal patterns using extreme value analysis with fuzzy clustering. Input features into the fuzzy clustering methods are parameter estimates of time varying location, scale and shape obtained from fitting the generalised extreme value (GEV) distribution to annual maxima or to the r-largest order statistics per year of the time series. An innovative contribution of the study is the development of new generalised fuzzy clustering procedures taking into account weights, and the derivation of iterative solutions based on the GEV parameter estimator. Simulation studies conducted to evaluate the methods, reveal good performance. An application is made to a set of daily sea-level time series from around the coast of Australia where the identified clusters are well validated and they can be meaningfully interpreted, before long-term projections are made.

Looking forward, back and sideways: the many directions of Bayesian forecasting

Kerrie Mengersen
Queensland University of Technology
Australia

Bayesian models are natural vehicles for forecasting and have a long, rich foothold in finance. Indeed, many of the historical and current advances in Bayesian statistics have arisen from this work. However, there are many other areas in which forecasting is required, including health and environmental applications. Moreover, there is an increasing range of models that are being applied to forecasting problems.

In this presentation, we review the role that Bayesian modelling has played, and is playing, in forecasting across a range of fields, and give some examples of our experiences in the practical application of these models. This work is being undertaken as part of the Australian Research Council Centre of Excellence in Mathematical and Statistical Frontiers: Big Data, Big Models and New Insights.

Making Lemonade Out of Lemons: Model Stacking for Forecasting

Konstantin Golyaev
Microsoft
United States

Jocelyn Barker
Microsoft
United States

Gagan Bansal
Microsoft
United States

The “No Free Lunch Theorem” of Wolpert and Macready (1997) formalizes a consistent finding that no single forecasting method consistently exhibits superior performance on a diverse set of problems. The most common approach to this problem is to try out several methods and select the best performer. Here we present model stacking, a powerful technique which combines forecasts from multiple methods into a single prediction, often outperforming each of the individual methods.

Model stacking is a two-stage process. In the first stage, time series cross-validation is used to obtain out-of-sample forecasts for several models. Since cross-validation is a required step for machine learning forecasts, this adds very little overhead to the modeling process. In the second stage, these out-of-sample forecasts serve as features in a linear regression model that yields the final ensemble forecast. While combining forecasts from multiple methods is a well-established technique, we are not aware of any forecasting applications that involve model stacking.

We evaluated our method using quarterly revenue for Microsoft’s consumer hardware devices across fifteen worldwide markets. We forecast the data using five different models three univariate time series models and two machine learning regression models (Elastic Net and K-Nearest Neighbors). All models were trained at weekly granularity and then forecasts for the next 13 weeks were summed to yield quarterly forecasts. We evaluated model performance by computing mean absolute percentage errors (MAPEs) over four quarters of 2016 and

found that no individual method performed well, with a mean standard deviation of 20.13% in MAPEs for Device 1 and 42.31% for Device 2, across the markets that we modeled.

The most common method for forecast aggregation, simple averaging of individual forecasts, did not produce any material improvement in errors. In contrast, the forecasts from the stacked models resulted in dramatically lower MAPEs: we found a 63.33% improvement in forecast accuracy for Device 1 and 76.42% improvement for Device 2, relative to the simple averaging method, as measured by the relative mean absolute forecast error. Model stacking's demonstrated ability to yield considerable improvements in accuracy with minimal incremental modeling overhead has made it a powerful tool for financial forecasting at Microsoft.

Maximising the likelihood function in a Markov switching model via modern optimisation techniques

Ali Eshragh
The University of Newcastle
Australia

In 1989, Hamilton developed a novel state-space model for a time series governed by the states of a system that may vary during the time. Hamilton used a Markov switching model to capture the impact of states on the time series, in particular, US GNP. One of the main contributions of that work was an iterative algorithm to derive the likelihood function to estimate the unknown parameters as well as the probability of occupying the states. All these can be achieved by maximising the likelihood function. However, as the likelihood function is a complicated nonlinear function in the Hamilton's model, current optimisation algorithms used in Markov switching models, both in theory and practice, may only be able to find a local optimal solution which can be far away from a global optimal solution. This implies that the maximum likelihood estimates highly depend on the quality of optimal solutions calculated via optimisation algorithms. We develop a new algorithm to maximise the likelihood function in the Hamilton's model by exploiting modern optimisation techniques. In this talk, the new optimisation algorithm is explained and numerical results are reported.

Maximum Entropy Extreme-Value Seasonal Adjustment

Richard Penny
Statistics New Zealand
New Zealand

Tucker McElroy
U.S. Census Bureau
United States

Some economic series in small economies exhibit meager (i.e., negligible non-positive) values, as well as seasonal extremes. For example, agricultural variables in countries with a distinct growing season may exhibit both of these features. Multiplicative seasonal adjustment typically utilizes a logarithmic transformation, but the meager values make this impossible, while the extremes engender huge distortions that render seasonal adjustments unacceptable. To account for these features we propose a new method of extreme-value adjustment based on the maximum entropy principle, which results in replacement of the meager values and extremes by optimal projections that utilize information from the available time series dynamics. This facilitates multiplicative seasonal adjustment. The method is illustrated on New Zealand agricultural series.

Measuring the productivity and the efficiency of the New Mexico City International Airport

Maria Rosa Nieto
Universidad Anahuac Mexico
Mexico

Rafael Bernardo Carmona-Benitez
Universidad Anahuac Mexico
Mexico

The contributions of this study are to assess the productivity and the efficiency of the New Mexico City International Airport (NAICM) againts the current Mexico City International Airport (AICM), London Heathrow Airport (HTW), Amsterdam Schiphol Airport (AMS), and the application of the ARIMA+GARCH+Bootstrap to forecast future efficiencies and productivities. We apply the Data Envelopment Analysis (DEA) methodology to compare these airports in terms of number of passenger transported and capacity infraestructure. We emphasize the analysis between the NAICM with the AICM. We contrast the actual and future situations by applying the ARIMA+GARCH+Bootstrap forecasting method. Additionally, we use a bootstrap correction for eliminating the bias in the efficiency factor. The results indicate, on one hand, that the worst efficiency factor through time is the one obtained for the AICM due to the increase in the number of passengers that has exceeded its capacity. Hence, the NAICM investment is validated. On the other hand, the results show that the NAICM is not going to achieve similar efficiencies and productivities as AMS and HTW. It can be due to the difference between the planned airport NAICM capacity infraestructura and the passenger demand forecasted. Apparently, the infraestructure is not going to be enough to attend the expected number of passengers.

Measuring the Effect of Regional Volatility Spillover to the Specific Market: An Evidence of Emerging Stock Markets

Tsung-Han Ke
National Chi Nan University
Taiwan

Henry Leung
Sydney University
Australia

There is a widespread belief that understanding the behavior of volatility spillover across stock markets is important for measuring and estimating the risk of internationally diversified portfolio. The volatility transmission can be of enormous influence to global stock markets, therefore, a more rigorous exam of this issue could be performed. As literature noted that there has been a dramatic proliferation of research concerned with the volatility spillovers from developed markets to emerging markets. However, there is a question worth noting, and more extensive research of this issue would be necessary to undertake. The primary research question to be addressed in this paper is whether there exist volatility spillovers from emerging markets to developed markets or specific markets. Based on the bivariate GARCH models, the purpose of this study is measuring the effect of regional volatility spillover to the specific market by adopting a new efficient method to construct a volatility spillover indicator from eight Asian emerging stock markets, including Hong Kong, Singapore, Indonesia, Korea, Malaysia, Philippines, Taiwan and Thailand. The specific stock markets are divided into three parts, which are four euro-zone countries (Netherlands, Austria, France, and Germany), three North American Free

Trade Agreement (NAFTA) countries (U.S., Canada, and Mexico), and one Asian country (Japan). The main goals of this study are measuring the effect of regional volatility spillover, and using the volatility spillover indicator to provide useful information for improving the volatility forecasting performance of stock market index in specific countries.

Monitoring forecasting performance

Allan Timmermann
University of California, San Diego
United States

Yinchu Zhu
University of California, San Diego
United States

We establish conditions under which instruments ("testers") can be used to monitor the performance of competing forecasts and address the scope for improving predictive accuracy through a conditional decision rule that, at each point in time, chooses the forecast with the smallest expected loss given the information contained in the testers. We show that testers can lead to improved forecasting performance in situations with weak predictors due to estimation error. Testers can also help track and improve conditional forecasting performance in situations with time-varying parameters, model misspecification, or serially correlated forecast errors. Through Monte Carlo simulations and empirical applications to predictability of stock market returns and inflation we demonstrate the gains from monitoring forecasting performance and choosing among forecasting models using conditional information on the models' relative expected loss.

Multi-horizon Forecast Evaluation of Electricity Demand in Australia

Mardi Dungey
University of Tasmania
Australia

Thomas Goodwin
University of Tasmania
Australia

Jing Tian
University of Tasmania
Australia

Australia's National Electricity Market (NEM) is characterized by some of the most volatile wholesale spot prices in the world. The highly volatile prices, together with insufficient infrastructure and environmental targets, are challenging Australian electricity sector to deliver affordable and reliable electricity. This paper is a part of the project that examines the contribution of the microstructure of the NEM to the extremely high price volatility. Electricity demand forecasts are provided by the AEMO to electricity wholesale suppliers in advance, and they are revised half hourly prior to dispatch orders. Electricity generators take these information and bid their supply schedule in the next 48 hours. A unique feature in the NEM is that the generators are allowed to revise their bids up to 5 minutes before dispatch, and it is observed that rebidding occurs fairly frequently. Why do the generators rebid? Are the rebids related to demand forecast revisions? If electricity demand forecasts are revised due to adapting newly available information (such as a sudden change in weather conditions), it is then rational for generators to adjust their schedule according to the revised demand forecasts and rebid quantity supply.

In this paper, we employ a newly developed unobserved component modelling framework to analyze the revision structure of electricity demand forecasts for the New South Wales region of the NEM over the period of July 2011 to July 2015. As far as we know, this is the first study on the performance of multi-horizon electricity demand forecasts for the purpose of price determination. The evaluation framework characterizes multi-horizon demand forecasts errors to three categories: 1) bias that captures systematic under- or over-forecasts; 2) rational forecasting errors that occur due to unanticipated information; and 3) implicit forecasting errors that are irrelevant to the realized electricity demand. Therefore, forecast revisions in general are made for 1) correcting bias made at longer horizons; 2) incorporating newly available information; and 3) adjusting for implicit forecasting errors. We find that AEMO's electricity demand forecasts consists of all three types of forecasting errors. Significant bias is seen in the forecasts that are made for the trading intervals in the late night and morning. The marginal information accumulation in forecast revisions depends on the pre-dispatch time rather than the length of forecast horizons. There are two peaks of information adaption in forecast revisions occur at around 7am and 9am regardless of forecast horizons.

Multi-Period Portfolio Selection with Drawdown Control

Peter Nystrup
Technical University of Denmark
Denmark

Stephen Boyd
Stanford University
United States

Erik Lindström
Lund University
Sweden

Henrik Madsen
Technical University of Denmark
Denmark

In this talk, model predictive control (MPC) is used to dynamically optimize an investment portfolio. The predictive control is based on multi-period forecasts of the mean and covariance of financial returns from a multivariate hidden Markov model with time-varying parameters. Estimation and forecasting are done using an online expectation--maximization algorithm. There are computational advantages to using MPC when estimates of future returns are updated every time new observations become available, since the optimal control actions are reconsidered anyway. Transaction and holding costs are important and are discussed as a means to address estimation error and regularize the optimization problem. A complete practical implementation is presented based on available market indices chosen to mimic the major liquid asset classes typically considered by an institutional investor. In an out-of-sample test spanning two decades, the proposed approach to multi-period portfolio selection successfully controls drawdowns with little or no sacrifice of mean--variance efficiency.

Multi-Step Ahead Forecasting and Signal Extraction of High Frequency Vector Time Series

Tucker McElroy
U.S. Census Bureau
United States

Recursive algorithms, based upon the nested structure of Toeplitz covariance matrices arising from stationary processes, are presented for the efficient computation of multi-step ahead forecasts (with mean squared error) and signal extractions. The methods are developed for nonstationary vector time series, and are applied to high frequency retail and immigration data exhibiting trend dynamics and multiple types of seasonality.

Natural Language Processing for Predicting Everyday Behavior with and without Time and Duration Information

Rafal Rzepka
Hokkaido University
Japan

Kenji Araki
Hokkaido University
Japan

Forecasting is a complicated process which is impossible without sufficient and reliable data. Researchers and data specialists have been predicting sales, epidemics, stocks, elections and many other possible outcomes which could be useful to know beforehand. People do not need to forecast phenomena that are characterized by easily predictable consequences but artificial agents are bad at forecasting very obvious human's everyday behavior. This is because data on everyday life is rarely available, being meant for commercial use or due to privacy issues, which makes machine learning difficult or very narrow. If such data is collected and shared, usually only a few specific aspects of behavior are recorded together with time stamps. In our approach we combine text mining and natural language processing for discovering order of acts with sentiment analysis to predict usual outcomes of human's deeds. In our presentation we introduce results of our experiments (Rzepka 2015) with predicting consequences of ethical and unethical actions from Japanese blogs (5,5 billion word raw text corpus, almost 80% agreement with human subjects). Then we report results of experiments with Twitter data and discuss a possibility of natural language processing techniques to automatically recognize and represent time duration in machine-readable format (Krawczyk 2013), which could be a helpful addition for many types of forecasting relying only on time-stamps. By presenting our findings in predicting judicial verdicts, we also share our suggestions on how text-similarity algorithms could be utilized by forecasting community. Then we compare our approach with a method for extracting future reference sentences using semantic role labeling (Nakajima 2016). We conclude with a discussion on how forecasting methods can be useful for common sense knowledge acquisition and how machines could become able to automatically discover trends worth forecasting, for example by machine reading of scientific papers.

(Krawczyk 2013) Marek Krawczyk, Yuki Urabe, Rafal Rzepka and Kenji Araki, "A-dur: Action Duration Calculation System", SIG-LSE-B301-7, pp.47-54, 2013. (Rzepka 2015) Rafal Rzepka, Marek Krawczyk and Kenji Araki, "Replacing Sensors with Text Occurrences for Commonsense Knowledge Acquisition”, Proceedings of IJCAI 2015 Workshop on Cognitive Knowledge Acquisition and Applications (Cognitum 2015), Buenos Aires, Argentina, July 25-31, 2015. (Nakajima 2016) Yoko Nakajima, Michal Ptaszynski, Hirotoshi Honma, Fumito Masui, "Automatic extraction of future references from news using morphosemantic patterns with application to future trend prediction", AI Matters 2(4): 13-15, 2016.

Nowcasting with large, international data sets: do sparse priors help?

Philipp Hauber
Kiel Institute for the World Economy
Germany

Christian Schumacher
Deutsche Bundesbank
Germany

Factor models can summarize the co-movements of a large number of variables and have proven useful in nowcasting and short-term forecasting of GDP growth. Factor models for nowcasting can tackle missing observations at the end of the sample, because business cycle indicators are subject to publication lags leading to di_erence in availability in real time. Typically, a large number of monthly business cycle indicators is available for nowcasting, and the models have to solve a mixed-frequency problem when nowcasting quarterly GDP growth. The main aim of the paper is to assess the importance of international variables for nowcasting national developments, an issue, which curiously has received relatively little attention in the academic literature so far. Given the large number of variables at both the national and international level, the question arises whether all this information is useful for nowcasting or not. As such, we also contribute to the continuing debate on variable selection and the optimal size of factor models for forecasting. Rather than choosing variables ad-hoc, we employ sparse priors on the factor model's loadings for Bayesian estimation. Sparse priors can help to identify those business cycle indicators that essentially determine the factors, whereas irrelevant variables are sorted out. In an empirical exercise, we evaluate nowcasts of GDP for the Euro area and the United States from models that use only the respective national data as well as the combined data sets comprising of roughly 150 variables.

On Bootstrapping Tests of Equal Forecast Accuracy

Firmin Doko-Tchatoka
School of Economics, The University of Adelaide
Australia

Qazi Haque
School of Economics, The University of Adelaide
Australia

This paper proposes a bootstrap procedure for tests of equal accuracy of out-of-sample forecasts in nested models, where errors may be arbitrary heteroscedastic and/or autocorrelated. The recursive F-statistic often used to assess this problem has a nonstandard asymptotic distribution and the computation of its critical values is burdensome. Recently, Hansen and Timmermann (2015) show that this statistic is equivalent to a Wald statistic whose asymptotic distribution is a convolution of possibly dependent $\chi^2(1)$-distributed random variables. We provide simulation evidence that under heteroscedastic and/or autocorrelated errors, both the recursive F-test and its Wald approximation do not have desirable size properties even with moderate sample sizes, particularly for multistep forecasts. We thus propose a moving block bootstrap (MBB) procedure that applies to both tests, whether the errors are heteroscedastic and/or autocorrelated or not. We establish uniform consistency of the MBB under the null hypothesis (size) and the alternative hypothesis (power). We show through Monte Carlo experiments that the MBB has substantially enhanced the small-sample properties of the tests. Finally, we compare the performance of our proposed bootstrap to other bootstrap procedures in the literature.

On Convexity of Weights in Optimal Forecast Combination

Anupama Lakshmanan
Indian Institute of Management Bangalore
India

Shubhabrata Das
Indian Institute of Management Bangalore
India

Forecast combining is a well-known method of improving forecast accuracy. The best linear combination in terms of minimising the mean squared error (MSE) may be selected from a few good forecasts. If this optimal combination entails negative weights for some forecasts, it may not be very appealing from the perspective of implementation and practice. In particular, while combining two forecasts, the optimal weights turn out to be positive when the two forecasts are negatively correlated. However, if they are positively correlated, the optimal weights are positive only if the correlation is less than a threshold value. In this work, we obtain a necessary and sufficient condition for the optimum linear combination of an arbitrary number of good forecasts to be a convex combination. We demonstrate this with forecast combination in multiple data sets.

Typically only a few (two to five) good forecasts are chosen to be combined; hence we attempt to arrive at a simpler set of sufficient conditions on correlation/covariance between such smaller number of forecasts, from which the best convex forecast combination is to be obtained.

If each of the original short-listed good forecasts is unbiased, then it is necessary for the weights to add up to one in order to ensure that the forecast combination yields an unbiased forecast. In that situation, the convexity of the linear combination is synonymous with the above mentioned non-negativity of the coefficients. We also explore a modified form of this when L^2-norm, as opposed to L^1-norm, of the coefficients is considered, leading to a reverse principal component like problem, as the objective is to minimise rather than maximise the variance of the linear combination.

One for all: forecasting intermittent and non-intermittent demand using one model

Ivan Svetunkov
Lancaster Centre for Forecasting
United Kingdom

John E. Boylan
Lancaster Centre for Forecasting
United Kingdom

Forecasting demand is one of the typical tasks in supply chain. However demand is usually not homogeneous and can exhibit different patterns and characteristics, which in practice implies usage of different models for different types of data. While non-intermittent demand can be efficiently predicted using conventional exponential smoothing (ETS), intermittent demand has been considered as a separate problem, needing separate set of forecasting methods. In this presentation we propose a statistical model, connecting intermittent and non-intermittent types of demand. We then demonstrate how the type of data can be detected automatically using information criteria. Finally, we show the efficiency of the proposed model and selection procedure using an example of real data.

Optimal retail cluster based Hierachical forecasting

Christoph Bergmeir
Monash University
Australia

Floris Padt
anonymous (CEO doesn't allow to use the name)
Netherlands

This presentation is about a global leader in optical retail delivering frames and sunglasses, amongst others, through a worldwide presence of dozens of retail banners. Tens of thousands of articles are sold in thousands of stores, resulting in millions of sparse time series.

Insight in future demand over the lead time is required to balance supply and demand in order to reduce lost sales and stock depreciation. Short product life cycles, slow – fashionable and seasonal sales in a big data environment imposes some additional challenges

We present a case study of a large-scale sales forecasting project. The data are daily sales across around 30,000 products in 1,000 stores of the company. We apply hierarchical time series forecasting with forecast reconciliation, and clustering of both the time series and the master data. Practical insights and lessons learned in this project will be presented, using R in a shiny business application.

Pareto analysis of judgmental adjustments in statistical forecasts

Fotios Petropoulos
University of Bath
United Kingdom

Paul Goodwin
University of Bath
United Kingdom

Robert Fildes
Lancaster University
United Kingdom

Brian Squire
University of Bath
United Kingdom

Demand planners frequently perform judgmental adjustments to the statistical output of forecasting support systems. The efficiency of such interventions has been studied in the literature, primarily with regards to their size and sign. Also, it is widely recommended that if judgment is to be applied in any form towards improving the forecasting process, this should be done for the most important (“A” items within a standard ABC classification) and least forecastable items (“Z” items of a XYZ classification). In this study, we try to map the value of judgmental adjustments to statistical forecasts with regards to these two dimensions: importance and forecastability. To this end, we distinguish between “optimal” forecastability, which assumes an optimal statistical forecast, and “experienced” forecastability, which is what the demand planners face due to potentially outdated forecast settings. Our empirical analysis involves real demand data, together with statistical and

judgmental forecasts obtained from a multibillion North American company in the consumer packaged goods industry. Based on our findings, we extend the discussion of the recent literature with regards to the effective design of forecasting support systems.

Periodically Correlated Model for Electricity Load Forecasting

Jesus Juan
Universidad Politecnica de Madrid
Spain

Eduardo Caro
Universidad Politecnica de Madrid
Spain

F. Javier Cara
Universidad Politecnica de Madrid
Spain

This work describes a time series model used by the Spanish System Operator (Red Eléctrica de España, REE) to make hourly forecasts of electricity demand from one to ten days ahead. This article presents a new approach for estimating multiequation models for hourly load forecasting that extends previous work in diferent important ways. In our version, the 24 hourly equations are assembled to form a periodic autoregressive moving-average model. This adjustment not only improves the short term predictions for the coming hour but, its effect is noticeable beyond 24 hours, horizon of enormous practical interest in the programming of an electrical system.

The model considers the hourly series as a periodically correlated process. It incorporates temperature as an explanatory variable and takes into account the effect on the demand of holidays. Direct estimation of the model by maximum likelihood is very complicated because it contains a large number of parameters. The problem is solved with a two-step estimate: first, an independent model is obtained for each hour (this is the usual strategy in most of the models proposed in the literature) and, secondly, a joint estimate of the parameters that relate each hour to the previous hours.

In this work it has been applied the methodology of regression splines to the problem of estimating the functional relationship between weather and electricity demand. The nonlinear relationship is observed graphically. The results agree with intuitive expectations, and the graphs clearly show how temperature influences demand changes throughout the day. The dynamic effect of the temperature on demand is incorporated by simply adding lagged temperature variables. The choice of the number of nodes is performed by cross-validation;in this implementation 3 nodes are used. These have been equispaced. It has been proven experimentally that the overall behavior of the method is quite robust.

Population Estimation Using Satellite Imagery

Aikaterini Koutsouri
National Technical University of Athens
Greece

Ilectra Skepetari
National Technical University of Athens
Greece

Konstantinos Anastasakis
National Technical University of Athens
Greece

Stefanos Lappas
National Technical University of Athens
Greece

Population data provides statistical information that can in turn support decision making processes. It provides essential information regarding many practices such as rescue operations or humanitarian actions, which require the estimation of local population. While traditional approaches are possible, they tend to be time consuming and expensive. Copernicus Earth Observation data (Sentinel 2 satellite images) provides high resolution satellite imagery that could be a useful alternative to collecting census data since it is significantly cheaper than traditional methods, although less accurate. This paper describes a methodology for exporting a population estimation using satellite imagery based on the use of classification techniques coupled with a statistical forecast on historical data.

Predict NBA 2016-2017 MVP Winner

Mason Chen
Mason Chen Consulting
United States

This project is to build a statistical model to predict who will win 2017 NBA Most Valuable Player (MVP) Award in the regular season. Team has collected three raw data from public Sports domain: (1) player statistics, (2) team win%, and (3) historical MVP winners. Before building the proto model, the player statistics have been standardized to the Z scale in each player statistics category in order to remove any mean and standard deviation effect. This Z transformation can eliminate any statistics bias or domination from any particular category. The Z scale will also analyze each player's performance as compared to the other top NBA players in the same year. The "MVP Index" has been derived from combining each player's Z statistics with equal weight as a "Uniform" model. To evaluate the model accuracy, team has derived another "Accuracy Index" of predicting the top five MVP players recognized annually. The first "Uniform" model can predict the top five winners at 47% accuracy. Team has further derived the "Weighted" model by adding the weight factor which was calculated based on the dispersion/separation between the top two MVP winners and the remaining players not in top 5. The weight factor will reflect which player statistics categories are more critically contributed to the MVP Award selection process. The "Weighted" model has further improved the Accuracy Index from 47% to 52%. To further optimize the prediction accuracy, authors have added the "Team Winning" factor. Most historical MVP winners were from the teams with best or better regular season records. Authors have assessed the team winning factor based on the "Power" model from power= 0 (equivalent to the Weighted Model), 1, 2, 3, 4, 5, 6 to power= infinity (MVP from the best Team). The MVP Index will be multiplied by the power of the team winning% in the Power model. Power transformation will improve the model accuracy but which may also create any over-fit concern when power level is getting higher. Based on the Power Model, team can improve the Accuracy Index to 70% at Power=3. There is little benefit but more over-fit risk to further increase the power level beyond 3. Team will use the Power=3 model to predict 2017 MVP on the first day of each month starting from Dec.1, 2016 until the season end in April, 2017. The final model prediction will be available around middle April, 2017.

Predicting Ordinary and Severe Recessions with a Three-State Markov-Switching Dynamic Factor

Model. An Application to the German Business Cycle

Kai Carstensen
University of Kiel, Ifo Institute, CESifo
Germany

Markus Heinrich
University of Kiel
Germany

Reif Magnus
Ifo Institute
Germany

We estimate a Markow-switching dynamic factor model with three states based on six leading business cycle indicators for Germany preselected from a broader set using the Elastic Net soft-thresholding rule. The three states represent expansions, normal recessions and severe recessions. We show that a two-state model is not sensitive enough to reliably detect relatively mild recessions when the Great Recession of 2008/2009 is included in the sample. Adding a third state helps to clearly distinguish normal and severe recessions, so that the model identifes reliably all business cycle turning points in our sample. In a real-time exercise the model detects recessions timely. Combining the estimated factor and the recession probabilities with a simple GDP forecasting model yields an accurate nowcast for the steepest decline in GDP in 2009Q1 and a correct prediction of the timing of the Great Recession and its recovery one quarter in advance.

Predicting the ‘great’ recession and the ‘poor’ recovery – results from a suite of non-linear models for Romania and the euro area

Anca-Adriana Gălăţescu
Banca Națională a României
Romania

Vincent Labhard
European Central Bank
Germany

Short-term forecasting of the main macroeconomic aggregates usually relies on statistical models which can exploit timely information and can deal with mixed frequencies and ragged edges. And, in most cases, short-term forecasting relies on a number of models and some kind of average across models, an approach often referred to as a suite-of-models approach. In the vast majority of cases, such model suites are linear. Non-linear suites are much less widespread, and often cast as a horse race with standard linear models as benchmarks. This paper assesses the predictive performance of a suite of non-linear single-indicator models (SIMs) in forecasting GDP and the main expenditure components for Romania and the euro area around the trough of the Great’ recession and during the subsequent recovery. This is done because those episodes have proven very challenging for forecasters, because the large downswings and upswings may have been better detected with non-linear models, and because there are few model suites, let alone non-linear ones, for both GDP and its main components. In the paper we provide and evaluate both point and density forecasts, based on two types of models from the nonlinear class of models, namely smooth transition autoregressive (STAR) models and Markov-switching (MS) models.

We forecast GDP, consumption, gross fixed capital formation, imports and exports, on the basis of 18 indicators, including industrial production (7 indicators), surveys (5), monetary (3) and foreign (3) indicators. We show that (i) in the case of both the single country and the country aggregate, depending on the indicator, non-linear models may have provided more timely signals of the major economic downturn ahead, thus given extra lead time to forecasters in predicting developments in GDP and the expenditure components, (ii) the nonlinear suite would have provided more accurate signals than the simple AR(1) benchmark and (iii) the Markov-switching models would have performed better than the smooth transition autoregressive models. While the evidence is encouraging as regards the regular use of non-linear model suites in forecasting, several issues that could strengthen the case have not been covered in this paper and are left for future work. This includes a more comprehensive evaluation also for other countries or country aggregates, illustration of the optimal integration and optimal weighting (over time and across countries) of model suites with linear and non-linear models, and practical ways to overcome the challenges of identifying optimal non-linear specifications and estimating them in real time.

Predicting the Proportion of a Population with Influenza-Like Illness Symptoms

Benjamin Moran
University of Newcastle
Australia

In this paper, we develop a statistical forecasting model for quantitative epidemiological data obtained from Flutracking, an online health surveillance system to detect epidemics of influenza. As the Flutracking time series data is gathered only for a certain period of time each year, there may be different number of lags in different years. This brings a high volatility in the time series data set and consequently a Seasonal Autoregressive Moving Average Model (SARIMA) model is not well-fitted. For this, we first refine the data to remove the trend in the variance and then fit an appropriate SARMIA model to the . The promising results show that our model performs well when compared to actual data by assessing the relative error.

Lastly, we make some suggestions as to how we can incorporate other time series into our model to improve its predictive performance and increase its utility when modelling epidemics of influenza, like including meteorological or ecological data to account for regional factors affecting epidemics, or by including similar data sets from influenza seasons in other regions.

Prediction Regions for Interval-valued Time Series

Gloria Gonzalez-Rivera
University of California, Riverside
United States

Yun Luo
University of California, Riverside
United States

Esther Ruiz
Universidad Carlos III, Madrid
Spain

We approximate a probabilistic forecast for an interval-valued time series by constructing bivariate prediction regions of the center and range of the interval. We dispense with the positive constraint on the range by estimating a bivariate system of the center/log-range. However, a forecast based on this system needs to be

transformed to the original units of center/range. Furthermore, even if log-range is normally distributed, the bivariate density of the center and range will not be bivariate normal, rendering the construction of a bivariate density forecast more difficult. For point forecasts of the center/range system, we examine the out-of-sample performance of parametric methods such as a naïve transformed forecast and a parametric bias-corrected forecast as well as nonparametric corrections like a smearing correction and a bootstrap forecast. We construct prediction regions based on the parametric and nonparametric corrected point forecasts with conservative Bonferroni bounds. Monte Carlo simulations show that biased correction methods do not generate forecasts that are uniformly superior and that a naïve transformed forecast performs well under Mean-Absolute-Error and Mean-Distance-Error loss functions. Bootstrap forecasts provide larger coverage rates of the realized intervals. Bootstrap ellipsoids and bootstrap Bonferroni rectangles are the prediction regions delivering the right probabilistic coverage. We apply these methods to the daily low/high intervals of the SP500 index and Google prices. We find one-step-ahead interval forecasts covering, on average, more than 50% of the realized intervals.

Probabilistic Forecasting of Wave Height for Offshore Wind Turbine Maintenance

James Taylor
Saïd Business School, University of Oxford
United Kingdom

Jooyoung Jeon
School of Management, University of Bath
United Kingdom

Wind power continues to be the fastest growing source of renewable energy. This paper is concerned with the timing of offshore turbine maintenance for a turbine that is no longer functioning. Service vehicle access is limited by the weather, with wave height being the important factor in deciding whether access can be achieved safely. If the vehicle is mobilized, but the wave height then exceeds the safe limit, the journey is wasted. Conversely, if the vehicle is not mobilized, and the wave height then does not exceed the limit, the opportunity to repair the turbine has been wasted. Previous work has based the decision as to whether to mobilize a service vessel on point forecasts for wave height. In this paper, we incorporate probabilistic forecasting to enable rational decision making by the maintenance engineers, and to improve situational awareness regarding risk. We show that, in terms of minimizing expected cost, the decision as to whether or not to send the service vessel depends on the value of the probability of wave height falling below the safe limit. We produce forecasts of this probability using time series methods specifically designed for generating wave height density forecasts, including ARMA-GARCH models. We evaluate the methods in terms of statistical probability forecast accuracy we well as monetary impact, and we examine the sensitivity of the results to different values of the costs.

Probabilistic Forecasts in Hierarchical Time Series

Puwasala Gamakumara
Monash University
Australia

Rob J Hyndman
Monash University
Australia

George Athanasopoulos
Monash University
Australia

Anastasios Panagiotelis
Monash University
Australia

Forecasting hierarchical time series has been of great interest in recent years. However, the literature has mainly focused on obtaining point forecasts that are aggregate consistent across a hierarchy. In this paper, we instead focus on producing probabilistic forecasts for hierarchical time series. We first observe that the trace minimization (MinT) approach introduced by Wickramasuriya, Athanasopoulos and Hyndman (2015) produces aggregate consistent mean and variance forecasts. If we assume Gaussianity, these will uniquely determine Gaussian forecast densities that are aggregate consistent across the hierarchy. We then relax the Gaussian assumption and propose a novel, non-parametric approach. This entails simulating future sample paths of the whole hierarchy by incorporating bootstrapped in-sample errors and reconciling these so that they become aggregate consistent. We evaluate both approaches via extensive Monte Carlo simulations. Results favor MinT(Shrink) and the conventional Bottom-up approach for producing both Gaussian forecast densities as well as aggregate consistent future sample paths using the non-parametric approach.

References: Wickramasuriya, S. L., Athanasopoulos, G., and Hyndman, R. J. (2015). Forecasting hierarchical and grouped time series through trace minimization. Working Paper 15/16. Department of Econometrics and Business Statistics, Monash University, Australia.

Product life cycle analysis as a tool for better demand forecast

Agata Chorowska
ProLogistica Soft
Poland

In this study we focus on the commonly neglected matter of including product life cycle analysis into a demand forecasting process.

Not only the demand pattern and its dynamics change between stages of product life cycle but also forecast error consequences are different in each stage. For example, while some excess inventory caused by overstated forecasts can be acceptable in the growth stage, it is usually very harmful in the decline stage. Similarly, understated forecasts causing stock-outs are much more dangerous at the beginning of the product life cycle than at its end.

With the recently observed shortening of product life cycles' lengths the discussed topic is becoming crucial in many business areas.

In this research we study the application of the product life cycle analysis in a forecasting process itself as well as in forecast results interpretation and their application to support business decisions.

We show how demand forecasts accuracy can be improved by using product life cycle analysis. We propose different algorithm choice strategies as well as different error measures suited for each product life cycle stage. We study various business areas with both short and long product life cycles and present numerical results to prove effectiveness of the proposed methods.

We discuss the obtained results from the business perspective covering forecasts applications in the inventory optimization and demand planning areas. In the discussion we take into account business characteristics of different product life cycle stages.

Random autoregressive models. A structured overview

Marta Regis
Eindhoven University of Technology
Netherlands

Nazanin Nooraee
Eindhoven University of Technology
Netherlands

Edwin R. van den Heuvel
Eindhoven University of Technology
Netherlands

Huge streams of values are nowadays recorded simultaneously from devices, often continuously in time, and data from multiple users are collected in clouds, rapidly growing into big data. The challenge consists in extracting as much information as possible, enabling the monitoring of the current status and the forecasting of future values. If on the one hand many data-driven techniques are rapidly developing, offering a wide range of flexible methods for interpretation and prediction, on the other hand these methods should still be supported and complemented by model-based approaches. Random autoregressive models (RAM), characterized by both the autoregressive structure and the inclusion of some random coefficients, are powerful tools in the analysis of highly frequent and volatile time series. Two popular examples are the random coefficient autoregressive model of Nicholls and Quinn (1982) for financial applications and the random coefficient autoregressive model of Liu and Tiao (1980) for biological studies. Despite the same definition, the authors refer to similar though different models by imposing peculiar sampling of the autoregressive coefficients: the first aims at capturing volatility in time for a single time series, while the goal of the second is representing heterogeneity among units under the same modeling structure, in the fashion of mixed models. The available literature is extremely broad, but often sectorial, overlapping and confusing. Furthermore, most of the available models focus on addressing one aspect of the data at a time, for example modeling heterogeneity between units, while much could be gained by combining the strength of various models and their sources of heterogeneity. The aim of the current work is first to structure the existing literature, showing in which aspects the various branches relate and differ from each other. On the other hand, this overview fills the gap existing between two main model families and by doing so introduces a unifying superstructure. Although only the basic properties of this general model are explored in the present work, the proposed random autoregressive model will enable the generalization of the one model in the direction of the other, opening a new route for further explorations and development of heterogeneous time series models for complex data.

Random Forest Density Forecasts in Retail

Stephan Kolassa
SAP Switzerland
Switzerland

Random Forests (RFs) are a nonparametric algorithm combining classification and regression trees (CARTs), bootstrapping and random selection of splitting variable candidates. RFs are especially useful in causal classification or regression with "many" correlated predictors. They have performed very well across many problem types (Fernández-Delgado et al., JMLR, 2014). However, RFs have apparently not yet been investigated in forecasting retail sales, which are characterized by many promotions, calendar effects and other causal drivers. We will present various ways of calculating density forecasts using RFs, apply these to a large dataset of daily

retail sales, and benchmark them against known high-performing density forecasting algorithms (Kolassa, IJF, 2016) using proper scoring rules.

Random matrix theory analysis of cross-correlation in the Nigerian Stock Exchange

Chimezie Nnanwa
Sheffield Halam University
United Kingdom

Thomas C. Urama
Sheffield Halam University
United Kingdom

Patrick O. Ezepue
Sheffield Halam University
United Kingdom

In this paper we use Random Matrix Theory to analyze the eigen-structure of the empirical correlations of 82 stocks which are consistently traded in the Nigerian Stock Exchange (NSE) over a 4-year study period 3 August 2009 to 26 August 2013. We apply the Marcenko-Pastur distribution of eigenvalues of a purely random matrix to investigate the presence of investment-pertinent information contained in the empirical correlation matrix of the selected stocks. We use hypothesised standard normal distribution of eigenvector components from RMT to assess deviations of the empirical eigenvectors to this distribution for different eigenvalues. We also use the Inverse Participation Ratio to measure the deviation of eigenvectors of the empirical correlation matrix from RMT results. These preliminary results on the dynamics of asset price correlations in the NSE are important for improving risk-return trade-offs associated with Markowitz's portfolio optimization in the stock exchange, which is pursued in future work.

Real-Time Density Nowcasts of U.S. Inflation: A Model-Combination Approach

Edward Knotek II
Federal Reserve Bank of Cleveland
United States

Saeed Zaman
Federal Reserve Bank of Cleveland
United States

We use mixed-frequency models to produce density nowcasts of headline and core inflation in the U.S. consumer price index (CPI) and price index for personal consumption expenditures (PCE). The empirical models we use have been shown by previous research to be highly accurate in generating point nowcasts of inflation. We use a linear opinion pool to combine the individual nowcast densities from a mixed-frequency dynamic factor model, a MIDAS model, and a parsimonious regression model. Both the individual and the combined density nowcasts of inflation are evaluated in real-time. Our results suggest that the combined density nowcasts of inflation generally improve upon individual models' density nowcasts, and that both approaches outperform the statistical benchmarks. As information accumulates over a month or quarter, the accuracy of both the individual and combined density nowcasts improve as judged by logarithmic scores and calibration tests. We evaluate various schemes to compute real-time recursive weights for combining density nowcasts.

Recovering the yield curve evolution

Bonsoo Koo
Monash University
Australia

Davide La Vecchia
University of Geneva
Switzerland

Oliver Linton
University of Cambridge
United Kingdom

We develop a novel nonparametric inferential procedure for the yield curve dynamics. Our methodology makes use of cross-sectional observations on bonds, either zero-coupon or coupon-paying, across several time periods. The novelty of our approach is the combination of two different techniques: cross-sectional nonparametric methods and kernel estimation for locally-stationary time series. The resulting nonparametric time-varying functional estimator is able to capture the yield curve shapes and dynamics commonly observed in the fixed income markets. Consistency, the rate of convergence and asymptotic normality of the proposed estimator are derived. A Monte Carlo exercise illustrates the good performance of the proposed method, under different scenarios.

Reduced Rank Regression in Large VARs

Kelly Trinh
The University of Queensland
Australia

Rodney Strachan
The University of Queensland
Australia

In this paper, we address the over-parameterization issue in large vector autoregressive (VARs) models by using reduced rank regression. Our aim is to provide a model specification which is invariant to the variable ordering. To achieve a computable and efficient algorithm for estimating such large VAR models, we adopt the parameter expansion approach, and a computation procedure exploiting a certain Kronecker structure of covariance matrices. We carry out an extensive Monte Carlo simulation to examine the performance of marginal likelihood approximated by using cross entropy, predictive likelihood, and Laplace approximation in selecting the number of ranks of the VAR matrices. The results reveal that these approaches underestimate the number of ranks when the dimension of VAR systems grows, and when the singular values of the VAR matrices are small (close to zero). We then go further to examine the forecast performance of the models with misspecified ranks relative to the model with the correct rank (the benchmark model) using the measures of point forecast (e.g., mean squared forecast error, weighted mean squared forecast error) and density forecast (e.g., log predictive likelihood). Our results suggest that the models with lower rank perform worse than the benchmark for short forecast horizons, however, they perform as well as or even beat the benchmark for long forecast horizons. These patterns are more evident when the magnitudes of the singular values of the VAR coefficient are small.

Renewable energy forecasting: nonlinearity, dimensionality, and sharing aspects

Pierre Pinson
Technical University of Denmark
Denmark

Renewable energy modelling and forecasting has come a long way since the first work performed and published in the early 80s. Forecasting is today recognized as a key element to an efficient integration of renewable energy generation in existing power systems and electricity markets. However, the underlying analytics has not progressed much, expect for a few breakthrough related to the transition from deterministic to probabilistic forecasting, or to the recent availability of new types of data in large quantity e.g. distributed surface observations, weather radar images, etc.

Some of the most fundamental challenges with renewable energy forecasting (e.g., wind and solar) relating to nonlinearity and dimensionality remain. The nonlinear nature of the processes involved originates in that of the weather, while being magnified through the conversion from meteorological information to power generation. Recent advances in accounting for those nonlinearities, using data transformation and Markov-switching dynamic modelling, will be covered. When it comes to dimensionality, new approaches have to be thought of for adaptive and recursive estimation in very large models, also encouraging and rewarding sparsity. This is while restrictions in data sharing ought to be considered, since most actors that collect renewable power forecasts do not want to share their private data and information in view of competing commercial interests. Consequently, new opportunities coming from the proposals and application of distributed optimization to learning problems will be presented and discussed. Finally, a vision for a future renewable energy forecasting framework will be described based on the concept of data, computation and forecast trading.

Retail Sales Forecasting at Walmart

Brian Seaman Walmart

There are many related uses of item level sales forecasts at Walmart including replenishing, pricing, and marketing, in which the inputs of inventory, price, and ad spend can all interact to impact sales. I will discuss how each of the uses can lead to different forecasting objectives and a corresponding focus on different error metrics. In addition, there are several tradeoffs in the forecasting methods used to best meet each objective including the kinds of models used, running time speed, and requirements of forecast accuracy, and I will speak through how we approach these decisions.

SARIMA Damp trend Grey Model forecasting method

Rafael Bernardo Carmona-Benitez
Universidad Anahuac Mexico
Mexico

Maria Rosa Nieto
Universidad Anahuac Mexico
Mexico

The contribution of this study is to propose a new model to forecast data by developing a stochastic damping factor based on the Seasonal Autorregresive Moving Average Model (SARIMA) able to estimate the trend of data and to track the seasonality factor with a dynamic seasonal component. The advantages are that this model is able to capture the dynamic trend of data, non taken into account by other Grey Models, and the seasonal signal is tracked from data rather than calculated with a static equation such as the Fourier series wich has been

used by other Grey Models. The main disadvantage is that this model needs a minimum data points depending on the data serie under study when other Grey Models work with only four data points. The validation of the model is done by creating several random scenarios assumming a data generation process with Monte Carlo simulation. We use a coverage of 95%.

Seasonal Adjustment Of Daily Data - First Results

Dominique Ladiray
INSEE
France

The progresses in information technology have fostered the availability of daily and weekly time series. The seasonal adjustment of high-frequency time series poses several challenges. First of all the seasonal period of the annual cycle is neither constant nor an integral. Secondly, in order to accommodate complex seasonal patterns many individual effects might be required. Thirdly, the need for robust methods is reinforced by the fact that the effects of outlying observations is not smoothed by temporal aggregation and that they are relatively more frequent.

Moving average methods (X-13ARIMA-SEATS) and ARIMA-based methods (TRAMO-SEATS) are recommended by Eurostat and are implemented in JDemetra+, the European software for seasonal adjustment. These methods only deal with monthly and quarterly series; they estimate first the outliers and calendar effects using a Reg-ARIMA model and then decompose the residual of the model into trend-cycle, seasonality and irregular component.

We present a first implementation of a Tramo-Seats algorithm and a X-12 algorithm to seasonally adjust high frequency data with multiple and non-integer seasonalities.

Seasonal data decomposition and forecasting using STR (method and R package)

Alexander Dokumentov
Monash University
Australia

Rob Hyndman
Monash University
Australia

New general methods for decomposing and forecasting seasonal time series are going to be discussed along with the corresponding R package. These new methods are much more general than alternative practical time series decomposition methods. They allow for multiple seasonal and cyclic components, multiple linear covariates with constant, flexible, seasonal and cyclic influences. The seasonal patterns, for both seasonal components and seasonal covariates, can be fractional and flexible over time. Moreover, the seasonal patterns can either be strictly periodic or have a more complex topology.

SEMI-PARAMETRIC EXTENSIONS OF THE CAIRNS-BLAKE-DOWD MODEL: A ONE-

DIMENSIONAL KERNEL SMOOTHING APPROACH

Han Li
University of New South Wales
Australia

Colin O'Hare
University of New South Wales
Australia

Over the last few decades, there has been an enormous growth in mortality modeling as the field of mortality risk and longevity risk has attracted great attention from academic, government and private sectors. In this paper, we propose a time-varying coefficient mortality model aiming to combine good characteristics of existing models with efficient model calibration methods. Nonparametric kernel smoothing techniques have been applied in the literature of mortality modeling and based on the findings from Li et al.'s (2015) study, such techniques can significantly improve the forecasting performance of mortality models. In this study we take the same path and adopt a kernel smoothing approach along the time dimension. Since we follow the model structure of the Cairns-Blake-Dowd (CBD) model, the TVC model we propose can be seen as semi-parametric extensions of the CBD model and it gives specific model design according to different countries' mortality experience. The fitting and forecasting results from empirical studies have shown superior performances of the model over a selection of well-known mortality models in the current literature.

Simulate Car Racing Video Game

Charles Chen
Morrill Learning Center
United States

Mason Chen
Morrill Learning Center
United States

This STEM project simulated the Video Gaming patterns to predict the gaming result. In today’s US education system, playing video games is becoming a critical part for High School and Middle School Students. Parents are suffering on kids’ time management and critical thinking development. In order to address these concerns, Hill Climb Car Racing video game was selected not based on commercial ranking/popularity, instead, due to its embedded comprehensive database. The database includes the highest score of each car for each field, and the four car upgrade technologies. Authors used STEM approach: Physics Science, Car Technology, Systematic Engineering Failure mode, and Statistics/Data Mining to select the right car, upgrade the right car technology for particular racing fields. The main objective of this project is to educate kids how to play video game smarter not harder by introducing the return of investment (ROI) metric. Return is the game score and investment is playing time and car upgrading cost. Regarding the analytics methodology, Z Transformation was applied to the collected raw data in order to eliminate the bias of sample mean and sample standard deviation when comparing different cars and different fields. Otherwise, the top 2-3 most difficult fields may dominate the predictive model. Hierarchical Clustering technique was applied to group the similar fields purely based on data. The clustering patterns aligned very well with the common failure modes and the relevant car racing physics science. To minimize investment cost, authors can just upgrade one particular car to support multiple fields with similar failure modes. Model can determine the right technology (such as Engine, Tire, Suspension, Fuel…) to upgrade the assigned Car for any field cluster. Multi-correlation was conducted to verify the car technology upgrading strategy and model. Authors also used simple linear regression model to predict the ROI performance (Y= Score,

X= Upgrade Units/Cost). Authors started from a very basic linear model with R-Square at 66% and slope of 147 meter/unit improved to R-Square at 92% and slope of 512 meter/unit. The higher R-Square will indicate the higher model accuracy. The higher slope will reflect the higher Return of Investment ratio. As compared to Quadratic model, and Cubic model, linear model was selected based on R-Square (Adj.) to avoid any over-fit concern. This Video Game STEM model was very successful to help families to manage students' video game playing time at home while developing their STEM concepts and skills.

SPECIFICATION OF GARCH MODEL UNDER ASYMMETRIC ERROR INNOVATIONS

OYEBIMPE ADENIJI
university of Ibadan
Nigeria

Ismail Shittu
University of Ibadan
Nigeria

An empirical analysis of the mean return and conditional variance of Nigeria Stock Exchange (NSE) index is performed using various error innovations in GARCH models. Conventional GARCH model which assumed normal error term failed to capture volatility clustering, leptokurtosis and leverage effect as a result of zero skewness and kurtosis respectively. We re-modify error distributions of GARCH (p,q) model inference using some thick-tailed distributions. Method of Quasi – Maximum Likelihood Estimation (MLE) was used in parameter estimation. The robust model that explained the NSE index is determined by log likelihood and model selection Criteria. Our result shows that GARCH model with fat-tailed densities improves overall estimation for measuring conditional variance. The GARCH model using Beta-Skewed-t distribution is the most successful for forecasting NSE index.

Statistical Challenges in Forecasting Revenue for a Large-Scale Hierarchically Structured Business

Julie Novak
IBM Research
United States

Beatriz Etchegaray Garcia
IBM Research
United States

Yasuo Amemiya
IBM Research
United States

Large-scale businesses need to have a clear vision of how well they expect to perform within all their different units. This information will directly impact managerial decisions that will in turn affect the future health of the company. In this talk, we focus on the statistical challenges that occur when implementing our revenue forecasting methodology on a weekly basis within a large business. We must provide reasonably accurate forecasts for all the geography/division combinations, which have fundamentally different trends and patterns over time. Our method must be robust to "oddities", such as typos in the input or unusual behavior in the data. In

addition, our forecasts must be stable over weeks, without sacrificing on accuracy. We describe the statistical methods used to maintain an efficient and effective operational solution.

STOCK PRICE PREDICTION USING DYNAMIC MODE DECOMPOSITION

DEEPTHI PRAVEENLAL KUTTICHIRA
AMRITA VISHWA VIDYAPEETHAM
India

Dr.GOPALAKRISHNAN E.A
AMRITA VISHWA VIDYAPEETHAM
India

Vijay Krishna Menon
AMRITA VISHWA VIDYAPEETHAM
India

Dr.Soman K.P
AMRITA VISHWA VIDYAPEETHAM
India

Stock Price Prediction Using Dynamic Mode Decomposition

Stock price prediction is a challenging problem as the market is quite unpredictable. We propose a method for price prediction using Dynamic Mode Decomposition assuming stock market as a dynamic system. DMD is an equation free, data-driven, spatial-temporal algorithm which decomposes a system to modes that have predetermined temporal behaviour associated with them. These modes help us determine how the system evolves and the future state of the system can be predicted. Is it usable for the predictive assessment of the stock market was our research question. We worked with the time series data of all the companies listed in National Stock Exchange. The granularity of time was minute. We have sampled a few companies across sectors listed in National Stock Exchange and used the minute-wise stock price to predict their price in next few minutes. The obtained price prediction results were compared with actual stock prices. We used Root Mean Square Error to calculate the deviation of predicted price from actual price for each company. Price prediction for each company was made in three different ways. In the first, we sampled companies belonging to the same sector to predict the future price. In the latter, we considered sampled companies from all sectors for prediction. In the first and second method, the sampling as well as the prediction window size were fixed. In the third method the sampling of companies was done from all sectors considered. The sampling window was kept fixed, but predictions were made until it crossed a threshold error. The results were presented in graphs. The graphs showed comparable results between predicted values and actual values. The lowest RMSE obtained was 0.05 and the highest was 18, for the third approach. Prediction was found to be more accurate when samples were taken from all the sectors, than from a single sector. When sampling window alone was fixed; the predictions could be made for longer period for certain instances of sampling. The spatial analysis identified stocks with larger weighting factors. These are the top stocks identified by DMD algorithm. For each sampling window, top stocks are identified. Using DMD gave promising results in stock market prediction. Better predictions were possible when different sectors were considered. This result support the inter-sector dependency in stock market. Current literature identifies two arguments. Technical analysts believe current price depend on past prices. Supporters of efficient market hypothesis opposes this view. The present study favours the former, but limited to certain times only. This opens to further research.

Strategic Operations Management Transitions in Food Processing SMEs :a case study in the north of UK

Konstantinos Nikolopoulos
forLAB Bangor University & Visiting Researcher Lancaster Centre for Forecasting
United Kingdom

Marc Christian Davies
forLAB Bangor University
United Kingdom

Siwan Mitchelmore
forLAB Bangor University
United Kingdom

Sarah Wale
forLAB Bangor University
United Kingdom

The forecasting group, forLAB, was invited to assess a local SME's planning and inventory systems with a view to improving their forecasting systems as well as implementing lean processes and knowledge management in their production methods. The company saw a need to alter their production scheduling so that it was less reactive and more anticipatory of orders. At the start of the project the company ran on a mainly make to order paradigm (MTO) and utilized little forecasting. One of the main challenges was the highly variable nature of the individual products demand patterns. forLAB proposed implementing a hybrid make to order/make to stock system (MTS) in order to both level production and provide better planning capacity for both raw materials as well as production scheduling. This is obviously a strategic decision, that can only be taken in the top of the hierarchy of the company and imposed downwards. In order to tackle the variable demand characteristics of each product a method was developed that competed 6 forecasting methods against each other to select the most suitable model for each product (see 'Horses for courses' in demand forecasting, Petropoulos et al, 2014) . The ADIDA (Nikolopoulos et al 2011) methodology was used to transform intermittent time series into continuous time series to increase the available forecasting methods. An intuitive and simple selection criteria was used to determine if a product was either make to order or make to stock using the average demand and coefficient of variation. A company specific software solution using the R statistical language was developed by forLAB to aid the managers in their production scheduling and raw material purchasing decisions. Ease of use was a main requirement, and the software solution meant the managers could use it without detailed statistical knowledge. This case study highlights the journey a company makes when changing to a new production paradigm, the selection of the correct forecasting method for each product time series, selecting the correct aggregation level, and the positioning of the order penetration point (see Olhager 2003).

Targeting Market Neutrality

John B. Lee
University of Auckland
New Zealand

Jonathan J. Reeves
University of New South Wales
Australia

Alice C. Tjahja
University of New South Wales
Australia

Xuan Xie
University of New South Wales
Australia

Neutralizing portfolios from overall market risk is an important issue in investment management particularly for hedge funds. In this paper we show an economically significant improvement in the accuracy of targeting market neutrality for equity portfolios. Key features of the approach are the relatively short forecast horizon of one week and forecasting with realized beta estimators computed using high frequency data provided by QuantQuote. We also find that too long and too short estimation windows result in poor beta forecasts and that the optimal length of estimation window depends on the frequency of return observations.

Testing and Adjusting for Sharp Breaks in Seasonal Patterns: The Case of U.S. GDP Data

Keith Phillips
Federal Reserve Bank of Dallas
United States

Michael Boldin
Federal Reserve Bank of Philadelphia
United States

In recent years several researchers have found the presence of residual seasonality in the official estimates of U.S. real gross domestic product (GDP). Prior to a revision in mid-2016, directly seasonally adjusting the official seasonally adjusted GDP would revise the first quarter growth since 2013 upward by an average of about 1.5 percentage points. In mid-2016 the U.S. Bureau of Economic Analysis (BEA) revised GDP during the period from 2013 – 2015, seasonally-adjusting more of the input data in an attempt to reduce the residual seasonality in the aggregate series. We find that, while this reduced the residual seasonality, the first quarter data were still on average about 0.6 percentage points too weak. Current U.S. GDP data contain a mix of seasonal factors and varying degrees of residual seasonality. In this paper we look at ways to test and adjust for shifting seasonal adjustment and residual seasonality so that analysts can best interpret what the quarterly changes in US GDP mean for the strength of the US economy.

The 2016 US presidential election and media on Instagram: Who was in the lead?

Harald Schmidbauer
ISCTE-IUL, Lisbon, Portugal
Portugal

Angi Rösch
FOM University of Applied Sciences, Munich, Germany
Germany

Fabian Stieler
FOM University of Applied Sciences, Munich, Germany

Germany

Social network services are used, among other things, to express political opinion. The present study is an effort to analyze the timing of media postings related to candidates Clinton and Trump on the platform Instagram before and after the 2016 US presidential election. A selected set of hashtags is used to determine whether a posting was intended to support or oppose either candidate. We thus obtain four time series of hourly readings of the number of Instagram postings: Clinton vs. Trump, supporters vs. opponents. We use cross-wavelet analysis to reveal the periodic structure of these series. It turns out that all four time series have significant 12- and 24-hour periods. Among our findings is that, at the 12-hour period, the time series of Trump supporters was leading Trump opponents as well as Clinton supporters the days before the election, while the series of Clinton opponents was often leading Clinton supporters: Trump supporters and Clinton opponents were eager to post media, while Trump opponents and Clinton supporters were sluggish. Considering the forecasts for this election, these results come as a surprise.

The Contribution of International Shocks in a Small Open Economy

Jamie cross
Australian National University
Australia

Aubrey Poon
Australian National University
Australia

he objective of this paper is to quantify the proportion of business cycle fluctuations that can be attributed to intertemporal international shocks in a small open economy. The methodology employs a novel modeling feature that allows for fat-tailed time varying shocks in a panel VAR with time varying country and variable specific dynamics as well as cross-country interdependencies. Using Australia as an exemplar of such an economy, the results indicate that whilst Student-t specified models provide better in-sample fit as compared to their Gaussian counterpart, allowing for time varying shocks reveals that otherwise determined outliers are actually significant structural instabilities within global business cycles. From a general macroeconomics perspective, we find strong evidence of a world business cycle amongst the large countries; namely China, Japan, The EU and The US, in and around the 2007/08 financial crisis, however distinct business cycles emerge for small economies; namely Australia, Canada, New Zealand and the Republic of Korea. As for Australia specifically, we show various interesting shifts in the contribution of international of shocks over time. A crude summary of the results is that both domestic and international shocks have contributed to around half of all business cycle fluctuations over the past two decades.

The Economic Value of Business Cycle Forecasts in Germany for Potential Investors

Karsten Mueller
University of Applied Science Merseburg
Germany

Joerg Doepke
University of Applied Science Merseburg
Germany

Lars Tegtmeier
University of Applied Science Merseburg
Germany

Several papers have called for a more economic approach in forecast evaluation (see, e.g., Granger and Pesaran, 2000). Hence, we evaluate the economic value of business cycle forecasts in Germany for potential investors on financial markets as opposed to statistical measures of forecasts accuracy. Based on annual data ranging from 1990 to 2016 covering 23 different forecasts stemming from 23 different public and private institutions, we update – in a first step – standard statistical measures of forecast accuracy and efficiency. In this respect, our findings broadly confirm results of previous studies: German growth forecast are unbiased, but still inefficient, while inflation forecast show even signs of bias, but appear to be efficient.

In a second step, analyses of the economic value of these forecasts based on evaluation of the directional change errors show that growth forecast have an information content and are significantly better than a coin flip, while the respective tests for inflation forecasts raise doubts on the usefulness of the predictions.

To gain more insights on the economic value of growth and inflation forecasts, we calculate hypothetical portfolios that are actively managed based on such predictions and find that they outperform passively managed benchmark portfolios. While this result might cheer up forecasters the results based on risk adjusted returns (Sharpe Ratios) are more mixed: Based on Gibbons, Ross, Shanken (1989) tests for equal Sharpe Ratios, we conclude that only some of the institutions under investigation provide forecasts that lead to a significantly higher Sharpe Ratio than the benchmarks. We check these results by a couple of robustness checks: Referring to portfolios built on simple trading rules that reflect the past errors made by the forecasters (the forecaster "credibility") does not change the main results. Also, a trading rule that incorporates growth and inflation forecasts jointly rather than independently leaves the main conclusions unaffected. Finally, we also document that a ranking of forecasts quality based on these measures differs markedly from rankings derived from purely statistical measures of accuracy.

The economic value of commodities in asset allocaton when returns are predictable

John Cotter
UCD Michael Smurfit Graduate School of Business
Ireland

Emmanuel Eyiah-Donkor
UCD Michael Smurfit Graduate School of Business
Ireland

Valerio Poti
UCD Michael Smurfit Graduate School of Business
Ireland

We evaluate the out-of-sample performance of commodities in portfolios composed of stocks, bonds, and T-bills. The evidence on their ability to generate economic value is mixed, with previous studies ignoring the potential role of asset return predictability, and the states of the business cycle. Using monthly data over the period 1976-2015, we document sizeable utility gains for a mean-variance investor who exploits predictability in asset returns. For example, an investor with moderate level of risk aversion, who imposes sensible restrictions on portfolio weights, can generate net-of-transactions-costs utility gains of over 130 basis points per annum. In addition, we find strong influence of the business cycle, as dated by the National Bureau of Economic Research, on portfolio performance. During the recessionary phase of the business cycle, commodities are shown to generate substantial utility gains of over 1338 basis points per annum. In expansionary periods of the business

cycle, however, commodities do not add economic value generating utility losses of over 50 basis points. These findings suggest that the ability of commodities to improve portfolio performance is countercyclical. Our findings are robust to varying levels of risk aversion, portfolio weight constraints, transactions costs and alternative performance metrics.

The impact of festivals on gold price expectation and volatility

Angi Roesch
FOM University of Applied Sciences, Munich
Germany

Harald Schmidbauer
BRU-IUL, ISCTE Business Research Unit, ISCTE-IUL, Lisbon
Portugal

Almost 60% of gold worldwide is bought as jewelry. China and India together account for about 60% of worldwide gold consumption. Gold is a traditional gift in many cultures and often given away on the occasion of festivals, such as Chinese New Year, Diwali, Ramadan Eid, and Christmas. These facts lead us to an obvious question: Do gold prices behave differently around festivals? This question has not yet been systematically discussed on the basis of stochastic models. The purpose of the present paper is to investigate the effects of a selection of festivals on the distribution of daily gold price changes. Dummy variables indicating the festival are modified to reflect anticipation and aftereffect of festivals, and they are used as covariates in a combination of regression and GARCH models to specify conditional expectation and volatility of daily gold price changes. After fitting this model to data, extensive robustness checks are undertaken to make sure our results are valid.

Differentiating between two periods (1991–2002 and 2003–2014) with distinct gold price characteristics, we find that several festivals (among them, the four mentioned above) are connected with gold price changes in different ways. This research contributes to the problem of understanding and forecasting gold prices.

The impact of macroeconomic leading indicators on inventory management

Yves R Sagaert
Ghent University
Belgium

Nikolaos Kourentzes
Lancaster University
United Kingdom

Stijn De Vuyst
Ghent University
Belgium

El-Houssaine Aghezzaf
Ghent University
Belgium

Bram Desmet
Solventure
Belgium

Forecasting tactical sales is important for long term decisions such as procurement and informing lower level inventory management decisions. Macroeconomic indicators have been shown to improve the forecast accuracy at tactical level, as these indicators can provide early warnings of changing markets while at the same time tactical sales are sufficiently aggregated to facilitate the identification of useful leading indicators. Past research has shown that we can achieve significant gains by incorporating such information. However, at lower levels, that inventory decisions are taken, this is often not feasible due to the level of noise in the data. To take advantage of macroeconomic leading indicators at this level we need to translate the tactical forecasts into operational level ones. In this research we investigate how to best assimilate top level forecasts that incorporate such exogenous information with bottom level (at Stock Keeping Unit level) extrapolative forecasts. The aim is to demonstrate whether incorporating these variables has a positive impact on bottom level planning and eventually inventory levels. We construct appropriate hierarchies of sales and use that structure to reconcile the forecasts, and in turn the different available information, across levels. We are interested both at the point forecast and the prediction intervals, as the latter inform safety stock decisions. Therefore the contribution of this research is twofold. We investigate the usefulness of macroeconomic leading indicators for SKU level forecasts and alternative ways to estimate the variance of hierarchically reconciled forecasts. We provide evidence using a real case study.

The influence of SOI index in Brazil streamflow using Generalized Additive Models

Reinaldo Castro Souza
PUC-Rio
Brazil

Paula Medina Maçaira
PUC-Rio
Brazil

Fernando Luiz Cyrino Oliveira
PUC-Rio
Brazil

Fábio Iha
CTG Brasil
Brazil

The Brazilian electrical energy matrix is essentially formed by hydraulic sources, which currently account for 70% of the installed capacity. One of the most important characteristics of a generation system with hydro predominance is the strong dependence on the streamflow regimes. Nowadays, the Brazilian Electrical System (BES) uses a linear autorregressive model to modelling and simulate synthetic scenarios for hydrological streamflow for the next 5 years on a monthly basis (60 months ahead). Meanwhile, this approach does not consider any exogenous information that may affect the hydrological regimes. The main objective of this paper is to infer on the influence of El Niño and La Niña phenomena into the streamflow process as a way to improve the model's performance. To do so, it was applied, as an innovative approach the so called, Generalized Additive Models considering, as an exogenous variable, the Southern Oscillation Index (SOI), a standardized index that gives an indication of the development and intensity of El Niño or La Niña events in the Pacific Ocean, to generate scenarios for the South Brazilian subsystem. The obtained results are quite promising, as the proposed method produced better performance metrics when compared to the official one. These findings indicate that the

inclusion of climate variables in this context is a promising research field, considering the importance of the scenarios accuracy in the long term energy planning.

The influence of the price variable in the forecast of electricity consumption in Brazil

Paula Medina Maçaira
PUC-Rio
Brazil

Reinaldo Castro Souza
PUC-Rio
Brazil

Fernando Luiz Cyrino Oliveira
PUC-Rio
Brazil

Vanessa Oliveira
Eneva
Brazil

The behavior of electricity consumption in Brazil is highly important, as it has a key role as a vector for economic, social development and life quality improvement, as well as on the estimation of future electricity demand. For planning purposes, it is important to design a tool able to integrate long-term electricity demand forecasting with the estimation of hourly demand and the feedback of the electricity price in the short term. In this paper, the annual electricity demand forecast is obtained via a bottom-up approach. Following, these annual forecasts are disaggregated into hourly load curves for each day of the year. These hourly forecasts are then used to generate the optimal marginal cost of generation, which is a proxy for electricity price. Once these future prices are generated, they feed the bottom up forecasting model to generate another run of annual and daily forecasts and so on and so forth, until convergence is achieved. The application of this procedure for Brazilian data produced, on the first round of the loop, an average growth of electricity consumption of 2% per year. The obtained consumption is lower than the predicted by the government over the entire study horizon, as expected, since the macroeconomic assumptions of government are much more optimistic than the considered in the bottom-up approach. In addition, the first round leads to lower prices than the original ones and, therefore, the predicted consumption obtained in the second round is larger than that obtained in the first one, however still lower than government forecasts. In other words, as one expects, the price variable has an inverse influence on the consumption of electricity.

The probability of remaining poor in Mexico

José Carlos Ramírez Sánchez
Universidad Anáhuac México, Campus Norte
Mexico

Vivian Daniela Kadelbach
Universidad Anáhuac México, Campus Norte
Mexico

Leovardo Mata Mata
Universidad Anáhuac México, Campus Norte

Mexico

This paper is aimed at estimating the probabilities of being poor in Mexico in the short term. By using a Markovian approach on the population's social deprivations we study Mexico´s poverty dynamics in the absence of any significant economic change. In doing so, we set a priori scenario of four states of social deprivations in which Mexico´s poverty evolves according to the economic conditions experienced by its population during the 2002-2009 period. Then we perform a hidden-Markov chain analysis in order to check the robustness of conclusions drawn from such scenario. After collecting results from both kind of analyses we find, on the basis of their ergodic vectors, the same pattern of impoverishment. So, far from fitting the ideas of the exit-time poverty theory, the Markovian approach developed reveals a darker view of Mexico's poverty evolution, especially when one observes that some probability transitions showed to be insensitive to changes in the GDP growth rate. For this reason, among others, we consider that using stochastic techniques provides a more realistic perspective of Mexico's poverty evolution than that offered by deterministic methods.

The time-varying nature of the predictors for Mobile Social Networking adoptions

Mohsen Hamoudia
Orange Labs
France

Emanuele Giovannetti
Anglia Ruskin University, Cambridge
United Kingdom

The diffusion of smartphones and of faster and more reliable mobile broadband has allowed the rapid adoption of multiple Mobile Social Networking applications. This paper develops a two-stage estimation strategy focusing first, on identifying early and late Mobile Social Networks adopters and then, on assessing how the indirect network externalities resulting from marketing strategies and affecting customers' usage time, pricing and smartphones penetration can be used as predictors for Mobile Social Networks adoptions. The paper's main finding is that the effects of these predictors can be radically different between early or late adopters. These differences are captured by estimating the interaction effects between prices, usage and smartphones diffusion with country-specific variables indicating whether a diffusion peak has been reached. The paper also shows that predictors based on penetration pricing strategies are more effective for early adopters, as the additional benefits experienced by later adopters makes them less sensitive to prices. The proposed modelling approach provides a useful framework for using two stage modelling into forecasting extensions of this work.

Time Series Forecasting at Scale by Blending Automatic Techniques with Non-Expert Interventions

Boris Roussev
NerdWallet Inc.
United States

Ramesh Subramonian
NerdWallet Inc.
United States

Indrajeet Singh
NerdWallet Inc.
United States

Time Series Forecasting at Scale by Blending Automatic Techniques with Non-Expert Interventions

Producing high quality forecasts at scale is a challenging task. In this paper, we present the design and features of NextPlay, a NerdWallet experimental project in time series analysis and forecasting at scale. NextPlay uses a combination of automatic forecasting techniques and manual user interventions to produce accurate forecasts at scale by non-experts in time-series analysis. We demonstrate how we collect, store, and process time series data as well as show a workflow for creating forecasts for a large financial advising domain space. We introduce a time series expression language for mixing and matching two or more time series to derive new ones to solve interesting problems in application domains. We show how users can share their results organization-wide and persist their settings and interventions to apply them automatically to future user sessions.

More than 100 million people in the US have used NerdWallet to get advice on a range of financial products including: healthcare; mortgages; life insurance; banking; credit cards; wealth management; small business tools, and college loans. At NerdWallet, we use forecasts for capacity planning, prioritization of new features, feature scoping and development, collaboration with partners, resource allocation, and goal setting. Our data analysts, business analysts, editors, writers, and managers need accurate forecasts on metrics such as user traffic, user engagement, user registrations, product applications, and transactions faceted by traffic source, attribution, device, financial verticals, and articles.

On the one hand, the large time series space necessitates the use of automatic forecast techniques. Scaling the execution of forecasting algorithms on Amazon EC2, S3, and RDS is relatively straightforward. However, automatic techniques cannot incorporate at scale ever-changing domain knowledge, such as market penetration percentages by region, local holidays, local events, US national holidays occurring at irregular intervals, and missing or erroneous observations due to system and/or human failures.

On the other hand, data scientists with highly specialized skills can help incorporate domain knowledge in forecast models, but they are few and far between, so this approach does not scale at all.

Based on the above observations and the organization's business needs, our goal with NextPlay is to provide non-experts with tools for generating interactively high-quality forecasts by using a combination of automatic forecasting techniques and domain expert interventions as well as to share their results company-wide using user-friendly web tools, which do not require training.

Uncertainty and forecasts in risk sharing through panel VARs

Filippo M. Pericoli
European Commission, Joint Research Centre (JRC)
Italy

Pilar Poncela
European Commission, Joint Research Centre (JRC)
Italy

Eduardo Rossi
European Commission, Joint Research Centre (JRC)
Italy

Panel VARs models are gaining popularity in macro analysis. For instance, within the European Union this model can be widely applied for policy evaluation. In macro problems, the time dimension T is usually long enough and in that case the homogeneity assumption about the coefficients of the VAR model might not be well suited and dynamics should be modeled for each unit. In heterogeneous panel VAR models, we specify each coefficient matrix as the sum of a common part to all units plus an idiosyncratic one. It could be of interest not only to estimate the common part but the idiosyncratic ones as well. The Bayesian approach is gaining popularity, however, the contributions in the frequentist domain are more scarce. This paper is devoted to fill this gap. Within the frequentist approach, the most popular estimator in macro analysis is the mean group estimator for the common part. Within this framework, it is of primary interest to be able to characterize the uncertainty around its impulse response functions as well as the mean square forecast errors of the model. Of course, they depend, among other things, on the variance-covariance structure assumed for the error term. We apply our model to estimate risk sharing for a set of developed countries. International risk sharing focuses on the cross border channels at work in smoothing income and consumption when a country is hit by a negative output shock. Indeed, in an ideal world of perfect risk sharing, countries are completely insured against bad events, and domestic consumption growth is independent from domestic output growth. In practice, this is hardly the case and our interest is to measure what percentage of GDP shocks are passed into consumption using panel VARs that can be possibly heterogeneous.

Uncertainty of budget and macroeconomic forecasting: asymmetric loss or unpredictability of recessions

Yoichi Tsuchiya
Tokyo University of Science
Japan

This study examines the budget forecasts to investigate whether uncertainty of those forecasts results from an asymmetric loss function and/or unpredictability of the state of an economy. This study covers various variables including real growth, inflation, financial indicators, and others for macroeconomic forecasts, and tax revenues including its main sub-components. It also examines forecasts made by government, governmental agencies, and professional forecasters. Empirical results generally show evidence to support both channels while some variables show different results. For example, examination of budget forecasts by the Congressional Budget Office (CBO) shows that (1) the CBO has an asymmetric loss for the budget forecasts, leading to underpredicted deficit, and overpredicted revenue and outlay forecasts, and (2) information of the state of the economy is not incorporated into revenue and deficit forecasts, and they result in underprediction of deficit and overprediction of revenue. In particular, the revenue forecast is crucial because the current- and next-year forecasts is overpredicted and do not incorporate information on the state of an economy. In contrast, the next-year outlay forecast incorporates those information while the current-year forecast does not efficiently use those information.

Understanding the Equilibrium Dynamics of European Broadband Diffusion

Gary Madden
Curtin University
Australia

Hiroaki Suenaga
Curtin University

Australia

This study examines European broadband market equilibria within a simultaneous demand and supply model. Unlike alternative reduced-form approaches, our model explicitly incorporates key features of the demand for network services: a bell-shaped distribution of heterogeneous consumers; and the presence of network effects. Our empirical analysis of European countries reveals that although the demand for broadband services exhibit strong network effects they are not 'strong enough' for critical mass to be achieved within these markets. Furthermore, the systems equilibrium dynamics show that even though the technology is introduced at very high prices initially, the expansion of suppliers' infrastructure investment reduces the equilibrium prices rapidly in the first five years. This price reduction, combined with strong network effects, stimulates expansion of the subscription base until consumers at the peak of the distribution enter the market. Convergence to the steady-state happens at around 8 years for service price and 15 years for subscription. Moreover, comparative-static analysis illustrates that steady-state equilibrium subscription is positively related to national population density, inter-platform concentration, and the share of unbundled local loop while equilibrium price is inversely related to these factors. Other supply and demand factors (viz., intra-platform concentration, income, and education) have minimal impact on the steady-state equilibrium.

Use and expectations of adoption forecasts in the planning of agricultural innovations

Oscar Montes de Oca
University of Western Australia
Australia

Rick Llewellyn
CSIRO
Australia

Larelle McMillan
CSIRO
Australia

John Moore
CSIRO
Australia

David Pannell
University of Western Australia
Australia

This paper explores the perceived usefulness of adoption forecasts in assisting planning in the Agricultural Research, Development and Evaluation (RD&E) sector. Accurate adoption forecasts are essential to ensure investments in RD&E achieve their goals (e.g. generating profit, or designing environmental policy). It is difficult to quantify the total investment from both the public and private sectors in Agricultural RD&E. However, according to the OECD and IFPRI statistics, public expenditure in agricultural research has increased constantly and considerably in both developed and developing countries over the last 20 years. The last available figures estimate a total investment of over US$ 13.6 billion made in OECD countries in 2014 and almost US$ 20 billion made in developing countries in 2012. Stakeholders representing both commercial and environmental interests in the agricultural RD&E sector include scientists, research funders, agricultural companies, government agencies, extension professionals, farmer organisations, NGOs, industry bodies, universities and other education institutions. This research aims at gaining a better understanding of the different uses of adoption forecasts within this diversity of stakeholders and objectives.

The adoption of innovations in agriculture is a complex problem with many different factors influencing it. Accurate adoption forecasting in this sector therefore relies on a multitude of methods, both statistical and judgemental, and considerable expert knowledge. In this research, we limited our investigation to the use of an established expert system by a cross-section of agricultural RD&E practitioners in Australia and New Zealand. We selected and interviewed 40 registered users of ADOPT (Adoption and Diffusion Outcome Prediction Tool), an agricultural innovation adoption forecasting model developed in Australia over the past nine years. ADOPT forecasts the adoption peak and the diffusion speed of an innovation before it is introduced to a population of potential users, calculating a sigmoidal curve and a sensitivity analysis of 22 drivers of adoption.

The interviews identified a variety of forecast uses in both individual and group decision-making across a wide range of organisational roles. The interviews also uncovered differences in opinion on the importance of forecast precision, sensitivity analysis, limitations and model use for learning and communicating adoption. Findings from the study also highlights the role of expert systems in forecasting adoption in agriculture. The results from this research can be used to better match forecasting methods to specialised user segments in the agricultural RD&E sector.

Using online search traffic and social network shares for operational demand forecasting

Oliver Schaer
Lancaster Centre for Forecasting, Lancaster University
United Kingdom

Nikolaos Kourentzes
Lancaster Centre for Forecasting, Lancaster University
United Kingdom

Robert Fildes
Lancaster Centre for Forecasting, Lancaster University
United Kingdom

Recently, there has been substantial research on augmenting forecasts with data from internet platforms, such as search traffic or social network shares. Although the majority of studies report increased accuracy, several exhibit various weaknesses, lacking adequate benchmarks or rigorous evaluation and experimental design. Furthermore, their usefulness over the product life-cycle has not been investigated, which may change, as initially, more consumers may search for pre-purchase information, but later for after-sales support. In this study we first review the relevant literature and subsequently we attempt to validate the key findings using two forecasting case studies. Our findings are in stark contrast with the literature and we find that established univariate forecasting benchmarks, such as exponential smoothing, consistently performed better. We argue that the limited usefulness of online platform data is due to the required lead time for taking operational decisions in organisations, an aspect typically overlooked in the existing literature.

Variance stabilizing transformations for electricity spot price forecasting

Bartosz Uniejewski
Wrocław University of Science and Technology
Poland

Rafal Weron
Wrocław University of Science and Technology
Poland

Florian Ziel
University Duisburg-Essen
Germany

Most electricity spot price series exhibit price spikes. These extreme observations may significantly impact the obtained model estimates and hence reduce efficiency of the employed predictive algorithms. For markets with only positive prices the logarithmic transform is the single most commonly used technique to reduce spike severity and consequently stabilize the variance. However, for datasets with very close to zero (like the Spanish) or negative (like the German) prices the log-transform is not feasible. What reasonable choices do we have then?

To address this issue, we conduct a comprehensive forecasting study involving 12 datasets from diverse power markets and evaluate 16 variance stabilizing transformations. We find that the probability integral transform (PIT) combined with the standard Gaussian distribution yields the best approach, significantly better than many of the considered alternatives.

Weather Normalization for Hourly Electric Energy Models

Bradley Lawson
SAS Institute
United States

INTRODUCTION The demand for electricity can vary widely during a given time because of temperature. This creates an erratic sales and revenue history that can make it difficult to understand the true underling growth trend. Electric utilities generally do some form of “weather adjustment” to smooth some of the variation. The concept of normal weather is not uniform across utilities and is often a standard determined by the state public utilities commission. This paper compares the methodology and results of current and two newer approaches. The purpose is not to determine whether any one of the three method is superior to another, but to examine the differences between the two methodologies.

Traditional Weather Adjustment All utilities historically relied on monthly grain data since most customers were served with a meter which was read once per month. Traditionally electric utilities have based their weather adjustment on the coefficients from a monthly regression model applied to the difference between actual and normal weather. Raw temperature data is difficult to use in this modeling and is usually converted into degree days. Hourly Weather Adjustment Automatic metering infrastructure (AMI) is now becoming widespread providing hourly or finer grain and allowing utilities to expand their modeling to the hourly grain. Models for hourly forecasting are more complex including not only hourly temperature variables but temporal interactions. Calculating a weather adjustment from an hourly forecast model requires an hourly normal temperature forecast for the target year. Normal temperature models are generated through a wide variety of methods ranging from simply selecting data from a given year to statistically based products. This paper investigates the range of models and the effect on an hourly weather adjustment where day of week and hour affects results. Weather Adjustment With Hourly Probabilistic Forecasts Many forecasters are moving toward probabilistic forecasts or scenario forecast based the historic variation of temperature across 30 to 50 years. While probabilistic forecasts offer the advantage of providing a range of possible forecasts they also offer the advantage of not requiring a normal temperature model. However not having the normal temperature model requires an alternate method of estimating temperature adjustment. Newer methods have been examined in the 2014 paper “Load Normalization Against Weather” by Hong, et al in their 2014 paper in IEEE. This paper reviews the results from probabilistic models against results for the models discussed prior.

Weight Selection for Forecast Combinations

Hong Wang
Monash University
Australia

Bonsoo Koo
Monash Universitu
Australia

Weight selection is central to forecast combinations. Although forecast combinations have been employed in many empirical studies, there is no clear theoretical ground for the problem of selecting combination weights. We study the conditions in which the \estimation free" equal weight yields an optimal forecast in a class of forecast combinations constructed from the forecasts of simple regression models. Since the conditions for the success of equal-weight schemes are limited, we propose a method of selecting combination weights by variants of the Least Absolute Shrinkage and Selection Operator (LASSO). In our Monte Carlo simulation experiments, we show that the proposed LASSO-based combination compares favorably with other forecast combinations in the literature. We demonstrate the performance of the proposed method with two empirical applications: (1) predicting stock returns using the financial predictors examined in Welch and Goyal (2008) and (2) forecasting US output growth using a large set of macroeconomic predictors in Stock and Watson (2004). Our findings show that the LASSO-based combination often outperforms the simple average of the forecasts, and in turn dominates the best ex-ante candidate forecast and other methods of forecast combination in the literature.

When last minute accommodation booking is really last minute. Short term booking forecasting based on big data analysis

Miriam Scaglione
Institute of Tourism, University of Applied Sciences and Arts Western Switzerland Valais
Switzerland

Favre Pascal
Institute of Tourism, University of Applied Sciences and Arts Western Switzerland ValaisInstitute of Tourism, University of Applied Sciences and Arts Western Switzerland Valais
Switzerland

One of the most important phases in planning a vacation is the booking activity process, this is not only relevant for demand but also for the supply side. The big question is to know when the optimal moment is to launch a “last minute” offer without losing revenues. The aim of this research is to answer this question using the analysis of the booking period (BP) by country of origin/season. The BP is the lapse of time between the booking moment and the arrival. The data used belongs to the largest self-catering online booking platform in the region of the Romand? Valais in Switzerland. The data set contains more than 141,000 transactions from 1st January 2010 to 1st February 2017. This research uses the Kaplan-Meier (KM) survival method for modelling the length of BP after the 10’000-resampling process. KM lifetable yields the timing where 95% of the observed arrival is reached for each sample giving the cut-time for lastminute booking. For practitioners, the results shed some light on planning behaviour across different markets and seasons. For scholars, beside methodological issues, the results show that countries of origin are less relevant than seasonality in the characterisation of the planning vacation process (PVP). The routines have been programmed in SAS Institute V9.4 macro environment using proc lifetest, proc surveyselect and SAS SQL.

Zero inflated GARMA model, with applications to infectious disease data

Siuli Mukhopadhyay
Indian Institute of Technology Bombay
India

Rashmi Tiwari
Indian Institute of Technology Bombay
India

Sheeba Sekharan
Indian Institute of Technology Bombay
India

The aim here is to develop a time series model for infectious disease prediction and forecasting disease counts in the future. The disease data collected is in the form of counts, with presence of trend, seasonality, overdispersion and zero inflation. Due to the data being counts, a Poisson generalized auto regressive moving average (GARMA) model is fitted to the data. However, due to the inability of the Poisson distribution to model the overdispersion present, a negative binomial GARMA has to be used. To take into account the excess of zero disease counts, a zero-inflated negative binomial GARMA model including both trend and seasonality, is proposed. Parameter estimation under a partial likelihood framework is discussed. Various forecasting techniques for the proposed model are discussed and compared.

The proposed methodology is illustrated with application to weekly dengue data. Due to dengue being more prevalent in monsoon and usually rare in other seasons, the weekly data has excess zero counts. Dengue counts are usually dependent on weather covariates, so along with trend and seasonality, weather covariates are also included in the model. Several GARMA models are compared and model selection criterion like mean squared error, deviance residuals, Akaike information criteria are used to select the final dengue model. Dengue counts are then forecasted for future years using the best model.

Regional Population Estimation Using Satellite Imagery

Aikaterini Koutsouri
National Technical University of Athens, Greece

Ilectra Skepetari
National Technical University of Athens, Greece

Konstantinos Anastasakis
National Technical University of Athens, Greece

Stefanos Lappas
National Technical University of Athens, Greece

Population data provides statistical information that can in turn support decision making processes. It provides essential information regarding many practices such as rescue operations or humanitarian actions, which require the estimation of local population. While traditional approaches are possible, they tend to be time consuming and expensive. Copernicus Earth Observation data (Sentinel 2 satellite images) provides high resolution satellite imagery that could be a useful alternative to collecting census data since it is significantly cheaper than traditional methods, although less accurate. This paper describes a methodology for exporting a population estimation using satellite imagery based on the use of classification techniques coupled with a statistical forecast on historical data.

Detecting and Adjusting Structural Breaks in Time Series and Panel Data

Rajesh Selukar
SAS, USA

Detection and adjustment of structural breaks is an important step in modeling time series and panel data. In some cases, such as studying the impact of a new policy or an advertising campaign, structural break analysis might even be the main goal of a data analysis project. In other cases, the adjustment of structural breaks is a necessary step to achieve other analysis objectives, such as obtaining accurate forecasts and effective seasonal adjustment. All change-point analysis methodologies make some assumptions about the data generation model of the observation process. This workshop describes a general-purpose change-point analysis methodology that assumes that the observation process follows a (linear) state-space model (SSM). This assumption is relatively weak: most data generation models that are considered in the change-point analysis literature—such as the multiple linear regression model, univariate and multivariate ARIMAX models, unobserved components models (UCMs), and a variety of panel data models—can be formulated as SSMs.