International Symposium on Forecasting 2018

Boulder, Colorado

June 16-20

Conference Proceedings

ISF2018
Boulder CO

General Chair **Len Tashman**

Program Chair **Fotios Petropoulos**

Contents

WORKSHOPS 4

KEYNOTES 5

PANEL DISCUSSIONS 8

FORECASTING PRACTICE TRACK 9

INVITED SESSIONS 12

- Macroeconomic forecasting 13
- ICT, Communications and Artificial Intelligence: in honor of Prof Gary Madden 15
- Integrated Energy Forecasting that Scales 17
- Tourism Forecasting - New Methods and Trends (I) 18
- From Forecasting to Policy: In Memory of Peg Young 20
- Forecasting in an uncertain environment 22
- Judgmental forecasting and adjustments 23
- Tourism Forecasting - New Methods and Trends (II) 26
- Predictability, Variable Selection and Causality 28
- Signals and shocks in business cycle 30
- Climate Econometrics 32
- Scientific Forecasting Methods for Climate and other Public Policy Issues 34
- Water Demand Forecasting 36
- Judgmental Forecasting with Structured methods 38
- Best Practices in State Budget Forecasting: Lessons Learned in Washington State 40
- Asset Return Forecasting 42

CONTRIBUTED SESSIONS 46

- Judgment 47
- Government forecasting 49
- Multivariate time series models 51
- State space models 53
- Inflation 55
- Nowcasting 57
- Density & anomaly 59
- Business cycles 61
- Finance (I) 63
- Demography 65
- Demand forecasting (I) 67
- Time series models (I) 70
- Macroeconomic forecasting (I) 72

Probabilistic energy forecasting 74
Supply chain forecasting (I) 77
Finance (II) 79
Tourism and transportation 81
Software and support systems 83
Time series models (II) 85
Macroeconomic forecasting (II) 87
Electricity price forecasting 89
Demand forecasting (II) 91
Combinations & Evaluation 93
Finance (III) 95
Energy forecasting 97
Bayesian models 99
Supply chain forecasting (II) 102
Commodity and consumption forecasting 104
Volatility 106
Climate and environment 108
Big data & data mining 111
Combinations 113
Forecast evaluation 115
Electricity load forecasting 117
Machine learning and neural networks 119
Hierarchies 122
Forecasting 125

WORKSHOPS

Time series forecasting with R

Sunday, 9:00 – 16:00

Nikos Kourentzes (Lancaster University, nikolaos@kourentzes.com)

This workshop will provide a full-day hands-on demonstration of R statistical software as a forecasting tool.

Business Forecasting: Techniques, Application and Best Practices

Sunday, 9:00 – 16:00

Eric Stellwagen (Business Forecast Systems, Inc., estellwagen@forecastpro.com)

This workshop surveys the most commonly implemented business forecasting methods, explains how they work conceptually, reveals their strengths and limitations, and offers best practices for applying them in a business environment.

Neural Networks and Artificial Intelligence in Forecasting

Sunday, 9:00 – 16:00

Hans Zimmermann (Fraunhofer Institute, hans.georg.zimmermann@scs.fraunhofer.de)

We will develop several small (open) systems and compare them to large (closed) systems. The small models are appropriate for demand or load forecasting while the large models fit to financial forecasting, e.g. commodity price forecasting. For such models we have to develop new ways of uncertainty analysis. The full day tutorial is not only about the handling of data, but about a thinking style in system identification and forecasting. The relevance will be shown in industry real world examples.

Large-Scale Automatic Forecasting with SAS Visual Forecasting

Sunday, 9:00 – 12:00

Chip Wells (SAS, chip.wells@sas.com)

This workshop introduces the TSMODEL procedure, released in 2017 and included in SAS Visual Forecasting and SAS Econometrics.

Implementing large-scale forecasting systems

Sunday, 13:00 – 16:00

Michele Trovero (SAS, michele.trovero@sas.com); Michal Kurcewicz (SAS, michal.kurcewicz@sas.com)

This practice-oriented workshop presents challenges in implementing large-scale forecasting systems, and discusses how modern in-memory architecture can address those challenges.

KEYNOTES

Why Economic Forecasting is Harder than Weather Forecasting (and Quantum Physics)

Monday, 8:30 – 9:30; Chair: Len Tashman

David Orrell (Writer and Mathematician, david@postpythagorean.com)

The quantum physicist Niels Bohr is attributed with the saying that “prediction is hard, especially about the future.” Still, economists seem to have more trouble than most. For example, mainstream economists uniformly failed to predict the global financial crisis that began in 2007, and have struggled since to predict things like inflation. Some economists have called for economic forecasting to become more like modern weather forecasting, which has a somewhat better track record at prognostication. This talk will compare the fields of weather and economic forecasting, and ask whether economists can learn from meteorologists – or is something more radical required?

Econometrics for Climate Change

Monday, 15:40 – 16:40; Chair: Gloria Gonzalez-Rivera

David Hendry (University of Oxford, david.hendry@nuffield.ox.ac.uk)

CO2 and other greenhouse gas emissions since the Industrial Revolution have been primarily anthropogenic. While the climate responses to anthropogenic emissions are governed by a range of physical processes, often modelled in process-based numerical models, alternative empirically-based modelling approaches can provide corroboration to assist communication. Further, the interaction of climate with socio-economic aspects inevitably requires empirical evidence. Consequently, Climate Econometrics has emerged as a field using empirical methods to disentangle complex relationships between human actions and climate responses and their associated economic effects. Econometric modelling of climate change, both estimating empirical climate models, as well as the economic impacts of climate change, has already become a sub-discipline, with annual conferences and a global network

The talk will explain how we model the non-stationary climate and economic time series involved and discuss the areas of application such as CO2 emissions, volcanic impacts, hurricane damage and others.

Challenges in Measuring the Modern Economy

Tuesday, 8:30 – 9:30; Chair: David Hendry

Diane Coyle (University of Cambridge, dc700@cam.ac.uk)

Just as 19th century agricultural statistics didn't capture the industrial revolution, our economic statistics today both fail to reflect activity in the modern economy and to properly measure economic welfare. An important part of the problem is the mismeasurement of digital activities and its implications for productivity growth estimates and price deflators, and through these, real output. The effects are wide in scope and reopen an old debate about GDP as a welfare measure. The scale of digital change is now so great that this debate must be revisited. The uncertainty about how digitalisation is affecting the figures means forecasters are not only driving in the fog using a rearview mirror, as always, but also know the road ahead is going in a different direction.

Challenges in Forecasting the Economy

Wednesday, 8:30 – 9:30; Chair: Rich Wobbekind

Azhar Iqbal (Wells Fargo, azhar.iqbal@wellsfargo.com)

An unsophisticated forecaster uses statistics as a drunken man uses lamp-posts - for support rather than for illumination." -- Andrew Lang

In a society with constantly changing human behavior, attempting to predict the path of the economy is like trying to hit a moving target. Over time, consumers are constantly changing their spending behavior while policy makers are regularly crafting new policies whose effects depend on the context in which they were implemented. For instance, the effectiveness of the Fed fund rates (a traditional monetary policy tool) to stimulate the economy raises questions in the current cycle compare to a few decades ago. Unforeseen shocks, such as inventions and natural disasters, also cause disruptions in economic and financial variables overtime. All these factors are potential sources of structural breaks, volatility and non-stationary behavior in economic and financial variables. Furthermore, volatility, non-stationary behavior and structural breaks pose serious challenges to economic and business forecasting.

This presentation addresses these issues (non-stationary behavior, structural breaks and volatility) and provides some practical solutions to build accurate and reliable models to forecast the economy/financial sector.

Judging forecasts and forecasters

Wednesday, 15:40 – 16:40; Chair: Rob Hyndman

Paul Goodwin (University of Bath, p.goodwin@bath.ac.uk)

Forecasting has had a negative press in recent years with surprise election results in the US, UK, and France, and the apparent failures to forecast the financial crisis of 2007-8 and the 40% fall in oil prices in 2014. However, assessments of forecasts are often based on misunderstandings of what forecasts are, confusion between forecast and decisions, inadequate data, and defective recall of past performance.

This talk will explore whether the way forecasts, and their associated performance measures, are presented and communicated contributes to mistaken perceptions amongst forecast users. Point forecasts dominate, but they fail to acknowledge uncertainty. Few users realise that they are, in general, merely a measure of central tendency from a probability distribution and therefore a single realised observation that deviates from this value usually cannot show that the forecast was wrong. Moreover, terms such as 'forecast error' imply incompetence, although filtering the noise from a forecast is a desirable attribute. In addition, hold-out samples are usually –and often of necessity – overly small, while error measures assume that particular loss functions apply, usually without stating this explicitly. Furthermore, published forecasts often neglect to reveal their underlying assumptions.

The talk will argue that we need to place a greater emphasis on probabilistic forecasting in research and practice and to educate users on how to interpret such forecasts. However, large hold-out samples are needed to measure the calibration of such forecasts reliably. Given the problems of assessing forecast performance relative to outcomes, greater weight should be attached to the quality of the forecasting process when judging forecasts - relative to measures of accuracy and calibration. Here the findings of the Principles of Forecasting project can provide guidance –but the principles need to be extended to give greater guidance on probabilistic forecasting. In decision-

making, good and bad decisions are distinguished from good and bad outcomes. The same should be true of forecasting.

PANEL DISCUSSIONS

Econometrics for Climate Change

Monday, 16:40 – 17:30

Panellists:

Claudia Tebaldi, Project scientist, NCAR

Douglas Nychka, Senior Scientist, NCAR

Brian O'Neill, Senior Scientist, NCAR

Jared Carbone, Associate Professor, Colorado School of Mines

Teaching Predictive Analytics and Forecasting

Tuesday, 14:00 – 15:20; Chair: Keith Ord

Panellists:

George Athanopoulos, Monash University

Nikolaos Kourentzes, Lancaster University

Gloria Gonzalez-Rivera, U.C. Riverside

Bob Nau, Duke University

Analytics is a fast-developing topic area for masters and undergraduates alike. It is all too often used as a catch-all heading that embraces much of Operations Research including forecasting. But both predictive analytics and forecasting have distinct educational requirements – are these conflicting or complementary, and can a syllabus be developed that meets the needs of the different disciplines as well as the needs of the students and employers. This panel, all teachers of forecasting, will discuss a number of questions arising from the growth of predictive analytics.

FORECASTING PRACTICE TRACK

Communicating Forecasts to the C-Suite: A Survival Guide

Monday, 9:40 - 11:00; Chair: Michael Gilliland

Todd Tomalak (John Burns Real Estate Consulting, ttomalak@realestateconsulting.com)

Most forecasting practitioners have extensive education and training in the technical skills but end up having to learn the hard way about presenting forecasts to the executive decision makers. This talk will present key tips for explaining the forecasts and gaining buy-in from above, while avoiding the temptation to show off your technical prowess.

The Future of Retail Forecasting

Monday, 11:20 - 12:40; Chair: John Boylan

Stephan Kolassa (SAP Switzerland AG, stephan.kolassa@sap.com)

The retail business is changing at breathtaking speed. Modern technologies, from Artificial Intelligence and Machine Learning, over the Internet of Things, to Social Media, shape an evolution of our everyday shopping experience that was unimaginable just ten or five years ago. In addition, product proliferation is accelerating, as are the possibilities for fulfillment.

Conversely, the retail business is driven by forecasts of consumer demand. There is very little make-to-order or order backlogging - if the customer doesn't find what he wants when he wants it, he will take his business elsewhere. Thus, demand forecasts are indispensable for planning in retail.

We will discuss the specific impact that recent developments in retail have on forecasting. Which existing challenges in retail forecasting will be exacerbated, which will be attenuated, and what new things are waiting just around the corner? Although we will focus mainly on retail, non-retailers will also find aspects that apply to their business, especially since many of the dynamics discussed above also apply to B2B and other markets.

A fully-automated “Reality Check” on your forecasts

Monday, 14:00 - 15:20; Chair: Eric Stellwagen

Allan Gray (End-to-End Analytics, allan@e2eanalytics.com)

Forecasting Managers are often responsible the quality of tens or even hundreds of thousands of forecasts. Especially where the forecasting process has a significant manual element, it is likely that some of these forecasts are completely unreasonable. Unreasonable forecasts can occur for a wide variety of reasons, such as:

- Salespeople either “sandbagging” or being blatantly over-optimistic, depending on their incentives.
- Failure to account for cannibalization among simultaneously promoted products.
- Data entry errors, such as accidental extra zeros or a decimal point in the wrong place.

The dilemma for the forecasting manager is how to detect these flawed forecasts before they are fed to supply chain planning systems, where they can wreak havoc in terms of excessive (or insufficient) material purchases and finished-goods production. The sheer volume of forecasts makes any kind of manual inspection impossible.

This talk will show how to solve the dilemma by conducting a fully-automated "Reality Check" on your forecasts. Using real examples from leading consumer goods companies, we will see how to:

- Use a variety of metrics to find and flag suspicious forecasts.
- Rank suspicious forecasts for review based on their $ impact.
- Show the right visualizations to let forecast stakeholders decide whether forecast is appropriate (and if not, how to change it).

This kind of automated reality check is often the easiest and fastest way to improve overall forecast accuracy. In many cases we have seen several points of improvement in weighted MAPE (MAD/Mean ratio) just by automatically finding and fixing the bad forecasts previously hidden among all the good ones.

Global Forecasting: The International Futures Model

Tuesday, 9:40 - 11:00; Chair: Ana Galvao

David Bohl (University of Denver, david.bohl@du.edu)

International Futures (IFs) is used by national governments, international governmental organizations, non-governmental organizations, and the private sector to provide forward-looking, policy-relevant analysis that helps to frame uncertainty around development trends and evaluate policy trade-offs across many significant development systems. It has played a significant role in multiple U.S. presidential briefs, has been established as a core policy analysis tool for multiple country development strategies, and to explore alternative pathways to achieving the SDGs.

IFs leverages historical data (over 4,000 historical series), identifies and measures trends, and models dynamic relationships to forecast hundreds of variables for 186 countries for every year from 2015 to 2100. It is a tool for thinking about long-term country-specific, regional, national, and global futures. IFs integrates forecasts across different sub-models, including: population, economy, agriculture, education, energy, sociopolitical, international political, environment, technology, infrastructure, and health. The deep interconnection between these sub-models allows IFs to simulate how changes in one system lead to changes across all other systems. As a result, IFs endogenizes more relationships from a wider range of key global systems than any other model in the world.

The M4 Competition in Progress

Tuesday, 11:20 - 12:40; Chair: Kostas Nikolopoulos

Evangelos Spiliotis (National Technical University of Athens, spiliotis@fsu.gr); Spyros Makridakis (Institute For the Future); Vassilios Assimakopoulos (National Technical University of Athens)

The M4 Competition is the continuation of three previous ones organized by Spyros Makridakis whose purpose has been to identify the most accurate forecasting method(s) for different types of predictions. M Competitions have attracted great interest in both the academic literature and among practitioners and have provided objective evidence of the most appropriate way of forecasting various applications of interest. The main purpose of M4, organized jointly by the University of Nicosia and the Forecasting and Strategic Unit of the National Technical University of Athens, is to replicate the results of the previous three Competitions and extend them into two directions. First by increasing the number of series to 100,000, and second by involving additional methods, including Machine Learning ones, to cover the entire range of possibilities. This session will present the preliminary results of the M4 Competition and discuss some initial conclusions from the analysis performed to the huge volumes of data being generated. The definite conclusions of the competition

will be presented later this year and in 2019, utilizing “Big Data” analytics in additional to standard statistical approaches, with the purpose of improving forecasting accuracy and being able to identify the most appropriate forecasting method(s) for various applications.

Building and Deploying Machine Learning Forecasting Models in the Cloud

Wednesday, 9:40 - 11:00; Chair: Stephan Kolassa

Jocelyn Barker (Microsoft, jocelyn.barker@microsoft.com)

In this session, we will review and demonstrate the framework that underlies the machine learning forecasting capabilities used by Microsoft’s finance organization. The forecasting framework enables professional data scientists and forecasters to produce a quality forecast using state-of-the-art machine forecasting techniques with a few lines of code, while also providing fine-grained control for further modeling as well as extensibility for wrapping custom models. The software is designed to support a formalized development process and to operate on large amounts of data. The participants in this session will learn how the framework enables one to compose time series forecasting pipelines, bring in additional predictor data, run learners and compare their accuracies, and deploy models as web services to be consumed easily in domain-specific applications. We will demonstrate in front of the live audience how our software helps forecasters get to a model quickly by eliminating the need for writing large amounts of “glue” code.

A Primer on Probabilistic Forecasts and Forecast Accuracy Measurement

Wednesday, 11:20 - 12:40; Chair: Shaun Snapp

Stefan De Kok (ToolsGroup, sdekok@toolsgroup.com)

The way we traditionally generate forecasts and measure forecast accuracy is generally unhelpful in telling us where to buffer and how much. A probabilistic forecast helps drive greater efficiencies in capacities and inventory, reducing both obsolescence and stock outs, and increasing customer service levels. This type of forecast provides the full range of uncertainty (i.e. probability distributions of outcomes) of future demand. To properly assess its value, and also measure it against traditional forecasts, a new set of metrics is required. In this session probabilistic forecasts will be explained and a new accuracy metric called Total Percentile Error (TPE) metric introduced.

A New Approach to Set Service Level and Safety Stock

Wednesday, 14:00 - 15:20; Chair: Carolyn Allmon

Shaun Snapp (Brightwork Research and Analysis, shaunsnapp@fastmail.com)

We all understand that accurate forecasts are a necessary ingredient for inventory management; however, the problem remains that if filtered through poor procedures for setting inventory targets, accurate forecasts by no means ensure desired inventory performance. It is often taken for granted that companies know what their service levels targets should be; that is, the desired probability that no shortages occur between the time you order more stock of the item and when it arrives on the shelf. However, it is too common that service level targets are set without analysis and validation, resulting in disappointing service, wasteful expenditure, and unnecessary investment in technology. This presentation will discuss the challenges in service level setting and offer new approaches for improving service level management.

INVITED SESSIONS

Macroeconomic forecasting

Monday, 9:40 - 11:00; Chair: Anna Galvao

An Empirical Investigation of Direct and Iterated Multistep Conditional Forecasts

Michael McCracken (Federal Reserve Bank of Saint Louis, michael.w.mccracken@stls.frb.org); Joseph McGillicuddy (Federal Reserve Bank of Saint Louis)

When constructing unconditional point forecasts, both direct- and iterated-multistep (DMS and IMS) approaches are common. However, in the context of producing conditional forecasts, IMS approaches based on vector autoregressions (VAR) are far more common than simpler DMS models. This is despite the fact that there are theoretical reasons to believe that DMS models are more robust to misspecification than are IMS models. In the context of unconditional forecasts, Marcellino, Stock, and Watson (MSW, 2006) investigate the empirical relevance of these theories. In this paper, we extend that work to conditional forecasts. We do so based on linear bivariate and trivariate models estimated using a large dataset of macroeconomic time series. Over comparable samples, our results reinforce those in MSW: the IMS approach is typically a bit better than DMS with significant improvements only at longer horizons. In contrast, when we focus on the Great Moderation sample we find a marked improvement in the DMS approach relative to IMS. The distinction is particularly clear when we forecast nominal rather than real variables where the relative gains can be substantial.

Macroeconomic forecasting with Bayesian time varying parameter VARs

Gianni Amisano (Federal Reserve Board, gianni.amisano1@gmail.com)

Bayesian time varying parameter models have been widely used for forecasting purposes and as exploratory devices. In spite of their obvious overparameterization, these models are capable of reproducing salient features of the data. The main challenge is to use them with many endogenous variables. The problem is that VAR models are over-parameterized and each coefficient is endowed with its own drift term. Therefore the state equation shock covariance matrix becomes huge. In addition, in order to contain the effects of random walk dynamics on coefficients it is important to specify a prior that greatly limits the amount of time variation and to provide a sensible initialization. Usually both tasks are performed by OLS estimation over a training sample but this is problematic to do in large dimensions, since OLS does not work well in these circumstances. In our paper we propose to use an informative Bayesian VAR on the pre-sample to obtain a sensible initialization and calibrate the prior in a conservative way. In addition, we use the VAR Kronecker structure to reduce the number of free parameters in the state equation. We also augment the model with a suitable stochastic volatility specification. In the paper we conduct two experiments, the first one in VARs with 7 and 20 variables. We show that our specification works very well in prediction, both in terms of point and density forecasts, and yields substantial gains especially in periods of recessions and economic turbulence.

Understanding the Common Dynamics in Corporate Events

Necati Tekatli (UNC-Chapel Hill and Elon University, necati.tekatli@icloud.com)

The paper analyzes the common dynamics of the cyclicality in major corporate events and the macro-financial factors that might drive these common dynamics. We perform a thorough statistical examination of the time-series properties of these corporate events. Building on these facts, we develop a novel Bayesian factor analysis method that performs common pattern and common

volatility extraction for five major corporate waves. We further study the macro-financial variables that can predict and explain the extracted common pattern and common volatility. Macro-financial variables explaining this pattern the best are the Fed's monetary policy variables and the business cycle indicators. The Factor Augmented Vector Autoregressive (FAVAR) analyses we conduct reveals that the factor showing the common pattern slows down by about 0.12% within six quarters after the Fed tightens credit through a one standard deviation increase in the policy rate. The common factor reacts faster to Fed's policy changes than most nominal macroeconomic indicators. The paper emphasizes the role of monetary policy, as well as the benefits of factor analysis, in studying the common dynamics of corporate events.

Data Uncertainty and Business Cycles: an Evaluation of the Bank of England's Probabilistic Backcasts

Ana Galvao (University of Warwick, ana.galvao@wbs.ac.uk); James Mitchell (University of Warwick)

Data uncertainty arises from the fact that macroeconomic aggregates are initially released by statistical agencies based on incomplete datasets and are subject to many rounds of data revisions. While the UK statistical office emphasises in their press releases that preliminary GDP estimates will be revised, it is the Bank of England that provides quantitative estimates of this data uncertainty. In this paper we evaluate how data uncertainty varies over time using different empirical evaluations. They include an evaluation of the accuracy of the backcasts in the Bank of England fan charts. We show that data uncertainty increases at the onset of recessions, in particularly for the 2008/2009 recession.

ICT, Communications and Artificial Intelligence: in honor of Prof Gary Madden

Monday, 9:40 - 11:00; Chair: Mohsen Hamoudia

Batch forecasting utilizing ANNs: best practices and ways forward

Artemios-Anargyros Semenoglou (National Technical University of Athens, artemis@fsu.gr); Evangelos Spiliotis (National Technical University of Athens); Vassilios Assimakopoulos (National Technical University of Athens)

Over the last years, an increasing number of researchers and practitioners support the use of Artificial Intelligence (AI) and Machine Learning (ML) in several fields of application. In the area of decision making, Artificial Neural Networks (ANNs) are being used for forecasting, although limited objective evidence is available regarding their relative performance as a standard forecasting tool, i.e. for large data collections of diverse time series. The purpose of this paper is to provide guidelines on how to effectively apply several types of ANNs for batch forecasting business time series. The dataset used to measure the accuracy of the competitors is a large subset of the M3-Competition, including the longest time series available. The types of ANNs discussed are the Multi-Layer Perceptron (MLP), the Simple Recurrent Neural Network (SRNN) and the Gated Recurrent Unit (GRU). For each of them, first, the optimal hyper-parameters (solver, activation function, learning rate and number of iterations) are determined using a predefined architecture for the network. Second, given the optimal hyper-parameters found previously, the best combination of number of input nodes and nodes in the hidden layer is defined. Finally, different preprocessing techniques are examined to further boost forecasting performance and provide a complete guide for time series extrapolation utilizing ANNs. A discussion on the results is also made in an attempt to explain why certain techniques and architectures work best given the nature of the time series examined and suggest potentially fruitful areas for future research.

Forecasting AI Performance

Lawrence Vanston (Technology Futures, Inc., lvanston@tfi.com)

Forecasts of AI performance have been widely discussed in the popular literature. Some of these forecasts violate fundamental forecasting principles and others are basically informal opinion polls of experts. Most people agree the potential impacts of AI are huge. It follows that forecasts of AI progress should be as solid as we can make them. We first critique prior AI performance forecasts. We then presents data, metrics, and trends of performance improvement for several popular AI applications. We discuss the relationships among improvement in computational power, advances in the art of AI, computational complexity, and AI performance. Finally, we discuss how these might relate to key questions regarding the future of AI.

A Cross-Country analysis of ICT Diffusion and Global Competitiveness

Paul Rappoport (Temple University, prapp4@gmail.com); James Alleman (University of Colorado); Andy Banerjee (SSRS)

This paper uses data from the World Economic Forum's Global Competitiveness Index (GCI) to identify drivers of ICT diffusion and investment. The paper looks at institutions, infrastructure, market size, market efficiency, technological readiness and innovation (components of the GCI) and their impact on ICT investment and diffusion. A dynamic panel model is utilized covering 137 countries and the period 2007 - 2017.

A choice-based diffusion model for multi-generation and multi-country data: Empirical analysis for 3G and 4G connection in 25 countries

Mohsen Hamoudia (France Telecom - Orange Labs, mhamoudia@orange.fr); Hyungsoo Lim (Kaist Business School, Seoul); Duk Bin Jun (Kaist Business School, Seoul)

We propose a framework covering multi-country and multi-generation diffusion processes. Our model focuses on choice-based diffusion model and we extend the model to incorporate heterogeneity of country. We decompose the choice probability for generation into two components; the first one is controlled by country-based variable and the other is affected by generation-based variable. This structure allows us to compare country specific differences. For empirical analysis, we estimate the non-linear random effect model for pooled data which contains 3G and 4G connection in 25 countries. Estimation results are better than several individual country models. We apply the correlated two random effects for 3G and 4G to the model. It shows the best performance for estimation period and forecasting period. The methodological contribution of our model is to simply connect both multi-country and multi-generation dimensions. This framework provides valuable information when determining policies for new country or new generation.

Integrated Energy Forecasting that Scales

Monday, 11:20 - 12:40; Chair: Jon Farland

Quantifying Forecast Variances of Updated Forecast Models

Daniel Burgess (Avista Corp., dan.burgess@avistacorp.com)

Forecasts that are updated hourly provide an opportunity to measure how variances increase the longer the forecast window. As intuitive as this seems, dissecting the variance characteristics provides a measure of confidence needed when using forecasts in other applications. The accuracy rate of change, seasonality, and biases can be parameterized so that they can be used as inputs in addition to the forecast. This paper reviews the methodology used to extract the variance characteristics and the corresponding approach on how these characteristics can be applied.

Solving the Utility Load Forecasting Conundrum

Ken Seiden (Navigant Consulting, ken.seiden@navigant.com); Brian Eakin

The utility industry is amid a global paradigm shift impacting not only distribution planning and grid modernization efforts, but also the underlying utility business model itself. Accuracy in long term energy demand forecasting is becoming increasingly more relevant and the proliferation of high-volume and high-frequency data can support new techniques to deliver forecasts that will be held to a higher standard. This paper describes why a traditional econometric/end-use forecast, where distributed energy resources (DER) are simply subtracted from the forecasting models, neither provides a reasonable future representation nor meets utility distribution planning needs. We present an alternative, bottom-up approach that takes advantage of customer characteristic, SCADA, AMI and other readily available data to forecast DER adoption, load shapes, and total loads nodally across the distribution system.

Integrated Energy Forecasting that Scales

Jonathan Farland (TROVE Predictive Data Science, jfarland@trovedata.com); Will Gifford (TROVE Predictive Data Science); Adam Stotz (TROVE Predictive Data Science)

Distribution grid operators at electric utilities are frequently tasked with understanding and managing the loads across a complex topology of devices. This task often requires the generation of accurate, real-time forecasts of amperage for distribution nodes across various hierarchal levels. This process is made more challenging as a significant number of these nodes are essentially ‘blind’ to operators as they lack sufficient historical SCADA telemetry from which predictive models can be trained. However, highly granular information is available from smart meter deployments at the terminal grid nodes and, through aggregation, can be made available to supplement the lack of SCADA telemetry at higher levels. Hourly and sub-hourly loads from smart meters can be aggregated across the hierarchy in a truly bottom-up fashion to provide increased coverage and visibility into loading patterns across the topology. This paper describes an approach to precisely forecast load across a utilities electricity distribution grid by integrating many data sources and algorithms developed by TROVE’s data science team. This approach is currently available as part of the Integrated Load Forecasting solver within the TROVE platform.

Tourism Forecasting - New Methods and Trends (I)

Monday, 11:20 - 12:40; Chair: Haiyan Song

Exploring the predictive ability of LIKES on the Facebook pages of four major city DMOs in Austria

Ulrich Gunter (MODUL University Vienna, ulrich.gunter@modul.ac.at); Irem Onder (MODUL University Vienna); Stefan Gindl (MODUL University Vienna)

The present research investigates the (pseudo) out-of-sample predictive ability of LIKES on the Facebook pages of four major city destination management organizations (DMOs) in Austria for the actual total tourist arrivals (total domestic and total foreign) to these destinations. The cities under scrutiny are Graz, Innsbruck, Salzburg, and Vienna. Data on LIKES were available on a daily basis between 2010M06 and 2017M02 and were retrieved from Facebook using its Graph Application Programming Interface (Graph API), while monthly total tourist arrivals data for the same period were obtained from the TourMIS database. In addition, monthly Google Trends indices for the 'travel' category were retrieved for the four cities as well since Google Trends have become the 'gold standard' regarding web-based indicators in a tourism demand forecasting context.

Due to presumably dynamic nature of the data-generating process and to allow for habit persistence and expectations in consumption, the autoregressive distributed lag (ADL) model class is employed after appropriate data treatment and statistical pre-testing. Taking into account the daily frequency of the original LIKES data and to fine-tune their lag structure, the mixed data sampling (MIDAS) model class is also employed. Representatives of the autoregressive moving average (ARMA), the state-space exponential smoothing (ETS), and the naïve model classes are included in the forecast evaluation exercise as pure time-series benchmarks.

Given the different possible combinations of LIKES and Google Trends in one ADL or MIDAS model, this results in a total of eight rival forecast models. The actual forecast evaluation is carried out in terms of the root mean square error (RMSE) and the mean absolute error (MAE) using expanding (or recursive) estimation windows to best mimic a natural forecasting situation. The forecast horizons h = 1, 2, 3, 6, 12, and 24 months ahead are evaluated separately. While the time-series benchmarks perform best for Graz and Innsbruck across forecast horizons and forecast error measures, the ADL models incorporating only LIKES as well as both LIKES and Google Trends outperform their seven competitors in most cases when forecasting total tourist arrivals to the city of Salzburg. For Vienna, instead, the MIDAS model including both LIKES and Google Trends produces the smallest RMSE and MAE values for most forecast horizons, which is followed by the ADL model with both LIKES and Google Trends.

Interval forecast combination in tourism

Chenguang WU (Sun Yat-Sen University, wucheng@mail.sysu.edu.cn); Gang Li (University of Surrey); Menglin Zhou (University of British Columbia); Xnyao Fan (University of British Columbia)

Forecast combination, as a forecasting technique, produces forecasts by combining the forecasts from different single models. Previous empirical studies have demonstrated that this technique is an effective tool to decrease the risk of forecasting failures and improve the forecasting accuracy (Wong et al. 2007; Shen et al. 2011). Among existing studies on tourism demand forecast combination, majority focused on point forecast combination, and combining methods vary from single average to optimal weighted methods.

Interval forecasts are assumed to be superior over point ones since a point forecast measures central tendency and provides no information to the variability associated with the forecast, whereas an interval forecast is capable of specifying the probability that the future outcome may fall within a stated interval and is more informative to decision-makers. However, the examination of forecast combination techniques on prediction intervals has received little attention. This study aims to examine the performance of interval forecast combination using different combination weighting methods. Particularly, three research questions are focused: first, can combined intervals outperform the intervals produced by single models? Second, is the efficiency of combined interval forecasts sensitive to the weighting method in use? And third, does the performance of interval forecast combination vary across different forecasting horizons?

In this study, four combination techniques are examined: simple average, Bayesian model averaging (BMA) (Garratt et al. 2003), recursive weights (Bao, Lee and Saltoglu, 2007) and Kullback-Leibler information criterion (KLIC) optimal weights (Hall and Mitchell, 2007). Except the simple average method, the above combination techniques have not been applied in the area of tourism forecasting. Since combining intervals directly is with problem due to the improper probability offered to the interval, the combination of interval forecasts of a given probability is achieved by combining density forecasts. The UK inbound tourism demand by its key source markets is chosen as an empirical case of this study to construct combined forecasts, and eight single models (Naïve, structural time series, exponential smooth, autoregressive integrated moving average, vector autoregressive, autoregressive distributed lag, error correction, and time varying parameter models) are used to produce single forecasts for combination. The coverage rates, widths and Winkler scores of intervals are employed to evaluate the forecasting accuracy. The findings of this study demonstrate the usefulness of interval forecasts in tourism practice.

Forecasting Tourism Growth Turning Points

Haiyan Song (The Hong Kong Polytechnic University, haiyan.song@polyu.edu.hk); Shui Ki Wan (Hong Kong Baptist University)

Tourism demand exhibits growth cycles, and it is important to forecast turning points in these growth cycles to minimize risks to destination management. This study estimates logistic models of tourism demand, which are then used to generate both short- and long-term forecasts of this demand growth. The performance of the models is evaluated using the quadratic probability score and hit rates. The results show that the ways in which this information is utilized are crucial to the models’ predictive power. Further, we investigate whether combining probability forecasts can improve predictive accuracy, and find that combination approaches, especially nonlinear combination approaches, are sensitive to the quality of forecasts in the pool. In addition, model screening can improve forecasting performance.

From Forecasting to Policy: In Memory of Peg Young

Monday, 11:20 - 12:40; Chair: Keith Ord

Peg Young: An Appreciation

Keith Ord (Georgetown University, keith.ord@georgetown.edu)

Peg Young: A Wonderful Friend and Colleague

Pamela Texter

Transportation Services and the Economy

Theresa Firestine (US Department of Transportation Bureau of Transportation Statistics, theresa.firestine@dot.gov); Ken Notis (US Department of Transportation Bureau of Transportation Statistics)

Transportation activities have a strong relationship to the economy. The Bureau of Transportation Statistics (BTS) developed the Transportation Services Index (TSI) to measure the volume of freight and passenger transportation services provided monthly by the for-hire transportation sector in the United States. BTS produces three indexes: a freight index, a passenger index, and a combined index. The indexes incorporate monthly data from multiple for-hire transportation modes. Each index shows the month-to-month change in for-hire transportation services. Monthly data on each mode of transportation is seasonally adjusted, then combined into the three indexes. The passenger index is a weighted average of data for passenger aviation, transit, and passenger rail. The freight index is a weighted average of data for trucking, freight rail, waterborne, pipeline, and air freight. The combined index is a weighted average of all these modes. These indexes serve both as multimodal monthly measures of the state of transportation and as indicators of the U.S. economic future. This presentation reviews Peg Young's contribution (former mathematical statistician for BTS) to measuring the relationship between the TSI and the national economy and presents the latest data. Additionally, the presentation includes ongoing work to provide a more timely data point. This work includes forecasting the input data series and then combining to form the TSI.

Forecasting and Marketing Insights from Modelling the Diffusion of Mobile Telephony

Nigel Meade (Imperial College, n.meade@imperial.ac.uk); Towhidul Islam (University of Guelph)

Mobile telephony is a major part of the Information and Communication Technologies (ICT) sector which accounts for 8.2% of the value added and for over 20% of employment in the OECD countries. Diffusion modelling, time series and technological forecasting have all been used for decision-making in this sector, however most research activity (measured by numbers of papers) has occurred in the modelling of diffusion of mobile telephony. After a brief review of forecasting in mobile telephony, we will examine how the use of diffusion models has led to insights into the different diffusion profiles across countries. We will discuss in more detail the segmentation of the international mobile telephony market and the effects of economic wealth on the time to market take-off.

The Policy Implications of Forecasting the Progress of the West African Ebola Epidemic, 2014-2015

Keith Ord (Georgetown University, keith.ord@georgetown.edu); Arthur Getis (San Diego State University)

The goal was to predict the progress and extent of the Ebola epidemic in West Africa, 2014-2015. Our analyses are based on the reported data, and the social, medical, and technological conditions that existed in real time during the epidemic. A modified, classical compartmental model is used to develop variations of Gompertz and logistic-type predictive models A map analysis strongly hints at the existence of an initial rural component of the transmission of the disease followed by an urban component. Our estimates of the toll of the disease, crude as they are, stand in marked contrast to some of the official estimates at the time, which greatly inflated the number of new cases.

Forecasting in an uncertain environment

Monday, 14:00 - 15:20; Chair: Dilek Onkal

The Wisdom of Committees

Neil Ericsson (Federal Reserve Board, ericsson@frb.gov); David Hendry (University of Oxford); Yanki Kalfa (SAIS, Johns Hopkins University); Jaime Marquez (SAIS, Johns Hopkins University)

Since 1992, participants of the Federal Open Market Committee have made twice-yearly forecasts of U.S. inflation, unemployment, and real growth. This paper assesses the value arising from individual forecasts in a committee-based policy process by employing forecast-encompassing tests in conjunction with saturation techniques. We find that the Federal Reserve Bank presidents and the Federal Reserve Board governors bring distinct complementary information about the U.S. economy to the FOMC meetings.

Selecting a Model for Forecasting

Jennifer Castle (University of Oxford, jennifer.castle@magd.ox.ac.uk); Jurgen Doornik (University of Oxford); David Hendry (University of Oxford)

The paper investigates the role of the significance level used to select models for the purpose of forecasting. The significance level used for model selection results in binary in-out decisions to retain variables, and therefore controls both the null retention frequency and the probability of retaining relevant variables. The paper identifies the optimal selection significance level in a bivariate model when there are structural breaks in the form of location shifts at the forecast origin. The results confirm the optimality of the Akaike Information Criterion for forecasting in completely different settings than initially derived. Simulation evidence explores the optimal selection significance level for 1-step ahead forecasts when there are unknown location shifts present under a range of alternative scenarios, using a multipath tree search algorithm, Autometrics, varying the target significance level for the selection of regressors. The results provide some support for a selection significance level of 16%, albeit with many highly non-linear interactions explaining the forecast performance. Correcting the biases due to pre-testing prior to forecasting is also considered in the simulation evidence, and although an operational bias correction formula can improve forecasts in some settings, it does not do so across all experimental designs. The paper recommends not implementing bias correction when selecting models for forecasting when the underlying data generating process is unknown.

Do forecasters target first or later releases of national accounts data?

Michael Clements (ICMA Centre, Reading University, m.p.clements@reading.ac.uk)

We consider whether it is possible to determine if macro forecasters are attempting to forecast first estimates of data, or revised estimates. Our approach requires that data revisions are predictable prior to the first estimate being released. There is some evidence that this condition is met for some series, and that some forecasters put some weight on later estimates for consumers' expenditure and the GDP deflator.

Judgmental forecasting and adjustments

Tuesday, 9:40 - 11:00; Chair: Anna Sroginis; Shari De Baets

The Ripple Effect Caused by Judgmental Forecast Adjustments

Jason Hurley (The University of Sydney, jason.hurley@sydney.edu.au); Behnam Fahimnia (The University of Sydney); Mohsen Reisi (The University of Sydney); Elliot Bendoly (The Ohio State University)

Demand forecasting is essential for businesses to optimally plan and manage many supply chain operations. It is common industry practice for statistically derived forecasts to be judgmentally adjusted based on expert knowledge and/or contextual information not incorporated into the quantitative forecasting model. Evidence suggests that this process can improve forecast accuracy, particularly in the presence of sales promotions, yet very little is known about the impact that it has on the supply chain holistically. In this study, a behavioural experiment is designed to assess the effects that judgmental forecast adjustments have on upstream supply chain operations including production and inventory replenishment decisions. We concurrently investigate the impact of transparent forecast information sharing and holding coordination stock throughout the supply chain as possible moderating factors. The findings of this study assist with bridging the current gap between judgmental forecasting and bullwhip effect streams of research, as well as provide managerial insights for more informed supply chain decision making. An empirical investigation is undertaken at Coca-Cola Amatil, Australia, consisting of observations of operations, discussions with management, and utilising archival data to design an externally valid controlled experiment.

The dynamics of judgmental adjustments

Robert Fildes (Centre for Marketing Analytics and Forecasting, Lancaster University, r.fildes@lancaster.ac.uk); Nikolaos Kourentzes (Centre for Marketing Analytics and Forecasting, Lancaster University)

Judgment plays a central role in forecasting, as statistical forecasts are often modified before informing user decisions. There is strong evidence that judgemental adjustments can be beneficial, yet inconsistent, often harming forecast accuracy. This has motivated research into how to best manage judgmental adjustments to maximise any benefits. However, a limitation of existing work is that there is little understanding of the dynamics of judgemental adjustments: how experts adjust forecasts across time and whether their effectiveness is consistent across forecast horizons. Furthermore, it is unknown whether experts incorporate the same information, and in the same way, for different horizons. This paper addresses these questions by investigating the dynamics of judgmental adjustments made by a UK manufacturer over a forecast horizon of 12 months. Issues include whether the forecasts are optimally updated as new information becomes available and the relationship between adjustments over the different time horizons. Our findings suggest that there is a clear shift in the type of information integrated in the statistical forecasts, particularly across forecast horizons where the magnitude, direction and the relative 'value added' of the adjustments are horizon dependent. This has implications for the design of forecasting processes.

Are people able to distinguish good from bad forecasting models? Effects of model quality on adjustment behaviour and forecasting accuracy

Shari De Baets (Ghent University, shari.debaets@gmail.com); Nigel Harvey (University College London)

Judgmental forecasting is pervasive in business practice (Fildes & Goodwin, 2007), with the most recent number indicating a combination approach of judgment and statistics of 55% and rising (Fildes & Petropoulos, 2015). While most studies have focused on the added value of judgment compared to statistical forecasting, few look at the effect of the statistical model itself on the combination of judgment and statistics. Sole exceptions are Lin and O'Connor (1995) and De Baets & Harvey (2018). Lim and O'Connor (1995) found that the effectiveness of adjustment was dependent on the initial quality of the statistical model. In contrast, De Baets & Harvey (2018) found that the quality did not matter for forecast accuracy. As a follow-up to these studies, we investigate (1) whether people are able to distinguish a 'good' from a 'bad' forecast, (2) whether this model selection outperforms the average, and (3) whether the quality of their model affects the participant's subsequent forecasting behavior (Experiment for (3) planned Spring 2018).

For the first experiment, with 191 participants, we manipulated between-subjects whether participants had to choose between a 'bad' (naive forecast) and a 'good' model (Autoregressive forecast), or between a 'medium' (exponential smoothing) and a good model (AR forecast). We manipulated noise (low, high) between subjects, and series type within-subjects (10 series each of AR(0), AR(0.8) and AR(-0.8)). We compared the Mean Absolute Error resulting from the forecast model choices. Analysis shows a main effect of Series type ($F(2,185) = 27.72$, $p < .001$), with the error of AR(0) series being higher than that of the AR(-0.8) series, which in turn was higher than that of the AR(-0.8) series; a main effect of Noise ($F(1,186) = 24.78$, $p < .001$), such that high noise series had higher error than low noise series; and a main effect of model choice ($F(1,186) = 56.56$, $p < .001$), such that choosing between a bad and good model led to higher error than choosing between a medium and a good model. While the first two main expects are as hypothesized, the latter is in the opposite direction than expected: it should have been easier to distinguish the 'bad' model from the 'good' model than the 'medium' model from the 'good'one. Because the higher error could result from the larger error associated with a wrong choice in case of the naive forecast, we looked at percentage correctly identified models. This average was slightly higher ($F(1,186) = 4.08$, $p = .045$) in the 'bad versus good' condition (Mean = 62.57%) than in the 'medium versus good' condition (Mean = 28.20%), confirming our initial hypothesis that people should be able to better distinguish bad from good forecasts than medium from good forecasts. Moreover, one-sample t-tests showed that, in all conditions, choosing the best model significantly outperformed ($p < .001$) a simple averaging of the models, confirming the findings of Fischer & Harvey (1999). Overall, these results suggest that a judgmental discrimination of forecasting models is indeed possible and provides potential compared to statistical averaging.

Interpreting algorithmic and qualitative information when making judgmental forecast adjustments

Anna Sroginis (Lancaster University, a.sroginis@lancaster.ac.uk); Robert Fildes (Lancaster University Management School); Nikolaos Kourentzes (Centre for Marketing Analytics and Forecasting, Lancaster University)

Despite the continuous improvements in statistical forecasting, human judgment remains essential in business forecasting and demand planning. Typically, forecasters do not solely rely on statistical forecasts, which are obtained from various Forecasting Support Systems (FSS); they also adjust forecasts according to their knowledge, experience and information that is not available to the statistical models. However, we do not have adequate understanding of the adjustment mechanisms, particularly how people use additional information (e.g. special events, promotions, strikes, holidays

etc.) and under which conditions this is beneficial. To investigate this, we conduct experiments that simulate a typical supply chain forecasting process that additionally provides qualitative and model-based information about past and future promotional periods for retail products. Using laboratory experiments, we find that when making adjustments people tend to focus on several anchors: the last promotional uplift, current statistical forecast and contextual statements for the forecasting period. At the same time, participants ignore the past baseline promotional uplifts and domain knowledge about the past promotions. They also discount statistical models with incorporated promotional effects, hence showing lack of trust in algorithms. These results highlight the need for more fundamental understanding of processes behind human adjustments and the reasons for them since it can help to guide forecasters in their tasks and to increase forecast accuracy.

Tourism Forecasting - New Methods and Trends (II)

Tuesday, 9:40 - 11:00; Chair: Haiyan Song

Forecasting International Tourism Demand Using a Spatiotemporal Autoregressive Model

Xiaoying Jiao (University of Surrey); Jason Chen (University of Surrey); Gang Li (University of Surrey, g.li@surrey.ac.uk)

Given the rapid development of the tourism industry and the perishable nature of tourism related goods and services, accurate tourism forecasting becomes increasingly important and attracts much attention from researchers. A range of different models have already been applied to tourism demand forecasting, but none of them has explicitly specified the spatial effect. This study aims to bridge the research gap by developing a spatial temporal autoregressive model to take account of the spillover effect of tourist flows and assessing its performance of forecasting in tourism.

Although spatial models have been applied in tourism demand modelling, this study presents the first attempt in applying a spatiotemporal model in tourism demand forecasting. To be more intuitive, for each destination, a spatial temporal model accounts for the spatial effect exerted by neighbouring destinations countries as an important factor that influences tourist inflows to a specific destination. A spatial temporal autoregressive model can be expressed as follows:

$Y_t = \tau Y_{t-1} + \rho W Y_{t-1} + \mu + \varepsilon$

where Y_t is the matrix expression of N countries' tourist arrivals at time t; W is a spatial weight matrix determined by the distance between destinations; μ is a constant vector term; ε the error term that is assumed to be independently and identically distributed.

This study uses annual data of international tourism arrivals to most European countries for the period of 1995-2016. The data are collected from the United Nations World Tourism Organization. The forecasting performance of the proposed model will be evaluated against three benchmark models including ARIMA, naive, exponential smoothing models. The forecasting accuracy is measured by the mean absolute percentage error (MAPE) and the root mean squared error (RMSE).

Nowcasting and Short-Term Forecasting Tourism Demand Using FaMIDAS Model: the Case of Mainland Chinese Outbound Travel to Hong Kong

Han Liu (Jilin University, hanliu@jlu.edu.cn)

This paper investigates the application of a dynamic factor-based mixed data sampling model, namely FaMIDAS, to predict tourist arrivals from mainland China to Hong Kong and evaluate the forecasting performance of this model using the web search data. MIDAS is a time series approach that is particularly useful for extracting information from high frequency indicators that are used as proxies of the target variables observed at lower frequency. This approach has been proved to be useful for various forecasting applications. The short-term forecasting performance of the FaMIDAS model is evaluated. The results show that the mixed frequency FaMIDAS model estimated using the Baidu index data generate relatively accurate nine-month out-of-sample forecasts of tourist arrivals to China as compared with the forecasts generated by alternative time series models. The contribution of this study is three folds. First, this is the first time a mixed frequency dynamic model is used in tourism forecasting; secondly, the study opens a new avenue for future tourism forecasting research using the search engine data; and thirdly, the empirical results of the study provide useful information for decision-makers in terms of formulating useful strategies for the future tourism development in China and Hong Kong.

Forecasting London Museum Visitation Using Google Trends Data

Katerina Volchek (The Hong Kong Polytechnic University); Anyu Liu (University of Surrey, anyu.liu@surrey.ac.uk); Haiyan Song (The Hong Kong Polytechnic University); Dimitrios Buhalis (Bournemouth University)

This study aims to forecast the monthly demand for the top 10 most visited museums in London. The forecasts are generated by an ARIMAX, an exponential smoothing state space (ESSS), an artificial neutral network (ANN) and a mixed frequency models. In addition, we also add the search queries data from Google Search in the above models except the ESSS model as an explanatory variable. Monthly data from January 2012 to December 2016 are used to estimate the models, based on which the forecasts for 2017 are generated. The actual data for 2017 are used to evaluate the forecasting performance of the various models mentioned above. This study is the first of its kind in museum visitation forecast and should have important policy and managerial implications for the museum sector in London.

Evaluating Probabilistic Forecasts of Global Tourism Demand

Chenguang WU (Sun Yat-Sen University); Zheng Cao (University of Surrey, zheng.cao@surrey.ac.uk); Gang Li (University of Surrey); Meini Tang (Sun Yat-Sen University)

Tourism demand forecasting is vital to tourism business planning and policy making. The current practice is dominated by point forecasting (Wu et al., 2017), which provides only one possible outcome of the future development. However, tourism demand is susceptible to volatile market conditions (Lewis & Pain, 2015; Ringbeck & Pietsch, 2013). Hence, it is considered increasingly important to convey the degree of variability or uncertainty surrounding point forecasts. Whilst interval forecasting is useful, one drawback is that the threshold values are often less intuitive to interpret. Neither do they allow for a full recovery of the forecast probability distribution function (Garratt et al., 2003). An intuitive means of conveying the degree of uncertainties is to use probabilistic forecasting, which represents “a statement of the likelihood of a specified event taking place conditional on the available information” (Garratt et al., 2012, p.145). It allows forecasters to specify the threshold values that define an event of their own interest. Then the probabilities of the specified events can be computed. Hence, probabilistic forecasting poses an appealing tool for tourism researchers to explore any range of outcomes in relation to future tourism demand. This study is the first to provide scenario-based probabilistic forecasts of tourism demand. We collect 1993Q1-2016Q4 data from open sources including IMF’s International Financial Statistics, Balance of Payments statistics, and the respective national statistics offices. A Global Vector Autoregressive (GVAR) model (Dees et al., 2007) is constructed as the underlying model and the probabilistic forecasting is based on the framework proposed by Garratt et al. (2003). The events of our interests concern the development of tourism growth and economic growth. We generate one-step ahead event probabilistic forecasts over 2012Q1-2016Q4 and evaluate the forecasting performance. The results from the study will provide policy makers and tourism businesses an outlook of the potential developments of major tourism markets, taking account of various uncertainties.

Predictability, Variable Selection and Causality

Tuesday, 9:40 - 11:00; Chair: Claudio Antonini

Compression and irreversibility in prediction

John Symons (University of Kansas, johnfsymons@gmail.com)

It is well-known that there are incompressible patterns and sequences that cannot be predicted independently of running a simulation. A range of phenomena are though to have similar properties and are known as weakly emergent. A less well-understood, but related phenomenon in the study of prediction, involves irreversible process. What are the consequences, for example, of the irreversibility of computational models on their predictive power? By irreversibility, we mean the fact that computational models can generally arrive at the same state via many possible sequences of previous states. Thus, while in the natural world, it is generally assumed that physical states have a unique history, representations of those states in a computational model will usually be compatible with more than one possible history in the model.

Prediction is an essential feature of non-arbitrary decision making but predictions at fine grain and coarse grain resolution involve different kinds of reliability. We can predict with relatively high confidence that all currently living cats will be dead in 2200. We cannot predict the date or even the year of my cats death with high degrees of confidence. Similarly, we can predict the weather within the next few days reasonably well, while the weather three weeks from now is very difficult to predict. This talk points to the relationship between the resolution of models, the compressibility of processes that we wish to understand, and the irreversibility of computational models.

I will refer to specific case studies in the talk and will refer to the recent literature on forecasting by way of illustration.

A modelled approximation to the ideal filter for nonstationary time series with application to business cycle fluctuations

Thomas Trimbur (U.S. Census Bureau, thomast2357@gmail.com); Tucker McElroy (U.S. Census Bureau)

This paper develops a representation of the "ideal" band-pass filter for nonstationary time series. The approach ties together frequency domain perspectives that involve periodicity and gain functions with a statistical modelling framework. The approximating filter has several advantages, for instance compared to existing methods, it has a more attractive gain profile that more accurately matches the targeted pass-band of the "ideal" filter when this is the desired gain. Our proposed filter also addresses the sample endpoint problem associated with previous representations and allows for evaluation of the "ideal" filter's implicit assumptions about trend-cycle dynamics. Further, it reveals how filtering errors can result from the indiscriminate use of the "ideal" filter and allows one to quantify such errors. A more flexible approach is to use a modelling framework and to design band-pass filters that adapt to series' properties -- consistent with how the trend and cycle components evolve and relate to each other -- rather than emulating a given gain function. Computer code is freely available for implementing the methodology in a way that avoids the need for an expert operator An application to cyclical fluctuations in macroeconomic time series is presented, showing how plausible and intuitive cycles are estimated via the ideal filter or with an adaptive framework.

The Effect of Ignoring Non-Stationarity In Forecasting Time Series Models: The High-Dimensional Case

Abolfazl Safikhani (Columbia University, as5012@columbia.edu)

Time series models are well--known tools applied in many scientific areas including economics, finance, climate, health sciences, etc. An important attraction of such models is the ability of utilizing them to perform forecasting. There are several developments in constructing univariate and multivariate time series models. A common assumption in such models is stationarity. However, there are many examples of real data sets which exhibit non-stationary behavior. The non-stationary behavior could be due to some parameter changes in the mean, variance or more general in the covariance structure over time. Ignoring such changes might lead to inconsistent estimation of parameters and hence, produces inefficiency in prediction. In this talk, we assume a high--dimensional piece--wise stationary vector autoregressive (VAR) model while allowing changes both in the mean and in the transition matrices. The effect of ignoring structural breaks in the model is investigated both theoretically and empirically. We show that in some special cases, a misspecified model might still produce consistent estimation with a small price which is vanishing by increasing the sample size. Also, comprehensive simulations are conducted to see the empirical effect of ignoring structural breaks in forecasting.

Prediction profiles

John Geweke (University of Washington, john-geweke@uiowa.edu)

The variance of a stationary multivariate time series can be decomposed into the variance of the n-step-ahead prediction and the variance of the n-step-ahead prediction error, for n = 1, 2, … , leading to a prediction profile for the time series. Starting with a univariate time series, the introduction of other time series changes the prediction profile. The paper derives the relationship between these changes and measures of feedback between time series (Geweke, 1982). The prediction profile is directly linked to the volatility of changes in forecasts from one period to the next, leading to diagnostics useful in applied forecasting. These developments are illustrated using economic time series and time series related to climate change.

Signals and shocks in business cycle

Tuesday, 11:20 - 12:40; Chair: Eva Senra

Propagation of shocks in international risk sharing: a panel VAR approach

Pilar Poncela (European Commission, Joint Research Centre, pilar.poncela@ec.europa.eu)

International risk sharing focuses on the cross border channels at work in smoothing income and consumption when a country is hit by an 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 idiosyncratic output growth. In practice, this is hardly the case. In the EU, the capability of the different countries to achieve risk sharing has shown to be different as evidenced in the last Great Recession and subsequent European sovereign debt crisis where risk sharing collapsed in South European countries. We propose a new empirical strategy for measuring international consumption risk sharing channels on a country by country basis, by adopting a Heterogeneous Panel VAR model. We also introduce new risk sharing channels as, for instance, government consumption, which turns to have a dis-smoothing effect. Our empirical analysis includes a set of 21 OECD countries over the time span 1960-2016. Overall, our results demonstrate that the amount of country heterogeneity is significant, ranging from below 10% to over 50%. In addition, dynamics are also quite diverse across countries. Finally, we compare the forecasting performance, in terms of the Root Mean Square Forecasting Error, of the channels from the Panel VAR approach with some benchmark forecasts of the channels. We believe that our approach is of particular interest for policy makers, as it allows identifying the strengths and the weaknesses of the risk sharing mechanisms at work in different countries and allows to deepen our understanding of the role of country governments when facing macroeconomic shocks.

Analyzing the Forecasting Performance of a Trimmed-Mean Stock Price Index

Keith Phillips (Federal Reserve Bank of Dallas, keith.r.phillips@dal.frb.org)

The use of stock price indexes as leading indicators of the U.S. economy has withstood the test of time. Despite the famous quote from Paul Samuelson that “the stock market has predicted nine of the last five recessions,” the Conference Board continues to include it in the U.S. leading index even as they have swapped out other variables in revisions to the index over time. Much of the criticism of stock price indexes as leading indicators of the business cycle rests in their high volatility, which makes it difficult to distinguish the signal from the indicator. In this paper we attempt to reduce the noise in the S&P 500 by each month trimming out the stocks that have largest absolute price movements. The distributions of stock returns are notoriously heavy-tailed, which make robust statistics like trimmed means particularly useful estimates of location. Similar to a trimmed-mean CPI or PCE which are designed to best measure the underlying movement in consumer prices, the trimmed mean S&P 500 is designed to best measure the underlying movement in stock prices. After calculating the index, we then test if this measure does a better job at anticipating turning points in the US economy and in forecasting U.S. GDP growth.

International Propagation of Shocks: A Dynamic Factor Model Analysis Using Survey Forecasts

Kajal Lahiri (University at Albany-State University of New York, klahiri@albany.edu); Yongchen Zhao (Towson University)

This paper attempts to quantify propagation of shocks across G7 and major Asia-Pacific countries using multi-horizon fixed-target forecasts of real GDP growth from 1995 to 2017. We show that if the forecasts are efficient in the long-run, results obtained from the forecasts are identical to those obtained from the actuals. Using a factor structural VAR model, we measure global business cycle connectedness and study the transmission of country-specific shocks as well as common international shocks. Our results suggest strong convergence of business cycles within the group of industrialized countries and the group of developing economies during non-crisis periods. However, transmission of shocks during crisis periods depends on the nature and origin of the crisis. In particular, we find increased decoupling between the industrialized countries and Asian developing economies after the 2008 recession.

Effects of the revisions in signal extraction when predicting the business cycle

Juan Bógalo (Universidad de Alcalá, juan.bogalo@telefonica.net); Pilar Poncela (European Commission, Joint Research Centre); Eva Senra (Universidad de Alcalá)

In this paper we analyze the effect of real time data revisions when estimating the business cycle. We compare two procedures, Circulant Singular Spectral Analysis, CSSA, which is a non-parametric technique in which we can extract components associated to desired frequencies, and the widely used Tramo-Seats based on ARIMA models. By a set of simulations we show that the magnitude of the revisions produced by CSSA is lower than those produced by Tramo-Seats. This would affect the forecasts of the cycle and the forecasts of observed time series when using unobserved component models.

Climate Econometrics

Tuesday, 11:20 - 12:40; Chair: Dilek Onkal

Beyond RCP8.5: Statistical Emulation of a Big Climate Model Using Quasi-Representative Concentration Pathways and Projecting Mortality Risk from Excessive Heat

William A. Brock (University of Wisconsin); J. Isaac Miller (University of Missouri, millerjisaac@missouri.edu)

Damage projections from climate change generally employ either detailed output from general circulation models (GCMs) under one of only four representative concentration pathways (RCPs) for CO2 and other emitted gases or else heavily aggregated output from simple statistical models. The complexity of the former prohibits comparisons of projections under alternative concentration pathways, while the simplicity of the latter - even if appropriately physically based - harms their credibility. Statistical emulation and pattern scaling aim to bridge the gap. We propose an approximate parameterization to the RCPs that allows the policymaker to consider intermediate pathways, and we employ a statistical emulator to downscale detailed GCM output under the RCPs to these pathways. We then employ spatially heterogeneous dose-response functions in order to project disaggregated relative mortality risk from excessive heat exposure throughout the Northern Hemisphere at the end of this century.

A False Sense of Security: The Impact of Forecast Uncertainty on Hurricane Damages

Andrew Martinez (University of Oxford, andrew.martinez@economics.ox.ac.uk)

The impact of uncertainty on economic activity is of growing interest. However, it is difficult to identify the effect of forecast uncertainty since it can both be driven by and drive economic outcomes. We use a new approach to identify the impact of forecast uncertainty by looking at damages associated with hurricane strikes where feedback is less of a concern. We develop a simple theoretical framework in which micro uncertainty can have a macroeconomic impact when operating through a short-term adaptation and beliefs channel. In order to test this, we embed measures of ex-post forecast uncertainty into a general empirical model of hurricane damages and then use automatic general-to-specific model selection methods to simplify. We find evidence of a positive relationship between ex-post forecast uncertainty and hurricane damages. Our results show that a 1 percent increase in the ex-post forecast error shock is associated with a 0.25 - 0.5 percent increase in damages on average. This translates roughly into a fall in local economic growth by 0.11 - 0.23 percentage points.

Analyzing Differences between Scenarios

David Hendry (University of Oxford, david.hendry@nuffield.ox.ac.uk); Felix Pretis (University of Oxford)

Comparisons between alternative scenarios are used in many disciplines from macroeconomics to climate science to help with planning future responses. Differences between scenario paths are often interpreted as signifying likely differences between outcomes that would materialise in reality. However, even when using correctly specified statistical models of the in-sample data generation process (DGP), itself an unlikely possibility in most observational disciplines, additional conditions are needed to sustain inferences about differences between scenario paths. Moreover, the calculation of uncertainties around scenario differences raises difficult issues since the underlying in-sample

distributions are identical for both `potential' outcomes, and the reported paths are often deterministic functions. Here we consider the conditions under which scenario projections can be informative, even if the underlying models are mis-specified. We show that tests for invariance based on the automatic detection of structural breaks can be used to identify in-sample invariance of models, which can subsequently provide information about their likely constancy in projected scenarios.

Scientific Forecasting Methods for Climate and other Public Policy Issues

Tuesday, 14:00 - 15:20; Chair: Kesten Green

The physical science of climate: A complicated and poorly understood discipline

David Legates (University of Delaware, legates@udel.edu)

Making scientific forecasts about changes in the Earth's climate requires knowledge about the state of climate science: what is known, and what is not. I will address some of the diverse assumptions made about the causes of climate change, discussing the estimates of effect sizes and strength of evidence. Most climate change scenarios are based on either (1) projecting recent trends into the future, or (2) simulations using models of assumptions about climate dynamics. I assess the validity of those approaches by examining the reliability of historical climate observations and of the physical science assumptions embedded in climate models.

Do Forecasters of Dangerous Manmade Global Warming Follow the Science?

Scott Armstrong (University of Pennsylvania, jscott@upenn.edu); Kesten Green (University of South Australia (School of Commerce and Ehrenberg-Bass Institute))

Problem: Climate change policies require scientific long-range forecasts of (1) a consistent and substantial long-term trend in global mean temperatures, (2) effects of the changes in temperature, and (3) costs and benefits of alternative policies and regulations including taking no action.

Methods: We rated the extent to which the procedures used to forecast dangerous manmade global warming comply with eight necessary criteria of the scientific method. We did so using a checklist based on a definition of the scientific method that drew on the writings of eminent scientists. Findings: We report on the extent to which forecasts of dangerous manmade global warming are scientific forecasts. We then assess the ex ante predictive validity of those forecasts—in particular, the UN Intergovernmental Panel on Climate Change's +3°C per century projection—relative to the accuracy of forecasting procedures that are consistent with the scientific method.

Originality: This is the first paper to provide an assessment of the extent to which procedures used to forecast dangerous manmade global warming comply with the scientific method. It extends our earlier work in assessing whether the procedures used to forecast dangerous global warming complied with evidence-based forecasting principles.

Usefulness: Ratings of the extent to which forecasting methods provide forecasts that are claimed to justify current and proposed climate policies and regulations comply with the scientific method can help political leaders and voters to make better decisions.

Testing the Predictive Validity of Multiple Regression Analysis

Kesten Green (University of South Australia (School of Commerce and Ehrenberg-Bass Institute), kesten.green@unisa.edu.au); Andreas Graefe (Macromedia University); Scott Armstrong (University of Pennsylvania)

Problem: Multiple regression analysis (MRA) is commonly used to develop forecasting models that inform policy and decision making, but the technique does not appear to have been validated for that purpose.

Methods: The predictive validity of published least squares MRA models is tested against naive benchmarks, alternative methods that are either plausible or commonly used, and evidence-based

forecasting methods. The out-of-sample errors of forecasts from the MRA models are compared with the errors of forecasts from models developed from the same data on the basis of cumulative relative absolute error (CumRAE), and the unscaled mean bounded relative absolute error (UMBRAE).

Findings: Results of the analysis are presented in the order of accuracy expectations from prior research, and from surveys of forecasting experts and econometricians.

Originality: This paper presents the most comprehensive test to date of the predictive validity of MRA models relative to models derived using diverse alternative methods.

Usefulness: The findings of this research will be useful whether they turn out to support or reject the use of MRA models for important policy and decision-making tasks. Validation of MRA for forecasting would provide a stronger argument for the use of the method than is currently available, while the opposite finding would identify opportunities to improve forecast accuracy and hence decisions.

Water Demand Forecasting

Tuesday, 14:00 - 15:20; Chair: Chris Meenan

One Forecast to Bind Them All

Shawn Stoddard (Truckee Meadows Water Authority, sstoddard@tmwa.com)

The Truckee Meadows Water Authority (TMWA) is in Washoe County, Nevada and serves approximately 401,000 persons or 90% of the county population which resides in cities of Reno and Sparks. The primary source of supply is the Truckee River that is also augmented with many groundwater wells located throughout the Truckee Meadows. Under average conditions, TMWA's water supply is 85% surface water and 15% groundwater. The local economy within the Truckee Meadows is booming. Currently, TMWA is growing at an average rate of 177 new, active water services per month. While new growth is tied to the transfer of water rights from developers to the utility, many stakeholder groups, agencies, and policy makers are questioning sustainable growth in the region and are looking closely at the published forecasting models.

There are several unique forecasting requirements centered around water demand, population, and TMWA's revenue requirements. The first requirement is a 20-year population projection with annual total water demand for TMWA's Water Resource Plan. The next requirement contains detailed water use coefficients and water resource projections for TMWA's Water Facility Plan. The third requirement is a detailed 10-year operating revenue projection at a monthly time scale for TMWA's Funding Plan. To satisfy these needs, TMWA has developed a comprehensive approach to water demand forecasting.

This presentation shows how TMWA developed a Vector Autoregression (VAR) model to estimate active water services, by customer class, and extended the VAR model to include water use to build a water demand model. This water demand model is further extended by including a rate component to develop an operating revenue model. Thereby, in a single effort, water resource planning, engineering facility planning, and finance revenue planning are bound together within one model - thus, one forecast to bind them all.

Understanding and Planning for the Uncertainties of Future Water Use

Ray Quay (Decision Center for a Desert City, Arizona State University, ray.quay@asu.edu); David Sampson (Decision Center for a Desert City, Arizona State University)

As climate change and drought have become a concern for many water utilities, much attention has been focused on understanding the factors of water supply uncertainty and their implications for water resource management. Unfortunately, less attention has been paid to the factors of uncertainty associated with demand and their impact on our ability to forecast or estimate future water demands. Over the past 50 years simple water demand forecasts based on estimates of population growth and of water efficiency, most notably gallons per capita per day, were adequate for most financial and water resource planning. However, over the last decade, increasing uncertainty in the factors driving demand have resulted in inaccurate forecasts that have created financing and system operations issues. For example, over the last 15 years most utilities have seen long term trends of water demand and wastewater production increasing at a pace equal to or greater than population growth suddenly shift to a trend of declines in unit water demand and wastewater production sufficient to keep total water demand and sewer flows flat or declining, even though communities were experiencing significant population growth. One result has been water planners over estimating future demand which led to an over building of infrastructure and over

estimating of water revenues. The factors that generated this shift are complicated and are still not necessarily stable. The uncertainties of this trend combined with the emerging uncertainty about the future of other factors affecting water demand will make water demand forecasting even more difficult creating new issues for traditional water resource, infrastructure, and financial planning. This presentation will examine the sources of future uncertainty for estimating short term and long term water demand and will discuss a new approach to planning under high uncertainty which moves away from forecasting to embrace anticipatory governance. Current research that is exploring these factors and their uncertainty will be presented, including the infrastructure of indoor and outdoor water use, customer behavior, land use and water interactions, and the potential impact of drought and climate on water use. Examples of utilities using the new planning paradigm of anticipatory governance will be presented, including the use of exploratory scenario analysis rather than best guess forecasting and the use of adaptive management to govern water planning and implementation.

Lessons Learned: 20 Years Making Daily Forecasts at the Las Vegas Valley Water District 1998 to 2018

Chris Meenan (Las Vegas Valley Water District, chris.meenan@lvvwd.com); William Murray (Las Vegas Valley Water District); Frank Kalenits (Las Vegas Valley Water District)

Most water demand forecasting papers and presentations focus on the process of building models, making forecasts, and determining appropriate error metrics. This presentation will provide a contextualized summary of experiences and lessons learned from the Las Vegas Valley Water District's (LVVWD) two decades of forecasting daily water demands. It will present the history and evolution of the Las Vegas water utility's forecasting process as customer needs have evolved. The LVVWD's forecasting system began as a simple set of regression models with lagged variables and weather forecasts, which progressed to a more complex combination of neural net and regression models before finally evolving into a highly-automated system of time series models designed to incorporate domain knowledge.

The primary lessons learned center around the importance of trust and clear communications between stakeholders. The forecaster must have an accurate understanding of what the customer is trying to accomplish, what constitutes success or failure for the customer, and what characteristics of the forecast create issues for the customer. Additionally, it is essential that forecasters support the success of the customer's process and always be perceived as part of the customer's team.

Some of the technical lessons learned include determining the appropriate error measure consistent with the customer's process and desired outcomes; the best way to track historical forecasts; the role of human oversight and intervention in forecasts; the usefulness of SAS and Excel; the impact of how the organization stores and makes data available; and ultimately, balancing timeliness, accuracy and complexity associated with the process. Where a record of historical forecasts is available, a summary of their accuracy and comparison to a random walk model will be made.

Judgmental Forecasting with Structured methods

Wednesday, 9:40 - 11:00; Chair: Kostas Nikolopoulos

Forecasting for Social Good: Relative performance of methods for forecasting major projects

Konstantia Litsiou (Salford University, k.litsiou@edu.salford.ac.uk); Yiannis Polychronakis (Salford University); Kostas Nikolopoulos (Bangor University)

Forecasting for Social good, most notably the socio-economic impact of major high-impact projects - like Olympic games or space exploration -is a very difficult but also extremely important task; not only for the resources allocated in such project but predominantly for the great expectations around them. This study evaluates the performances of Unaided Judgment (UJ), Structured Analogies (SA) and semi-Structured Analogies (s-SA) as well as Interaction Groups (IG) in forecasting the impact of such projects. The empirical evidence reveals that the use of s-SA Analogy leads to accuracy improvement compared to UJ. This improvement in accuracy is greater when introducing pooling of analogies through interaction in IG. A smaller scale experiment run to compare Delphi with IGs with inconclusive results.

Long-Term Economic Forecasting with Structured Analogies and Interaction Groups: The Case Study of Saudi Arabia

Waleed Alghassab (Bangor University, abp6d3@bangor.ac.uk); Kostas Nikolopoulos (Bangor University)

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

"Structured Superforecasting" Bringing together state of the art techniques in Judgmental forecasting.

Ilias Katsagounos (University of Peloponnese); Dimitrios Thomakos (University of Peloponnese); Kostas Nikolopoulos (Bangor University, k.nikolopoulos@bangor.ac.uk)

The success of the Good Judgmental Project in harnessing the power of superforecasting naturally leads to the question as to how one can implement that approach on a smaller scale with more limited resources and fewer participants. Small(er) corporate environments and SME-type decision structures are prime examples where the modified superforecasting approach can be used. In this research we focus on a hybrid approach of judgmental forecasting on special events where we combine training of superforecasters-to-be via the concept of a modified version of structured analogies, a staple of judgmental forecasting in the literature. We call the resulting approach structured superforecasting and illustrate its efficacy over samples of participants from the wider

public sector and the academic community. In particular, with a proper experimental design that includes a training and a control group, we apply the above methodology and compare performances. Our analysis of the results utilizes, beyond standard measurement concepts of judgmental forecasting, the methodology of stochastic dominance (SD) to evaluate the performance of participants -- for the first time to the best of our knowledge in this strand of the literature. The use of the SD concept is important for two reasons: first, it allows a complete/better view of outperformance compared to more traditional statistics and, second, it provides compelling visuals for the results across individual questions and across any sample split we wish. We find that, across most questions employed, analogies trained participants outperform the control group. This is an important and practical result which we validate for the first time. The implications of extending this research in other environments and different samples is obvious: expending effort and resources in training on analogies can super-charge super forecasting and thus tapping into the wisdom of the crowds does not need larger crowds but, importantly, smarter (by training) ones.

Best Practices in State Budget Forecasting: Lessons Learned in Washington State

Wednesday, 11:20 - 12:40; Chair: Elaine Deschamps

Collaborative and Consensus-Driven Government Budget Forecasting: A 20 Year Look Back

Elaine Deschamps (Washington State Caseload Forecast Council, elaine.deschamps@cfc.wa.gov)

"It's about time we break through organizational bias and balance judgmental forecasting with real-time data to create unbiased, transparent, and accurate forecasts," we thought twenty years ago when the statute creating the Washington State Caseload Forecast Council (CFC) was rolled out toward that purpose. There were bumps along the way and lessons learned. The CFC is an independent state agency that forecasts the demand for entitlement caseloads that drive 80 percent of the Washington state budget. In this session and in my segment, we'll discuss our challenges and lessons learned in forecasting: data mining, forecasting techniques, consensus forecasting, and translating legislative policy changes into the forecast. In my segment, I'll pay special attention to the importance of the CFC as an independent agency and the CFC's technical workgroup process, which is how we navigate outliers in the forecast data, agree upon our official forecasts, and weave the power of judgmental forecasting in with statistical models.

Forecasting US State and Federal Inmate Populations

Gongwei Chen (Caseload Forecast Council, gwchenn@live.com)

The United States has one of the highest ratios of inmates per capita among all countries. Justice and public safety demand the incarceration of certain criminals. However, there is considerable financial and social cost associated with mass incarceration, and building prisons requires years of advance planning and construction. Thus, it is of particular interest for policy makers to understand the causes of past inmate population growth and to have reliable future inmate population projections. This study reviews the drivers of the inmate population run-up in the 80's and 90's, and the reasons behind the recent decline. We explore the main types of models that are used in the inmate population forecasting and discuss the pros and cons of each model. Based on our experience in forecasting the Washington state's inmate population, we also discuss the unique challenges to forecasting inmate populations that are outside of the realm of statistics or mathematics, such as frequent legislative changes, discretion of police, prosecutors, and judges, and interaction between the inmate population level and legislators' and voters' choices.

A Table Driven, Module Based, Portable, Parallel, and Integrated Forecasting Application

Shidong Zhang (Washington State Caseload Forecast Council, shidong.zhang@cfc.wa.gov)

The Washington State Caseload Forecast Council (CFC) is required to finish all caseload forecasts in a short period of time which makes it necessary to develop an efficient forecasting system. The caseload forecasting application used in both medical assistance and some public assistance programs is a table driven and module based system designed to produce draft forecasts and various research datasets in a few hours after corresponding agencies update their administrative databases.

The features of the forecasting application:

Table driven: all metadata, system control parameters, forecasting model parameters, and forecast step parameters are stored in relational tables or independent tables. Changing metadata and system control parameters can change the whole forecasting structure to fit the forecasting needs of different programs.

Module based: the application consists of sequential related but independent modules on a variety of platforms. More functions can be easily incorporated by adding additional modules to allow unlimited expanding potentials.

Portable: the application's core can be easily deployed to different platforms and produce identical results.

Parallel: in one deployment, the application can produce many forecasts of different forecast cycles and tracking cycles, produce many independent forecast versions within one cycle, and freely and easily switch the current working cycle and version.

Integrated: all caseload forecast-related activities can be incorporated in this forecasting application.

Asset Return Forecasting

Wednesday, 14:00 - 15:20; Chair: David Rapach

What Firm Characteristics Drive US Stock Returns?

Yufeng Han (University of North Carolina at Charlotte); Ai He (Emory University); David Rapach (St. Louis University); Guofu Zhou (Washington University in St. Louis, zhou@wustl.edu)

Green, Hand, and Zhang (2017) conclude that twelve firm characteristics determine US stock returns prior to 2003, while only two characteristics affect returns after that. Instead of using unconstrained linear regressions, which are susceptible to overfitting, we apply robust forecast combination methods to the cross section of stock returns. We find that a large set of firm characteristics matter collectively before and after 2003. A spread portfolio formed on our method continues to generate significant economic gains after 2003. Machine learning approaches further verify our findings.

Return Predictability and Market-Timing: A One-Month Model

Blair Hull (Hull Investments); Xiao Qiao (SummerHaven Investment Management, xiaoqiao10@gmail.com); Petra Bakosova (Hull Investments)

We propose a one-month market-timing model constructed from 15 diverse variables. We use weighted least squares with stepwise variable selection to build a predictive model for the one-month-ahead market excess returns. From our statistical model, we transform our forecasts into investable positions to build a market-timing strategy. From 2003 to 2017, our strategy results in 16.6% annual returns with a 0.92 Sharpe ratio and a 20.3% maximum drawdown, whereas the S&P 500 has annual returns of 10%, a 0.46 Sharpe ratio, and a maximum drawdown of 55.2%. When our one-month model is used in conjunction with Hull and Qiao's (2017) six-month model, the Sharpe ratio of the combined strategy exceeds the individual model Sharpe ratios. The combined model has 15% annual returns, a Sharpe ratio of 1.12, and a maximum drawdown of 14%. We publish forecasts from our one-month model in our Daily Report.

Using copula function for assets multivariate distribution in portfolio optimization problem

Nataliya Pivnitskaya (National Research University Higher School of Economics, pivnitskaya.natali@gmail.com); Tamara Teplova (National Research University Higher School of Economics); Evgeniia Mikova (National Research University Higher School of Economics)

There is accumulated evidence that the most of stock returns are not distributed by normal distribution due to strong skewness and heavy tails (leptokurtic). The anomalous events of past years like the world financial crisis of 2008-2009 and European banking crisis of 2011 imposed on investors the necessary to reconsider their approach to risk estimation. These events make the precedents for observing the heavy tails. The asset returns are characterized by a dependence in the tails of the distribution: prices of financial assets are more likely to fall significantly in the same time than to grow. Thus, the current portfolio carries higher downside risk than many investors supposed. The modeling of this dependence structure can be carried out by constructing copulas functions that allow extracting the dependence structure from the joint distribution with the desired marginal distributions and the type of dependence structure.

The use of copula functions has found application in various areas, and recently there has been an increase in interest in modeling multivariate distributions in the task of optimizing the investment portfolio. The aim of this paper is to identify the optimal copula relationship between assets in the

portfolio consisting of MSCI Equity Indexes and determining the optimal weights of portfolio assets. The methodology of the study is based on the use of ARMA-GARCH models, copula functions, the Monte Carlo method, the nonparametric estimation of CVaR, the method of determining optimal portfolios according to the classical Markowitz theory.

Our work is further research on Sak et al (2010) and Autchariyapanitkul, Chanaim, Sriboonchitta (2015). These authors demonstrated the use of multidimensional Student's copula function that can take into account the tail dependence for modeling the risk structure in portfolio analysis.

We constructed optimal portfolios with different investment periods, using the copula function to estimate the multivariate distribution. For including autoregression and volatility clustering effects observed in market indexes we use the estimated parameters of the ARMA-GARCH model. There was revealed that the Student's copula from the family of elliptical copulas better than others describes the structure dependence in the market indexes. We simulate asset returns on 1 month and 6 months' time horizons based on the Student's copula whereas weights were randomly generated. Portfolio optimization was carried out according to the CVaR at different significance levels: 95% and 99%. We take the Markowitz optimal portfolio, which is based on the determination of structure between assets through a linear Pearson correlation coefficient, as a benchmark for evaluating the benefits of investing with the help of copula functions. We conduct testing the simulated portfolios on a control sample, the results were compared with those obtained in the classical MV optimization. The comparison was made using the risk-adjusted values. In our study, we found how chosen significance levels in the nonparametric estimation of CVaR can affect the results and answered what is the strengths of using copula functions at different time intervals.

Forecasting for social good

Wednesday, 14:00 - 15:20; Chair: Bahmna Rostami-Tabar; Tao Hong; Michael Porter

Forecasting Outcomes of the Refugee Crisis with Brainstormed Analytic Scenarios

Dilek Onkal (University of Bradford, dilekon@gmail.com); Mandeep Dhami (Middlesex University); Lars Wicke (London School of Economics and Political Science); Ian Belton (University of Strathclyde)

Forecasting with scenarios offers valuable insights for policy planning and decision-making by depicting alternative storylines about plausible futures. In an increasingly volatile world, these scenarios may be highly effective in stimulating foresight and enhancing visions of the future, thus constituting requisite tools for those informing policy-makers. This study focuses on the effectiveness of brainstormed analytic scenarios to investigate forecasts pertaining to a very important and real issue: the Syrian refugee crisis. We discuss the findings in the context of forecast quality and the training repercussions for strategic analysts.

Comparing forecasts of psycho-social anxiety, depression, and post-traumatic stress in Ukraine following the Chornobyl nuclear disaster

Robert Yaffee (New York University, ray1@nyu.edu); Monnie Mcgee (Southern Methodist University)

Background: In the 25-year review of the effects of the Chornobyl nuclear disaster, the mental health impact was found to be the largest public health consequence of the accident for Ukraine. However, due to various potentially confounding crises (e.g., Russian gas cut-offs to Ukraine in 2006 and 2009 plus the Great Recession), forecasting the intensity of psychosocial impact is difficult.

Objective: Our objective was to forecast anxiety, depression, and post-traumatic stress of the Chornobyl nuclear accident from 2006 to 2010. These forecasts were potentially confounded with other events that could also have a psychosocial impact. We compare forecasts over two horizons: one prior to, and the other, bracketing those events.

Methods: We conducted a survey of 702 residents of Kiev and Zhytomyr oblasts. We employed random selection of telephone numbers to obtain a representative sample of the Ukrainian residents of the oblasts. Interviews were conducted with willing respondents about their experiences and the results compiled for analysis.

Analysis: We compare the results of two different models to forecast results with ex post forecasts from 1998 through 2005, and ex ante from 2006 through 2010. The first model is a multivariate state space model with a common local level. The second model is a vector autoregressive model (VAR) with exogenous variables.

Implications: In modeling these post-nuclear disaster psycho-social sequelae, we demonstrate circumvention of potentially confounding crises by prior estimation termination and scenario forecasting for emergency medical analysis.

Forecasting in the emergency department

Tao Hong (University of North Carolina at Charlotte, hongtao01@gmail.com)

Emergency departments (EDs) are facilities specialized in offering medical services, including diagnosis and treatment of acute illness, to patients with urgent conditions. The heavy demand has made ED overcrowding a worldwide problem that prevents patients in need of emergency care from getting timely access to ED resources. ED patient arrival forecasting plays a critical role for designing

plans to avoid ED overcrowding. This presentation will introduce the ED forecasting problem to the forecasting community through a comprehensive review of the ED forecasting literature.

CONTRIBUTED SESSIONS

Judgment

Monday, 9:40 - 11:00; Chair: Roy Batchelor

Winning the Lottery, Mispredicting Utility: The Story of Unhappy Lottery Winners

Arie Sherman (Ruppin Academic Center, arie.sherman@gmail.com); Tal Shavit (The College of Management Academic Studies.); Guy Barokas (Ruppin Academic Center)

Recent literature suggests that individuals mispredict utility, and therefore make biased decisions. Specifically, some studies demonstrate that when individuals face change, they underestimate the future utility of consumption that satisfies intrinsic needs compared to consumption that satisfies extrinsic desires. We suggest a dynamic maximization model that takes into account the asymmetric adaptation to extrinsic and intrinsic attributes in order to explain the conditions under which exogenous monetary shock reduces wellbeing. Winning a lottery is an example of wealth shock that can change traditional preferences for extrinsic and intrinsic attributes. If the change in preferences leads the winner to reduce the amount of time and effort he or she allocates to preserving his or her stock of hedonic capital, his or her traditional hedonic capital might be destroyed. In turn, this leads to a process of depreciation that destroys the winner's means for non-pecuniary utility. As a result, life after wealth shock can be misery, with the lottery winners no happier after winning than they were before.

Generally higher but not generally better: How subjective expertise affects estimation accuracy

Laura Rebecca Rettig (University of Muenster, laura.rettig@wiwi.uni-muenster.de); Thomas Langer (University of Muenster)

In a laboratory experiment, we investigate the effect of subjective expertise (i.e. confidence intervals and competence) on accuracy. The experimental task was to estimate quantities on a bounded [0;100] scale with either high or low true values. In accordance with previous literature, we find that estimates tend towards the midpoint of the scale. We hypothesize that this is due to naive choice behaviour and lessens with expertise. Accordingly, more confident individuals in our study make more extreme, albeit not more accurate estimates. Competence however has a significantly positive effect on accuracy. Surprisingly, further analysis reveals that the magnitude of the true value is crucial for the influence of competence on estimation accuracy. The positive effect stems entirely from high true value questions. For questions with low true values, higher competence even shifts the estimate closer to the midpoint. In sum, our data suggest that higher subjective competence does not lead to generally better, but generally higher estimates. Applying our finding to a mathematical group setting, we obtain an improvement in group accuracy of 38% by asking for an identical estimation in a high vs. low true value question format. We further highlight implications for the assessment of overconfidence as more competent individuals are also significantly more (less) likely to be well calibrated for high (low) true value questions. We offer two explanations for our finding: First, the underlying mechanism might be similar to pseudodiagnosticity and confirmation bias, in that the specific question framing leads to biased information retrieval. Second, estimation magnitude might be an expression of individual confidence. Our research has practical relevance for decision making in a group and expert setting.

Anchoring in Share Price Charts

Roy Batchelor (Cass Business School, City, University of London, r.a.batchelor@city.ac.uk)

Forecasting the movements of prices in stock markets is challenging, not least because these prices are averages of the forecasts of market participants, and there is no agreed methodology for making forecasts. Some traders ("fundamentalists") react to news about events likely to affect company earnings, others ("technical analysts" - the focus of this study) look for patterns in charts of share prices.

The unpredictability of news, and the potentially complex interactions between groups of traders, mean that prices in financial markets do not show consistent short-term trends or cycles. The forecasting challenge is very different from that faced by subjects in earlier experimental studies on chart-based judgmental forecasting, where data are lower frequency, drawn from economic and business environments, and have potentially identifiable trend and seasonal features. There has been only limited research on how intuitive chart-based forecasts are generated from charts without such features.

This paper starts to fill this gap through a series of experiments testing for the presence of anchoring effects in forecasts made from share price charts. Specifically, we test whether judgmental forecasts are unduly influenced by salient features of visualisations used by technical analysts, and cluster around past peaks and troughs in price, Fibonacci retracements and projections from past price ranges, arbitrary grid lines, and realistic and arbitrary support and resistance lines overlaid on the main price chart. We also look at whether there are differences in the forecasts made by trained and untrained subjects.

Government forecasting

Monday, 9:40 - 11:00; Chair: Till Strohsal

Forecasting New York State Tax Revenue: A Factor MIDAS Approach

Kajal Lahiri (University at Albany-State University of New York); Cheng Yang (University at Albany-State University of New York, cyang6@albany.edu); Onur Bugdayci (University at Albany-State University of New York); John Delaney (University at Albany-State University of New York)

Models with mixed frequency and "big data" have been extensively used to forecast low-frequency variables such as GDP and inflation, but rarely used in government revenue forecasting. In this paper, we forecast annual New York State tax receipts using New York monthly tax data collected by us and a large number of U.S. macro variables and New York or northeast-region macro variables. We choose Factor MIDAS as our model. Our forecasts are compared with budget plans of the Division of the Budget (DOB). The results show although our model is not competitive in horizons shorter than or equal to 6 months as DOB has perfect tax information when fiscal years is about to end, it is doing better at longer horizons and most importantly, it does better at forecasting tax revenue in recessions. Since our model performs better at horizons of 1 year or above, it could be supportive to budget plan making as late budget is an issue in New York State.

Assessing Regulatory Impact Assessments (RIA) in Poland – government vs. private sector expectations on likely policy outcomes

Kamil Jonski (University of Lodz); Wojciech Rogowski (Warsaw School of Economics, wojciech.f.rogowski@gmail.com)

Regulatory Impact Assessment (RIA) – was designed to rationalize law-making process by providing policy makers with accurate summary of expected impacts of contemplated policies.

Thus, RIAs can be regarded as conditional forecasts (what outcomes should we expect, given policy implementation). However, unlike other economic forecasts (focused on single indicator), the scope of RIA is tremendously wide - it involves question of whether (and how much) 'the economy would be better off' after adoption of given law (implementation of the policy). Since there is no granular data on actual impacts of policies introduced in particular laws, it is impossible to assess RIAs like economic forecasts - using forecast errors.

To circumvent this problem, scholars attempting to evaluate RIAs tend to focus on indirect measures like compliance with formal requirements or expert assessment of underlying methodology. Thus, they overlook the information provided to the decision makers. This paper offers alternative approach, resorting to the comparison between government and private sector expectations. On the conceptual level, this corresponds to gauging forecast reliability by its agreement with consensus.

Specifically, using data on 129 laws passed by Polish Parliament during 2014-15 we compare government's (RIA) and business community expectations on likely policy impacts. To quantify the government expectations expressed in RIAs, each of them assessed by three independent coders. Results were aggregated using PCA into the index proxying government's expectations on likely policy outcomes. Business community expectations were proxied by index complied for each draft law by Polish Chamber of Commerce. To measure the degree of consensus between them, the latter index was regressed against the former.

Since our sample covers substantial reform of the Polish RIA process (involving institutional gatekeepers, staff training and RIA templates) we verified whether it affected the degree of

consensus between government's and business community views. We found that for pre-reform RIAs, government's expectations were reasonably good predictor of business community expectations (tendency to consensus). However, the link vanished after the reform. We link this surprising result to the bureaucratic response to growing pressure from reformers - excessive focus on details at the expense of the big-picture assessment.

Can the SPF and FOMC participants learn from each other? Using qualitative information from the FOMC minutes to elicit forecasts of the U.S. GDP growth

Olga Bespalova (International Monetary Fund, olgabesp@gwu.edu)

This paper utilizes textual analysis of deliberations in the Federal Open Market Committee (FOMC) minutes to analyze asymmetry of information on the US economy between the monetary policymakers and the Survey of Professional Forecasters (SPF). Building on Stekler and Symington (2016) and updated list of predominant recurring words, I convert FOMC qualitative statements into the current and next quarter outlook indexes and calibrate them to the US GDP growth forecasts. I extend their time-series (originally created for 2006-2010) adding 26 years of bi-quarterly observations to cover 1986Q1-2016Q4. Following Ericsson (2017), I interpret the derived calibrations (FMIs) as elicit-casts of the Greenbook forecasts. I show that these FMIs are unbiased, efficient, rational, and contain the Greenbook' informational advantage. Forecast encompassing tests suggest that both the FMIs and SPF forecasts contain unique knowledge and can learn from each other. I find evidence that the SPF forecasters efficiently use information from the FOMC minutes available to them before the forecast deadline (these cover second meeting from the previous quarter). Yet, they could improve their forecasts should the FOMC minutes from the first quarterly meetings become available without a publication lag. On the contrary, the FOMC policy-makers did not pay due attention to the SPF forecasts released prior to their second bi-quarterly meeting, accounting only for their own earlier assessment of the U.S. economy.

The Role of Data Revisions for Nowcasting German National Accounts

Till Strohsal (Federal Ministry for Economic Affairs and Energy, till.strohsal@bmwi.bund.de)

This paper examines the statistical properties of revisions to German National Accounts data. Except for benchmark revisions, the German Federal Statistical Office revises data 4 years backwards. Considering the data final after 4 years, I show that in comparison to the variation of final GDP data, revisions are quite large. Revisions are also significantly downward biased, even though the bias is fairly small. However, the news hypothesis that the initial release of GDP is an efficient forecast of the final value cannot be rejected. Also, no autocorrelation is present in revisions. The noise hypothesis that the initial release equals the truth plus measurement error is clearly rejected by the data.

Multivariate time series models

Monday, 9:40 - 11:00; Chair: Saeed Zaman

Handling Structural Breaks in Economic Forecasting: A Competition

Kandrika Pritularga (Lancaster University, kandrika.fp@gmail.com)

An unexpected shock or a structural break in time series close to where forecasts start damages forecasts generated by conventional models. This research initially focuses on two strategies to handle such a problem, namely the differencing strategy and the optimal window strategy.

A comparison between these two strategies yields an opportunity to develop an alternative strategy by combining the optimal window strategy and the rolling origin approach with re-estimation, named as the updating-optimal window strategy. The comparison between these three strategies was done by Monte Carlo simulation and an empirical analysis of the unleaded gasoline price in the U.S.

The results show that the updating-optimal window strategy outperforms others in most cases for the experimental and the empirical analysis. In addition, this research reveals that the differencing strategy is not robust toward different forecasting problems. In general, this research recommends employing the updating-optimal window strategy first to handle unexpected structural breaks while the differencing strategy needs to be considered in a particular situation.

Weekday-Weekend, Day of Week, and Prior Day Effects in Forecasting Daily Natural Gas Demand from Monthly Data

Maral Fakoor (Marquette University, maral.fakoor@marquette.edu); George F. Corliss (Marquette University); Ronald H. Brown (Marquette University)

Local distribution companies rely on accurate forecasts of daily natural gas demand for buying and delivering natural gas to their customers. To forecast daily natural gas demand, historical daily inputs such as temperature, wind, and gas flow are needed. However, in some situations, flow may be available only at the monthly level, perhaps in billing cycles. The aim of this work is to make effective daily models when only monthly flow data is available.

Previous research has shown that linear regression models of monthly measurements and aggregated inputs can be used to forecast daily flow. Multi-parameter linear regression base models trained with historical monthly weather and demand data are evaluated using daily weather data to forecast daily gas demand. Since information is lost during aggregation of daily flow to monthly flow, we introduce adjustments to the base daily models to account for the effects of weekday vs. weekend, day of the week, and prior day weather. Weekday vs. weekend and day of the week effects are not recoverable from monthly data, and not accounting for these assumes all days are the same. Prior day adjustment accounts for prior-day weather impacts including temperature and wind in addition to today's conditions.

We hypothesize that incorporating weekday vs. weekend, day of the week, and prior day weather improves the forecast performance of the base models. Weekday vs. weekend, day of the week, and prior day weather coefficients from available daily datasets are calculated to test the models.

In another test as in a real case, the information about weekday vs. weekend, day of the week, and prior day weather are not recoverable from monthly data, we use average values acquired from existing datasets and compare them to coefficients obtained from the specific dataset to examine the deleterious effect of using average. Extensive experiments with real data acquired from local

distribution companies show the validity of our proposed approach. Error decreases by including weekday vs. weekend, day of the week, and prior day weather effects as compared to a base model.

Hourly Hospitality Forecasting - Experiments using Random Forests to incorporate many complex regression variables

Aidan Morrison (Trendlock, aidan@trendlock.com.au); Ross Ireland (Trendlock)

Many bars, restaurants, and cafes depend upon walk-in customers for a large part of their trade. The timing and existence of lulls and rushes can have a very large impact on operations, and complicate the efficient rostering for staff and stock ordering. Most small businesses only have very limited capacity to conduct analysis of the data that is available, and existing Point-of-Sale (POS) systems tend not to incorporate any forecasting functionality. The very large number of potential external variables (such as weather, local events, holidays, school or university calendars) as well as complex intra-day and intra-week seasonality tend to complicate and frustrate many simple time-series forecasting techniques. For example, weather effects are often found to have different and even opposite regression co-efficients in different times of day, or different parts week, even in the same business. In this research, the performance and limitations of some forecasting techniques, including ARIMA and Bayesian Structured Time Series are examined. The potential for using Random Forests to quickly identify the most significant regression variables for a given business and produce an effective forecast quickly, or contribute to an ensemble to build on other methods are explored.

Combining Survey Long-Run Forecasts and Nowcasts with VAR Forecasts using Relative Entropy

Saeed Zaman (Federal Reserve Bank of Cleveland, saeed.zaman@clev.frb.org); Ellis Tallman (Federal Reserve Bank of Cleveland)

This paper constructs hybrid forecasts that combine both short and long-term conditioning information from external surveys with forecasts from a standard fixed-coefficient Vector Autoregression (VAR) model. Specifically, we use relative entropy to tilt one-step ahead and long-horizon VAR forecasts to match the nowcast and long-horizon forecast from the survey of professional forecasters. The results indicate meaningful gains in multi-horizon forecast accuracy relative to model forecasts that do not incorporate long-term survey conditions. The forecast accuracy gains for inflation and the interest rate are substantial, statistically significant, and are competitive to the forecast accuracy from both time-varying VARs and univariate benchmarks.

State space models

Monday, 9:40 - 11:00; Chair: John Boylan

Maximum Likelihood Estimation of a TVP-VAR

Guilherme V. Moura (UFSC, gvallemoura@gmail.com); Mateus Noriller (UFSC)

This paper discusses the maximum likelihood estimation of a vector autoregression with drifting coefficients and multivariate stochastic volatility. Dynamic covariance matrices are treated as Wishart processes, allowing a more natural modelling of time-varying covariances via a matrix-valued stochastic process, respecting positive definiteness and coupling nicely with the Normal distribution for the observable variables. Based on the works of Uhlig (1994), Uhlig (1997) and Kim (2014), and exploiting the conjugacy between the Normal and Wishart distributions, a closed form solution for this nonlinear filtering problem is presented, together with an analytical expression for the likelihood function.

Cogley and Sargent (2005) and Primiceri (2005) were responsible for introducing and popularizing TVP-VARs in macroeconomics, and they tackle the estimation problem of this nonlinear state space model using MCMC algorithms. However, their estimation approach can be cumbersome. We rely on the specific structure for the stochastic volatility process of Uhlig (1997) to exploits the conjugacy between Wishart and multivariate singular beta distributions. This choice allows us to analytically solve the nonlinear integrals in the filtering process of the TVP-VAR model with multivariate stochastic volatility driven by Wishart processes. A by-product of this result is an exact likelihood function, allowing maximum likelihood estimation. Thus, this paper proposes a TVP-VAR version of Uhlig's (1997) constant parameter VAR model, which can be estimated in one step by maximum likelihood, and whose stochastic volatility structure is modelled using matrix-valued stochastic processes. We then apply our method to a dataset of the U.S. economy which is similar to the one used in Cogley and Sargent (2005). Our model generates similar results for the evolution of the VAR coefficients, and also for the evolution of innovations' variances.

Forecasting with the Linear-Ellipsoidal-Bounded (LEB) estimator - comparison with the Kalman filter

Claudio Antonini (Bank of New York Mellon, cda@alum.mit.edu)

The Linear-Ellipsoidal-Bounded (LEB) estimator was originally developed in a non-probabilistic framework. Although its structure is comparable to the Kalman filter, its interpretation is substantially different. The estimator was created to work with inputs, errors, and outputs described only by their minimum and maximum temporal bounds (oftentimes the only description available of a time series) and, thus, it is not restricted to Gaussian environments. In addition to working with unknown distributions, the LEB estimator does not need to augment the number of states to account for 'colored' noises. Structurally, the estimator is expressed using the same matrix formulation than the Kalman filter and the processing follows the same traditional steps of extrapolation and update. Given the ellipsoidal description of the dynamic sets, however, the process incorporates two parameters that are used to build the bounding ellipsoids (one for extrapolation, another for the update). These two parameters are now estimated in a novel way.

We will derive a new way of solving the system equations (in special circumstances they can be solved analytically). At the same time, we will show different ways of finding the two optimum parameters that determine the extrapolation and update steps, by minimizing the volume of the bounding ellipsoid for the state estimates.

The behavior and performance of both filters (LEB and Kalman) will be shown in a reduced model of the economy (e.g., to forecast GDP via potential GDP and unemployment).

Multiple seasonal ETS models for multiple series

Huijing Chen (University of Portsmouth, huijing.chen1@port.ac.uk); John Boylan (Lancaster University Management School)

In recent years, there is a tendency to record data more frequently, which allows the presence of double or triple seasonality but also poses a challenge for estimating them accurately, especially when data histories are short. State space models with a single source of error are developed in this paper with multiplicative components. They underline the procedures for forecasting multiple series with individual levels and trends, but multiple seasonal factors, with the flexibility to assume some of them are common across series. Optimal forecasting procedures and prediction intervals are derived, analytically where possible, and through simulations.

These new methods are then compared to: (1) methods for multiple series with single seasonality, and (2) previous methods for a single series with multiple seasonality.

Inflation

Monday, 11:20 - 12:40; Chair: Michal Hulej

Improving Phillips Curve Inflation Forecasts Using A Robust Asymmetry Measure

Randal Verbrugge (Federal Reserve Bank of Cleveland, randal.verbrugge@clev.frb.org); Saeed Zaman (Federal Reserve Bank of Cleveland); Keerthana Nunna (Federal Reserve Bank of Cleveland)

Phillips curve-based forecasting models continue to be popular in inflation forecasting. This paper considers whether an aspect of the cross-sectional distribution of inflation rates across the 178 items in the Personal Consumption Expenditures (PCE) price index, namely a robust measure of its asymmetry, provides useful additional information to a Phillips-curve (PC)-based monthly forecast of headline PCE inflation. Our results show that, not only does asymmetry improve a PC-based inflation forecast, but also this variable actually has more predictive power than does the unemployment gap itself, in that replacing the unemployment gap in a PC model with this variable meaningfully improves the inflation forecast accuracy. We also show that model specifications that include the asymmetry measure generate forecast accuracy that rivals hard to beat benchmark models. We offer a conjecture as to why this information appears to have predictive value.

The role of inflation surveys for inflation forecasting

Raisa Basselier (National Bank of Belgium, raisa.basselier@nbb.be); David De Antonio Liedo (National Bank of Belgium); Jana Jonckheere (National Bank of Belgium); Geert Langenus (National Bank of Belgium)

In this paper, we examine whether forecasts of inflation can be improved by incorporating qualitative information from monthly surveys that are 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. This paper evaluates forecasts for both core and overall inflation in Belgium and the euro area for horizons ranging from one quarter to one year. First, these survey-based inflation expectations are directly exploited via simple regression equations. Second, we add economic activity variables into the regression, which thereby becomes an expectations-augmented Phillips curve. This formulation implicitly acknowledges the interaction between aggregate demand and inflation expectations in the determination of prices. We consider model uncertainty by exploring model specifications that also incorporate variables such as oil prices, the import deflator and alternative measures of economic activity, such as unemployment, GDP or the output gap.

Since the regression analysis does not allow us to conclude which one of the alternative activity indicators yields the best approximation of the true unobserved state of aggregate demand, and both surveys are not equally accurate across subsamples, we propose to incorporate all the variables in a single model. We develop a mixed-frequencies dynamic factor model with two underlying factors that represent real activity and inflation, respectively. Variables reflecting expectations, such as the above mentioned surveys, are defined in a model-consistent way. This approach allows to understand the precise role of each one of the indicators at defining expectations at both short and long-run horizons.

Inflation Expectations in a Highly Volatile Environment: Evidence from a Firms' Survey

Nathan Goldstein (University of California, Berkeley, nathan.goldstein@berkeley.edu); Ben-Zion Zilberfarb (Bar-Ilan University)

This study provides new evidence about inflation expectations formation in a highly volatile environment, in which inflation dropped dramatically from three-digits annual rate to less than 3%. Our findings, based on an Israeli survey of firms' expectations, support information rigidity theories, and at the same time point to the presence of asymmetric preference towards forecast errors. Most importantly, both inattentiveness and asymmetry depend on inflation, being more pronounced when inflation is sufficiently decreasing. This pattern applies both for high and low inflation regimes. Our evidence demonstrates the complexity of inflation expectations, having important implications for macroeconomic dynamics and policy issues.

Evaluating the NBP's density forecasts of inflation and GDP

Michal Hulej (Narodowy Bank Polski, michal.hulej@nbp.pl)

In this paper we evaluate the quality of density forecasts published by the Narodowy Bank Polski (NBP, central bank of Poland). In its minutes, the MPC put much emphasis on the uncertainty of future outcomes. It is thus relevant to assess whether it's communication embedded in the form of fan charts correctly leads the market. Two types of approaches to the probability forecast evaluation are used. First we test whether the dispersion of outturns relative to the fan chart percentiles is consistent with the nominal width of the bands. Second we check whether the probability of a certain event (CPI inflation falling into permissible fluctuation band around the inflation target, GDP growth running below its sample average) implied by the fan chart distribution function is a good indicator for the actual occurrence of these events. We found that the fan charts published by central bank of Poland in its Inflation Reports provide a good measure of the underlying forecast uncertainty. Both for GDP and CPI the fans are of correct width, though in case of inflation this conclusion is only valid in the sample constrained to exclude recent period of deflation. Thus, the uncertainty associated with the supply shocks related to falling energy commodity prices in the world markets was not accounted for to full extent in the NBP's communication. NBP's forecasts are relatively good at capturing probabilities of particular events of interest in GDP and CPI developments, which is particularly important under the inflation targeting regime.

Nowcasting

Monday, 11:20 - 12:40; Chair: Nuno Crato

The Montenegrin macroeconomic forecasting model

Michael Graff (ETH Zurich, KOF Swiss Economic Institute, graff@kof.ethz.ch); Yngve Abrahamsen (ETH Zurich, KOF Swiss Economic Institute); Vojin Golubović (University of Donja Gorica and Institute for Strategic Studies and Prognoses, Podgorica, Montenegro); Milika Mirković (University of Donja Gorica and Institute for Strategic Studies and Prognoses, Podgorica, Montenegro); Boriss Siliverstovs (The Bank of Latvia)

Measuring and forecasting GDP and its components form the core of a large number of applied economic analyses. For the small economy of Montenegro, which become independent in 2006, the last years revealed weaknesses in economic analyses and monitoring, which can largely be attributed to a of lack of appropriate data and forecasting models. Accordingly, a comprehensive macroeconomic model of the Montenegrin economy is called for. Initial attempts by the Montenegrin Institute for Strategic Studies and Prognoses (ISSP) to develop such a model (2004, 2006, 2008 and 2010) were mainly based on continuations of past trends. As this approach basically fails to predict anything apart from continuous evolutions, its practical usefulness was limited.

Recently, an institutional partnership of the ISSP and the Swiss KOF Economic Institute resulted in the first operational macroeconomic structural forecasting model for Montenegro.

The model is in the spirit of the demand driven approach and designed to provide timely information on the current state of the Montenegrin economy for different volumes and prices as well as to provide forecasts that give a consistent outlook. The theoretical foundations are the identities in national accounting and macroeconomics and the behavioural interrelatedness between important economic variables. The implementation is based on time series – mainly in terms of growth rates. This required significant effort in preparation an adequate data base due to the poor availability of data (both hard and soft), in particular for longer periods, given the only recent break-up of Serbia-Montenegro. Economic relations are modelled by multiple linear regressions.

The foci of the paper cover (1) the choice of the model type, (2) the creation of the data base including adjustments and computations of missing observations, (3) the preparation of a comprehensive code for the entire process, from specifying equation by equation to the formulation of the static and dynamic model solving process.

As a basis for measuring and forecasting of GDP and its expenditure components as well as their contribution to overall economic activity in the country, the model comprises nine groups of equations, reflecting the main behavioural equations for consumption, investment, exports and imports, divided into goods and services. Three additional equations relating to employment, net wages and foreign tourists' overnight stays are developed addressing the labour market and the tourism sector, which is particularly important for Montenegro.

Our model of the Montenegrin economy provides short and medium term forecasts of GDP and its expenditure components and allows for various model-based analyses of economic policy options. It is operational since early 2017, and four quarterly model runs have been conducted, and the results were released to the public. On this basis, the paper concludes by a look at the model's nowcasting and forecasting performance in a real-time framework, which provides first insights into its usefulness, although we are careful to stress the data limitations at this point.

Nowcasting Gross Domestic Output

Travis Berge (Federal Reserve Board, travis.j.berge@frb.gov); Kurt Lewis (Federal Reserve Board)

A recent literature in macroeconomic forecasting has developed models that estimate real GDP growth at high frequencies, known as nowcasts. However, there are two official measures of U.S. economic output, gross domestic product (GDP) and gross domestic income (GDI). Conceptually, GDP and GDI should be equal, but each is measured with error. We develop a model that produces high-frequency forecasts of the common component to the growth rates of real GDP and real GDI, which we denote real gross domestic output (GDO). The mixed-frequency dynamic factor model extracts the common component to GDP, GDI and roughly 30 additional monthly macroeconomic indicators. We produce, using a fully real-time dataset since 2006, GDO nowcasts. We compare the properties of GDO revisions to those of GDP and GDI, and evaluate the model's nowcasts of GDP and GDI themselves.

Nowcasting gross domestic product in Japan using professional forecasters' information

Nobuo Iizuka (Kanagawa University, nobuo-iizuka-0915@kanagawa-u.ac.jp)

This study proposes a new framework for nowcasting, the prediction of the present, the very near future, and the very recent past. Most previous research uses a dynamic factor model (DFM) and insists that it is the best method for nowcasting. However, the DFM takes considerable time and effort and for Japanese GDP, the forecast result with DFM is not necessarily good (Bragoli, 2017). We use professional forecasters' information instead. In this study, we combine professional forecasters' information with single-equation approaches, such as bridge equations (BEQ) and mixed-data sampling (MIDAS) regressions. We use cross-sectional disagreement among forecasters in the ESP forecast survey, which is the first monthly survey of macroeconomic forecasts conducted by professional forecasters in Japan.

A spectral clustering procedure with potential for nowcasting and forecasting risk in big data time series

Nuno Crato (University of Lisbon, ncrato@gmail.com); Jorge Caiado (University of Lisbon); Pilar Poncela (Universidad Autónoma de Madrid and EC-JRC)

Clustering time series may bring information about a set of time series and help forecasting or nowcasting a particular series of interest. Such procedures maybe particularly interesting but difficult in a big-data environment. We propose and study a frequency-domain procedure for characterizing and comparing large sets of long financial time series volatility. Instead of using all the information available from data, which would be computationally very expensive, we propose to select and summarize the most relevant information for clustering purposes. Essentially, we propose to use a fragmented periodogram, computed only at the driving seasonal components of interest. This procedure is computationally simple, but able to condense relevant second-order information to compare and cluster the volatility time series. We use this procedure to study the evolution of several stock markets indices, extracting information both on the European financial integration and on the recent worldwide financial crisis.

Density & anomaly

Monday, 11:20 - 12:40; Chair: Guillermo Carlomagno

Probabilistic Forecasting of Agricultural Yield

Heitor Arakawa (University of Virginia, hh5jx@virginia.edu)

Government agencies and companies usually omit the uncertainty associated with agricultural yield forecasts. In order to produce yield forecasts, conventional methodologies combine subjective assessments from farmers and analysts with a quantitative analysis of several predictors of agricultural production, such as weather conditions, crop measurements, satellite images, and records of past yields. This study expands upon a well-know Bayesian theory of probabilistic forecasting in the context of agricultural yields in order to quantify the uncertainty associated with public forecasts. Specifically, Bayesian Processor of Forecast (BPF) models are constructed for three cases: forecasting local yield, forecasting regional yield with local prior information, and forecasting regional yield with regional prior information. The predictand is the yield [production per area] after the harvest season is over. The predictors are the deterministic yield forecasts issued by public sources during the growing season.

Farmers provided their judgmental assessments in terms of quantiles of local and regional yield, from which prior density functions were constructed. These functions were combined to form an aggregated prior density function by expanding the theory of Bayesian Model Averaging (BMA). The BPF produced posterior density functions of the yield (local or regional) conditional on the realization of the predictors. A study case for each model is presented using data of soybean farmers from Mato Grosso, Brazil. The predictors were the deterministic forecasts issued in October, February, and May by Companhia Nacional de Abastecimento (CONAB) and Instituto Brasileiro de Geografia e Estatistica (IBGE). The predictand was observed in September and issued by IBGE. By explicitly quantifying uncertainty in the yield forecasts, the outputs of the BPF models have the potential to enhance decision making.

Anomaly Detection in Nonstationary Streaming Temporal Data

Priyanga Dilini Talagala (Monash University, Australia, dilini.talagala@monash.edu); Rob Hyndman (Monash); Kate Smith-miles (University of Melbourne)

This work develops a framework for detecting anomalous series within a large collection of time series in the context of non-stationary streaming data. We define an anomaly as an observation that is very unlikely given the forecast distribution for the corresponding time period. In this work we make two fundamental contributions. First, we propose a framework that provides early detection of anomalous behaviour within a large collection of streaming time series data. The proposed framework first forecast a boundary for the systems' typical behavior and then a sliding window is used to test for anomalous series within the newly arrived collection of series. An approach based on extreme value theory is used for the boundary prediction process. Second, we propose a novel approach for early detection of non-stationarity (also called "concept drift" in the machine learning literature.) The proposed algorithm uses time series features as inputs, and a density-based comparison to detect any significant change in the distribution of the features. Using various synthetic and real world datasets, we demonstrate the wide applicability and usefulness of our proposed framework. This framework is implemented in the open source R package oddstream. We show that the proposed algorithm can work well in the presence of noisy non-stationary data within multiple classes of time series.

Forecast densities for series contaminated with outliers

Guillermo Carlomagno (Center of Economic Research (CINVE), guillermocarlomagno@gmail.com); Rafael Pagaini (Center of Economic Research (CINVE))

When constructing unconditional point forecasts, both direct- and iterated-multistep (DMS and IMS) approaches are common. However, in the context of producing conditional forecasts, IMS approaches based on vector autoregressions (VAR) are far more common than simpler DMS models. This is despite the fact that there are theoretical reasons to believe that DMS models are more robust to misspecification than are IMS models. In the context of unconditional forecasts, Marcellino, Stock, and Watson (MSW, 2006) investigate the empirical relevance of these theories. In this paper, we extend that work to conditional forecasts. We do so based on linear bivariate and trivariate models estimated using a large dataset of macroeconomic time series. Over comparable samples, our results reinforce those in MSW: the IMS approach is typically a bit better than DMS with significant improvements only at longer horizons. In contrast, when we focus on the Great Moderation sample we find a marked improvement in the DMS approach relative to IMS. The distinction is particularly clear when we forecast nominal rather than real variables where the relative gains can be substantial.

Business cycles

Monday, 14:00 - 15:20; Chair: Marius Mihai

Modeling the business and financial cycle in a multivariate structural time series model

Jasper de Winter (De Nederlandsche Bank, j.m.de.winter@dnb.nl); Siem Jan Koopman (Vrije universiteit Amsterdam); Irma Hindrayanto (De Nederlandsche Bank)

We consider a multivariate unobserved component time series model to disentangle the short-term and medium-term cycle for the G7 countries and the Netherlands using four key macroeconomic and financial time series. The novel aspect of our approach is that we simultaneously decompose the short-term and medium-term dynamics of these variables by means of a combination of their estimated cycles. Our results show that the cyclical movements of credit volumes and house prices are mostly driven by the medium-term cycle, while the macroeconomic variables are equally driven by the short-term and medium-term cycle. For most countries, the co-movement between the cycles of the financial and macroeconomic variables is mainly present in the medium-term. First, we find strong co-cyclicality between the medium-term cycles of house prices and GDP in all countries we analyzed. Second, the relation between the medium-term cycles of GDP and credit is more complex. We find strong concordance between both cycles in only three countries. However, in three other countries we find 'indirect' concordance, i.e. the medium-term cycles of credit and house prices share co-cyclicality, while in turn the medium-term cycles of house prices and GDP share commonality. This outcome might indicate that the house price cycle is –at least partly– driven by the credit cycle. Lastly, the cross-country concordance of both the short-term cycles and the medium-term cycles of GDP, house prices and credit is low. Hence, the bulk of the cyclical movements seem to be driven by domestic rather than global factors.

Forecasting the best solution for hiring labor and purchasing materials by optimization with the Newton and Quasi-Newton Methods

Eduardo Pimentel (IFES - Instituto Federal do Espírito Santo, epimentel@outlook.com.br); Mário Mestria (IFES - Instituto Federal do Espírito Santo); Shirley Cani (IFES - Instituto Federal do Espírito Santo)

This work presents a nonlinear modeling to find the optimal solution in hiring local labor and in buying materials for concreting a factory floor. In this case of civil construction, it was observed that the decision variables - the choice between single hardware for assembly in the building site versus hardware already assembled in mesh form, and the choice between own labor force and subcontracted labor - presented a non-linear correlation between them, and that an appropriate way to predict the decision involved either a Newton or Quasi-Newton method. The best solution could be forecasted by both methods with similar results.

Do Credit Booms predict U.S. recessions?

Marius Mihai (City University of New York - Graduate Center, mmihai@gradcenter.cuny.edu)

This paper investigates the role of bank credit in predicting U.S. recessions since the 1960s in the context of a bivariate probit model. A set of interesting results emerge. First, credit booms are shown to have strong positive effects in predicting declines in the business cycle at the 6- and 9-month horizon. Second, I propose to trace out the effect of credit booms over the length of the time series by identifying the contribution of excess bank liquidity in the downturn of each cycle. Third, the out-

of-sample performance of the model is tested on the most recent credit-driven recession, the Great Recession of 2008. The model performs better than a more parsimonious version where we restrict the effect of credit booms on the business cycle in the system to be zero.

Finance (I)

Monday, 14:00 - 15:20; Chair: Maxwell Stevenson

Forecasting International Stock Market Returns Through Global Economic Activity

Elena Diaz (Universidad de Navarra, ediaza@unav.es); Gabriel Pérez-quirós (Bank of Spain)

This study examines the month-to-month predictability of international stock markets through global real economic activity. To measure the world's economic fluctuations, we develop a novel indicator based on selected commodity prices (iron ore, uranium and tin), which allows for a real-time knowledge of international business conditions. We find out-of-sample predictability of stock returns in world, developed, emerging and G7 stock markets in a sample between 2000M1 and 2016M12. One main finding is that predictability is especially significant during recession periods. Results indicate that during stable economic periods, stock markets efficiently incorporate global business conditions, but fail to do so during crisis periods. The use of a global economic activity indicator, therefore, becomes relevant during difficult times.

Have Earnings Momentum Anomalies in Stock Markets Disappeared? The Evidence from the U.S. and Non-US Markets Suggests No!

John Guerard (cin, jguerard@mckinleycapital.com); Sundaram Chettiappan (McKinley Capital Management, LLC)

Earnings momentum anomalies were identified in U.S. and Japanese stocks over 25 years ago. Have the anomalies been arbitraged away? We report evidence in U.S. and non-U.S. markets that these anomalies are still statistically significant during the 2003 - 2017 time period.

Time-Series Prediction via Principle Components: Predicting Stock Prices using Multiple Predictors

Mahsa Ghorbani (Colorado State University, mahsa.ghorbani@colostate.edu); Edwin Chong (Colorado State University)

The literature on financial forecasting documents evidence that prices can be predicted from past price data as well as other fundamental and macroeconomic predictors. We are interested in a general method for stock price prediction using covariance information. The basic goal is to predict future stock price values based on past values of stock prices and other relevant data. However, many prediction methods based on covariance information are numerically ill-conditioned. This problem becomes even more challenging when using multiple predictors because in that case the covariance matrix is likely to be very large, leading to severe ill-conditioning. In this paper, we propose a filtering operation based on principle component analysis, which mitigates the above ill-conditioning issue. The proposed method is easily implemented and can be configured to include an arbitrary number of predictors.

To improve the prediction accuracy we first use the multichannel cross-correlation coefficient (MCCC) as a measure for selecting the right predictors for each considered stock. Then we investigate the optimal number of predictors as the input to the proposed estimation framework. We illustrate our method on daily stock price values for General Electric, showing that incorporating additional predictors can improve prediction accuracy. We investigate the results based on two performance metrics: mean square prediction error and directional price movement.

Boom or Bust: Exploring the Impact of Risk Appetite and Valuation Ratio Measures on the Stock Market Outlook

Maxwell Stevenson (University of Sydney, maxwell.stevenson@sydney.edu.au); Maurice Peat (University of Sydney)

In this study we use standard valuation ratios and a measure of risk appetite proposed by Kumar and Persaud (2002) and formalised by Misina (2003, 2005, 2008) in a predictive regression analysis of stock market outlook. We examine two markets, the US market which is the largest and most liquid market and the Australian market, a smaller developed market. We find that valuation ratios are strongly predictive over a three month horizon in the Australian Market, but less so in the US market. In both markets a measure of global risk appetite is shown to have strong predictive value.

Demography

Monday, 14:00 - 15:20; Chair: Han Li

Divorces in Mexico: a rational choice?

Eduardo Loria Diaz (National Autonomous University of Mexico); Emmanuel Salas (National Autonomous University of Mexico, salas.emmanuel@gmail.com)

At least since 2005 the divorce rate has been raising sharply. Through the Economic Approach to Human Behaviour by Becker (1968) we find its economic determinants (2005.Q1-2016.Q1) and therefore we proceed to forecast its evolution (2016.Q2-2020.Q4). With a VEC(1) we find that the raise of the female rate of participation and the raise of female education affect positively the rate of divorce in Mexico. The impulse-response and variance decomposition analysis fairly prove the well-known 'Independence Effect' hypothesis.

Modeling survey time series data with flow-observed CARMA processes

Patrick Joyce (U.S. Census Bureau, patrickjoyce82@gmail.com); Tucker McElroy (U.S. Census Bureau)

Customized epoch estimates, projections made at arbitrary times and time intervals, are demanded by consumers of survey data. Because the granularity of survey data is unable to provide direct estimates at high frequencies (with requisite quality), we propose instead to utilize apply stochastic models to generate interpolations, projections, and temporal cumulations, and thereby publish customized epoch estimates that may also have improved properties over those reliant upon survey information alone. Our research focuses upon the class of continuous auto-regressive moving average (CARMA) processes, developing the inferential facets of using these models for flow time series, i.e., survey data that has been aggregated over a given epoch. We establish new theoretical, methodological, and computational results for CARMA processes, observed as either a stock and flow series, and study an application to the American Community Survey, providing customized estimates of the U.S. veteran population.

A flexible multi-factor age-period-cohort approach to mortality modeling with forecast reconciliation

Han Li (Macquarie University, han.li@mq.edu.au); Hong Li (Nankai University); Yang Lu (Paris 13 University)

Life expectancy has been increasing sharply in the U.S. since the second half of the 20th century. However, between 2014 and 2016, U.S. life expectancy has declined for two years in a row, for the first time since the early 1960’s. This phenomenon has attracted a lot of attention in U.S. society. As argued by various academic researchers and medical experts, potential factors that contribute to the decreasing trend may include a high obesity rate and drug overdoses. To gain better insights into the historical mortality patterns and to improve mortality forecasts, in this study, we analyze both aggregated national mortality data and cause-specific mortality data in the U.S. since 1950. First, we allocate cause-of-death mortality improvement into age, period, cohort, and residual components. Flexible multi-factor age-period-cohort models are proposed for a wide range of causes of death. We then link major causes of death back to extrinsic mortality drivers, such as anti-smoking campaigns, HIV/AIDS and pneumonia or influenza. Period effects that have been identified in the modeling process will be extrapolated to project future cause-specific mortality rates. It is expected that the aggregation of these individual cause-of-death forecasts will provide a more accurate projection for the national total death rate. In order to generate a set of coherent forecasts across all

causes and make sure they add up to national level mortality rates, we apply several advanced forecast reconciliation techniques such as the 'bottom up' method developed by Dangerfield and Morris (1992), Zellner et al. (2000), as well as the more recent optimal combination method developed by Wickramasuriya et al. (2017).

Demand forecasting (I)

Monday, 14:00 - 15:20; Chair: Harald Schmidbauer

Incorporating market sentiments into retail nowcasts and forecasts

Michał Chojnowski (SGH Warsaw School of Economics, michal.chojnowski@doktorant.sgh.waw.pl)

With Industry 4.0 researches are looking for more and more vast sources of data to analyse trends and predict the market performance. Among them Google Trends is most commonly used for forecasting purposes due to its availability (see: Choi & Varian (2012), D'Amuri & Marcucci (2012), Chen et al. (2015)). Especially when it comes to applying market sentiments, a factor more and more often described in recent macroeconomic literature (see: Benhabib (2013), Angeletos (2013), Feve & Guay (2016)), Google Trends seem to be a cheaper and more accessible source of customer preferences evaluation (see: McLaren & Shanbhouge (2011)).

This article proposes a novel approach of observing market sentiments in the real time on any market of choice. As opposed to commonly used sentiments indices based on surveys (such as PMI, ESI, PENGAB etc.), a method based on Bayesian SVAR model is presented, which allows to extract more specific market sentiments and implement them into forecasting models even if survey-based index does not exists for a given market.

Following 'news information' framework (see: Barsky & Sims (2012)), not sales itself but rather consumer's decision is a variable being directly influenced by market sentiments as they possess crucial information on current and future states of the market. Any change in information drives the economy to the newly established equilibrium, which implies inaccuracy of ARMA models. To overcome this issue LSTAR models were used with extracted sentiments as a threshold variable. The author shows in his previous work (see: Chojnowski (2018)) that implementing market sentiments can improve forecast accuracy by 25% with respect to best-performing ARMA model within 12-month horizon.

This speech focuses on capturing market sentiments in the real time by observing changes in Google Trends queries via artificial neural network (ANN). To prevent network from overfitting, queries are skimmed beforehand with the use of the algorithm based on slab-and-spike regression. By setting up the restriction in the algorithm, researchers are able to create Google Trends sets, which are consistent in time, but they update or remove queries if necessary. Results of ANN are implemented into LSTAR model with 2 regimes to provide the final forecast.

The model presented in the article provides for both business and researches an applicable tool to extract sentiments from historical data and observe them in the real time. The results obtained by the author suggest a significant improvement in forecasting capabilities: LSTAR model with ANN support improved forecast by 30% with respect to best-performing ARMA.

A Step Towards Demand Sensing: Using Open Orders in Demand Forecasting

Jente Van Belle (Vrije Universiteit Brussel, jente.van.belle@vub.be); Tias Guns (Vrije Universiteit Brussel); Wouter Verbeke (Vrije Universiteit Brussel)

The objective of this research is to investigate the added value of using open orders data in demand forecasting. Traditionally, demand forecasting relies on statistical methods such as ARIMA and ETS models to extrapolate historical sales into the future. Although these statistical models produce reasonably accurate forecasts in many cases, industry is looking for an increasingly higher level of forecast accuracy to further enhance supply chain efficiency and effectiveness. With the emergence

of the concept of demand sensing, one tendency is to attempt to use more downstream data (ideally from point-of-sale) in the forecasting process, as this data reflects more precisely the demand. Unfortunately, collecting good-quality downstream data such as from wholesalers can be a challenge. An alternative, explored in this research, is the use of open orders. Typically, open orders have the advantage of being more readily available since this data comes from within the organization. Our objective is to investigate the added value on forecasting accuracy of using open order data, next to sales history data, in demand forecasting. For this research, we focused on three years of sales and order data for fresh meat products. Forecasts were obtained from dynamic, regularized and support vector regression models. Adopting a proper machine learning approach (sliding window cross-validation) allowed to estimate the out-of-sample performance. Accuracy was compared to forecasts from univariate ARIMA and ETS models. The results show that forecast accuracy can be improved by using the data on open orders. However, as the forecast horizon increases, the improvement in accuracy fades as the added value of the open order data is proportional to the distribution and length of lead times.

Estimating the market potential pre-launch with search traffic

Oliver Schaer (Lancaster University Management School, o.schaer@lancaster.ac.uk); Nikolaos Kourentzes (Lancaster University Management School); Robert Fildes (Lancaster University Management School)

With shorter product life-cycles and increased competition, generating pre-launch forecasts is a vital task for companies. Forecasting the success of a new product is challenging as one need to estimate the market potential. In practice, this is typically done by expert judgment. However, there is substantial evidence that experts are biased when forecasting new products.

Alternative approaches which rely on large surveys or conjoint analysis can be expensive and provide limited data points before launch. Moreover, throughout the pre-launch phase, consumer preference for short life-cycle products change significantly. A potential solution for obtaining more timely data points is to use online user-generated information.

Pre-release buzz reflects the aggregate anticipation of consumers towards a new product. Various studies report improved forecast accuracy when incorporating pre-release information from sources such as search engines, blogs as well as discussions taking part in forums. However, a majority of them only investigate the forecasting potential for the first initial weeks of sales, which from an operational point of view might not suffice, given that traditional time series models a require reasonable size of sales history.

In this research, we provide insights on whether search traffic information from Google Trends is useful in improving the estimation of the market potential before product launch. Search traffic is particularly interesting as it is regarded as a proxy for consumer interest in a product, but it also captures partially marketing expenditures. We (i) develop an approach to augment analogy based information from previous generations with pre-launch search traffic; (ii) compare the forecast performance against forecast from traditional analogy based methods and (iii) investigate the leading properties of pre-launch buzz.

Patterns of interest in industrial brands: Instagram media uploads and sentiment

Harald Schmidbauer (Shanxi University of Finance and Economics, harald@hs-stat.com); Angi Roesch (FOM University of Applied Sciences); Fabian Stieler (University of Augsburg)

Active interest in brands of potential customers manifests itself in their upload behavior to social networks. In the context of this study, the content uploaded consist of media annotated with certain brand-related hashtags, or comments on suchlike media. We analyze the timing of uploads of media or comments and the sentiment they reflect to Instagram with hashtag annotation related to industrial brands in the automotive, food or fashion sector. Using wavelet analysis, we can identify periodic patterns in individual as well as in pairs of hourly upload series and thus obtain insight into leading and lagging properties of the series relating to competing brands (for example, BMW and Mercedes-Benz) and detect changes in patterns as response to the launch of a new product. The approach we suggest is thus able to identify which among several brands is leading at a given time with respect to media postings by Instagram users. These insights can be useful in the context of business analytics, as they provide a means for analyzing, as well as short-term predicting, the attention potential customers pay to industrial brands, and it permits an assessment of how energized potential customers are with respect to media postings on Instagram about brands.

Time series models (I)

Monday, 14:00 - 15:20; Chair: Siuli Mukhopadhyay

Disease Incidence Forecasting using Long Short Term Memory and Autoencoder

Jerelyn Co (Ateneo de Manila University, jerelynco@gmail.com); Maria Regina Justina Estuar (Ateneo de Manila University); Kennedy Espina (Ateneo de Manila University)

Deep learning methods have been gaining popularity for challenging state-of-the-art Artificial Intelligence methods on diverse domains. Despite its promising capability of capturing nonlinear properties on real world datasets, application of these methods to infectious disease forecasting is still minimal. As such, this study presents an investigation on the applicability of Long Short Term Memory (LSTM) in forecasting Dengue Fever, an endemic infectious disease in the Philippines, using historical disease incidence alongside weather conditions datasets. Evaluations were done on three prediction cases: Naive or the baseline prediction, LSTM Prediction Network only (LSTM), and LSTM Autoencoder + LSTM Prediction Network (AE + LSTM). Results showed that the model for LSTM Prediction Network did not significantly outperform Naive Predictions. On the other hand, AE + LSTM obtained the harmonic mean (NRMSE and Correlation) score of 0.712, almost a three-fold jump over the performance scores of Naive at 0.2333 and two-folds over the LSTM models at 0.321. Although this study focuses on the multivariate evaluation of these models, short experiment using the same procedures on univariate data resulted to better harmonic mean performances for AE + LSTM (0.7927) and LSTM (0.6953). This finding suggests that no improvement was made by including the weather conditions as predictors of Dengue Fever incidence.

The Estimation of Long-Memory Spectrum in High Frequency Financial Data

Md Al Masum Bhuiyan (The University of Texas at El Paso, mbhuiyan@miners.utep.edu); Maria C. Mariani (The University of Texas at El Paso); Ionut Florescu (Stevens Institute of Technology)

This paper analyzes the high-frequency financial time series in a long-memory pattern in order to forecast the indices of the stock market. To do so, it studies a sequence of two types of return values generated from stock companies in each minute and on each day. A class of auto-regressive models is presented to study the time series in a stationary environment. The modeling of stationary time series with consistent properties helps to characterize the key variables that incorporate long-rang dependence which has a slowly decaying auto-covariance function. Analysis is made of the long-memory spectrum of each financial time series corresponding to its fractionally integrated parameter. The results suggest that the incorporation of long-memory pattern into the high frequency financial data significantly improves the predictive performance via Maximum Likelihood Estimation. Furthermore, the estimation algorithm proposed in this paper is feasible with large data sets and has good convergence properties.

Multi-step Forecasting with Partial Least Squares

Guillaume Chevillon (ESSEC); Joonsuk Kwon (ESSEC, joonsuk.kwon@essec.edu)

The two main forecasting methods at multistep horizons, say h > 1, consist in the iterated multistep (IMS) and the direct multistep (DMS) techniques. Partial least squares (PLS) estimator, used to forecast using many predictors, was shown to produce an intermediate method between IMS and DMS by Franses and Legerstee (2010). Here, we analyze the relationship between IMS, DMS and PLS

estimators and compare their forecasting power at multi-period horizons analytically, by simulations as well as empirically.

First, we show theoretically how PLS estimators combine the properties of IMS and DMS. We also show that PLS can be superior to both IMS and DMS when the model is misspecified for the dynamics of the errors, in particular when it omits negative serial correlation.

To support our findings, we conduct Monte Carlo simulations and provide analytical results concerning the predictive power of IMS, DMS, and PLS under well- and mis-specified ARMA settings. The types of misspecification we study include under- or over-specification of the dynamics of the error process as well as deterministics breaks. Empirical applications confirm our analytic and simulation results that shows how PLS can yield superior forecast accuracy at multistep horizons.

A new method for forecasting discrete valued time series

Siuli Mukhopadhyay (Indian Institute of Technology Bombay, siuli@math.iitb.ac.in); Vurukonda Sathish (Indian Institute of Technology Bombay)

A new forecasting method based on the concept of profile predictive likelihood function is proposed for discrete valued processes. In particular, generalized auto regressive and moving average models for Poisson distributed data are explored in details. Large sample results are derived for the forecasting distribution. Numerical studies using simulations and a real data set are used to establish the performance of the proposed forecasting method. Highest density regions are used to construct forecasting regions. Robustness of the proposed method to model misspecification is also studied.

Macroeconomic forecasting (I)

Tuesday, 9:40 - 11:00; Chair: George Athanasopoulos

Forecasting GDP Growth with NIPA Aggregates

Christian Garciga (Federal Reserve Bank of Cleveland); Edward Knotek (Federal Reserve Bank of Cleveland, edward.knotek@clev.frb.org)

Beyond GDP, which is measured using expenditure data, the U.S. national income and product accounts (NIPAs) provide an income-based measure of the economy (gross domestic income, GDI), a measure that averages GDP and GDI, and various aggregates that include combinations of GDP components. This paper compiles real-time data on NIPA aggregates and uses these measures in simple time-series models to construct out-of-sample forecasts for GDP growth. Over short forecast horizons, NIPA aggregates—particularly consumption and GDP less inventories and trade—together with these simple time-series models have historically generated more accurate forecasts than a canonical AR(2) benchmark. This has been especially true during recessions, although we document modest gains during expansions as well.

Looking for the stars: unobserved components model approach to estimate natural real rate

Irma Hindrayanto (De Nederlandsche Bank, a.i.w.hindrayanto@dnb.nl); Mengheng Li (De Nederlandsche Bank)

The natural rate of interest is a long-run equilibrium or steady-state concept in the DSGE literature. It is clear that although such long-run equilibrium are treated as fixed in the DSGE framework, natural rates may be subject to low-frequency fluctuations due to advancement of technology and changing time preference of the representative agent that are difficult to detect. Holston, Laubach and Williams (2017) provide a New Keynesian modeling framework based on a Phillips curve and an intertemporal IS curve to describe the stochastic driving forces behind output gap and real interest rate gap. They also allow for low-frequency gradual shifts in the potential growth rate of output and therefore in the natural rate of interest. Empirically, however, their model does not seem to capture the output gap satisfactorily compared with estimates produced by government agencies such as Congressional Budget Office for the US or the OECD for some European countries. We propose a method that is stable and relatively robust for different economies. Our model is a modification of Holston et al (2017) where we first pin down the potential growth rate of output via Okun's law. Using the estimated potential growth rate of output, we then estimate the Phillips curve, IS curve, potential output and natural real rate simultaneously using a full-fledged unobserved components model with similar cycles. We test our model for the US, UK and EA data.

Structural Transformation in Africa: A Predictive View

Adusei Jumah (Central University, aduseijumah@gmail.com); Robert Kunst (University of Vienna)

Structural transformation refers to the systematic changes in sector proportions as economies grow. Initially, at low levels of income, agriculture dominates both as a share of GDP and as a share of employment. As economies transition toward middle levels of income per capita, agriculture accounts for a smaller share of both GDP and employment (replaced by industry and services). This pattern is a central feature of economic development and both a cause and an effect of economic growth. Yet, even for middle-income countries of Africa, there is the tendency to prescribe agriculture as the sector with the most potential for economic growth.

We adopt both a dualistic perspective in which an economy consists of agriculture and non-agriculture, and a multi-sectorial perspective in which an economy consists of agriculture, industry and services to determine the engine of growth in the respective African countries within a vector autoregressive framework using panel data. The model is evaluated by means of impulse response functions and various forecasting experiments.

Macroeconomic forecasting for Australia using a large number of predictors

Bin Jiang (Monash University); George Athanasopoulos (Monash University, Australia, george.athanasopoulos@monash.edu); Rob Hyndman (Monash University, Australia); Anastasios Panagiotelis (Monash University, Australia); Farshid Vahid (Monash University, Australia)

A popular approach to forecasting macroeconomic variables is to utilize a large number of predictors. Several regularization and shrinkage methods can be used to exploit such high-dimensional datasets, and have been shown to improve forecast accuracy for the US economy. To assess whether similar results hold for economies with different characteristics, an Australian dataset containing observations on 151 aggregate and disaggregate economic series as well as 185 international variables, is introduced. An extensive empirical study is carried out investigating forecasts at different horizons, using a variety of methods and with information sets containing an increasing number of predictors. In contrast to other countries the results show that it is difficult to forecast Australian macroeconomic variables more accurately than some simple benchmarks. There is little to no improvement in forecast accuracy when the number of predictors is expanded beyond 20-40 variables and international factors do not help.

Probabilistic energy forecasting

Tuesday, 9:40 - 11:00; Chair: Jooyoung Jeon

On the Creation of Quantile Regression based Scenario and Parametric Probabilistic Forecasts in the Context of Energy Forecasting

Jorge Ángel González Ordiano (Karlsruhe Institute of Technology, jorge.ordiano@kit.edu); Riccardo Remo Appino (Karlsruhe Institute of Technology); Timm Faulwasser (Karlsruhe Institute of Technology); Veit Hagenmeyer (Karlsruhe Institute of Technology); Ralf Mikut (Karlsruhe Institute of Technology)

The current transition from traditional energy supply (as e.g., coal power plants) to a decarbonized energy system is accompanied by new and challenging problems. For instance, weather-dependent renewable energy systems, as e.g., wind and solar power, complicate with their volatility the necessary balancing of energy demand and supply. Therefore, forecasting models able to accurately predict the future generation and load have become essential in the planning and scheduling of energy grids. However, most forecasting methods are unable to quantify the forecast uncertainty, which if known could significantly influence optimal decision making. Probabilistic forecasts are able to describe this uncertainty, hence increasing the interest of the energy forecasting community towards them. Non-parametric probabilistic forecasts based on quantile regressions have become one of the preferred methods in the literature, as they avoid assumptions regarding the distribution of the forecast uncertainty. Nonetheless, there are several optimization and scheduling algorithms (as., e.g. model predictive control) that require either scenario forecasts or parametric probabilistic forecasts. The former are comprised of a collection of possible time series scenarios, for which an explicit description of the correlation structures between neighboring time series values is required, while the latter assume that the uncertainty follows a given parametric distribution (e.g., Gaussian).

For this reason, the present contribution describes two methods that convert quantile regressions into scenario or parametric probabilistic forecasts. An important advantage of the presented method consists in the creation of scenario forecasts without the explicit description of the correlation between neighboring forecast values. The second method estimates a parametric distribution that best fits the forecast uncertainty, a goal that is achieved without the need of a distribution assumption. The simplicity and straightforwardness of the presented methods allow for their easy integration into the forecasting service of the EnergyLab 2.0, in which an automated and computationally efficient forecasting of load and volatile renewable power is of interest. Finally, the accuracy of the obtained forecasts is evaluated in the present work using the data provided for the Global Energy Forecasting Competition of 2014 (GEFCom14).

Time Series Clustering with Dynamic Time Warping for Energy Forecasting Data Cleansing

Garrett Frere (SAS Institute, garrett.frere@sas.com); Tae Yoon Lee (SAS Institute)

Load forecasting commonly uses hierarchical forecasting to group time series based on disaggregation rules. In the case of load forecasting a single hierarchy can be used to describe the physical inclusiveness of various customer classes. At the highest layers of the hierarchy distinctions are frequently tangible: total transmission with substations and circuit breakers underneath. There can also be cases where secondary hierarchies are implemented to group time-series based on other descriptors, these can be non-tangible such as rate classes.

Data cleansing is the process of finding and correcting incorrect, inaccurate, and incomplete records. Load forecasting as well as other forms of hierarchical forecasting can leverage the formalized

hierarchy to impute and correct data. For instance a naïve model without segmentation for a circuit breaker may use only the one pattern derived off the substation to impute data. This is despite the fact that group time-series frequently show patterns beyond the single pattern described in the hierarchy. Dynamic Time Warping is a form of sequence alignment that calculates the optimal match between two temporal signals. It can be used for time-series classification. DTW "warps" time sequences such that sequences of different lengths are stretched or compressed such that they can be compared. Time series clustering is similar to clustering but the similarity measure incorporates the temporal nature of time sequences. Time series clustering algorithms are commonly variations of hierarchical clustering and partitional clustering. In the case of load forecasting our method uses DTW to build a distance matrix which can be fed to hierarchical clustering algorithms.

Time series clustering with Dynamic Time Warping can be used instead of a naive model to segment data within a circuit breaker and capture patterns by segment. When data imputation occurs the imputed value is derived from the segmented pattern instead of the naive model that depends on the single pattern from the entire circuit breaker. This can result in more accurate identification of outliers and missing value replacement.

Load forecasting has four decomposition types: seasonality, cycle, trend, and irregularity. Time series clustering with Dynamic Time Warping can be used to segment by those four types. In our approach, rather than applying all types of models for forecasting the forecast can be derived more specifically from the selected pattern segment. Such an approach can speed up the forecasting process and reduce fitting incorrect models to segmented data. Granularity of forecasting segments will inform the granularity of the data cleansing segments.

Forecasting Design Day Demand Using Extremal Quantile Regression

Jarrett Smalley (Marquette University GasDay Laboratory, jarrett.smalley@marquette.edu); David Kaftan (Marquette University GasDay Laboratory); Ronald H. Brown (Marquette University GasDay Laboratory); George F. Corliss (Marquette University GasDay Laboratory); Richard Povinelli (Marquette University GasDay Laboratory)

Extreme events occur rarely, making them difficult to predict. Extreme cold events strain natural gas systems to their limits. Natural Gas distribution companies need to be prepared to satisfy demand on any given day that is at or warmer than an extreme cold threshold. The hypothetical day with temperature at this threshold is called the Design Day. To guarantee Design Day demand is satisfied, distribution companies need to determine the demand that is unlikely to be exceeded on the Design Day.

We approach determining this demand as an extremal quantile regression problem. We review current methods for extremal quantile regression. We implement a quantile forecast to estimate the demand that has a minimal chance of being exceeded on the design day. We show extremal quantile regression to be more reliable than direct quantile estimation. We discuss the difficult task of evaluating a probabilistic forecast on rare events.

Probabilistic forecasting is a quickly growing research topic in the field of energy forecasting. Our project contributes to this field in three ways. First, we forecast quantiles during extreme cold events where data is sparse. Second, we forecast extremely high quantiles that have a very low probability of being exceeded. Finally, we provide a real world scenario on which to apply these techniques.

Probabilistic Forecasting for Aggregated Demand Optimised for Peer to Peer Energy Markets

Estêvão Alvarenga (University of Bath); Jooyoung Jeon (University of Bath, j.jeon@bath.ac.uk); Ran Li (University of Bath); Fotios Petropoulos (University of Bath)

To minimise forecast uncertainty in Peer to Peer energy trading, which aims to utilise surplus in renewable energy generation locally, we develop probabilistic forecasts for aggregated demand rather than individual households. Our study compares different ways of composing customer aggregation, which include random sampling and stochastic optimisation under forecast uncertainty. The empirical study involves hourly smart meter times series from Ireland and Korea, and evaluate the methods using forecasts up to 24 hour ahead.

Supply chain forecasting (I)

Tuesday, 9:40 - 11:00; Chair: Yves Sagaert

Forecasting Spare Part Demand using Service Maintenance Information

Sarah Van der Auweraer (KU Leuven, sarah.vanderauweraer@kuleuven.be); Robert Boute (KU Leuven)

We focus on the inventory management of spare parts that are used for service maintenance. These parts are commonly characterized by a large variety and an intermittent demand pattern. When a service part is required but not immediately available, the incurred shortage costs may be substantial. Specialized service parts models should therefore focus on improving the availability of parts whilst limiting the investment in inventories. We develop a method to forecast the demand of these spare parts by linking it to the service maintenance policy (either preventively or upon failure). As the demand of these parts originates from the maintenance activities that require their use, it is related with the number of machines in the field that make use of this part (known as the installed base), in combination with the part's failure behaviour and its maintenance policy. By tracking the installed base (through machine sales and discards) and estimating the part failure behaviour, we provide a forecast of the future spare parts demand during the upcoming lead time. Our work is validated by a numerical experiment. We show that our policy has the potential to provide cost savings compared to the traditional forecasting techniques for intermittent spare part demand.

Forecasting the Time to Repair for Automotive Parts: An Ordinal Logit Model using LASSO Selection Techniques

Shixuan Wang (Cardiff University, shixuan_wang@hotmail.com); Ying Liu (Cardiff University); Carla Di Cairano-gilfedder (BT TSO, Research & Innovation); Gilbert Owusu (BT TSO, Research & Innovation); Aris Syntetos (Cardiff University)

The Time to Repair (TTR) is crucial for the garage management in terms of scheduling and planning. Availability of reliable TTR forecasts when the vehicles arrive may largely improve the efficiency of the garage and reduce the waiting of drivers. It is common that the distributions of TTR are multi-modal, which are associated with different failure modes. However, the failure modes cannot be identified before the (preliminary or full) diagnosis on faulty parts. In this work, we developed a forecasting method based on the ordinal logit model to predict the TTR with the consideration of a large number of exogenous variables, such as condition information, manufacturing information, and geographical information related ones. The forecasting performance has been examined by an empirical dataset of actual hours on repairing automobile parts (engines and transmissions), provided by a sizeable fleet company in the United Kingdom.

The purpose of our work is to support garage managers to make better scheduling. It is more natural to forecast the range of the TTR, rather than the precise point prediction, because the scheduling is usually made in a block of time. Our forecasting method is based on the ordinal logit model, a technique used frequently in studies of dependent variables taking on only a finite number of values under a natural ordering. In our case, the TTR can be classified as ordinal categories by the number of multiples of 0.5 hours. Since we consider a large number of exogenous variables and their interactions, the standard estimation techniques are inefficient due to the "curse of dimensionality". To tackle the dimensionality problem, we propose to employ the least absolute shrinkage and selection operator (LASSO) method to regularise the estimation results and automatically choose the best combination of exogenous variables and their interactions.

The existed literature on forecasting TTR is insufficient and most of the relevant work is designed for very special systems, such as petroleum production facilities (Gao et al., 2009) and mining equipment (Barabadi et al, 2011). This research fills the literature gap by proposing a novel forecasting method for the TTR of the automobile parts and providing an empirical analysis based on an actual dataset. Specifically, our contribution is threefold: i) we show the merits of our forecasting method for the TTR of automobile parts; ii) we identify the important exogenous variables affecting the TTR for our empirical dataset; iii) our forecasting results have been compared with the benchmark of the proportional repair models and show superior predictive power and outperform the benchmark in out-of-sample predictions. Our developed method can support garage managers to make better scheduling decisions and facilitate the operational management of the garage.

The inventory impact of including macroeconomic leading indicators in global supply chain management

Yves R. Sagaert (Ghent University, yves.sagaert@ugent.be); Nikolaos Kourentzes (Lancaster University); Stijn De Vuyst (Ghent University); El-houssaine Aghezzaf (Ghent University)

Supply chain management is increasingly performed at a global level. The decision process is often based on tactical sales forecasts, which has been shown to benefit from including relevant exogenous information. Leading indicators that cover different aspects of macroeconomic dynamics are appealing in this context, as macroeconomic dynamics in target countries can affect companies end markets. Even though this information can be beneficial on a tactical level, it remains unclear how this information can impact sales forecasts at Stock-Keeping-Unit (SKU) product level, due to increased levels of noise and products having differing demand patterns and dynamics, masking macro-effects. Nonetheless, hierarchical forecasting can be used to reconcile macroeconomic leading indicators from tactical level forecasts to detailed SKU levels, and vice versa. In this paper, we evaluate the feasibility and benefits of merging tactical and operational forecasting, where higher level forecasts include leading indicators, in contrast to univariate SKU operational predications. We present a framework that identifies automatically the most relevant leading indicators on global sales level, and by exploiting the hierarchical product structure, carries this information to sales forecasts at SKU product level. For our evaluation we rely on inventory metrics obtained from simulation experiments, reflecting the associated supply chain risk.

Finance (II)

Tuesday, 11:20 - 12:40; Chair: Maria Rosa Nieto Delfin

Exchange rate forecasting for BRICS: The role of fundamental predictors

Ilse Botha (University of Johannesburg, ilseb@uj.ac.za); Marinda Pretorius (University of Johannesburg)

Forecasting currency exchange rates is important in business risk management tasks. Some empirical evidence does exist that monetary fundamentals may contain predictive power for exchange rate movements in the long run although they do not have short-run predictive power.

The purpose of this paper is to provide evidence of the predictability of exchange rates among the BRICS (Brazil, Russia, India, China and South Africa) countries using fundamental predictors. We compare a fundamental model based on the cointegration-ECM framework to the random walk model using traditional statistical measures of forecasting error at various horizons.

Forecasting stock yearly performance after three months - recurrent neural network approach

Adam Chudziak (SGH Warsaw School of Economics, achudz@sgh.waw.pl)

One of the successfully evaluated investment strategies are momentum strategies. Their simple, yet statistically verified, underlying concept is that stocks which performed well over one period, will perform well in the following period. This allows for a period-to-period qualitative behavior prediction. This work examines the idea further, aiming at identification of other statistical patterns which foreshadow subsequent rise or fall of the price. Using the LSTM networks and other statistical learning techniques we try to find other patterns which allow for qualitative period-to-period predictions. The preliminary results of this study will contribute to the ongoing debate surrounding the Efficient Markets Hypothesis.

Forecasting Based Metal Prices with Commodity Currencies

Pablo Pincheira (Universidad Adolfo Ibáñez, pablo.pincheira@uai.cl); Nicolás Hardy (Universidad Adolfo Ibáñez)

In this paper we show that the Chilean exchange rate has the ability to predict the returns of the London Metal Exchange Index and of the six primary non-ferrous metals that are part of the index: aluminum, copper, lead, nickel, tin and zinc. The economic relationship hinges on the present-value theory for exchange rates, a floating exchange rate regime and the fact that copper represents about a half of Chilean exports and nearly 45% of Foreign Direct Investment. Consequently, the Chilean peso is heavily affected by fluctuations in the copper price. As all six base metal prices show an important comovement, we test whether the relationship between copper prices and Chilean exchange rates also holds true when it comes to the six primary non-ferrous metals. We find interesting evidence of predictability both in-sample and out-of-sample. Our paper is part of a growing literature that in the recent years has evaluated and called into question the ability of commodity currencies to forecast commodity prices.

The effect of capitalization levels on systemic risk for Mexican banks

Maria Rosa Nieto Delfin (Universidad Anahuac Mexico, maria.nieto@anahuac.mx)

In this paper we propose to measure the effect of capitalization levels of a financial institution on the systemic risk through the SRISK index. The aim is to analyze that systemic risk decreases as banks maintain higher levels of capitalization with respect to the value of its assets. Systemic risk is defined as the possibility that a negative event of a certain bank could cause instability of the finance system or the entire economy. In this research we measure systemic risk of Mexican banks using an index called SRISK that is defined as the expected capital shortfall of a financial entity conditional on a prolonged market decline. It estimates the amount of capital that a bank need to operate when it faces a financial crisis. In this paper, the GJR-GARCH model is used for the estimation of SRISK. The model is set up by analyzing two of the most important banks in Mexico, Grupo Financiero Inbursa and Grupo Financiero Banorte, from December 2007 to December 2016. We conclude that these two Mexican banks have low exposure to systemic risk since they maintain capital levels above minimum requirement. Mexican regulation should consider SRISK as an alternative measure to assess the capital adequacy of banks.

Tourism and transportation

Tuesday, 11:20 - 12:40; Chair: Miriam Scaglione

Modeling Airbnb demand to New York City employing panel data at the listing level

Irem Onder (MODUL University Vienna, irem.onder@modul.ac.at); Ulrich Gunter (MODUL University Vienna); Bozana Zekan (MODUL University Vienna)

This research employs a panel data set on Airbnb listings in New York City, which is sourced from AirDNA and comprises NT = 478,386 observations for the time period 2014M09 to 2016M06. After seasonal adjustment and statistical pre-testing, a standard tourism demand model with the occupancy rate as dependent variable as well as the relative listing price and source-market-weighted real gross domestic product as explanatory variables is estimated. To this end, a one-way fixed effects panel regression model with cluster-robust standard errors is employed. Since all variables have been transformed to natural logarithms before estimation, the regression coefficients can be interpreted as price and income elasticities of demand, respectively. In addition to an overall satisfying model fit, price and income elasticities are highly statistically significant and feature the algebraic signs expected from microeconomic theory. However, in order to be able to exploit the good predictive ability of panel-data models in forecasting in terms of producing accurate out-of-sample forecasts, sub-samples of the panel data set should be created according to the different characteristics of the listings. While Airbnb demand in Manhattan, for instance, is generally price-inelastic, Airbnb demand outside this borough is price-elastic and comparatively more income-elastic. Similar findings also hold for those listings, which are not entire homes or apartments, for those which are not offered by commercial Airbnb providers, as well as for combinations of these characteristics.

Probabilistic Analysis of Cycle Length for Signalized Intersections in Transportation Engineering

Chandra Putcha (California State University, Fullerton); Brian Sloboda (University of Phoenix, School of Advanced Studies, Center for Management and Entrepreneurship, bsloboda@email.phoenix.edu)

This paper deals with application of principles of probability theory to the problem of determining cycle length for signalized intersections in the field of transportation engineering. Determination of signal timings is an important aspect dealt with by transportation engineers. An important parameter that determines the actual signal timings is the selection of cycle-length (CL). The appropriate cycle length is determined by Webster formula (Papacostas and Prevedouros, 2001). As per that, the cycle length is a function of total lost time, L, and CS (sum of the flow ratios of critical movements). These critical movements typically involve the estimated volume of traffic in various directions at a signal intersection (through traffic, Through + Right, Through + Left , Left only and Right only). In this formula, L can be considered as a deterministic quantity while CS is essentially a probabilistic quantity. While a traffic designer estimates the % of volume of traffic in any of the possible directions stated above, these are not exact. A method is suggested in this paper to perform probabilistic analysis of CS. This is done by using the well-known technique of Monte Carlo simulation (Ang and Tang, 2007). Monte Carlo simulation involves creating a probabilistic equivalence to a deterministic parameter. parameters are obtained, a statistical analysis of these parameters is performed and the necessary probability values can be calculated for various limit states. Using that, a set of random numbers are generated (typically 100). These random numbers are then associated with the deterministic estimated % volume in each of the traffic directions an correct values of the critical sum (CS) are then calculated and the revised cycle lengths are calculated. In this study, only one random variable of critical sum (CS) is considered while other variable is treated as a deterministic quantity as

it is essentially a two-variable problem. This paper will give both the traditional values and revised values of cycle lengths (CL) using the present study based on the application of probability theory. Application of probabilistic principles to transportation problems has been done by many authors (Putcha, 2007). The results will be very useful to traffic engineers in better designing the traffic signal timings for optimum efficiency.

Estimating the optimal momentum for launching last minute deals in self-catering accommodation. The case of Valais Switzerland

Miriam Scaglione (Institute of Tourism, University of Applied Sciences and Arts Western Switzerland Valais, miriam.scaglione@hevs.ch); Pascal Favre (Institute of Tourism, University of Applied Sciences and Arts Western Switzerland Valais)

The perishable nature of travel and tourism services is one of the principal drivers for suppliers to offer “last-minute deals”. Timing and pricing are central to an optimization strategy in order to develop last-minute deals for the accommodation service providers. A previous study (Scaglione, Johnson, & Favre, 2017) analysed the median booking period (BP), namely the time lapse between the booking and the actual travel date. The results showed that the segmentation of the booking transaction based solely on the country of origin of the client and the season in which he/she travels gives sound estimates of median BP. The study used 141,000 transactions spanning from 1st January 2010 to 1st February 2017 on 19 different destinations in Valais. Using Kaplan-Meier (KM) survival modelling and resampling techniques, BP yields the median estimation by looking to the corresponding life tables. In IFS 2017, the segmentation based on origin/season that had yielded satisfactory estimations of the median time of booking period (Scaglione et al., 2017) fails in the case of 95% quantile estimation using the same set of data and methodology. Scaglione, Johnson, and Favre (2018) showed, using logistic regression of the binary variable “last minute transaction or not”, that other transactions variables such as the length of stay, the chosen destination in Valais or the party composition (couples, family, group) are significant (cfr.: https://www.tourobs.ch/5700.aspx). The aim of this research is to estimate 95% quantiles using more fine grain clustering of the raw data. .Also, we will explore the quantile regression forest methods and a comparison of both methods will be showed. The authors used SAS Institute 9.4 software (proc survey select) and programmed a customized routine as well as R with the package 'quantregForest' in order to produce and analyze the simulated results.

Software and support systems

Tuesday, 11:20 - 12:40; Chair: Rob Hyndman

Data-cube forecasting: forecasting level selection

Igor Gusakov (GoodsForecast, igusakov@gmail.com)

In this study, I develop ideas, described in my Foresight article "Data-Cube Forecasting for the Forecasting Support System" (Foresight, Issue 35, Fall 2014). In this article I propose to use OLAP-cubes as a base for the FSS development, because most OLAP systems already have everything necessary for the perfect FSS - except forecasting itself. To compensate this drawback, I suggested to forecast entire cube (that is - forecast every possible combination of time series inside the cube and then align all forecasts using special technique). Although theoretically interesting, this idea is not applicable in practice due to the huge number of time series to forecast and necessity to solve optimization task afterwards, which ends up with such resource consuming operation as matrix inversions.

After several years of FSS development, I want to share simple and effective methodology for the data-cube forecasting, which is applicable in practice and is not resource consuming. The main idea is to divide original data set, used for the cube creation, into an arbitrary number of sub-sets, thus creating a selection of sub-cubes. Inside every sub-cube arbitrary forecasting level can be defined, which allows flexible time series selection inside the cube.

Although this approach seems to be quite simple, its implementation in the real FSS has some nontrivial aspects. I want to share some real cases our team had to solve during FSS development.

RegressIt: An Interface Between Excel and R

Robert Nau (Duke University, robert.nau@duke.edu)

Excel and R have their own advantages as environments in which to work with data. Excel is almost universally understood and its files are easily navigated and shared and used as platforms for decision models and presentations, but it comes with only primitive statistical tools. R has an almost unlimited toolkit for analysis, but it is not accessible to non-programmers. This talk will demonstrate a free Excel-in, RegressIt, that performs graphics-rich linear and logistic regression analysis within Excel and also includes a 2-way interface with R. It provides a menu-driven front end to R that makes its regression and time series modeling tools accessible without any writing of code, and it also allows Excel to be used as a back end to R for producing additional presentation-ready output in spreadsheet format for linear and logistic regression models. Only a few keystrokes are required to go in either direction. For practitioners, this can allow Excel and R to more easily leverage each other. For educators, it can lower the barrier to introducing larger populations of students to the use of R, and the add-in also includes extensive built-in teaching notes. These features are controlled from a novel ribbon interface that greatly simplifies navigation of the output space and the model space, making analysis and model comparisons more efficient. The table and chart output that is produced is highly interactive, and the interface allows a fully detailed analysis to be driven on a touchscreen.

Tidy forecasting in R

Rob Hyndman (Monash, rob.hyndman@monash.edu); Mitchell O'hara-wild (Monash University); Earo Wang (Monash University)

The forecast package in R is widely used and provides good tools for monthly, quarterly and annual time series. But it is not so well-developed for daily and sub-daily data, and it does not interact easily with modern tidy packages such as dplyr, purrr and tidyr.

I will describe our plans and progress in developing a collection of packages to provide tidy tools for time series and forecasting, which will interact seamlessly with tidyverse packages, and provide functions to handle time series at any frequency.

Time series models (II)

Tuesday, 11:20 - 12:40; Chair: Clara Cordeiro

Demand forecasting: an industry case study

Inga Maslova (University of Southern California, imaslova@marshall.usc.edu)

Deep learning techniques are rapidly gaining popularity in a variety of industries. Developing accurate forecasting models can help ensure success in a corporate environment. Specifically, accurate demand prediction can empower decision making and expand vital planning capabilities of a company. These models can allow better negotiation on flexible shipping rates, as well as avoiding premium fees.

This talk will go over an industry use case that focuses on weekly demand forecasting using both classical and modern machine learning techniques. It was performed following the CRISP-DM approach, utilizing a recurrent neural network model, and launched to production. As the result, the developed forecasting model was used to propose a flex-allowance shipping contract with estimated savings of $2M over a 9-week period.

Partial varying-coefficient regression and autoregressive model for nonlinear time series

Zhiqiang Cao (The Hong Kong University of Science and Technology, zcaoae@connect.sut.hk); Man-Yu Wong (The Hong Kong University of Science and Technology)

Regression and autoregressive mixed model is a popular model to analyze the relationship between response variable and its covariables of time series. The coefficients in classical regression and autoregressive mixed are constants. However, in some complicated cases, the coefficients of covariables may not be constant, it changes with some explored variable. In this paper, we propose a kind of partial varying-coefficient linear regression and autoregressive model for nonlinear time series.

Our proposed model can be regarded as the extension of the functional-coefficient autoregressive regression (FAR(p)) (Cai, et al., 2000) model and the classical regression and autoregressive mixed (CRAM) model, so it has both advantages of the FAR(p) model and the CRAM model. Moreover, this model can handle autocorrelation and heteroscedasticity of a time series. The B-spline approach is used to estimate parameters in the model, which is fast and easy to implement. And we use a modified multifold cross-validation method to determine the optimal degree of freedom and number of knots in the B-spline. Simulation studies show that the performance of the B-spline approach is very satisfactory.

To illustrate the application of the proposed model, a real data about Lake Shasta inflow is analyzed, and this data can be downloaded from the R package "astsa". According to the performance of 3-step forward predictions of the inflow Lake Shasta based on the climatic factors, our proposed model performs better than the functional-coefficient autoregressive regression model as well as the classical regression and autoregressive mixed model.

Polynomially Adjusted GARCH Financial Risk Forecasting

Luca Bagnato (Università Cattolica S.C.); Maria Grazia Zoia (Università Cattolica S.C.); Valerio Poti (University College Dublin, valerio.poti@ucd.ie)

Orthogonal polynomials can be employed to tailor finite-moment symmetric parent probability density functions to fit target densities. In this paper, we propose a framework that exploits this possibility for the purpose of estimating and forecasting the volatility of financial asset returns. Our proposed forecasting models are GARCH models with innovations having polynomially modified distributions. We demonstrate the advantages of this approach in the context of an empirical application to forecasting the volatility of the return on range of currencies, in which we test for the superior out-of-sample predictive ability (SPA) of this specification compared to extant GARCH models.

Xtreme time series modelling

Clara Cordeiro (University of Algarve, ccordei@ualg.pt); Manuela Neves (University of Lisbon)

Statistical analysis of extreme values was traditionally applied to hydrology and insurance. Nowadays, there is a quite large variety of fields of application of extreme value theory such as Climatology, Oceanography, Environment and Biology.

Unlike most traditional central statistical theory, which typically examines the usual (or the average) behavior of a process, extreme value theory deals with models for describing unusual behavior or rare events. The heart of extreme value theory is the reliable extrapolation of values beyond the observed range of sample data. Modeling rare events of univariate time series is an area of important research. In classical time series modeling, a key issue is to determine statistically how many parameters have to be included in the model. However, special care must be given to extreme events in the series that need specific statistical procedures based on the behavior of extremes.

Extreme value models were initially obtained through arguments that assumed an underlying process consisting of a sequence of independent and identically random variables. However, in many situations where extreme value models are of great interest to be applied, temporal independence is unrealistic. The most natural generalisation of a sequence of independent random variables is a stationary setup. In the last decades, many progresses have been made in parameter estimation of extreme values in time series, with relevance to asymptotic results. However, for finite samples, limiting results provide approximations that can be poor. Computer intensive methods, among which we refer to Generalised Jackknife and Bootstrap methodologies, have recently shown to improve results in parameter estimation in statistics of extremes.

Resampling techniques, such as Bootstrap, has revealed to be a good option in the forecasting context. Therefore, its application to time series extremes in order to obtain good resampled extreme values has become a challenge. The idea is to build bootstrap estimators versions for parameters of extreme events in order to overcome the difficulties that classical estimators have shown. The objective of this new approach is to improve the performance of the estimators addressed in extreme value theory for dependent sequences through more efficient bootstrap procedures.

Acknowledgments: Research partially supported by FCT- Fundação para a Ciência e Tecnologia, through the project FCT Portugal UID/MAT/00006/2013.

Macroeconomic forecasting (II)

Tuesday, 14:00 - 15:20; Chair: Sylwia Roszkowska

A New Macroeconomic Risk Indicator: Differences between Developed and Developing Countries

Gloria González-Rivera (University of California, Riverside); Esther Ruiz (Universidad Carlos III, ortega@est-econ.uc3m.es); Javier de Vicente (Universidad Carlos III)

In this paper we construct a time-varying macroeconomic risk index that could serve as an early warning signal for macroeconomic recessions. These indicators are constructed using annual data on macroeconomic growth, observed from 1985 to 2015 for 87 countries with different characteristics (industrial, emerging and other developing countries). The proposed methodology is based on a factor augmented predictive regression model, extracting the factors by Principal Components, and computing its uncertainty using subsampling procedures.

In order to construct the risk index, we consider the value-in-stress risk measure proposed by Gonzalez-Rivera (2003). We find that developing and developed countries have different risk structures and are differently affected by the global macroeconomic risks.

Macro Forecasting of National Health Expenditure in Palestine

Rabeh Morrar (An-Najah National University, rabeh.morrar@najah.edu)

Over the past decades, several health outcomes have improved in Palestine. Life expectancy has increased, infant mortality rates have decreased, vaccination rates remain sustainable high and higher than many regional neighbors. Nevertheless, Palestine faces evolving health challenges. The Palestinian population is undergoing rapid epidemiological transition: Non-communicable diseases, such as cardiovascular diseases and hypertension, diabetes, and cancer, are the main causes of morbidity and mortality. The availability of many needed health services remains sub-optimal, and access issues due to movement restrictions imposed by the Israeli Government remain.

The logic behind forecasting the health care expenditure in Palestine is that it is very important for health care managers and policy makers in order to make decisions about future service. Forecasts enable them to anticipate the future, to estimate future demand in order to build strategies and allocate resource efficiently, and to manage risks and plan accordingly. Good forecasts are essential for planning (short, medium, and long-term), and are essential inputs for all types of service production systems.

This paper tries to forecasting the national health expenditure indicators and to examine the relationship between national health expenditure and its determinants, in order to help decision makers to analyze health expenditures and to set the relevant policies toward rationalizing the expenditure on health. Also, the study aims to analyze the relationship between macro and socioeconomic factors as exogenous variables and health expenditures. Two methods were used in this study, forecasting health expenditures and Granger-Causality Relations. The method for forecasting is mainly based on the E-Views software in order to forecast the health expenditure from financing agencies perspective. To assess the causality between health spending and exogenous variables, the test involves estimating the simple vector auto regressions (VAR). Data was obtained from the Ministry of Health reports and National Health Accounts reports between 2000-2014. The determinants of health expenditure were based on the prevalence rates of selected NCDs from the studies obtained from Institute of Community and Public Health (ICPH) at Birzeit University. We found that national health expenditure was estimated in 2015 to be $1.450 billion USD, growing at 7%

annually, this due to expected increasing government health expenditure and household spending 5% and 7% respectively compared with 2014. It is anticipated that this will reach to more than 2000 million USD in 2020. Household spending is expected to increase whereas government health spending and nonprofit institute are expected to decrease. Chronic diseases comprise one of the most important variables which might significantly lead to increases in total health expenditures. Macroeconomics variables such as change in GDP per capita and population growth positively affect the overall total health expenditure. Other socioeconomic variables such as number under the age of 15 and above 65, numbers of physicians per 1000 population and beds per 1000, show no relationship with national health expenditures. For this, reviewing and modifying health financing from the three-main functions is needed ae well as reforming health insurance scheme.

Forecasting regional quarterly GDP using CVAR framework

Sylwia Roszkowska (University of Lodz / Narodowy Bank Polski, sylwiaroszkowska@gmail.com)

The analyses of economic differences between regions commonly need data of higher frequency than that offered by official statistics. The main goal of the paper is to create higher frequency data at regional level. The article presents the application of cointegrated VAR model to temporal disaggregation of regional GDP.

The forecasts of the disaggregated observed time series will be prepared using a CVAR model and the additional macroeconomic indicators (including labour market outcomes, price indices, gross fixed capital formation, industry production and other as necessary).

The proposed method consists of two stages. In the first step the parameters of the system of equations linking the GDP at national level with the structural macroeconomic variables are estimated. In the second step, on the basis of the estimates from the first step, the estimated quarterly GDP for individual regions is treated as a function of estimated parameters.

The data covers Polish economy and the period from 1995:01 to 2017:04. The analysis (besides forecasting reasons) allows for identification of the long run relations, common stochastic trends and the estimation of the adjustment dynamics in the system. Having estimated parameters of aforementioned CVAR, we use the structural parameters to get quarterly GDP at regional level.

Macroeconomic data obtained as a result of the forecasting process indicate that economic dynamics of Polish regions is diversified. Moreover, analysed fluctuations at the national level and based on them fiscal and monetary policy seem to be ineffective for certain regions.

Electricity price forecasting

Tuesday, 14:00 - 15:20; Chair: Angi Roesch

Mid-Term Stochastic Modeling of Energy Markets and its Applications

Eina Ooka (The Energy Authority, eooka@teainc.org)

Historically energy utilities have focused on developing short-term market models for operational purposes, and long-term models for investment and regulatory purposes. However, mid-term models ranging from a month to several years has been less explored. I will present an approach to building an all-encompassing stochastic mid-term model for markets, resources and portfolios. This is based on a model that has been successfully operational in the real world for active portfolio management and trading.

Some of the challenges we faced include multiple collinear price nodes, heteroscedasticity, non-normal distributions, different time granularities, and non-linear variable dependencies. We addressed these issues using a combination of multiplicative seasonal decomposition, adjustment through hierarchical models, seasonal bootstrapping, and use of predictive models such as Random Forest and KNN.

We will end the talk with examples of real-world applications such as delta-hedging and budgeting.

Day-ahead electricity price forecasting under changing liquidity conditions: a case study of the Japan Electric Power Exchange

Oliver Senter (PwC Advisory LLC, oliver.db.senter@pwc.com); Arthur El Medioni (PwC Advisory LLC)

Since the deregulation of the retail power market in April 2016, the traded volume of day-ahead electricity on the Japan Electric Power Exchange (JEPX) has grown rapidly from around 2% of power consumption before deregulation to 7% in late 2017. In addition, changing economic liquidity, as measured by methods such as the Amihud price impact, can be observed over the same period. These rapidly changing market conditions create an ideal scenario for analysing the impact of liquidity on optimal model selection for short-term electricity price forecasting (EPF).

Firstly, the JEPX spot price time series is segmented into two distinct periods with significant differences in liquidity. Secondly, several commonly used models for EPF, such as the naive method, ARIMA and ANN, are applied and the forecasting performance, measured by RMSE and MAPE, of these models is compared in the periods of relatively low and high liquidity. Furthermore, given that the JEPX has not been widely studied in the EPF literature, other relevant features of the market for forecasting are identified.

Short-Term Electricity Price Forecasting Related to Locations With Different Congestion Characteristics

Michael Stanek (American Electric Power); Jenny Zhao (American Electric Power, jzhao@aep.com); Joseph Sheridan (American Electric Power); John Wayman (American Electric Power)

Forecasting energy prices for day-ahead electricity markets is difficult for a number of reasons. This is primarily because prices can be very volatile and history is often not representative of the future. In addition, different pricing points have characteristics associated with different physical realities. The price for a load zone that covers a geographically large area can be relatively stable while a price received by a generating unit that is located in an electrically congested location is often volatile. Numerous methods have been used in this effort, including physical models of the power

system, statistical techniques, and machine learning models. This paper focuses on two specific insights that have been gained in our efforts to forecast different pricing points using neural network models. The impact of these issues is investigated for different types of pricing points within the PJM market that are impacted by different levels of electrical congestion. The first issue investigated is the objective function used in the neural network training process. Although minimization of squared errors is generally used as the objective function, we find that this leads to a material upward bias. The second issue investigated is optimally incorporating recent history into the network. We find that the method for doing this has a significant impact on forecast accuracy. Finally, we show that these issues are more material for prices at isolated and congested locations compared to prices associated with geographically large and uncongested locations.

Asymmetry in volatility spillovers between energy and agricultural commodity prices

Angi Roesch (FOM University of Applied Sciences, angi@angi-stat.com); Harald Schmidbauer (Shanxi University of Finance and Economics); Sebastian Birkett (FOM University of Applied Sciences)

The 2007/8 food crisis is suspected to have been aggravated by the concurrent surge in oil prices. A recent strand of research is concerned with this issue, often identifying substitution effects between biofuels and fossil fuel, together with a growing focus on renewable energy, as the major determinants of this food crisis. Bans on using certain crops in biofuel production have been called for, and concerns on farmland being diverted for biofuels production and its impact on greenhouse gas emissions are being expressed. The purpose of our study is to investigate the relation between energy prices and a selection of agricultural commodity prices after the 2007/8 food crisis. We use a combination of VARMA and the mGJR-GARCH, which allows for asymmetry in the spirit of the GJR-GARCH, to model the dynamics of co-movements and potential asymmetry in volatility spillovers between returns on energy and agricultural commodity prices. Our findings give rise to interpretations in terms of news impact characteristics lending themselves to welfare implications.

Demand forecasting (II)

Tuesday, 14:00 - 15:20; Chair: Ivan Svetunkov

What Management Must Know About Forecasting

Michael Gilliland (SAS Institute, mvgilliland@gmail.com)

Why are business forecasts so frequently wrong, and what can we do about it? This presentation explores fundamental issues in forecasting that may not be apparent to those managers and executives who oversee or rely on the forecasting process. It will expose common worst practices that politicize the forecasting process, or add costs and complexity that fail to improve forecasting results. This material can be used to help educate management on the limitations of forecasting, and make them aware of alternative approaches to their business problems. Topics include:

- Common forecasting practices that make the forecast worse.
- How demand volatility impacts to the ability to forecast, and how typical business practices add to this volatility.
- Ways to reduce demand variability, making the business easier to forecast and manage.
- Managing accuracy expectations, and how to set reasonable forecast accuracy objectives.
- How to use forecast value added (FVA) analysis to identify and eliminate non-value-adding activities from the forecasting process.

Attendees will come away with a new way to talk about business forecasting with their management – it isn't all about the statistical models – and will have new tools like FVA analysis to address their forecasting process issues.

How to prevent stock-outs from biasing demand forecasts?

Agata Chorowska (ProLogistica Soft, agata.chorowska@prologistica.pl)

One of difficulties in forecasting demand is that very often it is not directly observable. Typically, companies collect sales data. However sales may not reflect the real demand, and stock-outs are one of the most common reasons of this phenomenon. In such cases, with a part of the demand being unfulfilled (and unobserved), basing demand forecasts merely on the sales data can lead to underestimated forecasts and in consequence cause further stock-outs. Such problems can be difficult to spot as forecast accuracy (when comparing forecasts with actual sales) is high. Business model of one of our clients allows this company to collect accurate unfulfilled demand data. This creates a unique opportunity to examine stock-outs in the context of forecasting. In this study we concentrate on how stock-outs information can be harnessed to improve demand forecasts accuracy. First, we discuss how stock-outs can be automatically identified based on stock levels and sales data. Second, we present a comparison of different approaches to incorporate stock-out information into the forecast data. In particular we consider 1) treating stock-out times as outliers and removing them from the forecast data, 2) defining correlated series or calendar events reflecting stock-out times, 3) using expectation maximization techniques to impute the missing demand information. We evaluate those approaches in the light of forecast accuracy as well as computational complexity. Finally, we discuss the impact of incorporating stock-out information into demand forecasts on companies key performance indicators such as sales value, inventory value and service levels.

Retail forecasting – The challenges and opportunities

Jagtej Bewli (WalmartLabs, jbewli@walmartlabs.com)

Efficient retail supply chains are dependent on having accurate demand forecast for items at each store. Improved forecast accuracy enables better supply chain and logistical planning and translates to better in-stocks, reduced excess inventory or waste in case of perishable items.

Retail forecasting problems have been solved over the last several decades with forecasting techniques getting more robust yielding continually better results. The industry has predominantly relied on time series models. While time-series models do a great job in capturing year over year seasonality and trends, they lack the flexibility to incorporated hundreds of external factors such as weather, local events etc that shape customer demand. There is an opportunity to improve forecast accuracy by supplementing time series models with machine learning algorithms.

Additionally, optimal forecasting results aren't just achieved by building better models, it requires the forecasting solution to have the ability to incorporate the human intuition coming from merchants and demand managers who have years of experience running the business.

This talk will cover the above topics and discuss factors to consider when building a forecasting solution for a retail business.

Forecasting intermittent data with complex patterns

Ivan Svetunkov (Lancaster University Management School, i.svetunkov@lancaster.ac.uk); John Boylan (Lancaster University Management School); Patricia Ramos (Institute for Systems and Computer Engineering, Technology and Science); José Manuel Oliveira (Institute for Systems and Computer Engineering, Technology and Science)

Several intermittent demand forecasting methods have been developed over the years (e.g. Croston, SBA, TSB). These methods produce level forecasts for both demand sizes and demand intervals / occurrence, which correspond to a straight line in a holdout sample. However, the reality may be more complicated and more advanced models could help to reveal the other time-series components.

Using an intermittent state-space model and the principles of the logistic models, we introduce a new, more general framework that allows not only capturing the complex dynamics in both demand sizes and demand occurrences (with potential trends, seasonality and special events), but also producing forecasts that incorporate these features. The proposed approach is tested on a large dataset of sales of a retailer evaluating its performance and comparing with the existing forecasting methods.

Combinations & Evaluation

Tuesday, 14:00 - 15:20; Chair: Pablo Pincheira

An alternative approach to combine weights of ensemble models

Rodrigo Heldt (Universidade Federal do Rio Grande do Sul - UFRGS); Carla Freitas Silveira Netto (Universidade Federal do Rio Grande do Sul - UFRGS, carla.netto@gmail.com); Cleo Schmitt Silveira (Universidade Federal do Rio Grande do Sul - UFRGS)

Access to bigger data sets has given researchers the possibility to apply methods from computer science, machine learning techniques. Marketing modelling has given, as far as we know, little attention to these techniques and to forecasting accuracy as well, since the focus is on modelling with the purpose to explain or describe consumer choice more then to predict it. Some efforts were made (e.g. Ali, Sayın, Van Woensel, & Fransoo, 2009; Sun, Choi, Au, & Yu, 2008) applying CART and extreme learning machine to forecasting sales, but marketing practice is far more advanced in its use than marketing research (Chintagunta et al., 2016; Cui & Curry, 2005; Wedel & Kannan, 2016). Wedel & Kannan (2016) state that future studies and models in marketing should use machine learning approaches that are easier to implement, such as ensemble models.

The use of ensemble models to improve forecasts is very popular in machine learning competitions such as Kaggle. They proved to be more accurate (Coussement & De Bock, 2013; De Bock, Coussement, & Van den Poel, 2010; Lemmens & Croux, 2006) and have much stronger generalization ability than single models (Wang, Sun, Ma, Xu, & Gu, 2014). Ensemble models are those that combine the results of more than one base model by giving each of them weights. One common question concerns those weights. Some authors state that giving equal weights or averaging the results is the default and has not been outperformed yet (i.e. Barrow & Crone, 2016; Graefe, Kuchenhoff, Stierle, & Riedl, 2015; Krawczyk, 2015).

We compare two alternative techniques: stacking, that optimizes the models based on the mean absolute error; and quadratic optimization that optimizes both the mean absolute error and the error variance of the different base models used, applied for several subsamples. The quadratic optimization alternative was inspired in the 'vogging' method proposed be Derbeko, El-Yaniv, & Meir (2002) in which the authors apply the technique, adapted from finances (portfolio optimization) to a classification problem (regression with a binary response variable) to find weights to different applications of only one base model (support vector machine) for several subsamples.

The paper focus on the use of ensemble learning and develops an alternative approach to the stacking method and the equal weights method. As far as we know, this still has not been applied in marketing, for forecasting a continuous response variable, such as sales. To accomplish that, the single base models used were decision tree, KNN, random forest, support vector machine and linear regression, combined in an ensemble model in which the base models' weights are defined using quadratic optimization. This approach can contribute to establishing an alternative method to find an optimum combination of weights and improve marketing models' accuracy to predict sales.

Meta-learning how to forecast time series

Thiyanga Talagala (Monash University, Australia, thiyanga.talagala@monash.edu); Rob Hyndman (Monash University, Australia); George Athanasopoulos (Monash University, Australia)

A crucial task in time series forecasting is the identification of the most suitable forecasting method. We present a general framework for forecast model selection using meta-learning. A Random Forest

is used to identify the best forecasting method using only time series features. The proposed framework has been evaluated using time series from the M1 and M3 competitions, and is shown to yield accurate forecasts comparable to several benchmarks and other commonly used automated approaches of time series forecasting. A key advantage of our algorithm is that the time-consuming process of building the random forest can be handled in advance of the forecasting task.

On combining forecasting methods using time series features

Pablo Montero-Manso (University of A Coruña, p.montero.manso@udc.es); Thiyanga Talagala (Monash University, Australia); Rob Hyndman (Monash); George Athanasopoulos (Monash University, Australia)

It is well known that ensemble approaches produce improvements over single methods in statistical learning. Nevertheless, when calculating predictions over a large dataset, computation time for the whole ensemble can be prohibitive, so individual model selection becomes the preferred approach.

We present a method for combining forecasting models by posing it as a gradient tree boosting classification problem using features extracted from the time series. Unlike regular classification problems, we minimize the average forecast error of the selected method rather than the classification error. Not only does this address the aim of accurate forecasting, it also provides measures of relative method accuracy across the time series, and relative difficulty across time series. In contrast, a classic classification approach would give the same importance to all series and methods.

The presented classifier is compared with state-of-the-art approaches to forecasting and time series classification. The results show an improvement of error over alternative approaches.

These experiments allow us to show the relevance of both the feature set and the proposed optimization approach to several collections of time series.

The scalability of the approach allows it to be applied to forecasting a large collection of time series. It can also be efficiently trained to tailor specific domains or datasets.

Small sample adjustment for out-of-sample tests of predictability

Pablo Pincheira (Universidad Adolfo Ibáñez, pablo.pincheira@uai.cl)

In this paper we introduce a small sample adjustment for out-of-sample tests of predictability. The relevant econometric environment is one in which the econometrician wants to compare the population Mean Squared Prediction Errors (MSPE) of two models: one big nesting model, and another smaller nested model. Although our adjustment can be used to improve the properties of many out-of-sample tests of predictability, in this paper we focus on the ENC-T test of Clark and McCracken (2001) and the widely used test developed by Clark and West (2006, 2007). Our new test multiplies the Clark and West t-statistic by a factor that should be close to one under the null hypothesis that the short nested model is the true model, but that should be greater than one under the alternative hypothesis that the big nesting model is more adequate. We use Monte Carlo simulations to explore the size and power properties of our approach. Our simulations reveal that the new test is well sized and powerful. In particular, it tends to be less undersized and more powerful than the tests by Clark and West (2006, 2007). Although most of the gains in power are associated to size improvements, we also obtain gains in size-adjusted power. Finally we present an empirical application in which many more rejections of the null hypothesis are obtained with our new test.

Finance (III)

Wednesday, 9:40 - 11:00; Chair: Jonathan Reeves

Forecasting financial time series with stable GAS models

Daniel Takata Gomes (Brazilian Geography and Statistics Institute, daniel.gomes@ibge.gov.br); Chang Chiann (University of São Paulo); Clélia Maria de Castro Toloi (University of São Paulo)

GARCH models with normal and Student's t conditional distributions are widely used for volatility modeling in financial data. However, such distributions may not be suitable for some heavy-tailed and leptokurtic series. The stable distributions may be more adequate to fit such characteristics, as already exploited in the literature. On the other hand, the recently developed GAS (Generalized Autoregressive Score) models are dynamic models in which the updating mechanism of the time-varying parameters is based on the score function (first derivative of the log-likelihood function). This provides the natural direction for updating the parameters, based on the complete density. We propose a new GAS model with symmetric stable distribution for volatility modeling. The model can be interpreted as a generalization of the GARCH models, since the classic Gaussian GARCH model is derived from it by using particular choices of the stable distribution and the model structure. There are no closed analytical expressions for general stable densities in most cases, hence its numeric computation and derivatives are detailed for the sake of complete development of the estimation process. Simulation studies, as well as an application to real data, are presented for comparisons between the usual models and the proposed model, illustrating the effectiveness of the latter.

The Dog Has Barked for a Long Time: Dividend Growth is Predictable

Jack Strauss (University of Denver, jack.strauss@du.edu)

Motivated by the Campbell-Shiller present-value identity, we propose a new method of forecasting dividend growth that combines out-of-sample forecasts from 14 bivariate predictive regressions based on common return predictors and the logged dividend payout. Combination forecast methods generate robust out-of-sample predictability of annual dividend growth over the entire post-war period as well as most sub-periods with out-of-sample R2 up to 21%. The dividend-growth forecasts significantly forecast economic activity and coupled with the dividend-price ratio also forecast excess returns. The dividend-growth forecasts combined with forecast combination methods produce out-of-sample R2 that exceed 10% at the one year-horizon and 18% at the two-year horizon. In spite of robust dividend predictability, we find that most variation in the dividend-price ratio is still attributable to variation in expected returns.

Mortgage rate forecasting

Robertas Gabrys (University of Southern California, gabrys@marshall.usc.edu)

Online modern financial advising platforms have been enjoying an increased popularity over the last decade. Their main goal is to utilize analytics capabilities to help their customers make financial decisions at every stage of life. This talk will present a consulting project on the development of a mortgage rate forecasting model conducted by USC MSBA students for an online financial advising startup. The startup's business model is rather simple: the company makes revenue by being a lead generator, i.e. by directing potential customers from its site to the site of a bank who will then give out the loan. By accurately predicting what the mortgage rate will be in the future and relaying that information to customers, the company will increase the total leads generated, therefore increasing

their revenue. The goal of the project is to build a model to predict 30-year fixed mortgage rate, 15-year fixed mortgage rate, and 5/1 arm mortgage rate, 4 periods into the future. The project accomplishes this by utilizing classical techniques (regression, smoothing, Box-Jenkins) and modern machine learning techniques models (xgboost, RNNs, LSTM) .

Targeting Market Neutrality and Volatility

Bao Doan (University of New South Wales); Jonathan Reeves (University of New South Wales, reeves@unsw.edu.au)

Advances in volatility and beta forecasting are extended to the setting of volatility timing of market neutral portfolios. Key features of the study include short horizon forecasting from models with higher accuracy levels than previously documented in the literature. A trade-off in the joint targeting of portfolio beta and volatility is also identified and appropriate balancing of these two dimensions of risk is demonstrated.

Energy forecasting

Wednesday, 9:40 - 11:00; Chair: Fernando Cyrino

Modeling and Forecast of Brazilian Reservoir Inflows under ENSO and Precipitation Influence

Paula Medina Maçaira (Pontifical Catholic University of Rio de Janeiro (PUC-Rio), paulamacaira@aluno.puc-rio.br); Fernando Cyrino (Pontifical Catholic University of Rio de Janeiro (PUC-Rio))

The Brazilian energy matrix is mainly composed by hydroelectric plants and one of the main characteristics of this kind of generation systems is the strong dependence on hydrological regimes. Currently, the Brazilian electric sector uses the stochastic variable know as Natural Energy Inflow to generate hydrological scenarios from a Periodic Autoregressive model. Such model does not consider any exogenous information that could affect the hydrological regimes and, consequently, the energy production. Recent studies indicate that the use of climate variables may serve as a factor to reduce uncertainties. Moreover, decomposing the hydrological series into signal and noise before modelling may also improve the fitting stage. In this context, the objective of this work is the development of hybrid models that combine techniques of decomposition of hydrological series and time series models. The new approach contemplates the use of Singular Spectrum Analysis (SSA) and Multichannel-SSA decomposition techniques combined with Periodic Autoregressive (PAR) and PAR with one exogenous variable models. As conclusion, the applied models have proved to be efficient for the proposed objectives, but there is still room for improvement.

A Load-Based Temperature Prediction Model for Anomaly Detection

Masoud Sobhani (University of North Carolina at Charlotte, masood.sobhani@gmail.com); Tao Hong (University of North Carolina at Charlotte)

Electric load forecasting, as a basic requirement for the decision-making in power utilities, has been improved in various aspects in the past decades. Many factors may affect the accuracy of the load forecasts, such as data quality, goodness of the underlying model and load composition. Due to the strong correlation between the input variables (e.g., weather and calendar variables) and the load, the quality of input data plays a vital role in forecasting practices. Even if the forecasting model were able to capture most of the salient features of the load, a low quality input data may result in inaccurate forecasts. Most of the data cleansing efforts in the load forecasting literature have been devoted to the load data. Few studies focused on weather data cleansing for load forecasting. This presentation proposes an anomaly detection method for the temperature data. The method consists of two components: a load-based temperature prediction model and a detection technique. The effectiveness of the proposed method is demonstrated through two case studies: one based on the data from the Global Energy Forecasting Competition 2014, and the other based on the data published by ISO New England. The results show that by removing the detected observations from the original input data, the final load forecast accuracy is enhanced.

Probabilistic Short-term Wind Forecasting Based on Pinball Loss Optimization

Jie Zhang (University of Texas at Dallas, jiezhang@utdallas.edu); Mucun Sun (University of Texas at Dallas); Cong Feng (University of Texas at Dallas)

Probabilistic wind power forecasts that quantify the uncertainty in wind output have the potential to aid in the economic grid integration of wind power at large penetration levels. This talk will introduce

a novel probabilistic wind forecasting approach based on pinball loss optimization, in conjunction with a multi-model machine learning based ensemble deterministic forecasting framework. This is a two-step probabilistic forecasting method, consisting of deterministic forecasts generation and predictive distribution (type and parameters) determination. A machine learning based multi-model forecasting framework (MMFF) is first developed to generate short-term deterministic wind forecasts. By assuming the deterministic point-forecasted value as the mean at each point, one unknown parameter (i.e., standard deviation) of a predictive distribution at each forecasting point is determined by minimizing the pinball loss. A surrogate model is developed to represent the unknown distribution parameter as a function of deterministic forecasts. This surrogate model can be used together with deterministic forecasts to predict the unknown distribution parameter and thereby generate probabilistic forecasts. Numerical results of case studies show that the proposed method has improved the pinball loss by up to 35% compared to a baseline quantile regression forecasting model.

An Approach to Spatial Wind Speed Prediction with Ensemble of Decision Trees

Bruno Bastos (Pontifical Catholic University of Rio de Janeiro (PUC-Rio), brunoq.bastos@gmail.com); Fernando Cyrino (Pontifical Catholic University of Rio de Janeiro (PUC-Rio))

The increasing worldwide penetration of wind energy introduces a great challenge in reliable operation and planning of modern power systems. Accurate wind speed prediction is of great importance in this context, because it helps decrease the uncertainty related to the supply side, and hence provides, for example, a means for reducing costs of contracting operational reserves.

Many studies have addressed the issue of wind speed and wind power forecasting. In this work, we investigate the use of ensemble of decision trees (DTs) to solve the task of one-hour-ahead wind speed prediction for a given region in Brazil. DT is one of the most popular algorithms adopted for inductive inference; it provides the advantages of being relative fast to construct and of performing feature selection internally. Ensemble techniques, on the other hand, may be applied to improve accuracy performance for a task by combining predictions of single models. In this work, we investigated both AdaBoost technique and Random Forests to create ensemble of DTs.

To solve our task, we constructed models which relate meteorological data (e.g., temperature, pressure at sea level, relative humidity) – in present and past time – and geographical data (e.g., latitude, longitude) to one-hour-ahead wind speed. This way, the models are able of inferring wind speed for any given location (within the region of study), if meteorological information is available.

The database used in our work is from the Climate Forecast System Reanalysis (CFSR) dataset, produced and updated by the National Centers for Environmental Prediction (NCEP). The CFSR dataset contains hourly time series data for a different set of variables, such as air temperature, humidity, precipitation, wind speed, etc., which are produced by the Climate Forecast System (CFS).

The results show that the Random Forest algorithm with 10 trees of maximum depth of 20 provided the best accuracies for the task, slightly better than the Adaboost of DTs. Overall, the ensemble techniques greatly outperformed the baseline model (Linear Regression), but were only slightly better than the single DT. The proposed approach is very promising and needs more attention. Future works should include the use of attributes which might capture more regional aspects which may be more general, so that, if a given location does not have any meteorological information, one could still infer wind speed at that location.

Bayesian models

Wednesday, 9:40 - 11:00; Chair: David Suda

Bayesian Shrinkage Estimation Of Logistic Smooth Transition Autoregressions

Mario Giacomazzo (Arizona State University, mgiacoma@mail.bw.edu); Yiannis Kamarianakis (Arizona State University)

The logistic smooth transition autoregressive (LSTAR) model captures nonlinear, regime-dependent associations; LSTAR is formulated as a weighted combination of linear autoregressive (AR) processes. In this work, sparse estimation of LSTAR is attained using Bayesian shrinkage (Laplace and Horseshoe) priors on the autoregressive coefficients of each regime. Furthermore, Dirichlet priors are employed to identify composite threshold variables in the transition function. The proposed specification provides a flexible alternative to computationally intensive stepwise model building procedures and reversible jump Markov chain Monte Carlo (RJMCMC) schemes. Simulation experiments demonstrate the efficacy of the methodology. Application to sunspot data illustrates the ability to achieve superior forecasting performance. Finally, the capability to handle exogenous inputs is exemplified through daily maximum water temperature forecasting: in this application Bayesian linear and nonlinear river-specific models are evaluated with regard to their 3-day and 7-day ahead forecasting performance.

Alternative benchmark models for a large-scale BVAR: application to Russian macrodata

Oxana Malakhovskaya (National Research University Higher School of Economics, University of Paris-Saclay, omalakhovskaya@hse.ru); Boris Demeshev (National Research University Higher School of Economics)

Accurate macroeconomic forecasts are extremely important for policymaking. Central banks and government bodies monitor many available series to increase the accuracy of forecasts. It is also a wide-spread belief that large-scale models generally outperform univariate or low-dimensional models as they do not lose any potentially relevant information.

Recently, many papers have shown a good forecasting performance of large-scale Bayesian vector autoregressions (BVARs) for forecasting. A standard procedure is to compare an accuracy measure (for example, RMSE) for a Bayesian model with a random walk model that serves as a simple benchmark.

This paper enlarges the standard procedure by comparing the accuracy of out-of-sample forecasts obtained from BVARs of different sizes with forecasts obtained from small-scale frequentist VAR as well as automatic ARIMA and ETS models.

First of all, we show that many Russian macroeconomic indicators can be forecast by BVARs more accurately than by frequentist VARs and random walk models. However, contrary to several other studies, we do not confirm that the relative forecast error monotonically decreases with increasing the cross-sectional dimension of the sample. Two clear exceptions here are the agriculture production index and PPI. The most accurate forecasts for these variables are obtained with a small-scale unrestricted VAR model.

Then we compare out-of-sample forecasts made with BVAR models with those obtained from simple benchmarks: automatic ARIMA model and ETS. Though in most cases BVARs give the most accurate forecasts, there are some exceptions. For example, in CPI forecasting the univariate models allow the researcher to diminish the RMSE more than twice in comparison with the forecasts obtained from the BVAR model.

For some variables and some forecasting horizons, neither of considered models outperform the random walk. For example, in all specifications we consider, the nominal exchange rate cannot be forecast by univariate or multivariate models better than by the random walk, and it is a long-held consensus in economics. Our results confirm that for each macroeconomic variable a separate forecasting model should be fit and used for forecasting. In some cases, an automatically-chosen univariate model might outperform a thoughtfully elaborated large-scale BVAR.

A hybrid simulation - empirical Bayes forecasting algorithm for geographical demand estimation

John Bowman (Walmart Labs, jbowman@walmartlabs.com); Jagtej Singh (Walmart Labs)

Walmart eCommerce has several warehouses in the U.S., each stocking hundreds of thousands of distinct items. Item demand forecasting is done at a U.S. level, but the forecasts then have to be allocated to the warehouses for use by inventory control and capacity planning systems. This task is complicated by three factors. First, many items are very low demand, so a specific warehouse / geographical region may see no demands for an item for several months. Second, items are often not supplied from the geographically nearest warehouse, as, if multiple items are on the same purchase order (which is how most items are actually purchased), the shipment will often come from the nearest warehouse that has all of the items in stock to save outbound shipping costs - which may not be the geographically closest site. Third, many orders are for more than a single unit, which makes it difficult to evaluate how much information is contained in a single week's sales. We have developed a geo-demand forecasting system that is a hybrid of simulation and empirical Bayes approaches that addresses all three of these problems, and present a comparison with the previous Bayesian framework's accuracy.

Threshold regime-switching models for football predictions: a frequentist and empirical Bayes approach

David Suda (University of Malta, david.suda@um.edu.mt); Lawrence Grech (University of Malta)

We contribute to the discussion of whether football betting markets are efficient by devising appropriate adaptive models which respond to changes over time. For this purpose, we look at two football teams of different ability in the German Bundesliga: FC Bayern Munich and FSV Mainz 05 . We use three seasons starting from August 2012 as our training set, and two seasons starting from August 2015 as our test set, and focus on the Asian Handicap market. We implement a time series regression approach on goal differences which allows for the inclusion of past lags and exogenous input, with the possibility of a threshold regime-switching mechanism based on an exogenous or endogenous input (or a hybrid of both). As exogenous input, we include data about the team, the opponent, home advantage and data from betting markets available prior to each game, to aid us in more reliable forecasts. We also consider two parameter estimation approaches. The first is the frequentist approach via ordinary least squares, which can be implemented within a time series regression context subject to certain assumptions on the data, and provided the normal distribution remains a reasonable approximation. The second is an empirical Bayes approach which we specifically devise for this setting. The Bayesian aspect of regression is well-studied and the empirical Bayes setting could make sense in this context because it allows the model to adapt to recent changes. In our method we propose to optimize hyperparameters for the prior using information in a recent time window. We finally compare outcomes of different models, to determine whether the inclusion of regime-switching dynamics and the implementation of the empirical Bayes approach improves predictions. We also test hypotheses related to predictability and profitability for the different modelling approaches and teams.

Supply chain forecasting (II)

Wednesday, 9:40 - 11:00; Chair: Mohammad Ali

Forecasting Spare Parts Demand with Clustering Approaches

Christian Menden (Fraunhofer-Gesellschaft, Germany, christian.menden@scs.fraunhofer.de); Julius Mehringer (Fraunhofer-Gesellschaft, Germany)

Predicting spare parts demand for the time period after production ceased is a central issue in supply chain management in order to reduce costs in terms of storage, transportation, disposal, finance and to maintain a desired level of customer satisfaction. This paper proposes a forecasting approach that uses clustering and nonlinear regression techniques to find patterns in a large dataset of master and consumption data from a big manufacturer of household goods from 1985-2017. More specifically, we apply different clustering methodologies with varying distance measures and parameter calibrations to determine products that share similarities in terms of (i) master data and (ii) consumption patterns. We use these clusters to construct predictions for “new” products where historical data is scarce using the data from similar (older) products where more data is already available by means of Bayesian clustering and regression. Our results indicate that this step-wise approach of combining clustering and Bayesian prediction methods yields significantly better forecasting results than a baseline model and improves the spare parts planning and controlling process.

Information Sharing in the Presence of Promotions in a Supply Chain

Patrick Saoud (Lancaster University, patricksaoud@gmail.com); Nikolaos Kourentzes (Centre for Marketing Analytics and Forecasting, Lancaster University); John Boylan (Lancaster University Management School)

Many supply chains experience the Bullwhip effect, defined as the upstream amplification of demand variability. This information distortion results in a misalignment of forecasts, generating expensive business costs. A proposed remedy in the literature is the sharing of Point of Sales information data among the members of the supply chain. The theoretical and empirical results have pointed in different directions, with the empirical evidence suggesting that information sharing helps achieve better forecasting accuracy. A less studied facet of the Bullwhip is the effect of promotions on it, which was highlighted as one of its four original sources. This research is dedicated to examining the effect of promotions and other demand shocks on the performance of the different tiers of the supply chain. In particular, it will study the impact of promotions on forecasting accuracy, Bullwhip propagation and safety stocks for the participants of the supply chain. Furthermore, it will also investigate the impact of different types of information sharing in this context on the Supply Chain, and compare their performance in terms of gains in forecasting accuracy.

Effectiveness of Non-overlapping Temporal Aggregation on Forecast Accuracy

Bahman Rostami-Tabar (Cardiff University, rostami-tabarb@cardiff.ac.uk); Zied Babai (Kedge Business School); Mohammad Ali (Coventry University); John Boylan (Lancaster University)

Temporal aggregation is an intuitively appealing and effective approach to deal with demand uncertainty for fast moving and intermittent moving items. The non-overlapping temporal aggregation has been shown to be an effective approach that may improve forecast accuracy under auto-correlated stationary demand processes. However, its impact on nonstationary auto-correlated

processes has not been investigated. The aim of this study is to bridge this gap by investigating the effectiveness of the non-overlapping temporal aggregation on forecast accuracy, supply chain demand and orders. To do so, we consider a two-stage supply chain (e.g. a retailer and a manufacturer) where the retailer faces an integrated moving average demand process of order 1, ARIMA(0,1,1), that is forecasted by using Single Exponential Smoothing which is an optimal forecasting method for the demand process. We derive the analytical expressions of the mean squared forecast error at the retailer and the manufacturer levels as well as the bullwhip ratio when the aggregation approach is used. We discuss the merit of using the aggregation approach and we highlight its implications for business practice.

Commodity and consumption forecasting

Wednesday, 9:40 - 11:00; Chair: Marinda Pretorius

How important are common factors in forecasting non-energy price changes?

Lya Sierra (Pontificia Universidad Javeriana of Cali, lyap@javerianacali.edu.co); Pilar Poncela (European Commission, Joint Research Centre); Eva Senra (Universidad de Alcalá)

Commodity prices influence price levels of a broad range of goods and, in the case of some developing economies, production and export activity. Therefore, information about future commodity inflation is useful for central banks, forward-looking policy-makers, and economic agents whose decisions depend on their expectations of future commodity inflation. After 2004, we have witnessed the so-called financialization of the commodity markets, which might induce greater communalities among commodity prices. The predictive content of the co-movement either of a large range of commodities, or the co-movement within a specific category of raw material prices is evaluated. This paper reports success in using small-scale factor models in forecasting the nominal price of non-energy commodity changes in most of commodities into the categories: Beverages, Agricultural Raw Materials and Vegetable Oils and Protein Meals.

Income and wealth effects: a thick modelling approach for euro area private consumption

Arne Gieseck (European Central Bank, arne.gieseck@ecb.europa.eu); Gabe Jacob De Bondt (European Central Bank); Zivile Zekaite (European Central Bank)

This study develops a tool for forecasting real private consumption in the euro area based on a thick modelling approach in the spirit of Granger and Jeon (2004). To the best of our knowledge, this approach has not yet been adopted for private consumption. In addition, we make another important contribution by disaggregating not only total wealth in a consumption equation but also real disposable income.

A large number of error correction models (ECM) are estimated by extending a standard consumption equation along several dimensions. Firstly, we examine the effects of different types of income on consumption in the short and long run by decomposing disposable income into labour and non-labour components in three alternative ways. Secondly, total wealth is split into financial net wealth and non-financial assets. Third, we add a large number of other variables to explain private consumption in the short run. These variables are drawn from five groups in various combinations: (i) interest rates and spreads; (ii) measures of consumer indebtedness; (iii) measures of government indebtedness; (iv) income uncertainty; (v) other variables.

In the next step of the exercise, in-sample and out-of-sample criteria are used to identify well-specified ECM equations and eliminate the rest. The pre-selected models are estimated using data until 2015Q3 and evaluated in terms of average out-of-sample Root Mean Squared Error (RMSE) on the basis of 1 to 8 quarters ahead forecasts over the period 2015Q4 - 2017Q3. The equations are sorted according to the RMSE from the smallest to the largest and compared to the benchmark model, i.e. ECM with only total disposable income and two wealth components as determinants in both the short and long run. In the final selection step, we check whether the signs of the estimated coefficients are economically correct starting from a model at the top of the list and continuing down the list as long as a relative RMSE against the benchmark is less than 0.85.

Full-sample estimation results of the selected ECMs allow making several interesting observations. First, the long run estimates indicate that labour income has a significantly stronger

effect on consumption as compared to a non-labour component, albeit the latter is also an important factor. The magnitude of the coefficient on non-labour income is about half as large as that on labour income. The long-run coefficients also show that financial net wealth has a much more pronounced impact on consumption and contributes significantly more to consumption growth as compared to non-financial assets. With respect to the short run estimates, the evidence of differential income effects is weaker and mixed. Finally, we show that independently of income decomposition net financial wealth and non-financial assets are typically equally relevant determinants of private consumption in the short run. The final set of ECM equations is used to provide a range of forecasts for real private consumption. Based on the pseudo out-of-sample 8 quarters ahead forecasts, we show that the mean forecast of our selected models is reasonably close to the actual consumption series.

Forecasting commodity currencies considering commodity prices as predictor

Ilse Botha (University of Johannesburg, ilseb@uj.ac.za); Marinda Pretorius (University of Johannesburg)

Researchers have long been concerned about the determinants of exchange rate movements. Recent findings argue that the failure to uncover predictive power in fundamentals such as commodity prices is the low frequency data used in the samples. The purpose of this paper is to consider higher frequency data to determine whether commodity prices can forecast exchange rates in commodity driven countries. We focus on commodity prices and currency pairs of several major commodity exporting countries and compare the commodity-based forecasts to the random walk model.

Volatility

Wednesday, 11:20 - 12:40; Chair: Abril Imelda Rosen Esquivel

Volatility Forecasting of Real Estate Firms in Malaysia

Leong Mow Gooi (University Putra Malaysia); Wei-Chong Choo (University Putra Malaysia, wcchoo@upm.edu.my)

Many previous studies have focused on forecasting the real estate returns but not on its volatility. A reliable volatility forecast in the real estate stock market may provide important information in decision making for the policymaker, central banker, developer, lender and investor. Hence, this paper investigates the performance of various methods in forecasting the Malaysian real estate stock volatility. Thirty-Three Malaysian real estate companies individual stock price daily return will be employed. Forecasting methods used are ad-hoc methods, exponential weighted moving average, generalized autoregressive conditional heteroscedasticity (GARCH) models, and newly proposed Smooth Transition Exponential Smoothing (STES) method. Using both mean absolute error and root means squared error as the evaluation criteria, STES method outperforms other methods in providing volatility forecasting accuracy.

Structural breaks and GAS models of exchange rate volatility

Nyamekye Asare (University of Ottawa, math_fan_85@hotmail.com)

Empirical evidence suggests that structural breaks are present in data on macro-financial variables such as exchange rates. The potentially serious consequences of ignoring a break in GARCH parameters motivated Rapach and Strauss (2008) to study the empirical relevance of structural breaks in the context of GARCH models. However, the literature does not address the empirical relevance of structural breaks in the context of Creal et. al's (2008) generalized autoregressive score (GAS) models. This paper contributes to this literature by extending Rapach and Strauss's (2008) study to include the Student's t-GAS model and by comparing its performance to that of the GARCH model and two non-GAS models, the t-GARCH and stochastic volatility (SV) models. The empirical relevance of structural breaks in the models of volatility are assessed using the formal test of Dufour and Torres (1998) to determine whether the estimated parameters change across sub-periods. The performance of all the models is analyzed using the weekly USD trade-weighted index between January 1973 and October 2016. Through this analysis, this paper also addresses whether accommodating structural breaks in the unconditional variance of both GAS and non-GAS models will improve forecasts. The results show that structural breaks are empirically relevant in GAS models of USD volatility in terms of modeling. They also indicate that using models that accommodate breaks can improve forecasts only in the short run.

Forecasting the Realized Volatility of Global Commodity returns

Abril Imelda Rosen Esquivel (University of Essex, airose@essex.ac.uk)

The paper focuses on the ability to forecast the daily Realized Volatility of the Bloomberg Commodity Index Excess return (BCOM) which has not previously been examined. It uses an Heterogeneous Autoregressive model (HAR) and competing models that include an Implied Volatility (IV) measure from either the Commodity or US Stock Market; for the former the IV for at the money call options of the Dow Jones-UBS Commodity Index published by DataStream, and for the latter the US. Stock Market (VIX). Realized Volatility is measured by three different proxies: absolute returns, and two

range-based estimators, Parkinson's (1980) and Rogers and Satchell's (1991), constructed with open, close, high and low daily prices. In-sample results for the period from 28/07/2011 to 31/10/17 evidence that the IV measure estimates are small but statistically significant, suggesting the IV is a biased estimator of future Realized Volatility. The models used to obtain the one-day-ahead out of sample forecasts from 17/07/15 to 31/10/17, were estimated dynamically using a rolling window. To compare the forecasting accuracy of the models their respective Root Mean Squared Error (RMSE) were computed, which show that in most cases the HAR specification outperforms the alternatives including the IV measures in forecasting Realized Volatility using a RMSFE metric.

Climate and environment

Wednesday, 11:20 - 12:40; Chair: P Geoffrey Allen

Climate Change Forecasting in the Service of Policy and Practitioners - Are We on the Right Track?

Kevin Murphy (Arizona State University, kwmurphy@asu.edu)

Although the physics underlying earth's radiative thermal balance were discovered over a century ago, it was not until 30 years ago that indications of a global warming trend had emerged above natural variability following a period of heightened concerns from major weather variations that had included a period of global cooling in the 1970s. The topic came to prominence in June 1988 with testimony to the United States congress utilizing a well-documented global temperature forecast with a time horizon up to the present day. That event also introduced the political shroud which envelopes climate change to this day, thickening resistance to examination of forecast accuracy and suitability of methods used to inform climate policy.

As scientists geared up three decades ago for intensified study and the attendant research funding, the climatology community was faced with two methodological paths by which to pursue its investigations. The first was conventional empirical statistical methods, and the second was employment of complex climate models built off weather models that were coming into wide use at the time. A struggle ensued for which would gain prominence and hence the majority of research funding. Climate modeling prevailed and has been the basis for projections to the end of this century that informed the Intergovernmental Panel on Climate Change (IPCC) for formulation of climate policy. As use of climate models became ubiquitous they were also promoted for climate change impact assessments at regional levels.

As we reach the 30 year mark since the onset of controversy over this topic, a data record is now available from which to assess forecast accuracy in comparison to alternative methods. The greenhouse gas instrumental record now exceeds 60 years, providing a basis upon which to assess the greenhouse gas emission projections driving current forecasts. Many regional climate records now exceed 100 years by which to assess impact forecasts originating in the climate modeling approach. This presentation will review climate forecast accuracies at the global and a regional level to gauge adequacy of the climate modeling approach for informing policy and conducting impact assessments. The results are not encouraging.

Understanding the effects of dry bulb temperature and wind speed on electric load in Barbados

Adanna Robertson-Quimby (University of the West Indies/Caribbean Institute for Meteorology and Hydrology, adannarobertson@gmail.com); Andrea Sealy (Caribbean Institute for Meteorology and Hydrology/University of the West Indies)

Studies conducted worldwide assessing the impacts of weather and climate on energy systems are country, scale, climate and resource specific. Similar studies are quite limited in the Caribbean. While load forecasting is being undertaken in the Caribbean, there is a need to better understand the impacts due to climate change and variability on load forecasts, in order for the energy sector to become more resilient and adaptive to weather and climate events there.

This study focuses on providing a better understanding of the existing and future relationships between climatic variables and energy demand for applications such as electric load forecasting. Analyses of monthly, quarterly, seasonal and yearly data were performed applying regression techniques. Climatological statistics such as the hottest and coolest months on average were also

used as a basis for analysing the possible relationship of the climate variables with the electric load. These highlighted the significance of the dry bulb temperature and wind speed with the electric load. A stepwise regression process was also conducted in order to determine the possible relationship amongst a combination of these weather variables and the electric load. Examination of the relationships highlighted that the combined effects of wind speed and dry bulb temperature were statistically significant and the coefficient of determination increased from that which was obtained with the respective independent variables. Yearly and seasonal data exhibited the strongest correlations.

In anticipation of climate change and variability and the expected rise in surface temperatures and increase in wind speed, it would be worthwhile to conduct a sensitivity analysis to better understand and predict the associated impacts on energy demand. The development of a framework for forecasting the future electric load due to changes in Caribbean climate is therefore underway.

Regularized Cascade Correlation Networks for Particulate Matters and O3 In Medellín, Colombia.

Fernan Villa Garzón (National University of Colombia, favilla0@unal.edu.co); Alvaro Alejandro Villa Garzón (National University of Colombia); Juan David Velásquez Henao (National University of Colombia)

Modeling and forecasting the average daily concentrations for Particulate Matters (PM_{10}, $PM_{2.5}$) and O_3 has been considered an arduous task not suited for linear models. The concentrations are influenced by complex factors like human activity patterns, industrial pollution, and weather conditions. The air contaminants forecasting is a relevant topic because helps to governments to define air quality regulations, politics, and monitoring. In this work, we proposed the Regularized Cascade Correlation Network (RCCN) architecture has an imputation method and we use it to forecast the daily average concentration for PM_{10}, $PM_{2.5}$, and O_3 air contaminants, the available data are from Medellín (Colombia) between 2008 and 2016. The Cascade Correlation architecture was proposed in 1991 by Fahlman and Lebiere, they build dynamically a multilayer structure and present theoretical advantages over the multilayer perceptrons (MLP). However, like other architectures, the cascade correlation can suffer overfitting. To control it, in 2013 Villa and Velasquez was incorporated three regularization strategies in this architecture: weight decay, weight elimination, and ridge regression. We compared the air contaminants forecasting with ARIMA and MLP models. The forecasting performance of the RCCN was better than the other models. In conclusion, the RCCN can capture the intrinsic dynamics of the air contaminants time series better than other traditional models, and it is able to give a more accurate forecast for a 30 days horizon. In addition, RCCN are an accurate method to impute missing data values in air contaminants data.

Decadal Forecasting Performance of Some Published Simple Models of Global Temperature

P Geoffrey Allen (University of Massachusetts Amherst, allen@resecon.umass.edu)

The present state of climate science might be described as excessive reliance on complex earth system models. These have been used to make century-scale projections based on standard scenarios and to perform experiments on the abrupt vast increase in carbon dioxide. Their outputs have been emulated by simpler box models. A few authors have shown that even simpler, often single equation, models perform about as well suggesting that aggregate global surface temperature dynamics can be captured by a few dominant modes. Earth system modelers have also produced decadal forecasts, or hindcasts, starting in 1960. These have the twin purposes of improving model structure, particularly in dealing with initialization, and demonstrating the forecast skill of the forecaster. Comparison of these

with the simpler models has received less attention. The present paper takes some simple published models, re-estimates them on data up to the forecast origin, makes unconditional decadal forecasts and compares the results with selected complex model conditional forecasts. I also discuss why simple models perform comparatively well and the implications for climate forecasting.

Big data & data mining

Wednesday, 11:20 - 12:40; Chair: Ozden Gur Ali

Finding a scapegoat of the exchange rate anomaly with machine learning methods

Levent Bulut (valdosta state university, lbulut@valdosta.edu); Zeliha Gormez (Istinye University)

At times when exchange rates are driven by unobservable shocks and macro fundamentals fail to explain the disconnect, the scapegoat theory of exchange rates justifies this with market participants' changing expectations about the structural parameters as they might put excessive weights to a specific exchange rate fundamental, as a scapegoat of the anomaly, that was off-the-track at the time. In empirical testing of the scapegoat theory of exchange rates, identifying (and the timing of being) a scapegoat factor necessitates modelling the exchange rate fundamentals and measuring the unobservable shocks. In this paper, we propose to use machine learning techniques to pinpoint the macro fundamentals that can serve as scapegoat factor and match the market expectations of these macro fundamentals by using Consensus Economics survey data on these variables for 25 currency pairs.

Selecting an Optimal Modeling Hierarchy Using Machine Learning Methods

Yue Li (SAS Institute Inc, yue.li@sas.com)

Hierarchical Forecasting has been proven to be an effective method to improve forecast accuracy for large time series data with hierarchical structure, due to its capability of pooling information at different aggregation level to reduce noise and enhance signal. This method usually requires a good hierarchy; however, in reality the hierarchy is usually manually provided by the users based on planning structure, which might not be the best one for modeling purpose. This paper presents a data driven framework to derive the best hierarchy that not only organize the data in an easy-to-interpret way, but also improve the forecast accuracy from the modeling perspective.

Scalable Cloud-Based Time Series Analysis and Forecasting

Thiago Quirino (SAS Institute Inc, thiago.quirino@sas.com); Michael Leonard (SAS Institute Inc); Ed Blair (SAS Institute Inc (Retired))

Many organizations need to process large numbers of time series for analysis, decomposition, forecasting, monitoring, and data mining. This paper presents a resilient, distributed, optimized generic time series analysis scripting environment for cloud computing. The framework comes equipped with capabilities such as automatic forecast model generation, automatic variable and event selection, and automatic model selection. It also 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 and provides examples that use this technique.

Automatic Forecasting of Category-Store Sales Considering Cross-Category Effects

Ragıp Gürlek (Koc University); Ozden Gur Ali (Koc University, oali@ku.edu.tr)

Sales forecast is a crucial part of retail operations. At the category-store level short to medium term multi-period ahead forecasts are needed. A typical retail chain needs to forecast thousands of related

time-series, making it necessary to use a robust automated method. In addition to seasonality and random disturbances, category-store level sales are affected by marketing instruments in the focal category and potentially other categories. Further, considering lags and distributional characteristics of marketing instruments creates a large set of features that can be used for better forecasts, while adding to the complexity of the forecasting task.

Sales, marketing efforts, and random events in one category may affect the sales of another. These cross-category interactions may be due to substitution, income, or complement effects, or an indirect effect through store traffic.

We propose an automated method that forecasts multi-period-ahead category-store-level retail sales considering cross-category effects, and features that capture the SKU level distribution of marketing instruments. We apply and test the accuracy of the method on a retail dataset supplied by IRI for 3 chains with 154 stores and 31 categories with multiple horizons within the ten years of data. Further, we investigate under which circumstances the cross-category effects and distributional features contribute the more accurate forecasts.

The method consists of three stages: In the first stage we deseasonalize sales and marketing variables. In the second stage, we regress the deseasonalized sales on deseasonalized marketing variables of all categories in the store by using the Elastic Net, which combines Lasso and ridge regression methods, selecting variables or shrinking variable coefficients as appropriate. The parameters are selected through a cross-validation method we tailored for time-series. The third stage extrapolates the residuals of the second, with lead time specific models that incorporate the effect of random disturbances in both the focal category and cross-categories.

Our method significantly outperforms the benchmarks STL (Seasonal and Trend decomposition using Loess) and ARIMA (Autoregressive Integrated Moving Average) in forecast accuracy. We find that disaggregate promotional information improves forecast accuracy for lead times of 1-4 and 5-9 weeks, however, the extra information hurts the forecast accuracy for longer lead times. We do not find evidence that using lags of promotional variables improve predictive accuracy. The findings show that own-marketing, cross-category marketing, and own-disturbance components individually improve the forecast accuracy overall. On the other hand, use of cross-category disturbance contributes to the forecast for specific categories in two chains. Interestingly these are occasional and convenience categories.

Combinations

Wednesday, 11:20 - 12:40; Chair: Nikolaos Kourentzes

The combined performance of the theta method with multiple temporal aggregation

Bonan Wang (University of Bath, b.wang2@bath.ac.uk); Fotios Petropoulos (University of Bath); Jooyoung Jeon (University of Bath); Gunes Erdogan (University of Bath)

Extracting more information from data for forecasting is a popular practice for forecasters. The theta method, a decomposition method used within this practice, has attracted great attention of researchers due to its simplicity and superior performance in M3-competition (Makridakis and Hibon, 2000). The theta method extracts information from data by generating various theta lines which are curvature-modified series from the original data. In the standard theta method, two theta lines are used. One is for modelling short-range behaviour and the other is for capturing long-range behaviour. Another method of extracting information from data is Multiple Temporal Aggregation (MTA). Using MTA, we can obtain information from data at each temporal aggregation level. In this study, we aim to investigate the performance of combining the theta method with MTA. We only focus on the standard theta method, Multiple Aggregation Prediction Algorithm (MAPA, one of the MTA approaches) and their combinations. We apply cross-validation for method selection to derive weights for different combination schemes. The empirical results with the M3-competition indicate the combination of the theta method and MAPA outperform the individual methods overall in terms of forecasting accuracy. Particularly, the simple combination is appealing as it demonstrates forecast accuracy comparable to sophisticated combinations and requires less computational cost.

STL-Bagging vs. Combining for extrapolative time series forecasting: An empirical analysis using M, M3 and NN3 competition data

Srihari Jaganathan (UCB Inc, sriharitn@gmail.com)

Bergmeir, Hyndman and Benitez (2016, Bagging exponential smoothing methods using STL decomposition and Box-Cox transformation, International Journal of Forecasting 32, 303-312) proposed a computationally intensive bootstrap aggregating (bagging) procedure using STL decomposition, box-cox transformation and exponential smoothing model for extrapolative time series forecasting. The STL-bagging procedure was effective on monthly series but not on quarterly and yearly frequencies in M3 forecasting competition dataset. The main objective of this study is twofold. Firstly, we extend STL-bagging procedure to ARIMA and Theta forecasting model. Secondly, we propose an approach of simple combination of diverse forecasting models and compare it with STL bagged forecasts from ETS, ARIMA and Theta models. We empirically demonstrate that simple combination of diverse forecasting models is superior in terms of accuracy as well computationally less expensive compared to bagging procedure across all frequencies – monthly, quarterly and yearly data. All empirical analysis are conducted using M, M3 and NN3 forecasting competition datasets. All programs are conducted using R statistical software program and are made available for replication analysis.

Higher moments of information criteria for forecast selection and combination

Nikolaos Kourentzes (Centre for Marketing Analytics and Forecasting, Lancaster University, nikolaos@kourentzes.com); Ivan Svetunkov (Lancaster University Management School); Stephan Kolassa (SAP Switzerland AG)

In doing forecast selection or combination we often rely on some performance metric. For example, that could be Akaike Information Criterion or some cross-validated accuracy measure for forecast selection, or respectively the analogous weights for combination purposes. There is ample empirical evidence demonstrating the reliability of such metrics, both in terms of forecast accuracy and automation of the forecasting process. Yet, these performance metrics are summary statistics, that do not reflect higher moments of the metrics. This poses similar issues to analysing only point forecasts to assess the risks associated with a prediction, instead of looking at prediction intervals as well. Looking at summary statistics does not reflect the uncertainty in the ranking of alternative forecasts, and therefore the uncertainty in selection and combination of forecasts. We propose a modification in the use of the AIC and an associated procedure for selecting a single forecast or constructing combination weights that aims to go beyond the use of a simplistic summary statistic to characterise each forecast. We demonstrate that our approach does not require an arbitrary dichotomy between forecast selection, combination or pooling, and switches appropriately depending on the time series on hand and the pool of forecasts considered. We proceed to show that the same procedure can be done with a wide variety of metrics, such as cross-validated errors and demonstrate its empirical performance on a large number of real time series from various sources.

Forecast evaluation

Wednesday, 11:20 - 12:40; Chair: Julie Bessac

Analysis of the Forecasting Performance of the Threshold Autoregressive Model

Paola Vaca (National University of Colombia, pavacag@unal.edu.co); Fabio Nieto (National University of Colombia)

In this investigation, we analyze the forecasting performance of the threshold autoregressive (TAR) model. To this aim, we find the Bayesian predictive distribution from this model, based on Nieto's (2008) and Calderón's (2014) methodologies and then, we conduct an out-of-sample forecasting exercise, where we compare forecasts from the TAR model with those from a linear model and nonlinear smooth transition autoregressive, self-exciting threshold autoregressive and Markov-switching autoregressive models. For this empirical forecast evaluation, we: i) use the U.S. and Colombian GDP, unemployment rate, industrial production index and inflation time series, which lead us to estimate and forecast forty models; and ii) use evaluation criteria and statistical tests that are mostly used in literature. We also compare the in-sample properties of the estimated models. For the overall comparison, we find a satisfactory performance of the TAR model in forecasting the chosen economic time series, and a shape changing characteristic in the Bayesian predictive distributions of this model that may capture the cycles in the economic time series. This gives important signals about the forecasting ability of the TAR model in the economic field.

How Good is Your Forecast? The Nuances to Prediction Evaluation Across Time

Aric LaBarr (Elder Research, Inc., aric.labarr@elderresearch.com)

When predicting across time, typical methodologies of prediction evaluation no longer hold true. It is not practical to take a hold-out sample randomly from observations in the data set or even use a typical k-fold cross-validation structure. Even newer methods of prediction evaluation in cross-sectional data like target shuffling should not just be applied to data where a temporal structure is inherent. How then can we determine if we have a good forecast? This talk highlights advantages and disadvantages to techniques evaluating predictions when forecasting future observations. It also discusses possible biases arising from time structures of data that should be considered.

Forecast evaluation with imperfect observations and imperfect models

Julie Bessac (Argonne National Laboratory, jbessac@anl.gov); Philippe Naveau (Laboratoire des Sciences du Climat et de l'Environnement - CNRS)

A key challenge associated with probabilistic forecast is to evaluate the quality of these datasets and the ability of the underlying model to produce relevant predictions. In statistics, one way to quantitatively evaluate and rank forecast models is through the use of scoring rules. This is typically based on scalar metrics and takes as inputs verification data and outputs from the forecasting model to be evaluated. Statistical properties of scores such as propriety insures mathematically consistent ranking between forecasts. However verification data are often tainted with errors due to for instance measurement errors or model errors. Consequently, forecaster's decisions are affected by these errors. In particular, we build on the work of Ferro (2017) to propose statistical models that incorporate errors of the verification data, mainly a Gaussian additive noise model and a Gamma multiplicative noise one. These models allow to express proper scoring rules such as the log-score or the Continuous Rank Probability Score (CRPS) while accounting for the error in the verification data.

We discuss the statistical properties of these proposed scores, and we illustrate the benefit of accounting for errors in the forecasts and the verification data in terms of decision-making process.

Electricity load forecasting

Wednesday, 14:00 - 15:20; Chair: Xiaoyin Wang

IoT Driven Analytics Will Define the Future Electric Utility

Bradley Lawson (SAS Institute Inc, bradley.lawson@sas.com)

Electric utilities are facing a rapidly changing landscape fueled in large part by the falling price of batteries. Lower priced batteries will allow Individual homes and businesses to couple batteries with their solar panels, but the biggest effect will come from the growth of electric vehicles(EVs). In many areas of the US and the world, one new EV charger can add as much demand as a new customer. Fast chargers that recharge EVs in just a few hours can add three times that demand. Too many EVs in a localized area could overload equipment, reducing reliability and requiring costly equipment upgrades.

Analytics and the Internet of Things (IoT) will become partners to create systems that enable the customer and utility to continually communicate their status and options. This will give customers maximum flexibility to charge their EVs without increasing the utility's costs. With a two-way flow of information, the utility will be able to take advantage of excess energy from EVs to maintain grid stability and reduce costs. Robust communications would allow utilities and customers to coordinate demand response to lower the utility demand at critical times. Historically, actions would be to simply disconnect a water heater or air conditioner for a few hours. Advanced analytics coupled with robust communications will allow complex actions as instructing a home thermostat to change settings.

This paper reviews the individual systems already in place, their integration and the analytics needed for a system that helps keep utility costs low while giving customers maximum flexibility.

Forecasting energy consumption combining Bagging and Clusters

Tiago Dantas (Pontifical Catholic University of Rio de Janeiro (PUC-Rio), t.mendesdantas@gmail.com); Fernando Cyrino (Pontifical Catholic University of Rio de Janeiro (PUC-Rio))

In a recent paper, Oliveira and Cyrino Oliveira (2018) have used Bootstrap Aggregating (Bagging) to generate highly accurate energy consumption forecasts for developed and developing countries. To do so, the authors have generated Bootstrap versions using Sieve Bootstrap (RSB) and Moving Block Bootstrap (MBB) and produced forecasts using Exponential Smoothing and ARIMA. This paper investigates the effects of incorporating a clustering phase to the approach used by the authors in order to reduce variance and, therefore, the forecast error. The results are validated using time series from the energy sector and compared to approaches using Bagging and other benchmarks, such as Artificial Neural Networks (ANN), Support Vector Regression (SVR), Theta method and Singular Spectrum Analysis (SSA).

To retrain or not to retrain – data cleaning vs. online learning in energy forecasting

Gregory Merkel (Marquette University GasDay Laboratory, gregory.merkel@marquette.edu); Richard Povinelli (Marquette University GasDay Laboratory)

In academic forecasting it is often an unstated assumption that forecasting models are retrained for every individual forecast to account for recent trends. However, in many domains, the immediate on-line data quality may not be good enough to support new models for each forecast point, and models are instead retrained monthly or annually. This paper seeks to determine which of these retraining schedules works best for short-term load-forecasting of natural gas.

Natural gas is the fuel of choice for many applications, including space heating and industrial processes. It is becoming the number one fuel for electricity generation. One of the problems that must be overcome in building accurate natural gas forecasters is the issue of data quality. Currently, when we build models a substantial effort is made to clean the data by removing outliers, identifying naive disaggregation, and correcting conversion factors. For models trained once a year this process is justified, and adequate time is available for data cleaning. However, for models retrained daily, this level of data cleaning is not possible.

We compare two forecasting strategies. The first involves models retrained yearly with extensive off-line cleaning and some data correction as needed. The second involves models retrained daily with rudimentary cleaning of on-line data and some seasonality adjustments. The tradeoff is that the models retrained yearly have greater opportunity for data cleaning and correction but can't incorporate recent trends, while models retrained daily can incorporate recent trends but some of the 'recent trends' may be due to data quality issues. For both strategies, we consider linear regression, artificial neural network, and ensemble models. The primary metric considered is weighted mean absolute percent error (WMAPE). We compare the forecasting strategies on more than 60 data sets.

Methodology for Natural Gas Demand Forecasting

Xiaoyin Wang (Towson University, xwang@towson.edu); Yunwei Cui (Towson University)

Accurate forecasting of the United States natural gas consumption is of crucial importance to energy providers for decision making regarding natural gas purchasing and energy pricing to increase profits and efficiency. Energy forecasting covers a wide range of forecasting problems, such as generation forecasting, load forecasting, price forecasting, demand response forecasting, and so on. The project aims to seek effective methods to improve the accuracy of natural gas demand forecasting in the US. To match the EIA monthly data, we establish a parallel method to make a natural gas demand forecast. The method models the scrape data and ratio data individually in two modules and then combines the outputs of the two modules to generate the forecast.

An assortment of models to forecast the natural gas scrape data are investigated. In particular, we create a multi-scale method to treat the long term trend, medium and short term parts respectively. Two models of the multi-scale method are developed based on the semi-parametric technique and time series method. We test the proposed methods on the residential and commercial natural gas consumption data in the New England area for one year. The proposed multi-scale method consistently generates models which are among the top performers. Moreover, the proposed parallel method performs very well for generating forecasts that closely match the published EIA data.

Machine learning and neural networks

Wednesday, 14:00 - 15:20; Chair: Yingying Chen

Incorporating structure into neural forecasts

Tim Januschowski (Amazon); David Salinas (Amazon, dsalina@amazon.com); Jan Gasthaus (Amazon); Valentin Flunket (Amazon); Yuyang Wang (Amazon); Ali Caner Turkmen (Amazon)

Deep-learning has been highly successful in industrial forecasting applications, in particular when millions of observations from related time series are available. However, many forecasting use-cases contain much fewer observations, and classical methods such as ARIMA, Exponential Smoothing (ETS) or Innovation State Space Models are more successful in this regime. These methods provide reliable forecasts with less data, and typically work well with only a few hundreds of observations on a single time-series. This is partly due to their ability to encode structural information on time series as modelling assumptions, such as temporal continuity, trend and seasonality. This paper aims to combine the strengths of classical approaches with state of the art recurrent neural network models. We extend an earlier model, DeepAR, an auto-regressive recurrent neural network using Long Short Term Memory along several directions using structure that encodes knowledge about the problem. We demonstrate that our extensions improve the data-efficiency: with proper regularization and the incorporation of these structural assumption, our models are competitive with e.g. ARIMA and ETS in the small-data regime and can outperform them with as little as a few hundred observations for training.

Forecasting Using Dynamic Graph Neural Network Systems

Slawek Smyl (Uber Technologies, slaweks@hotmail.co.uk)

Dynamic Graph Neural Network (DGNN) systems, like Dynet and Pytorch, offer imperative programming paradigm, instead of a typical configure-and-run approach. This makes developing (and especially debugging) of non-standard NN systems much easier, but more importantly, it allows developing hybrid blackbox-algorithmic, hierarchical, and explainable models.

These capabilities will be illustrated with time series forecasting models that achieve superior results on M3 and M4 Competition data sets.

Cross-Validation based Forecasting Method: A Machine Learning Approach

Jeronymo Pinto (CEMAP-EESP-FGV, jeronymomp@gmail.com); Emerson Marçal (CEMAP-EESP-FGV e CCSA-Mackenzie)

Advance on information science allows analysts to develop a plethora of forecasting models, as well as automatic algorithm to choose different specifications. Our paper aims to evaluate two novel methods on selecting the best forecasting model or its combination based on a Machine Learning approach. The methods are based on the selection of the "best" model, or combination of models, by cross-validation technique, one of the most used techniques in time series forecasting.

The first one is based on the seminal paper of Granger-Bates (1969), but weights are estimated by a process of cross-validation applied on the training set. This is a modified version of Granger and Bates procedure (MGB). The second one selects the model with the best forecasting performance in the process described above, which we called CvML (Cross-Validation Machine Learning Method).

Analytically, both forecasts are constructed based on models obtained in the training set to assess the best model to forecast. We opt to use an expanding window on cross-validation exercise.

The following models are used: exponential smoothing, SARIMA, artificial neural networks and Threshold autoregression (TAR). Model specification is chosen by R packages: forecast and TSA.

Both methods – CvML and MGB - are applied to these models to generate forecasts from one up to twelve periods ahead. Frequency of data is monthly. We run the forecasts exercise to monthly series of Industrial Product Indices for seven countries: Canada, Brazil, Belgium, Germany, Portugal, UK and USA. The data was collected at OECD data, with 504 observations.

We choose as naïve benchmark double seasonal difference model (SDD), as suggested by Clements and Hendry (2001). To compute forecast combinations, we also run model confidence set (MCS) developed by Hansen et al. (2011) in the training set. Only the best models were used to compute combined forecasts, similar to Garcia et al. (2017).

Our complete set of models contains CvML, MGB, SDD, exponential smoothing, SARIMA, artificial neural networks, TAR, and combined models – simple forecast averages, Granger-Bates method and MCS Combination.

Our results suggest that MGB did not performed well. However, CvML had lower mean absolute error and mean squared error for most of countries and forecast horizons, particularly at longer horizons.

CvML forecast were compared to all models forecast at pairs using Diebold Mariano test. We collected evidence that CvML outperformes rival models in most cases. Best results are obtained for Belgium, Canada, United Kingdom, and United States.

Demand forecasting using machine learning models

Yingying Chen (Microsoft, yinchen@microsoft.com); Xin Li (Microsoft); Kaan Katircioglu (Microsoft)

Demand forecasting is a universal problem that has been seen in a lot of the industrial supply chain practices. Traditionally, a variety of time series based statistical models have been studied extensively to forecast demand based on the historical knowledge of the demand. However, with the advancements in machine learning models, understanding how to apply machine learning models in demand forecasting and how its performance compares with time series-based forecasting models, and under what circumstances they can be used successfully are questions all worthy investigating.

In our study, we apply different machine learning models in demand forecasting with a selection of variety of predictor variables as well as dependent variables. In selecting predictor variables, we consider multiple factors that may explain the future demand. This includes different number of lagged historical observations of demand, the number of distinct user accounts, as well as the user profile information such as their geographical regions. We also compare the performance by using different dependent variables, including the use of demand itself, the use of multiplicative growth rate (defined as the future demand divided by the historical demand), and the use of additive growth rate (defined as the difference between the future demand and the historical demand).

We applied the methods described above in regional usage demand forecasting for our Microsoft's cloud-based compute resources. Our study suggests that proper choice of predictor variables affects forecast accuracy significantly. Increasing the number of predictor variables not only decreases the forecast accuracy but also increases the variance in forecasts. Using demand itself as the dependent variable in the model training significantly under-forecasts demand in large regions, while over-forecasting demand in small regions. On the other hand, using growth rate as the dependent variable provides more accurate forecasts across all regions. We also compared our machine learning models

with autoregressive integrated moving average (ARIMA) and exponential smoothing (ETS) models. The results indicate that with the proper choice of predictor variables and dependent variables, the forecast accuracy of machine learning models is superior or equivalent to the time-series models. We believe the lessons learned from this study and the outcome can be easily applied to other industrial demand forecast scenarios.

Hierarchies

Wednesday, 14:00 - 15:20; Chair: Anastasios Panagiotelis

Optimal Transport and Spatio-Temporal Forecasting

Lucas Roberts (Amazon); Leo Razoumov (Amazon); Lin Su (North Carolina State University); Yuyang Wang (Amazon, wangyuyang1028@gmail.com)

Demand forecasting plays a central role in supply chain management, precipitating automatic ordering, in-stock management, and facilities location planning amongst others. The classical forecasting problem in this context constitutes predicting sales, given a feature set, regardless the location of each sale. New product lines and services may exhibit rapid growth which requires a customized perspective in the forecasting domain. These challenges call for a localized perspective which requires that we go beyond the coarser granularity and forecast the demand of a product for different locales. Given a coarser granularity forecast for each product, one sensible approach to the spatio-temporal (local) forecast is to generate a forecast for each locale, a partition of the coarser granularity forecast. Despite recent work on spatiotemporal and related hierarchical forecasting, assessing the quality of such forecasts remains an open question in the supply chain optimization setting. Standard forecast evaluation criteria such as mean absolute percentage error (MAPE) or quantile loss (QL) fail to take into account the main purpose of the spatio-temporal forecast: deciding the where to store on-hand inventory to minimize the cost of satisfying the customers' demand. To this end, any orthodox accuracy criteria should consider both the location of a product's demand and the topology of the distribution network. In this talk, we propose a metric based on optimal transport (OT) distance, which respects the intrinsic geometry of the underlying transportation network. To deal with the high computation complexity of solving the linear program associated with OT, we propose a regularized OT based on Gini impurity function and we demonstrate through simulations both the qualitative and the quantitative characteristics of regularized OT problem. Comparing to the popular entropic regularized OT formation, we show that, empirically, its convergence to the classical OT solution is much faster in terms of regularization parameter and thus results in a numerically more stable algorithm. Finally, we propose spatiotemporal forecasting algorithms based on deep learning model, with its simplest case multinomial logistic regression, that directly optimizes the introduced OT metric.

Hierarchical forecasting from a parameter estimation perspective

Daniel Waller (Centre for Marketing Analytics and Forecasting, Lancaster University, d.waller1@lancaster.ac.uk); John Boylan (Centre for Marketing Analytics and Forecasting, Lancaster University); Nikolaos Kourentzes (Centre for Marketing Analytics and Forecasting, Lancaster University)

Retailers are frequently faced with extremely variable data series when analysing sales of individual stock-keeping units (SKUs). The low signal-to-noise ratio typically displayed by these series not only makes forecasting challenging, but also causes problems in the estimation of certain quantities or parameters. For example, estimates of price elasticity often exhibit an excessive magnitude, or even the wrong sign. Whilst inaccurate parameter estimation might also lead to poor forecasts, the estimates themselves are often of further interest to a company, and should not be neglected, as they carry market intelligence, for example, elasticities of price and promotions.

We examine ways to improve estimation and forecasting by using hierarchical information, which is common in situations of this type. In the hierarchical forecasting area, the optimal combination of

forecasts at all levels has spurred a recent surge of interest in the topic. A strength of this combination method is that no assumption is made on the forecasting model that generates the inputs. We explore whether including the assumption that the forecasting model takes a specific form, can lead to improvements not only in how the forecasts are adjusted, but also in how the parameter estimates of those models might be simultaneously adjusted alongside them. The aim is a consolidated hierarchical forecasting approach which emphasises improved parameter estimation, but also reconciles forecasts in line with those adjustments. The evaluation examines the impact on both estimation accuracy and forecast accuracy.

Improving Regional Revenue Forecasts using Product Hierarchy

Amita Gajewar (Microsoft Corp., amitag@microsoft.com)

With the increasing trend of machine learning (ML) influencing various business processes within Finance domain, one of its impactful applications is to forecast revenue very efficiently and accurately. In [1,2], we described the ML models that we built to forecast quarterly revenue for Microsoft's Finance organization. Using methodologies described in [1,2], we had developed ML models to forecast enterprise products' revenue for thirteen regions across the globe. These models used data available at individual region-level. As a next step, we wanted to forecast revenue for individual products within each region and use their sum to compute the region-level forecasts. The potential benefits include: more accurate forecasts per region, enable more people in the field to consume ML forecasts, and an ability to explain the region-level forecasts.

To develop the aforementioned models, we use some of the existing framework in [1,2]. However, we needed to address the challenge that the amount of historical revenue available for various products varied depending upon how old/new the product is, and for some of the products the revenue seemed to be diminishing/increasing in the most recent quarters. Therefore, we apply different methods based on the amount of historical data available. If there is sufficient historical data, then we apply ML models (e.g., RandomForest, Elastic-Net, SVM) in addition to the traditional time series models (e.g., ARIMA, ETS and STL). Otherwise we rely only on time series models. The methodology used to select the best forecast from these models also vary depending upon the availability of the cross-validation dataset which in turn depends on historical data availability. In this paper, we describe our modelling methodologies, features constructed for ML models, and final model selection criteria used. We also present results evaluated using real-world Finance data.

We ran the experiments for approximately thirty enterprise products in each of thirteen regions world-wide, constituting the yearly revenue in the order of tens of billions of dollars (USD) as of latest fiscal-year. The maximum historical data available was for about eight years and we used the last one year as a test dataset. We observed that using these models, the accuracy of forecasts improved for majority of the regions. This also helped in improving the overall world-wide forecast's accuracy by about 30 percent. (Note: due to sensitivity of the data, actuals errors cannot be disclosed). Seeing the promising results, these ML models are adopted by Microsoft's Finance team, deployed into production environment, and being used as part of the ensemble of models that generate quarterly enterprise revenue forecasts. Further, these forecasts are distributed to hundreds of controllers in the field. In future, it would be of interest to explore other ML techniques, e.g., LSTM, to forecast revenue of products that are deeper in hierarchy.

Probabilistic Forecasts in Hierarchical Time Series

Puwasala Gamakumara (Monash University); Anastasios Panagiotelis (Monash University, Australia, anastasios.panagiotelis@monash.edu); George Athanasopoulos (Monash University, Australia); Rob Hyndman (Monash)

Multiple time series that are characterized by an aggregation structure - often referred to as hierarchical time series - are common in practice. In this context we extend existing concepts from the point forecasting literature to probabilistic forecasts. We provide definitions of coherent probabilistic forecasts - that is probabilistic forecasts that satisfy aggregation constraints. We introduce a framework for methods that adjust incoherent probabilist forecasts in a way that they become coherent - a process we call probabilistic reconciliation. We also discuss the suitability of commonly used multivariate scoring rules in particular the log score and energy score to the hierarchical context.

Forecasting

Wednesday, 14:00 - 15:20; Chair: Emmanuel Silva

Outcome Prediction in the Practice of Law

Michael Gilliland (SAS Institute, mvgilliland@gmail.com); Mark Osbeck (University of Michigan Law School)

Lawyers are best recognized for their role as client advocate -- arguing the client's interests in legal proceedings (and, if necessary, in the media). But an equally important role is as advisor -- analyzing the client's various options and the likely outcome of each. Outcome prediction, therefore, is an essential lawyering skill. So how do lawyers make their predictions? And can outcome prediction be improved with data and analytics?

This presentation looks at traditional methods of outcome prediction used by lawyers, and examines the prospect of using data science to make better predictions. We find that predictive analytics can be expected to complement the traditional tools of outcome prediction. However, for reasons to be discussed, it is doubtful predictive analytics will totally replace traditional tools in the near future.

Combining Prediction Markets and Forecasting Contests

Alasdair Brown (University of East Anglia); James Reade (University of Reading, j.j.reade@reading.ac.uk)

Two popular methods for aggregating individual forecasts are prediction markets, where participants bet on the outcome of future events, and forecasting contests, where participants are ranked according to the accuracy of their forecasts. Can these methods be used in concert to produce more accurate forecasts? We analyse 1.79 million forecasts on oddsportal.com, a social network for sports tipsters. Tipsters are ranked according to the betting return on their tips. We find that an aggregation of these tips predicts sporting outcomes, after controlling for betting/prediction market prices. Rank-order forecasting contests, even without tangible prizes, are useful tools for eliciting crowd forecasts.

Forecasting Interval-Valued Time Series Based on a Truncated Mixture Transition Model

Yun Luo (University of California, Riverside, yluo021@ucr.edu); Gloria González-Rivera (University of California, Riverside)

We propose a new model to forecast the interval-valued time series data. This type of data is observed as an interval at a time, such as daily interval high/low stock price, and daily high/low temperature. The model specifies the conditional joint distribution of the upper and lower bounds of the interval to be a mixture of truncated bivariate normal distribution, with each component containing only one lag information. The truncated distribution guarantees that the forecast of the upper bound is always greater than the lower bound. Flexibility of mixture models allows for a better approximation to underlying true distribution, and provides a direct density forecast. We investigate the identification problem. For estimation, the traditional EM algorithm for mixture models does not provide closed-form solutions and the monotonicity of the likelihood cannot be preserved. We propose a new EM algorithm that encloses a higher level of data completeness and provide closed-form solutions. Monte Carlo simulations show convergence and monotonicity of the new EM algorithm.

Forecasting with Auxiliary Information in Forecasts using Multivariate Singular Spectrum Analysis

Emmanuel Silva (University of the Arts London, e.silva@fashion.arts.ac.uk); Hossein Hassani (OPEC Vienna); Mansi Ghodsi (Energy Community)

There exists a variety of published forecasts representing auxiliary information and freely accessible online. This increasing availability motivates the search for a possibility of exploiting the auxiliary information contained within a forecast to generate a new and improved forecast. The proposed theoretical concept can be exploited by any model which can consider data with different series lengths and accordingly, predictions into the future. The initial results are promising, following applications considering official forecasts generated via unknown time series models, whereby the theoretical approach succeeds in extracting the auxiliary information in the forecast for generating a new and more accurate forecast, along with statistically significant results in certain cases. The impact of filtering and the use of Google Trends in combination with the proposed methodology is also considered. We conclude that there is indeed a sound opportunity to exploit the forecastability of auxiliary information contained within existing forecasts.