ITISE-2024

PROGRAM and ABSTRACTS

15th-17th JULY, 2024
Gran Canaria (SPAIN)

ITISE-2024 Program

Sunday, July 14th, 2024	
18:30-20:00	**REGISTRATION DESK** ***(start at 18:30h but it is open during all the conference)***
18:30-20:00	**Upload the presentations to the room's computer (in case you haven't sent them by email).**

NOTES:

- All **Sessions A** will be held in Hotel Lopesan Villa del Conde Resort. They are **face-to-face sessions**, and they will also be shared on-line by Zoom. The **plenary lectures** are in **Session A.**
- All **Sessions B** will be held on-line (virtual) using Zoom.
- **Oral** Presentation: **15 minutes** (including questions). Short Presentation: **10 minutes** (including questions). Depending on whether there are absent speakers, times may be adjusted.
- **Poster** authors are requested to place their posters on the panels before the start of the poster session (e.g. morning posters can be placed before 10 o'clock, before the coffee break). It is recommended to use **A0 size** and large fonts.

Session A: Located on the last floor of the main building

Monday, July 15, 2024		
8:30	**REGISTRATION DESK** ***(start at 8:30h but it is open during all the conference)***	
	All **Sessions A: Oral face-to-face sessions**. All **Sessions B:** Oral (will be held on-line by Zoom)	
9:00-10:15	**Session A.1: Econometric Modelling of Financial Market Trends (Part I)**	**Session B.1: Econometric Models and Forecasting (Part I)**
10:15- 11:00	**Session A.2: Time Series Analysis for Sustainable and Resilient Infrastructure Systems**	**Session B.2: Computational Intelligence methods for Time Series (Part I).**
11:00-11:40	**COFFEE BREAK**	
11:00-13:00	**Session A.3: POSTER SESSION**	
11:40-12:55	**Session A.4: Time Series Analysis with Computational Intelligence in Energy Forecasting**	**Session B.3: Time Series Analysis for Sustainable and Resilient Infrastructure Systems**
13:00-14:00	**Session A.P1: Opening & Plenary Lecture.** **Prof. Faramarz F. Samavati** **Department of Computer Science, University of Calgary**	
14:00-16:00	**REST BREAK**	
16:00-17:45	**Session A.5: New Advances in Time Series Analysis and Forecasting (Part I)**	**Session B.4: Forecasting Financial Markets**
17:50-19:50	**Session A.6: Applications in: energy, finance, transportation, networks…** Short Presentation	**Session B.5: New Advances in Time Series Analysis and Forecasting (Part I)**

<table>
<tr><th colspan="3">Tuesday, July 16th, 2024</th></tr>
<tr><td rowspan="2">8:45</td><td colspan="2">REGISTRATION DESK
(start at 8:45h but it is opened during all the conference)</td></tr>
<tr><td colspan="2">All Sessions A: Oral <u>face-to-face sessions</u>.
All Sessions B and C: Oral (will be held on-line by Zoom)</td></tr>
<tr><td>9:00-9:45</td><td rowspan="2">Session A.7: Computational Intelligence methods for Time Series.</td><td></td></tr>
<tr><td>9:45- 10:45</td><td>Session B.6: Econometric Models and Forecasting (Part II)</td></tr>
<tr><td>10:45-11:15</td><td colspan="2">COFFEE BREAK</td></tr>
<tr><td>11:15-12:00</td><td>Session A.8: Econometric Modelling of Financial Market Trends (Part II)</td><td>Session B.7: Time Series Analysis with Computational Intelligence in Energy Forecasting</td></tr>
<tr><td>12:00-13:00</td><td>Session A.9: Forecasting theory, adjustment and data preprocessing methods.</td><td>Session B.8: Computational Intelligence methods for Time Series (Part II).</td></tr>
<tr><td>13:00-14:00</td><td colspan="2">Session A.P2: Plenary Lecture.
Prof. Hossein Bonakdari
P.Eng., esteemed professor at the University of Ottawa, Canada.
Assoc. Editor Highlights of Sustainability. Editorial board, Natural Resources</td></tr>
<tr><td>14:00-16:00</td><td colspan="2">REST BREAK</td></tr>
<tr><td>16:00-17:25</td><td>Session A.10:
Economics as a Methodological Progression: The Pipeline from Conventional Methods to Supervised and Unsupervised Machine Learning</td><td rowspan="2">Session B.9: New Advances in Time Series Analysis and Forecasting (Part II)</td></tr>
<tr><td>17:30-18:45</td><td>Session A.11: New Advances in Time Series Analysis and Forecasting (Part II)
Short Presentation</td></tr>
<tr><td>20:30</td><td colspan="2">GALA DINNER
Hotel Lopesan Baobab 5*
(15 minutes walking from Hotel Lopesan Villa del Conde Resort)</td></tr>
</table>

Wednesday, July 17th, 2024		
8:45	**REGISTRATION DESK** ***(start at 8:45h but it is opened during all the conference)***	
	All **Sessions A: Oral face-to-face sessions**. All **Sessions B:** Oral (will be held on-line by Zoom)	
9:00-10:15	**Session A.12: Advanced Applications in Time Series Forecasting**	**Session B.10: New Advances in Time Series Analysis and Forecasting (Part III)**
10:15- 11:00	**Session A.13: Forecasting Financial Markets**	**Session B.11: Data preprocessing methods in Time Series**
11:00-11:30	**COFFEE BREAK**	
11:30-13:00	**Session A.14: Advanced econometric methods**	
13:00-13:45	**Session A.15: Functional Time Series Analysis and Application**	
13:45	**Closing ceremony**	

ITISE 2024 PROGRAM

Monday, July 15, 2024

(9:00-10:15) Session A.1: Econometric Modelling of Financial Market Trends (Part I)

Chairman: Dr. Marta Tolentino and Dr. Maria de la O Gonzalez Perez

Testing for asymmetric correlations between US sector returns and interest rate changes **(Ref: *66*)**

Francisco Jareno, María de La O González and José M^{a} Almansa

Gender as a tool for diversification **(Ref: *2061*)**

Ana Escribano, Antonio Díaz and Rocio Hidalgo

Measuring the impact of climate risk in financial markets: a joint quantile and expected shortfall regression mode **(Ref: *2819*)**

Lidia Sanchis-Marco and Laura Garcia-Jorcano

Gauging Growth Risk in an International Financial Centre: Some Evidence from Singapore **(Ref: *3699*)**

Hwee Kwan Chow

Which is the Significance of Environmental Awareness in Energy Investment Decisions? A Study on Portfolio Rebalancing with an Environmental Focus **(Ref: *5286*)**

Carlos Esparcia Sanchís and Antonio Diaz

(10:15-11:00) Session A.2: Time Series Analysis for Sustainable and Resilient Infrastructure Systems

Chairman: Dr. Maria Luisa Villani and Dr. Ebrahim Ehsanfar

Towards an automatic tool for resilient waterway transport: the case of the Italian river PO **(Ref: *8495*)**

Maria Luisa Villani, Ebrahim Ehsanfar, Sohith Dhavaleswarapu, Alberto Agnetti, Luca Crose and Sonia Giovinazzi

Energy Efficiency Evaluation of Frameworks for Algorithms in Time Series Forecasting **(Ref: *9117*)**

Sergio Aquino-Britez, Pablo García Sánchez, Andrés Ortiz and Diego Aquino-Britez

Detecting Trend Turning Points in PS-InSAR Time Series: Slow-Moving Landslides in Province of Frosinone, Italy **(Ref: *9712*)**

Ebrahim Ghaderpour, Benedetta Antonielli, Francesca Bozzano, Gabriele Scarascia Mugnozza and Paolo Mazzanti

(11:00-13:00) Session A.3: POSTER SESSION

Chairman: Dr. Fernando Rojas and Dr. Alberto Guillen

Assessing Urban Bicycle Traffic Using a Forecasting Model **(Ref: *1603*)**

Anamaria Ilie, Eugen Roșca, Cristina Oprea, Aura Ruscă, Oana Dinu and Valentina Radu

Comparative methods for structural breaks in time series **(Ref: *1786*)**

Dulce Gomes

Prediction of the characteristics of concrete containing crushed brick aggregate **(Ref: *2679*)**

Marijana Hadzima-Nyarko, Miljan Kovacevic, Ivanka Netinger Grubesa and Silva Lozancic

Chlorophyll-a time series study on a saline Mediterranean lagoon: the Mar Menor case **(Ref: *3946*)**

Arnau Garcia i Cucó, José Gellida Bayarri, Beatriz Chafer-Dolz and José M. Cecilia

On the Use of Preprocessing and a Consensus Model to Calculate the Social Security Contributions of Spanish Self-Employed Workers Based on Adjusted Revenue Estimate **(Ref: *4241*)**

Luis Palomero López de Armentia, Vicente Garcia and Jose Salvador Sanchez

Developing a forecasting model for allergenic nettle pollen in North-Eastern Croatia **(Ref: *4352*)**

Edita Stefanic, Sanda Rasic, Pavo Lucic, Marin Lukacevic and Slavica Antunovic

Simulating the Aerial Ballet: The Dance of Fire-Fighting Planes and Helicopters **(Ref: *4747*)**
Kalle Saastamoinen, Juha Alander and Lauri Honkasilta

Estimation of the Gini coefficient using incomplete data **(Ref: *4797*)**
Szabolcs Kelemen, Máté Józsa and Zoltán Néda

Studying LF and HF time series to characterize cardiac physiological responses to mental fatigue **(Ref: *5133*)**
Alexis Boffet, Veronique Deschodt Arsac and Eric Grivel

Exploring Different Modelling Approaches to Forecast Acute Respiratory Infections: An Italian Epidemiological Time Series Study **(Ref: *5280*)**
Riccardo Boracchini, Benedetta Canova, Pietro Ferrara, Elisa Barbieri, Pietro Giorgio Lovaglio, Giovanni Corrao, Daniele Donà, Carlo Giaquinto, Costanza Di Chiara and Anna Cantarutti

Forecasting methods for road accidents, the case of Bucharest city **(Ref: *5548*)**
Cristina Oprea, Eugen Rosca, Ionut Preda, Anamaria Ilie, Mircea Rosca and Florin Rusca

A time series analysis of provincial growth in Spain in the 20th century **(Ref: *5821*)**
Rafael González-Val and Miriam Marcén

Forecasting at Scale: An AutoML Framework for Time-Series Data Challenges **(Ref: *6063*)**
Siddharth Chatterjee, Sourav Banerjee and Divyananda Dileep Aravapalli

Forecasting energy poverty in national energy and climate plans **(Ref: *6482*)**
Renata Slabe Erker, Montserrat González Garibay, Kaja Primc, Darja Zabavnik and Miha Dominko

The Cassandra Method: Dystopian Visions As A Basis For Responsible Design. **(Ref: *6494*)**
Sarah Diefenbach and Daniel Ullrich

Modelling the Daily Concentration of Airborne Particles Using 1D Convolutional Neural Networks **(Ref: *7014*)**
Ivan Gudelj, Mario Lovrić and Emmanuel Karlo Nyarko

Modeling the future of hydroelectric power: a cross-country study **(Ref: *7238*)**
Farooq Ahmad, Livio Finos and Mariangela Guidolin

Sentiment Dynamics and Volatility: a Study Based on GARCH-MIDAS and Machine Learning **(Ref: *7371*)**
Gianmarco Vacca and Luigi Riso

Spectral Characteristics of Strong Ground Motion Time Series for Low to Medium Seismicity Regions with Deep Soil Atop Deep Geological Sediments—An Example of the City of Osijek, Croatia **(Ref: *8118*)**
Silva Lozancic, Borko Bulajic, Gordana Pavic, Ivana Bulajic and Marijana Hadzima-Nyarko

Minimal Reservoir Computing **(Ref: *8321*)**
Haochun Ma, Davide Prosperino and Christoph Räth

Smart Belay Device for Sport Climbing - an Analysis about Falling **(Ref: *9345*)**
Heiko Oppel and Michael Munz

Legendre polynomial modelling-based permutation entropy to analyse encrypted Time series **(Ref: *9491*)**
Meryem Jabloun

Prediction and estimation of yield cereal production using NDVI time series (Sentinel-2) data in central Spain **(Ref: *9973*)**
César Sáenz, Alfonso Bermejo-Saiz, Víctor Cicuéndez, Tomás Pugni, Diego Madruga, Javier Litago and Alicia Palacios-Orueta

Assessing the Pre-processing Benefits of Data-Driven Decomposition Methods for Phase Permutation Entropy - Application to econometric Time-series. **(Ref: *9974*)**
Meryem Jabloun

(11:40-12:55) Session A.4: Time Series Analysis with Computational Intelligence in Energy Forecasting

---

Chairman: Dr. Peter Glösekötter, Dr. Moerschell Joseph and Dr. Ignacio Rojas

Towards Resolving the Ambiguity in Low-Field All-Optical Magnetic Field Sensing With High NV-Density Diamonds **(Ref: *1520*)**
Ludwig Horsthemke, Jens Pogorzelski, Dennis Stiegekötter, Frederik Hoffmann, Ann-Sophie Rösner, Markus Gregor and Peter Glösekötter

Optimizing Biogas Power Plants through Deep Learning-Aided Rotor Configuration **(Ref: *1964*)**
Andreas Heller, Héctor Pomares and Peter Glösekötter

Advancing Sustainable Mobility and Transport through Predictive Multi-Level Control of Fuel Cell Electric Vehicles **(Ref: *2713*)**
Christoph Hametner and Stefan Jakubek

Forecasting Electricity Prices in Times of Distress using Bid Data **(Ref: *7840*)**

Aitor Ciarreta and Blanca Martínez-Gonzalo

Neural Network Estimator Performance and Electric Energy Network Optimization **(Ref: *6315*)**

Joseph Moerschell, Fereshteh Jafari and Charles Praplan

(13:00-14:00) Opening Ceremony. Plenary Talk: Prof. Faramarz F. Samavati

Department of Computer Science, University of Calgary

(16:00-17:45) Session A.5: New Advances in Time Series Analysis and Forecasting (Part I)

Chairman: Dr. Emmanuel Karlo Nyarko and Dr. Diefenbach Sarah (tentative)

Evaluating leading and coincident indicators of regional business cycles, a closer look into the Spanish scenario **(Ref: *27*)**

Marco Aurelio Pérez Navarro and Aránzazu De Juan Fernández

Application of the optimised Pulse Width Modulation (PWM) based encoding-decoding algorithm for forecasting with Spiking Neural Networks (SNN) **(Ref: *1448*)**

Sergio Lucas and Eva Portillo Pérez

Modelling Explosive Non-stationarity of Ground Motion Shows Potential for Landslide Early Warning **(Ref: *9458*)**

Michael Manthey, Guoqi Qian and Antoinette Tordesillas

A Global Deep Learning Perspective on Australia-Wide Monthly Precipitation Prediction **(Ref: *9610*)**

Luyi Shen, Guoqi Qian and Antoinette Tordesillas

A comparative analysis of forecasting models for CO2 prediction in school classrooms in Navarra **(Ref: *9781*)**

Peio Garcia, Aranzazu Jurio and Daniel Paternain

(17:50-19:50) Session A.6: Applications in: Energy, Finance, Networks, Meteorology, Health, etc (Short Presentation)

Chairman: Dana Bryazka and Dr.Benjamin Ott

Dynamic Maps Powered by Machine Learning and Time Series Classification for Wildfire Risk Management **(Ref: *298*)**

Nicolò Perello, Giorgio Meschi, Andrea Trucchia, Mirko D'Andrea, Silvia Degli Esposti and Paolo Fiorucci

Exploring Optimal Strategies for Small-Hydro Power Forecasting: Training Periods and Methodological Variations **(Ref: *2318*)**

Duarte Lopes, Isabel Preto and David Freire

Enhanced Renewable Power Forecasting through NWP and Historical Power Data Integration **(Ref: *2434*)**

Isabel Preto, António Couto, Ricardo Faria, Hugo Algarvio, Duarte Lopes and Ana Estanqueiro

Big Data Techniques Applied to Forecast Photovoltaic Energy Demand in Spain **(Ref: *9369*)**

Jorge Tapia García, Luis G. Baca Ruiz, David Criado Ramón and María del Carmen Pegalajar Jiménez

Enhanced Volcanic Signal Detection in Santorini: Implications for Sustainable Tourism Development **(Ref: *2301*)**

Antonios Marsellos, Katerina Tsakiri, Stelios Kapetanakis, Nick Ptak and Faith Renner

Forecasting the impacts of smoking prevalence scenarios from 2022 to 2050 **(Ref: *9384*)**

Dana Bryazka, Marissa Reitsma, Natalia Bhattacharjee, Stein Emil Vollset and Emmanuela Gakidou

Signal Detection in High-Noise Time Series Data Using R **(Ref: *4392*)**

Katerina Tsakiri and Antonios Marsellos

A quantum current sensor for energy flow estimation in smart grids **(Ref: *3380*)**

Frederik Hoffmann, Ann-Sophie Bülter, Ludwig Horsthemke, Dennis Stiegekötter, Jens Pogorzelski, Peter Glösekötter and Markus Gregor

Calibration Free Current Measurement with Integrated Quantum Sensor **(Ref: *7470*)**
Jens Pogorzelski, Jonas Homrighausen, Ludwig Horsthemke, Dennis Stiegekötter, Frederik Hoffmann, Ann Sophie Bülter, Markus Gregor and Peter Glösekötter

Digital Boxcar Averager on a Microcontroller for Pulsed ODMR Measurements of NV Centers **(Ref: *9480*)**
Dennis Stiegekötter, Ludwig Horsthemke, Ann-Sophie Bülter, Frederik Hoffmann, Jens Pogorzelski, Peter Glösekötter and Markus Gregor

Tuesday, July 16th, 2024

(9:00-10:45) Session A.7: Computational Intelligence methods for Time Series.

Chairman: Dr. Héctor Pomares

Automatic selection of methods to perform combination of predictions in monthly series **(Ref: *828*)**
Carlos García-Aroca, Asuncion Mayoral, Javier Morales and José Vicente Segura

A Hybrid, Computer Intensive Approach Integrating Machine Learning and Statistical Methods for Fake News Detection **(Ref: *1528*)**
Livio Fenga

Continual Learning for Time Series Forecasting **(Ref: *2474*)**
Quentin Besnard and Nicolas Ragot

Hybridizing Machine Learning with Time Series Analysis for Enhanced Forecasting in Management Science and Operational Efficiency: A Systematic Review **(Ref: *2919*)**
Aydin Teymourifar and Maria A. M. Trindade

Neural Networks Global Models with Context **(Ref: *3460*)**
Slawek Smyl

Drought forecasting for agricultural insurance payments in grasslands using LSTM neural networks **(Ref: *5058*)**
Tom Vanwalleghem and Filippo Milazzo

Modeling algorithms for anomaly detection based on real-time sensor data: The case of a Desalination Plant **(Ref: *8850*)**
Ron Weitzman and David Gavbriel Pinto

(11:15-12:00) Session A.8: Econometric Modelling of Financial Market Trends (Part II)

Chairman: Dr. Ana Escribano and Dr. Rafael González-Val

Impact of interest rate fluctuations on traditional, environmental and stable cryptocurrency returns **(Ref: *5562*)**
Maria de La O Gonzalez, Francisco Jareño and Jose M^{a} Almansa

From Upside and Downside Risks to Resilience: The ESG Imperative in Energy Market Sustainability **(Ref: *6812*)**

Carlos Esparcia, Mariya Gubareva and Francisco Jareño Cebrián

Examining the performance of some stock market indices: before and after the Russia Ukraine conflict (2021-2023) **(Ref: *8685*)**

Marta Tolentino, María del Valle Fernández and Laura Prieto

(12:00-13:00) Session A.9: Forecasting theory, adjustment and data preprocessing methods.

Chairman: Dr. Kalle Saastamoinen and Dr.Ebrahim Ghaderpour

Ordinary and robust NIPALS decompositions to explore time series **(Ref: *1971*)**

Adelaide Freitas, Alberto da Silva and Filipa Santana

State-Space Modeling in Time Series Analysis: A Data Assimilation Approach and Challenges in Parameter Estimation **(Ref: *3097*)**

Marco Costa, Magda Monteiro, A. Manuela Gonçalves and F. Catarina Pereira

Statistical analysis of longest dry spells phenomenon in Northern Tunisia based on probability laws **(Ref: *4802*)**

Majid Mathlouthi and Fethi Lebdi

Robust Multitaper Tests for Detecting Frequency Modulated Signals **(Ref: *8767*)**

Benjamin Ott, Glen Takahara and Wesley Burr

(13:00-14:00) Plenary Talk: Prof. Hossein Bonakdari

P.Eng., esteemed professor at the University of Ottawa, Canada.
Assoc. Editor Highlights of Sustainability.
Editorial board, Natural Resources

(16:00-17:25) Session A.10: Economics as a Methodological Progression: The Pipeline from Conventional Methods to Supervised and Unsupervised Machine Learning

Chairman: Dr. James Chen

Relevance, Redundancy, and Regularization: Penalized Regression and the Quest for the Lo Quasi-Norm **(Ref: *6188*)**
James Chen

Shadows of Resilience: Exploring the Impact of the Shadow Economy on Economic Stability in Times of Crisis **(Ref: *2242*)**
Charalampos Agiropoulos, James Chen, Thomas Poufinas and George Galanos

Feature presentation: The proper sampling of feature space as a reasonable response to the unreasonable ineffectiveness of mathematics in economics **(Ref: *2484*)**
James Chen

Exploring Regional Determinants of Tourism Success in the Eurozone: An Unsupervised Machine Learning Approach **(Ref: *4123*)**
George Galanos, Charalampos Agiropoulos, James Ming Chen and Thomas Poufinas

Multi-Source Big Data for Fine-Scale Population Distribution Simulation in Rural Areas **(Ref: *2533*)**
Li Li, Ning Niu and Xiaojian Li

(17:30-18:45) Session A.11: New Advances in Time Series Analysis and Forecasting (Part II) (Short Presentation)

Chairman: Dr. Pablo García Sánchez

The relationship between corporate social responsibility and ethical leadership: the role played by corruption in countries. **(Ref: *1594*)**
Isidro Peña, Rosa M. Muñoz, Silvia M. Andrade and Graca Silva

Boosted Value at Risk and Expected Shortfall – application of gradient boosting machine learning models in market risk estimation problem **(Ref: *2094*)**
Michał Woźniak

Promoting Electric Vehicle Growth Through Infrastructure and Policy: A Forecasting Analysis **(Ref: *2579*)**
Anuva Banwasi, Brennan McManus and Adele Sinai

Forecasting Hungarian FX Bond Yields **(Ref: *5226*)**
Dávid Tran

Multicriteria Forecast Combination with Non Linear Programming **(Ref: *7031*)**
Oscar Generoso Gutiérrez, Clara Simón de Blas and Ana Elisabeth García Sipols

Forecasting stock market dynamics from market cap time series of firms through fluctuating selection **(Ref: *7207*)**
Hugo Fort

Wednesday, July 17th, 2024

(9:00-10:15) Session A.12: Advanced Applications in Time Series Forecasting

Chairman: Dr. Antonios Marsellos and Dr. Christoph Hametner

Optimal Forecast Combination for Japanese Tourism Demand **(Ref: *776*)**
Saeed Heravi, Bo Guan, Yongmi Fang, Emmanuel Silva and Hossein Hassani

Self-normalizing Tests Using the Cauchy Distribution **(Ref: *2366*)**
Uwe Hassler and Mehdi Hosseinkouchack

Crisis and Youth Inactivity: Central and Eastern Europe During the Financial Crisis of 2008 and the COVID-19 Outbreak of 2020 **(Ref: *6602*)**
Nataša Kurnoga, Tomislav Korotaj and James Ming Chen

Detecting Change Points in Temporally Correlated Data: An Environmental Time Series Application **(Ref: *7967*)**
Magda Monteiro and Marco Costa

Modeling and forecasting hourly ozone concentration using functional data approach **(Ref: *8519*)**
Ismail Shah

(10:15-11:00) Session A.13: Forecasting Financial Markets

Chairman: Dr. Katerina Tsakiri and Dr.Saeed Heravi

A Time Series Dashboard for Decision Makers in Higher Education Institutions **(Ref: *946*)**
Gabriel David Pinto and Ron Weitzman

Harnessing Manager Sentiment and Ordinal Regression for Corporate Credit Rating Forecasting **(Ref: *4296*)**
Petr Hajek

GARCH Models Under Additive Outliers: A Robust M-quantile Approach **(Ref: *9812*)**
Patrick Patrocinio, Valderio Reisen and Pascal Bondon

(11:30-13:00) Session A.14: Advanced econometric methods

Chairman: Dr. Pascal Bondon and Dr. Sophie Brana (tentatives)

European electricity market integration and volatility spillovers **(Ref: *1658*)**
Evelyn Chanatásig-Niza, Aitor Ciarreta, Cristina Pizarro-Irizar and Ainhoa Zarraga

Monetary policy and bank risk taking: can macroprudential policy interact ? **(Ref: *3251*)**
Sophie Brana and Stéphanie Prat

High-Dimensional Mean-Variance Spanning Tests **(Ref: *4872*)**
David Ardia, Sébastien Laurent and Rosnel Sessinou

Income inequality and credit cycles: booster or anchor? **(Ref: *8043*)**
Alessandra Centinaio, Fausto Pacicco, Andrea Venegoni and Massimiliano Serati

Testing for Nonlinear Cointegration under Heteroskedasticity **(Ref: *8361*)**
Christoph Hanck and Till Massing

Transmission of oil price volatility to MENA stock markets: A comparative study between oil exporters and importers **(Ref: *8563*)**
Khalil Mhadhbi

(13:00-13:45) Session A.15: Functional Time Series Analysis and Application

---

Chairman: Dr. Till Massing and Dr. Adelaide Freitas

New algorithm for detecting weak changes in the mean in a class of CHARN models with application to welding electrical signals **(Ref: *1112*)**
Youssef Salman, Anis Hoayek and Mireille Batton-Hubert

Do Chinese foreign investment boost the Latin American countries? **(Ref: *1796*)**
Nadia Urriola Canchari and Abdul Samad Ibrahimi

Multiscale Hierarchical Forecasting of Urban Footfall **(Ref: *7501*)**
Tom Komar and Philip James

VIRTUAL SESSION ITISE - 2024

SESSIONS B

Monday, July 15, 2024

(9:00-10:15) Session B.1: Econometric Models and Forecasting (Part I)

Chairman: Dr. Yuvraj Sunecher and Dr. Boris Kozyrev (tentative)

Using the dichotomous variable to model structural change in time series: an application to international trade **(Ref: *435*)**
José Gerardo Covarrubias López and Xuedong Liu Sun

Forecast Combination and Interpretability Using Random Subspace **(Ref: *1209*)**
Boris Kozyrev

Comparison of Inferential Methods for a Novel CMP Model. **(Ref: *4508*)**
Yuvraj Sunecher and Naushad Mamode Khan

New Approaches in Demand Forecasting: Tobit Exponential Smoothing with Time Aggregation Constraints **(Ref: *5679*)**
Diego J. Pedregal and Juan R. Trapero

(10:15-11:30) Session B.2: Computational Intelligence methods for Time Series (Part I)

Chairman: Dr. Diego J. Pedregal and Dr. Eugene Pinsky (tentative)

Deep Leaning for crime forecasting of multiple regions, considering the spatial temporal correlations between regions **(Ref: *180*)**
Martín Solís and Luis-Alexander Calvo-Valverde

Enabling Diffusion Model for Conditioned Time Series Generation **(Ref: *783*)**
Frédéric Montet, Benjamin Pasquier, Beat Wolf and Jean Hennebert

A Machine Learning-based Approach to Analyze and Visualize Time-Series Sentencing Data **(Ref: *917*)**
Eugene Pinsky and Pironavakumar Kandaswamy

Analyzing Patterns of Injury in Occupational Hand Trauma Focusing on Press Machines: Registry Based Study and Machine Learning Analysis **(Ref: *1056*)**
Sarthak Pattnaik, Sagar Mandiya, Parita Danole, Ghazal Mashhadiagha, Ali Foroutan, Yousef Shafaei Khangdah, Khatereh Sazadehfar and Eugene Pinsky

Foreign Exchange forecasting models: LSTM and BiLSTM comparison **(Ref: *1328*)**
Fernando García, Francisco Guijarro, Javier Oliver and Rima Tamošiūnienė

(11:40-12:55) Session B.3: Time Series Analysis for Sustainable and Resilient Infrastructure Systems

---

Chairman: Dr. Maria Luisa Villani and Dr. Ebrahim Ehsanfar

Assessing Global Wildfire Dynamics and Climate Resilience: A Focus on European Regions Using the Fire Weather Index **(Ref: *3468*)**
Ayat-Allah Bouramdane

Reservoir neural network computing for time series forecasting in Aerospace and beyond. Potential applications to predictive maintenance. **(Ref: *4692*)**
Juan Manuel Rodríguez Riesgo and Juan Luis Cabrera Fernández

Missing data imputation for heat pump performance analytics **(Ref: *9025*)**
Sandhya Patidar, Chakron Dechkrut, Andrew Peacock, Abhinanda Roy and Yigit Duveroglu

Short-term real-time forecasting during turbulent times. A model for the Spanish GDP after the pandemic **(Ref: *5889*)**
Matias Pacce, Ana Gomez Loscos and Miguel Angel Gonzalez Simon

---

(16:00-17:45) Session B.4: Forecasting Financial Markets

Chairman: Dr. Anna Łyczkowska-Hańćkowiak

The Interconnected Generalized Autoregressive Conditional Heteroskedasticity Model **(Ref: *2687*)**
Javier Sánchez García, Salvador Cruz Rambaud and Paula Ortega Perals

Fuzzy portfolio selection based on 5-days recommendations – a case study **(Ref: *6415*)**
Anna Łyczkowska-Hańćkowiak and Aleksandra Wójcicka-Wójtowicz

The Intraday Dynamics Predictor: A TrioFlow Fusion of Convolutional Layers and Gated Recurrent Units for High-Frequency Price Movement Forecasting **(Ref: *6541*)**
Ilia Zaznov, Atta Badii, Julian Kunkel and Alfonso Dufour

Do Professional Forecasters Believe in Uncovered **(Ref: *7263*)**
Mengdi Song and Constantin Burgi

Measuring the efficiency of introducing businesses‘ digitalization elements over time in relation to their performance **(Ref: *9372*)**
Jarmila Horváthová and Martina Mokrišová

(17:50-19:50) Session B.5: New Advances in Time Series Analysis and Forecasting (Part I)

Chairman: Dr. Arzu Sardarli

New Studies on Birth, Death, Temperature Time Series and Their Correlation **(Ref: *669*)**
Arzu Sardarli

Identification and Forecasting of Bull and Bear Markets Using a Mixture of Markov switching Models and Rule-based Dating Algorithms **(Ref: *761*)**
John Maheu and Hamidreza Masoumi

Detecting short-notice cancellation in hotels with machine learing **(Ref: *1323*)**
Eleazar C-Sánchez and Agustín J. Sánchez-Medina

Assessing asymptotic tail independence: a simulation study **(Ref: *2078*)**
Marta Ferreira

The impact of the geometry and location of storm surge barriers on storm surge propagation: Great Egg Inlet, New Jersey, USA case study **(Ref: *2598*)**
Gregory Slusarczyk, Mary Cialone and Robert Hampson

Performance of an End-to-End Inventory Demand Forecasting Pipeline Using a Federated Data Ecosystem **(Ref: *3689*)**
Henrique Moura, Els de Vleeschauwer, Gerald Haesendonck, Ben De Meester, Lynn D'Eer, Tom De Schepper, Siegfried Mercelis and Erik Mannens

A System for Efficient Detection of Forest Fires Through Low Power Environmental Data Monitoring and AI **(Ref: *8587*)**
Ipek Uremek, Paul Leahy and Emanuel Popovici

Tuesday, July 16th, 2024

(9:45-10:45) Session B.6: Econometric Models and Forecasting (Part II)

Chairman: Dr. Paula Ortega Perals and Dr. José Gerardo Covarrubias López

Standard errors in nonlinear and time series models **(Ref: *232*)**
Guy Mélard

The role of healthcare quality and macrofinancial factors determining hospital mortality in Spain. Dynamic panel data evidence **(Ref: *7504*)**
Paula Ortega Perals, Salvador Cruz Rambaud and Javier Sánchez García

Modeling a set of variables with different attributes on a quantitative dependent variable: an application of dichotomous variables **(Ref: *9515*)**
José Gerardo Covarrubias López and Xuedong Liu Sun

(11:15-12:00) Session B.7: Time Series Analysis with Computational Intelligence in Energy Forecasting

Chairman: Dr. Jose Aureliano Martin Segura and Dr. David Esneyder Bello

A Simple Computational Approach
to Predict Long-Term Hourly Electric Consumption **(Ref: *1837*)**
Etinenne Meunier, Pierre Moreau, Tanvi Sharma and Eugene Pinsky

The impact of greenhouse gas emissions on global economic activity and health **(Ref: *2265*)**
Jose Aureliano Martin Segura and Cesar Pérez López

Analyzing the impact of Spanish electricity hourly prices over the consumption patterns **(Ref: *9730*)**
Eduardo Caro and Jesús Juan Ruiz

(12:00-13:00) Session B.8: Computational Intelligence methods for Time Series (Part II).

Chairman: Dr. Phillip Wenig (tentative)

Multi-objective Optimization for Selection of Clusterings Across Time **(Ref: *2604*)**
Sergej Korlakov, Gerhard Klassen, Luca T. Bauer and Stefan Conrad

Explaining when Deep Learning models are better for time series forecasting **(Ref: *6055*)**
Martín Solís and Luis-Alexander Calvo-Valverde

JET: Fast Estimation of Hierarchical Time Series Clustering **(Ref: *7729*)**
Phillip Wenig, Mathias Höfgen and Thorsten Papenbrock

Multi-Modal Model based on LSTM for Production Forecasting in Oil Wells with Rod Lift System **(Ref: *9771*)**
David Esneyder Bello Angulo and Elizabeth León Guzmán

(16:00-19:30) Session B.9: New Advances in Time Series Analysis and Forecasting (Part II)

Chairman: Dr. Evangelos Ioannidis and Dr. Santiago Zazo del Dedo

Impact of Solar Activity on Snow Cover Variation Over the Tibetan Plateau and Linkage to the Summer Precipitation in China **(Ref: *211*)**
Yan Song, Zhicai Li and Yu Gu

Time Series Forecasting on Electrochemical Curve in Battery Cycling using Multivariate LSTM and Attention Mechanism **(Ref: *392*)**
Florent Magaud, Xu Zhang and Arnaud Demortiere

A spatio-temporal study of tourist flow in Italian museums **(Ref: *744*)**
Leonardo Moreno, Gabriel Brida and Bibiana Lanzilotta

The stock market reaction to bank Mergers and Acquisitions (MAs): Do ESG factors and Climate Policy Uncertainty affect the wealth effects? **(Ref: *1210*)**
Ioannis Tampakoudis and Nikolaos Kiosses

Deep and Interpretable Probabilistic Forecasts **(Ref: *2016*)**
Philipp Baumann

Landscape Transformation & GLOF's Risk in Climate Changing Scenario in Hindukush Himalayan Region of Pakistan **(Ref: *3084*)**
Muhammad Qasim, Kainat Nafees and Mudassir Khan

Modelling Asymmetric and Time-Dependent Volatility of Bitcoin: An Alternative Approach **(Ref: *4195*)**
Abdulnasser Hatemi-J

Evaluation of Economic Interventions in economic blocks during an Economic and Sanitary Crisis **(Ref: *4767*)**
Carmin Montante and Clemente Hernandez-Rodriguez

Saving attitudes and behaviors in an emerging economy: Assessing the moderating effect of household size **(Ref: *5114*)**
Fatima Zahra Bendriouch and Harit Satt

Forecasting the realized volatility of stock markets: The roles of jumps and asymmetric spillovers **(Ref: *5569*)**
Abdelrazzaq Alrababa'A

Annual runoff forecasting through Bayesian causality **(Ref: *5731*)**
Santiago Zazo del Dedo, José Luis Molina González, Carmen Patino Alonso, Fernando Espejo Almodovar and Juan Carlos García Prieto

Ownership Identity and Stock Price Crash Risk: Evidence from the Mena Region **(Ref: *5948*)**
Harit Satt

A Python Module for Implementing Cointegration Tests with Multiple Endogenous Structural Breaks **(Ref: *6019*)**
Abdulnasser Hatemi-J and Alan Mustafa

A New Non-parametric Cross-spectrum Estimator **(Ref: *6283*)**
Evangelos Ioannidis

Energy prices impact on inflation per household category in Greece **(Ref: *7792*)**
Panagiotis Delis, Stavros Degiannakis and George Filis

Evaluation of clustering methods for the unsupervised classification of ground deformation time series **(Ref: *7542*)**
Jhonatan Rivera Rivera, Héctor Aguilera, Marta Béjar-Pizarro, Carolina Guardiola-Albert, Alejandra Staller, Pablo Ezquerro, Riccardo Palamà and Oriol Monserrat

Wednesday, July 17th, 2024

(9:00-10:15) Session B.10: New Advances in Time Series Analysis and Forecasting (Part III)

Chairman: Dr. Neelam Sinha

Evaluation of the University of Lagos Waste Generation Trend **(Ref: *6304*)**
Charles A. Mbama, Austin Otegbulu, Iain Beverland and Tara K. Beattie

Cellular Automata Framework for Dementia Classification using Explainable AI **(Ref: *6437*)**
Siva Manohar Reddy Kesu, Neelam Sinha and Hariharan Ramasangu

Catalyzing Supply Chain Evolution: A Comprehensive Examination of Artificial Intelligence Integration in Supply Chain Management **(Ref: *6533*)**
Sarthak Pattnaik, Natasya Liew, Ali Ozcan Kures, Kathleen Park and Eugene Pinsky

Decomposing the Sri Lanka yield curve using principal component analysis to examine the term structure of the Interest rate **(Ref: *8995*)**
Sanjeewa Dayarathne and Uthayashnger Thayasiwam

(10:15- 11:00) Session B.11: Data preprocessing methods in Time Series

Chairman: Dr. Charles Mbama

A Contextual Genetic algorithm for weighting Salient Crypto assets **(Ref: *1560*)**
Moulay Amine Jaidi

Short Term Forecasting of Nonstationary Time Series **(Ref: *3079*)**
Amir Aieb, Antonio Liotta, Alexander Jacob and Muhammad Azfar Yaqub

Extraction and forecasting of trends in cases of signal rank overestimation **(Ref: *3136*)**
Nina Golyandina and Pavel Dudnik

Abstracts book ITISE-2024

[0027] *Evaluating leading and coincident indicators of regional business cycles, a closer look into the Spanish scenario*

Marco Aurelio Pérez Navarro (Universidad Autónoma de Madrid) and Aránzazu De Juan Fernández (Universidad Autónoma de Madrid).

Using the Linear Dynamic Harmonic Regression algorithm developed by Bujosa, García-Ferrer and Young (2007), we obtain the Composite Leading Indicator (CLI) and Composite Coincident Indicator (CCI) for the Spanish Regions (Comunidades Autónomas) using the same variables used for the CLI and CCI for the Spanish economy as a whole (Garcia-Ferrer et al, 2014 and García-Ferrer et al, 2020) and we analyze if there exists a pattern in leading the Spanish recession periods. In order to build our CLI and CCI indicators we have used two different approaches. On the one hand, we have built them using the same list of variables as in Bujosa, García-Ferrer and de Juan (2013), who estimate a CLI and CCI indicator for the Spanish economy as a whole. On the other hand, we have estimated the regional CLI and CCI indicators considering different variables for each region. The reasoning behind this second approach is to take into consideration the different structure of each regional economy, which can help us to better forecast their behavior. For instance, when building the indicators of coastline regions with important ports, we take into account their merchandise traffic. We find that almost all Spanish regions with a coastline show a similar behavior. Specifically, they lead more than a year of the Spanish recession troughs. Therefore, the presence of international flows, either in the form of goods trade or tourism, may be playing an important role in the composition of our CLI and CCI indicators. Nevertheless, some heterogeneity can still be found among these coastal regions. For instance, the Canary Island's case can be noticeable since it is almost three years leading on average the peaks of the Spanish economic cycle but coincides on average with the troughs of the recessions. Therefore, this region seems to be very good at anticipating the recession periods, which could be, again, related with the importance of tourism in its economy, but it is not able to anticipate recoveries. Inland regions tend to anticipate peaks and throughs a few quarters later than the coastal ones, although we still find some exceptions. For instance, Castilla y León, is on average two and a half years ahead of the peaks of the recessions and more than one year ahead of the troughs of the recessions. Finally, we believe that our results could be useful for regional and national policymakers. A better understanding of the regional business cycles should lead them to a better timing and design of policies.

[0066] *Testing for asymmetric correlations between US sector returns and interest rate changes*

Francisco Jareno (Universidad de Castilla-La Mancha), María de La O González (Universidad de Castilla-La Mancha) and José Mª Almansa (Universidad de Castilla-La Mancha).

This paper analyses how US sector equity portfolios respond to fluctuations in nominal interest rates. By decomposing nominal interest rates into real interest rates and inflation expectations, following Jareño (2006 and 2008) and Jareño and Navarro (2010), this research investigates relevant non-linear and asymmetric interdependencies between eleven US sector portfolios and changes in interest rates using the NARDL methodology. Moreover, another contribution of the paper is to use US market inflation swaps as a proxy for inflation expectations, following Jimeno and Ibáñez (2018). Specifically, the study uses two models in three sample periods from April 2019 to January 2024, capturing the pre- and

post-COVID-19 periods. Our results confirm that the relevant linkages (in terms of cointegration, short- and long-run asymmetry, and persistence of lags) between US sector returns and interest rates are period and sector dependent, in line with previous literature. Finally, the study shows that some US sectors could serve as diversifying assets or even as safe havens in investment strategies under certain scenarios.

[0180] *Deep Leaning for crime forecasting of multiple regions, considering the spatial temporal correlations between regions*

Martín Solís (Tecnológico de Costa Rica) and Luis-Alexander Calvo-Valverde (Tecnológico de Costa Rica).

Crime forecasting has gained popularity in recent years, however, the majority of studies have been conducted in the United States, which may result in a bias towards areas with a substantial population. In this study, we generated different models capable of forecasting the number of crimes in 82 regions of a small country. These models include the spatial-temporal correlation between regions. The findings indicate that the architecture based on LSTM encoder-decoder achieved superior performance. The best model achieved the best performance in regions where crimes occurred more frequently, however, in more secure regions, the performance decayed.

[0211] *Impact of Solar Activity on Snow Cover Variation Over the Tibetan Plateau and Linkage to the Summer Precipitation in China*

Yan Song (China Meteorological Administration Training Centre), Zhicai Li (Shanxi Climate Center) and Yu Gu (University of California).

Solar activity is one of the main external forcing factors driving the Earth's climate system to change. The snow cover over the Tibetan Plateau is an important physical factor affecting the East Asian climate. At present, insufficient research on the connection between solar activity and snow cover over the Tibetan Plateau has been carried out. Using Solar Radio Flux (SRF), Solar Sunspot Number (SSN), and Total Solar Irradiance (TSI) data, this paper calculated the correlation coefficients with snow indices over the Tibetan Plateau, such as winter and spring snow depth (WSD/SSD) and snow day number (WSDN/SSDN). These snow indices are obtained from the daily gauge snow data in the Tibetan Plateau. Through correlation analyses, it is found that there are significant synchronous or lag correlations between snow indices and solar parameters on multi-time scales. In articular, the Spring Snow Day Number (SSDN) is of significant synchronous or lag correlation with SRF, SSN, and TSI on multi-time scales. It is further found that SSDN over the Tibetan Plateau has more stable positive correlations with SRF by using the 21-year running mean and cross spectrum analyses. Therefore, SSDN can be ascertained to be the most sensitive snow index to the solar activity compared with other snow indices. Moreover, its influence on summer precipitation of China is strongly regulated by solar activity. In high solar activity years (HSAY), the significant correlated area of summer precipitation in China to SSDN is located further north than that in low solar activity years (LSAY). Such impact by solar activity is also remarkable after excluding the impact of ENSO (i.e., El Niño–Southern Oscillation) events. These results provide support for the application of snow indices in summer rainfall prediction in China.

[0232] *Standard errors in nonlinear and time series models*

Guy Mélard (Université libre de Bruxelles, Faculty Solvay Brussels School of Economics and Management & ECARES).

Estimation of the parameters in non-linear statistical models is usually per- formed by non-linear least-squares or maximum likelihood using numerical optimization. Then, standard errors are often derived by using numerical derivatives. An empirical study of time series modelling through several statistical packages reveals that these standard errors are not very accurate, sometimes with only 2 or 3 correct digits. A further investigation on much simpler models determines the main reason for such a lack of accuracy. Some suggestions are given to developers of statistical software in order to improve their products. The present communication is inline with a sequence of works from different authors who warn users that are too confident in the relevance of their results.

[0298] *Dynamic Maps Powered by Machine Learning and Time Series Classification for Wildfire Risk Management*

Nicolò Perello (University of Genoa), Giorgio Meschi (CIMA Research Foundation), Andrea Trucchia (CIMA Research Foundation), Mirko D'Andrea (CIMA Research Foundation), Silvia Degli Esposti (CIMA Research Foundation) and Paolo Fiorucci (CIMA Research Foundation).

In recent decades wildfires are becoming more frequent and severe, posing significant risks to human lives, thus emphasizing the need to explore new technologies for wildfire risk management. Severe wildfires are linked to the development of meteorological conditions that make vegetation prone to ignition and fire spread. Identifying these conditions helps forecast high danger wildfires conditions.

In this study, the use of Machine Learning techniques in Time Series classification for the production of decision support tool is tested. Several maps are produced to provide support to Civil Protection systems, offering valuable information for the management of firefighting resources. The model has been tested as a case study in Italy.

A Time Series Forest classifier is used to assign a probability of wildfire occurrence and the probability of receiving aerial support requests for each administrative unit in Italy, at province level. The model has been trained using data from the past five years of wildfire events and aerial support requests. The last three days' time series of the three-hourly indexes of the Forest Fire Danger Rating system RISICO are used to identify trends related to wildfire occurrence and potential aerial support requests. The geographic location of each administrative unit is also considered to address spatial wildfire patterns and regional-level wildfire risk management policies in Italy.

Although the model has been implemented only for 2023, it has shown good performance during the fire season, highlighting the potential of using Machine Learning techniques in Time Series classification to provide operational decision support tools.

[0392] *Time Series Forecasting on Electrochemical Curve in Battery Cycling using Multivariate LSTM and Attention Mechanism*

Florent Magaud (CNRS), Xu Zhang (UPJV) and Arnaud Demortiere (CNRS).

Long Short-Term Memory (LSTM) is well known model for forecasting multivariate time series. It is already used for the estimation of the state of charge [1], it is also used for the prediction of the remaining useful life [2,3] or to predict the capacity fading [4]. Here, we use it to predict the evolution of the cycling curve of battery to reduce the time, the cost, and the energy of the concept testing for the battery manufacturer.

We work on data from TIAMAT [5] a French start up working on sodium-ion battery. These data consist in the evolution of the voltage during the cycling (in charge and discharge), the current applied on the battery and the number of cycles. The voltage and current exhibit clear seasonality; however, as the number of cycles increases, the duration of this seasonality decreases, indicating non-stationary data. Extensive data preprocessing is required to adjust the time values. Data acquisition lacked a constant time step, necessitating interpolation to standardize the time intervals. Once the data preprocessing is done, we can train our model which consist of 2 layers of LSTM with a multi-head attention (and a dropout layer) between them and a fully connected layer for the output. To identify the optimal hyperparameter combination, we conducted extensive testing across various parameters. We evaluated grids of parameters, including input size, target size, batching overlap, and diverse input data. Attention mechanisms are widely used to improve performances of time series prediction [6], so we add it to our model to improve the information processing in our model.

We achieved highly accurate results, with an accuracy of 2x10-5 on the training data set and 5x10-5 on the validation data set, utilizing the Mean Squared Error as the loss function. Our model effectively comprehends and accurately interprets complex datasets, capturing the curve's deformation and predicting the reduced time required for each cycle. It is working well for long term predictions there is no divergence of the model the farther we predict. For now, the limit of our model is the instability on precise part of a cycle, which decrease the quality of the predictions and prevent us to have a precise understanding of the result.

References:

[1] Junxiong Chen, Yu Zhang, Ji Wu, Weisong Cheng, Qiao Zhu, “SOC estimation for lithium-ion battery using the LSTM-RNN with extended input and constrained output”, Energy, Volume 262, Part A, 2023, 125375, ISSN 0360-5442, https://doi.org/10.1016/j.energy.2022.125375.

[2] Y. Zhang, R. Xiong, H. He and Z. Liu, "A LSTM-RNN method for the lithuim-ion battery remaining useful life prediction," 2017 Prognostics and System Health Management Conference (PHM-Harbin), Harbin, China, 2017, pp. 1-4, doi: 10.1109/PHM.2017.8079316.

[3] K. Park, Y. Choi, W. J. Choi, H. -Y. Ryu and H. Kim, "LSTM-Based Battery Remaining Useful Life Prediction With Multi-Channel Charging Profiles," in IEEE Access, vol. 8, pp. 20786-20798, 2020, doi: 10.1109/ACCESS.2020.2968939.

[4] A. Gupta, M. Sheikh, Y. Tripathy and W. D. Widanage, "Transfer learning LSTM model for battery useful capacity fade prediction," 2021 24th International Conference on Mechatronics Technology (ICMT), Singapore, 2021, pp. 1-6, doi: 10.1109/ICMT53429.2021.9687230.

[5] http://www.tiamat-energy.com

[6] Youru Li, Zhenfeng Zhu, Deqiang Kong, Hua Han, Yao Zhao, "EA-LSTM: Evolutionary attention-based LSTM for time series prediction", Knowledge-Based Systems, Volume 181, 2019, 104785, ISSN 0950-7051, https://doi.org/10.1016/j.knosys.2019.05.028.

[0435] *Using the dichotomous variable to model structural change in time series: an application to international trade*

José Gerardo Covarrubias López (Faculty of Higher Studies Aragon, National Autonomous University of Mexico) and Xuedong Liu Sun (Faculty of Higher Studies Aragon, National Autonomous University of Mexico).

This research aims to elucidate the methodology employed in econometric estimation utilizing dichotomous variables. These variables serve a dual purpose: firstly, they denote an attribute designed to discern structural changes within a linear relationship, and secondly, they quantify the impact and statistical significance of one or more quantitative independent variables on a dependent variable in the presence of such structural changes. The core of the document not only encompasses the presentation of the methodology but also includes an applied analysis of Mexico's foreign trade. This analysis delves into estimating the impact and statistical significance of exports on economic growth across different periods, reflecting significant structural changes in the development of the Mexican economy.

[0669] *New Studies on Birth, Death, Temperature Time Series and Their Correlation*

Arzu Sardarli (First Nations University of Canada).

Human birth seasonality was first observed in the 17th century in France. Although the seasonality of birth remains an object of intensive study by various disciplines for centuries, a consensus about the determinants of seasonality has not been achieved by researchers. Some authors insist on the dominance of social demographic characteristics, while others emphasize biological and environmental factors. An extensive review of the available literature shows that researchers usually analyze the monthly birth time series, perhaps because of the unavailability of daily birth data. At the Annual Meeting of the Canadian Population Society, I am going to present the preliminary results of my studies of monthly and daily birth, death, and temperature time series in eight Canadian provinces (New Brunswick, Nova Scotia, Quebec, Ontario, Manitoba, Saskatchewan, Alberta and British Columbia) for 2011 – 2019. I used a novel approach to collecting temperature data to improve the accuracy of measurements. Dissimilar to previous studies, where the researchers used temperature values averaged over the region, I determined the coordinates of centres of populations of provinces and used the exact temperature measurements conducted at weather stations close to these centres. The Fourier analysis of monthly birth time series indicated annual cycles, which is in good agreement with results previously reported in the literature. However, the Fourier analysis of daily birth time series allowed me to discover the new cycles, which, to my knowledge, had never been reported in previous works. I also studied the correlations between birth, death and temperature time series. Although the correlation between birth, death and time series is not strong, the signs of cross-correlation coefficients were the same for all eight observed provinces. The birth and death data were obtained from the Regina and Saskatoon Research Data Centres of the Canadian Research Data Centre Network of Statistics Canada. The historical temperature data was collected from publicly available resources of the Canadian Government. The project was supported by the First Nations University of Canada and the University of Regina within the President's Research Seed Grant & Social Sciences and Humanities Research Council (SSHRC) Explore Grant in 2021-2023.

[0744] *A spatio-temporal study of tourist flow in Italian museums*

Leonardo Moreno (University of the Republic), Gabriel Brida (University of the Republic) and Bibiana Lanzilotta (University of the Republic).

The importance of tourism as a relevant engine of a country's economy is well known. In particular, cultural tourism contributes to the preservation of cultural heritage, promoting the exchange and understanding of different cultures. It also stimulates national art, giving a greater visualization of local art.

The offer of museums occupies a substantial place in attracting this type of tourists. Therefore, being able to predict the flow of tourists about their visits to the different locations of the museums is of relevance for the generation of policies for the location of new museums, or modifications in the transportation system of the region.

To study the dynamics of tourist flow, the use of spatio-temporal models is appropriate. This work uses a methodology based on copula functions. Two copulas are considered, and each of them contains temporal and spatial information respectively. Unifying both and using the Wasserstein distance, a dissimilarity measure is constructed. This allows us the construction of clusters of museums in different regions of Italy. The characterization of these clusters is a relevant input for decision-making in the implementation of policies that preserve and stimulate cultural tourism.

[0761] *Identification and Forecasting of Bull and Bear Markets Using a Mixture of Markov switching Models and Rule-based Dating Algorithms*

John Maheu (McMaster University) and Hamidreza Masoumi (McMaster University).

The acceptance of different stock market patterns is widespread among academics and practitioners. Fluctuations in the stock market often occur before changes in the overall economy and are influenced by varying expectations of future cash flows and risk premiums,see Haase & Neuenkirch (2023) and Gordon & St-Amour (2004).In bull markets, prices generally rise often with minimal fluctuations, while bear markets are characterized by price declines and mostly increased volatility. Therefore, it is important for investors and corporate decision-makers to anticipate shifts in market patterns, especially bear trends.

The inherent nature of the equity market poses a clear challenge as its state remains hidden or latent. Within the current body of literature, two prominent approaches have emerged for the identification and prediction of past and future market states: dating algorithms and Markov switching models. These approaches diverge in their treatment of state characteristics employed for identification and forecasting. Dating algorithms or rule-based methods focus mainly on examining the price changes linked to different states, employing algorithms that follow a set of classification rules (evident in studies such as Gonzalez et al.(2005) ;Lunde & Timmermann (2004) ; and Pagan & Sossounov (2003)). In contrast, Markov switching models adopt a more inclusive approach by considering the whole distribution of returns.

The literature contains a debate regarding the optimal method for dating business cycle regimes. Harding & Pagan (2003) advocate for the use of simple classification rules, whereas Hamilton et al. (2003) propose the application of Markov switching models. One of the primary limitations of dating algorithms, as evidenced in the literature by (Haase & Neuenkirch(2023)), is that they can only identify a turning point after several observations following its occurrence. Moreover, ex-post dating methods cannot be employed for statistical inference on returns or investment decisions requiring a more comprehensive understanding of the return distribution, such as assessing changing risks (Maheu et al. (2012)). In order to facilitate effective risk management and informed investment decisions, it is necessary to have a probability model for returns that accounts for the changing distribution of returns over time.

According to Kole & Van Dijk (2017), rule-based methods are preferable for in-sample identification of the market state, while Markov switching models are more suitable for out-of-sample forecasting. In the in-sample period, the mean return of the market index is the primary factor, precisely captured by rule-based methods. However, Markov switching models incorporate both the mean and variance to infer the state, resulting in superior forecasts and significantly better out-of-sample performance compared to rule-based methods. They conclude that considering the variance is crucial for forecasting the market state, particularly in the out-of-sample context. Furthermore, incorporating a greater number of states within Markov switching models allows for the inclusion of intra-regime dynamics in the market.This enhanced modeling capability can lead to improved forecasting accuracy and enable the development of more effective market timing strategies. (Maheu et al. (2012)).

The main drawback of Markov switching models is that the identification process is mainly derived from volatility dynamics. As a result, these models may classify volatile periods as bear markets, even if prices are rising, and classify calm periods as bull markets, even if prices are falling as documented by Kole & Van Dijk (2017).

We introduce a fully specified model called MSLT-4 with four states, which integrates the Lunde and Timmerman (LT) dating algorithm and Markov switching models into a comprehensive framework for forecasting. This model combines the advantages of considering both the mean and variance in forecasting. Our design intentionally incorporates the accurate in-sample sorting provided by the LT algorithm to improve the identification process of Markov switching models.

Our proposed model makes a notable contribution by addressing a limitation of Markov switching models, specifically their heavy reliance on volatility as the primary driver for capturing trends. Our model exploits and expands the distribution-based approach of Markov switching models, accurately classifying high-volatility upward price trends as bull

states,thus effectively resolving this drawback. This enhanced identification capability leads to a significant improvement in the model's out-of-sample market timing strategy.

Another important contribution of this model is the introduction of two additional states specifically designed to capture abrupt price spikes or sudden market crashes. By incorporating these states, our model provides a more comprehensive and nuanced representation of market dynamics, offering a richer understanding of the underlying behavior and fluctuations in prices.

Our paper is, to the best of our knowledge, the first one to mix two main approaches in bull and bear market identification in one fully specified model with four states. The confirmed consensus in the literature suggests that predictability in returns is more pronounced during periods of unfavorable market conditions (Haase & Neuenkirch (2023)). The proposed MSLT-4 model empowers investors to recognize a market state characterized by a declining price trend and relatively low volatility. This enables investors to potentially avoid participating in the market during periods with a higher likelihood of incurring losses, while also avoiding missing out on potential market growth.

We have analyzed four market-timing investment approaches using MSLT-4, and we demonstrate that employing simple timing rules for entering and exiting the market leads to notable improvements in all aspects such as Sortino ratios, Sharpe ratios, utility-derived performance fees, and mean returns, compared to buy-and-hold strategy. This finding is intriguing since almost all existing literature on market timing models fails to enhance both utility and risk-adjusted returns, as well as mean returns simultaneously. What sets MSLT-4 apart is its ability to significantly outperform the market in terms of mean returns when utilized in market-timing strategies. This is in contrast to the results of previous models like those presented in Haase & Neuenkirch (2023) and Kole & Van Dijk (2017) that underperform the market in this regard.

Notably, when compared to the benchmarks LT-2, MS-2 and MSMS-4, our proposed MSLT-4 consistently achieves superior performance with the same strategy. These results hold up across various timing strategies and can be attributed to our model's accurate identification and forecasting of turning points in the market. An interesting aspect is that these market-timing strategies can be effectively applied in real time and would have been accessible to investors using our model for forecasts.

[0776] *Optimal Forecast Combination for Japanese Tourism Demand*

Saeed Heravi (Cardiff Business School), Bo Guan (Cardiff Business school), Yongmi Fang (South China Normal University), Emmanuel Silva (Glasgow Caledonian University) and Hossein Hassani (Tehran University).

Japan's tourism industry is becoming a major economic engine since the introduction of the Visit Japan campaign in 2003. However, academic research on forecasting Japanese tourism demand is sparse. Accordingly, this study proposes a combination forecast for Japanese tourism demand at an aggregated and disaggregated level based on the purpose of travel (tourism, business and other). We compare the forecasting performance of Ensemble Empirical Mode Decomposition (EEMD) methods, Support Vector Machine, Neural Network and ARIMA models. We find that combination forecasts using EEMD results in significantly better forecasts in most cases. Therefore, we conclude that forecast combination methods related to EEMD are better than individual forecasts (in most cases) when they are decomposed to high and low frequencies data.

[0783] *Enabling Diffusion Model for Conditioned Time Series Generation*

Frédéric Montet (iCoSys Institute, University of Applied Sciences Western Switzerland, Fribourg), Benjamin Pasquier (iCoSys Institute, University of Applied Sciences Western Switzerland, Fribourg), Beat Wolf (iCoSys Institute, University of Applied Sciences Western Switzerland, Fribourg) and Jean Hennebert (iCoSys Institute, University of Applied Sciences Western Switzerland, Fribourg).

Synthetic time series generation is an emerging field of study in the broad spectrum of data science, addressing critical needs in diverse fields such as finance, meteorology and healthcare. In recent years, diffusion methods have shown impressive results for image synthesis thanks to models such as Stable Diffusion and DALL·E, defining the new state-of-the-art methods. In time series generation, their potential exists but remains largely unexplored. In this work, we demonstrate the applicability and suitability of diffusion methods for time series generation on several datasets and a rigorous evaluation procedure. Our proposal, inspired from an existing diffusion model, obtain better performance than a reference model based on generative adversarial networks (GANs). We also propose a modification of the model to allow guiding the generation with respect to conditioning variables. This conditioned generation is successfully demonstrated on meteorological data.

[0828] *Automatic selection of methods to perform combination of predictions in monthly series*

Carlos García-Aroca (Universidad Miguel Hernández de Elche), Asuncion Mayoral (Universidad Miguel Hernandez de Elche), Javier Morales (Universidad Miguel Hernández de Elche) and José Vicente Segura (Universidad Miguel Hernández de Elche).

In the era of optimal use of information, it is logical to expect the availability of optimal tools or procedures to efficiently produce forecasts that anticipate demand in the industry and, in general, the behavior of time series with a reasonable amount of historical data. Efforts in the last years have been made in developing algorithms that combine forecasts so as to improve accuracy compared to individual forecasting methods [1,2], but the choice of weights has been crucial in deciding how good the combined estimation method is [3,4].

Combination schemes have evolved from simple methods without estimation to sophisticated approaches involving time-varying weights, nonlinear combinations, correlations among components, or cross-learning. These schemes encompass combining point forecasts and probabilistic forecasts [5,6]. Ensemble methods, expert aggregation, and forecast combination have proven highly effective in time series forecasting across various real-world domains, including industrial production, sales and demand, oil prices, energy, pollution levels, health or car sales [7-13].

Our research issue concerns developing an efficient method that automatically: 1) combine available or chosen methods, 2) assess the optimal weights for the estimates provided by these methods, as well as the number of methods to combine, 3) improves efficiency (in terms of time) and accuracy (in terms of prediction error) over the existing methods based on combining forecasts, 4) so requiring as less as possible decisions of the user to get an efficient forecast.

Our objectives are partially achieved. To date we have developed a new algorithm that: 1) combines relevant forecasting methods available in R, 2) sorts these methods in terms of smaller global error, with a global error obtained from a principal component analysis on the usual error measures, 3) assesses weights for the estimates provided by the individual methods, in several ways, 4) gets the combined forecast from different amounts of the methods given in terms of percentiles, and 5) provides the common error measures evaluated in the training, validation and test samples considered for the analysis. This work has allowed us to prove that our method performs better than other methods of combined estimation, in a large number of time series with different characteristics [14].

The aim of this paper is to further automate to avoid the user having to make decisions on the number of methods to combine. By means of statistical tests we select the final number of configurations of the prediction methods that should participate in the combination of predictions. This allows us to go further in obtaining more accurate forecasts.

The implementation of the algorithm is done in R language in order to have direct access to time series prediction methods that have an associated function in that language.

References

[1] Thomson, M.E.; Pollock, A.C.; Önkal, D. & Gönül, M.S. (2019). Combining forecasts: Performance and coherence. International Journal of Forecasting, Volume 35, Issue 2, 2019, Pages 474-484. 10.1016/j.ijforecast.2018.10.006. [2] Lichtendahl, K.C. & Winkler R.L. (2020). Why do some combinations perform better than others? International Journal of Forecasting, Volume 36, Issue 1, 2020, Pages 142-149. 10.1016/j.ijforecast.2019.03.027. [3] Timmermann, A (2006). Forecast Combinations. In: Handbook of Economic Forecasting. Ed. by G Elliott, CWJ Granger, and A Timmermann. Vol. 1. Elsevier. Chap. 4, pp.135–196. 10.1016/S1574-0706(05)01004-9. [4] S. Cang, H. Yu. (2014). A combination selection algorithm on forecasting. European Journal of Operational Research, 234 (1), pp. 127-139, 10.1016/j.ejor.2013.08.045 [5] Li, X., Y. Kang, Y., Li, F. (2020). Forecasting with time series imaging. Expert Systems with Applications, 160 (113680) 10.1016/j.eswa.2020.113680 [6] Montero-Manso, P., Athanasopoulos, G., Hyndman, R.J., Talagala, T.S. (2020). FFORMA: Feature-based forecast model averaging. International Journal of Forecasting, 36 (1), pp. 86-92, 10.1016/j.ijforecast.2019.02.011 [7] Stock and Watson (2006). Handbook of Economic Forecasting, Volume 1, Elsevier. G. Elliott, C.W.J. Granger, A. Timmermann (Eds.). [8] Gurnani M., Korke Y., Shah P., Udmale S., Sambhe V., & Bhirud S. (2017). Forecasting of sales by using fusion of machine learning techniques. International Conference on Data Management, Analytics and Innovation (ICDMAI). [9] Manescu, C. & Van Robays, I. (2014). Forecasting the brent oil price addressing time-variation in forecast performance. Working paper series NO 1735 / September 2014, European Central Bank. [10] Gaillard, P., Goude, Y., Nedellec, R. (2016). Additive models and robust aggregation for gefcom2014 probabilistic electric load and electricity price forecasting. International Journal of forecasting, 32 (3, pp. 1038-1050, 10.1016/j.ijforecast.2015.12.001 [11] Auder, B., Bobbia, M., Poggi, J.-M., Portier, B. (2016). Sequential aggregation of heterogeneous experts for pm10 forecasting. Atmospheric Pollution Research, 7 (6), pp. 1101-1109, 10.1016/j.apr.2016.06.013 [12] Perone G. (2022). Comparison of ARIMA, ETS, NNAR, TBATS and hybrid models to forecast the second wave of COVID-19 hospitalizations in Italy. The European Journal of Health Economics Aug;23(6): 917-940. 10.1007/s10198-021-01347-4. [13] Fortsch, S.M., Choi, J.H. & Khapalova, E.A. (2021). Competition can help predict sales. Journal of Forecasting. First published: 19 August 2021. Wiley Online Library. 10.1002/for.2818. [14] García-

Aroca, C., Martínez-Mayoral, M.A., Morales-Socuéllamos, J., Segura-Heras, J.V. (2024). An algorithm for automatic selection and combination of forecast models. Expert Systems With Applications, Vol. 237-C. DOI https://doi.org/10.1016/j.eswa.2023.121636

[0917] *A Machine Learning-based Approach to Analyze and Visualize Time-Series Sentencing Data*

Eugene Pinsky (Boston University metropolitan College) and Pironavakumar Kandaswamy (Boston University Metropolitan College).

Analyzing time-series sentencing data presents many challenges. The data has many dimensions and changes with time. This makes it difficult to identify patterns and discuss their similarities over time. This work proposes a Machine Learning approach to associate patterns with clusters. This allows a representation of sentencing data regarding trajectories in the appropriate (time, cluster) space. We propose to use the Hamming distance of trajectories to measure the similarity of sentencing data across districts. For any offense, we can define the average Hamming distance that has a simple interpretation as the average period when sentencing patterns are different. We introduce simple statistical measures on trajectories to show similarities and changes in sentencing behavior over time. We illustrate our approach by analyzing sentencing data for narcotics and retail theft.

[0946] *A Time Series Dashboard for Decision Makers in Higher Education Institutions*

Gabriel David Pinto (Azrieli - College of Engineering Jerusalem) and Ron Weitzman (Azrieli - College of Engineering Jerusalem).

During the past two decades, there were significant changes in the number of engineering students in institutions of higher education in Israel. The total number of engineering students grew each year up to 2015, and from 2015 it follows a stationary pattern. Accordingly, the number of higher education institutions that provide engineering programs was also stabilized. During the demand growth period, a few leading institutions that did not teach engineering have opened new tracks in the field of engineering. For example, Bar-Ilan University responded to the growing demand in students studying engineering and opened a bachelor's degree program in industrial engineering. The growing variety of engineering programs caused an increase in competition between the institutions. Based on this evolving institutional competition the motivation for this work arises. Our first goal is to assist decision makers in academic institutions, by creating an individualized Business Intelligence (BI) portal (AcaData) with time series data analyses and visualizations to serve as a decision support tool. The second goal of this work is to allow for the discovery of insights based on indicators which may predict success or failures of academic programs.

The time series data includes financial, operational and throughput data. Furthermore, presented in the dashboards are calculated parameters that combine and show the relationship between the expenses and throughput, as well as the distribution of institutional resources and their consequential results. The dashboard may allow academic institutions to compare themselves to others and answer questions such as: 1. Does our current resources allocation produce a good result regarding our outputs (research, # of students,)? 2. Do the expenses of the salary for the academic staff lead to an increase in their academic results (teaching and research)? 3. Did the advertising and marketing expenses lead to an increase in the number of students the following year and are therefore justified? 4. What are the strengths of our institutions? As well as what are the weaknesses that need to be improved? 5. What predicts the success or failure of academic institutions in the field of engineering?

AcaData (https://aca-data.com) is a specialized website developed to efficiently display the dashboards and to effectively present the data collected and calculated from the institutions of higher education in Israel. The displayed dashboards are divided into the following two categories. The public accessible data center, which presents the information on all institutions with engineering degrees, and an individualized portal developed for each institution for their eyes only. The institutional dashboards are customized areas that contain data and indications that are of interest, as well as financial data and the impact of the institution's expenses on their throughput results.

[1056] *Analyzing Patterns of Injury in Occupational Hand Trauma Focusing on Press Machines: Registry Based Study and Machine Learning Analysis*

Sarthak Pattnaik (Boston University metropolitan College), Sagar Mandiya (Boston University Metropolitan College), Parita Danole (Boston University Metropolitan College), Ghazal Mashhadiagha (Boston University Metropolitan College), Ali Foroutan (Shiraz University of Medical Sciences), Yousef Shafaei Khangdah (Shiraz University of Medical Sciences), Khatereh Sazadehfar (Ardabil University of Medical Sciences) and Eugene Pinsky (Boston University metropolitan College).

Objectives: The aim of the project is to analyze the data of patients who have been admitted to the emergency room due to severe hand and palm injuries.

Methods: We have used data visualization and statistical analysis to observe trends in various factors pertaining to the patients, such as place of injury, machine-causing injuries, date and time of the injury, amputation, fracture, etiology, distribution of the injured hand, etc.

Results: There is a significant difference between age and gender groups across various injuries. Most of the injuries in the dataset are occupational injuries caused by press machines. Most injuries take place in the later half of the week on Thursdays and Saturdays.

Conclusion: There were 1,676 patients who were reported at the medical emergency center. Out of these, only a handful of them have undergone extremely painstaking injuries where there was uncontrolled bleeding and Hemi-amputation. We can also surmise the same by looking at the data that provides the summary of the number of fingers injured. Most patients have either 1 or 2 fingers injured. Very few patients had more than two fingers injured

---

[1112] *New algorithm for detecting weak changes in the mean in a class of CHARN models with application to welding electrical signals*

Youssef Salman (Mines Saint-Etienne, Univ Clermont Auvergne, CNRS, UMR 6158 LIMOS, Institut Henri Fayol, F - 42023 Saint-Etienne France), Anis Hoayek (Mines Saint-Etienne, Univ Clermont Auvergne, CNRS, UMR 6158 LIMOS, Institut Henri Fayol, F - 42023 Saint-Etienne France) and Mireille Batton-Hubert (Mines Saint-Etienne, Univ Clermont Auvergne, CNRS, UMR 6158 LIMOS, Institut Henri Fayol, F - 42023 Saint-Etienne France).

In this paper, we propose a new automatic algorithm for detecting weak changes in the mean of a class of piece-wise CHARN models. Through a simulation experiment, we demonstrate its efficacy and precision in detecting weak changes in the mean and accurately estimating their locations. Furthermore, we illustrate the robust performance of our algorithm through its application to welding electrical signals (WES).

[1209] *Forecast Combination and Interpretability Using Random Subspace*

Boris Kozyrev (Halle Institue for Economic Research (IWH)).

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

Nowcasting as well as short-term forecasting current economic activity is crucial for decision-makers such as firms and governments. Over the last couple of decades, the amount of data used for (improving) macroeconomic forecasting has been drastically increasing. This means that one may end up with hundreds of different forecasts of a given macroeconomic indicator, for example, GDP. These individual predictions may be derived from different methods or data sources. A natural question arises: how to combine and/or select numerous forecasts to produce a final prediction?

Another strand of research has lately been dealing with unraveling the black box critique of machine learning (ML) models. Oftentimes, it is almost impossible to study the relationship between the dependent and independent variables within the ML framework or to explain why the estimations are as such. In particular, the "interpretability" aspect is crucial for decision-makers in economics, when some sort of a structure is desired for a couple of reasons.

Not only do economists and policy makers care about precision, but also they are keen on compre- hending the driving forces behind a given phenomenon. Forecasting models are supposed to motivate some economic or public policies instead of giving out an intractable – yet, often precise – estimation (Burgess et al., 2013). Let alone even the "best" single model can from time to time produce imprecise estimations. In theory, understanding a model's mechanics allows for the identification and explanation of its odd behaviors, a task often unfeasible with some ML models.

The noticeable increase in using predictive ML in macroeconomics drove researchers to try to add some sort of "interpretability" and to introduce a universal method, which explains the relative importance of any given feature. The Shapley value decomposition (SVD) is one of the most popular tools, allowing for decomposition into the contributions of individual variables. Despite its popularity, there are some disadvantages. For example, the SVD can handle high-dimensional data poorly and demands a lot of computing time. Computational burdens force researchers to restrict their attention to a smaller number of indicators or shift to specific types of models, for which shortcuts have been developed.

In this article, I propose a framework that connects two seemingly independent fields, i.e., forecast combination and model interpretability. First, I introduce a forecast combination method based on Random subspace regression (RS), proposed by Boot and Nibbering (2019), which makes it possible to work within a linear framework even if the number of covariates (in this case, individual forecasts) is rather huge. This aggregation scheme produces a new set of forecasts, based on the initial values. RS helps smooth the distribution of predictions if there are severe outliers. The forecasting performance of this aggregation method is compared to common benchmarks. Additionally, I explore a synergy between RS and the SVD. Due to the fact that the SVD is easily computed in a standard linear framework, one can decompose the prediction obtained by RS into each forecast contribution using well-known shortcuts. Furthermore, this union of the methods allows researchers to handle even high-dimensional data. Thus, under this framework, it is possible to circumvent the challenges of using many variables while dealing with the SVD. I make sure that the main SVD advantage – the difference between the final prediction and the historical average prediction is fairly distributed among the covariates – is fulfilled.

I argue that RS used for forecast aggregation is a powerful tool, whose forecasting performance is superior to both AR(1) and taking a median of a given set of forecasts. The number of individual forecasts to include in one draw plays a significant role. My empirical results indicate that the optimal number of the latter may lie between 10 and 20. Then, the contribution of each individual forecast (which may exceed 100) to the total prediction is observed and analyzed over time.

[1210] *The stock market reaction to bank Mergers and Acquisitions (M&As): Do ESG factors and Climate Policy Uncertainty affect the wealth effects?*

Ioannis Tampakoudis (University of Macedonia) and Nikolaos Kiosses (University of Macedonia).

This study investigates the stock market reaction to Mergers and Acquisitions (M&As) in the U.S. banking sector using alternative asset pricing models. The aim is to disentangle the valuation effects of M&A announcements by separately estimating Cumulative Abnormal Returns (CARs) for acquirers, targets, and their close peers. Using a sample of M&A deals in the U.S. banking sector, we first conduct an event-study analysis to evaluate the market reaction using alternative asset pricing models over various event-windows. Specifically, we utilize time-series data of stock market returns for acquirers, targets, and their peers, applying different asset pricing models to measure the wealth effects following M&A announcements. The statistical significance of CARs is assessed using both parametric and non-parametric tests. In the second stage, each deal is matched with data on Environmental, Social, and Governance (ESG) scores and the level of Climate Policy Uncertainty to analyze the effect of these factors on the market reaction. The results suggest that bank M&As convey information that leads to abnormal returns for both bidders and targets. Specifically, during the three-day event window surrounding the deal announcement, acquiring banks exhibited a negative reaction, experiencing a loss of approximately -1.4%. In contrast, target firms demonstrated positive and statistically significant abnormal returns, indicating value creation for their shareholders. However, when considering the peers of acquirers and targets, there appears to be an unclear pattern in the market reaction over the entire sample period. The results from cross-sectional analyses indicate that various factors, including bank-specific factors, deal-specific factors, ESG issues, and the level of Climate Policy Uncertainty, influence the valuation effects of M&As in the banking sector.

[1323] *Detecting short-notice cancellation in hotels with machine learing*

Eleazar C-Sánchez (Instituto Universitario de Ciencias y Tecnologías Cibernéticas (IUCTC), University of Las Palmas de Gran Canaria, Spain) and Agustín J. Sánchez-Medina (Instituto Universitario de Ciencias y Tecnologías Cibernéticas (IUCTC), University of Las Palmas de Gran Canaria, Spain).

Cancellations play a critical role in the lodging industry. Considering the time horizon, cancel-lations placed close to check-in have a significant impact on hoteliers, who must respond promptly for effective management. In recent years, the introduction of Personal Name Records (PNR) has brought innovative approaches to this domain, but short-notice cancellation prediction is still underdeveloped. Using real PNR data from a four-star hotel, this research aims to combine un-supervised and supervised machine learning methods to forecast cancellations placed close to the entry day, slightly improving the performance of individual techniques.

[1328] *Foreign Exchange forecasting models: LSTM and BiLSTM comparison*

Fernando García (Universitat Politècnica de València), Francisco Guijarro (Universitat Politècnica de València), Javier Oliver (Universitat Politècnica de València) and Rima Tamošiūnienė (Vilnius Gediminas Technical University).

Knowledge of currency prices and their evolution is important for firms and investors, both for hedging exchange rate risk and for investment and trading. Classical time series models, such as ARIMA models, were relevant until the advent of neural networks. Several studies show that recurrent neural networks, such as Long Short-Term Memory (LSTM), significantly outperform ARIMA models in predicting short-term stock prices. This paper presents a comparative analysis between the LSTM model and the BiLSTM models. There is evidence for an improvement of the hybrid model in predicting time series such as stock prices. In this case, we analyse whether this efficiency is consistent in predicting different currencies as well as the bitcoin futures contract

---

[1448] *Application of the optimised Pulse Width Modulation (PWM) based encoding-decoding algorithm for forecasting with Spiking Neural Networks (SNN)*

Sergio Lucas (University of the Basque Country (UPV/EHU)) and Eva Portillo Pérez (University of the Basque Country (UPV/EHU)).

Spiking Neural Networks (SNN) are recognised for processing spatiotemporal information with ultra-low power consumption. However, applying a non-efficient encoding-decoding algorithm can counter the efficiency advantages of the SNN. In this sense, this paper presents one-step ahead forecasting centered on the application of an optimised encoding-decoding algorithm based on Pulse Width Modulation (PWM) for SNN. The validation is carried out with sine-wave, 3 UCI and 1 available real-world datasets. The results show the practical disappearance of the computational and energy costs associated with the encoding and decoding phases (less than 2% of the total costs) and very satisfactory forecasting results (MAE lower than 0.0357) for any dataset.

[1520] *Towards Resolving the Ambiguity in Low-Field All-Optical Magnetic Field Sensing With High NV-Density Diamonds*

Ludwig Horsthemke (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Jens Pogorzelski (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Dennis Stiegekötter (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Frederik Hoffmann (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Ann-Sophie Rösner (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Markus Gregor (Department of Engineering Physics, FH Münster—University of Applied Sciences) and Peter Glösekötter (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences).

In all-optical magnetic field sensing, using nitrogen-vacancy center rich diamonds, an ambiguity in the range of 0-8 mT can be ob- served. We propose a way to resolve this ambiguity using the magnetic- field-dependent fluorescence lifetime. We therefore record the frequency response of the fluorescence upon modulation of the excitation intensity in a frequency range of 1–100 MHz. The magnetic-field-dependent decay dynamics lead to different response characteristics for magnetic fields below and above 3 mT, allowing to resolve this ambiguity. We use a physics-based model function to extract fit parameters, which we use for regression and compare it to an approach, purely based on an artificial neural network.

[1528] *A Hybrid, Computer Intensive Approach Integrating Machine Learning and Statistical Methods for Fake News Detection*

Livio Fenga (university of Exeter).

The challenge of early identification of fake news is addressed within the context of anomaly detection for time-dependent data. The presented approach entails a computationally intensive method based on a resampling scheme utilizing Maximum Entropy principles. It is hybrid in nature, employing a sophisticated machine learning algorithm alongside two time series statistical models, complemented by bootstrapped versions of two statistical tests (namely, Binomial and χ^2). In the devised framework, the detection of fake news through an anomaly detection system involves pinpointing patterns that sharply deviate from the norm, indicative of a significant and temporary shift in the underlying data-generating process. After selecting a trending topic based on expert opinions, a set of keywords extracted from diverse online sources (e.g., Twitter, Google, Facebook) undergoes outlier testing. The results are then cross-checked both between and within sources to unveil any inconsistencies.

[1560] *A Contextual Genetic algorithm for weighting Salient Crypto assets*

Moulay Amine Jaidi (Semantichasm).

Introduction

Crypto assets surged in popularity during the 2020 pandemic and their profitability attracted several traders and investors into the market. The following three years saw significant market movements and transitions and many assets dropped in value incurring significant losses to investors. In 2023 crypto assets experienced a slow-moving bear market when in February 2024 Bitcoin (BTC) became an Exchange-Traded Fund (ETF) allowing institutions to invest in the asset in the US, Hong Kong and Australia pushing BTC to new all-time highs. Bitcoin's influence on market sentiment generally leads other assets to follow its price trend. As exchanges around the world list new assets with a potential salient upside, exploring the ones to likely yield high returns and less risk can improve trading strategy.

"Salience theory naturally lends itself to the analysis of the demand for risky assets. After all, risky assets are lotteries evaluated in a context described by the alternative investments available in the market" [1].

This study outlines a Contextual approach that leverage Genetic algorithms (GAs) to search for salient assets using data from crypto exchanges. Our study defines Contexts as a frame of discernment to describe market states that contain events and situations of interest.

Method

Genetic algorithms are applied extensively in portfolio analysis and optimise a variant of the Markowitz model on daily price data. The conventional Markowitz Expected-Value model assumes a market composed of 'n' different assets with corresponding expected returns 'x', and asset covariances Qij. Applying a GA to this model can find the optimal weights 'Wi' for the covariance matrix Qij to solve the equation Wi * Qij * Wi. The Markowitz Model's expected return can be used as a parameter to explore positive values of expected return 'x' and minimize risk quantified by the variance as explained in [2]. This study focuses on assets bringing 3% and above as this was Bitcoin's minimum daily growth during the selected Contexts of the 2024 bull market.

The approach adopted in this study follows a three-step process for Context identification, selecting salient assets, and aggregating search results. Contexts are defined as price growth periods in Bitcoin's daily candles. As such Context will frame price periods starting with a minimum to a desired percentage increase or maximum. The aim is to explore these periods at a granular level for many assets to highlight which ones are salient and follow in Bitcoin's price trajectory.

The Context periods are not set in number of days with selected periods varying between 3 to 7 days. In this case a Genetic Algorithm is used to search asset returns from 15mins data at scale across Contexts and output a vector of weights Wi assigned to crypto assets. A final step of aggregating the results for all Contexts by calculating an average weight will be used to select the highest averages for further technical analysis.

Findings

The method was applied to search for salient assets in data collected from Binance and Crypto.com during the recent 2024 bull market. Aggregating the results from the search on all Contexts proved the influence of Bitcoin on certain asset classes such as meme-coins and NFTs. These assets weighted by the algorithm are salient and reoccur across the searched Contexts. Our study suggest that these can be considered as part of a portfolio to trade in the short-term given that Contexts cover periods of 3 to 7 days. Crypto.com generally lists a varied set of tokens exhibiting high volatility and returns makes it a good candidate exchange for finding salient assets to trade.

Conclusion

Using Genetic Algorithms to search for salient assets in selective Contexts was successful in identifying the ones to trade when considering high upsides and lower risk. Searching across two exchanges was enough to get conclusive results on the influence of Bitcoin and finding enough salient assets to explore for further analysis. Additionally, the method suggests that the weighting of salient assets in selected Context periods is more aimed at short-term trading as opposed to long-term holding. However, the method is flexible enough to apply for long-term asset holding by testing the different parameters of return and reducing the number of assets to search covering longer Context periods.

References

[1] Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. "Salience and asset prices." American Economic Review 103.3 (2013): 623-628.

[2] Yahaya, A. "Portfolio selection using genetic algorithms." International Journal of Natural and Applied Sciences 5.4 (2009): 365-373.

[1594] *The relationship between corporate social responsibility and ethical leadership: the role played by corruption in countries.*

Isidro Peña (University of Castilla La Mancha), Rosa M. Muñoz (University of Castilla La Mancha), Silvia M. Andrade (University of Castilla La Mancha) and Graca Silva (University of Lisbon).

This study focuses on the relationship between ethical leadership and corporate social responsibility. The objectives of the research include providing a more in-depth under-standing of the relationship between ethical leadership and corporate social responsibil-ity, which is studied within different dimensions, and clarifying the role played by cor-ruption in countries. The CSRHUB database was used to analyze 340 private organizations in the European financial industry. The research employed a longitudinal design and the partial least squares structural equation modeling (PLS-SEM) technique to test hypotheses related to the impact of EL on CSR and the moderating role of corruption in countries. The results support the positive relationship between ethical leadership and corporate social responsibility, with ethical leadership being more effective as regards enhancing corporate social responsibility performance when directed toward the community and employees. The study also confirms the moderating role of corruption in countries in the ethical leadership- corporate social responsibility relationship.

[1603] *ASSESSING URBAN BICYCLE TRAFFIC USING A FORECASTING MODEL*

Anamaria Ilie (National University for Science and Technology POLITEHNICA Bucharest), Eugen Roșca (National University for Science and Technology POLITEHNICA Bucharest), Cristina Oprea (National University for Science and Technology POLITEHNICA Bucharest), Aura Ruscă (National University for Science and Technology POLITEHNICA Bucharest), Oana Dinu (National University for Science and Technology POLITEHNICA Bucharestt) and Valentina Radu (National University for Science and Technology POLITEHNICA Bucharest).

The city seen as urban expansion, which depends to a large extent on the increase in the population share, is a sensational and colossal phenomenon, but not a starting point in the history of mankind. Like all events in life, this is only a phenomenon in time, which must be researched in relation to the past in order to be fully understood. A sustainable transport system must provide mobility and access to all urban inhabitants, offering a safe, cozy and comfortable transport. Bicycles serve as a sustainable and efficient mode of transportation, offering numerous benefits ranging from reduced traffic congestion to improved public health. Forecasting bike use is crucial for urban planners and policymakers to optimize infrastructure investments and enhance accessibility. However, accurately predicting bike usage requires a comprehensive understanding of both physical infrastructure and socio - economic factors influencing individuals' travel behaviours. The authors will use the previously described forecasting models in a case study applied to a small city in southern Romania, which aims to promote cycling as a sustainable mode of transport. Starting from the current situation, the forecast model is used in the prediction of future bicycle traffic considering the values of local factors.

[1658] *European electricity market integration and volatility spillovers*

Evelyn Chanatásig-Niza (University of the Basque Country, UPV/EHU), Aitor Ciarreta (University of the Basque Country, UPV/EHU), Cristina Pizarro-Irizar (University of the Basque Country, UPV/EHU, and Basque Center for Climate Change, BC3) and Ainhoa Zarraga (University of the Basque Country, UPV/EHU).

The creation of a single European market for electricity, which formally began in 1996, was expected to enhance competition and bring efficiency gains for European consumers and industries. To that aim, several measures were adopted in the past, and the European Commission is currently working on reforming the European Union's electricity market to protect consumers from electricity price volatility and boost investment in renewable energy, among others. Therefore, the analysis of volatility spillovers between the European electricity markets is crucial and provides useful information for market participants in order to anticipate the expected outcomes of further integration initiatives. This work analyzes volatility spillovers between nine European electricity markets for the period 2015-2021 using the forecast-error variance decomposition in a VAR model approach including both variances and covariances. The results show that the markets under study are highly integrated although there are differences in the degree of integration across markets. The volatility spillovers are explained to a large extent by the interconnections and harmonization in market regulation. Moreover, the inclusion of covariances prevents the underestimation of spillovers. The results are important for market participants to implement optimal policies including electricity price risk management strategies and an efficient use of interconnections.

[1786] *Comparative methods for structural breaks in time series*

Dulce Gomes (Dept. of Mathematics, School of Science and Technology,, University of Évora).

The aim of this work is to assess whether, when and how the structure of the data generation process underlying a set of observations changes over time. For that purpose, different methods in detecting breakpoints (changepoints or structural breaks), were studied through simulation studies. In particular, methods based on probability density ratio and methods based on model fitting, with the appropriate autocorrelation structure. Due to the specific nature of simulated time series, a form-based time series clustering approach using elastic measures such as Dynamic Time Warping (DTW) distance as a measure of dissimilarity and the hierarchical clustering algorithm to separate groups according to data similarities, were used to group the series according to similar patterns. This approach is extremely important when we want to evaluate structural breaks in many time series simultaneously, since the methods for identifying breakpoints require a priori certain specifications about the time series.

[1796] *Do Chinese foreign investment boost the Latin American countries?*

Nadia Urriola Canchari (University of San Marcos) and Abdul Samad Ibrahimi (University of San Marcos).

The world is increasingly getting urbanized and globalized, and even there have been a covid pandemic since 2020, the Asian countries are still having an economic growth, having a significant impact in Latin American economies. Since most Latin American countries have proximity to the Amazon, they, therefore, have vast agricultural, fuel and other products, which are mainly exported to China for its exponential demand growth. Therefore, to the best of our knowledge, this study is the first to investigate the link between Chinese investment in Latin American countries and their economic growth through a panel data model, using as parameters the inflation index, the depreciation, the exportation and importation to China, and the index of industrialization. Furthermore, a feedback causality was found between the studied variables. This study argues for the increase of the Chinese investment in strategic sectors through the introduction of sustainability practices in the natural resource sector. More so, this study has as primordial objective to analyze the different reception of the Chinese investment in Latin American countries to proceed with some political strategies between them to have, in a long term, a sustainable economic growth as Alliances.

[1837] *A Simple Computational Approach\\ to Predict Long-Term Hourly Electric Consumption*

Etinenne Meunier (Boston University Metropolitan College), Pierre Moreau (Boston University Metropolitan College), Tanvi Sharma (Boston University Metropolitan College) and Eugene Pinsky (Boston University Metropolitan College).

By exploiting the patterns in past data points, we could forecast long-term consumption with a computationally simple algorithm. Our approach is simple to interpret. It incorporates the seasonality of past consumption and can predict power consumption for any time scale. The algorithm can be easily implemented directly in SQL. It can run sub-second long-term predictions on large-scale data marts. The proposed method scored a Mean Absolute Percentage Error (MAPE) of just 5.88% when predicting hourly values for France's electric consumption in 2017 based on hourly data from 2008-2011.

[1964] *Optimizing Biogas Power Plants through Deep Learning-Aided Rotor Configuration*

Andreas Heller (FH Münster), Héctor Pomares (Universidad de Granada) and Peter Glösekötter (FH Münster).

The increasing demand for sustainable energy sources has in- tensified the exploration of biogas power plants as a viable option. In this research, we present a novel approach that leverages deep learning tech- niques to optimize the performance of biogas power plants through the strategic placement configuration of rotors within the fermentation ves- sel. Our study involves the simulation of a diverse range of biogas power plant scenarios, each characterized by varying rotor locations and rotat- ing speeds, influencing the agitation levels of the biogas substrate. The simulation results, encompassing multiple performance metrics, serve as input data for an artificial neural network (ANN). This ANN is trained to learn the intricate relationships between rotor placement, rotor speed, agitation levels, and overall system efficiency. The trained model demon- strates predictive capabilities, enabling the estimation of plant efficiency based on specific rotor configurations. The proposed methodology pro- vides a tool for both optimizing existing biogas power plants and guid- ing engineers in the design and setup of new facilities. Our model offers valuable insights for engineers in the initial planning stages of new biogas power plants, enabling them to make informed decisions that contribute to sustainable and efficient energy generation.

[1971] *Ordinary and robust NIPALS decompositions to explore time series*

Adelaide Freitas (Department of Mathematics, University of Aveiro. CIDMA-UA), Alberto da Silva (CIDMA, University of Aveiro) and Filipa Santana (Department of Mathematics, University of Aveiro. CIDMA-UA).

The difference matrix between the ordinary and robust NIPALS decompositions of the trajectory matrix associated with a time series is computed. We investigate this approach concerning to its advantages for detecting temporal patterns and facilitating clustering tasks in time series analysis.

[2016] *Deep and Interpretable Probabilistic Forecasts*

Philipp Baumann (ETH Zurich).

Quantifying uncertainty plays a crucial role in many high-stakes decision-making processes. The advent of deep learning has led to a proliferation of novel probabilistic forecasting tools. However, the increasing complexity of deep learning models has compromised their interpretability. Moreover, the question of whether deep learning is truly beneficial for probabilistic forecasting remains unanswered. In light of these developments, we propose an extension to autoregressive transformation models –a semi-parametric probabilistic forecasting method– using deep learning. Our approach aims to improve predictive performance while enhancing interpretability with a newly developed interpretability score. To achieve this, we embed our new model class in a multi-objective optimization framework and tackle the optimization problem using a modified version of NSGA-2, an evolutionary algorithm. We demonstrate the effectiveness of our approach by applying it to widely used time series benchmark datasets.

[2061] *Gender as a tool for diversification*

Ana Escribano (Universidad de Castilla-La Mancha), Antonio Díaz (Universidad de Castilla-La Mancha) and Rocio Hidalgo (Universitat de València).

This study investigates the impact of managerial gender on portfolio risk, based on the assumption that women are more risk averse than their male counterparts. The research focuses on a key question: Does this risk aversion manifest itself in the stock market performance of firms that are majority-led by women? To answer this question, we constructed two distinct portfolios based on the gender of top managers: one composed of predominantly female-led companies and one composed of exclusively male-led companies. The selection of companies was based on MSCI's annual global gender diversity reports, and the analysis was conducted over a sample period from 2018 to 2022. Portfolio risk was assessed using the value-at-risk (VaR) metric, with dynamic optimisation procedures applied through an AR(1)-EGARCH(1,1) model to minimise variance. Our results reveal that portfolios composed of predominantly female-managed companies exhibit lower risk, as measured by VaR, compared to portfolios of male-only managed companies.

[2078] *Assessing asymptotic tail independence: a simulation study*

Marta Ferreira (Centro de Matemática, Universidade do Minho).

The occurrence of extreme values in one variable can trigger the same in other variables, making it necessary to assess the risk of contagion. Usual dependence measures based on the central part of the data typically fail to assess extreme dependence. Within the scope of EVT, tail dependence measures were developed, such as the Ledford and Tawn coefficient that we discuss here. This is a measure of residual dependence that is particularly important when it comes to analyzing at the tail level where data is scarce. We will consider different estimation methodologies and compare them based on a simulation study. We finish with an application to real data.

[2094] *Boosted Value at Risk and Expected Shortfall – application of gradient boosting machine learning models in market risk estimation problem*

Michał Woźniak (University of Warsaw).

The accurate estimation and management of market risk is of paramount importance for financial institutions, especially with regard to the commonly used risk metrics of Value at Risk (VaR) and Expected Shortfall (ES). The absence of a universally ideal model for estimating VaR and ES, particularly in the context of global economic crises that often distort the performance of classical statistical models, underscores the need for novel and more robust models. This becomes particularly relevant with the increasing intersection of machine learning and statistics, which has given rise to a multitude of advanced models for VaR and ES estimation.

This study seeks to contribute to this growing body of research by focusing on the development and evaluation of shallow statistical learning models for market risk estimation. These models leverage the capabilities of Gradient Boosting Machines (GBM), which have been observed to dominate classical econometric models in many forecasting tasks, for instance please see M5 forecasting competition (Januschowski et.al., 2022). Moreover, the study is motivated by the growing prevalence of machine learning model architectures that embrace the concept of probabilistic modeling (Duan et al., 2020; März & Kneib, 2022; Rigby & Stasinopoulos, 2005), as well as the increasing utilization of transfer learning in contemporary machine learning, exemplified by foundation models like Nixtla TimeGPT (Garza & Mergenthaler-Canseco, 2023).

The research objectives have been systematically divided into three distinct phases. The first phase involves identifying the most effective approach for market risk models from classes of probabilistic learning architectures. These include but are not limited to: 1) quantile forecast: Quantile Linear Regression, Quantile K-Nearest Neighbors, Quantile Random Forest, Quantile Extra Trees, Quantile Gradient Boosting Machine, Quantile LightGBM, Quantile CatBoost, 2) distribution forecast: Natural Gradient Boosting for Probabilistic Prediction (NGBoost), Probabilistic Gradient Boosting Machines (PGBM), XGBoost for probabilistic prediction (xgboost-distribution), XGBoost for Location, Scale, and Shape (XGBoostLSS), LightGBM for Location, Scale and Shape (LGBoostLSS), 3) variance forecast: ARCH-like Random Forest, ARCH-like Extra trees, ARCH-like LightGBM, 4) interval forecast: Model Agnostic Prediction Interval Estimator (MAPIE). The aforementioned models are examined and implemented.

The second phase is dedicated to enhancing the best machine learning model and its probabilistic architecture obtained from the first phase. This involves the utilization of advanced machine learning techniques such as transfer learning, the introduction of high-parameter probability distributions, extensive tuning of hyperparameters, feature engineering/selection, and smoothing of the output layers of models.

The third and final phase consists of comparing the effectiveness of the model developed in the second phase with state-of-the-art classical econometric models.

The research questions guiding this study are fourfold. (RQ.1) Firstly, is the model utilizing the Gradient Boosting Machine concept the top-performing choice among shallow probabilistic machine learning models for market risk? (RQ.2) Secondly, is the distributional forecast the most effective probabilistic statistical learning architecture for market risk modeling? (RQ.3) Thirdly, do techniques such as transfer learning, the introduction of high-parameter probability distributions, extensive tuning of hyperparameters, feature engineering/selection, and smoothing of the output layers of models yield improved machine learning models for VaR and ES? (RQ.4) Lastly, are the developed shallow statistical learning models for market risk estimation capable of outperforming state-of-the-art econometric models in terms of effectiveness?

The study's results are evaluated using a custom Light Benchmark (Woźniak & Ślepaczuk, 2023), encompassing 375 validation paths, with two-thirds dedicated to VaR and one-third to ES. This benchmark covers several categories of validation techniques during the scoring procedure, including fulfilling regulatory requirements, forecasting adequacy, and capital effectiveness. The benchmark paths are constructed on various dimensions, such as asset categories (commodities, cryptocurrencies, currencies, equity indices and stocks), representatives of each asset category (the 5 most liquid representatives in each category), testing periods (5 periods, where each period has different volatility

characteristics for the training and evaluation sets e.g. high volatility on training period and low volatility on testing period), testing period sizes (250 days per each testing period), testing horizons (one-step-ahead forecast), and VaR/ES confidence levels (VaR at 1% and 2.5%, and ES at 2.5%).

The findings reveal that the model utilizing the Gradient Boosting Machine concept – the XGBoostLSS, is the top-performing choice among shallow probabilistic machine learning models for market risk estimation. Furthermore, the distributional forecast is identified as the best probabilistic statistical learning architecture for market risk modeling. However, the study does not rule out other probabilistic architectures in future stages of research. The utilization of advanced machine learning techniques such as transfer learning, the introduction of high-parameter probability distributions, extensive tuning of hyperparameters, feature engineering/selection, and smoothing of output layers of models has led to significant improvements in the machine learning models for VaR and ES. Finally, the shallow statistical learning models developed in this study have demonstrated their capability to outperform state-of-the-art econometric models in terms of effectiveness.

Overall, this research contributes to the advancement of market risk estimation by developing and evaluating shallow statistical learning models based on the Gradient Boosting Machine concept and various probabilistic statistical learning architectures. The findings of this study hold significant implications for financial institutions by providing efficient and robust models for VaR and ES estimation.

References:

Duan, T., Anand, A., Ding, D. Y., Thai, K. K., Basu, S., Ng, A., & Schuler, A. (2020, November). Ngboost: Natural gradient boosting for probabilistic prediction. In the International conference on machine learning (pp. 2690-2700). PMLR.

Garza, A., & Mergenthaler-Canseco, M. (2023). TimeGPT-1. arXiv preprint arXiv:2310.03589.

Januschowski, T., Wang, Y., Torkkola, K., Erkkilä, T., Hasson, H., & Gasthaus, J. (2022). Forecasting with trees. International Journal of Forecasting, 38(4), 1473-1481.

März, A., & Kneib, T. (2022). Distributional gradient boosting machines. arXiv preprint arXiv:2204.00778.

Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society Series C: Applied Statistics, 54(3), 507-554.

Woźniak, M., & Ślepaczuk, R. (2023). LIGHT Benchmark-comprehensive backtesting framework for market risk models comparison. Available at SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4586897

[2242] *Shadows of Resilience: Exploring the Impact of the Shadow Economy on Economic Stability in Times of Crisis*

Charalampos Agiropoulos (University of Piraeus), James Chen (Michigan State University), Thomas Poufinas (University of Thrace) and George Galanos (University of Piraeus).

In this study, we embark on a comprehensive analysis of the shadow economy within the European Union, examining its influence on the economic resilience of member countries. Our dataset spans over two decades, covering a broad spectrum of EU nations, and provides a unique opportunity to explore the relationship between the size of the shadow economy and economic stability during various economic cycles. To delve into this intricate relationship, we employ a diverse array of econometric approaches. OLS and Fixed effects methodologies serve to mitigate omitted variable bias, providing a means to control for unobserved heterogeneity within data. Conversely, techniques such as Lasso and Automatic Relevance Determination (ARD) tackle the challenge of overfitting caused by the inclusion of numerous irrelevant and redundant variables within the design matrix, which introduce noise and obscure clear scientific insights. Although such noise does not necessarily impede accurate predictions, it significantly hinders our ability to understand the underlying reasons for these predictions. These rigorous statistical techniques reveal how the shadow economy proxied by Schneider's shadow economy index interacts with key economic resilience indicators, such as GDP, national debt, and population figures, across different phases of economic stability and turbulence. This research aims to uncover the nuanced role that the shadow economy plays in the economic fabric of the EU, considering the diversity of economic contexts within the EU. Preliminary findings suggest a complex and varied relationship between the shadow economy and economic resilience in the EU. The implications of these findings are significant, offering fresh insights for economists, policymakers, and international bodies. This study not only enhances our understanding of the shadow economy's role in economic resilience but also provides a valuable foundation for policy decisions aimed at fostering sustainable economic growth and stability within the European Union.

[2265] *The impact of greenhouse gas emissions on global economic activity and health*

Jose Aureliano Martin Segura (University of Granada) and Cesar Pérez López (University Comllutense of Madrid).

Climate change is considered one of the most complex challenges of this century.

Our main objective is to estimate the impact of greenhouse gas emis-sions on the economic activity and health of the general public, and this study will update the findings of our previous research presented at the 2015 ITISE congress, which used the data available at that time.

All the econometric estimations made using the data panel model indicate that the environmental impact index contributes to the increase in the global mortality rate.

[2301] *Enhanced Volcanic Signal Detection in Santorini: Implications for Sustainable Tourism Development*

Antonios Marsellos (Hofstra University), Katerina Tsakiri (Rider University), Stelios Kapetanakis (Distributed Analytics Labs), Nick Ptak (Hofstra University) and Faith Renner (Hofstra University).

We introduce an application of wavelet and spectral analysis methodologies aimed at detecting volcanic temperature signals with an application on Santorini Island, a renowned active European volcano. Despite its idyllic allure attracting millions of tourists annually, the latent volcanic threat underscores a critical need for advanced monitoring techniques to ensure the safety and continuity of both tourism and local businesses. By monitoring ground temperature variations directly attributable to volcanic activity from those influenced by atmospheric conditions, our research delineates an understanding of the island's geothermal dynamics. Employing high-frequency ground temperature data alongside meteorological observations, we deploy wavelet analysis to identify temperature anomalies across varied timescales, effectively deciphering the volcanic signatures. The integration of the Kolmogorov-Zurbenko filter enhances this analysis, enabling a separation of atmospheric effects and highlighting volcanic temperature patterns. This insight into volcanic temperature oscillations facilitates the construction of an early warning system, instrumental in fortifying the region's resilience against potential natural hazards and strengthening a sustainable economic and touristic development.

[2318] *Exploring Optimal Strategies for Small-Hydro Power Forecasting: Training Periods and Methodological Variations*

Duarte Lopes (ENLITIA), Isabel Preto (ENLITIA) and David Freire (ENLITIA).

The training period of a small-hydro power forecast model is crucial, as its sensitivity depends on the variability of hydrological conditions influenced by factors such as rainfall, streamflow, and environmental dynamics. Opting for the appropriate training duration is critical; a longer period provides a wider range of data for the model to learn from, potentially improving accuracy. In contrast, a shorter period may restrict exposure to various hydrological scenarios, diminishing predictive reliability. Moreover, Transmission System Operators (TSOs) strategies impact hydroelectric power generation, affecting reservoir management and grid balancing. Considering these factors helps tailor forecasting models to specific system conditions, improving accuracy.

The study investigates ideal training intervals for a small-hydro power regression model, essential for reliable forecasts under diverse conditions, specifically targeting Portugal's total small-hydro portfolio.

The small-hydro forecasting model, developed by a Portuguese renewable energy provider, is a regression model based on kernel density estimation. The model is fine-tuned using historical hourly production values, as well as calendar variables, based on a historical dataset covering one month leading up to the day of forecasting. To generate the forecast for a specific hour of the day, the model is applied to a daily average forecast. This daily average forecast is calculated by combining the hydro potential, represented by the average of the latest real production values, with the average forecasted precipitation for the day being forecasted. Notably, the model undergoes training every time it runs. Although this method provides satisfactory results, relying solely on the preceding month may oversimplify complex hydrological patterns and strategies. Therefore, alternative approaches were investigated to determine the optimal historical period for specific forecasting conditions, aiming to enhance accuracy and customization.

Three distinct time series clustering approaches were evaluated: one based on Dynamic Time Warping (DTW), another using a method like K-means, labeled "K-Means Alike," and a final approach based on traditional K-mean clustering.

Dynamic Time Warping (DTW) serves as a distance metric to compare two-time sequences, facilitating the evaluation of their similarity or dissimilarity. Acknowledging the potential limitations of daily data in representing a specific period adequately, this approach uses the seven days leading up to the day of forecast as the Reference Time Series Information. Historical small-hydro power production data is partitioned into one-month time sequences spaced two

weeks apart, resulting in 24 distinct segments along one year of data. Each segment undergoes evaluation to identify the one that most closely resembles the Reference Time Series Information. To perform that decision, a DTW computation is executed between the Reference Time Series Information and each one-month time sequence, with the sequence yielding the lowest DTW distance chosen as the historical data.

For the "K-Means Alike" approach, a variation of K-means clustering is employed. Using the same Reference Time Series Information, a centroid resembling a typical K-means cluster is calculated based on power production and the hour variable. Using the Euclidean distance between this centroid and each point in the historical data, the n-nearest points closest to the centroid are identified. After, to optimize this process and find the optimal length for our training history, this process was tested using different values of n between one and four months. Finally, it is important to highlight that the selected points, unlike the DTW approach, do not need to be sequential and are chosen solemnly based on their Euclidean distance to our centroid. A version was also tested where the centroid was a medoid.

Traditional K-Means clustering, in contrast with the previous approaches, not only takes advantage of the value of daily average of power, but also the average precipitation as variables. For this method, various numbers of clusters (k) were tested to see, once again, the optimal length for our history training set. Once the clusters are defined, the day of forecast – represented by a tuple of daily average power and precipitation - is assigned to a specific cluster, and all days corresponding to that cluster are included as the history training data.

To evaluate each approach, their results were compared against those of the operational approach, utilizing the normalized mean squared error (NRMSE) as the metric. This involved obtaining a small-hydro power forecast by training the regressor with the selected historical data obtained from each approach. The testing period ranged from early November 2023 to mid-February 2024, while the historical data available for the search spanned from January 2023 to the moment when the prediction is being made.

The results of the approaches show varying levels of performance measured by NRMSE: the operational approach achieved 9.65%, DTW method attained 10.88%, K-means Alike with a 2-month length exhibited 9%, and traditional K-means with 6 clusters reached 9.67%. The results show that the "K-Means Alike" variant is the top-performing method, showcasing significant improvements over the operational approach. The optimal cluster length is 2 months of hourly data. This approach is also notable for its reduced computational and time requirements. It was also interesting to note that the selected points predominantly corresponded to the month of the day when the forecast was performed. For instance, predictions made for January 2024 were based on data collected in January 2023, underscoring the significance of acquiring the most relevant data that accurately represents the forecasting period

Dynamic Time Warping (DTW) exhibits promising results, albeit with increased computational time compared to operational and K-means methodologies. Despite its computational demands, DTW does not consistently outperform other methods. Traditional K-Means can potentially yield better results if applied to different variables built as a function of precipitation, such as cumulative precipitation data, capturing trends within the dataset. This approach might enhance the identification of 'typical days,' aiding in the selection of optimal training history. Moreover, both K-Means Alike and Traditional approach reach similar optimal month length. Both approaches seem to reach training dimensions of two months of data, highlighting the existence of an ideal size for training set which can be of the utmost importance for future works.

The dynamic history approach emphasizes the importance of selecting appropriate historical data to achieve optimal outcomes. Proper clustering obviates the need for temporal dependencies in the data, instead prioritizing the alignment of chosen days with the behavior of the Reference Time Series Information. Overall, this work demonstrates that dynamically choosing data improves the quality of the final results. Moreover, diverse methodologies emphasize the strengths and weaknesses of different approaches, each producing compelling outcomes in comparison to the baseline

---

[2366] *Self-normalizing Tests Using the Cauchy Distribution*

Uwe Hassler (Goethe University Frankfurt) and Mehdi Hosseinkouchack (European Business School).

A testing principle is introduced where the statistic is the ratio of two weighted averages. Under general assumptions both averages are asymptotically normal. Upon normalization and orthogonalization the ratio thus converges to the standard Cauchy distribution under the null hypothesis. Critical values and p-values are hence readily available. At the same time a potential nuisance scaling parameter cancels from the ratio without having to be estimated, making these Cauchy tests self-normalizing (scale-invariant). Self-normalization is particularly useful when analyzing time series where the nuisance parameter typically is the so-called long-run variance that is notoriously cumbersome to estimate. These tests are not directed against specific alternatives but rather belong to the toolkit of general specification testing. Still, assuming specific local alternatives, asymptotic power can be determined since the ratio converges to a non-centered Cauchy distribution. Power crucially hinges on the weighting scheme when computing the weighted averages.

Three examples are studied: tests against a Wiener prcocess, against a Brownian bridge, and under long memory. We discuss the choice of weighting schemes building on eigenvalues and eigenfunction of the Karhunen-Loeve (KL) expansions under H_0 by means of asymptotic power analysis. Further, we provide simulated finite sample evidence when testing for the null hypothesis of a zero mean of a time series integrated of order zero against a break in mean or against a local linear time trend. The asymptotic scale-invariance turns out to be effective in finite samples: The Cauchy test has better size control than competing procedures, and this is even more true under long memory of the time series. This nice property comes at the price that the Cauchy test is less powerful than competitors directed against specific alternatives.

Finally, we indicate how Cauchy tests can be carried to a multivariate framework of correlated samples, e.,g., to panel data under cross-dependence. Here, the cross-covariances between the time series reduce to one scaling parameter that cancels from the Cauchy ratio by virtue of self-normalization: The tests are robust with respect to cross-dependence without need to explicitly account for it.

[2434] *Enhanced Renewable Power Forecasting through NWP and Historical Power Data Integration*

Isabel Preto (ENLITIA), António Couto (Unidade de Energias Renováveis e Eficiência Energética, Laboratório Nacional de Energia e Geologia I.P.), Ricardo Faria (ENLITIA), Hugo Algarvio (Unidade de Energias Renováveis e Eficiência Energética, Laboratório Nacional de Energia e Geologia I.P.), Duarte Lopes (ENLITIA) and Ana Estanqueiro (Unidade de Energias Renováveis e Eficiência Energética, Laboratório Nacional de Energia e Geologia I.P.).

Introduction In recent years, considerable efforts have been directed towards developing systems capable of predicting power output from wind or solar power plants, as well as electricity demand. A significant challenge in the energy sector centers on the widespread integration of variable renewable energy sources (vRES) such as wind and solar, into electrical power systems in a manner that is both economically feasible and environmentally sustainable. Transmission system operators (TSOs) continuously work to keep a precise balance between electricity generation and demand. The safe and robust operation of a power system depends on highly accurate forecasts of both variable renewable energy sources' power production and consumption. This precision is critical for minimizing the need for energy balancing in reserve markets, which is often associated with high costs. Furthermore, as we transition towards near 100% renewable power systems, the importance of forecast systems and their accuracy will only grow. In this work three forecasting approaches were simulated, involving the exploration of models based on Numerical Weather Prediction (NWP) and investigating the potential integration of real-time data from power production. Specifically, the scenarios included: i) an NWP-based model using exclusively meteorological data derived from an NWP, ii) a power-based model, using historical power series data, and iii) a hybrid model combining the two previous approaches. The NWP methodology was developed as part of the EU-funded project TradeRES - Tools for the Design and Modelling of New Markets and Negotiation Mechanisms for a ~100% Renewable European Power System.

Data and methods The simulations were geared towards forecasting wind and solar power generation for different spatial levels: national level - Portugal, Spain, and Germany, and wind/power park level covering a total of 27 wind parks and 8 solar parks in Portugal. Wind farm capacities ranged from 17 MW to 255 MW, with an average capacity of 68 MW, while solar farms varied from 8 MW to 40 MW, with an average capacity of 17 MW. For the national cases, data from Ninja.Renewables database was used as the target variable, while for the power plants, the target was the hourly observed power production. Simulations addressed two main forecasting horizon objectives: i) providing a 24-hour forecast crucial for the Day-Ahead Market (DAM), considering both existing and new gate closure hours, and ii) exploring an innovative electricity market design developed in TradeRES project called Period-ahead Market (PAM), which required updated forecasts at 00:00, 06:00, 12:00, and 18:00. The models were trained on data from 2018 and then evaluated on the complete dataset from 2019. The developed NWP methodology optimally combines a large number of meteorological parameters, collected across a national grid of points, using statistical and machine learning approaches, such as dimensionality reduction and feature selection algorithms, prior to applying several regression algorithms calibrated for different weather regimes. One of the core characteristics of this methodology is the integration of NWP-based features often overlooked in literature despite their known influence in wind and solar variablity. Meteorological data, sourced from Global Forecast System (GFS) – a global NWP model, encompasses not only common variables like wind speed and direction but also parameters such atmospheric boundary layer and geopotential height for different pressure levels. PCA is applied to each meteorological indicator, retaining components explaining 85% or more of variance, to identify dominant synoptic variations. Unsupervised machine learning techniques identify weather regimes, with LASSO regression and hybrid methods like Sequential Forward Feature Selection utilized for feature selection. The regression algorithms applied were Random Forest, Support Vector Machines, Extreme Gradient Boosting and LightGBM. During the tests period LightGBM obtained the best performance. Different regression algorithms, including traditional

time series models like ARIMA, were utilized for the power-based model. Ultimately, LightGBM was employed to maintain consistency and based on the performance results. For the power-based model, the power timeseries were iteratively integrated, covering power from the last hour up to the previous 24 hours and, additionally, power from multiples of 24 hours up to a week was considered due to the high autocorrelation for these periods. The hybrid approach combines NWP features, represented by principal components retained for weather indicators, with historical power series data, including various time lags from the prior hour to a week. Furthermore, the performance of the different simulation models was analyzed against a short-term horizon, comparing them to a persistence model that maintains the power value for the initial forecast hour. This analysis was conducted for each national case study and for the 27 wind parks and 8 solar parks.

Results The results of the new NWP methodology developed within TradeRES indicate a significant reduction in normalized root mean square error (NRMSE) compared to the operational wind GFS model from a Portuguese forecast provider. On average, there was a decrease of 5.5%, with a maximum reduction reaching 10.1%, observed across the 27 wind power plants during testing. This underscores the significance of selecting relevant meteorological features for each power park or aggregated country to enhance power forecast accuracy, alongside the crucial role of employing a grid of meteorological points and incorporating non-conventional variables like the atmospheric boundary layer. Regarding the scenarios simulations, results showed that for both technologies the use of NWP had a clear positive impact on the forecast for the next 24 hours compared to the model based solely on the historical power series. This was evident across national simulations, where the hybrid model incorporating NWP proved advantageous not only for PAM but also for DAM periods, as the power-based model exhibited very poor performance for this period. For instance, for the wind case study, in Portugal, for DAM period the power-based model registered the highest value of NRMSE - 70.50%, significantly declining to 31.16% when considering the hybrid model with NWP. The NWP-based model yielded a slightly higher NRMSE value of 31.86%. In the PAM period, the power-based model presents a better performance with an NRMSE of 24.70%, decreasing to 21.99% with NWP integration and the NWP-based model achieved a NRMSE marginally lower value of 24.02%. Comparatively, for Spain, the NRMSE for the power-based model in the DAM period was 51.78%, declining to 21.29% in the hybrid model, while the NWP-based model presented a NRMSE of 21.79%; in the PAM period, the power-based model performed better, with an NRMSE of 20.37%, falling to 13.39% with NWP integration and the NWP-based reaching a NRMSE of 18.15%. In Germany's DAM period, the power based model had a very poor performance represented by a NRMSE of 73.19%, dropping to 23.72% with the NWP incorporation in the hybrid model, while the NWP-based obtained an error of 21.67%; in the PAM period, NRMSE for power-based model was 28.75%, declining to 18.04% in the hybrid model and 20.10% in the NWP-based model. The aforementioned conclusions prevail for individual wind and solar parks. Generally, the hybrid model, combining NWP and historical power series, demonstrates superior performance compared to the NWP-based model, despite their similar results. In contrast, the power-based model consistently performs poorly in the DAM period, especially when compared to those derived from the NWP-based and hybrid models. Notably, in the PAM periods, its performance improves, nearing that of the NWP-based model. With respect to the comparison with the persistence model, the behavior in the three national cases is very similar. The persistence model exhibits the best performance up to the next two hours. The NWP-based model shows nearly constant error behavior, and the hybrid model outperforms the NWP model in the first four hours after which they become very analogous in terms of error. The model based solely on power also experiences a degradation in performance as the time horizon increases, while maintaining the power value. Similarly, for the 27 wind parks, the logic is analogous. The persistence error progressively increases with the time horizon, as does the error of the model based on historical power series. As the horizon extends, the hybrid model slightly degrades in performance, becoming like the NWP-based model. As observed in the wind power model, for the solar power model errors increase as the forecasting horizon extends. However, in this case, the persistence model error exhibits a significant increase, even for 1-hour ahead forecast. The performance of the hybrid model shows a slight decline with a longer horizon, eventually reaching a level comparable to that of the NWP-based model. In the case of solar parks, the same pattern is observed.

Final remarks Despite the improvements achieved in the forecasts for the day-ahead market, high power forecast errors are still observed (a NRMSE of nearly 30% for wind and solar in Portugal and nearly 20% for Spain). Market designs featuring shorter forecast timeframes have the potential to decrease power forecast errors substantially. These findings underscore the importance of assessing the most appropriate forecasting method based on the forecast duration. Consequently, power producers should prioritize facilitating real-time access to observed power data for forecast providers to improve forecast accuracy. Despite adjustments such as delaying bids closer to real-time, wind and solar PV technologies in the day-ahead market still exhibit significant errors. Additionally, for forecast horizons below six hours in alternative market designs, NWP-based results demonstrate inferior performance compared to approaches forecasting future values based on past data.

Acknowledgments. This work has received funding from the EU Horizon 2020 research and innovation program under project TradeRES (grant agreement No 864276)

[2474] *Continual Learning for Time Series Forecasting*

Quentin Besnard (Université Tours, LIFAT EA6300) and Nicolas Ragot (Université Tours, LIFAT EA6300).

Deep learning has brought about significant advancements in the field of artificial intelligence, particularly in robotics, imaging, and sound processing. However, a common major challenge faced by all neural networks is their substantial demand for data during the learning process. The required data must be both quantitative and stationary to ensure the proper functioning of classical models. Nevertheless, adhering to these constraints is quite often impossible for many real life applications because of dynamic environments. Thus, data can evolve over the time, which is known as data and concept drift, where modifications occur in the distribution of data or even in the goals to pursue within these environments. Research in the field of Continual Learning seeks to address these challenges by implementing progressive models capable of evolving over the time. Exploratory efforts are evident in all applications of deep learning (graphs, RL, etc.), but to date, there is still a limited amount of work in the case of time series, specifically in the context of regression and forecasting.This paper aims to provide a first survey on the field of continuous learning applied to time series forecasting.

[2484] *Feature presentation: The proper sampling of feature space as a reasonable response to the unreasonable ineffectiveness of mathematics in economics*

James Chen (Michigan State University).

The intrinsic parsimony of mathematics, remarkably or even “unreasonably” effective in the nat-ural sciences, notoriously fails when applied to economics. One solution to this mystery lies in the inevitably subjective process of experimental design in the social sciences. Predictive variables must be chosen from an intractably large number of candidates. Feature selection, particularly through machine-assisted methods of regularized regression, can ameliorate collinearity as the primary source of Type I error in experimental design.

[2533] *Multi-Source Big Data for Fine-Scale Population Distribution Simulation in Rural Areas*

Li Li (Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences), Ning Niu (School of Resources and Environment, Henan University of Economics and Law) and Xiaojian Li (School of Resources and Environment, Henan University of Economics and Law).

This study presents an innovative approach to simulate fine-scale population distribution ub rural areas using multi-source big data. Leveraging advanced simulation techniques, the research focuses on the dynamic and complex population patterns in mountainous rural regions that are often challenging to depict due to geographical constraints, and integrates diverse data types, including terrain and climate data, land use, road networks, hydrographic data, and Points of Interest (POI), to achieve a more accurate and granular population simulation at the building scale. The implications of this research are vital for rural planning, public health, disaster risk assessment, and environmental conservation in underdeveloped rural areas.

[2579] *Promoting Electric Vehicle Growth Through Infrastructure and Policy: A Forecasting Analysis*

Anuva Banwasi (Columbia University in the City of New York), Brennan McManus (Columbia University in the City of New York) and Adele Sinai (Columbia University in the City of New York).

This study examines electric vehicle (EV) adoption in the United States, specifically the interconnected relationship between EV-promoting policies, EV charging infrastructure, and registrations of EVs. Gasoline-powered vehicles make up a significant portion of the US's carbon emissions, and increasing the use of EVs is a way to decrease this footprint. Over the past decade, there have been many incentives and policy-driven changes to propel electric vehicle adoption forward. The focus of this study is to identify if there is a significant relationship between these three factors, and the extent to which these factors are significant predictors for each other. To do so, we applied statistical tests such as Granger Causality on time-series data to analyze the forecasting effect of changes in EV policies on EV infrastructure, and changes in infrastructure on EV registrations. Through our investigation, we discovered compelling evidence for a significant relationship between EV incentive policies and the growth of charging infrastructure. Furthermore, it is possible to accurately forecast changes in EV charging stations over time using time-series data of previous EV charging stations and

policies. The Vector Autoregression (VAR) shows that time-series data for EV stations and EV policies can be utilized to successfully forecast percentage change in EV stations over time. We also analyzed on a policy category basis and found that laws/regulations as well as private incentives were the most significant for forecasting. These findings can be especially useful in the development of future policies to further encourage EV adoption. This strong forecasting relationship between EV incentive policies and the expansion of charging infrastructure provides valuable insights for policymakers, industry stakeholders, and researchers attempting to understand and promote EV adoption.

[2598] *The impact of the geometry and location of storm surge barriers on storm surge propagation: Great Egg Inlet, New Jersey, USA case study*

Gregory Slusarczyk (US Army Corps of Engineers), Mary Cialone (US Army Corps of Engineers) and Robert Hampson (US Army Corps of Engineers).

The focus of this paper is on numerical modeling and analysis of two alternatives of Storm Surge Barriers (SSBs) or gates closing Great Egg Inlet, New Jersey, USA. The inlet forks into two channels that in turn connect the Atlantic Ocean with two major bays: Great Egg Harbor to the South and Lakes Bay to the North. Therefore, due to the specific geometry of the inlet, the following two alternatives were investigated: 1) a single closure at the entrance of the inlet and 2) two closures, each closing one of the channels. A set of 10 hypothetical storm events was statistically selected to the range of storm conditions in the study region and then numerically simulated for all closure alternatives utilizing the ADvanced CIRCulation hydrodynamic model (Luettich et al. 1992). While investigating the storm surge propagation in the embayment, four gate scenerios were taken into the account: 1) a static gate condition (the gate closed for the entire simulation time), 2) a dynamic gate closure (the short gate closed 2 hours before the 50% of the annual exceedance probability (AEP) water level was reached and re-opened after the storm passed on the falling tide when the head difference at the gate was less than 0.33 m, 3)the long gate closed at low tide 6 to 12 hours before the storm arrived and reopened about 12 hours after the short gate closure), and 4) a Base Condition (the inlet fully open for the comparison reason).

The importance of the location of the closures was proven based on the analysis of the simulation results. The single inlet closure exhibits superior advantage in storm surge mitigation and back-bay flooding since the water behind the structure baywise had free passage between the northern and southern channels allowing equal distribution of surge water. In the case of two storm surge barriers, closing the northern and southern channels separately, there is no flow exchange between the channels which results in excessive water surface elevation in the northern channel and Lakes Bay when the wind direction is N-to-S, trapping water behind the northern channel closure.

As SSBs become more commonly considered to reduce back-bay flooding and to manage coastal storm flood risk, the importance of increased understanding of hydrodynamic processes associated with the implementation of storm surge barriers that is presented in this paper will become increasingly beneficial to the international coastal community.

Luettich, R.A., Westerink, J.J. and Scheffner, N.W., 1992. ADCIRC: an advanced three-dimensional circulation model for shelves, coasts, and estuaries. Report 1, Theory and methodology of ADCIRC-2DD1 and ADCIRC-3DL.

[2604] *Multi-objective Optimization for Selection of Clusterings Across Time*

Sergej Korlakov (Heinrich-Heine-University), Gerhard Klassen (University of Bremen), Luca T. Bauer (University of Bremen) and Stefan Conrad (Heinrich-Heine-University).

Nowadays, time series data is ubiquitous, encompassing various domains like medicine, economics, energy, climate science and the Internet of Things. One crucial task in analysing these data is clustering, aiming to find patterns that indicate previously undiscovered relationships among features or specific groups of objects. In this work, we present a novel framework for clustering of multiple multivariate time series over time, that utilizes multi-objective optimization to determine the temporal clustering solution for each time point. To highlight the strength of our framework, we conduct a comparison with alternative solutions using multiple labeled real-world datasets. Our results reveal that our method not only provides better results, but also enables a comparison between datasets with regard to their temporal dependencies.

[2679] *Prediction of the characteristics of concrete containing crushed brick aggregate*

Marijana Hadzima-Nyarko (Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek, Osijek), Miljan Kovacevic (Faculty of Technical Sciences, University of Pristina, Kosovska Mitrovica), Ivanka Netinger Grubesa (Department of Construction, University North, Varazdin) and Silva Lozancic (Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek, Osijek).

Efficient utilization of limited natural resources is paramount for environmental preservation in the construction industry. This study explores the potential of in-corporating recycled materials, specifically crushed clay bricks, as substitutes for traditional aggregates in concrete production. The research emphasizes the envi-ronmental benefits of reducing natural aggregate consumption, addressing waste management concerns, and conserving natural resources. Drawing from recent studies, the article discusses the mechanical properties of concrete containing recycled brick aggregates. It highlights that incorporating re-cycled brick can lead to concrete with mechanical properties comparable to con-ventional concrete. Achieved precision in predicting concrete characteristics, in-cluding compressive strength at different ages of samples, demonstrates the via-bility of utilizing recycled brick aggregates in concrete production. Notably, recy-cled brick concrete exhibits enhanced fire resistance and favorable thermal prop-erties, contributing to its potential in the construction industry. Furthermore, the study investigates the optimization of both individual regression tree models and ensembles of regression trees for predicting concrete characteris-tics, focusing on parameters such as the minimum leaf size. By experimenting with different minimum leaf sizes and ensemble techniques, such as Random Forest or TreeBagger, the research evaluates performance metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R). Feature im-portance analysis revealed that the age of the concrete samples (in days), super-plasticizer, and crushed bricks larger than 4 mm are the three most important vari-ables influencing concrete characteristics. The water-to-cement ratio also played a notable role in predicting concrete properties. Regression tree models with a minimum leaf size resulting in an RMSE of 4.00, MAE of 2.95, MAPE of 0.10, and R of 0.96 exhibited the best accuracy in pre-dicting concrete characteristics. Overall, the study demonstrates that both individual regression tree models and ensembles of regression trees are researched in the article. The interdisciplinary research integrates insights from materials engineering, environmental science, and predictive modeling techniques to advance sustainable construction practices. The findings provide valuable guidance for industry professionals and research-ers seeking to enhance concrete structures' environmental sustainability and per-formance.

[2687] *The Interconnected Generalized Autoregressive Conditional Heteroskedasticity Model*

Javier Sánchez García (University of Almeria), Salvador Cruz Rambaud (University of Almeria) and Paula Ortega Perals (University of Almeria).

We suggest a novel model, named Connected Generalized Autoregressive Conditional Heteroskedasticity (CN-GARCH), which introduces a connectedness parameter to the conditional variance of a financial time series. This connectedness coefficient measures the degree to which the variance of a financial element is linked with other elements in a financial network. The CN-GARCH model is crafted to improve the effectiveness of GARCH models in real-time systemic risk assessment and the immediate prediction of financial stress. We contend that by integrating a concise metric that captures the speed at which a shock to one component spreads across interconnected financial entities, our model provides a more nuanced understanding of the dynamics inherent in financial markets. Monte Carlo simulations will be conducted to evaluate the model's theoretical foundations and performance across various real-world financial structures. An empirical application is intended to demonstrate the model's efficacy in measuring financial risk compared to a range of GARCH variants.

[2713] *Advancing Sustainable Mobility and Transport through Predictive Multi-Level Control of Fuel Cell Electric Vehicles*

Christoph Hametner (TU Wien) and Stefan Jakubek (TU Wien).

Battery and fuel cell electric vehicles play an important role in mitigating global greenhouse gas and pollutant emissions. Particularly for heavy-duty vehicles, the efficient deployment of fuel cell systems represents a critical element in the transition towards environmentally friendly transportation. However, to achieve reliable and efficient operation of such alternative powertrains, their control architecture and operating strategy are key factors.

In order to address the inherent complexity of fuel cell powertrains, this study focuses on highly integrated concepts for control and diagnosis, ensuring that the vehicle performance meets the high expectations of customers and legislation. The main contribution of this work is centered around optimal and predictive control concepts at different levels of fuel cell electric vehicles. An important aspect of this work is to incorporate forecasts on, for example, route, traffic, and road

conditions, enabling an accurate power demand prediction. By directly taking this forecast into account in the vehicle operating strategy, a significant increase in efficiency can be expected.

The study presents recent investigations and results on predictive control at multiple levels. At the lower level, individual components, such as the fuel cell system control, are considered. At the higher level, a predictive energy management and cooling strategy ensures that the required power is provided even in demanding driving scenarios and that various constraints (e.g., state of charge, power, and temperature constraints) are met at any point in time. In particular, this involves optimizing control of the battery's state of charge and battery temperature along the entire route. Furthermore, improving efficiency and component lifetime at the vehicle by means of a health-conscious adaptation of the energy management strategy and also the impact of optimal speed planning is discussed.

The results highlight how forecasts and predictive control concepts, implemented at different levels, contribute to improving the performance, efficiency, and component lifetime of fuel cell electric vehicles, paving the way for sustainable transport in the future.

[2819] *Measuring the impact of climate risk in financial markets: a joint quantile and expected shortfall regression mode*

Lidia Sanchis-Marco (Universidad de Castilla-La Mancha) and Laura Garcia-Jorcano (Universidad de Castilla-La Mancha).

We propose a novel approach to the measurement and forecasting of financial risk conditioned to carbon risk exposure using the joint quantile and expected shortfall semiparametric methodology based on Dimitriadis and Bayer (2019), and Fissler and Ziegel (2016). This method allows us to estimate and forecast two climate risk measures called CoClimateVaR and CoClimateES which capture the dependence of the bank returns on extreme changes in CO2 returns at extreme quantiles representing green and brown states. We further built related measures, DeltaCoClimate and ExposureClimate for VaR and ES specifications. The first one, appraise the effect on the tail risk of the bank to an improvement/deterioration in climate risk that leads to entering a green or brown state. The second one, allows us to establish each bank's exposure to climate risk and to identify differences between banks. We find that the banks appear to be highly vulnerable to climate risks, whereas there is a higher potential for large losses in crisis periods, given the different pace of banking sector restructuring.

[2919] *Hybridizing Machine Learning with Time Series Analysis for Enhanced Forecasting in Management Science and Operational Efficiency: A Systematic Review*

Aydin Teymourifar (Católica Porto Business School, Centro de Estudos em Gestão e Economia, Porto, Portugal) and Maria A. M. Trindade (SDA Bocconi, School of Management, Milano, Italy).

Abstract: In the dynamic landscape of management science, this systematic review provides a comprehensive exploration of the amalgamation of machine learning techniques with traditional time series analysis methods. As time series analysis continues to play an increasingly pivotal role in enhancing managerial decision-making processes by offering insights derived from sequential data points, this study endeavors to shed light on the multifaceted applications and synergistic benefits resulting from the integration of time series analysis with machine learning. By scrutinizing a diverse array of studies and practical implementations, the study aims to illuminate the rich potential of this hybrid approach across various domains, including market trend forecasting, inventory management, financial management, and operational efficiency. Through an in-depth analysis, this review elucidates how the fusion of machine learning and time series analysis contributes to heightened forecasting accuracy and operational efficacy, thus empowering decision-makers with more robust insights and strategies. Keywords: Machine learning, Time series analysis, Forecasting, Management science, Operational efficiency. Introduction The integration of machine learning with traditional time series analysis methods has become pivotal in the field of management science. Time series analysis provides crucial insights into sequential data points, aiding managerial decision-making. This study explores the symbiotic relationship between machine learning and time series analysis, investigating its impact on market trend forecasting, inventory management, financial management, and operational efficiency. Through a systematic review, we uncover how this hybrid approach enhances forecasting accuracy and operational efficacy in management practices. In market trend forecasting, hybrid models that combine machine learning algorithms, such as Neural Networks and Support Vector Machines, with traditional time series models like ARIMA, have proven effective in addressing complex nonlinear patterns and seasonality. The incorporation of feature extraction techniques, such as Fourier or Wavelet Transforms, further enhances forecasting capabilities by extracting meaningful insights from time series data (Cheng Zhang et al., 2022; Dama & Sinoquet, 2021, Ghaderpour et al., 2021).

Inventory management experiences enhanced benefits through the integration of machine learning with time series forecasting, facilitating precise predictions of product demands and the optimization of inventory levels. The utilization of techniques like reinforcement learning serves to further refine inventory optimization strategies, resulting in substantial cost savings and efficiency improvements. This integration of machine learning with time series forecasting not only aids in accurately predicting product demands but also in optimizing inventory levels by considering factors like promotional activities and historical sales data. The incorporation of reinforcement learning techniques further refines the optimization of inventory levels based on predictions, ultimately leading to significant cost savings and efficiency improvements (Seyedan & Mafakheri, 2020). In financial management, the utilization of machine learning models for analyzing financial time series data plays a pivotal role in risk management and portfolio optimization. The integration of hybrid models, which combine predictive analytics with traditional financial theories, empowers more informed decision-making and enhances risk assessment strategies. Specifically, machine learning models are employed to thoroughly analyze financial time series data, contributing to the refinement of risk management and portfolio optimization. The synergy achieved in hybrid models, blending predictive analytics with traditional financial theories, further elevates decision-making processes and refines strategies for risk assessment (Cheng Zhang et al., 2022). Operational efficiency receives a substantial boost from the application of machine learning to time series data, particularly in predictive maintenance and process optimization. Hybrid modeling, a combination of machine learning and other techniques, plays a crucial role in identifying performance bottlenecks and predicting future system behaviors, ultimately contributing to enhanced operational strategies. The application of machine learning to time series data specifically facilitates predictive maintenance, effectively reducing downtime by forecasting equipment failures. Concurrently, hybrid modeling proves instrumental in process optimization, where it identifies performance bottlenecks and predicts future system behaviors, ensuring a comprehensive approach to operational enhancement (Fan et al., 2021; Yi et al., 2023).

Methodology The methodology of this study is the systematic review, guided by established frameworks such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), which involves a comprehensive and structured process of collecting, analyzing, and synthesizing existing literature on the topic. This approach follows predetermined criteria to ensure rigor, transparency, and reproducibility in the review process (Moher et al., 2009).

Conclusion Based on the methodology and delving into subjects such as data preprocessing, feature extraction, model selection, and performance evaluation, we leverage the strengths of both machine learning and traditional time series analysis. The integration of machine learning with time series analysis presents a promising avenue for enhancing forecasting accuracy and operational efficiency across diverse domains within management science. This review underscores the potential of this integrated approach in refining decision-making processes and operational strategies. Synthesizing existing literature and practical applications, this study contributes to the advancement of management practices, ultimately fostering improved outcomes in various organizational contexts. References Cheng Zhang, Nilam Nur Amir Sjarif, Roslina Ibrahim. (2022). Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020-2022. arXiv preprint. Dama, F., & Sinoquet, C. (2021). Time series analysis and modeling to forecast: A survey. arXiv preprint arXiv:2104.00164. Fan, C., Xiao, F., Madsen, H., & Wang, D. (2021). A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data. Frontiers. Ghaderpour, E., Pagiatakis, S. D., & Hassan, Q. K. (2021). A Survey on Change Detection and Time Series Analysis with Applications. MDPI Applied Sciences. Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Medicine, 6(7). Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. Journal of Big Data, 7(1), 1-22. Yi, K., Zhang, Q., Wang, S., He, H., Long, G., & Niu, Z. (2023). Neural Time Series Analysis with Fourier Transform: A Survey. arXiv preprint arXiv:2302.02173.

[3079] *Short Term Forecasting of Nonstationary Time Series*

Amir Aieb (Faculty of Engineering, Free University of Bozen-Bolzano, 39100 Bolzano, Italy), Antonio Liotta (Faculty of Engineering, Free University of Bozen-Bolzano, 39100 Bolzano, Italy), Alexander Jacob (Eurac Research, Institute for Earth Observation, Bozen-Bolzano, 39100 Bolzano, Italy) and Muhammad Azfar Yaqub (Faculty of Engineering, Free University of Bozen-Bolzano, 39100 Bolzano, Italy).

Forecasting climate events is crucial for mitigating and managing risks related to climate change; however, the problem of nonstationarity in time series (NTS) arises, making it difficult to capture and model the underlying trends. This task requires a complex system to address the challenge of creating a strong model that can effectively handle the nonuniform variability provided in various climate datasets. In this work, we use a daily standardized precipitation index dataset as NTS, whereby the heterogeneous variability of daily precipitation poses complexities for traditional machine learning models in predicting future events. To address these challenges, we introduce a novel approach, aiming to adjust the non-uniform distribution and simplifying detection time lag using autocorrelation. Our study employs a range of

statistical techniques, including sampling-based seasonality, mathematical transformation, and normalization, to preprocess data for increasing the time lag window. Through the exploration of linear and sinusoidal transformations, we aim to assess their impact on increasing the accuracy of forecasting models. A strong performance is effectively observed by using the proposed approach in capturing more than one year of time delay, across all the seasonal subsets. Furthermore, improved model accuracy is observed, notably with K-Nearest Neighbors (KNN) and Random Forest (RF). The study underscores RF's consistent strong performance across all transformations, while KNN demonstrates optimal results only when the data has been linearized.

[3084] *Landscape Transformation & GLOF's Risk in Climate Changing Scenario in Hindukush Himalayan Region of Pakistan*

Muhammad Qasim (Department of Environmental and Conservation Sciences, University of Swat), Kainat Nafees (Abdul Wali Khan University Mardan) and Mudassir Khan (Quaid-e-Azam University Islamabad).

Mountains are highly subject to land cover changes, which are driven by human induced factors and climate change. This study is an attempt to assess the landscape transformation particularly changes in forest, agriculture, waterbodies and snow cover. The study takes in consideration the spatial and temporal dynamics of LULC in Upper Swat Pakistan, using satellite imageries for the years 1988 and 2015. Supervised classification was applied to demonstrate land use changes during the mentioned period. We found drastic changes in this high mountainous landscape, as in 1988 the forest cover was 361.1 Km2, which decreased to 339.9 Km2 in 2015. Major causes of shrinking greenery are attributed to high fuel wood demand of growing population, lack of alternative energy resources, disputed ownership and timber mafia. The built-up area increased by100% in 27 years consuming 6 Km2 of agricultural land and 1.5 Km2 of rangeland. The significant and eye-catching decrease took place in snow cover, which was 1431.6 Km2 in 1988, while in 2015 it decreased just to 370.6 Km2. With melting glaciers formation of new lakes is obvious and our maps analysis showed 66 new lakes in the study area. The altitudinal range of these newly formed lakes is 3350 to 4606 m2 above sea level. The largest among new lakes has an area of 490.65 Km2, while the smallest one is just 1.4 Km2. Our maps analysis further showed that glacial lakes at high altitude are more vulnerable that might burst causing havoc in the area. These results are valuable for many line agencies, disaster management departments working on climate change, environmental risk assessment, climate change impacts and adaptation strategies.

[3097] *State-Space Modeling in Time Series Analysis: A Data Assimilation Approach and Challenges in Parameter Estimation*

Marco Costa (University of Aveiro), Magda Monteiro (UA), A. Manuela Gonçalves (University of Minho) and F. Catarina Pereira (University of Minho).

State-space models are increasingly popular in time series modeling in several areas due to both their flexibility and intuitive structure, describing the different individual processes of a complex system, and are considered in various areas, from economics to the environment or neurology. This type of stochastic model is particularly useful for modeling data assimilation processes. Data assimilation process is, in general, a sequential time-stepping procedure, in which forecasts from a source are compared with newly received observations to produce optimal forecasts. It can occur in meteorology, where meteorological forecasts are improved by combining forecasts and observations or by combining several remote sensing data sources that might be used for automated prediction of the location of objects. In this context, state space models with a stationary state have particular application. Despite often being presented only as a numerical procedure, the state space models adjustment has highlighted challenges in its implementation. In this context, the estimation of parameters by both maximum likelihood and distribution-free estimators based on the method of moments is discussed. One of the aspects addressed in this presentation is the impact of replacing the unknown parameters with their estimates, even if obtained through consistent estimators, on the loss of the optimal properties of the Kalman filter predictors. These methodological aspects are demonstrated by modeling real data, namely precipitation and maximum temperature data.

[3136] *Extraction and forecasting of trends in cases of signal rank overestimation*

Nina Golyandina (St.Petersburg State University) and Pavel Dudnik (St.Petersburg State University).

Singular spectrum analysis allows automation for the extraction of trends of arbitrary forms. Here we study how the estimation of the signal-rank estimates influences the accuracy of trend extraction and forecasting. It is numerically

shown that the trend estimates and their forecasting are slightly changed if the signal rank is overestimated. If the trend is not of finite rank, the trend estimates are still stable, while forecasting may be unstable. The method of automation of trend extraction includes an important step of improving the separability of time series components to avoid their mixture. However, the better the separability improvement, the larger the forecasting variability, since the noise components become separated and can be similar to trend ones.

[3251] *Monetary policy and bank risk taking: can macroprudential policy interact ?*

Sophie Brana (BSE University of Bordeaux) and Stéphanie Prat (INSEEC Grande Ecole).

Our article focuses on the interaction between monetary policy and macroprudential policy in the context of the transmission channels of monetary policy, and in particular the channel of risk taking that came to light in the wake of the subprime crisis. This is all the more important as macroprudential and monetary policies affect the financial conditions of an economy or zone, which calls for a debate on the relationship between these two types of policy and, more broadly, on the objectives of these types of policy. For that purpose, we use a panel dataset of 126 banks from 20 European countries from 2000 to 2015 and implement both linear static and dynamic multiplicative interaction models. We show that monetary policy and macroprudential variables interact to explain the bank risk-taking channel and that interaction is not linear. The results shed further light on the complex transmission mechanisms of monetary policy and their impact on financial stability.

[3380] *A quantum current sensor for energy flow estimation in smart grids*

Frederik Hoffmann (FH Münster), Ann-Sophie Bülter (FH Münster), Ludwig Horsthemke (FH Münster), Dennis Stiegekötter (FH Münster), Jens Pogorzelski (FH Münster), Peter Glösekötter (FH Münster) and Markus Gregor (FH Münster).

The integration of renewable energy sources, electric vehicles, and storage systems into power grids necessitates advanced control and monitoring systems. Among the key components of these systems are precise current sensors, which are crucial for ensuring the reliable and efficient distribution of electricity. Quantum sensors, particularly those based on nitrogen vacancy centers in diamond, offer a promising approach for high-precision current measurement. This document presents an innovative method for generating microwave signals using direct digital synthesis and field programmable gate arrays to drive multiple signals.

[3460] *Neural Networks Global Models with Context*

Slawek Smyl (Walmart).

Modern ML models used for time series forecasting are global ones, i.e., they are trained across all or at least large portions of all time series in the dataset. However, the preprocessed input is typically based on the history and some other attributes of each time series only, as it is difficult to add inputs derived directly from other time series. When such enrichment is attempted in neural networks models, it is done through Attention or using ideas of Graph Neural Networks. In all cases, it is computationally expensive, typically O(N2) per time step, where N is the number of series. This paper proposes another approach called Context, which is competitive in terms of accuracy, but much less computationally demanding. We provide an example source code in https://github.com/slaweks17.

[3468] *Assessing Global Wildfire Dynamics and Climate Resilience: A Focus on European Regions Using the Fire Weather Index*

Ayat-Allah Bouramdane (LERMA, College of Engineering and Architecture, International University of Rabat).

Wildfires pose significant threats to ecosystems, human safety, and socio-economic stability, necessitating a deep understanding of fire-prone landscapes for effective management. This study globally assesses and focuses on European regions, utilizing the Fire Weather Index (FWI) as a crucial indicator of landscape flammability. Historical FWI data from the European Forest Fire Information System (EFFIS) under the Copernicus Emergency Management Service (CEMS) was analyzed using tools like the Climate Data Store (CDS) API. The results reveal spatial patterns, highlighting regions with heightened wildfire risk and those with reduced fire danger. Southern and Southeastern Europe face elevated danger, driven by factors like high temperatures, low humidity, and reduced precipitation, while Northwestern and Northeastern

Europe exhibit lower risk due to milder conditions. The study also discusses the impacts of these patterns on agrivoltaic systems. The findings inform resilient strategies for policymakers, land managers, and communities, contributing valuable insights for proactive and sustainable wildfire mitigation.

[3689] *Performance of an End-to-End Inventory Demand Forecasting Pipeline Using a Federated Data Ecosystem*

Henrique Moura (University of Antwerpen), Els de Vleeschauwer (University of Ghent), Gerald Haesendonck (University of Ghent), Ben De Meester (University of Ghent), Lynn D'Eer (University of Antwerpen), Tom De Schepper (University of Antwerpen), Siegfried Mercelis (University of Antwerpen) and Erik Mannens (University of Antwerpen).

One of the key challenges for (fresh produce) retailers is achieving optimal demand forecasting, as it plays a crucial role in operational decision-making. Improved forecasts holds the potential to harmonize minimizing waste and avoiding shortages. Different retailers have partial views on the same products, which – when combined – can improve the forecasting of individual retailer's inventory demand. However, retailers are hesitant to share all their individual data. Therefore, we propose an end-to-end graph-based time series forecasting pipeline using federated data ecosystem to predict inventory demand for supply chain retailers. While graph deep learning forecasting has the ability to comprehend intricate relationships, it harmonizes seamlessly with the diverse, multi-retailer data present in a federated setup. The system aims to create a unified data view without centralization, addressing technical and operational challenges, which are discussed through out the text. We test this pipeline using real-world data across large and small retailers, and discuss the performance obtained and how it can be further improved.

[3699] *Gauging Growth Risk in an International Financial Centre: Some Evidence from Singapore*

Hwee Kwan Chow (Singapore Management University).

This paper applies the growth-at-risk framework proposed by Adrian et al. (2019) to Singapore, an international financial centre whereby financial shocks are intermediated away quickly. We gauge near-term risks around growth projections taken from the Survey of Professional Forecasters by accounting for financial stress in both local and global financial markets, as well as worldwide economic uncertainty. The conditioning variables are first linked to future growth through quantile regressions, and the estimated quantiles are fitted with skew t-distributions to produce full predictive distributions. Scenario analysis reveals that greater local financial strain widens the uncertainty of growth outlook, higher global financial stress portends more severe recessions, while increased uncertainty in the economic policy environment diminishes the intensity of economic booms. We also document higher predictive scores and more accurate probabilities of contraction for conditional distributions relative to unconditional ones. Our results highlight the importance of incorporating the influence of foreign vulnerabilities in addition to domestic ones to ascertain near-term growth risk in an international financial centre like Singapore.

[3946] *Chlorophyll-a time series study on a saline Mediterranean lagoon: the Mar Menor case*

Arnau Garcia i Cucó (Universitat Politecnica de Valencia), José Gellida Bayarri (Universitat Politecnica de Valencia), Beatriz Chafer-Dolz (Biologiccropscienc) and José M. Cecilia (Universitat Politecnica de Valencia).

The Mar Menor is the biggest saline lagoon in Europe. Like many lagoons, it has been suffering a eutrophication process. Chlorophyll-a concentration in water is a good estimator of the eutrophication process evolution and could be used to set alarms when there could happen a massive fish-kill. As such, the main objective of this project was to forecast this time series. Several models were applied, but the best-performing were the multivariates, such as LSTM and specially ARIMAX, that took into account the chlorophyll-a concentration relationship with the humidity, the water temperature, the conductivity and the turbidity. However, these predictions contained an error cascade, so univariate models such as Prophet, triple exponential smoothing and ARIMA must be taken into account also.

[4075] *test*

Olga Valenzuela (University of Granada).

test

[4123] *Exploring Regional Determinants of Tourism Success in the Eurozone: An Unsupervised Machine Learning Approach*

George Galanos (University of Piraeus), Charalampos Agiropoulos (University of Piraeus), James Ming Chen (Michigan State University) and Thomas Poufinas (University of Thrace).

This paper presents an initial analysis of the factors influencing tourism success at the NUTS 2 regional level across the Eurozone from 2010 to 2019. Utilizing an extensive dataset that includes economic, demographic, and tourism-specific indicators, we employ unsupervised machine learning techniques, primarily K-means clustering and Principal Component Analysis (PCA), to unearth underlying patterns and relationships. Our study reveals distinct clusters of regions characterized by varying degrees of economic prosperity, infrastructure development, and tourism activity. Through K-means clustering, we identified optimal groupings of regions that share similar characteristics in terms of GDP per capita, unemployment rates, tourist arrivals, and overnight stays, among other metrics. Subsequent PCA provided deeper insights into the most influential factors driving these clusters, offering a reduced-dimensional perspective that highlights the primary axes of variation. The findings underscore significant disparities in tourism success across the Eurozone, with economic robustness and strategic infrastructural investments emerging as key drivers. Regions with higher GDP per capita and lower unemployment rates tend to exhibit higher tourism metrics, suggesting that economic health is a substantial contributor to regional tourism appeal and capacity. This paper contributes to the literature by demonstrating how machine learning can be applied to regional tourism data to better understand and strategize for tourism development. The insights garnered from this study are poised to assist policymakers and tourism planners in crafting targeted interventions aimed at enhancing tourism competitiveness in underperforming regions.

[4195] *Modelling Asymmetric and Time-Dependent Volatility of Bitcoin: An Alternative Approach*

Abdulnasser Hatemi-J (Department of Economics and Finance).

Volatility as a measure of financial risk is a crucial input for hedging, portfolio diversification and the calculation of the value at risk. In this paper we estimate the asymmetric and time-varying volatility for bitcoin as the dominant cryptocurrency in the world market. A novel approach that explicitly separates the falling markets from the rising ones is utilized for this purpose. The empirical results have important implications for investors and financial institutions. Our approach provides a position dependent measure of risk. This is essential since the source of risk for an investor with a long position is the rising prices while the source of risk for an investor with a short position is the falling prices. Thus, providing a separate risk measure in each case is expected to increase the efficiency of the underlying risk management in both cases compared to the existing methods in literature.

[4241] *On the Use of Preprocessing and a Consensus Model to Calculate the Social Security Contributions of Spanish Self-Employed Workers Based on Adjusted Revenue Estimate*

Luis Palomero López de Armentia (Universitat Jaume I), Vicente Garcia (División Multidisciplinaria Ciudad Universitaria de la Universidad Autónoma de Ciudad Juárez) and Jose Salvador Sanchez (Universitat Jaume I).

The method currently used by the Spanish Tesorería General de la Seguridad Social(TGSS) to calculate the monthly contributions paid by the self-employed is based on forecasting the full-year revenues. In the event of observable changes in anticipated revenue, the system allows self-employed workers to adjust the contribution amount. Finally, when the actual revenue is known at the end of the fiscal year, the difference between the actual contributions to be paid and those paid throughout the year is computed, and the difference is accordingly cancelled.

The Spanish company Declarando has developed a methodology that automatically suggests contributions by combining a Simple Moving Average (SMA) model with an adjustment method to compensate for previous deviations, which consists of three main parts: the revenue forecast for the entire year, the monthly contribution inference and an adjustment process. A simulation study revealed that this methodology significantly outperformed the baseline method used by the TGSS. In addition, Exponential Time Series (ETS), SMA, ARIMA, Facebook Prophet, Linear Regression and Deep Forest (DF) were also tested. Based on these results, the present work aims to improve the effectiveness of the

proposed methodology by examining the use of preprocessing techniques and a forecasting method based on a consensus approach.

[4296] *Harnessing Manager Sentiment and Ordinal Regression for Corporate Credit Rating Forecasting*

Petr Hajek (University of Pardubice).

To address the challenges of overfitting and imbalanced class distribution in corporate credit rating forecasting, we propose here a novel model integrating minority class oversampling and ensemble learning tailored for imbalanced ordinal regression. Using a dataset of 1,138 U.S. companies and 35 financial indicators, including FinBERT-based manager sentiment, our approach addresses over-fitting and class imbalance issues common in traditional ordinal class models. Empirical results reveal significant improvements in forecasting accuracy, achieving a 65.3% accuracy rate, a Spearman R of 0.83, and a mean absolute error of 0.39, providing a more effective tool for credit market participants.

[4352] *Developing a forecasting model for allergenic nettle pollen in North-Eastern Croatia*

Edita Stefanic (Faculty of Agrobiotechnical Sciences in Osijek), Sanda Rasic (Faculty of Agrobiotechnical Sciences in Osijek), Pavo Lucic (Faculty of Agrobiotechnical Sciences in Osijek), Marin Lukacevic (Faculty of Agrobiotechnical Sciences in Osijek) and Slavica Antunovic (University of Slavonski brod).

Pollen grains are airborne allergen and the most common trigger of various allergic reactions to sensitive persons, like allergic rhinitis, allergic conjunctivitis, or asthma. These biological particles are microscopic, fine, powdery, male reproductive organs of flowering plants. They are transferred between flowers by wind (anemophily pollen) or insects (entomophily pollen) ensuring plant reproduction. Different plants pollinate at different period of the vegetation season, so the flowering time that affect allergic reaction depends on what sort of pollen they are allergic to. It is estimated that prevalence of pollen caused allergic reaction among European population is approximately 15-20%, having a direct impact on the overall quality of life of sensitive people (Huynen et al., 2003). Moreover, research reports highlighted an increasing trend of people triggering allergic symptoms. Nettle (genus Urticaceae) are among the prolific herbaceous plants in north-eastern Croatia, growing in many natural environments like fields, pastures, roadsides, ditches and even in the urban areas. Even though their pollen is found in the atmosphere with low concentrations, its allergenic proteins can cause significant allergic reaction (Gioulekas et al., 2004). The study was carried out in North-Eastern Croatia, a part of a large sedimentary Pannonian Plain with moderately warm and rainy continental climate. A Hirst-type of sampler (Burkard 7-day volumetric spore trap) was used to collect the airborne pollen from 2005 to 2023. Daily values of were expressed as number of pollen grains per cubic meter of air. Selected taxa (Urticaceae) are ranked among the most allergenic ones (D¬'Amato et al., 2007), therefore having a high interest for the public health in the region. The pollination period was calculated according to Andersen (1991) as 95% of the annual total pollen. The first days producing up to 2,5% of total production, as well as the final days from 97,5 to 100% of total production were removed from the calculations. Furthermore, meteorological parameters from the nearby monitoring station (Croatian Meteorological and Hydrological Service) which correspond to the same time period were used. The following variable were analyzed: air temperature (minimum, maximum and mean), soil temperature (5 and 10 cm) precipitation, relative humidity, solar radiation and wind speed. The objective of this research was the prediction of nettle pollen concentration which is important for allergic persons regarding taking preventive measures to protect themselves from the severity of pollen season. For that purpose, the artificial neural network (ANN) represent a good tool for modelling and forecasting airborne pollen concentrations (Puc, 2012). During the training process, this model enable relation among different income variables with an already registered outcome value. Then, during the validation process, with the new values at the entry, the system provides the exit values. Obtained results show that ANN model provides good results and can be used with accuracy in allergology for taking preventive measures in population suffers from pollinosis. Presence of nettle pollen in the air is ranged from mid-April to mid-October, and characterized with an oscillation in pollen concentrations. The annual total amount of nettle pollen in the air ranges from 1289 in 2007 to 5966 in 2013. Correlation obtained between the selected meteorological variables and the daily pollen concentration showed positive impacts with the maximum and mean air temperature, and the soil temperature at 5 cm. By contrast, a negative correlation was obtained for rainfall, relative humidity and wind speed. The combination of meteorological and pollen data in neural network model showed that after a trial-and-error process, the best values of fitting parameter R2 between real and predicted values of nettle pollen concentration 9-6-3 (nine inputs meteorological values and previous nettle pollen concentration, and the outputs of the predictive values for the next 1, 2 and 4 days). Although, meteorological variability in dataset were detected, this model enables good advance prediction.

[4392] *Signal Detection in High-Noise Time Series Data Using R*

Katerina Tsakiri (Rider University) and Antonios Marsellos (Hofstra University).

This study presents a new methodology for enhancing the detection of stationary sinusoidal signals within noisy time series data, utilizing a suite of advanced computational techniques within the R programming environment. We address the challenge of detecting sinusoidal signals hidden with a noise of up to 20 times that of the signal, corresponding to a very low signal-to-noise ratio (SNR) of 1/20. Our experimental design consists of two distinct paths: one series of tests applied our method-ology for spectral analysis to each noisy dataset, while the other series proceeded without any such enhancements, applying standard analysis packages. The proposed ensemble methodolo-gy includes spectral differentiation, the Kolmogorov-Zurbenko filter application, dataset repe-tition, and zero tapping to refine the frequency analysis process. Comparative analysis with ten well-maintained R packages in CRAN demonstrated our meth-odology's improved performance in identifying the presence of signal frequencies across all tested noise levels. Results indicated a significant improvement in frequency identification, detecting the presence of frequencies at a consistent and linear smaller offset from the true values than conventional approaches. This research underscores the potential of combining advanced filtering, data processing, and spectral analysis techniques to improve signal detection in time series data, particularly in noisy environments.

[4508] *Comparison of Inferential Methods for a Novel CMP Model.*

Yuvraj Sunecher (UNIVERSITY OF TECHNOLOGY MAURITIUS) and Naushad Mamode Khan (UNIVERSITY OF MAURITIUS).

In many real-life instances, time series of counts are often exposed to the dispersion phenomenon while at the same time being influenced by some explanatory variables. This paper takes into account these two issues by assuming that the series of counts follow an observation-driven first order integer-valued moving average structure (INMA(1)) where the innovation terms are COM-Poisson (CMP) distributed under a link function characterized by time-independent covariates. The second part of the paper constitutes of estimating the regression effects, dispersion and the serial parameters using popular estimation methods mainly Conditional Least Squares (CLS), Generalized Method of Moments and a Generalized Quasi-Likelihood (GQL) approach. In addition, this paper also compares the asymptotic relative efficiency of the CMP with the Negative Binomial model in the fitting of MA(1) time series of over-dispersed counts. The performance of these estimation methods is compared via simulation experiments under different levels of dispersion and same is carried out to assess the asymptotic relative efficiency (ARE) of CMP with Negative Binomial. These models are also applied to analyze real-life time series accident data in some regions of Mauritius.

[4692] *Reservoir neural network computing for time series forecasting in Aerospace and beyond. Potential applications to predictive maintenance.*

Juan Manuel Rodríguez Riesgo (Universidad Politécnica de Madrid) and Juan Luis Cabrera Fernández (Universidad Politécnica de Madrid).

Predictive maintenance is a budding field in the aerospace industry with certain caveats that have made implementation sluggish, especially considering the need for data availability and sensor technologies which lack wide commercial implementation industry-wide. This work proposes Reservoir computing optimized using a Grey Wolf algorithm (RC-GWO) to explore the hyperparameters space in search of the configuration best suited to forecast an input signal. It is expected that a generated accurate prediction could be applied to determine whether maintenance is required. The necessary requirements for the system to generate these predictions are defined, with specific suggestions as to how a prediction can be improved through reservoir computing. Three different datasets are studied: the Mackey-Glass equation, the CMAPSS dataset and the power consumption from a resort facility. Results from these diverse set of time series are used to determine certain common rules that improve the quality of the predictions and focus the optimization towards hyperparameter solutions that may allow for a faster approach to predictive maintenance. This research is a starting point to develop methods that could inform accurately on the remaining useful life of a component in aerospace systems.

[4747] *Simulating the Aerial Ballet: The Dance of Fire-Fighting Planes and Helicopters*

Kalle Saastamoinen (National Defence University of Finland/Military Technology), Juha Alander (National Defence University of Finland/Military Technology) and Lauri Honkasilta (National Defence University of Finland/Military Technology).

This study introduces a simulation model to analyze the efficacy of different aerial firefighting strategies in Finland, focusing on the comparative water production capacity and associated costs of firefighting aircraft versus helicopters of varying sizes. By utilizing publicly available data and direct inquiries, the model evaluates the impact of water collection distance on the volume of extinguishing water procured and its costs. The simulation reveals that firefighting aircraft offer a cost-effective solution, particularly when collecting water from distances of thirteen kilometers, where their cost per liter of water aligns with that of smaller helicopters operating closer to the fire zone. The study underscores the importance of precise data input into the calculator, highlighting the potential of aerial firefighting strategies in enhancing wildfire suppression efforts.

[4767] *Evaluation of Economic Interventions in economic blocks during an Economic and Sanitary Crisis*

Carmin Montante (Universidad de Guadalajara) and Clemente Hernandez-Rodriguez (Universidad de Guadalajara).

The purpose of this study is to evaluate the economic interventions that took place during the initial stages of the pandemic in 2020 in the US, Mexico, and Canada. These countries share a free trade agreement that would indicate their willingness to cooperate in economic terms with each other and should adopt similar economic policies considered best practice due to both their shared agreement and the willingness of finding an efficient management of the crisis (Kandogan, 2008). Results show that regardless of each countries individual makeup, neither countries' approach actually presented adequate response towards mitigating the crisis however there are key take aways from this study. First that the Interrupted Time Series models (ITS) show a way to measure interventions and in the future can be used as a measurement to study best policy economic routes. Second, the model adequately measured issues with the policy that were later observed, and lastly this can be seen as a tool to forecast and predict the results of a policy implementation.

[4797] *Estimation of the Gini coefficient using incomplete data*

Szabolcs Kelemen (Department of Physics, Babeş-Bolyai University, Cluj-Napoca, Romania), Máté Józsa (Department of Physics, Babeş-Bolyai University, Cluj-Napoca, Romania) and Zoltán Néda (Department of Physics, Babeş-Bolyai University, Cluj-Napoca, Romania).

In this study, we introduce an innovative methodology for Gini coefficient computation in scenarios where individual-level income data is absent within large, well-defined social systems. Our approach utilizes average income and population metrics at subordinate hierarchical tiers, supported by a specific distribution assumption derived from a novel master equation-based model, the Bíró Néda model. We validate our technique through comparisons with diverse data sources, encompassing comprehensive income data for Cluj county, average income statistics at the settlement level, and percentile and average income data for US, Arizona cities. This approach facilitates a more subtle assessment of inequality, particularly when confronted with datasets providing solely average data at lower hierarchical levels, as opposed to individual-level details. The study underscores the limitations of our model and underscores potential drawbacks in conventional Gini coefficient calculation methods, which may oversimplify aspects of inequality.

[4802] *Statistical analysis of longest dry spells phenomenon in Northern Tunisia based on probability laws*

Majid Mathlouthi (LRSTE at INAT, Tunisia) and Fethi Lebdi (INAT, Tunisia).

It is likely that climate change increases the frequency and duration of droughts. This contribution focuses on dry event analysis from series of observations of the daily rainfall. The approach has been illustrated on a case study catchment localized in Northern Tunisia where the average rainfall is about 600 mm. The dry events are constituted of a series of dry days framed by the rainfall event. Rainfall events are defined themselves in the form a uninterrupted series of rainfall days understanding at least a day having received a precipitation superior or equal to a threshold value. The rainfall events are defined by depth and duration, which are found to be correlated. An analysis of the depth per event conditioned on the event duration has been undertaken. The negative binomial distribution appears the best overall fit for the depth per event. The duration of the rainfall event follows a geometric distribution while that the dry event

follows the negative binomial distribution. The length of the climatically cycle adjusts to the Incomplete Gamma. Event based analysis was used to study of the effects of climate change on water resources and crops and to calibrate precipitation models with little rainfall records.

[4872] *High-Dimensional Mean-Variance Spanning Tests*

David Ardia (HEC Montreal), Sébastien Laurent (Aix Marseille Université) and Rosnel Sessinou (HEC Montreal).

We introduce a new framework for the mean-variance spanning (MVS) hypothesis testing. The procedure can be applied to any test-asset dimension and only requires stationary asset returns and the number of benchmark assets to be smaller than the number of time periods. It involves individually testing moment conditions using a robust Student-t statistic based on the batch-mean method and combining the p-values using the Cauchy combination test. Simulations demonstrate the superior performance of the test compared to state-of-the-art approaches. For the empirical application, we look at the problem of domestic versus international diversification in equities. We find that the advantages of diversification are influenced by economic conditions and exhibit cross-country variation. We also highlight that the rejection of the MVS hypothesis originates from the potential to reduce variance within the domestic global minimum-variance portfolio.

[5058] *Drought forecasting for agricultural insurance payments in grasslands using LSTM neural networks*

Tom Vanwalleghem (Universidad de Córdoba) and Filippo Milazzo (Universidad de Córdoba).

Drought is one of the major meteorological disasters that impacts on agricultural production. In Spanish grasslands, agricultural insurance schemes use remote sensing observations of NDVI or vegetation greenness to quantify payments to farmers. Payments for drought damage due to reduced grassland production occur when NDVI is below historical means, with different loss levels depending on the premium paid. The objective of this paper is to predict these NDVI time series in Mediterranean grasslands using long short-term memory (LSTM) neural networks at different lead times, and compare these to conventional random forest machine learning methods. We use two soil moisture products as predictors, simulated soil moisture values and satellite-based Soil Water Index (SWI) values. The training and validation of the NDVI forecast models were performed from 21/07/2015 to 30/12/2021, using a 70-30% data split. Results show that both products can be used as reliable predictors of permanent grassland in Mediterranean areas. Predictions at 7 days are more accurate and better forecast the negative effect of drought on vegetation dynamics than 30 days.

[5114] *Saving attitudes and behaviors in an emerging economy: Assessing the moderating effect of household size*

Fatima Zahra Bendriouch (Al Akhawayn university in ifrane) and Harit Satt (al akhawayn university in ifrane).

Purpose – The aim of this research is to investigate the mechanisms that explain saving attitudes and behavior among citizens in a majority-Muslim country context. The authors propose and test an integrated model explaining the key individual, social and economic factors explaining saving attitude and behavior.

Design/methodology/approach – Based on data collected from a sample of 618 Moroccan respondents, we empirically tested the proposed conceptual model using a partial least squares (PLS) estimation.

Findings – First, saving attitude has a significant impact on individuals' saving actual behavior; at the same time, our findings supported the substantial effect of all proposed background factors on saving attitude except financial literacy. Finally, our findings confirm the moderating effect of household size, e.g., saving attitude is more likely to convert into a saving behavior for individuals belonging to households of smaller size.

Practical implications – The current research allows a better understanding of the antecedents of individuals' saving attitude and behavior in an emerging country context. Our results provide valuable insights for both financial authorities and institutions. Financial authorities can design improved fiscal programs to assist citizens in organizing their savings in an institutionalized manner and therefore improve their economic conditions.

Originality/value – The current research fills a gap in literature by investigating individuals' saving attitudes and behavior by addressing factors that have been analyzed independently in prior studies. This study combines individual, social and economic factors to explain saving attitude and behavior in a context marked by informal saving habits.

[5133] *Studying LF and HF time series to characterize cardiac physiological responses to mental fatigue*

Alexis Boffet (IMS laboratory), Veronique Deschodt Arsac (University of Bordeaux, IMS Laboratory) and Eric Grivel (Bordeaux INP, IMS Laboratory).

Heart rate variability (HRV) was largely used to evaluate psychophysiological status of Human at rest as well as during cognitive tasks, for both healthy subjects and patients. Among the approaches used for assessing cardiac autonomic control from HRV analysis, biomarkerssuch as the power in low and high frequencies (LF-HF) are often extracted from short-term recordings lasting 2 to 5 minutes. Although they correctly reflect the average psychophysiological state of a subject in situation, they fail to analyse cardiac autonomic control over time. For this reason, we suggest investigating the LF-HF biomarkers over time to identify mental fatigue and determine different physiological profiles. The following step consists in defining the set of parameters that characterise the LF-HF time series and that can be interpreted easily by the physiologists. In this work, polynomial models are considered to describe the trends of the LF-HF time series. The latter are then decomposed into decreasing (d) and increasing (i) parts. Finally, the proportion of the i parts of the polynomial trends of the LF and HF powers over time are combined with classically-used metrics to define individual profiles in response to mental fatigue.

[5226] *Forecasting Hungarian FX Bond Yields*

Dávid Tran (Government Debt Management Agency Pte. Ltd.).

This paper presents a new approach to model and forecast the prices and yields of Hungarian foreign currency bonds. Hungary is a frequent issuer in several currencies (EUR, USD, JPY, CNY), therefore efficient modelling of Hungarian foreign currency yield curves is essential to formulate cost-risk efficient compositions of financing. A longstanding policy of the Hungarian Government Debt Management Agency Pte. Ltd. (AKK) is to keep the foreign currency exposure 100% euro, achieved via swaps. Therefore, this paper focuses on forecasting the yields and prices of Hungarian euro-denominated bonds using bond and cross-currency interest rate swap data.

[5280] *Exploring Different Modelling Approaches to Forecast Acute Respiratory Infections: An Italian Epidemiological Time Series Study*

Riccardo Boracchini (Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Milan, Italy;), Benedetta Canova (Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Milan, Italy;), Pietro Ferrara (Center for Public Health Research, University of Milan-Bicocca, Via Cadore 48, 20900 Monza, Italy.), Elisa Barbieri (Department for Woman and Child Health, University of Padua, Italy.), Pietro Giorgio Lovaglio (Department of Statistics and Quantitative Methods and CRISP, University of Milan-Bicocca, Milan, Italy.), Giovanni Corrao (Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Milan, Italy;), Daniele Donà (Department for Woman and Child Health, University of Padua, Italy), Carlo Giaquinto (Department for Woman and Child Health, University of Padua, Italy.), Costanza Di Chiara (Department for Woman and Child Health, University of Padua, Italy. Penta – Child Health Research, Padua, Italy) and Anna Cantarutti (Department of Statistics and Quantitative Methods, University of Milan-Bicocca, Milan, Italy;).

INTRODUCTION Acute respiratory infections in young children pose a significant global health challenge, leading to high rates of illness and death [1]. They are estimated to be the fourth leading cause of mortality worldwide, particularly impacting children under five [2]. To better predict and prepare for future outbreaks, this study aimed to identify the most effective time series model for forecasting the current 2023/2024 epidemiological season of ARIs. METHODS Data on Acute Respiratory Infections (ARIs) were collected from Pedianet, a network of over 400 Italian family paediatricians that collects clinical and demographic information about their children. We analyzed monthly incidence rates of ARIs from September 2010 to September 2023, following the typical seasonal pattern of these infections. Several forecasting models were compared to predict future ARI cases: Error, Trend and Seasonality (ETS); Seasonal Autoregressive Integrated Moving Average (SARIMA); Unobserved Components (UCM); and Trigonometric seasonality, Box-Cox transformation, ARMA errors and Trend and Seasonal components (TBATS). We evaluated each model's accuracy by examining the residuals and the Mean Absolute Percentage Error (MAPE). Due to public health measures implemented during the COVID-19 pandemic, that significantly reduced the spread of seasonal respiratory viruses, the period between March 2020 and February 2022 was modeled using data from September 2010 to February 2020. Model parameters

were estimated using the in-sample and out-of-sample approach. RESULTS Our analysis included data on over 1.2 million cases of ARIs. On average, there were 1.8 ARI cases per 1,000 children per month, with rates varying between 0.48 and 2.87 across different years. The data showed a clear seasonal pattern, with higher cases early in the year followed by a gradual decline (Augmented Dickey-Fuller Test, p=0.01). We used the specific forecasting model Exponential Smoothing (ETS), chosen based on its accuracy in both in-sample and out-of-sample testing, to predict the pandemic period between February 2020 and February 2022. Overall, our findings suggest that exponential smoothing models as the ETS (MAPE=7.09) and the TBATS (MAPE=7.17) were most effective in predicting future trends in monthly ARI cases compared to other methods (i.e., UCM MAPE=11.42, and SARIMA MAPE=26.45) (Figure 1). CONCLUSION These findings suggest that the exponential smoothing models are preferable in forecasting ARIs historical series in Italy. Despite this, data about the ongoing epidemiological season are crucial to understanding if the residual pandemic effect still impacts the pattern of respiratory infections.

REFERENCES 1) World Health Organization. "Global Health Estimates: Global burden of disease estimates." Web. Access Date (30/01/2024) http://www.who.int/healthinfo/global_burden_disease/estimates/en/index1.html. 2)GBD. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 396, 1204–1222 (2020)

[5286] *Which is the Significance of Environmental Awareness in Energy Investment Decisions? A Study on Portfolio Rebalancing with an Environmental Focus*

Carlos Esparcia Sanchís (UNIVERSIDAD DE CASTILLA-LA MANCHA (UCLM)) and Antonio Diaz (UNIVERSIDAD DE CASTILLA-LA MANCHA).

This empirical investigation delves into the financial performance of energy asset allocation strategies concerning green and brown preferences. Distinct behaviors in investment decisions are presumed between green investors and traditional or brown energy investors. Consequently, optimal portfolio choices are devised through a multi-step process for each energy investor profile and during each monthly rebalancing period. Initially, companies within the energy sector undergo screening based on quartiles of their Environmental (E) ratings. Two categories are then formed, encompassing companies rated above the first quartile (Q1) and those below the third quartile (Q4). Subsequently, a dynamic minimum variance (MV) problem is optimized for each category, employing univariate and multivariate GARCH models. The overall performance of the investment strategies is subsequently evaluated. This contribution to financial literature establishes empirical evidence that green investments not only contribute to reducing the carbon footprint but also reveals a risk-return spread between green and brown energy investments, thereby adding value for environmentally conscious investors.

[5548] *Forecasting methods for road accidents, the case of Bucharest city*

Cristina Oprea (National University for Science and Technology Politehnica Bucharest), Eugen Rosca (National University for Science and Technology Politehnica Bucharest), Ionut Preda (Doctoral School of Transport, National University for Science and Technology Politehnica), Anamaria Ilie (National University for Science and Technology Politehnica Bucharest), Mircea Rosca (National University for Science and Technology Politehnica Bucharest) and Florin Rusca (National University for Science and Technology Politehnica Bucharest).

This paper aims to emphasize the necessity for policy reform, improvements in vehicle design and enhanced public awareness through its projection of future trends in road accidents, injuries and fatalities. The statistical methods that are being used in this paper are the empirical laws of Smeed and Andreassen. The main gap that the researchers identify is the lack of a standardized method-ology with the help of which the appropriate forecasting method can be chosen in the field of road accidents. In the current research, the authors propose such a methodology that can be generalized, being able to be used for any urban agglomeration at the micro and macro level.

[5562] *Impact of interest rate fluctuations on traditional, environmental and stable cryptocurrency returns*

Maria de La O Gonzalez (University of Castilla-La Mancha), Francisco Jareño (University of Castilla-La Mancha) and Jose Mª Almansa (University of Castilla-La Mancha).

The aim of this research is to analyse the sensitivity of traditional, “environmental” (or “green”) and “stable” cryptocurrency returns to changes in nominal interest rates during the period between April 2019 to April 2023. Specifically, this study proposes a model to examine the relation between these three types of cryptocurrency returns and variations in the SP500 index and in the nominal interest rates. The regression technique used to estimate this model is the quantile regression approach to check whether the sensitivity of the returns of these three types of cryptocurrencies is affected by the state of the economy. The main results of this study highlight that stablecoins such as Tether are more sensitive to changes in nominal interest rates at economic peaks and also there is a positive and significant relation between the SP500 index and all the cryptocurrencies in bearish periods. Furthermore, this model shows greater explanatory power in the lower and higher quantiles, showing a U-shaped pattern over the different quantiles and reaching its maximum values in bearish periods of the crypto market (at quantile 0.04), followed by bullish periods (at quantile 0.96). Therefore, the suitability of the quantile regression approach for the estimation of the model is demonstrated, as its explanatory power is more relevant in the extreme quantiles, related to extreme phases of the economy. In conclusion, the sensitivity of the three types of cryptocurrencies’ returns to changes in the SP500 index and in the nominal interest rates is higher in extreme market conditions.

[5569] *Forecasting the realized volatility of stock markets: The roles of jumps and asymmetric spillovers*

Abdelrazzaq Alrababa'A (Yarmouk university).

This paper evaluates the roles of jump and sign asymmetry spillovers in forecasting the realized volatility in a large sample of 20 stock markets. We compare for the first time whether controlling for either the jumps or asymmetric spillovers into the Heterogeneous Autoregressive-Realized Volatility (HAR-RV) model of Corsi (2009) improves the forecasts over 1-day and 5- and 22-days. Before doing so, the spillovers predictors are generated using the Diebold and Yilmaz (2012)’s approach. In analyzing the spillover process, we find that the US stock market remains the main net transmitter of shocks and while China is relatively detached from the spillover linkages, such effects may be transmitted through Hong Kong, which is a significant receiver of shocks. The out-of-sample results reveal that the incorporation of jump spillovers improves forecast performance the most across a range of measures. This is more clearly demonstrated at the 22-day forecasting horizon more notably in Europe, France, Germany, India, and the UK. Lastly, irrespective of the forecasting horizon, performing the predicting stability test uncovers significant improvements in the jumps spillover-based model during periods of notable market stress such as the 2014-2016 oil price crash and COVID-19. Overall, results suggest paying more attention to jumps spillover while constructing international portfolios based on the realized volatility.

[5679] *New Approaches in Demand Forecasting: Tobit Exponential Smoothing with Time Aggregation Constraints*

Diego J. Pedregal (Universidad de Castilla-La Mancha) and Juan R. Trapero (University of Castilla-La Mancha).

This paper introduces the Tobit Exponential Smoothing model (TETS) with time aggregation constraints, providing a flexible framework for handling censored time series data, particularly useful in demand forecasting scenarios. The TETS model, a variant of the Tobit Innovations State Space system, is formulated to address censoring from above in observed output time series, such as sales data, where a known censoring level is imposed. The model's versatility extends to accommodating various modifications, including exogenous regression terms and heteroscedastic noise. Furthermore, the paper proposes a method for incorporating time aggregation into Tobit Innovations State Space systems, enabling the handling of aggregated data without loss of information. The aggregation period is defined, and the state equation is augmented to include observation equations reflecting cumulative aggregated data. This enables the modelling of aggregated censored variables while maintaining the underlying structure of the state space framework. Case studies demonstrate the efficiency of the proposed approach in practical scenarios. In a simulated demand setting, the TETS model effectively captures the true demand, even under stringent censoring conditions. Comparisons with standard methods highlight the TETS model's ability to mitigate bias and produce more accurate forecasts, crucial for inventory management decisions. Moreover, the paper addresses challenges such as the spiral-down effect, where biased estimation leads to suboptimal inventory policies over time. Strategies for mitigating this effect, including replacing lost sales observations and employing the TETS model, are discussed. Results indicate that the TETS model significantly outperforms conventional methods in terms of minimizing lost sales and excess stock, thus offering substantial cost savings. In conclusion, the proposed Tobit Exponential Smoothing model with time aggregation constraints presents a

robust framework for handling censored time series data, particularly in demand forecasting applications. Its flexibility and effectiveness make it a valuable tool for improving inventory management practices and decision-making processes.

[5731] *Annual runoff forecasting through Bayesian causality*

Santiago Zazo del Dedo (IGA Research Group. Salamanca University, High Polytechnic School of Engineering Avila), José Luis Molina González (IGA Research Group. Salamanca University, High Polytechnic School of Engineering Avila), Carmen Patino Alonso (IGA Research Group. Salamanca University, Department of Statistics), Fernando Espejo Almodovar (IGA Research Group. Salamanca University, High Polytechnic School of Engineering Avila) and Juan Carlos García Prieto (CIDTA. Salamanca University, Faculty of Pharmacy).

This contribution is focused on forecasting ability of Bayesian Causality (BC) on annual runoff series. For that, the time series is synthetized through a Bayesian Net, in which the probability propagation over the time is done. The BC analytical ability highlighted the logical temporal structure latent in the hydrological records that defines its general behavior. This allowed to quantify, through a novel dependence matrix, the two temporal fractions within the runoff, one due to time (Temporally Conditioned Runoff) and one not (Temporally Non-conditioned Runoff). Based on this condition-ality, a predictive model is implemented for each temporal fraction, and its reliability is analyzed from a dual probabilistic-metrological perspective.

[5821] *A time series analysis of provincial growth in Spain in the 20th century*

Rafael González-Val (Universidad de Zaragoza & IEB) and Miriam Marcén (Universidad de Zaragoza).

This paper analyses regional growth in Spain at the provincial level (NUTS III regions) during the twentieth century, focusing on the second half of the period that includes the aftermath of the Spanish Civil War (1936-1939), the Franco Dictatorship (1939-1975), and the restoration of the democracy. We consider several economic variables, covering different time periods: population (in relative terms over the total population of the country) and public investment from 1900 to 2011 (with missing data for the investment series during the civil war period), and per capita GDP, employment, and industrial employment in a 5-year basis from 1930 to 2000. All variables are considered in logs. We use time series analysis to test whether each province's path is stationary or follows a unit root process. To test for the presence of unit roots, we run the Augmented Dickey-Fuller test. We find that the null hypothesis of a unit root in the population share is not rejected for most of the provinces in the sample (6% of provinces at the 5% significance level). For public investment, if we focus on the subperiod after the civil war (1940-2011), the null hypothesis of a unit root is not rejected for any province. Regarding the 5-year variables, results also show that a unit root is not rejected for most of the provinces in the sample; the percentage of rejections at the 5% confidence level is 0% for per capita GDP, 14% for employment, and 22% for industrial employment. Nevertheless, the latter results are less reliable because the observations are more spaced in time and, thus, sample size is lower. We also test for unit roots allowing for the presence of one (Perron and Vogelsang, 1992) and two (Clemente et al., 1998) structural breaks, and evidence supporting unit roots persists with slight changes in the percentages of rejection of a unit root in all variables. Finally, we also run Pesaran's panel unit root test. In line with the results obtained using time series analysis, panel results support a unit root in all variables. We run two robustness checks to validate our results. First, we consider alternative datasets of data for per capita GDP, employment, and industrial employment in a 2-year basis from 1955 to 1997. This increases the frequency of the time series and the sample size. Results hold similar. Second, we re-run the time series analysis considering the generalised supremum ADF (GSADF) test statistic for explosive behaviour proposed by Phillips et al. (2011, 2015). This recursive test does not assume exogenous structural breaks. The percentage of unit roots rejected at the 5% level increases for all variables except for the relative population, but the new percentages are still below the 50% of provinces for most variables. Finally, we are also interested in analysing whether the path of those regions that aligned themselves with the losing side of the war was significantly different during the dictatorship period in the aftermath of the Civil War. Those are the regions identified by González-Val and Silvestre (2023) as 'republican' regions. We also interpret the break dates detected in light of the major historical events that took place in the period: the Franco Dictatorship, and the restoration of the democracy.

[5889] *Short-term real-time forecasting during turbulent times. A model for the Spanish GDP after the pandemic*

Matias Pacce (Bank of Spain), Ana Gomez Loscos (Bank of Spain) and Miguel Angel Gonzalez Simon (Bank of Spain).

Following the outbreak of the COVID-19 pandemic, most economic indicators experienced an increase in observed volatility, reducing the accuracy of nowcasting models. In this paper, we propose a new specification for a mixed-frequency dinamic factor model commonly used to nowcast the quarterly GDP growth rate of the Spanish economy --the Spain-STING--. With the aim of improving the predictive capacity of the model after the pandemic, we consider three proposals: (i) the relationship between the indicators and the estimated common factor is now contemporaneous, and not leading for some of the indicators; (ii) the variance of the common component is estimated by a stochastic process to allow it to vary over time; (iii) the set of variables is revised with the aim of including only those that add the most relevant information to the nowcast of the quarterly GDP growth rate after the pandemic. All these three modifications imply a notable improvement in the nowcasting performance during the period after the COVID-19 pandemic, while maintaining the accuracy obtained before it.

[5948] *OWNERSHIP IDENTITY AND STOCK PRICE CRASH RISK: EVIDENCE FROM THE MENA REGION*

Harit Satt (Al Akhawayn university in ifrane).

Purpose: This paper aims to examine the relationship of Ownership Identity and Stock Price Crash Risk in the MENA region to decide whether ownership type investors may influence the stock price crash risk. Methodology: A sample panel of 5126 year-observation for non-financials publicly traded companies in the MENA region was generated to run our analysis covering the period between 2008 and 2021. The GMM -Generalized Methods of Moments- model is applied to investigate the correlation between stock price crash risk and ownership identity. Findings: This study used the Panel data analysis to explore the relationship between the stock price crash risk and ownership identity. It shows that there is a positive correlation between these two variables but with little significance. This research incorporates variables adjusting the results. As an example of those control variables, we find size, leverage, market to book to assist the hypothesis. Originality/Value: The research emphasizes the concept of stock price crash risk jointly with the ownership identity. Its purpose is to understand the if strategic entities ownership and institutional ownership influence the stock price crash risk considering many other measures that are incorporated in each company.

[6019] *A Python Module for Implementing Cointegration Tests with Multiple Endogenous Structural Breaks*

Abdulnasser Hatemi-J (Department of Economics and Finance) and Alan Mustafa (IEEE).

Testing for long-run relationships between time series variables with the short-run adjustments is an integral part of many empirical studies nowadays. Allowing for structural breaks in the estimations is a pertinent issue within this context. The purpose of this paper is to provide a consumer-friendly module that is created in Python for implementing three residuals based cointegration tests with two unknown regime shifts. The timing of each shift is revealed endogenously. The software is easy to use via a Graphical User Interface (GUI). In addition to implementing cointegration tests, the software also estimates the underlying parameters along with the standard errors and the significance tests for the parameters. An application is also provided using real data to demonstrate how the software can be used. To our best knowledge, this is the first software component created in Python that implements cointegration tests with structural breaks.

[6055] *Explaining when Deep Learning models are better for time series forecasting*

Martín Solís (Tecnológico de Costa Rica) and Luis-Alexander Calvo-Valverde (Tecnológico de Costa Rica).

There is a gap of knowledge about the conditions that explain why a method has best forecasting performance than another. Specifically, this research aims to find factors that can influence deep learning models to work better with time series. We generated linear regression models to analyze if 11-time series characteristics influence the performance of deep learning models versus statistical models and other machine learning models. For the analyses, 2000 time series of M4 competition were selected. The results show findings that can help explain better why a pre-train deep learning model is better than another kind of model.

[6063] *Forecasting at Scale: An AutoML Framework for Time-Series Data Challenges*

Siddharth Chatterjee (Mathco), Sourav Banerjee (Mathco) and Divyananda Dileep Aravapalli (Mathco).

Businesses across sectors like retail, finance, and healthcare face a critical scalability challenge as they manage and analyze large time-series datasets. These datasets, essential for informed strategic decision-making, exhibit diverse characteristics—such as smoothness, intermittency, trends, and seasonality—and their volume grows exponentially over time. Without effective scalability, companies struggle with maintaining model performance, skyrocketing server costs, and extended processing times. We present a comprehensive Automated Machine Learning (AutoML) framework designed to integrate multiple forecasting methodologies, from exponential smoothing models to deep learning models, aiming to achieve a harmonious balance between accuracy, computational efficiency, and cost-effectiveness across data types. The framework advances model training efficiency through sophisticated noise reduction and time-series clustering techniques, while employing an innovative hyperparameter search strategy to adeptly balance the trade-off between exploration and exploitation in model tuning. Furthermore, we outline the development of a dynamic, on-demand cloud infrastructure tailored to support the operational needs of this framework. A detailed case study on the M5 dataset, focusing on point-wise forecasting at the item level across varied time intervals, validates our AutoML framework. Our scalability benchmarks highlight our ability to produce highly accurate daily forecasts for 29,000 individual items, navigating millions of potential models in less than 10 hours at a peak cloud cost of 29 dollars per hour. Additionally, we evaluate our method's adaptability to scale based on varying business and resource constraints. This research aims to make a significant contribution to the forecasting domain by offering a scalable, precise, and cost-effective solution that facilitates superior decision-making and resource optimization applicable across different industries.

[6188] *Relevance, Redundancy, and Regularization: Penalized Regression and the Quest for the ℓ_0 Quasi-Norm*

James Chen (Michigan State University).

A linear model’s coefficients ideally support causal inference. Collinearity among variables, however, undermines model interpretation. A large positive coefficient on one variable may be offset by a comparable negative coefficient on a collinear variable.

Collinearity can be remedied through regularized regression. ℓ_2 (Ridge) and ℓ_1 (Lasso) regularization can shrink or eliminate irrelevant and redundant variables. This article also describes ElasticNet (a hybrid of Lasso and Ridge), Bayesian Ridge, automatic relevance determination, and orthogonal matching pursuit. Methods incorporating the ℓ_1 norm can induce sparsity. The resulting vector of nonzero coefficients delivers the ℓ_0 quasi-norm of the model’s causal inferences.

The application of these methods to death rates from cancer in 3,047 counties in the United States strengthens inferences regarding incidence, education, and public health care. Conscious departures from least-squares estimates remain stable and generalizable when regularized regression methods are extended to previously unseen data. These methods should be deployed whenever panel data is analyzed.

[6283] *A New Non-parametric Cross-spectrum Estimator*

Evangelos Ioannidis (Department of Statistics, School of Information Sciences and Technology, Athens University of Economics and Business.).

A new non-parametric estimator of the cross-spectrum of a bivariate stationary time series is proposed, which is non-quadratic in the observations. This estimator is an extension of the Capon-estimator of the spectrum of a univariate time series. The proposed estimator is designed so as to cope with the leakage effect induced by strong peaks of the marginal spectra by utilizing adaptive windowing. We study the asymptotic bias and covariance structure of the proposed estimator and prove a central limit theorem for its distribution. We also obtain a result of independent importance for the consistency rate of the cross-covariance matrix of the two series. The performance of the estimator in comparison to more traditional ones is demonstrated in a simulation study under a model exhibiting extreme characteristics, such as strong peaks in its marginal spectra.

[6304] *EVALUATION OF THE UNIVERSITY OF LAGOS WASTE GENERATION TREND*

Charles A. Mbama (University of Strathclyde), Austin Otegbulu (University of Lagos), Iain Beverland (University of Strathclyde) and Tara K. Beattie (University of Strathclyde).

This study examines waste generation patterns at the University of Lagos (UoL), Nigeria, to inform decision-making towards improving the efficiency of the university's management strategies in-line with Sustainable Development Goal 12, target 12.5 to reduce waste generation through prevention, reduction, recycling, and reuse by 2030. The moving average of the waste generation was studied using time series data. During October 2014 to October 2016 the UoL generated an average of 877.5 tons of waste every month, with the lowest observed value being 496.6 tons and the highest recorded value being 1250.5 tons. The trend result indicates a gradual decrease in the generation of waste over time. There is also a noticeable negative cyclical pattern with seasonal variations, where the highest generation point is observed in March and the lowest point is observed in June, particularly in the latter half of the second quarter, as time progresses. Although, there is a reduction in the amount of waste generated over time, it is crucial to persist in evaluating diverse waste management strategies that could further reduce the amount of waste generated in the case study area.

[6315] *Neural Network Estimator Performance and Electric Energy Network Optimization*

Joseph Moerschell (HES-SO), Fereshteh Jafari (HES-SO) and Charles Praplan (HES-SO).

Neural networks (NN) are more and more used to generate forecasts of power generation and load of electric networks. Production of renewable sources like photovoltaics or wind is weather dependent and cannot be freely steered. Loads are user dependent. Such forecasts can be used to optimize the degree of auto-consumption of an electric network including renewable power generation, to maintain network voltage within defined tolerances, or to maximize revenue from energy exchanges with other networks. If the network contains energy storage, like batteries, water dam or power-to-gas coupling, the size of the energy storage to be provided is an important driver of investment costs. Usually, neural network based forecasts of sun radiance and wind are evaluated statistically over a significant period. e.g. calculating a least means square error (LMSE) of prediction during a day, a week or a year. LMSE depends on the length of history considered for the prediction, and on the future horizon of forecasting. In this study we undertake to link such statistical bench marks to characteristics that are of interest to electric network operators, like the degree of auto-consumption, worst-case voltage excursion and battery size. We consider low-voltage (LV) grids only, with 3-phase 400V nominal voltage. Use cases considered are (a) the network of a family home and (b) a district network connected to one LV/MV (medium voltage) transformer. The network characteristics are not only function of LMSE of NN based predictors, but also depend on the possibility to steer loads in the network, like heating and warm water preparation, or electric car battery charging. We take into account the influence of load steering on resulting network performance. Our aim is to provide a simple design tool to help the network operator to determine the influence of NN based estimator performance on network characteristics. Therefore, as a result of our study, we present a set of equations expressing battery capacity, voltage excursion and degree of auto-consumption as functions of LMSE, lengths of history and of forecasting.

[6415] *Fuzzy portfolio selection based on 5-days recommendations – a case study*

Anna Łyczkowska-Hańćkowiak (WSB MERITO University in Poznan) and Aleksandra Wójcicka-Wójtowicz (Poznań University of Economics and Business).

Investors usually leverage their individual knowledge, experience and intuition as they navigate the complexities of stock investment market. Investment recommendations, often communicated through natural language labels such as "buy", "sell", or "hold", provide investors with a broader perspective. Beyond the subjective personal judgment, the opinions of experts bring an element of objectivity to the decision-making process. The paper proposes the case study comparing 5-day consecutive recommendations established by the procedure basing on fuzzy approach (with the use of trapezoidal oriented fuzzy numbers) with those issued by experts. The analysis covers chosen companies included in DAX index (German Stock Exchange).

[6437] *Cellular Automata Framework for Dementia Classification using Explainable AI*

Siva Manohar Reddy Kesu (M S Ramaiah University of Applied Sciences, Bangalore), Neelam Sinha (Centre for Brain Research, Indian Institute of Science, Bangalore) and Hariharan Ramasangu (Relecura, Bangalore).

Clinical dementia rating scale has been used for analyzing the dementia severity based on cognitive impairments. Many researchers have introduced various statistical methods and machine learning techniques for the classification of dementia severity. Feature importance with deep learning architecture can give a better analysis of the dementia severity. A CA framework has been proposed for the classification of cognitive impairment and LIME has been used for explaining the local interpretability. Feature vectors for healthy and unhealthy classes have been converted to

redistributed CA images. These CA images have been classified using deep learning architecture and promising results have been achieved. GRAD-CAM and LIME explainer have captured the feature importance of cognitive impairment of CA images.

[6482] *Forecasting energy poverty in national energy and climate plans*

Renata Slabe Erker (Institute for Economic Research, Ljubljana), Montserrat González Garibay (European Union Agency for the Cooperation of Energy Regulators), Kaja Primc (Institute for Economic Research, Ljubljana), Darja Zabavnik (Institute for Economic Research, Ljubljana) and Miha Dominko (Institute for Economic Research, Ljubljana).

The study emphasizes the increasing role of long-term planning in the European Union member states' policies on energy poverty and the need for accurate forecasting tools to support it. Machine learning and econometrics are given attention as potential techniques promising improvements in the field, provided certain inherent challenges are overcome. With the National Energy and Climate Plans due for a revision in 2023, energy poverty forecasts have become a necessity, considering they play a pivotal role in EU's strategic planning, even though the specific models and methods are not yet finalized. This study delineates current challenges and the way forward using and critical literature review combined with bibliometric analysis. Utilizing machine learning and econometric methods could aid in efficiently predicting energy poverty by observing patterns of micro and macro-level conditions, which would allow policymakers to devise more effective strategies.

[6494] *The Cassandra Method: Dystopian Visions As A Basis For Responsible Design.*

Sarah Diefenbach (LMU Munich) and Daniel Ullrich (LMU Munich).

Innovative technologies often have unforeseen negative consequences on an individual, societal, or environmental level. To minimize these, the Cassandra method aims to foresee such negative effects by systematically investigating dystopian visions. Starting with the activation of a (self-)critical mindset, the next steps are collecting a maximum number of negative effects, and assessing their relevance. Finally, the envisioned impairments are used to improve the product concepts in a responsible way. This paper broadly sketches the method, its applications during product development and research and reports experiences from an expert workshop.

[6533] *Catalyzing Supply Chain Evolution: A Comprehensive Examination of Artificial Intelligence Integration in Supply Chain Management*

Sarthak Pattnaik (Boston University Metropolitan College), Natasya Liew (Boston University Metropolitan College), Ali Ozcan Kures (Boston University Metropolitan College), Kathleen Park (Boston University Metropolitan College) and Eugene Pinsky (Boston University Metropolitan College).

The integration of Artificial Intelligence (AI) into Supply-Chain Management (SCM) has revolutionized operations, offering avenues for enhanced efficiency and decision-making. AI has become pivotal in tackling various Supply-Chain Management challenges, notably enhancing demand forecasting precision and automating warehouse operations for improved efficiency and error reduction. However, a critical debate arises concerning the choice between less accurate explainable models and more accurate yet unexplainable models in Supply-Chain Management applications. This paper explores this debate within the context of various Supply-Chain Management challenges and proposes a methodology for developing models tailored to different Supply-Chain Management problems. Drawing from academic research and modelling, the paper discusses the applications of AI in demand forecasting, inventory optimization, warehouse automation, transportation management, supply chain planning, supplier management, quality control, risk management, and customer service. Additionally, it examines the trade-offs between model interpretability and accuracy, highlighting the need for a nuanced approach. The proposed methodology advocates for the development of explainable models for tasks where interpretability is crucial, such as risk management and supplier selection, while leveraging unexplainable models for tasks prioritizing accuracy, like demand forecasting and predictive maintenance. Through this approach, stakeholders gain insights into Supply-Chain Management processes, fostering better decision-making and accountability.

[6541] *The Intraday Dynamics Predictor: A TrioFlow Fusion of Convolutional Layers and Gated Recurrent Units for High-Frequency Price Movement Forecasting*

Ilia Zaznov (University of Reading), Atta Badii (University of Reading,), Julian Kunkel (University of Göttingen) and Alfonso Dufour (University of Reading).

This paper presents a novel deep-learning approach for intraday stock price direction prediction, motivated by the need for more accurate models to enable profitable algorithmic trading. The key problems addressed are effectively modeling complex limit order book (LOB) and order flow (OF) microstructure data and improving prediction accuracy over current state-of-the-art models. The proposed deep learning model, TrioFlow Fusion of Convolutional Layers and Gated Recurrent Units (TFF-CL-GRU), takes LOB and OF features as input and consists of convolutional layers splitting into three channels before rejoining into a Gated Recurrent Unit. Key innovations include a tailored input representation incorporating LOB and OF features across recent timestamps, a hierarchical feature learning architecture leveraging convolutional and recurrent layers, and a model design specifically optimized for LOB and OF data. Experiments utilize a new dataset (MICEX LOB OF) with over 1.5 million LOB and OF records. Comparative evaluation against the state-of-the-art model demonstrates significant performance improvements from the TFF-CL-GRU approach. Through simulated trading experiments, the model also demonstrates practical applicability, yielding positive returns when used for trade signals. This work contributes the new dataset, performance improvements for microstructure-based price prediction, and insights into effectively applying deep learning to financial time series data. The results highlight the viability of data-driven deep learning techniques in algorithmic trading systems.

[6602] *Crisis and Youth Inactivity: Central and Eastern Europe During the Financial Crisis of 2008 and the COVID-19 Outbreak of 2020*

Nataša Kurnoga (University of Zagreb, Faculty of Economics and Business), Tomislav Korotaj (University of Zagreb, Faculty of Economics and Business) and James Ming Chen (Michigan State University, College of Law).

This paper analyzes eleven central and eastern European countries after the financial crisis of 2008 and the COVID-19 pandemic of 2020. It investigates heterogeneity in the labor market among selected countries based on youth inactivity, secondary education attainment, and income share of the bottom fifty percent of the population. Hierarchical cluster analysis with Ward's method and k-means clustering generated diverse cluster solutions. Comparative analysis of four-cluster solutions for 2008 and 2020 showed multiple changes in cluster composition. Joint groupings of geographically and historically close countries, such as the Baltics, the former Czechoslovakia, and the former Yugoslav republics of Croatia and Slovenia, were identified for 2008. Lithuania emerged as a singleton in 2020. Youth inactivity, educational levels, and income inequality reveal the status of youth in central and eastern Europe during these crises.

[6812] *From Upside and Downside Risks to Resilience: The ESG Imperative in Energy Market Sustainability*

Carlos Esparcia (Universidad de Castilla-La Mancha), Mariya Gubareva (ISEG – Lisbon School of Economics and Management, SOCIUS/CSG, Universidade de Lisboa) and Francisco Jareño Cebrián (Universidad de Castilla-La Mancha).

In this study, we use a time-varying parameter vector autoregression model to explore how changes in ESG ratings and in ESG sub-pillar scorings impact the portfolio risk exposure in the energy sector. Our research sheds light on the skewed risk profile often overlooked by traditional measures, emphasizing the significance of upside and downside risks. By analyzing environmental, social, and governance dimensions, we provide novel insights on their varying impacts on different risk levels in energy portfolios. Our findings highlight the importance of considering both upside and downside risks, with ESG rating changes playing a transmitting role while environmental, social, and governance ratings act as receivers. Notably, ESG rating changes have a prolonged effect on downside risk exposure, echoing investor risk aversion.

[7014] *Modelling the Daily Concentration of Airborne Particles Using 1D Convolutional Neural Networks*

Ivan Gudelj (Faculty of Electrical Engineering, Computer Science and Information Technology Osijek), Mario Lovrić (Centre for Bioanthropology, Institute for Anthropological Research, Zagreb) and Emmanuel Karlo Nyarko (Faculty of Electrical Engineering, Computer Science and Information Technology Osijek).

This paper focuses on improving the prediction of the daily concentra-tion of the pollutants, PM10 and nitrogen oxides (NO, NO2) in the air at urban monitoring sites using 1D convolutional neural networks (CNN). The results show that the 1D CNN model outperforms the other machine learning models (LSTM and Random Forest) in terms of coefficients of determination and abso-lute errors.

[7031] *Multicriteria Forecast Combination with Non Linear Programming*

Oscar Generoso Gutiérrez (Higher School of Computer Science, University of Rey Juan Carlos, Madrid, Spain), Clara Simón de Blas (Higher School of Computer Science, University of Rey Juan Carlos, Madrid, Spain) and Ana Elisabeth García Sipols (Higher School of Experimental Sciences and Technology, University of Rey Juan Carlos, Madrid, Spain).

Improving predictions computation for time series analysis is still a challenge. Finding a method that combines the benefits of different methodologies is still an open problem. Besides the very efficient prediction combinations techniques proposed, there is still a lack in procedures that considers jointly error measures combinations and model constraints. In this work, we propose a new forecast combination procedure based in multicriteria methods that allows the assignment of weights to different error measures in the objective function and the incorporation of constraints. A real case is presented in the pharmaceutical industry for a probiotic product sale to illustrate the performance of the proposal. This method is capable to consider different error measures and non distance based errors, is enriched by the consideration of constraints that contemplate desirable properties of the solution and is robust with respect to different time series characteristics such as trends, seasonality, etc. Results shows similar accuracy to the best known forecasting methods to the date.

[7207] *Forecasting stock market dynamics from market cap time series of firms through fluctuating selection*

Hugo Fort (Universidad de la República).

Evolutionary economics has been useful to explain the nature of the processes of innovation, as well as providing useful heuristics for applied research. However, quantitative tests are rare and in general fail to capture the observed dynamics of firms in real markets. A main problem is how to estimate the fitnesses of companies. We present a quantitative implementation of fluctuating selection to explain and forecast the dynamics of firms' market capitalizations (MC). We start by proposing a recipe to estimate the financial fitness of a company from its MC time series, which is completely analogous to the Malthusian fitness of evolutionary biology. Using Malthusian fitnesses in terms of MC implies that all companies, through their stocks, irrespective of their industry or sector, are competing for investors' money. These fitnesses, in turn, allow to forecast the companies' proportions of total MC through a natural selection equation. We validate the method using the daily MC of public-owned Fortune 100 companies across 2000-2021. This Fluctuating Selection from Market Capitalization (FSMC) formula is particularly accurate in forecasting the MC proportions of companies. It outperforms the geometric random walk stochastic benchmark for all forecasting instances across validation periods ranging from a month to a year.

[7238] *Modeling the future of hydroelectric power: a cross-country study*

Farooq Ahmad (Department of Statistical Sciences, University of Padua), Livio Finos (Department of Statistical Sciences, University of Padua) and Mariangela Guidolin (Department of Statistical Sciences, University of Padua).

To solve the problems of energy security, stability, and climate change, countries are realizing energy transition programs and turning to renewable energy sources like hydroelectric power, solar energy, wind power, and geothermal energy. Much of the scientific interest so far has focused on studying the growth of more recent energy sources like solar and wind energy, while less interest has been dedicated to more traditional renewable sources, such as hydroelectric power. On the other hand, according to the International Energy Agency (2023), hydropower currently generates more electricity than all other renewable technologies combined and is expected to remain the world's largest source of renewable electricity generation into the 2030s. Therefore, it will continue to play a critical role in decarbonizing the power system and improving system flexibility. However, it is also acknowledged that, without major policy changes, global hydropower expansion is expected to slow down this decade. This contraction results from slowdowns in the development of projects

in China, Latin America, and Europe. At the same time, increasing growth in Asia Pacific, Africa, and the Middle East partly offset these declines. Increasingly erratic rainfall due to climate change is also disrupting hydro production in many parts of the world. Starting from these pieces of evidence, in this paper we analyze the role of hydroelectric data in countries' energy transition using the approach of innovation diffusion models. The literature on the application of innovation diffusion models for energy growth processes has been developing significantly in recent years and we aim to contribute to this direction. To do so, the time series of hydroelectric energy generation are considered and modeled with the Bass Model, BM, (Bass, 1969). Based on the parameters of the estimated models, the analysis seeks to identify some regularities and similarities in the studied countries. Whenever the simple BM is not enough suitable to describe the observed data, some convenient extensions of it are used for the purpose, such as the GGM (Guseo and Guidolin, 2009). This extension proves satisfactory in all those cases where the observed data depart from the classical bell-shaped structure typical of the BM. To provide predictions on the future evolution of hydroelectric energy, we also propose some further extensions to the BM and GGM that, by relaxing the typical assumption of these models of a saturating behavior of the growth process, can allow more flexibility to data description and provide more reliable predictions for the future evolution of hydroelectric power.

[7263] *Do Professional Forecasters Believe in Uncovered*

Mengdi Song (University College Dublin) and Constantin Burgi (University College Dublin).

No, not according to our data. Using a unique data set, we run panel regressions to test whether professional forecasters believe in uncovered interest rate parity (UIP). Specifically, we test whether the interest rate expectations for individual forecasters are in line with their exchange rate expectations using the UIP condition. This new approach allows us to test directly whether forecasters believe in UIP. We find that professional forecasters generally do not believe in UIP across a range of currencies and horizons. Given the prevalence of the UIP condition in our international macro models, these results reiterate the importance of finding the drivers for these deviations.

[7371] *Sentiment Dynamics and Volatility: a Study Based on GARCH-MIDAS and Machine Learning*

Gianmarco Vacca (Università Cattolica del Sacro Cuore) and Luigi Riso (Università Cattolica del Sacro Cuore).

This work investigates the relationship between investor sentiment and volatility of stock indexes. A sentiment proxy is constructed via a machine learning approach from the consumer confidence indexes of four countries. Granger causality tests highlight the influence of sentiment on volatility. This impact is quantified via GARCH-MIDAS models that, retaining variables in their sampling frequency, allow the estimation of the long-run volatility without information loss. Sentiment is finally used to predict long-run volatility. Thus, further insights into the relationship between investor sentiment and return volatility are provided, helping investors to stabilize the former and contain its effect on market uncertainty.

[7470] *Calibration Free Current Measurement with Integrated Quantum Sensor*

Jens Pogorzelski (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Jonas Homrighausen (Department of Engineering Physics, FH Münster—University of Applied Sciences), Ludwig Horsthemke (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Dennis Stiegekötter (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Frederik Hoffmann (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Ann Sophie Bülter (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences), Markus Gregor (Department of Engineering Physics, FH Münster—University of Applied Sciences) and Peter Glösekötter (Department of Electrical Engineering and Computer Science, FH Münster—University of Applied Sciences).

This paper presents the application of a compact and fully integrated LED quantum sensor based on NV centers in diamond for current measurement in a busbar. The magnetic field measurements from the sensor are directly compared with measurements from a numerical simulation, eliminating the need for calibration. The sensor setup achieves an accuracy of 0.28 % in the measurement range of 0 − 30 A DC. The integration of advanced quantum sensing technology with practical current measurement demonstrates the potential of the sensor for applications in electrical and distribution networks.NV center · ODMR · current sensing · simulation · calibration fre

[7501] *Multiscale Hierarchical Forecasting of Urban Footfall*

Tom Komar (Newcastle University) and Philip James (Newcastle University).

Footfall patterns in urban environments show regularity at various scales. Quality of forecasts can improve performance of urban operations in a smart city context. Broader patters can be missed when forecasting footfall is done separately for related locations at a single temporal scale. Applying hierarchical modelling provides us with enhanced consistency, reduction of overfitting and improved accuracy by regularising more detailed predictions with higher-level trends, and reduced complexity by assigning the tasks of capturing trends apparent on different levels to separate models.

[7504] *The role of healthcare quality and macrofinancial factors determining hospital mortality in Spain. Dynamic panel data evidence*

Paula Ortega Perals (University of Almeria), Salvador Cruz Rambaud (University of Almeria) and Javier Sánchez García (University of Almeria).

Hospital mortality is a topic of major interest for every government as patient mortality rates are a sign of the efficiency of the health system. Among the factors most affecting mortality, the quality of care has been shown to be of great relevance due to its consequences over economic decisions. The objective of this paper is to analyze the effect of quality of care on hospital mortality and how the economic circumstances determine such relationship. The methodology used is a regression analysis with dynamic panel data by using the Generalized Method of Moments (GMM) estimator in two steps, employing a panel of seventeen Spanish regions for the period 2007-2019. The results reveal that public debt, consumer price index, and readmissions are the variables most influencing hospital mortality. The main contribution of this paper is to empirically test the relationship between quality of care, certain macrofinancial factors with hospital mortality. The main limitation is the lack of data availability for the years 2020 and 2021. For further research, an international perspective can be included to find differences between developed and developing countries, and whether cultural and social determinants provide different results.

[7542] *Evaluation of clustering methods for the unsupervised classification of ground deformation time series*

Jhonatan Rivera Rivera (Instituto Geológico y Minero de España (CN IGME-CSIC)), Héctor Aguilera (Instituto Geológico y Minero de España (CN IGME-CSIC)), Marta Béjar-Pizarro (Instituto Geológico y Minero de España (CN IGME-CSIC)), Carolina Guardiola-Albert (Instituto Geológico y Minero de España (CN IGME-CSIC)), Alejandra Staller (ETSI Topografía, Geodesia y Cartografía, Universidad Politécnica de Madrid (UPM)), Pablo Ezquerro (Instituto Geológico y Minero de España (CN IGME-CSIC)), Riccardo Palamà (Geomatics Research Unit, Centre Tecnológic Telecomunicacions Catalunya (CTTC/CERCA)) and Oriol Monserrat (Geomatics Research Unit, Centre Tecnológic Telecomunicacions Catalunya (CTTC/CERCA)).

One of the most significant products of Synthetic Aperture Radar Interferometry (InSAR) is the generation of Deformation Time Series (DTS). Its analysis enables the identification, characterization, and monitoring of ground deformation caused by both natural processes (e.g., landslides or earthquakes) and anthropogenic activities (mining settlements or subsidence due to groundwater extraction). However, InSAR datasets are often voluminous, requiring semi-automated clustering or classification methods to reduce computational cost and facilitate interpretation of results. In this study, we propose a two-step unsupervised classification methodology: (1) identification of stable temporal series using thresholds of minimum average velocity for detecting ground movements from satellite data and (2) identification and characterization of deformation patterns using clustering techniques for temporal series, such as K-means and K-shape executed for different distance metrics. Additionally, we conduct a performance evaluation of four classification approaches, including the proposed methodology and three recently published approaches: statistical classification approach by Berti et al. (2013), dimensionality reduction (PCA and UMAP) and unsupervised classification (K-means and HDBSCAN). To carry out this evaluation, we collect thousands of DTS from the European Ground Motion Service (EGMS) in different areas of Spain and Italy representing various deformation processes and temporal behaviors. Finally, we discuss the advantages and disadvantages of each InSAR classification approach in terms of computational efficiency, clustering/classification coherence, method scalability, and result interpretability. This work has been developed thanks to the pre-doctoral grant for the Training of Research Personnel (PRE2021-100044) funded by MCIN/AEI/10.13039/501100011033. It is also supported by “FSE invests in your future” within the framework of the

SARAI project “Towards a smart exploitation of land displacement data for the prevention and mitigation of geological-geotechnical risks” PID2020-116540RB-C22, funded by MCIN/AEI/10.13039/501100011033.

[7729] *JET: Fast Estimation of Hierarchical Time Series Clustering*

Phillip Wenig (Hasso Plattner Institute, University of Potsdam, Germany), Mathias Höfgen (Rolls Royce Deutschland Ltd. & Co. KG) and Thorsten Papenbrock (Philipps University of Marburg).

Clustering is an effective, unsupervised classification approach for time series analysis applications that suffer a natural lack of training data. One such application is the development of jet engines that involves numerous test runs and failure detection processes. While effective data mining algorithms exist for the detection of anomalous and structurally conspicuous test recordings, these algorithms do not perform any semantic labelling. So data analysts spend many hours connecting the large amounts of automatically extracted observations to their underlying root causes. The complexity, amount and variety of extracted time series make this task hard not only for humans, but also for existing time series clustering algorithms: These algorithms either require training data for supervised learning, cannot deal with varying time series lengths, or suffer from exceptionally long runtimes. In this paper, we propose JET, an unsupervised, highly efficient clustering algorithm for large amounts of variable-lengths time series. The main idea is to transform the input time series into a metric space, then apply a very fast conventional clustering algorithm to obtain an effective, but rather coarse-grained pre-clustering of the data; this pre-clustering serves to subsequently estimate the more accurate but also more costly shape-based distances of the time series and, thus, enables JET to apply a highly effective Hierarchical Clustering algorithm on the entire input time series collection. Our experiments demonstrate that JET is highly accurate and much faster than its competitors.

[7792] *Energy prices impact on inflation per household category in Greece*

Panagiotis Delis (Department of Economics, University of Patras, Rio, Greece), Stavros Degiannakis (Department of Economic and Regional Development, Panteion University of Social and Political Sciences, Athens, Greece) and George Filis (Department of Economics, University of Patras, Rio, Greece).

Since the early 2021 the European economies are suffering from unprecedented inflation pressures, which are primarily fuelled by price spikes in the energy markets, due to the Russian invasion in Ukraine. Extreme inflation due to energy price spikes can cause significant loss of income, which can subsequently cause reduction in household consumption with detrimental effects on well-being, redistribution of income and economic development. There are studies that show that the distribution of inflation rates across different households are significantly different and they tend to vary over time. More specifically, they find that lower-income households face remarkably higher inflation, which provides evidence that different policy measures should be taken across households. The aim of the current study is twofold, namely, to investigate the impact of energy prices on disaggregated inflation by household type, as well as to assess the predictive ability of the former to the latter inflation rate. The research focuses on Greece for the period 2009 – 2022 (168 monthly observations). More specifically, in our study, we use monthly data from the Household Budget Survey (HBS) of the Hellenic Statistical Authority to generate the inflation rates by household types. We further use a set of explanatory variables, which are grouped into three blocks, namely, the Greek-specific (domestic) factors, the global factors and the energy specific factors. For the energy-specific factors we consider the Brent crude oil prices, the EU natural gas prices and wholesale electricity prices. We start our analysis from the in-sample estimation by employing a Dynamic Model Averaging (DMA) approach, which is based on combinations of time-varying parameter (TVP) model specifications. In this regard, we develop different model specifications that take into consideration solely the domestic, global or energy factors, as well as specifications that combine all these factors. Based on the Dynamic Model Averaging (DMA) methodology, we assess both the probability that the best model that explains household inflation includes the energy variables, as well as the magnitude of the effect of the energy factors, so to identify any cross-household inflation differences. Tentative results show that models which include the energy factors present a persistently higher probability to explain the household inflation rates. Even more, we find that there are important cross-household differences, with households at the lower income levels, as well as with more dependent members (children under 16 years old, 1 parent with children under 16 years old, etc.) to receive a more severe impact from the energy prices. By implementing an out-of-sample forecasting framework, we find that the predictive gains from the energy prices are higher for households that are impacted more severely by these prices, as these were identified from the in-sample analysis.

[7840] *Forecasting Electricity Prices in Times of Distress using Bid Data*

Aitor Ciarreta (The University of the Basque Country, UPV/EHU) and Blanca Martínez-Gonzalo (Universidad Complutense de Madrid).

We propose to forecast hourly prices in the Iberian Day-ahead electricity market using publicly available information on demand and supply offers for the period 2021 to 2023. We use a novel iterative procedure where we fit fractional polynomial functions and logistic functions. Then, we use the estimated coefficients to predict prices using different forecasting models. Our approach shows how modeling using functional forms fitted on observed curves improves forecast not only of prices but also quantities. The former are relevant for market makers who need to hedge against excesive price volatility. The latter for small market participants who behave as price-takers.

[7967] *Detecting Change Points in Temporally Correlated Data: An Environmental Time Series Application*

Magda Monteiro (University of Aveiro) and Marco Costa (University of Aveiro).

Changes in time series often arise due to various factors like financial crises, government regulations, natural disasters, and technological advancements. These structural changes, referred to as change points, can significantly influence the behavior and characteristics of time series data, leading to shifts in trends, seasonality, volatility, or other important features. Failure to detect these change points can result in misleading analyses and conclusions. Therefore, the investigation of change point detection methods has been extensively explored in recent decades. Initially developed within the framework of independent and identically distributed random variables sequences, most of these methods relied on the assumption of Gaussian distribution. Procedures such as CUSUM-type and likelihood ratio tests for changes in mean and/or error variance have been prominently featured in the literature. However, when the distributional form of the phenomenon is unknown, non-parametric methods are employed for studying change points. This study aims to compare various approaches to detect multiple change in long environmental time series. The series are modeled using state space models and change point detection is performed by applying two approaches: one using the innovations series, from parametric and non-parametric perspectives, and the other using Kalman filter smoothers that represent the trend prediction of the series under analysis, from a parametric perspective.

[8043] *Income inequality and credit cycles: booster or anchor?*

Alessandra Centinaio (LIUC Università Cattaneo), Fausto Pacicco (LIUC Università Cattaneo), Andrea Venegoni (LIUC Università Carlo Cattaneo) and Massimiliano Serati (LIUC Università Cattaneo).

The 2008 Financial Crisis reminded economists that financial frictions could be relevant in understanding modern business cycle dynamics, and since then several studies uncovered the role of credit as one of the main drivers of financial instability. Credit expansion during booms, indeed, increases the likelihood of financial crises as well as the severity of the following recessions (Jordà, Schularick, and Taylor, 2013). While the existing literature recognizes the centrality of the “credit-driven household demand channel” in explaining the real effects of credit expansions, the mechanisms underlying the build-up of credit are still debated. There are several sources of credit expansion, especially when dealing with open economies (Sufi and Taylor, 2021). Among them, the rise in income inequality has recently emerged as a promising explanation (Baziellier, Héricourt, and Ligonnière, 2021; Mian and Sufi, 2018; Sufi and Taylor, 2021). Our study sits in this burgeoning literature. As the related empirical literature has so far failed to reach a unanimous agreement regarding the extent and direction of the relation between inequality and credit expansion (Bazillier and Hericourt, 2017), we aim to investigate the impact – if any – of income inequality on household credit, and whether household credit reacts differently depending on the source and nature of the income inequality change. Specifically, we explore if, given an increase (decrease) in income inequality, household credit reacts differently when it comes to a rising (shrinkage) in the top decile of the income distribution or in the bottom decile. To assess the effects of income inequality variations on household credit, we employ a Bayesian Factor-Augmented Mixed-Frequency VAR (FA-MFVAR) model in the 10 main Euro Area (EA henceforth) countries in terms of GDP. The EA economies, indeed, provide an interesting setting as they have witnessed increasing income and wealth inequalities in the last decades. Our contribution to the existing literature is twofold. First, we consider changes at both the top and the bottom of the income distribution. As Atkinson and Morelli (2011) and Baziellier, Héricourt, and Ligonnière (2021) point out, income distribution as a whole needs to be accounted for since inequality can arise from variations in the tails as well as in the middle of the distribution. This distinction is important because the implications for credit expansion may differ depending on the source of inequality increase (decrease). Most of the existing studies, however, focus on the top income share changes (i.e. top 10%, top 5%, top 1%). Secondly, since data on inequality are usually available on an annual basis, which does not fit the requirements of a VAR analysis, we employ a Mixed-Frequency model. This allows us to

observe how the variations simulated on annual variables affect the target dimensions quarterly or even monthly. Results confirm that the source of income inequality variation influences the reaction of household credit, which varies depending on whether the shift happens in the top or the bottom decile of income distribution. Following a rise in top incomes, household credit volumes increase. However, there are exceptions, such as Germany, France, and Spain, where this does not hold. On the other hand, an increase in income inequality due to a shift in the lowest decile of the distribution leads to a more fragmented picture, with some countries (i.e. Germany, Spain, and Ireland) experiencing higher household credit while others lower. These findings shed light on the importance of setting up different reaction strategies to income distribution modifications according to their nature.

[8118] *SPECTRAL CHARACTERISTICS OF STRONG GROUND MOTION TIME SERIES FOR LOW TO MEDIUM SEISMICITY REGIONS WITH DEEP SOIL ATOP DEEP GEOLOGICAL SEDIMENTS—AN EXAMPLE OF THE CITY OF OSIJEK, CROATIA*

Silva Lozancic (Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek), Borko Bulajic (Faculty of Technical Sciences, University of Novi Sad), Gordana Pavic (Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek), Ivana Bulajic (Institute of Transportation CIP, Belgrade, Serbia) and Marijana Hadzima-Nyarko (Faculty of Civil Engineering and Architecture Osijek, Josip Juraj Strossmayer University of Osijek).

Time series analysis plays a crucial role in understanding earthquake ground motion' temporal patterns, trends, and dynamics, ultimately aiding in more reliable strong ground motion prediction, risk assessment, and disaster management efforts. We examined features of both horizontal and vertical strong motion for deep soil sites above deep geological deposits, in low-to-medium seismicity zones. We selected Osijek, Croatia—a city on the right bank of the Drava River and in the southern portion of the Pannonian Basin—for our case study since previous research revealed that the city's building stock is very vulnerable. Using only the strong-motion time series (i.e. surface ground acceleration time histories) that were recorded in the region of the north-western Balkans, new empirical regional equations were derived for spectral and peak ground acceleration attenuation. The total of 436 horizontal and 218 vertical strong strong-motion accelerogram components from 112 earthquakes with magnitudes ranging from 3 to 6.8 that were recorded in the northwest Balkans are included in the strong-motion database that was utilized to develop the predictive equations. The proposed equations simultaneously account for the influence of deep geological conditions and local soil characteristics.

[8321] *Minimal Reservoir Computing*

Haochun Ma (Allianz Global Investors, LMU Munich), Davide Prosperino (Allianz Global Investors) and Christoph Räth (German Aerospace Center (DLR)).

Reservoir computers are powerful machine learning algorithms for predicting nonlinear systems. Unlike traditional feedforward neural networks, they operate on small training data sets, use linear optimization, and therefore require minimal computational resources. However, the traditional reservoir computer uses random matrices to define the underlying recurrent neural network and has a large number of hyperparameters to optimize. Recent approaches show that the randomness can be removed by running regressions on a large library of linear and nonlinear combinations constructed from the input data and their time lags and polynomials. For high-dimensional and nonlinear data, however, the number of such combinations explodes. Here, we show that a few simple modifications to the traditional reservoir computer architecture, which further minimize computational resources, lead to significant and robust improvements in short- and long-term predictive performance compared to similar models, while requiring minimal sizes of training data sets. We quantitatively evaluate the predictions using the maximum Lyapunov exponent, the correlation dimension, and the forecast horizon. For certain parameterizations, we find that the predictions are accurate for more than 10 Lyapunov times. We also find that, under certain conditions, ordinary least squares regression directly on the embedded data can predict the long-term climate of chaotic systems.

[8361] *Testing for Nonlinear Cointegration under Heteroskedasticity*

Christoph Hanck (University of Duisburg-Essen) and Till Massing (University of Duisburg-Essen).

See PDF

[8495] *Towards an automatic tool for resilient waterway transport: the case of the Italian river PO*

Maria Luisa Villani (ENEA - CR Casaccia), Ebrahim Ehsanfar (Fraunhofer Institute for Material Flow and Logistics), Sohith Dhavaleswarapu (Fraunhofer Institute for Material Flow and Logistics), Alberto Agnetti (Agenzia Interregionale per il Fiume Po, AIPO), Luca Crose (Agenzia Interregionale per il Fiume Po, AIPO) and Sonia Giovinazzi (ENEA - CR Casaccia).

Improved navigability can enhance inland waterways transportation efficiency, contributing to synchro-modal logistics and promoting sustainable development in regions that can benefit from the presence of considerable waterways. Modern technological solutions, such as digital twins in corridor management systems, must integrate functions of navigability forecast that provide timely and reliable information for safe trip planning. This information needs to account for the type of vessel and for the environmental and geomorphological characteristics of each navigation trait. The paper presents a case study, within the EU project CRISTAL, focusing on the Italian river PO, of which the navigability forecast requirements of a digital twin are illustrated. Preliminary results to deliver navigability risk information have been obtained. In particular, the statistical correlation of water discharge and water depth computed from historical data, suggests that efficient forecast models for navigability risk given some water discharge forecast might be built. To this aim the LSTM (long-short-term-memory) technique has been used on the same data to provide models linking water discharge and water depth predictions. Future work involves fur-ther testing these models with updated real data and integrating outcomes with climatic and infrastructure management information to enhance the accuracy of the risk information.

[8519] *Modeling and forecasting hourly ozone concentration using functional data approach*

Ismail Shah (Department of Statistical Sciences, University of Padua, Italy).

Air pollution, especially ground-level Ozone, poses severe threats to human health and ecosystems. Accurate forecasting of ozone concentrations is essential for reducing its adverse effects. This study aims to use the functional time series approach to model ozone concentrations, a method less explored in the literature, and compare it with traditional time series and machine learning models. To this end, the ozone concentration hourly time series is first filtered for yearly seasonality using smoothing splines that lead us to the stochastic (residual) component. The stochastic component is modeled and forecasted using a functional autoregressive model (FAR), where each daily ozone concentration profile is considered a single functional datum. For comparison purposes, different traditional and machine learning techniques, such as ARIMA, VAR, NNAR, RF, and SVM, are also used to model and forecast the stochastic component. Once the forecast from the yearly seasonality component and stochastic component are obtained, both are added to obtain the final forecast. For empirical investigation, data consisting of hourly ozone measurements for the Los Angeles ranges from 2013 to 2017 is used, and one-day-ahead out-of-sample forecasts are obtained for a complete year. Based on the evaluation metrics, such as R^2, RMSE, and MAE, the forecasting results indicate that the FAR outperforms the competitors in most scenarios, with the SVM model performing the least favorably across all cases.

[8563] *Transmission of oil price volatility to MENA stock markets: A comparative study between oil exporters and importers*

Khalil Mhadhbi (Faculty of Economic Sciences and Management of Nabeul).

This paper explores the transmission of volatility from Brent oil price evolution to the stock returns of 7 MENA countries, encompassing 3 importers and 4 exporters, after excluding 4 initial countries using the ARCH test. Employing the GARCH-BEKK estimation method, we detect this transmission from January 2008 to September 2022. The results reveal significant volatility persistence across 6 stock markets, with 3 importer countries and 3 exporters. These findings align with Shiller's theory, indicating high volatility in financial markets. Tunisia's stock market shows sensitivity to oil market developments, while the Omani market demonstrates volatility transfer from Brent oil prices. However, Morocco's market exhibits resilience, with no significant transmission from international oil prices. Exporting countries, except the UAE, display significant and positive coefficients, indicating volatility transmission. The study suggests further research into underlying mechanisms and recommends policymakers and investors implement strategies to mitigate volatility effects. Advanced modeling and behavioral insights can enhance risk management strategies.

[8587] *A System for Efficient Detection of Forest Fires Through Low Power Environmental Data Monitoring and AI*

Ipek Uremek (Department of Electrical and Electronic Engineering University College Cork), Paul Leahy (Department of Energy Engineering University College Cork Cork) and Emanuel Popovici (Department of Electrical and Electronic Engineering University College Cork).

Abstract. Forest fires are a significant environmental concern worldwide, leading to immense ecological damage. The 2019 Amazon wildfires witnessed a dramatic increase in fires across Brazil, Bolivia, Paraguay, and Peru, with Brazil's INPE (National Institute for Space Research) reporting a 77% rise to over 80,000 fires, including 40,000 in the Amazon. This led to 906,000 hectares of the Amazon biome being lost to fires. The disaster highlighted the urgent need for improved fire management strategies, including advanced warning, early detection, and rapid response to protect communities and minimize damage [1]. Extinguishing fires proves challenging due to various factors, including limited firefighter access to forested areas, the influence of wind on fire behaviour, and delayed intervention from specialized firefighting teams. A potential solution lies in the timely detection of fire outbreaks facilitated by modern wireless systems. Today, the planet's best surveillance system is the one made with artificial geostationary satellites, which instantly signals the outbreak of a fire. NASA's satellite tools are often the first to detect fires that are burning in distant regions and new fire locations are sent directly to field managers around the world within a few hours of satellite travel [2]. However, despite these advancements, the devastation caused by wildfires underscores the inadequacy of current detection and response measures in effectively mitigating their impact. The urgency to detect forest fires is paralleled by the growing emphasis on environmental data acquisition, which is gaining momentum in efforts to monitor ecosystems, track CO2 levels and air quality, optimize machine learning for forecasting, integrate system dynamics and forecasting models. Weather is everything that affects our daily lives, society, economy, and environment. To what extent weather will affect our lives depends on various factors such as type of event, timing, duration, severity, location, etc. According to Weather Analytics, 33% of worldwide GDP is affected by the weather [3]. In response to these environmental challenges, we propose a versatile system that not only focuses on detecting forest fires using low-cost sensors but also serves broader environmental monitoring purposes leveraging lightweight AI algorithms and large weather forecast models for predictive analysis and forecasting complex/big data. The proposed system presents a sensor technology to gather critical environmental data, offering insights into various ecological parameters beyond fire detection, such as temperature fluctuations, humidity levels, wind direction, atmospheric CO2 concentrations, etc. This system incorporates large weather forecast models to predict environmental conditions for specific locations by using an AI model. This approach enables the detection of fires as anomalies in environmental data time series, enhancing the accuracy and timeliness of fire detection. Despite the challenges of deploying a sensor network in forests, including ensuring reliability and managing energy resources efficiently, this system provides fire prevention efforts while supporting vital environmental conservation objectives. By balancing immediate fire detection needs with long-term environmental monitoring goals, our solution represents a forward-thinking approach to environmental management and protection. This paper presents an alternative approach to environmental monitoring utilizing LoRa (Long Range IOT) technology coupled with BME280 and wind sensors for data collection and machine learning algorithms for predictive analysis. By integrating a network of low-cost, high-efficiency sensors, this system is designed to provide a comprehensive view of the forest ecosystem, enabling not just early fire detection but also contributing valuable data for anomalies in environmental data time series enhancing climate forecasting capabilities. The methodology involves deploying LoRa Dragino sensor nodes in remote forest areas to gather environmental data, which is then transmitted to a central server via the LoRa Multitech gateway and The Things Network (TTN). Enhancements to this system include the ability to collect and analyse location-specific data, enabling targeted monitoring and predictive analysis of environmental conditions. By identifying anomalies in the data, such as sudden temperature spikes, humidity drops, or unusual wind patterns, the system can alert to potential fire outbreaks or environmental disturbances, offering a proactive approach to disaster prevention and environmental anomalies. Through the collection of comprehensive, location-specific environmental data and the identification of anomalies, this approach offers a forward-thinking solution to environmental monitoring and protection. By reference to IoT, understood as a network of physical devices that are connected to the Internet and capable of communicating with each other, it is breaking into different industries and markets. One of the applications is in the field of Smart Cities, to monitor different variables such as traffic, pollution, water systems, etc [4]. To support the Internet of Things (IoT), recent developments in communication technologies have given rise to a Low-power Wide-area Network (LPWAN) [5]. LPWAN provide a wide range of communication, consume a small amount of power and prove to be very appropriate to applications requiring low data rates of transmission [6]. Among the LPWAN technologies, LoRa stands out for its exceptional ability to facilitate communication over vast distances while maintaining low power consumption, making it highly suitable for environmental monitoring tasks, including forest fire detection. Sensors equipped with LoRa can gather a wide range of environmental data, including temperature, humidity, wind speed, and more, from vast and inaccessible areas. The integration of BME280 sensors across a network of nodes allows for the collection of detailed, real-time

environmental data. These sensors, when deployed throughout a forest, can continuously monitor conditions of forest fire risk by sending this data to a central system for analysis, they enable real-time monitoring and early warning capabilities. This technological synergy allows for a comprehensive and dynamic approach to forest management and fire prevention, offering a crucial tool in the fight against the devastating impact of wildfires, and also environmental analysis. A decision tree-based machine learning algorithm is developed and evaluated for its effectiveness in predicting forest fires, offering a nuanced approach to environmental monitoring. This model employs a hierarchical structure of decisions based on an extensive set of variables, including temperature, humidity, wind speed, and other environmental factors, not only to predict the fire risk with high precision but also for broader environmental monitoring applications. The sequential application of the decision tree model exemplifies a strategic deployment of data mining techniques in our research, highlighting our primary analytical strategy. Data mining, as an iterative process of extracting hidden patterns from large data sets and a critical component of the knowledge discovery process, consists of a collection of automated and semiautomated techniques for modelling relationships and uncovering hidden patterns in large data repositories [7]. This framework is designed to analyse the environmental data, identifying potential fire hazards, and anomalies in the time series with high accuracy. Also, local sensors will provide feedback on the accuracy of meteorological models, enhancing the system's predictive capabilities. The paper thoroughly explores the technical deployment of the sensor network, data transmission challenges, communication protocols, low power adjustments, and the integration of sensor data with a predictive model. The calibration customization of the decision tree algorithm to suit the specificities of the environmental data collected through LoRa technology is discussed in depth. Moreover, the paper navigates through the challenges encountered during the implementation phase, including sensor node energy optimization, data reliability under varying environmental conditions, and the scalability of the network to cover large forest areas effectively. The potential impact of this integrative approach is profound, offering a scalable, efficient solution for forest fire prediction that could significantly mitigate the frequency and severity of fires through timely interventions. Addressing the challenge of data reliability in variable environmental conditions was one of the key focuses of our study. Weather changes, such as shifts in temperature, humidity, and wind speed, can affect sensor accuracy, making it crucial for our system to adapt for consistent data collection. To ensure prediction reliability, we implemented a prediction method. This method helps correct anomalies, normalize data, and enhance the trustworthiness of the information our sensors collect. Additionally, local sensors provide feedback on the accuracy of meteorological models, further improving our system's ability to offer accurate forest fire predictions despite the unpredictable nature of weather patterns. By applying these strategies, we improved our system's ability to provide accurate predictions and analyses, despite the unpredictable nature of weather patterns. In conclusion, this study presents a comprehensive approach that combines LoRa technology with machine learning algorithms to enhance our capabilities in forest fire detection and broader environmental monitoring. By utilizing an integrated network of low-cost sensors and advanced predictive analytics, this system not only addresses the urgent need for effective forest fire management but also contributes significantly to the collection and analysis of environmental data. This methodology enables the identification of anomalies within environmental time series, facilitating early warnings and proactive interventions for various ecological threats like forest fires. Through the strategic deployment of sensor technology and the innovative application of data mining techniques, this research provides an alternative approach to environmental analysis, forecasting complex/big data. Future work will explore the integration of satellite communications to enhance coverage in areas lacking GSM connectivity, aiming to broaden the system's applicability and ensure more inclusive environmental surveillance.

[8685] *Examining the performance of some stock market indices: before and after the Russia Ukraine conflict (2021-2023)*

Marta Tolentino (Universidad de Castilla-La Mancha), María del Valle Fernández (Universidad de Castilla-La Mancha) and Laura Prieto (Universidad de Castilla-La Mancha).

The Russian invasion of Ukraine has marked a turning point in the global economy, which has been reflected in the main financial markets and, therefore, in their most representative indices. A lot of studies have analysed the contagion effect (Toribio, 2012), that is, the way in which events in one stock market affect other markets. The main objective of this paper is to analyse and understand in depth the impact of the conflict between Russia and Ukraine on some of the most representative stock market indices worldwide. This objective was chosen because of the influence of geopolitical events on financial markets. To achieve this objective, the following specific objectives are addressed. First, to identify trends and patterns of behaviour in the stock markets under study, which will provide a clearer picture of how these markets react to a major international crisis. The second objective is to analyse the behaviour of stock market indices in the period from one year before the start of the conflict to one year after. The Russian invasion began on 24 February 2022, so the period selected for the study is 24 February 2021 (one year before the start of the invasion) and 24 February 2023 (one year after the start of the conflict). Similarly, the results will be analysed in two sub-periods: the first will comprise the

pre-conflict period and the second the conflict period. The representative indices to be analysed are IBEX35 (Spain), the main indicator of the Spanish stock market; S&P 500 and Nasdaq100 (United States), Nikkei225 (Japan), CAC 40 (France), DAX 40, the most important stock market index of the Paris Stock Exchange in France; the DAX 40 (Germany); FTSE MIB (Italy); HANG SENG, (Republic of China); PFTS, (Ukrainia) and MOEX, (Russia). The weekly quotes of the indices were obtained from the "Investing" platform. Once the data had been downloaded, the weekly returns of each of the indices were calculated in an Excel spreadsheet and then imported into R, specifically its interface known as R-Studio. The results show that the stock index whose returns are most highly correlated with the returns of the Russian and Ukrainian indices, over the full period, is the FTSE MIB (Italy). This could be indicative of the economic and trade relations between these financial markets before the war broke out. However, when analysing the results in the two sub-periods mentioned above, it is observed that this trend only occurs in the pre-war period. This would suggest that the conflict is having a considerable impact on the relationship between these markets, which could be due to factors such as economic sanctions, commodity price volatility and geopolitical uncertainty, among others.

[8767] *Robust Multitaper Tests for Detecting Frequency Modulated Signals*

Benjamin Ott (Queen's University), Glen Takahara (Queen's University) and Wesley Burr (Trent University).

In this paper, we propose a new test for the detection of polynomial modulated signals, as well as an aggregated test based on the new test. The test statistics are developed under the multitaper spectral framework and are designed to be robust to the choice of the multitaper order K. The proposed tests are based on a modification of a previous test statistic in Ott et al.(2023), denoted by F4. We review the F4 test and discuss some of its shortcomings, then propose our 2 tests. We illustrate performance via simulations and apply our tests to SoHO GOLF optical solar time series.

[8850] *Modeling algorithms for anomaly detection based on real-time sensor data: The case of a Desalination Plant*

Ron Weitzman (JCE - Azrieli Academic College of Engineering) and David Gavbriel Pinto (Azrieli Collage of Engineering).

The challenge in anomaly detection is to find patterns in data that do not conform with a priori expected behavior, and to apply the best corrective actions. In today's industrial era, where large volumes and a variety of sensor data are collected, we face an opportunity to identify nontrivial, valid, and novel patterns in data stored in production databases. Pattern detection has many known applications and is now receiving special attention in the field of predictive maintenance both in academia and industrial communities. This work focuses on early detection of abnormal patterns in data collected from high-pressure pumps (HPP) in desalination plants. Specifically, data was collected from vibration sensors of these pumps. The desalination process is highly dependent on the availability and reliability of its HPPs. Therefore, early detection of abnormal patterns is of great value to the operations managers who may apply timely corrective actions to prevent equipment breakdown. In this work, we propose a novel approach which combines time series models combined with statistical process control (SPC) for early and efficient anomaly detection. The approach proposed uses unsupervised learning due to the lack of sufficient samples of labeled training data. The developed models are now tested in three Israeli desalination plants. Preliminary results show very high sensitivity in the detection of abnormal vibration patterns. From these abnormal patterns, a maintenance prioritization scheme was developed. The early abnormal pattern detection exposed early-stage damage in several pumps and damaged components were replaced, thus preventing breakdown. According to the plants' experts, this early detection and prioritization is expected to lower the plant's downtime and maintenance costs.

[8995] *Decomposing the Sri Lanka yield curve using principal component analysis to examine the term structure of the Interest rate*

Sanjeewa Dayarathne (University of Moratuwa) and Uthayashnger Thayasiwam (University of Moratuwa).

In this study, we delve into the dynamics of the Sri Lankan government bond market, building upon prior research that focused on the application of prin-cipal component analysis (PCA) in modelling sovereign yield curves. Our analysis encompasses data spanning from January 2010 to August 2022. The study applied several PCA variants such as multivariate PCA, Randomized PCA, Incremental PCA, Sparse PCA, Functional PCA, and Kernel PCA on smoothed data. Kernel PCA was found to be explaining the majority of the variation associated with the data. Findings reveal that the first principal component accounted for a substantial 97.69% of the variations in yield curve movements, 2nd PCA 1.88%, 3rd

0.42%. These results align with previous research, which generally posits that the initial three principal components tend to elucidate around 95% of the fluctuations within the term structure of yields. Our results question the empirical findings, which state the 1st PCA represents the longer tenor of the yield curve. In Sri Lanka, in-stead, the 1st PCA represents the 3-year bond yields. It may be because of the liquidity constraints in underdeveloped frontier markets, where longer tenor yields do not react fast enough to reflect the movement of the yield curve. 2nd PCA represents the slope of the yield curve which is the yield difference of a 10-year T-Bond and 3 months T-Bill. 3rd PCA which represents the curvature of the yield curve attributed to 2x3 years T-Bond yield – 3 months T-bill – 10-year T-Bond.

[9025] *Missing data imputation for heat pump performance analytics*

Sandhya Patidar (Heriot-Watt University), Chakron Dechkrut (Heriot-Watt University), Andrew Peacock (Heriot-Watt University), Abhinanda Roy (Heriot Watt University) and Yigit Duveroglu (DV Energy).

Missing data imputation for heat pump performance analytics Dr Sandhya Patidar;

Missing data is one of the most common and a critical problem occurs in the application of data science approaches, specifically in time series modelling, where complete dataset is often a prerequisite for developing a robust and reliable model [1]. Missing data could lead to biased results and if not infilled using appropriate approaches could significantly effect the statistical significance and prediction power of the model/algorithms. In this context, missing value imputation is an essential pre-processing step that involves estimating or infilling missing data points with some predicted or substitute values. Recently, a range of statistical and machine learning based approaches has been proposed for data imputation, each with some advantages and limitations [2]. Nevertheless, it is important to assess the suitability of approach, which depends on type of data sets, ratio of missingness and not just on the problem under investigation but also on the purpose of investigation. In addition, there are several other important aspects that should be considered in the selection/developing of missing data imputation strategy. For developing a statistically significance missing data imputation strategy, it is crucial that method is suitable for the problem under investigation, accounts in any data related constrains, and most essentially preserves the key statistical properties of the dataset. Additionally, in case of time series data imputation, underlying patterns, and temporal dynamics (e.g. autocorrelation, stationarity, etc.) should remain unaltered.

This paper presents an overview of widely applied missing data imputation approaches and conduct an intensive comparative analysis of a logical algorithm (developed by authors) underpinned by domain knowledge along with some widely used machine learning based data imputation methods [3]. The paper will conduct the investigation (focusing on large gaps) using a large portfolio of dataset (182 dwelling) collected as part of the Electrification of Heat (EoH) demonstration project, funded by the Department for Business, Energy and Industrial Strategy (BEIS), to understand the feasibility of a large-scale rollout of heat pumps across the UK [4]. Each dwelling contains approximately 1 year of data, available at a temporal resolution of 2 mins. Table 1 provide a quick statistical percentile distribution of percentage (%) of missing values across nine variables for 182 dwelling (See attached paper pfd).

Table 1: Percentile distribution of missing values for nine variables measuring heat pump operation (Air source Heat pumps). Percentiles are measure for a dataset of 182 dwellings.

The paper will review and scope for some key algorithms to assess their suitability for different variables which includes):

1. Interpolation with scipy.interpolate.griddata; ['linear','nearest','cubic'] [5] 2. Regression a. Simple Linear Regression, Multiple Linear Regression, Polynomial Regression, Support Vector Regression (SVR), Decision Tree Regression, Random Forest Regression with scikit-learn [6] b. Gaussian Processes with scikit-learn [7] 3. Signal processing - Spectral interpolation with scipy.fftpack [8]- using Fourier transform and inverse average-frequency domain pre/post-missing data scipy.signal.detrend(np.real(scipy.fftpack.ifft(avg_pre_post))) 4. Long Short-Term Memory (LSTM) - Recurrent neural networks (RNN) with Tensorflow [9]and keras [10]- used to prediction time-series data 5. Artificial Neural Network (ANNs) could be used with unix timestamp converted to 24hr timestamp (hh.fraction(mm)) **include other data features dimensions and potential for further implementation 6. Convolution Neural Networks (CNNs) like AutoEncoder and Generative Adversarial Network (GAN) are some of the advanced approaches that also allow inclusion of other data features dimensions and need systematic approach for practical implementation [11]. The performance of various imputation approaches are comparatively assessed using key statistical indicators involving discrete percentile distribution, continuous density distribution, and pattern analysis of large gaps in missing data using autocorrelation and partial-autocorrelation properties. References

[1] S. v. Buuren, Flexible Imputation of Missing Data, 2nd Edition, A Chapman & Hall Book, 2018. [2] T. Thomas and E. Rajabi, “A systematic review of machine learning-based missing value imputation techniques,” Data Technologies and Applications, Vols. ISSN: 2514-9288, 2021. [3] S. Patidar, D. P. Jenkins, A. Peacock and A. Lotfipoor, “Missing data imputation for community energy demand modelling,” in 3rd IBPSA-Scotland Conference: USim2022: Urban Energy in a

Net Zero World - University of Strathclyde , Glasgow, United Kingdom, 2022. [4] "Electrification of Heat - Home Surveys," [Online]. Available: https://es.catapult.org.uk/report/electrification-of-heat-home-surveys-and-install-report/. [5] scipy.interpolate.griddata. [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.griddata.html. [6] https://scikit-learn.org/stable/supervised_learning.html, "Supervised learning," [Online]. Available: https://scikit-learn.org/stable/supervised_learning.html. [7] G. Processes. [Online]. Available: https://scikit-learn.org/stable/modules/gaussian_process.html. [8] scipy.fftpack.ifft. [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.fftpack.ifft.html. [9] TensorFlow, "Working with RNNs," [Online]. Available: https://www.tensorflow.org/guide/keras/working_with_rnns#rnn_layers_and_rnn_cells. [10] Keras, "LSTM layer," [Online]. Available: https://keras.io/api/layers/recurrent_layers/lstm/. [11] A. Lotfipoor, S. Patidar and D. P. Jenkins, "Transformer network for data imputation in electricity demand data," Energy and Buildings, Vols. 300, 113675, 2023.

[9117] *Energy Efficiency Evaluation of Frameworks for Algorithms in Time Series Forecasting*

Sergio Aquino-Britez (Universidad de Granada), Pablo García Sánchez (Universidad de Granada), Andrés Ortiz (Universidad de Málaga) and Diego Aquino-Britez (Universidad de Granada).

In this study, the energy efficiency of time series forecasting algorithms is addressed in a broad context, highlighting the importance of optimizing energy consumption in computational applications. The purpose of the study is to compare the energy efficiency and accuracy of algorithms implemented in different frameworks, specifically Darts, TensorFlow, and Prophet, using the ARIMA technique. The experiments were conducted on a local infrastructure. The Python library CodeCarbon and the physical energy consumption measurement device openZmeter were used to measure energy consumption. The results show significant differences in energy consumption and algorithm accuracy depending on the framework and execution environment. We conclude that it is possible to achieve an optimal balance between energy efficiency and accuracy in time series forecasting, which has important implications for developing more sustainable and efficient applications. This study provides valuable guidance for researchers and professionals interested in the energy efficiency of forecasting algorithms.

[9345] *Smart Belay Device for Sport Climbing - an Analysis about Falling*

Heiko Oppel (Ulm University of Applied Sciences) and Michael Munz (Ulm University of Applied Sciences).

As sport climbing is a comparatively safe sport, injuries can be more severe as in other sports. With the increasing popularity over the last decades, it started to attract people unknown to the sport. Especially as a belayer, one is responsible for the safety of the climber. This makes it necessary to address the issue of safe belaying. By using an instrumented belay device, we were able to record a climbers fall into the rope with varying configurations. In combination with machine learning algorithms, we could then categorize the recordings and further extract the information about the severity of the fall for the climber.

[9369] *Big Data Techniques Applied to Forecast Photovoltaic Energy Demand in Spain*

Jorge Tapia García (Department of Computer Science and Artificial Intelligence, University of Granada, Granada, 18014, Andalucía, Spain), Luis G. Baca Ruiz (Department of Software Engineering, University of Granada, Granada, 18014, Andalucía, Spain), David Criado Ramón (Department of Computer Science and Artificial Intelligence, University of Granada, Granada, 18014, Andalucía, Spain) and María del Carmen Pegalajar Jiménez (Department of Computer Science and Artificial Intelligence, University of Granada, Granada, 18014, Andalucía, Spain).

Renewable energies play an important role in our society's development, addressing the challenges presented by climate change. Specifically, in countries like Spain, technologies such as solar energy assume a crucial significance, enabling the generation of clean energy. This study addresses the critical need to accurately predict Photovoltaic (PV) energy demand in Spain. By using the data collected from the Spanish Electricity System, four models (Linear Regression, Random Forest, Recurrent Neural Network, and LightGBM) were implemented, with adaptations for Big Data. The LR model proved unsuitable, while LGBM emerged as the most accurate and timely performer. The incorporation of Big Data adaptations amplifies the significance of our findings, highlighting the effectiveness of LGBM in forecasting PV energy demand with both accuracy and efficiency.

[9372] *Measuring the efficiency of introducing businesses` digitalization elements over time in relation to their performance*

Jarmila Horváthová (Faculty of Management and Business, University of Prešov) and Martina Mokrišová (Faculty of Management and Business, University of Prešov).

The introduction of digitalization elements into the life of companies is significant in terms of achieving better economic results. The aim of the research was to determine the technical efficiency, as well as the change in efficiency and technological change of the digital transformation of com-panies in EU countries in relation to their performance. The Malmquist index was used to measure these parameters over time. The results of the research indicate the significance of the dynamic measurement of the efficiency of digital transformation. It brings interesting results that point to the importance of evaluating the efficiency of the use of already established elements, as well as eval-uating the introduction of new technological changes.

[9384] *Forecasting the impacts of smoking prevalence scenarios from 2022 to 2050*

Dana Bryazka (Institute for Health Metrics and Evaluation; University of Washington), Marissa Reitsma (Stanford University), Natalia Bhattacharjee (Institute for Health Metrics and Evaluation; University of Washington), Stein Emil Vollset (Institute for Health Metrics and Evaluation; University of Washington) and Emmanuela Gakidou (Institute for Health Metrics and Evaluation; University of Washington).

Background

Smoking is the leading behavioral risk factor for mortality in adults over age 20, causing 205 million deaths and 4.70 billion years of life lost (YLLs) since 1990 worldwide. While global smoking prevalence has declined 29% since 1990, the number of smokers has increased by 5.28% and progress has slowed for many countries. Despite legislation proposed in several countries to eliminate tobacco initiation among youth to achieve a tobacco-free generation, none have been implemented yet. Understanding what might occur if current smoking trends persist, as well as what is possible under smoking prevalence reduction scenarios provides critical evidence to lawmakers about the costs of inaction. Our study presents a novel framework for forecasting the health burden of tobacco use to more accurately translate the impacts of smoking prevalence on mortality. Using this framework, we forecast life expectancy at birth and all-cause and cause-specific YLLs under three smoking prevalence scenarios.

Methods

In this analysis, we utilize the Institute for Health Metrics and Evaluation's Future Health Scenarios platform to forecast the effects of three smoking prevalence scenarios from 2022 to 2050 for 204 countries and 32 smoking-associated diseases and injuries by sex and 5-year age group. This platform forecasts major "drivers of health" to 2050 and uses these, along with past mortality trends, to forecast mortality for 365 diseases and injuries. Drivers of health include 68 risk factors, interventions such as vaccine coverage, and covariates such as the education, income, and fertility. Estimates of these drivers of health, as well as fertility and migration, were obtained from the Global Burden of Disease Study (GBD) 2019, for every location, age-group, and sex, from 1990 to 2019.

The full time series from 1990 to 2050 for each independent driver is used to forecast cause-specific mortality rates to 2050. Future age- and sex-specific population is computed for each location based on future all-cause mortality, fertility, and net migration. Finally, YLLs are computed using mortality forecasts and the GBD reference life table. Importantly, all measures, such as mortality, population, life expectancy, and YLLs are forecast for each scenario, allowing us to capture the effects of changing risk factor exposure levels on mortality rates, population age structure, and population size.

Most risk factors in our platform are forecast using past estimates of a summary measure of risk exposure. However, the complexity inherent in smoking exposure required a modified approach, wherein we estimate future smoking prevalence and exposure intensity, and then use these factors to directly compute the summary measure. Using this novel framework, we created three smoking prevalence scenarios: the reference, immediate elimination, and accelerated progress scenarios. Our reference scenario forecasts what might occur if past risk factor and other health driver trends continue into the future, and if relationships between risk factors and causes of disease stay the same. To construct this scenario, we forecast current and former smoking prevalence from 2022 to 2050 using an ensemble model based on projections of past annualized rates of change (ARC). This ensemble model is comprised of six sub-models with unique recency weights, allowing trends in recent years to be weighted more heavily in the final model. The final model is a weighted average of these six sub-models, with weights determined by out-of-sample predictive validity on a 10-year holdout. The immediate elimination scenario reflects the maximum health benefits possible from smoking elimination, by forecasting what will occur if smoking prevalence were eliminated starting in 2023. The accelerated progress scenario

offers a benchmark for countries planning to implement policies to meet tobacco-free targets within the coming decades. In this scenario, we linearly decrease smoking prevalence between 2022 and 2050 and assume no new smoking initiation. Together, these scenarios illustrate the significant health gains that could be achieved by 2050 if decisive measures are taken globally to put an end to the tobacco epidemic.

Results

Our reference scenario suggests that if past trends persist, global age-standardized smoking prevalence will decline by 25.9% among males and 30.0% among females from 2022 to 2050. In contrast, age-standardized smoking prevalence declined by 30.2% and 40.0% between 1990 and 2022. While we have seen an increase of 8.40 years of additional life expectancy at birth between 1990 and 2022, these gains would also slow, resulting in just 5.00 years of additional life expectancy globally by 2050 relative to 2022. Under this scenario, there would be a cumulative 50.0 billion YLLs over this time period.

If smoking were eliminated in 2023, we forecast 2.17 billion fewer cumulative YLLs by 2050, and an additional 1.50 years of life expectancy for males and 0.50 years for females, relative to the reference scenario. Just over 50% of the YLLs averted under the immediate elimination scenario would be averted under the accelerated progress scenario. Under that scenario, life expectancy would increase by 1.12 years for males and 0.43 years for females relative to the reference scenario. Of those YLLs averted under the accelerated progress scenario, 65.3% of which would be in lower- and middle-income countries. In this scenario, cumulative YLLs due to lung cancer and chronic obstructive pulmonary disease will be reduced by 13.5% and 11.4%, respectively.

Conclusion

This analysis represents an exciting advancement in our ability to translate forecasted changes in smoking prevalence to smoking attributable mortality. Our results suggest that smoking will continue to cause enormous health burden in the coming decades. Accelerated progress in tobacco control is necessary to maximize the prevention of premature, smoking-attributable deaths.

[9458] *Modelling Explosive Non-stationarity of Ground Motion Shows Potential for Landslide Early Warning*

Michael Manthey (University of Melbourne), Guoqi Qian (University of Melbourne) and Antoinette Tordesillas (University of Melbourne).

This work applies the rarely seen explosive version of auto-regressive modelling to a novel practical context - geological failure monitoring. This approach is more capable than standard ARMA or ARIMA methods in that it allows the underlying data process to be explosively non-stationary, which is often the case in real-world slope failure processes. We develop and test our methodology on a case study consisting of high quality (in-situ) line-of-sight radar displacement data from a slope which undergoes a failure event. Specifically, we first optimally estimate the characteristic roots of the auto-regressive processes underpinning the displacement time-series preceding the failure at each monitoring location. We then construct a Wald-type statistic based on the maximum characteristic root estimate and use this statistic to test the explosiveness of the corresponding true characteristic roots. Concluding that a true characteristic root becomes explosive at some significance level implies the underlying displacement process is explosively non-stationary, and hence local geological instability is suspected at this significance level. We found that the actual location of failure (LOF) was identified well in advance of the time of failure (TOF) by flagging those locations where explosive root(s) were concluded from the Wald-type maximum root test. This statistical feedback model for ground motion dynamics presents an alternative and/or addition to the velocity threshold approach to early warning of impending failure.

[9480] *Digital Boxcar Averager on a Microcontroller for Pulsed ODMR Measurements of NV Centers*

Dennis Stiegekötter (FH Münster), Ludwig Horsthemke (FH Münster), Ann-Sophie Bülter (FH Münster), Frederik Hoffmann (FH Münster), Jens Pogorzelski (FH Münster), Peter Glösekötter (FH Münster) and Markus Gregor (FH Münster).

Renewable energy sources account for a steadily growing share of energy production. Weather-dependent energy sources make it difficult to develop forecasting models due to a lack of data collection. Reliable power measurement in power supply grids with easy handling therefore plays a major role. This work focuses on the development of a current sensor based on NV centers in diamonds for the application in power grids. A complete demonstrator for the measurement of magnetic resonances of NV centers is presented, with the potential to detect current-induced magnetic

fluxes of electrical conductors. For this purpose, a digital boxcar averager is implemented on a microcontroller for signal processing of pulsed optical detection of magnetic resonances. An approach to optimize noise suppression is proposed, and it is shown that the ODMR contrast is robust against the timing of the boxcar window on the fluorescence response of the NV diamond.

[9491] *Legendre polynomial modelling-based permutation entropy to analyse encrypted Time series*

Meryem Jabloun (PRISME laboratory).

Securing sensitive data, particularly healthcare information transmitted through cloud services, is mandatory. However, preserving the privacy of medical data during analysis poses a significant challenge. Therefore, it is of paramount importance to develop tools capable of evaluating the robustness of encryption techniques and determining the level of confidence in privacy preservation.

Objectives This paper presents a comparative analysis between two entropy measures: Legendre polynomial modelling-based permutation entropy (LPPE), recently introduced in as a complexity measure for time series, and established variants of permutation entropy (PE), such as multiscale PE (MPE) and phase PE (PPE). The study focuses on their application to encrypted data as indicators of privacy preservation.

State-of-the-art Information entropy can serve as a quantitative measure to assess the disorder or randomness in data. While few entropy-based methods have been utilised to evaluate the resilience, efficacy, or vulnerability of encryption techniques, they are less commonly employed to detect corrupted or compromised data .

The Shannon entropy (SE) is the most commonly employed entropy for assessing encryption robustness, with a high SE value serving as a relevant indicator of a secure encryption technique. Recently, entropies based on ordinal pattern theory, such as PE and its variants MPE, have emerged as alternative metrics for this purpose. However, there has been limited discussion in the literature regarding the application of the newest PE variants, namely PPE and LPPE, as indicators of encryption robustness.

Methods PE is a concept that applies the SE to the probability distribution of ordinal patterns (OP) extracted from the data, rather than the original data. An OP represents the permutation in the increasing order of segments of successive samples. Classical PE, MPE and PPE, unfortunately, exhibit sensitivity to factors such as sampling frequency, noise presence and downsampling preprocessing. In contrast, the recently introduced LPPE benefits from applying SE to the probability distribution of OP constructed using a robust Legendre polynomial fitting of the sample segments. LPPE demonstrates superior performance compared to PE, MPE and PPE, making it a promising candidate for evaluating encryption robustness.

Results and conclusion We conduct an analysis of homomorphically encrypted data* using the complexity measures: PE, MPE, PPE and LPPE. Preliminary results depicted in fig. 1, 2 and 3 indicate that LPPE performs better than PE as an indicator of well-performed encryption on these data. Other medical datasets will be considered to consolidate our findings.

[9515] *Modeling a set of variables with different attributes on a quantitative dependent variable: an application of dichotomous variables*

José Gerardo Covarrubias López (Faculty of Higher Studies Aragon, National Autonomous University of Mexico) and Xuedong Liu Sun (Faculty of Higher Studies Aragon, National Autonomous University of Mexico).

This study outlines the methodology employed to model the relationship among a set of dichotomous variables on a nominal scale, which represent attributes. The objective is to elucidate their influence on a quantitative dependent variable measured on a ratio scale. This approach allows for the quantification of the impact of these attributes and their significance in shaping the behavior of the entity possessing them. The resolution method employed for estimation is ordinary least squares. How-ever, it is crucial to note that interpreting the estimators in the resulting model requires a nuanced perspective, distinguishing it from the conventional interpre-tation of slope or rate of change in a classic model. To clarify, these estimators align with the average behavior of the dependent variable concerning binary characteristics, and the outcomes are in harmony with the analysis of variance.

[9610] *A Global Deep Learning Perspective on Australia-Wide Monthly Precipitation Prediction*

Luyi Shen (School of Mathematics and Statistics, University of Melbourne, Parkville, VIC 3010, Australia), Guoqi Qian (School of Mathematics and Statistics, University of Melbourne, Parkville, VIC 3010, Australia) and Antoinette Tordesillas (School of Mathematics and Statistics, University of Melbourne, Parkville, VIC 3010, Australia).

Gaining a deep understanding of precipitation patterns is beneficial for enhancing Australia's adaptability to climate change. Driven by this motivation, we present a specific spatiotemporal deep learning model that well integrates matrix factorization and temporal convolutional networks, along with essential year-month covariates and key climatic drivers, to analyze and forecast monthly precipitation in Australia. We named this spatiotemporal TCN-MF method. Our approach employs the precipitation profiler-observation fusion and estimation (PPrOFusE) method for data input, synthesizing monthly precipitation readings from the Bureau of Meteorology (BoM)'s gauge measurement, JAXA's Global Satellite Mapping of Precipitation (GSMaP), and NOAA's Climate Prediction Center Morphing (CMORPH) technique. The input dataset spans from April 2000 to March 2021 and covers 1391 Australian grid locations. To evaluate the model's effectiveness, particularly in regions prone to severe flooding, we employed the empirical dynamic quantiles (EDQ) technique. This method ranks cumulative rainfall levels, enabling focused analysis on areas most affected by extreme weather events. Our assessment from April 2021 to March 2022 highlights the model's proficiency in identifying significant rainfall, especially in flood-impacted locations. Through the analysis across various climatic zones, the spatiotemporal TCN-MF model contributes to the field of continent-wide precipitation forecasting, providing valuable insights that may enhance climate change adaptability strategies in Australia.

[9712] *Detecting Trend Turning Points in PS-InSAR Time Series: Slow-Moving Landslides in Province of Frosinone, Italy*

Ebrahim Ghaderpour (Sapienza University of Rome), Benedetta Antonielli (Sapienza University of Rome), Francesca Bozzano (Sapienza University of Rome), Gabriele Scarascia Mugnozza (Sapienza University of Rome) and Paolo Mazzanti (Sapienza University of Rome).

Detecting slow-moving landslides is a crucial task for mitigating potential risk to human lives and infrastructures. In this research, Persistent Scatterer Interferometric Synthetic Aperture Radar (PS-InSAR) time series, provided by European Ground Motion Service (EGMS), for province of Frosinone in Italy are employed, and the Sequential Turning Point Detection (STPD) is applied to them to estimate the dates when the displacement rate change. The estimated dates are classified based on the land cover/use of the province. Moreover, local precipitation time series are employed to investigate how precipitation rate change might have triggered the landslides.

[9730] *Analyzing the impact of Spanish electricity hourly prices over the consumption patterns*

Eduardo Caro (Universidad Politécnica de Madrid) and Jesús Juan Ruiz (Universidad Politecnica de Madrid).

The restructuring of electricity fees in Spain on June 1, 2021, ushered in a new era of tariff distinctions, dividing the day into peak, flat, and valley periods. This shift introduced variability into the hourly pricing of electricity, impacting consumers' conventional consumption patterns and reshaping the load curve. A substantial portion of consumers, now subject to variable hourly prices, have adapted their consumption behaviors to capitalize on lower-priced periods, such as nights and weekends, in a bid to mitigate monthly expenses. Consequently, the utilization patterns of electrical appliances, including washing machines, dishwashers, and electric vehicle charging stations, have undergone subtle shifts, influencing the load profile. The dependency of electricity prices on time of day suggests a corresponding variation in consumption patterns, with a propensity to migrate towards lower-cost periods. Incorporating econometric variables into predictive models of electricity consumption stands to enhance forecasting accuracy, as fluctuations in hourly consumption correlate with the prevailing electricity prices. This interplay between hourly electricity prices and consumption dynamics underscores the need for a nuanced understanding of how market forces shape consumer behavior and load profiles, offering insights into optimizing energy management strategies for a sustainable future. This paper quantifies and analyzes the impact of the varying electricity hourly prices over the patterns of electricity consumption.

[9771] *Multi-Modal Model based on LSTM for Production Forecasting in Oil Wells with Rod Lift System*

David Esneyder Bello Angulo (Universidad Nacional de Colombia) and Elizabeth León Guzmán (Universidad Nacional de Colombia).

This paper introduces a novel multi-modal recurrent model for time series forecasting based on LSTM, focusing on production forecasting in oil wells with rod lift systems. The model is designed to process and integrate time series with heterogeneous data types, where each time step has multimodal data from images and numerical sources, enabling comprehensive analysis over specified temporal windows. The architecture comprises distinct submodels for handling different data modalities. These submodels generate a unified concatenated feature vector, capturing a holistic view of the well's operational status, which is further processed through a dense layer for non-linear transformation and integration. The core of temporal analysis is conducted via a Long Short-Term Memory (LSTM) layer, adept at capturing information across extended periods. This is complemented by a fully connected layer with linear activation output for one-shot multi-step forecasting necessary for the different modalities between input and output. The results show that the multimodal model achieved the best fit in the studied cases with a Mean Absolute Percentage Error (MAPE) of 8\%, compared to ARIMA implementations, as well as univariate and multivariate Deep Learning-based models that obtained results over 9\%.

[9781] *A comparative analysis of forecasting models for CO2 prediction in school classrooms in Navarra*

Peio Garcia (Department of Statistics, Computer Science and Mathematics, Public University of Navarre), Aranzazu Jurio (Department of Statistics, Computer Science and Mathematics, Public University of Navarre) and Daniel Paternain (Department of Statistics, Computer Science and Mathematics, Public University of Navarre).

In the context of the COVID-19 pandemic, the attention to indoor air quality (IAQ) has increased significantly. Due to the considerable amount of time people spend indoors during their lives, the control and monitoring of air parameters has become more critical. This interest focuses not only on the transmission of viruses, but also on the risks associated with exposure to pollutants present in IAQ, where carbon dioxide (CO2) is a key ventilation indicator, and which negatively affects productivity and academic performance. In this context, monitoring technologies based on Internet of Things (IoT) and the current development of low-cost IoT de-vices that provide real-time data are a powerful source of data to improve IAQ. In this work, we focus on CO2 prediction in schools, where avoiding exposing chil-dren to CO2 pollution peaks is very important. Specifically, we have installed devices in 15 classrooms of schools in Pamplona (Navarra) and have collected data with an interval of 10 minutes between measurements during the period from 10 January 2022 to 3 May 2022. These devices, called MICA and provided by the company Inbiot Monitoring (https://www.inbiot.es/productos/dispositivos-mica/mica), can measure multiple IAQ parameters, including temperature, humidity, PMs (particulate matter) and CO2. The objective of this work is to analyze the performance of differ-ent forecasting methods for predicting CO2 at different time horizons. For compari-son, different model architectures from the literature related to different methodolog-ical approaches have been used and trained. Firstly, as baseline models we have se-lected Shifted and Moving Average models. Then, from statistical models we have chosen ARIMA. In the context of Machine Learning (ML), we have divided into classical ML models, with Random Forest (RF) and XGBoost, and deep learning-based models, with Long Short-Term Memory Network (LSTM) and N-HiTS. For evaluating the prediction models, a validation strategy called Rolling Cross Validation has been proposed, where we split the whole period of data into 8 different intervals of train and test sets, to obtain more realistic results and less biased to a single period. In parallel, the models will be trained and evaluated to predict CO2 levels at different prediction instants using the backtesting process. We are not only interested in evaluating the prediction capacity of the models at a given instant, but we are also studying the usefulness of the models as we extend the prediction instant. Finally, after analyzing the results, we have seen that most of the models show a similar performance when the prediction horizon is very close. However, as we in-crease the prediction horizon, ML-based models show evidence of better results than the rest of the models, except for LSTM. This study is useful as a starting point for future solutions to predict CO2 levels inside school classrooms.

[9812] *GARCH Models Under Additive Outliers: A Robust M-quantile Approach*

Patrick Patrocinio (Université Paris-Saclay (CentraleSupélec)), Valderio Reisen (Federal University of Espirito Santo) and Pascal Bondon (Université Paris-Saclay (CentraleSupélec)).

M-regression and quantile methods have been suggested to estimate generalized autoregressive conditionally heteroscedastic (GARCH) models. In this study, we propose an M-quantile approach, which combines quantile and M-regression to obtain a robust estimator of the conditional volatility when the data have abrupt observations or heavy-tailed distributions. Monte Carlo experiments are conducted to show that the M-quantile approach is more resistant

against additive outliers than M-regression and quantile methods. The usefulness of the method is illustrated on two financial datasets.

[9973] *Prediction and estimation of yield cereal production using NDVI time series (Sentinel-2) data in central Spain*

César Sáenz (Departamento de Ingeniería Agroforestal, ETSIAAB, Universidad Politécnica de Madrid (UPM)), Alfonso Bermejo-Saiz (Departamento de Ingeniería Agroforestal, ETSIAAB, Universidad Politécnica de Madrid (UPM)), Víctor Cicuéndez (Departamento de Física de la Tierra y Astrofísica, Universidad Complutense de Madrid (UCM)), Tomás Pugni (Departamento de Ingeniería Agroforestal, ETSIAAB, Universidad Politécnica de Madrid (UPM)), Diego Madruga (Departamento de Ingeniería Agroforestal, ETSIAAB, Universidad Politécnica de Madrid (UPM)), Javier Litago (Departamento de Economía Agraria, Estadística y Gestión de Empresas, ETSIAAB, Universidad Politécnica de Madrid (UPM)) and Alicia Palacios-Orueta (Departamento de Ingeniería Agroforestal, ETSIAAB, Universidad Politécnica de Madrid (UPM)).

Agriculture in Spain is an important economic sector with a usable agricultural area of approximately 23 million hectares, where 17 million hectares are used and dedicated to crops. The area is divided into rainfed (76%) and irrigated (24%) crops. One of the most important agricultural areas is Castilla y León (CyL), where 3.5 million hectares are sown with rainfed arable crops corresponding 2.04 million hectares to cereal crops. Due to the importance of agriculture, predicting crop yields is essential for optimizing farming practices and improving the financial management of agri-food farms, as well as for policy makers and farmers' organizations to develop and implement appropriate management policies. Production estimation in non-irrigated crops is especially difficult in areas with a mediterranean type of climate, because of the high interannual variability of rainfall. Nowadays Earth Observation sensors provide access to a large amount of data with sufficient temporal and spatial resolutions for assessing many land surface processes. Specifically, data obtained from the Sentinel-2 mission, (Copernicus program of the European Space Agency) with 10m spatial resolution and 5 days temporal frequency are extremely useful for monitoring crop evolution and dynamics. Time series of the vegetation index NDVI (Normalized Difference Vegetation Index) contain highly relevant information on crops temporal dynamics, making it possible to monitor and understand vegetation growth as well as the development of effective predictive models. The aim of this work is to develop time series methodologies to estimate crop yield based on precipitation and maximum NDVI values during the growing period. For that, it is essential to develop effective NDVI forecasting models to predict future NDVI values during the year. NDVI time series along several years has been used to estimate and forecast future NDVI values at a pixel level in several cereal crops. The study area comprises the provinces of Soria, Burgos and Palencia, located in the autonomous community of Castilla y León, in the northern part of the central plateau. The NDVI time series from the Sentinel-2 tiles 30 TWM (Soria), 30 TVM (Burgos) and 30 TUN (Palencia) for the period 2017-2023 were compiled. Precipitation data for the period 2017-2021 was provided by the State Meteorological Agency (AEMET). Production data from the “Encuesta de Superficies y Rendimientos de Cultivos” (ESYRCE) in the period 2017-2022 have been used for validation purposes. The NDVI time series of a rainfed cereal crop have been analyzed using statistical time series methods and their values were predicted using the Box-Jenkins methodology. After that, a multiple linear regression model was applied using historical plot yield data, accumulated precipitation and maximum NDVI.

[9974] *Assessing the Pre-processing Benefits of Data-Driven Decomposition Methods for Phase Permutation Entropy - Application to econometric Time-series.*

Meryem Jabloun (PRISME laboratory).

This paper investigates the efficacy of various data-driven decomposition methods combined with Phase Permutation Entropy (PPE) to form a promising complexity metric for analyzing time series. PPE is a variant of the classical permutation entropy (PE), while the examined data-driven decomposition methods include Empirical Mode Decomposition (EMD), Variational Mode Decomposition (VMD), Empirical Wavelet Transform (EWT), Seasonal and Trend decomposition using Loess (STL), and Singular Spectrum Analysis-based decomposition (SSA). To our knowledge, this combination has not been explored yet. Our primary aim is to assess how these preprocessing methods affect PPE’s ability to capture temporal structural complexities within time series. This evaluation encompasses the analysis of both simulated and econometric time series. Our results reveal that combining SSA with PPE produces superior advantages for measuring the complexity of seasonal time series. Conversely, VMD combined with PPE proves to be the less advantageous strategy. Overall, our study illustrates that combining data-driven preprocessing methods with PPE offers greater benefits compared to combining them with traditional PE in quantifying time series complexity.

doble folio
EDITORIAL
ISBN: 979-13-87522-03-2
GR 1304-2024