# 7ο Ετήσιο Συνέδριο της Κυπριακής Στατιστικής Εταιρείας

Παρασκευή, **19 Ιουνίου 2026**

Πανεπιστήμιο Frederick
Λευκωσία

Διοργάνωση

ΚΥΠΡΙΑΚΗ ΣΤΑΤΙΣΤΙΚΗ ΕΤΑΙΡΕΙΑ
CYPRUS STATISTICAL SOCIETY

Frederick University
1965 2025

Χορηγοί

Clinbay

Alliott Partellas Kiliaris Ltd
Certified Public Accountants

Grant Thornton

pulse
market research

# 7th Annual Conference of the Cyprus Statistical Society
# 19th of June 2026, Nicosia Cyprus

## Book of Abstracts

### 1. Estimation for Multivariate Integer-Valued Time Series

Konstantinos Fokianos

University of Cyprus

We study inference and modelling techniques for multivariate count time series. In particular, we focus on autoregressive, network, and space–time models. The development of both the underlying theory and its applications depends critically on the choice of an appropriate response distribution for multivariate count data, as well as on the framework of observation driven models. Our aim is to highlight recent methodological advances in this area and to outline several directions for future research.

## 2. Making Gender Perspective Visible: Data-Driven Insights Into Academic Communication And Teaching Practices

Reda Adomaitienė, Lijana Stundžė, Aurelija Novelskaitė

Gender Studies Centre at Faculty of Communication, Vilnius University, Lithuania

Although the integration of a gender perspective in research and higher education institutions is becoming an increasingly relevant and widely discussed topic and represents a necessary step toward achieving gender equality, there is still a lack of a systemic and integrated approach to the inclusion of the gender dimension and the application of an intersectional perspective in teaching and learning practices. Since study programmes are related not only to factual knowledge but also to cultural attitudes, societal values, and the overall organizational culture, gender issues inevitably need to be addressed within them (Grunberg, 2010; Lift Anis Ma'shumah & Chamami, 2021). Gender is closely interconnected with other socially constructed identities (such as race, age, ethnicity, sexuality, disability, etc.); therefore, an intersectional perspective should also be integrated into curricula, taking into account how the combinations and intersections of different forms of inequality and oppression (sexism, class-based inequality, racism, homophobia, and ableism) shape specific forms of discrimination (Crenshaw, 1989). Moreover, communication plays a crucial role in integrating both the gender perspective and intersectionality into study programmes, as teaching in higher education is inherently a communicative activity (Werner et al., 2023).

This presentation introduces the results of a student survey conducted by a Lithuanian team of researchers within the international Erasmus+ KA2 project *TINGLE – Academia* (*Through Intersectional and Gendered Lens to Equity in Academia*, https://tingleacademia.org). The aim of the survey was to collect data on the integration of the gender perspective and the application of an intersectional approach in academic practices at the university. The questionnaire survey, conducted in 2024 among 146 students, revealed that students evaluate their knowledge of the gender perspective as more sufficient than their knowledge of intersectionality. It was also observed that a considerable proportion of respondents do not pay attention to manifestations of the gender perspective in the university environment. This suggests that one of the possible reasons is the absence of appropriate academic communication, which is associated with a lack of gender sensitivity and insufficient knowledge of gender-related issues.

The presented research findings aim to highlight the importance of statistics and data collection, as empirical data are essential and relevant for assessing the integration of the gender perspective and the application of intersectionality in the academic field, for creating a more inclusive teaching and learning environment, and for developing institutional strategies.

### 3. Adaptive Bayesian inference with series and shallow neural network p-exponential priors with p<1

Sergios Agapiou, Ismael Castillo, Paul Egels

University of Cyprus

I will discuss recent work on $\alpha$-smooth $p$-exponential series priors and their posterior contraction rates over $\beta$-Sobolev truths. Earlier results for $p\in[1,2)$ showed that when the prior is oversmoothing ($\alpha>\beta$), smaller $p$ leads to faster rates, while for $\alpha<\beta$ the rate is independent of $p$ and suboptimal, with the minimax rate achieved for $\alpha=\beta$. These findings had inspired the construction of oversmoothed heavy-tailed priors, which achieve full adaptivity without sampling any hyperparameters. In this talk, I will overview new results extending this behaviour to $p<1$, showing that the improvement continues, and discuss how these ideas allow us to introduce adaptive series priors with $p$-exponential weights—again without the need for sampling hyperparameters. The improvement for smaller $p$ holds beyond series priors, in particular, for overparametrized shallow neural network priors with $p$-exponential weights. Similarly to the series priors case, this allows us to design adaptive shallow neural network priors.

## 4. On the Stochastic Modelling of Daily Rainfall Intensity: A Censored Autoregressive Approach with Transformed Marginals

Andrea Marina Alexiou

University of Cyprus

Daily rainfall records are characterised by a large proportion of dry days and highly skewed positive rainfall amounts, features that are difficult to capture simultaneously using conventional time series models.

This work proposes a flexible stochastic framework for modelling daily rainfall intensity that combines rainfall occurrence and rainfall magnitude within a unified probabilistic structure.

The proposed model is based on a latent Gaussian autoregressive process that is censored at zero to represent dry-day occurrence and subsequently transformed to a target marginal distribution through the probability integral transform. This construction decouples temporal dependence from the distribution of wet-day intensities, allowing alternative rainfall distributions to be incorporated without modifying the underlying dependence structure.

A moment-based estimation procedure is developed, enabling direct estimation of rainfall occurrence and temporal persistence from the observed process. The performance of the methodology is investigated through an extensive simulation study under Gamma, Weibull, Lognormal and Kaniadakis ω-Weibull marginals. Results demonstrate accurate parameter recovery and reliable identification of the underlying distributional form.

The methodology is applied to daily precipitation time series representative of different climatic regimes, demonstrating its ability to capture rainfall occurrence, persistence, and intensity while accommodating diverse distributional characteristics. The results illustrate the effectiveness and flexibility of the approach in reproducing key features of precipitation data and support its use as a general framework for stochastic rainfall modelling.

Overall, the proposed approach provides a computationally efficient and interpretable framework for precipitation modelling, with potential applications in environmental statistics, hydrology, and stochastic time series analysis.

## 5. Sparse Sufficient Dimension Reduction

Andreas Artemiou

University of Limassol

We propose a sparse sufficient dimension reduction method within the framework of Least Squares Support Vector Machines (LSSVM) that achieves dimension reduction and feature selection simultaneously. By incorporating sparsity-inducing regularization, the proposed approach estimates the central subspace while selecting relevant predictors. An efficient proximal gradient algorithm is developed that avoids covariance matrix inversion, making the method suitable for high-dimensional settings. Compared with existing SVM-based dimension reduction methods, including principal SVM and principal LSSVM, the proposed approach is more scalable with respect to the dimensionality of the input features. Simulation studies and a real data application demonstrate the effectiveness of the proposed method.

## 6. The role of Artificial Intelligence for employed people with disabilities

Samer Assaf, Haritini Tsangari

University of Nicosia

Artificial Intelligence (AI) is currently transforming contemporary workplaces. While technological innovations promise improved efficiency, they also highlight the critical need for fostering belongingness within organizational settings. The call is becoming even more pronounced for people with disabilities, who are still evidenced to face barriers. The current research aims to examine if AI can serve as a tool for enhancing belongingness in the workplace. A comprehensive conceptual framework is developed and tested, depicting the synergy between AI technologies and workplace inclusion, for people with disabilities.

## 7. Redefining Pedestrian Motion: A Novel, Data-Driven Approach via Inverse Reinforcement Learning

Rafael Athanasiades, Andrew Duncan

Imperial College London

Pedestrian motion displays a complex relationship between intentions and the influence of the surrounding environment. The depth of this dynamic environment has implications not only for urban planning and crowd management, but also for the in current rapid development of autonomous vehicles.

This research explores pedestrian motion with the goal of formulating a data driven approach to develop an improvement of current approaches which are not backed up by data. Such example is the Social Force Model (SFM) which stands as baseline to our approach.

We employed Inverse Reinforcement Learning (IRL) to infer the underlying reward function that describes pedestrian behaviors and decisions in particular scenarios, based on real-world data. By integrating the extracted reward function into a custom OpenAI Gym environment we ensure adaptability for subsequent research in the future. With the application of Reinforcement Learning (RL) in this environment, along with the inferred reward function, we obtain a stochastic policy which describes pedestrian motion to an individual level. This policy results trajectory predictions, providing a comprehensive representation of potential pedestrian paths. Our understanding of their decision-making of pedestrians becomes clearer setting the stage for further investigation of pedestrian motion using IRL.

## 8. Diameter Thresholds of Random Cayley Graphs - I

Demetres Christofides[1], Klas Markström[2] & Christina Savvidou[1]

[1]INSPIRE Research Center and UCLan Cyprus, Cyprus, [2] Umeå University, Sweden

**Abstract**. For any fixed integer $d \geqslant 2$, the asymptotic threshold for the binomial random graph $\mathcal{G}(n,p)$ to achieve diameter at most $d$ is a classical result in probabilistic combinatorics, exhibiting a sharp phase transition driven by independent edge exposures. While random regular graphs share identical thresholds due to asymptotic universality, this universality breaks down when structural dependencies are introduced. In this first of a two-part talk, we review the classical diameter thresholds for the binomial random graph model along with the core probabilistic techniques behind their proofs. We then introduce the model $\mathcal{G}(G,p)$ of random Cayley graphs, where edges are governed by an underlying group structure. Finally, we highlight the unique probabilistic challenges that arise when dealing with algebraically dependent edge distributions and present our main results establishing new upper and lower bounds for the diameter.

## 9. Diameter Thresholds of Random Cayley Graphs - II

Demetres Christofides[1], Klas Markström[2] & Christina Savvidou[1]

[1]INSPIRE Research Center and UCLan Cyprus, Cyprus, [2] Umeå University, Sweden

In the random Cayley graph model $\mathcal{G}(G,p)$, edges are determined by choosing elements from an underlying group $G$ independently at random with probability $p$. Unlike classical independent models like the Erdős-Rényi random graph, the geometry of $\mathcal{G}(G,p)$ is heavily constrained by the group's algebraic structure. This creates complex structural dependencies and forced path intersections, causing a breakdown of the standard universality constants typically observed in diameter phase transitions.

In this second of a two-part talk, we focus on the probabilistic machinery and core proof strategies developed to determine these diameter thresholds for both constant and growing diameters $d$. We will first provide the combinatorial intuition behind why different group families shift the threshold constants by factors related to commutativity or the density of involutions. We then present our unified framework designed to overcome these algebraic dependencies, highlighting an optimized dependency graph construction and a refined application of Janson's inequality over intersecting, non-independent paths. While complementing the first of a two-part talk, this talk is self-contained and emphasizes modern probabilistic tools for constrained random structures.

## 10. Data-Driven Evaluation of an AI E-Learning Programme for Retail Professionals: Evidence from the INAIR Pilot Study

Thomas Photiadis, Alexandros Yeratziotis

University of Cyprus

Artificial Intelligence is progressively reshaping retail operations, placing growing pressure on businesses — particularly micro, small and medium-sized enterprises — to develop relevant competences among their workforce. This paper presents a data-driven evaluation of the INAIR E-Learning Programme on AI for Retail, developed under the Horizon Europe project *Increasing the Uptake of AI in Retail* (Grant Agreement No. 101133847). The programme offers a multilingual digital learning environment structured around personalised learning pathways, sixteen modular units across three proficiency levels, and embedded assessment mechanisms tailored to retail professionals across five European countries: Cyprus, Germany, Italy, Poland and Romania. The pilot study assessed the programme's usability, relevance, perceived learning value, user satisfaction and real-world applicability. Data were collected from platform participation records and two structured evaluation surveys. By the time of analysis, 242 users had registered, 69 had completed the final assessment, 75 retail workers and business owners had submitted end-user survey responses, and 61 external experts had completed a parallel expert survey. The analysis combined descriptive statistics from Likert-scale items with systematic interpretation of open-ended qualitative responses.

Results were broadly positive across all evaluation dimensions. Among end-user respondents, 84% considered the educational resources reflective of real-world retail scenarios, 87% found the modules clear and well-structured, 88% reported improved understanding of AI in retail, and 85% indicated increased confidence in considering or adopting AI-based solutions. A further 72% expressed intention to integrate AI tools or approaches into their work. Expert evaluations were consistently stronger, with 97–98% affirming content relevance and 96% stating they would recommend the programme to retailers. Qualitative feedback highlighted the practical orientation of the content and the clarity of the learning pathway as particular strengths, while pointing to opportunities for improvement in interactivity, module pacing, mobile usability, accessibility and multimedia integration.

The study contributes to the emerging literature on evidence-based evaluation of digital upskilling initiatives targeting the retail workforce. The findings suggest that structured, role-based, learner-centred and context-relevant e-learning can meaningfully support AI literacy and adoption readiness. Nonetheless, as the evaluation relies on self-reported data from a voluntary pilot sample, future research should examine longer-term behavioural outcomes and actual organisational uptake of AI practices following programme completion, including measurable impacts on workplace performance and digital transformation readiness.

## 11. Integrating Data Science through Interdisciplinary STEAM Practice: A School-Based Implementation for Civic Engagement and Social Justice

Michalis Gavrielides, Leonie Hadjithoma

The English School, Nicosia

This presentation examines the implementation of the DataScEd4CiEn project at The English School in Nicosia, Cyprus, with particular emphasis on the integration of data science education within interdisciplinary STEAM learning as a means of fostering civic engagement and social justice. The project is situated within the broader educational imperative to cultivate students' data science literacy, critical thinking, ethical awareness, and capacity to engage with complex societal issues through evidence-informed reasoning.

The presentation considers both the initial pilot phase and the subsequent round of implementation, tracing the project's development from teacher professional learning to classroom enactment. Through a school-based professional development model, teachers were supported in engaging with data science concepts, digital tools, artificial intelligence, ethical inquiry, and inquiry-based pedagogical approaches. Particular emphasis was placed on collaborative curriculum design, cross-departmental planning, and the development of learning scenarios that connected disciplinary knowledge with socially relevant contexts.

Across the two phases of implementation, students engaged with authentic data, formulated investigable questions, interpreted patterns, considered ethical and civic dimensions, and developed informed responses to issues related to sustainability, responsibility, and social justice. These experiences illustrate the potential of data-rich STEAM activities to transcend traditional subject boundaries and to strengthen the relationship between classroom learning and civic participation.

The presentation discusses the contribution of the project to teacher professional learning, interdisciplinary collaboration, and the practical enactment of data science education in a school setting. It further reflects on implementation challenges and on the institutional conditions required to sustain and extend such approaches within formal education.

**12. Teaching Statistics in an AI world**

Thanasis Hadzilacos

Frederick University

It was 7:30 pm. Students had already completed a full workday. In a computer lab they were facing screens, ChatGPT open, phones competing for attention. I started with a joke: "A recent study showed that most fires occur where there are the most fire engines. Therefore, fire engines cause fires - let's eliminate them to reduce fires." They laughed. Ten years from now, this may be the only thing they remember from my AI course. But after the joke comes the hard part: how do you teach Cronbach's alpha and variable correlation to tired evening students who barely remember mean versus median? I needed an example connecting Maritime Operations, AI and variable correlation, all serving my hidden curriculum: correlation is not causation. So I used a synthetic dataset from an imaginary port, showing that when Thanasis inspected vessels, ships under Liberian flag took longer to unload - except on weekends when Elena managed customs, when unloading became faster. Did Thanasis dislike Liberia? Were Thanasis and Elena bribed on weekends? "The data showed that…" -our entry point into AI's ability to detect hidden patterns, and into the next question: how does AI decide that variables are correlated at all?

**From "statistics as mathematics" to "statistics as judgement":** Many students think Statistics means formulas. In practice, the key decisions are conceptual: what counts as a variable, what data to collect, whether the sample is biased. Mathematics comes later.

**LLMs as cognitive prostheses, not replacements for understanding:** Weak students can now produce statistically-looking work without understanding it. Teaching shifts from calculation to critical interpretation.

**The danger of "black-box statistics" has increased, not decreased:** Before LLMs, weak students avoided statistics. Now they produce statistically-looking work without understanding. Teaching shifts: calculation to critical interpretation.

**Synthetic datasets are pedagogically powerful:** When students invent hidden patterns and datasets, they experience directly: noise versus signal, accidental correlations, multivariable interactions, and fragile conclusions.

**Visualization before formalization:** Humans often understand patterns visually before mathematically. Scatterplots, timelines, heatmaps, and grouped tables help students "see" concepts such as variance, clustering, outliers, and correlation.

**The hidden curriculum: intellectual humility:** "The data showed that…" is not the same as "Reality is…". AI, statistics, and humans all infer patterns from incomplete observations.

**AI has unexpectedly created motivation for statistics:** Students who "hate statistics" often become interested when AI produces surprising or contradictory results from realistic data.

## 13. Oncology endpoints and adaptive dose-finding designs: a statistical perspective

Κατερίνα Ιωαννίδου, Σωτήρης Προίσκος

ClinBAY

This presentation provides a statistical perspective on oncology clinical trial design, focusing on the selection of clinically meaningful endpoints and the use of adaptive dose-finding methods in early-phase studies. Key oncology endpoints, including overall survival, progression-free survival, and surrogate measures, will be discussed in the context of their methodological challenges and practical interpretation. The talk will also highlight modern adaptive and model-based dose-escalation designs that improve trial efficiency, patient safety, and decision-making compared with traditional approaches. The presentation aims to provide an introduction to the methodologies and standard practices used in oncology drug development.

## 14. Prevalence of HPV vaccination among young women in Cyprus and its association with knowledge of HPV infection and cervical cancer.

Katerina Kyprianou[1,2], Paraskevi A. Farazi[3], James Hebert[4], Costas Christophi[1,2]

1. Cyprus International Institute for Environmental & Public Health, Cyprus University of Technology, Cyprus.

2. Department of Rehabilitation Sciences, School of Health Sciences, Cyprus University of Technology, Limassol, Cyprus.

3. European University Cyprus in Nicosia, Cyprus

4. Arnold School of Public Health, University of South Carolina in Columbia, South Carolina, United States

**Introduction**: Vaccination against human papillomavirus (HPV) effectively protects against HPV infection, a major cause of cervical cancer. Despite its benefits, knowledge gaps and limited awareness remain barriers to vaccine uptake. This study assessed HPV vaccination prevalence among women aged 20-30 years in Cyprus and examined the association between HPV and cervical cancer knowledge and vaccination status. **Methods**: A cross-sectional study was conducted using a self-administered online questionnaire distributed through REDCap. A total of 134 complete responses were analyzed. The primary outcome was HPV vaccination status. Knowledge regarding HPV infection and cervical cancer was categorized into low and high levels using Youden's index derived from ROC analyses. Results are presented as proportions with 95% confidence intervals (CIs), while associations were evaluated using odds ratios (ORs) with corresponding 95% CIs.

**Results**: Among participants, 35.8% (95% CI: 28.2%-44.2%) reported receiving the HPV vaccine. High HPV knowledge was observed in 44.0% (95% CI: 35.9%-52.5%) of women, while 55.2% (95% CI: 46.8%-63.4%) demonstrated high cervical cancer knowledge. Women with high knowledge in both domains had substantially greater odds of vaccination (OR=6.39, 95% CI: 2.52-16.18) compared with women with low knowledge in both domains.

**Conclusions**: HPV vaccination uptake among young women in Cyprus was low. Limited knowledge and insufficient information were identified as key barriers to vaccination. Improving awareness may contribute to increased vaccine uptake in this population.

## 15. A Multivariate State-Space Framework for Modeling Complex Extreme Event Processes

Andreas Makrides[1], Alexandra Papadopoulou[2] and Nikolaos G. Galanopoulos[1]

[1]University of the Aegean, Greece, [2]University of Thessaloniki, Greece

A novel idea for modelling extreme events especially focused on seismic activity is presented in this paper. Unlike classical models, the proposed framework incorporates a multivariate state space where each state is defined as a vector encoding both the magnitude of the mainshock and temporal or numerical features of the associated foreshock and/or aftershock sequences. This structure enables a dynamic and data-informed correlation between the mainshock characteristics and the surrounding seismic activity, enhancing the descriptive and predictive power of the model. The generalized Markov structure allows flexible sojourn time distributions, capturing the non-memoryless nature of seismic transitions. Empirical validation using real earthquake catalogs demonstrates that the model effectively captures key patterns in seismic sequences, offering new insights into the temporal evolution and interdependence of seismic events. This new approach might reveal promising paths for probabilistic seismic hazard analysis and provide additional stochastic tools for thorough investigation of extreme phenomena attributes.

## 16. Preparing Teachers for Responsible AI and Data Science Education: Pilot Testing DataSETUP Modules in a Master's Course on AI in STEAM Education

Maria Meletiou-Mavrotheris[1], Dionysia Bakogianni[2,3], Yianna Danidou[1], Efi Paparistodemou[1,4], Alexandros Kofteros[1,5]

[1]European University Cyprus, [2]National Kapodistrian University of Athens, [3]Karlstad University, [4]Cyprus Pedagogical Institute, [5]Cyprus Ministry of Education

As Artificial Intelligence (AI) and Machine Learning (ML) increasingly shape educational, social, and civic life, statistical literacy must expand to include critical engagement with data-driven algorithms, assumptions, and consequences. This presentation reports the first pilot implementation of two modules developed within the Erasmus+ DataSETUP project and embedded in the Master's-level course "AI in STEAM Education" at European University Cyprus. The modules formed a scaffolded sequence for 36 student teachers: *Responsible AI & Data Science: Ethics, Society, and Citizenship* which introduced AI/ML-related data science concepts through hands-on classification tasks, while *Integrating AI/ML-Related Data Science in STEAM Education to Foster Responsible and Active Citizenship* shifted focus toward classroom-ready STEAM scenarios. The empirical material analysed in this study consisted of student-produced artefacts, including group worksheets and projects, written reflections, presentations, scenario-based tasks, and pre/post and satisfaction survey responses. These materials were analysed thematically around three dimensions: participants' conceptualisation of data science in relation to statistics and AI/ML; their reasoning about data quality, bias, privacy, fairness, and model-based decision-making; and their translation of these ideas into possible STEAM classroom activities. Initial analysis suggests that participants broadened their understanding of data science beyond statistics, increasingly recognised data and algorithms as non-neutral, and began to connect AI/ML-related data science with teaching practice. At the same time, the move from ethical awareness to concrete, age-appropriate pedagogical design remained uneven. Overall, the pilot suggests that integrating AI, data science, and statistics in teacher education can support a shift from treating AI as a technical tool toward approaching it as a data-driven, socially situated, and ethically consequential object of inquiry.

## 17. Inconsistent Responding on a Mixed-Worded Scale in the PISA 2022 Questionnaire: Prevalence and Predictors Across Seven Countries

Michalis P. Michaelides[1], Evi Konstantinidou[1], Militsa Ivanova[1], Isa Steinmann[2]

[1] University of Cyprus

[2] Oslo Metropolitan University

Mixed-worded questionnaire scales, containing both positively and negatively phrased items, are intended to encourage thoughtful responses, but may result in inconsistent responses and wording effects and can reduce reliability and validity. This study examined the prevalence and predictors of inconsistent responding to a mixed-worded scale in the 2022 Programme for International Student Assessment (PISA) across seven countries. Using factor mixture analysis, findings revealed a low but variable rate of inconsistency (2–8%) across countries. Separate confirmatory factor analysis models on the effective and on filtered samples of consistent respondents showed improved fit and reliability indices for the filtered samples for all countries, suggesting that inconsistent respondents had unexpected response patterns. In logistic regression models, inconsistent responding was associated with lower PISA reading scores and often with reduced self-reported effort, suggesting that inconsistent responses may stem from lower reading ability and/or carelessness. The results contribute to ongoing discussions about the use of mixed wording in questionnaires and highlight the need to account for respondent behavior when analyzing survey data.

## 18. A big team science study on the robustness of findings from the social and behavioral sciences

Michalis Michaelides

University of Cyprus

Within a research study a dataset can be analysed in alternative, justifiable ways. A recent large-scale, collaborative project approached the variability of findings emerging from alternative analytical decisions (e.g. in data cleaning, variable definitions, software used, modelling choices, interpretation). I will present results from a study of the SCORE program I participated in, where 457 independent analysts examined the analytical robustness in the social and behavioural sciences. One hundred papers published between 2009 and 2018 were randomly selected. At least five independent researchers reanalysed the original data for one claim from each paper. Their analysis reports were then peer-reviewed and compared.

Results showed that 34% of reanalyses produced results similar to the original findings (within ±0.05 Cohen's d), increasing to 57% when a broader tolerance range was used. Overall, 74% of reanalyses reached the same conclusion as the original study, 24% found no effect or inconclusive evidence, and 2% reported the opposite effect. While this speaks to the credibility of the published research, the common practice of relying on a single analytical approach in social and behavioural research may not adequately capture uncertainty arising from alternative, equally justifiable analyses. Therefore, a wider use of methods that explore and communicate this source of uncertainty is recommended.

**19. Deep Learning for Stochastic Stackelberg Competition in Electricity Markets**

Sebastian Niehaus

Needlestack Technologies Ltd., Larnaca, Cyprus and Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

The increasing penetration of decentralized and renewable generation has transformed electricity markets into highly stochastic and data-rich environments, where strategic interactions among heterogeneous agents are subject to significant uncertainty. In particular, large market participants operate in settings where exogenous volatility and the endogenous reactions of competing agents across coupled futures and spot markets must be accounted for. We model these interactions within a stochastic Stackelberg framework, focusing on the follower's optimization problem under uncertainty. Given a leader strategy *x*, the follower's decision is characterized by the stochastic program

$$y^*(x,\omega) = \arg\max_{y \in \mathcal{Y}} U_F(x,y,\omega), \tag{1}$$

where $\omega$ denotes random market conditions such as demand fluctuations and renewable generation. To address the computational complexity of repeatedly solving this problem, we propose a deep learning–based approach to approximate the follower best-response mapping $y^*(x,\omega)$. Neural networks are employed to learn a surrogate model $\hat{y}(x,\omega)$ from high-dimensional, multimodal inputs, including price trajectories, load profiles, and renewable generation signals. This data-driven approximation replaces the need for repeated optimization and enables efficient evaluation of follower responses across a wide range of scenarios. We further analyze key architectural challenges, including stability under nonstationary data, representation of stochastic dependencies, and consistency with economic rationality. The proposed framework provides a scalable surrogate for the follower problem in stochastic bilevel optimization and supports data-driven analysis of strategic behavior in modern electricity markets.

**20. On the Dynamics of Spatio-Temporal Random Fields**

Marios Papadopoulos, Anastassia Baxevani

University of Cyprus

Many real-world phenomena evolve randomly in time and space, such as signals, environmental data, and wave propagation. These are naturally described using stochastic processes and random fields. Classical Gaussian random fields often rely on separability between space and time, which limits their ability to capture the transport of spatial structures over time.

To address this limitation, we consider spatio-temporal fields in which space and time are coupled through transport mechanisms. In this work, we investigate the evolution of spatio-temporal fields constructed through spectral representations and examine how they propagate under prescribed dynamics.

A flow-based formulation is introduced, linking the evolution of the field to particle trajectories and an underlying velocity field. A connection with AR(1)-type temporal dynamics and Ornstein–Uhlenbeck-type decay is established.

Finally, a one-dimensional spatio-temporal field is simulated using backward recursion along flow trajectories, illustrating how independent innovation fields transported by the flow generate the evolving random field.

## 21. Data literacy as a key twenty-first century competence: the European project DATA-READY

Efi Paparistodemou, Marina Michael, Eleni Papageorgiou

Cyprus Pedagogical Institute

This paper presents the European project DATA-READY (ERASMUS-EDU-2024-POL-EXP-DIGITAL), which aims to develop an evidence-informed framework for integrating Data Literacy into compulsory education. Data literacy is increasingly recognised as a key twenty-first century competence, closely linked to critical thinking, informed decision-making, responsible citizenship, and the ability to engage critically with data-driven and AI-mediated societies. Despite its growing importance, data literacy remains insufficiently embedded in school curricula, while many teachers report limited preparation for integrating data-related practices into teaching and learning.

The project developed a Data Literacy Competence Framework for compulsory education. The framework is organised around seven competence domains, three progression bands, and three proficiency levels, supporting the gradual development of learners' data-related knowledge, skills, and dispositions from primary to lower secondary education. Emphasis is placed on inquiry-based learning, ethical and critical engagement with data, interpretation and communication of information, and the cross-curricular integration of data literacy across subject areas.

The framework was further refined and validated through workshops and focus groups involving teachers and educational stakeholders, aiming to enhance its relevance, validity, and applicability within authentic school contexts. The DATA-READY project seeks to contribute to the development of a sustainable European approach to data literacy education by supporting curriculum innovation, teacher professional learning, and the preparation of students for active and responsible participation in increasingly data-driven societies.

## 22. Preparing Educational Research Methodology Students for a Data-Rich World: Development of a Theory-Grounded Module and Validation of a Data Science Attitudes Scale

Ioulia Televantou, Maria Meletiou, Georgia Solomonidou, Daniel Frischemeier

European University, Cyprus

Educational research is entering a new era, driven by the rabid expansion of large-scale survey and log data. Harnessing these data requires data science skills, yet such approaches remain underrepresented in most higher education research methodology curricula. This study addresses this gap through two complementary objectives:

First, it presents the design and implementation of a sequence of two data science modules within the DataSETUP project: A theoretical module, providing conceptual clarification, introducing prospective and in-service teachers to the breadth of large-scale educational data, including survey, process, and log data, as well as their methodological and ethical implications, and an applied module. The latter, grounded in the EDUCATE framework, engages participants in hands-on analysis of large-scale datasets to investigate policy-relevant issues such as educational inequalities.

Second, grounded in Expectancy-Value Theory, the study examines the psychometric properties of the Greek adaptation of the Student Survey of Motivational Attitudes toward Data Science (S-SOMADS), administered to graduate education students following module participation. The study evaluates whether the scale's properties hold in this population.

By linking structured curriculum design with measurement, the study contributes to both data science education and the validation of attitudes towards data science.

## 23. Autoimmune comorbidity doubles in Hashimoto thyroiditis: A retrospective case control study in Cyprus

Andriana M. Till[1,2], Konstantinos C. Makris[1,2], Georgios S. Papaetis[3,4], Costas A. Christophi[1,2,5†], Michalis K. Picolos[6,7†]

1. Cyprus International Institute for Environmental & Public Health, Cyprus University of Technology, Cyprus.

2. Department of Rehabilitation Sciences, School of Health Sciences, Cyprus University of Technology, Cyprus.

3. K.M.P THERAPIS Paphos Medical Center, Internal Medicine and Diabetes Clinic, Cyprus.

4. CDA College, Paphos, Cyprus.

5. Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA

6. Alithias Endocrinology Centre, Nicosia, Cyprus.

7. European University of Cyprus, Nicosia, Cyprus.

†: These authors contributed equally to this work.

Introduction: Autoimmune diseases frequently co-exist within individuals. Increasing evidence suggest a higher burden of autoimmune comorbidities in people with Hashimoto's thyroiditis. This study aimed to estimate the prevalence of autoimmune diseases in patients with Hashimoto's thyroiditis compared to age- and sex-matched controls with thyroid nodules.

Methodology: This retrospective case-control study was conducted at an endocrinology center in Cyprus. Adult patients (≥18 years) attending the center over a one-year period were screened. A total of 1525 patients with Hashimoto thyroiditis were matched by age and sex with 1525 controls. Demographic and clinical data were extracted from medical records. Conditional logistic regression analyses were performed to evaluate associations with autoimmune comorbidity.

Results: The study population was predominantly female (85%) with a median age of 50.5 years. Autoimmune comorbidities were more common in patients with Hashimoto's thyroiditis compared to controls (16.1% vs. 7.6%, $p<0.001$), corresponding to more than two-fold higher odds of additional autoimmune diseases (OR=2.44; 95% CI: 1.80-3.31). Rheumatoid arthritis, autoimmune gastritis, vitiligo, Sjögren's syndrome, celiac disease, systemic lupus erythematosus and mixed or undifferentiated connective tissue disease were significantly more prevalent in the Hashimoto's thyroiditis group. Rheumatologic/systemic diseases were the most prevalent system specific autoimmune disorders (OR=3.45; 95% CI: 1.93-6.19) followed by endocrine (OR=2.44; 95% CI: 1.29-4.60), dermatologic (OR=2.34; 95% CI: 1.24-4.46), and gastrointestinal (OR=1.92; 95% CI: 1.07-3.44).

Conclusion: Hashimoto's thyroiditis was associated with an increased risk of additional autoimmune diseases. These findings highlight the importance of clinical awareness and consideration of screening for coexisting autoimmune conditions in patients with Hashimoto's thyroiditis.

## 24. Estimating Heterogeneous Heat Stress Effects in Dairy Systems with Double Machine Learning and Causal Forests

Charalambia Varnava, Leonidas Christodoulou, Vasilis Symeou, Constantine Dovrolis

The Cyprus Institute

Heat stress poses a major challenge for dairy production, reducing milk yield and increas-ing economic losses under rising climate variability. Quantifying its effects is statistically challenging, because responses are dynamic, nonlinear and highly heterogeneous across animals and production environments. Conventional regression approaches often impose restrictive functional-form assumptions and are limited in their ability to capture complex treatment heterogeneity. This study demonstrates how modern causal machine learning methods can be used to estimate heterogeneous and time-varying effects of heat stress on milk production using high-frequency dairy farm data from Cyprus.

The empirical framework combines double machine learning (DML) with causal forests to estimate both average and conditional treatment effects in a high-dimensional setting. DML enables flexible nuisance estimation, while preserving valid causal inference through orthogonalization and cross-fitting, reducing bias from regularization and model selection. Causal forests are then employed to recover heterogeneous treatment effects without imposing parametric interaction structures, allowing treatment responses to vary flexibly across animal characteristics and farm environments. The proposed framework accommodates nonlinearities, complex interactions and high-dimensional confounding, while maintaining interpretability of estimated causal effects.

The results show persistent negative effects of HS on milk production across several lag periods, consistent with cumulative physiological stress. Estimated losses increase non-linearly with heat intensity and exhibit substantial heterogeneity under severe conditions. Important differences emerge across parity, lactation stage, productivity groups and barn environments with different cooling regimes. The heterogeneity estimates suggest that vulnerable subpopulations experience disproportionately larger production losses during extreme heat events.

Beyond the empirical findings, the study highlights the methodological advantages of causal machine learning for agricultural and environmental applications characterized by dynamic exposures, heterogeneous responses and observational data structures. The integration of DML, causal forests and distributed lag modeling provides a flexible and scalable framework for causal estimation in complex production systems. These results demonstrate the potential of modern statistical learning methods to support precision live-stock management and climate adaptation under increasing environmental stress.