

---

---

# 26th Young Statisticians Meeting

## YSM 26

---

---

PROGRAMME — ABSTRACTS — PARTICIPANTS

SEPTEMBER 30 - OCTOBER 2, 2022  
BOHINJ, SLOVENIA



SUPPORTED BY  
CENTER FOR METHODOLOGY AND INFORMATICS,  
FACULTY OF SOCIAL SCIENCES, UNIVERSITY OF LJUBLJANA

## **Local Organizer**

Luka Kronegger  
University of Ljubljana, Faculty of Social Sciences  
Kardeljeva pl. 5, 1000 Ljubljana, Slovenia  
luka.kronegger@fdv.uni-lj.si

## **International Program Committee**

Austria: Peter Filzmoser, Wien  
Croatia: Nenad Šuvak, Zagreb  
Hungary: David Simon, Budapest  
Italy: Paola Berchiolla, Torino  
Slovenia: Luka Kronegger, Ljubljana

Published by: CMI, FDV, University of Ljubljana

Edited by: Luka Kronegger  
Marjan Cugmas

Logo Design by: Andy Posner  
L<sup>A</sup>T<sub>E</sub>X template: Vladimir Batagelj

# Programme

# Programme of the 26th Young Statisticians Meeting

## Friday, September 30, 2022

17:30 – 19:00 Registration at Hostel pod Voglom

## Saturday, October 1, 2022

8:00 – 9:00 Registration at Hostel pod Voglom

9:15 – 9:30 Welcome address

9:30 – 11:15 **Session 1**

Chair: L. Kronegger

**Erich Batzella** (Italy): Associations between mixture of perfluoroalkyl substances and lipid profile in a highly exposed adult community in the Veneto Region.

**Montserrat González Garibay** (Slovenia): Does Machine Learning Offer Added Value Vis-à-Vis Traditional Statistics?

**Petra Parrag** (Hungary): The characteristics of Hungarian lung cancer patients discovered in 2018

11:15 – 11:45 **COFFEE BREAK**

11:45 – 13:15 **Session 2**

Chair: M. Cugmas

**Nikolina Milinčević** (Croatia): Extremes of random clusters with application to some marked renewal cluster processes

**Roman Parzer** (Austria): Bayesian Dimension Reduction for Regressions in High Dimensions

**Kristjan Bolčič** (Slovenia): A Bayesian approach to non-life insurance premium calculation

13:15 – 14:15 **LUNCH**

14:15 – 15:45 **Session 3**

Chair: D. Simon

**Máté Szalai** (Hungary): Stochastic modelling of the bactericidal potency

**Ana Martinčić Špoljarić** (Croatia): Modelling of Telomere Loss Process

**Vladimir Brajković** (Croatia): The importance of the mitogenome effect on milk traits in cattle

15:45 – 16:10 **COFFEE BREAK**

16:10 – 17:40 **Session 4**

Chair: P. Filzmoser

*Jonas Brugger* (Austria): Methodology of survival analysis under non-proportional hazards

*Špela Orehek* (Slovenia): Multilevel modelling in social science research

19:00 **DINNER**

## **Sunday, October 1, 2022**

9:30 – 10:20 **Session 5**

Chair: L. Kronegger

*Emanuele Koumantakis* (Italy): Suicide spectrum among young people during COVID-19 pandemic: a systematic review and meta-analysis

*Veronica Casotto* (Italy): Excess mortality in Veneto (Italy) during COVID-19 pandemic 2020-2021: time-series analysis

10:20 – 10:50 **COFFEE BREAK**

9:30 – 10:20 **Session 6**

Chair: N. Šuvak

*Ivan Ivkovic* (Hungary): Simulating Fractal Noise Driven Stochastic Processes

*Maximilian Ofner* (Austria): Reconstruction of functional data with missing fragments via factor models

11:00 **CLOSING**

# Abstracts

## Life after discharge from Pediatric Intensive Care Units: effects of air pollution on hospital readmission in high-risk children

Erich Batzella, Maryam Zare Jeddi, Gisella Pitter, Francesca Russo, Tony Fletcher and Cristina Canova

**Background and aim:** Residents of a large area in the Veneto Region (North-Eastern Italy) were exposed to drinking water contaminated by perfluoroalkyl substances (PFAS) for decades. While exposure to PFAS has been consistently associated with elevated serum lipids, combined exposures to multiple PFAS have been poorly investigated. Utilising different statistical approaches, we examined the association between chemical mixtures and lipid parameters.

**Methods:** Cross-sectional data from the Regional health surveillance program (34,633 individuals aged 20-64 years) were used to examine the combined effects of PFAS mixture (Perfluorooctanoic acid (PFOA), perfluorooctane sulfonate (PFOS), perfluorononanoic acid (PFNA), perfluorohexane sulfonate (PFHxS)) on total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C). Weighted Quantile Sum (WQS) regression, Quantile-based G-computation (Q-Gcomp) and Bayesian Kernel Machine Regression (BKMR) were used based on their ability to handle highly correlated chemicals.

**Results:** We observed that each quartile increase in the WQS index was associated with an increase in the levels of TC ( $\beta$ : 4.09, 95% CI: 3.47-4.71), HDL-C ( $\beta$ : 1.13, 95% CI: 0.92-1.33), and LDL-C ( $\beta$ : 3.14, 95% CI: 2.65-3.63). Q-Gcomp estimated that a quartile increase in the PFAS mixture was associated with increased TC ( $\psi$ : 4.04, 95% CI 3.5-4.58), HDL-C ( $\psi$ : 1.07, 95% CI 20.87-1.27) and LDL-C ( $\psi$ : 2.71, 95% CI 2.23-3.19). In the BKMR analysis, the effect of PFAS mixture on serum lipids increased significantly when their concentrations were at their 75th percentiles or above, compared to those at their 50th percentile. All methods revealed a major contribution of PFOS and PFNA, although the main exposure was due to PFOA. We found suggestive evidence that associations varied when stratified by gender.

**Conclusions:** The PFAS mixture was positively associated with lipid parameters, regardless of the applied method. Very similar results obtained from the three methods may be attributed to the linear positive association with the outcomes and no interaction between each PFAS.

## **Does Machine Learning Offer Added Value Vis-à-Vis Traditional Statistics? An Exploratory Study on Retirement Decisions Using Data from the Survey of Health, Ageing, and Retirement in Europe (SHARE)**

Montserrat González Garibay, Andrej Srakar, Tjaša Bartolj and Jože Sambt

Do machine learning algorithms perform better than statistical survival analysis when predicting retirement decisions? This exploratory article addresses the question by constructing a pseudo-panel with retirement data from the Survey of Health, Ageing, and Retirement in Europe (SHARE). The analysis consists of two methodological steps prompted by the nature of the data. First, a discrete Cox survival model of transitions to retirement with time-dependent covariates is compared to a Cox model without time-dependent covariates and a survival random forest. Second, the best performing model (Cox with time-dependent covariates) is compared to random forests adapted to time-dependent covariates by means of simulations. The results from the analysis do not clearly favor a single method; whereas machine learning algorithms have a stronger predictive power, the variables they use in their predictions do not necessarily display causal relationships with the outcome variable. Therefore, the two methods should be seen as complements rather than substitutes. In addition, simulations shed a new light on the role of some variables—such as education and health—in retirement decisions. This amounts to both substantive and methodological contributions to the literature on the modeling of retirement.

## **The characteristics of Hungarian lung cancer patients discovered in 2018**

Petra Parrag

According to international studies on malignant tumours, lung cancer is the most frequently diagnosed and has the greatest mortality rate. As to the comparative international estimates of GLOBOCAN conducted by the World Health Organisation (WHO), Hungary has the greatest number of newly diagnosed patients (incidence) and death rate (mortality) of lung cancer. However, regarding Hungary the methodology does not use the population-based incidence data for the estimates, but it is calculated from the ratio of mortality and incidence of neighbouring countries. The mortality is estimated from the mortality data provided by the Hungarian Central Statistical Office (HCSO).

The study explains the characteristics of Hungarian lung cancer patients based on a cleaned database of the year 2018. We used a system of rules of criteria in which we involved the institutes who reported the patients to exclude those, who were falsely reported to have lung cancer. The study also gives descriptive data on the excluded patients. Moreover, to answer the questions of how are the error rates of the doctoral reports change over the years and how can the incidence between the years of 2001 and 2020 be revised, based on the measured and cleaned data of 2018, the study also uses the databases of the years between 2001 and 2020, but these were not cleaned with the rules of criteria.

Last but not least, based on global observations, the operation of population-based cancer registries was fundamentally affected by the COVID-19 pandemic, which resulted in the decline of the number of patients discovered. Therefore, the study also addresses the impact of the pandemic on the number of lung cancer patients of the country in 2020.

Beside descriptive statistics for the year of 2018, direct, age-standardised incidence and mortality are presented for the period of the last two decades. Furthermore, survival analysis is conducted by using Kaplan-Meier curves (with Log-rank and Bonferroni correction to compare the curves within a variable) and a sample of the patients are analysed with Cox Proportional-Hazard Model. Hot deck imputation with the Predictive Mean Matching method is used to deal with the item non-response cases. The proportional hazard assumption is tested with the Schoenfeld test. Finally, the annual number of newly reported cases are presented for the last two decades: those being measured, those being amended based on the cleaned database of 2018 and those which would have been reported if the pandemic did not take place.

## **Extremes of random clusters with application to some marked renewal cluster processes**

Nikolina Milinčević, Bojan Basrak, Petra Žugec

We consider i.i.d. random variables  $X_i$ ,  $i \in \mathbb{N}$ , that belong to maximum domain of attraction  $\text{MDA}(G)$  for some extreme value distribution  $G$  and impose conditions on  $X_i$  and integer-valued positive random variable  $K$  so that the maximum of a random cluster  $\max_{i \in \{1, \dots, K\}} X_i$  belongs to the same  $\text{MDA}(G)$ . We emphasize that  $K$  and  $X_i$  are not necessarily independent. We apply our result to variants of Neyman-Scott, Bartlett-Lewis and randomly marked Hawkes process.

# Bayesian Dimension Reduction for Regressions in High Dimensions

Roman Parzer, Laura Vana Gür

We aim to find a dimension reduction technique in high-dimensional ( $p \gg n$ ) and correlated data settings to explain and predict relevant quantities. A common approach in these high-dimensional settings is to use some frequentist dimension reduction method first and then fit a predictive model on the reduced space. However, many reduction methods have poor performance or are computationally infeasible for high dimensions.

Additionally, such two-step approaches fail to keep track of the uncertainty in the reduction for the predictions. We propose several adaptations and extensions to an existing Bayesian method called Targeted Random Projection (TARP, Mukhopadhyay and Dunson 2020). It combines a probabilistic variable screening step with a random projection step to obtain a sparse set of reduced variables. These are used to fit a simple linear model for prediction. This procedure of screening, projection and prediction is repeated several times to explore different reductions. The obtained set of smaller predictive models is then aggregated by model averaging to get overall predictions.

New insights into this method are pointed out including suggested adaptations in the targeting step and possible robust extensions to handle response outliers. In extensive simulations and a real data application we show the advantages of these adaptations and extensions over the original method and other competitors.

## References:

- 1 Mukhopadhyay, M., & Dunson, D. B. (2020). Targeted random projection for prediction from high-dimensional features. *Journal of the American Statistical Association*, 115(532), 1998-2010.

## **A Bayesian approach to non-life insurance premium calculation**

Kristjan Bolčič

In the field of non-life insurance pricing, ratemaking for a new insurance product is often difficult without some prior knowledge and enough data. Even then, the final result is often subject to a strong actuarial judgment under uncertainty. Actuaries in these circumstances seek external information from other industries or similar lines of business, but often fail to incorporate uncertainty of external information in their models.

For this purpose we show an approach using methods from Bayesian statistics to model risk premiums i.e. the minimal amount needed to cover basic insurance expenses (claims) for small sample size insurance products. We present a method using generalized linear models to effectively incorporate prior external knowledge using informative and power priors into risk premium models, giving us a tool to describe uncertainty under different sets of priors. We further develop a method for sampling from the posterior predictive distribution of risk premiums, representing the distribution of profit and loss statements.

## Stochastic modelling of the bactericidal potency

Máté Szalai

Branching processes are classical tools to model cell proliferation. However, for estimation of bactericidal potency of antibiotics only deterministic models were used. We develop a stochastic model for bacterial growth, where the size of the bacterial population follows a Galton-Watson branching process, and the offspring distribution depends on the concentration of the antibiotic.

For a fixed antibiotic concentration  $c > 0$  we provide an estimator for the offspring mean  $m(c)$ , and show that the estimator is weakly consistent and asymptotically normal. Assuming that the offspring mean has the form  $m(c) = 2/(1 + \alpha c^\beta)$  with a suitable  $\alpha > 0, \beta > 0$ , we obtain an estimator for the parameters  $(\alpha, \beta)$  and investigate its properties. We provide a consistent estimator for the minimal inhibitory concentration (MIC), which is the smallest concentration preventing bacterial growth, an important parameter in microbiology.

As a real data we used measurements of *Chlamydia trachomatis* growth which was analyzed by a novel quantitative PCR method treated by 4 different antibiotics at 12 different antibiotic concentrations. We found that our model fits very well to this data.

The talk is based on ongoing joint work with Péter Kevei and Dezső P. Virok (University of Szeged).

## Modelling of Telomere Loss Process

Ana Martinčić Špoljarić

In 1961 Moorhead and Hayflick first described the molecular mechanisms that control the limited number of human cell divisions, i.e., the limited growth capacity of the cell. Numerous studies point to the telomeres, the end of chromosomes, as the underlying cause of the cellular (or replicative) senescence, which is considered to be the basis of organismal aging. In 1980 Smith and Whitney conducted the experiment in which they described the heterogeneity in proliferative potential of a single clone of normal human fibroblast. They also showed that the degree of difference in doubling potential between two sister cells coming from the same mitosis event may vary significantly.

Studying the telomere shortening mechanism is important for understanding cellular aging in biology. Presented model is trying to capture telomere-controlled cell senescence by considering two modes of telomere shortening – gradual and abrupt shortening. In order to investigate how the stochastic behaviour of telomeres affects the cell with time, time – discrete processes of telomere length are approximated by the time – continuous processes. Dependent on sources of randomness included in the model, we arrive at limit processes with continuous trajectories and processes with jumps. Their limit distributions and asymptotic properties are analysed. By computational simulations of telomere dynamics, we replicate Smith and Whitney's original experiments and demonstrate that this research sets a valid biological and mathematical model that is able to explain stochastic nature of cell aging.

## The importance of the mitogenome effect on milk traits in cattle

Vladimir Brajkovic, Ivan Pocrnic, Miroslav Kaps, Marija Špehar, Maria Lie Selle, Vlatka Cubric-Curik, Strahil Ristov, Ivica Medugorac, Dinko Novosel, Gregor Gorjanc, Ino Curik

Bovine mitochondrial DNA (mitogenome) contains 16338 base pairs encoding the synthesis of 13 proteins essential for cellular energy production and inherited exclusively through the maternal line, providing a genetic mechanism for cytoplasmic inheritance. Intensive efforts have been made to improve livestock production by understanding the relationship between DNA sequence and economically important production traits through genomic selection. Here we present the effects of the mitogenome on milk production traits in Croatian Holstein cattle, which is rarely studied compared with the nuclear genome. Complete mitogenomes representing 109 maternal lines were sequenced using NGS technology. This allowed the assignment of mitogenome sequence information to 2393 cows with 7115 milk production records (milk, fat and protein yield) using the in-house software MaGelLAN 1.0. A quantitative genetic model was applied to estimate the proportion of total variance explained by mitogenome inheritance ( $m_2$ ) using five different models: (1) cytoplasmic model with maternal lineages ( $m_2$ CYTO), (2) haplotypic model with mitogenome haplotypes ( $m_2$ MITO), (3) amino acid model with unique amino acid combinations ( $m_2$ AMIN) and (4) evolutionary model based on BEAST phylogenetic analysis ( $m_2$ EVOL). The effects of animal, sex chromosome, parity, calving season, region, year, herd and age at first calving were also included in each model. The estimated proportions of phenotypic variances explained by mitogenome inheritance ( $m_2$ CYTO,  $m_2$ MITO,  $m_2$ AMIN) ranged from 4% to 7% for all three milk traits. At the same time, the estimated proportions of phenotypic variances explained by  $m_a$  and by sex chromosomal loci were negligible. The obtained results indicate that a considerable proportion of phenotypic variance in milk traits is explained by mitogenome variation. A better understanding of the use of mitogenome inheritance can play an important role in practical breeding and remains a challenge.

## **Methodology of survival analysis under non-proportional hazards**

Jonas Brugger, Martin Posch, Franz König

Time-to-event analysis or survival analysis refers to a set of methods that examine the time passed until the occurrence of an event. It is unique in the sense that in general events will not be observed for all instances due to follow up losses or trial termination and therefore the true survival time is unknown for a subset of observations. This phenomenon is commonly referred to as censoring. Traditional methods like the Cox proportional hazards model and the Log-rank test base inference on the assumption that the effect of a risk factor on the failure rate of instances is constant over time. More specifically, the survival curves of two or more groups are proportional at any time  $t$ .

In novel immuno-oncology trials this assumption often does not hold, due to the unique mechanism of action of these drugs. In this scenario methods that assume proportionality are no longer optimal in terms of power. A wide variety of methods specifically designed to take this behaviour into account have been proposed. Their use depends largely on the intended interpretation and the specific scenario at hand.

## **Multilevel modelling in social science research**

Špela Orehek

Recently, the multilevel modelling approach has been an increasingly popular statistical method for analysing survey data, especially in the field of social sciences. The terms hierarchical linear modelling, linear mixed model or random coefficient model are also often used in the literature, referring to the same method. Multilevel modelling means an analysis of data that is structured on two or more levels (groups), for example, individual level (socio-psychological characteristics, e.g. attitudes, norms etc.) and organizational level (organizational processes, e.g. communication). Basically, the multilevel model is derived from a simple (or multiple) linear regression and the purpose of the analysis is the explanation of the dependent variable, combining the influence of independent variables from different levels with their possible interactions (called cross-level interaction; for example, norms and communication). Thus, it allows us to more accurately and adequately interpret the relationships and influences between the studied variables. The presentation will include a brief introduction to multilevel modelling with its assumptions and requirements, based on an example of information security behaviour model.

# Suicide spectrum among young people during COVID-19 pandemic: a systematic review and meta-analysis

Emanuele Koumantakis

**Background:** There are concerns that suicidal behaviors are arising among adolescents. The COVID-19 pandemic could have worsened the picture, however, studies on this topic reported contrasting results. This work aimed to summarize findings from the worldwide emerging literature on the rates of suicidality among young people during the COVID-19 pandemic.

**Methods:** A systematic review and meta-analysis was performed, searching five electronic databases for studies published from January 1st, 2020 until July 27th, 2022. Studies reporting rates for each of the three considered outcomes (suicide, suicidal behaviors, and suicidal ideation) among young people under 19 years old during the COVID-19 pandemic were included. Random-effects meta-analyses were conducted, and the intra-study risk of bias was assessed. When pre-COVID-19 data were available, incidence rate ratio (IRR) and prevalence ratio (PR) estimates were calculated between the two periods. The review protocol was registered on PROSPERO (CRD42022308014).

**Findings:** 47 observational studies were selected for more than 65 million subjects. The results of the meta-analysis show a pooled annual IR of suicides of 4.9 cases/100,000 during 2020, accounting for a non-statistically significant increase of 10% compared to 2019 (IRR 1.10, 95% CI 0.94-1.29). The suicidal behaviors pooled prevalence during the COVID-19 pandemic was higher in the psychiatric setting (25%; 95% CI: 17-36%) than in the general population (3%; 1-13%) and ED (1%; 0-9%). The pooled rate of suicidal ideation was 17% in the general population (11-25%), 36% in psychiatric setting (20-56%) and 2% in ED (0-12%). The heterogeneity level was over 97% for both outcomes in all settings considered. The comparison between before and during COVID-19 periods highlighted a non-statistically significant upward trend in suicidal behaviors among the general population and in ED setting. The only significant increase could be observed for suicidal ideation in psychiatric setting, among studies conducted at a greater temporal distance from the pandemic outbreak (PR 1.15; 95% CI: 1.04-1.27), not observed among those exploring 2020 alone.

**Interpretation:** During the pandemic, suicide spectrum issues seemed to follow the known pattern described in previous studies, with higher rates of suicidal ideation than of suicidal behaviors and suicide events. Governments and other stakeholders should be mindful that youth may have unique risks at the outset of large disasters like the COVID-19 pandemic and proactive steps are necessary to address the needs of youth to mitigate those risks.

## Excess mortality in Veneto (Italy) during COVID-19 pandemic 2020-2021: time-series analysis

Veronica Casotto, Claudio Barbiellini Amidei, Nicola Gennaro, Laura Cestari, Elena Schievano, Manuel Zorzi, Ugo Fedeli

**Objectives:** Excess overall mortality is one of the most reliable approaches for measuring the impact of the COVID-19 pandemic. The excess can be estimated by measuring the difference between mortality from all causes that is observed during the pandemic and mortality from all causes that would be expected if the pandemic had not occurred (baseline mortality). This study focuses on examining the excess of all-cause mortality during the 2020 COVID-19 pandemic in the Veneto Region, in the North-East of Italy.

**Methods:** Data on all-cause mortality from January 1, 2008 to December 31, 2020 were obtained from the Veneto Region mortality register that includes all death certificates of residents in Veneto. We analyzed monthly age-standardized death rate, applying population weights from the 2013 European standard.

We explored four approaches for estimating expected mortality: (1) the average count of deaths observed during 2018-2019; (2) mean age-standardized death rate registered in 2015-2019 (3) Seasonal Autoregressive Integrated Moving Average (SARIMA) model and (4) Generalized Estimating Equations (GEE) model for grouped data; the latter two approaches were based on monthly age-standardized mortality rates registered in 2008-2019.

**Results:** In 2020 there were overall 56,973 deaths; the percentage increase with respect to the average number of deaths observed in 2018-2019 was 17.2%. The increase in the age-standardized mortality rate with respect to 2015-2019 average was +9.5%. A SARIMA (1,0,0)(2,1,0)<sub>12</sub> with drift model was selected as the best model which minimized the value of the Akaike Information Criteria: the observed age-standardized rate in 2020 was higher than that predicted, corresponding to a +15.2% of percentage change. The age-standardized rate in 2020 was 15.8% higher than that predicted by the GEE model.

**Conclusions:** Excess mortality in 2020 estimated based on models taking into account the long-term decreasing trend was substantially higher than simple comparison with the average age-standardized rate of the previous five years.

# Simulating Fractal Noise Driven Stochastic Processes

Ivan Ivkovics

There is now a great deal of interest in parameter estimation of Ornstein-Uhlenbeck processes with fractional driven noise in finance because of its modelling capability. That is why the ELTE AI research group, which I am a member of, aims at investigating the prediction of the unknown parameters of certain transformed fractional Ornstein-Uhlenbeck processes, e.g. Stochastic Correlation Processes, with neural networks in the hope that more accurate estimators will be obtained than the ones using classical statistical methods. The importance of an efficient data generator system for each analyzed process rises high in this case, since if complex neural network structures are applied in the learning procedure, then a big amount of data is needed for a good performance.

I have developed an efficient generator system for stochastic integrals with respect to fractional Wiener process, which contains many models used in mathematical finance, e.g. the fractional Ornstein-Uhlenbeck processes. Since the integrand functions are assumed to be Hölder-continuous to avoid the task of simulating Skorohod-integrals, the obtained processes are path-wise Riemann-Stieltjes integrals, i.e. the core of the simulation procedure is the subroutine of generating the fractal noise. This observation let me generalize this class of stochastic processes to the isonormal processes, which can be simulated efficiently by the fastest exact generator methods, e.g. by the circulant matrix embedding based methods.

Isonormal processes are Gaussian processes, where the covariance structure is defined by a kernel space endowed with an inner product. I have developed a class for generating isonormal processes determined by an arbitrary inner product structure and to simulate the special cases, i.e. the stochastic models. In case of fractional Ornstein-Uhlenbeck processes with zero initial value, 10000 sequences over 1500 time grid points can be simulated more than five times faster than using the best open-source codes, which computation is also necessary for the Heston pricing model with stochastic correlation. This library is being developed in Python and will be published open-source.

## **Reconstruction of functional data with missing fragments via factor models**

Siegfried Hörmann, Maximilian Ofner

Functional data analysis gained popularity in the last decades, not least due to modern devices and an enormous increase in densely measured data. Empirical work, however, shows that such data are often prone to missing values. In this talk, we propose an easy-to-use algorithm to re-impute the missing data.

Our method is based on factor models which naturally arise from discretely sampled functions with additive noise. We comment on this relationship and construct estimators for the factors and loadings in the presence of missing data. The corresponding estimator for the common component is then used to recover the missing fragments. It is shown that this procedure allows for consistent reconstruction. The established convergence rates are uniform and hold without restrictive smoothness assumptions.

# Participants

Arslan Talha (Turkey): mstalhaar@gmail.com  
Barnet Viktoria (Hungary): barnetviki@gmail.com  
Batagelj Vladimir (Slovenia): vladimir.batagelj@fmf.uni-lj.si  
Batzella Erich (Italy): erich.batzella@unipd.it  
Bersia Michela (Italy): michela.bersia@edu.unito.it  
Bolčič Kristjan (Slovenia): kristjan.bolcic@windowslive.com  
Brajkovic Vladimir (Croatia): vbrajkovic@agr.hr  
Brugger Jonas (Austria): jonas.brugger@meduniwien.ac.at  
Casotto Veronica (Italy): veronica.casotto@studenti.unipd.it  
Comoretto Rosanna (Italy): rosannairene.comoretto@unito.it  
Cugmas Marjan (Slovenia): marjan.cugmas@fdv.uni-lj.si  
Ferligoj Anuška (Slovenia): anuska.ferligoj@fdv.uni-lj.si  
Filzmoser Peter (Austria): Peter.Filzmoser@tuwien.ac.at  
Gonzalez Garibay Montserrat (Slovenia): gonzalezm@ier.si  
Grimaldi Piercesare (Italy): piercesare.grimaldi@unito.it  
Ivkovic Ivan (Hungary): ivkovic@renyi.hu  
Košuta Tina (Slovenia): tina.kosuta@mf.uni-lj.si  
Koumantakis Emanuele (Italy): emanuele.koumantakis@unito.it  
Martinčić Špoljarić Ana (Croatia): ana.martincic@rimac-technology.com  
Milinčević Nikolina (Croatia): nikolina.milincevic@math.hr  
Ofner Maximilian (Austria): m.ofner@tugraz.at  
Orehek Špela (Slovenia): spela.orehek@fdv.uni-lj.si  
Parrag Petra (Hungary): petfun96@icloud.com  
Parzer Roman (Austria): roman.parzer@tuwien.ac.at  
Srakar Andrej (Slovenia): srakara@ier.si  
Szalai Máté (Hungary): szalai.mate21@gmail.com  
Šuvak Nenad (Croatia): nsuvak@mathos.hr  
Uhljar Janka (Hungary): jannuska6@gmail.com  
Vratanar Bor (Slovenia): bor.vratanar@gmail.com