22w5079 Combining Causal Inference and Extreme Value Theory in the Study of Climate Extremes and their Causes

Johanna G. Nešlehová (McGill University),
Marloes Maathuis (Eidgenössische Technische Hochschule Zürich),
Linda Mhalla (École Polytechnique Fédérale de Lausanne),
Philippe Naveau (Laboratoire des Sciences du Climat et de l'Environement)

June 26, 2022 –July 1, 2022

1 Overview of the Field

In a changing climate, the frequency and severity of climate extremes is of prime importance. It is essential to understand and quantify how the distributional behavior of extremes will differ from the past and how it might evolve in the future. To answer these questions, a rigorous mathematical and statistical framework needs to be developed to assess and quantify causal effects on and of extremes. The emerging area of causal inference for extremes calls for a combined research effort of three scientific fields: causal inference, extreme-value theory, and climatology.

Traditional approaches based on correlation, regression, and Gaussianity are ill-suited for this purpose; they do not provide sufficient insights into the causal mechanisms underlying the dynamics of a system, and underestimate the frequency and severity of extreme phenomena. In contrast, causal inference and extreme-value theory constitute the foundations on which a suitable statistical methodology can be built. On the one hand, causal inference goes beyond the assessment of association between variables in that it can also predict the behavior of the outcome variable when the system is changed, i.e., when some of its drivers have been changed by external manipulations. On the other hand, extreme-value theory allows to extrapolate beyond the observed data, and as such is indispensable for the study of extreme phenomena.

Quantification of causal effects on and of extremes is particularly pressing in the face of global warming. Extreme event attribution in the climate sciences is at the crossroad of statistics, atmospheric dynamics, and human sciences. Its goal is to assess whether and how the probability or magnitude of an event, typically a damaging extreme phenomenon such as a hurricane or a heatwave, has been altered by anthropogenic climate change. Combining recent advances in causality and extreme-value theory allows to study the mathematical and statistical problems around extreme event attribution. This includes causal discovery, causal structure learning for extremes, causal evidencing and quantifying causal effects involving extremes, as well as uncertainty quantification of the resulting estimates.

2 Overview of the Workshop

The aim of the BIRS workshop on *Combining Causal Inference and Extreme Value Theory in the Study of Climate Extremes and their Causes* was to bring together junior researchers and senior experts from the three scientific communities, climatology, causal inference, and extreme-value theory, to develop new collaborations, identify new problems and advance the state of the art of the budding area of causal inference for extremes.

3 Presentation Highlights

3.1 Overview Talks

Climatology

To set the tone of this workshop, it was important to start with the main domain of application of this event. As any research field, climatology has its own topics, vocabulary and data analysis tools. It was key that statisticians working on causality and/or extreme value theory possessed the main ingredients of the climate research in terms of extreme events at the beginning of this workshop.

Dr. **Aurélien Ribes** (CNRM, Meteo France, Toulouse), leads a team of climatologists at the forefront of the study of climate change. He was able to clearly explain the main aspects of the fields of the so-called climate detection and attribution. As greenhouse gazes have been steadily increasing during this last century, a recurrent question after any large scale extreme event, e.g. the European 2003 heatwave, is to wonder if the probability of such extreme events is *due* to the increase of anthropogenic forcing or to other causes. In this context, Aurélien Ribes presented different statistical concepts to answer this difficult question. Modelling uncertainties is paramount in this problem. Observational measurements are tainted by instrumental errors, numerical climate models are approximation of our true climate, natural variability of the climate system is also key. In addition, the treatment of extremes is challenging due to their rare nature by definition. Also, framing causality is difficult here as no counterfactual Earth, i.e. observations without the impact of increasing greenhouse gazes, exists. Aurélien Ribes reviewed in details all of these uncertainties and he offered a blueprint to integrate all these issues under a Bayesian regression model that was based on the generalized extreme value distribution. He also proposed different perspectives and mathematical challenges to improve the statistical techniques used in his domain. His talk spurred various questions and exchanges among the participants.

Extreme Value Theory

Prof. **Antony Davison** (EPFL, Lausanne) presented in a didactical way the fundamental elements of univariate and multivariate Extreme Value Theory and their applications to geosciences. His talk started by recalling the probability backbone of this theory, following by a presentation of a variety of statistical models based on this theory. The inference strategy to estimate the parameters of such models was also at the core of his talk. As the goal of this workshop was to mix communities, his presentation highlighted a few cases studies focusing on geophysical applications, in particular extremes of river networks. For such climate and hydrological data analysis, a key element is the spatial aspect of extremes events. The modeling question is to how to capture adequately the spatial dependencies among different locations. The statistical treatment of such random fields differs from the classical Gaussian spatial analysis. A large part of the presentation was dedicated to this important question, theoretically challenging and important for practitioners. This vast panorama of the multivariate Extreme Value Theory field ended with questions and interactions among participants.

Causal Inference

In his talk, Prof. **Linbo Wang** (University of Toronto, Canada), gave an introduction to causal inference through two formal frameworks: the counterfactual framework [4, 6] and the framework of the structural equations models [3, 5]. He first set up the terminology and notations to facilitate communication between participants from different fields. Then, he outlined the assumptions under which causal effects are identifiable within both frameworks. Inference for causal effects in then tackled. Under the assumption of absence of hidden confounders, identifiability conditions allow us to estimate average causal effects through a regression model or using the propensity score under the inverse probability weighting approach. Finally, he introduced the notion of instrumental variables that in central for causal inference in the presence of latent confounders.

Under additional assumptions on the instrumental variables, identifiability and estimation of the causal effect is made possible in settings where common causes (of the treatment and the effect) are unobservable.

3.2 Climatology

This list of highlighted talks, although non-exhaustive, represents a sample of the main topics addressed by the applied scientists at this workshop.

In her talk, **Prof. Gabi Hegerl** (Edinburgh University) considered changes in temperature extremes over the historical period and into the future. Hence, this presentation nicely complemented the introductory talk by A. Ribes on climate change. The topic moved from the global scale towards the analysis of extremes at the regional level. Historical regional temperature extremes show strong variability in the past. Over many regions, climate model simulated future changes in extreme temperature events seem to most clearly show a shift towards warmer extremes without much evidence for a change in other aspects of the distribution, yet there appear to be some exceptions. The team of Prof. Gabi Hegerl has experimented with observational constraints on changes in extremes and results suggest that selecting more realistic models can influence simulated future changes, particularly in the tropics. This analysis brought the important statistical question of blending datasets with different uncertainties.

Dr. **Dáithí Stone** (NIWA, Wellington, NZ) also focused about the role of anthropogenic emissions in specific extreme weather events, ranging from methods based on free-running atmosphere-ocean climate models through to methods based on highly constrained numerical weather forecasts. He explained that, while technically these different experiment designs are each addressing different particular questions, in practice some methods may be more feasible or scientifically defensible than others. In this talk, five different extreme events occurring over Aotearoa New Zealand during the past four years, using atmosphere-ocean climate model experiments, atmosphere-only model experiments, numerical weather forecast experiments, and reanalysis experiments were examined. He concluded that there is a strong dependence on experiment design for extreme hot events, but that any dependence for extreme wet events is small in relation to various sampling uncertainties. This talk emphasized the rich potential of blending statistical experimental design methods and climate analysis.

Dr. **Thordis Thorarinsdottir** (Norwegian Computing Center, Oslo) shifted from climate issues to impact studies and hazards. In particular, she focused on the consistent estimation of extreme precipitation and flooding across multiple durations, an important topic for infrastructure design that commonly requires assessments of extreme quantiles of precipitation and flooding, with different types of infrastructure requiring estimates for different durations. Her modeling strategy nicely combined mathematical developments based on generalized extreme value (GEV) distribution and time series analysis with multiple durations constraints to ensure that, e.g., the 0.99 quantile of annual maxima of 2 hour precipitation is larger than that for 1 hour precipitation.

Prof. **Manuela Brunner** (ETH Zurich) brought another way to understand climate extremes. She analyzed in detail extreme droughts in Europe. This work opens a lot of links between the causality and climate communities. She proposed drivers of severe and moderate hydrological droughts in Europe. She demonstrated that moderate droughts are mainly driven by rainfall deficits while severe events are mainly driven by snowmelt deficits in colder climates and by streamflow deficits transitioning from the wet to the dry season in warmer climates. From a statistical perspective, the proposed classification scheme provides a template that can be expanded to include other climatic regions and human influences.

3.3 Extremes

Several advances in extreme-value theory were presented at the workshop, some but not all related to causality and climate extremes. Below are a few highlights, by no means exhaustive.

In her presentation, Prof. **Anna Kirliouk** (University of Namur, Belgium), presented a novel approach to inference for extreme regions of interest. Under the assumption of multivariate regular variation, the approach relies on the closed-form expression of the Tail Pairwise Dependence Matrix (TPDM) for max-linear models. For instance, she proposes to approximate failure probabilities, probabilities of critical regions, by approximating the observed process by a max-linear model with coefficients obtained though a matching procedure of the empirical version of the TPDM with the one induced by the max-linear model. The proposed methodology works for fairly high dimensions and is not computationally costly. The challenging part of the proposed algorithm resides in finding a completely positive decomposition of the TPDM, which the authors show to be exact when the process of interest follows a max-linear model. This work sets the stage for interesting settings where one is concerned with inference for failure probabilities when the underlying process is believed to be induced by a max-linear Bayesian network, for example.

Dr. **Gloria Buriticá** (Université Sorbonne) presented a novel statistical analysis, the so-called stable sums method, that takes into account the spatio-temporal dependencies of extreme rainfall, or more generally of a heavy-tailed process. Building on large deviations of regularly varying stationary time series, she proposed to infer marginal tail behavior at an individual site from a multivariate stable distribution model for appropriately scaled sums over disjoint blocks. This approach leads to improved estimation of high quantiles and return levels, as she illustrated on simulated data and employed to analyze heavy rainfall in northeastern France.

Prof. **Jenny Wadsworth** (University of Lancaster, United Kingdom) presented her ongoing, innovative work with her PhD student Ryan Campbel and UBC colleague Prof. Natalia Nolde on statistical inference for the geometric representation for multivariate extremes pioneered in [1], which is based on the shapes of sample clouds in light-tailed margins and their limit sets. She outlined a method for parametric estimation of the limit set shape, which includes a useful non/semi-parametric estimate as an initial step. She illustrated how appropriate parametric forms of the limit set lend flexibility to the model and showed how the approach can be used to extrapolate into the tail of the distribution via simulation from the fitted model.

In his presentation, Prof. **Richard Smith** (UNC Chapel Hill, North Carolina) first discussed a Bayesian model of temperature measurements in Kelowna, BC with global mean temperature outputted from a climate model as a covariate, in order to address how the return level of the June 2021 heatwave evolves in a changing climate. He then moved on to questions surrounding the characterization of extreme event probabilities and return levels spatially across a wide region. Focusing on an ongoing study of hurricane Harvey, he again contemplated the approach of treating global meteorological indicators, such as the mean sea surface temperature over a relevant area, as covariates and using existing work on causal determination for large-scale meteorological variables for causal inference about more localized extreme events like hurricane Harvey.

3.4 Causal Inference with Focus on Extremes

In his presentation, Dr. **Nicola Gnecco** (University of Geneva, Switzerland) presented a causal discovery method tailored for extremes events, bridging therefore two of the main topics of the workshop. He highlighted the need for such a method as causal relations may differ in the body of the distribution from the tail. Relying on observational data, he proposed a causal tail coefficient to uncover causal effects in the upper tail region of a distribution. Properties of this coefficient are developed under a structural causal model with heavy-tailed noise variables and an extension of its use in the multivariate setting was proposed through a greedy algorithm.

In his presentation, Prof. **Sebastian Engelke** (University of Geneva, Switzerland) presented a novel approach to modelling treatment effects under the potential outcome framework combined with the extreme value setting. The focus of his work being on the tail region of the distribution of the outcome, he investigated the size of the causal effect on the extreme quantiles of the potential outcomes. In addition to the assumption of regular variation of the potential outcomes that is necessary in such settings, he assumed the presence of observed confounders and proposed to take it into account through propensity scores. For instance, inference for extreme quantiles and thus the quantile treatment effect is performed based on estimates of intermediate quantiles and an adjusted (through propensity scores) Hill estimator of the tail index parameter. Asymptotic properties of the resulting estimator of the quantile treatment effect are discussed and the methodology is used to assess the effect of college education on wage distribution.

In his presentation, Dr. **Leonard Henckel** (University of Copenhagen, Denmark) discussed a novel approach to exploiting instrumental variables in causal effect estimation. He started with an overview on how and why instrumental variables are very useful in estimating average treatment effects in settings where counfounding is unobserved or cannot be measured. He then described the two-stage least square estimator under some validity assumptions on the instrumental variable. The goal of his work is to modify/relax these assumptions such that causal effects can still be identifiable and measured under a so-called independence condition rather than the usual covariance condition. Then, he proposed an algorithm to estimate the causal effect under unobserved confounding and using an instrumental variable, by maximizing the validity of the independence condition through its quantification by the Hilbert–Schmidt Independence Criterion (HSIC). The performances of the proposed algorithm are assessed in a simulation study and future expansions to measure, for instance, the quantile treatment effect are discussed.

In his presentation, Dr. **Jakob Runge** (German Aerospace Center, Germany) presented the work of his research group, which is driven by inspiration from various collaborations, and which aims to provide theory, methods, and software to run causal inference for data from Earth Systems Science [7]. He highlighted various challenges posed by the processes, data, and methods, such as autocorrelation, non-linear dependencies, time delays, different time scales, unobserved variables, discrete data, high-dimensionality and so on. He also explained the challenges with the definition of extremes and how they can possibly be represented in structural causal models. The final part of his presentation was devoted to the *causeme* platform, available at

`https://causeme.uv.es/`

The idea for it grew out of a competition at neurIPS; it provides researchers with the opportunity to upload their own method, test it on synthetic data provided on the platform, and have it automatically evaluated and compared with competing approaches. The ultimate goal is to help end-users to better choose methods suitable for their specific type of data.

4 Activities for Early-Career Researchers

One of the aims of the workshop was to put more focus on early-career researchers whose career paths and goals might have been affected more than expected by the covid period. We thus included a number of activities to help them grow scientifically, create a large and rich network, and of course benefit from the expertise and guidance of senior researchers among the participants.

With the perspective of allowing all participants to engage equally in the workshop, the (rather) usual poster session was removed from the program and senior and early-career researchers were given the opportunity to present and discuss their work during a 30mn talk. Regular talks were grouped in theme-based sessions that were chaired by both senior and junior researchers.

In the same train of thought, a mentoring session centered around some of the main topics related to scientific research was organised. With a list of panelists including four senior researchers, Valérie Chavez Demoulin (Université Lausanne, Switzerland), **Christian Genest** (McGill University, Canada), **Anthony Davison** (EPFL, Switzerland), and **Claudia Klüppelberg** (TU Munich, Germany), the early-career participants had the chance to prepare and ask questions related to different stages of a career in academia, such that networking and collaboration as well as publishing. Prof. Christian Genest gave a comprehensive overview of the various phases, aspects, and motivations behind the exercise of publishing scientific work. He discussed his personal view on good practices in publishing, gave advice to participants with the aim of increasing awareness about the goal of this exercise, its positive and negative impact on the scientific community as well various ways of maintaining honesty and integrity within the community.

5 Recent Developments and Open Problems

The scientific program of the workshop was concluded by a roundtable on future challenges, led by three early-career researchers present on site: **Emma Simpson** (University College London), **Leonard Henckel** (University of Copenhagen), and **Sebastian Engelke** (University of Geneva), as well as climatologist **Dáithí Stone** (NIWA). The four panelists identified various future challenges in the three scientific areas, climatology, causal inference, and extreme-value theory, and at their intersections:

- **Climatology**: The limitations of regional climate models were discussed, particularly in their ability to predict extreme weather, and the importance of statistical approaches to understand, model, and predict regional weather extremes. Another issue is the role of statistics in weather forecasting, particularly for the long-term. The discussions further highlighted the challenge of merging information from climate model outputs and observational data.

- **Extremes**: The roundtable discussion first evolved around stopping rules which were touched upon in several talks [2], as well as structural models for extremes, such as graphical models, max-linear structures, and interactions between covariates. Future challenges focused on consistent estimation across different scales, variables, and dimensions, the need for flexible tools for complex data, and machine learning techniques for extremes.

- **Causal Inference with focus on Extremes**: One of the main challenges are assumptions underlying existing causal inference methods, e.g., modularity or invariance, which may be hard to verify and indeed violated in applications in climatology. The definition of exposure may also be unclear. Synthetic controls were mentioned as a possible future avenue to explore. Furthermore, there is need for smooth transition in the causal mechanism to account for a changing climate: what is extreme today may no longer be extreme tomorrow. Questions also arise how to best combine synthetic and observational data, as well as a statistical vs. climate model based approaches to causal effect quantification. At last but not least, causal discovery problems remain difficult and there is much room for further research.

6 Outcomes of the Meeting

The workshop has been particularly valuable in that it helped to establish links between climate scientists and statisticians. Both communities sometimes use different terminology to describe the same phenomenon, and there is need to learn from each other to be able to communicate effectively and to truly profit from the scientific advances at both ends. Dáithí Stone's offer to organize a follow-up workshop in Aotearoa New Zealand was met with enthusiasm.

6.1 Scientific Outcomes

Several new collaborations have started at the workshop, and fruitful multiple discussions took place, both on site and on the slack platform. The workshop has been particularly valuable in that it helped to establish links between climate scientists and statisticians. Both communities sometimes use different terminology to describe the same phenomenon, and there is need to learn from each other to be able to communicate effectively and to truly profit from the scientific advances at both ends. Dáithí Stone's offer to organize a follow-up workshop in Aotearoa New Zealand was met with enthusiasm.

6.2 Reflections on the Hybrid Format

Of the 60 participants, 11 were onsite and 49 online. Despite the fact that 20 in-person slots were available, many participants were reluctant to travel due to various travel obstacles due to COVID-19. In addition, as one main topic of this workshop was climate change, a lot of researchers attach a great of importance at reducing their carbon footprint, and consequently, on-line attendance offer a great way to deal with this important issue. Also, several junior researchers could not come due to family constraints. This shows that going forward, hybrid format workshops have benefits, notably greater inclusiveness for female early-career researchers with young children, and much smaller environmental impact.

The superb, modern facilities and infrastructure at UBC Okanagan made the hybrid format very easy to implement. We found the lecture room to be very well equipped with a state-of-the-art recording system, high-resolution projector and microphones; this makes the interaction between on-site and on-line participants as easy as it gets. Listening to talks on site that were delivered on zoom was just as easy and engaging as following talks given in-person. The only downside we found was that the recordings of the talks were available online with a delay, which made it difficult for online participants to catch up with talks they missed. Furthermore, there seems to be no way to make the recordings available to the workshop participants only. This lead to some of the recordings not being shared at all, notably talks that discussed results in progress, the round table discussions, and the mentoring panel. In our opinion, it would have helped to make the recordings available quickly with restricted access to the workshop participants, and to publish those that are suitable for broad dissemination on the public workshop website later on.

These benefits are however offset by a number of challenges that are not easy to resolve. A particular challenge we faced was to accommodate multiple time zones, spanning from Saudi Arabia to British Columbia.

We felt compelled to make the schedule as inclusive as possible, so that everyone can attend all activities. This is why we compressed the entire scientific program to an approximately four hour window between ca. 8:30 and 12:30 local time. This constraint limited the number of talks we could accommodate and made for a cramped schedule with little room for breaks. Luckily, the local staff at UBCO was very helpful in providing a flexible format of coffee breaks. The compressed morning schedule had the benefit of free afternoons for local participants. This gave ample time for focused scientific discussions in small groups. Participants could also discover the fantastic environment of the Okanagan Valley through two excursions, to the Knox Mountain and the Frind Winery, respectively, take walks in nature around the campus, and visit Kelowna. This created a wonderful atmosphere and very friendly atmosphere between all on-site participants. One of them, Professor Richard Smith (UNC Chapel Hill) summarized his reflections in his own report which can be found on his website at

`http://rls.sites.oasis.unc.edu/UBCO.htm`

In contrast, engaging the on-line participants in the same way as those who attended in person was much harder. Although we created a platform on Slack for virtual discussions, it was clearly much less effective than personal discussions and spontaneous interactions onsite. The human aspect in scientific workshops is very important indeed and we did not find a good way to reproduce it online.

Despite these challenges, we succeeded at holding a vibrant workshop and received many positive comments from in-person as well as online participants. We are very grateful to BIRS for having given us the opportunity and for having provided a superb support, facilities, and assistance on site.

References

[1] G. Balkema and P. Embrechts, *High Risk Scenarios and Extremes: A Geometric Approach*, European Mathematical Society, 2007.

[2] L. Belzile and A.C. Davison, Improved inference on risk measures for univariate extremes, *arXiv* (2020), 2007.10780.

[3] T. Haavelmo, The statistical implications of a system of simultaneous equations, *Econometrica* **11** (1943), 1–12.

[4] J. Neyman, On the application of probability theory to agricultural experiments. Essay on principles, *Annals of Agricultural Sciences* (1923), 1–51.

[5] J. Pearl, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, 2000.

[6] D. B. Rubin, Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology* **66** (1974), 688–701.

[7] J. Runge, S. Bathiany, E. Bollt, G. Camps-Valls, D. Coumou, E. Deyle, C. Glymour, M. Kretschmer, M.D. Mahecha, J. Muñoz-Marí, E.H. van Nes, J. Peters, R. Quax, M. Reichstein, M. Scheffer, B. Schölkopf, P. Spirtes, G. Sugihara, George, J. Sun, K. Zhang, and J. Zscheischler, Inferring causation from time series in Earth system sciences, *Nature Communications* **10**(1) (2019), 2553.