Report for Project 1083 "Climate Informatics: New Machine Learning Methods for Climate Data and Climate Model Evaluation"

Project: **1083**
Project title: **Climate Informatics: New Machine Learning Methods for Climate Data and Climate Model Evaluation**
Project leader: **Prof. Dr. Jakob Runge**
Report period: **2021-05-01 to 2022-04-30**

Overview

In this period we have worked on a variety of aspects within the tasks proposed under the project. We have on the one hand analyzed and evaluated climate model data from the Coupled Model Intercomparison Project (CMIP) with reanalysis data focusing on causal networks and extreme events like droughts and heatwaves. On the other hand we have worked to create models for synthetic datasets that could serve as benchmark datasets for various causal discovery methods. Before we delve into the tasks individually, we would like to state that the ongoing global pandemic affected the use of our granted computing resources. Additionally, the hiring of two postdocs, for whom a part of the resources was intended, at the chair of Climate Informatics headed by Prof. Runge at the TU Berlin faced another hurdle as the IT-system of the university suffered a hack in March 2021, from which it is still recovering.

Report for Task (i) Development and application of latent causal discovery methodology for observations CMIP model evaluation

The ESMValTool (Eyring et al., 2020) preprocessor as well as existing and newly created diagnostics are used to preprocess ECMWF ERA-5 reanalysis and CMIP simulations. These data are used to study causal networks using PCMCI. Several topics are addressed including the evaluation of CMIP6 models based on causal networks related to Arctic-midlatitude teleconnections (Galytska et al., 2022 in prep.), the analysis of modes of climate variability and the evaluation of CMIP6 models based on the relationship between causal networks for sea level pressure data, which allows to define "fingerprints" of causal links for each model and precipitation patterns. The latter is based on a work from Nowack et al. (2020) for CMIP5 data and example of recent results can be seen in Fig. 1.

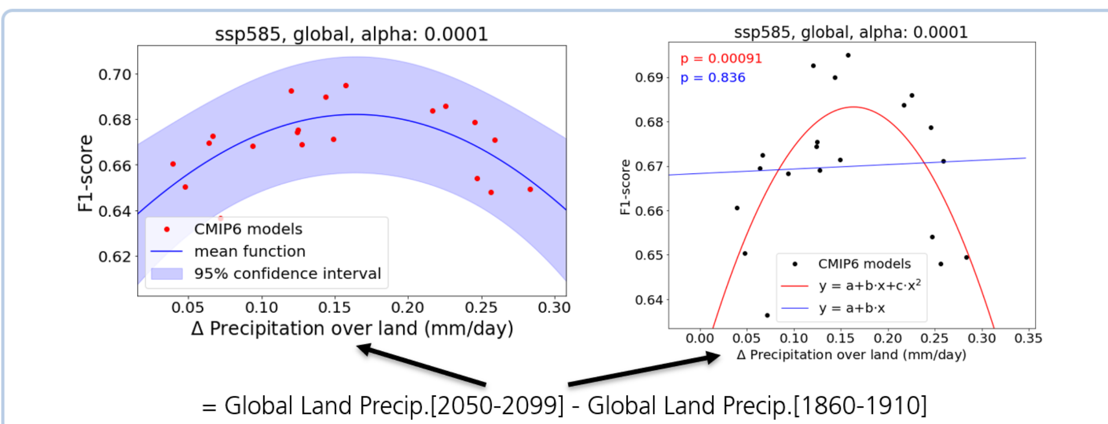

Figure 1: A parabolic relationship between the F1-score for causal networks of sea level pressure from CMIP6 historical runs and reanalysis data and precipitation change over land under SSP5-8.5 scenario is found (similar to Nowack et al. (2020) for the CMIP5 RCP8.5 scenario) This can in turn help to estimate future precipitation change over land under climate change more accurately.

Report for Task (ii) Development and application of mixed-type causal discovery for observations and CMIP model evaluation

In causal discovery tasks, the need to deal with mixed-type data is important because the choice of variables ranges from continuously-valued variables, such as temperature and pressure, to discretely-valued variables, such as climate indices with values representing different regimes (e.g. the Madden-Julien Oscillation). An essential subtask of causal discovery is conditional independence testing to determine whether the correlation between variables is explained by confounding due to a third variable. We constructed models for mixed-type data generation, and worked on the theoretical development of a conditional independence test. The testing of this model and the independence testing will consume most of the remaining node-h before the current reporting period ends. In particular, exploring and developing on existing Kernel-based methods for independence testing, as stated in the proposal, has been embarked on but has not used as many computational resources as requested. We attribute it to the delay in the hiring process for the postdoc positions in the Climate Informatics branch at TU Berlin caused by COVID-19 and the IT-hack of the entire university intranet system.

The analysis of modes of climate variability from task (i) was taken further to include non-continuous data due to distinct regimes with different causal networks using the Regime-PCMCI based on (Saggioro et al., 2020).

Report for Task (iii) Application of Causal Discovery to systematically find emergent constraints

Recent causal discovery methods have been developed to understand cause-effect phenomena in the climate as a two step procedure. One estimates the modes of variability through a dimensionality reduction technique, such as Varimax, since the modes or teleconnections are emergent properties of models. Subsequently one applies the specific causal discovery technique, such as PCMCI, (Runge et al., 2019),(Runge, 2020), (Gerhardus and Runge, 2020) that fits the best to the specific setting. A major challenge that lies here is the lack of sufficient ground truth datasets that can help benchmark the casual discovery techniques. We have worked on a simplified stochastic climate model that outputs gridded data and represents climate modes and their teleconnections through a Spatially Aggregated Vector-Autoregressive (SAVAR) model (Tibau et al., 2021). This model is used to construct benchmarks and evaluate the strengths and weaknesses of causal discovery techniques on spatiotemporal data. In this work, we also present a novel causal discovery method at the grid level that has orders of magnitude better performance compared to current approaches, called MappedPCMCI. Additionally, before the current reporting period ends we would still be using some of the remaining computational resources to generate about 5000 models.

Report for Task (iv) Development and application of causally explainable deep learning techniques for observations and CMIP model analysis

In the task of causal discovery even the atomic bivariate case, seemingly the simplest, is challenging and requires further assumptions to be identifiable at all. In recent years, a variety of approaches to address this problem has been developed, each with its own assumptions, strengths, and weaknesses. In machine learning, common benchmarks with real and synthetic data have been a main driver of innovation. Synthetic benchmarks can explicitly model data characteristics such as the underlying functional relations and distributions to assess how methods deal with these. However, a systematic assessment of the state-of-the-art of methods was missing so far. In (Käding and Runge, 2022) a detailed and systematic comparison of a range of methods on a novel collection of datasets that systematically models individual data challenges is presented. Further, we evaluate more recent methods missing in previous benchmarks. The novel suite of datasets will be contributed to the `causeme.net` benchmark platform to provide a continuously updated and searchable causal discovery method inter-comparison database. Our aim is to assist users in finding the most suitable methods for their problem setting and for method developers to improve current and develop new methods.

Report for Task (v) Application of extreme event detection machine learning techniques for CMIP model analysis

Classical drought indices (standardized precipitation index SPI and standardized precipitation evapotranspiration index SPEI) were computed with the ESMValTool for CMIP5 and CMIP6 data, see (Weigel et al., 2021). These indices where analyzed and are the basis for further investigations and comparisons to ML methods for drought detection. For this, experiments where done using an detection algorithm for anomalous interval (MDI, Barz et al. 2019). Additionally, several ML algorithms were compared for emulating the soil moisture index (SMI) obtained from a hydrological model with variables from ERA5-Land reanalysis data and land use information from MODIS satellite observation, including ablation study for coarser resolutions (Gottfriedsen et al., 2021). The publications Gottfriedsen et al. (2021) and Weigel et al. 2021 include results based on calculations for Project 1083. Heat waves were analyzed using Gaussian mixture models to detect extreme events and the change in their frequency using CMIP6 data sets under different global warming level scenarios (Paçal et al., 2022 in prep.).

References

B. Barz, E. Rodner, Y. G. Garcia, and J. Denzler. Detecting regions of maximal divergence for spatio-temporal anomaly detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1088–1101, 2019. doi: 10.1109/TPAMI.2018.2823766.

V. Eyring, L. Bock, A. Lauer, M. Righi, M. Schlund, B. Andela, E. Arnone, O. Bellprat, B. Brötz, L.-P. Caron, N. Carvalhais, I. Cionni, N. Cortesi, B. Crezee, E. L. Davin, P. Davini, K. Debeire, L. de Mora, C. Deser, D. Docquier, P. Earnshaw, C. Ehbrecht, B. K. Gier, N. Gonzalez-Reviriego, P. Goodman, S. Hagemann, S. Hardiman, B. Hassler, A. Hunter, C. Kadow, S. Kindermann, S. Koirala, N. Koldunov, Q. Lejeune, V. Lembo, T. Lovato, V. Lucarini, F. Massonnet, B. Müller, A. Pandde, N. Pérez-Zanón, A. Phillips, V. Predoi, J. Russell, A. Sellar, F. Serva, T. Stacke, R. Swaminathan, V. Torralba, J. Vegas-Regidor, J. von Hardenberg, K. Weigel, and K. Zimmermann. Earth system model evaluation tool (ESMValTool) v2.0 – an extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of earth system models in CMIP. *Geoscientific Model Development*, 13(7):3383–3438, July 2020. doi: 10.5194/gmd-13-3383-2020. URL `https://doi.org/10.5194/gmd-13-3383-2020`.

E. Galytska, K. Weigel, J. Runge, V. Eyring, D. Handorf, R. Jaiser, and R. Köhler. Causal model evaluation of arctic-midlatitude teleconnections in cmip6, 2022 in prep.

A. Gerhardus and J. Runge. High-recall causal discovery for autocorrelated time series with latent confounders. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 12615–12625. Curran Associates, Inc., 2020. URL `https://proceedings.neurips.cc/paper/2020/file/94e70705efae423efda1088614128d0b-Paper.pdf`.

J. Gottfriedsen, M. Berrendorf, P. Gentine, M. Reichstein, K. Weigel, B. Hassler, and V. Eyring. On the generalization of agricultural drought classification from climate data, 2021. URL `https://arxiv.org/abs/2111.15452`.

C. Käding and J. Runge. Distinguishing cause and effect in bivariate structural causal models: A systematic investigation. *under review at Journal of Machine Learning Research*, 2022.

P. Nowack, J. Runge, V. Eyring, and J. D. Haigh. Causal networks for climate model evaluation and constrained projections. *Nature Communications*, 11(1), Mar. 2020. doi: 10.1038/s41467-020-15195-y. URL `https://doi.org/10.1038/s41467-020-15195-y`.

A. Paçal et al. Detecting extreme temperature events using gaussian mixture models, 2022 in prep.

J. Runge. Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. In J. Peters and D. Sontag, editors, *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, volume 124 of *Proceedings of Machine Learning Research*, pages 1388–1397. PMLR, 03–06 Aug 2020. URL `https://proceedings.mlr.press/v124/runge20a.html`.

J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic. Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, 5(11):eaau4996, 2019. doi: 10.1126/sciadv.aau4996. URL `https://www.science.org/doi/abs/10.1126/sciadv.aau4996`.

E. Saggioro, J. de Wiljes, M. Kretschmer, and J. Runge. Reconstructing regime-dependent causal relationships from observational time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 30(11):113115, Nov. 2020. doi: 10.1063/5.0020538. URL `https://doi.org/10.1063/5.0020538`.

X.-A. Tibau, C. Reimers, A. Gerhardus, J. Denzler, V. Eyring, and J. Runge. A spatio-temporal stochastic climate model for benchmarking causal discovery methods for teleconnections. *under review at the Journal of Environmental Data Science*, 2021.

K. Weigel, L. Bock, B. K. Gier, A. Lauer, M. Righi, M. Schlund, K. Adeniyi, B. Andela, E. Arnone, P. Berg, L.-P. Caron, I. Cionni, S. Corti, N. Drost, A. Hunter, L. Lledó, C. W. Mohr, A. Paçal, N. Pérez-Zanón, V. Predoi, M. Sandstad, J. Sillmann, A. Sterl, J. Vegas-Regidor, J. von Hardenberg, and V. Eyring. Earth system model evaluation tool (ESMValTool) v2.0 – diagnostics for extreme events, regional and impact evaluation, and analysis of earth system models in CMIP. *Geoscientific Model Development*, 14(6):3159–3184, June 2021. doi: 10.5194/gmd-14-3159-2021. URL `https://doi.org/10.5194/gmd-14-3159-2021`.