Identifying the atmospheric drivers of drought and heat using a smoothed deep learning approach

Magdalena Mittermeier
Department of Geography
LMU Munich
m.mittermeier@lmu.de

Maximilian Weigert
Statistical Consulting StaBLab
Department of Statistics
LMU Munich
maximilian.weigert@stat.uni-muenchen.de

David Rügamer
Department of Statistics
LMU Munich
david.ruegamer@stat.uni-muenchen.de

Abstract

Europe was hit by several, disastrous heat and drought events in recent summers. Besides thermodynamic influences, such hot and dry extremes are driven by certain atmospheric situations including anticyclonic conditions. Effects of climate change on atmospheric circulations are complex and many open research questions remain in this context, e.g., on future trends of anticyclonic conditions. Based on the combination of a catalog of labeled circulation patterns and spatial atmospheric variables, we propose a smoothed convolutional neural network classifier for six types of anticyclonic circulations that are associated with drought and heat. Our work can help to identify important drivers of hot and dry extremes in climate simulations, which allows to unveil the impact of climate change on these drivers. We address various challenges inherent to circulation pattern classification that are also present in other climate patterns, e.g., subjective labels and unambiguous transition periods.

1 Introduction

In recent summers such as those of 2003, 2010 and 2018, Europe has been subject to particularly outstanding summer drought and heat events, which caused large economic and societal damage including heat-related deaths [1, 2]. The frequency and intensity of hot and dry extremes has recently increased and is projected to further increase due to climate change and rising global mean temperatures [3, 4].

Drivers of hot and dry extremes There are two key processes leading to drought and heat events: thermodynamic and dynamic factors. Thermodynamic factors involve, e.g., evaporation and the feedback between soil moisture and air temperature. Dynamic factors on the other hand describe the atmospheric drivers of heat and drought, which are mainly anticyclonic conditions and blocking [3]. While anticyclonic conditions go along with various high-pressure systems, blocking describes a particular, persistent high-pressure situation that is associated with the displacement of westerly winds and their accompanying weather systems [3, 5]. These atmospheric drivers of hot and dry extremes are part of the large-scale atmospheric circulation in the mid-latitudes, which control the weather and climate over Europe [6, 7]. Changes in the atmospheric circulation are complex with opposing processes and thus many open research questions remain [5, 8].

Classification of circulation patterns Objectively classifying the circulation patterns that are associated with hot and dry extremes is an important step towards a better understanding of how climate change affects the atmospheric drivers of potentially disastrous extreme events. Previous studies, e.g., [9, 10, 11] have shown that deep learning approaches are powerful tools for the detection of extreme weather in climate simulations. In this study, we use a subjective catalog of circulation type classifications over Europe by Hess & Brezowsky [12, 13]. Our goal is to learn the categorization of six circulation types with anticyclonic conditions over Europe, which are associated with dry and hot summer conditions in Central Europe [14]. The classification of circulation types comes with various challenges that need to be addressed with appropriate modeling strategies. Challenges include noisy labels due to subjective expert choices in ambiguous climate situations [15], an imbalanced class distribution of labels, undefined transition days between successive circulation patterns, and a fixed dwell time of a circulation pattern of at least three consecutive days by its definition [13].

Our contribution In this work we propose a novel modeling procedure to address existing challenges in classifying anticyclonic circulation patterns. Especially in times of large ensembles of climate simulations [16] that consist of dozens of model runs and thousands of model years, our study can help to efficiently analyse large climate simulations and be another piece of the puzzle to better understand changes in the atmospheric drivers of drought and heat.

2 Data

The Hess & Brezowsky catalog contains a subjective categorization of circulation patterns created by experts manually labelling air pressure patterns over Europe into 29 classes. In this way, daily air pressure constellations are retrospectively assigned to one of these classes. A circulation pattern is, by definition, required to last at least three days. The six circulation patterns associated with heat and drought are (abbreviations originate from German): *Zonal ridge across Central Europe (BM), Norwegian Sea-Iceland high, anticyclonic (HNA), North-easterly anticyclonic (NEA), Fennoscandian high, anticyclonic (HFA), Norwegian Sea-Fennoscandian high, anticyclonic (HNFA), and South-easterly anticyclonic (SEA)* [17, 14]. The mean air pressure patterns for the six circulation patterns of interest are given in Figure 1 for the variables *sea level pressure* and *geopotential height at 500 hPa* (average values at roughly 5500 meters height). For the analysis of heat and drought, the remaining 23 circulation types are assigned to a residual class. The frequencies of the six anticyclonic patterns are between 8.5% (BM) and 1.4% (HNFA), whereas the residual class comprises about 80% of the days.
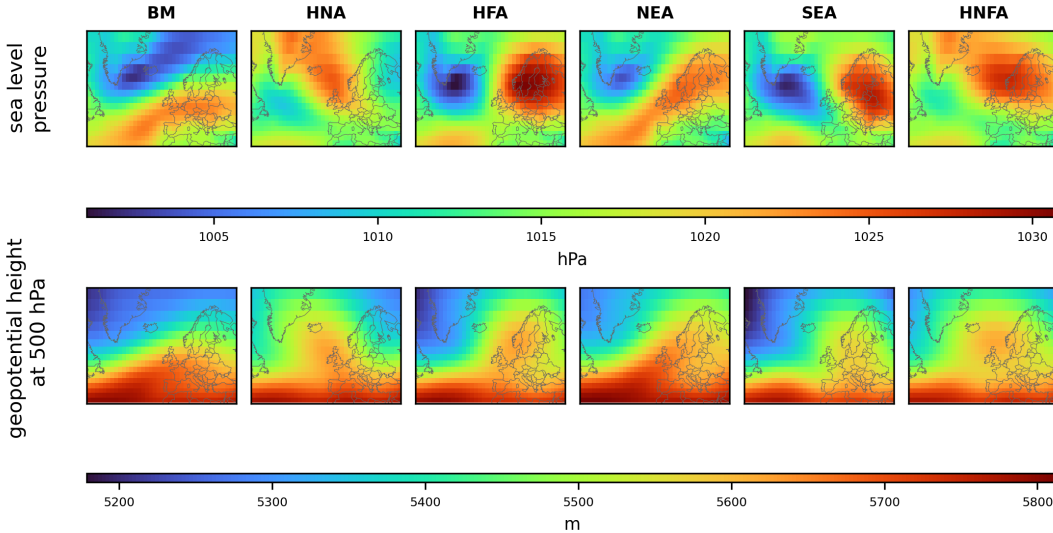

Figure 1: Mean air pressure patterns of the six anticyclonic circulation types BM, HNA, NEA, HFA, HNFA, and SEA (columns) averaged over all days in the period between 1900 and 2010. The plots are shown for the variables (rows) sea level pressure [hPa] and geopotential height at 500 hPa [m].

Next to the catalog by Hess & Brezowsky, we supplement our data base with the ERA-20C reanalysis data by the European Centre for Medium-Range Weather Forecasts [18]. The long record ERA-20C data set contains global spatial information on various climate parameters from 1900 to 2010. In accordance with the definition by [13], we use the two atmospheric variables, sea level pressure and geopotential height at 500 hPa as predictors for circulation patterns. The spatial domain of our data set is defined over a region covering Europe and the North Atlantic [13] (see Figure 1) with a spatial resolution of 5°x 5° resulting in a grid of 16 x 29 pixels. The resulting data set contains daily information over 111 years, i.e., T = 40541 observations.

3 Methods

Model definition, training, tuning and evaluation   To account for the spatial information and the specific characteristics of circulation patterns, we use a convolutional neural network (CNN) following [19], who propose a network architecture for climate pattern detection problems. Since air pressure patterns of atmospheric features are comparatively simple, our chosen architecture consists of only two convolutional layers with larger kernels (5x5-8 and 5x5-16), a dropout layer and two fully-connected layers as well as individual channels for both climate parameters. The two different atmospheric variables are included as individual channels in the CNN as in [2, 10]. While there is reason to believe that accounting for the temporal structure of our data, e.g., through a structured model [20] or ConvLSTM [21] improves the model, a previous study [22] showed no improvement in the classification of circulation patterns when using a temporal-aware architecture. As explained in the next paragraph, our approach instead smoothes predicted labels to account for their temporal nature.

The model is trained using Adam optimization with a batch size of 128 for 35 epochs and early stopping based on a validation set of size 3650 with patience of 6 epochs. Hyperparameter tuning for learning rate and dropout rate is performed using Bayesian optimization [23]. We evaluate the model using overall accuracy and macro F1-score. For class-specific evaluations, we consider recall and precision. To obtain performance estimates that are as unbiased as possible, a nested cross-validation with ten inner and eleven outer folds is used. In order to not leak intra-year information, observations within the same year are required to belong to the same fold.

Modeling challenges   Our approach takes into account several data-specific characteristics for circulation pattern data. First, we employ a loss-weighting scheme to account for imbalanced classes by weighting the classes with their inverse frequencies. Moreover, the assigned categories in the Hess & Brezowsky catalog can be noisy, in particular for transition days between two subsequent circulation patterns. This is due to the continuous movement of pressure systems while circulation types are discrete by definition and in-between states do not fit in one or the other class. We address this problem by using label-smoothing [24] for the first and last day of each occurrence of a specific circulation pattern. Finally, our target variable must adhere to the aforementioned definition of a circulation pattern, implying a pattern to last at least three days. A transition-smoothing step ensures that this three day rule is respected. In this step, the final predicted class $\tilde{y}_t$ at time point $t = 2, ..., T - 2$, is given by

$$
\tilde{y}_t = \begin{cases}
\hat{y}_{t-1} & \text{if } \hat{y}_{t-1} = \hat{y}_{t+1} \text{ (Neighborhood Consistency)}, \\
\hat{y}_{t-1} & \text{if } \hat{y}_t = \hat{y}_{t+1} \wedge \hat{y}_{t-1} = \hat{y}_{t+2} \text{ (2-days Consistency)}, \\
m(\hat{\boldsymbol{\pi}}_{t-1}, \hat{\boldsymbol{\pi}}_{t+1}) & \text{if } \hat{y}_t \neq \hat{y}_{t+1} \wedge \hat{y}_{t-1} \neq \hat{y}_{t+1} \text{ (Transition Membership)}, \\
m(\hat{\boldsymbol{\pi}}_{t-1}, \hat{\boldsymbol{\pi}}_{t+2}) & \text{else},
\end{cases}
$$

where $\hat{\boldsymbol{\pi}}_t$ denotes the predicted probability vector at time t, $\hat{y}_t = \arg\max \hat{\boldsymbol{\pi}}_t$ the predicted class prior to the transition-smoothing step, and

$$
m(\boldsymbol{\pi}_s, \boldsymbol{\pi}_t) = \arg\max\{\boldsymbol{\pi}_{u^*}\} \text{ with } u^* = \arg\max_{u \in \{s,t\}}\{\max(\boldsymbol{\pi}_u)\}.
$$

This guarantees consistency with the required three day rule and systematically replaces isolated single or two day-type predictions.

4 Results

Taking into account the aforementioned subjectivity of the circulation pattern catalog and the noisy labels, the overall performance of the proposed model is satisfactory and our proposed smoothing approaches consistently improve the model across all classes. Our smoothed convolutional neural network classifier achieves an macro F1-score of 38.4% and an overall accuracy of 59.9% averaged over the test sets during nested cross-validation. Table 1 shows the corresponding confusion matrix together with the precision and recall. The best performance in terms of recall is achieved for the circulation patterns HNA and BM, the lowest performance for NEA. In absolute numbers, misclassifications mainly occur for residual class observations. Due to the proposed error weighting technique, we obtain larger recall than precision values except for the residual class.

Table 1: Confusion matrix of our proposed smoothed approach, averaged over the test sets in the nested cross-validation. Correctly classified classes are highlighted in bold.

		BM	HNA	HFA	NEA	SEA	HNFA	Residual	$\sum$	Precision
					LABELS					
OUTPUTS	BM	**208.8**	4.0	7.7	6.5	1.7	5.0	477.4	711.2	0.29
	HNA	11.9	**75.8**	3.7	4.5	7.1	3.7	204.9	311.5	0.24
	HFA	22.9	3.5	**41.9**	14.6	3.6	1.5	138.6	226.6	0.18
	NEA	10.4	2.8	10.4	**61.5**	6.3	11.7	85.9	188.9	0.33
	SEA	3.2	11.6	3.6	15.0	**25.2**	5.5	78.4	142.5	0.18
	HNFA	9.1	3.8	2.4	20.5	5.5	**44.7**	185.7	271.7	0.16
	Residual	43.3	7.0	6.6	5.2	2.7	5.3	**1729.9**	1800	0.96
	$\sum$	309.6	108.6	76.2	127.7	52.1	77.4	2900.7	**3652.4**	–
	Recall	0.67	0.70	0.55	0.48	0.48	0.58	0.60	–	–

Table 2 shows an ablation study of our modeling procedure. Including label- and transition-smoothing improves the overall accuracy by 4 percentage points and the macro F1-score by 2 percentage points compared to a model without any smoothing steps. The class specific F1-scores also considerably increase for all patterns. A comparison of the networks without label-smoothing and without transition-smoothing indicates that the label-smoothing step has rather little impact while the proposed transition-smoothing is the key to our observed performance gains.

Table 2: Comparison of class-specific F1-scores (first 7 columns), accuracy and macro F1-Score (last two columns) for the final smoothed model (Final), a model without label-smoothing (NO LS), a model without transition-smoothing (No TS) and a model without label-smoothing (No LS and TS). Best results are highlighted in bold.

		BM	HNA	HFA	NEA	SEA	HNFA	Residual	**Accuracy**	**F1-score**
MODEL	Smoothed network	**0.41**	**0.36**	**0.28**	0.39	**0.26**	**0.26**	**0.74**	**0.60**	0.38
	No LS	**0.41**	**0.36**	**0.28**	**0.40**	**0.26**	**0.26**	**0.74**	**0.60**	**0.39**
	No TS	0.39	0.33	0.25	0.36	0.23	0.23	0.70	0.56	0.36
	No LS and TS	0.39	0.33	0.25	0.37	0.23	0.23	0.70	0.56	0.36

5 Conclusion and Outlook

Our results indicate the high potential of deep learning-based methods in classifying the atmospheric drivers of drought and heat. We also demonstrate the effectiveness of our smoothed approach to deal with typical challenges in circulation type classifications, e.g., transition-smoothing for historical

dwell time definitions. To the best of our knowledge, we are the first to use air pressure patterns over Europe to classify circulation patterns associated with drought and heat as given in the Hess & Brezowsky catalog. While the proposed approach can potentially also be used for other circulation patterns associated with different kinds of extreme climate events, our goal was to establish a baseline model for this specific and highly relevant circulation pattern categorization. Although our approach relies on a network architecture developed for climate applications [19], there is room for improvement in modeling the analyzed patterns. As an alternative to the transition-smoothing step, we will investigate a deep hidden Markov model that accounts for the state dwell times by assuming a latent process that emulates the data-inherent three day transition rule.

Acknowledgements

References

[1] A. Bastos, Z. Fu, P. Ciais, P. Friedlingstein, S. Sitch, J. Pongratz, U. Weber, M. Reichstein, P. Anthoni, A. Arneth, V. Haverd, A. Jain, E. Joetzjer, J. Knauer, S. Lienert, T. Loughran, P. C. McGuire, W. Obermeier, R. S. Padrón, H. Shi, H. Tian, N Viovy, and S. Zaehle. "Impacts of extreme summers on European ecosystems: a comparative analysis of 2003, 2010 and 2018". In: *Philosophical Transactions of the Royal Society B* 375.1810 (2020).

[2] B. Liu X.and He, L. Guo, L. Huang, and D. Chen. "Similarities and Differences in the Mechanisms Causing the European Summer Heatwaves in 2003, 2010, and 2018". In: *Earth´s Future* 8.4 (2020), p. 5.

[3] L. Suarez-Gutierrez, W. A. Müller, C. Li, and J. Marotzke. "Dynamical and thermodynamical drivers of variability in European summer heat extremes". In: *Climate Dynamics* 54.9-10 (2020), 4351–4366.

[4] J. Spinoni et al. "Future Global Meteorological Drought Hot Spots: A Study Based on CORDEX Data". In: *Journal of Climate* 33.9 (2020), pp. 3635–3661.

[5] V. Masson-Delmotte, P. Zhai, S.L. Connors A. Pirani, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change". In: *EGU General Assembly Conference Abstracts*. Cambridge University Press, 2021.

[6] M. F. Huguenin, E. M. Fischer, S. Kotlarski, S. C. Scherrer, C. Schwierz, and R. Knutti. "Lack of Change in the Projected Frequency and Persistence of Atmospheric Circulation Types Over Central Europe". In: *Geophysical Research Letters* 47.9 (2020).

[7] T. Woollings, A. Hannachi, and B. Hoskins. "Variability of the North Atlantic eddy-driven jet stream". In: *Journal of the Royal Meteorological Society* 136 (2010), pp. 856–868.

[8] M. Stendel, J. Francis, R. White, P. D. Williams, and T. Woollings. "The jet stream and climate change". In: *Climate Change* (2021), pp. 327–357.

[9] M. Mittermeier, M. Braun, M. Hofstätter, Y. Wang, and R. Ludwig. "Detecting Climate Change Effects on Vb Cyclones in a 50-Member Single-Model Ensemble Using Machine Learning". In: *Geophysical Research Letters* 46.24 (2019), pp. 14653–14661.

[10] E. Racah, C. Beckham, T. Maharaj, S. E. Kahou, and C. Prabhat Pal. "ExtremeWeather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events". In: (2016).

[11] T. Kurth, J. Zhang, N. Satish, I. Mitliagkas, E. Racah, M. A. Patwary, T. Malas, N. Sundaram, W. Bhimji, M. Smorkalov, J. Deslippe, M. Shiryaev, S. Sridharan, Prabhat, and P. Dubey. "Deep Learning at 15PF: Supervised and Semi-Supervised Classification for Scientific Data". In: (2017).

[12] P. Hess and H. Brezowsky. *Katalog der Großwetterlagen Europas (1881-2009)*. 1969.

[13] P. C. Werner and F. W. Gerstengarbe. *Katalog der Großwetterlagen Europas (1881-2009) nach Paul Hess und Helmut Brezowsky*. Nr. 119, 7. Aufl. PIK Report. Potsdam, 2010.

[14] H. Caspary. *Die Entwicklung von trockenen Großwetterlagen mit Auswirkungen auf den süddeutschen Raum*. Heft 18. Arbeitskreis KLIWA, 2012.

[15] R. Huth, C. Beck, A. Philipp, M. Demuzere, Z. Ustrnul, M. Cahynová, J. Kysel, and O. E. Tveitof. "Classifications of atmospheric circulation patterns: recent advances and applications". In: *Annals of the New York Academy of Sciences* 1146 (2008), pp. 105–152.

[16] N. Maher, S. Milinski, and R. Ludwig. "Large ensemble climate model simulations: introduction, overview, and future prospects for utilising multiple types of large ensemble". In: *Earth System Dynamics* 12.2 (2021), pp. 401–418.

[17] P. Sýkorová and R. Huth. "The applicability of the Hess–Brezowsky synoptic classification to the description of climate elements in Europe". In: *Theoretical and Applied Climatology* 142.3-4 (2020), 1295–1309.

[18] P. Poli, H. Hersbach, D. P. Dee, P. Berrisford, A. J. Simmons, F. Vitart, P. Laloyaux, D. G. H. Tan, C. Peubey, J.-N. Thépaut, Y. Trémolet, E. V. Hólm, M. Bonavita, L. Isaksen, and M. Fisher. "ERA-20C: An atmospheric reanalysis of the twentieth century". In: *Journal of Climate* 29.11 (2016), pp. 4083–4097.

[19] Y. Liu, E. Racah, Prabhat, J. Correa, A. Khosrowshahi, D. Lavers, K. Kunkel, M. Wehner, and W. Collins. "Application of deep convolutional neural networks for detecting extreme weather in climate datasets". In: *arXiv preprint arXiv:1605.01156* (2016).

[20] D. Rügamer, C. Kolb, and N. Klein. "Semi-Structured Deep Distributional Regression: Combining Structured Additive Models and Deep Learning". In: *ArXiv e-prints* (2021). eprint: 2002.05777.

[21] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-k. Wong, and W.-c. Woo. "Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting". In: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*. NIPS'15. Montreal, Canada, 2015, 802–810.

[22] H. Funk, C. Becker, A. Hofheinz, G. Xi, Y. Zhang, F. Pfisterer, M. Weigert, and M. Mittermeier. "Towards an automated classification of Hess & Brezowsky's atmospheric circulation patterns Tief and Trog Mitteleuropa using Deep Learning Methods". In: *Environmental Informatics: A bogeyman or saviour to achieve the UN Sustainable Development Goals?* Ed. by V. Wohlgemuth, S. Naumann, H.K. Arndt, and G. Behrens. Shaker, forthcoming.

[23] J. Snoek, H. Larochelle, and R. P. Adams. "Practical bayesian optimization of machine learning algorithms". In: *Advances in neural information processing systems* 25 (2012).

[24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. "Rethinking the inception architecture for computer vision". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2818–2826.