Towards dynamical adjustment of the full temperature distribution

Edoardo Vignotto*
edoardo.vignotto@unige.ch
Research Center for Statistics, University of Geneva
Geneva, Switzerland

Sebastian Sippel
Institute for Atmospheric and Climate Science, ETH Zurich
Zurich, Switzerland

Flavio Lehner
Institute for Atmospheric and Climate Science, ETH Zurich
Zurich, Switzerland

Erich M. Fischer
Institute for Atmospheric and Climate Science, ETH Zurich
Zurich, Switzerland

ABSTRACT

Internal variability due to atmospheric circulation can dominate the thermodynamical signal present in the climate system for small spatial or short temporal scales, thus fundamentally limiting the detectability of forced climate signals. Dynamical adjustment techniques aim to enhance the signal-to-noise ratio of trends in climate variables such as temperature by removing the influence of atmospheric circulation variability. Forced thermodynamical signals unrelated to circulation variability are then thought to remain in the residuals, allowing a more accurate quantification of changes even at the regional or decadal scale. The majority of these methods focus on climate variable's averages, thus discounting important distributional features. Here we propose a machine learning dynamical adjustment method for the full temperature distribution that recognizes the stochastic nature of the relationship between the dynamical and thermodynamical components. Furthermore, we illustrate how this method enables evaluating how specific events would have unfolded in a different, counterfactual climate from a few decades ago, thereby characterizing the emergent effect of climatic changes over decadal time scales. We apply our method to observational data over Europe and over the past 70 years.

KEYWORDS

dynamical adjustment, machine learning, quantile regression, heatwaves

ACM Reference Format:
Edoardo Vignotto, Sebastian Sippel, Flavio Lehner, and Erich M. Fischer. 2020. Towards dynamical adjustment of the full temperature distribution. In *10th International Conference on Climate Informatics (CI2020), September 22–25, 2020, virtual, United Kingdom.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3429309.3429317

*Corresponding author

CI2020, September 22–25, 2020, virtual, United Kingdom

ACM ISBN 978-1-4503-8848-1/20/09…$15.00
https://doi.org/10.1145/3429309.3429317

1 INTRODUCTION

Internal atmospheric variability can mask or amplify the near-term and medium-term effects of forced climate changes and thereby affect the public perception of climate change [11]. Moreover, internally generated climate variability is particularly large both at the local and short-term scale, which hinders the detectability of forced long-term trends [2]. Externally forced changes in the atmosphere can be conceptually decomposed into a thermodynamical component, related to changes in the system's energy balance, and a dynamical component, related to changes in atmospheric circulation. While many studies have emphasized the impact of climate change on the thermodynamical component, forced changes in atmospheric dynamics remain uncertain [15] and have been argued to be weak [8]. In fact, local and short-term internal variability is associated mostly with unforced atmospheric circulation-induced variability. Removing the impact of atmospheric circulation on a target variable of interest can consequentially enhance the thermodynamical signal as a residual and thus enhance the detectability of a climate change signal [4, 12].

It is possible to evaluate internal variability within a model ensemble by running multiple realizations of the same climate model starting from slightly different initial conditions but forced with the same external forcing. The average of several of these realizations represents an estimate of the forced response to the radiative forcings while the spread is a quantification of unforced internal variability [6]. Unfortunately, this is not a viable solution when dealing with observational data, since in this case only one of the many possible realizations of the climate system has been recorded.

Dynamical adjustment can help in tackling this issue by enhancing the signal-to-noise ratio of observational data. It aids isolating the forced response to external forcing by quantifying and separating the contribution of atmospheric circulation to a target variable from thermodynamical change that characterizes forced climate change [19]. Operationally speaking, the first step is to construct a statistical model for the relationship between the target variable and a proxy of atmospheric circulation that is ideally not influenced by thermodynamical changes. In this way, it is possible to estimate the response of the target variable given a specific atmospheric circulation pattern in a climate comparable to the one of the training set. Thus, from an analytical point of view, dynamical adjustment aims to determine the *conditional behaviour of a target*

variable T, e.g., temperature, conditioned on an atmospheric circulation pattern S. The time series of the residuals then can be thought to contain the forced thermodynamic response. This approach assumes non-significant forced changes in atmospheric circulation itself (or would require more assumptions such as detrending). This assumption is fundamental in this context since it is difficult to attribute changes in circulation to forcing or internal variability in observations alone. For mid-latitude locations and temperature, climate models lend strong support for this assumption over the historical record [3]. Different statistical models have been used for dynamical adjustment, from analogues based methods [7] to regression based ones [16, 17].

The vast majority of these studies focus on a specific summary statistic of the target variable of interest, usually a temporal or spatial average. This approach obviously simplifies the analytical framework and the interpretation and communication of the results, but it has also some drawbacks. In fact, modelling only a specific summary statistic, it misses capturing other important distributional features that can be highly relevant for understanding climate change. For example, increases in the variance of the temperature distribution, even with no changes in the average, are associated with increased chances of extreme heat waves. Methods that do not account for distributional changes are clearly not suited for such cases.

To overcome such limitations, we propose in this work a novel methodology based on quantile regression forests to dynamically adjust the full temperature distribution given a specific sea level pressure pattern (as a proxy for atmospheric circulation). This approach is more informative than the standard techniques mentioned because it takes into account the full probability distribution and hence allows probabilistic inferences. Furthermore, it permits to evaluate how a specific extreme event would have unfolded in another climate, e.g., a few decades ago or in a pre-industrial world not affected by climate change, as the method illustrates how probabilities have changed. The new methodology, therefore, can be seen as a generalized analogue-based algorithm. We illustrate how the method can be used to build a counterfactual estimate of how an extremely hot summer would have been, under the same atmospheric circulation, in earlier periods. In particular, we assess the contribution of circulation variability to temperature extremes in Europe and over the past 70 years using observational datasets. Furthermore, we show how our dynamical adjustment technique can be used to detect changes in temperature extremes. Lastly, we show how accounting for the influence of atmospheric circulation can give new insights on the relationships that link temperature extremes with other climatological drivers such as soil moisture. The proposed methodology will bridge the gap between standard dynamical adjustment approaches and more advanced quantile regression techniques and shed new light on circulation-induced variability and anthropogenic forcing influence on temperature changes over the last half of the 20th century.

The remainder of this work is organized as follows. In section 2 we describe the methodology and the datasets that we use. In section 3 the results of our application are presented and discussed. Finally, in section 4 we state some conclusions and implications.

2 MATERIALS AND METHODS

2.1 Marginal and conditional distributions

A fundamental distinction for the approach proposed in this work is between the marginal distribution of a random variable T, namely temperature, and the conditional distribution of T given a set of covariates $S = (S^1, \dots, S^p)$, e.g., sea level pressure fields in our case. Loosely speaking, while the former is related to the stochastic behaviour of T unconditional on any other variable, the latter is associated with the stochastic behaviour of T given a specific configuration of S.

Random variables can be described with different probabilistic tools. One of the most common descriptions in this regard is the cumulative distribution function. For the following, it is important to distinguish between the cumulative distribution function associated with the marginal distribution of T

$$\begin{aligned} F_T(t) =& P(T \leq t) \\ =& E(1_{\{T \leq t\}}) \end{aligned}$$

and the one associated with the conditional distribution of T given S

$$\begin{aligned} F_{T|S}(t \mid s) =& P(T \leq t \mid S = s) \\ =& E(1_{\{T \leq t\}} \mid S = s). \end{aligned}$$

The inverse operation is given by the quantile function and it is related to the extremeness of an event given its stochastic distribution. That is, it outputs the value of a random variable (e.g. temperature anomaly) given its probability and its statistical distribution. Also in this case we made a distinction between the marginal version

$$Q_T(\tau) = \inf\{t \in \mathbf{R} : \ \tau \leq F_T(t)\} = F_T^{-1}(\tau)$$

and the conditional version

$$\begin{aligned} Q_{T|S}(\tau \mid s) =& \inf\{t \in \mathbf{R} : \ \tau \leq F_{T|S}^{-1}(t \mid s)\} \\ =& F_{T|S}^{-1}(\tau \mid s). \end{aligned}$$

2.2 Quantile regression forests

The goal of quantile regression forests [9] is to estimate the conditional cumulative distribution function $F_{T|S}(t \mid s)$ of a target univariate random variable T, i.e., daily temperature at a specific location in this work, given a set of covariates $S = (S^1, \dots, S^p)$, i.e., a proxy of the atmospheric circulation pattern (that is, mean sea level pressure) around that location in our case, starting from n training observations (t_i, s_i), $i = 1, \dots, n$.

A large number B of decision trees is then grown. For each tree, a bootstrap sample of the original training data is used and at each node the splitting procedure considers only m randomly selected covariates. The m value is the only hyperparameter of the algorithm and is chosen through cross-validation or out-of-bag optimization. However, results from the two approaches are usually consistent with respect to m and thus, for simplicity, we use the standard value of $p/3$, where p is the number of covariates.

Given a new set of covariates s, for each tree G_b, $b = 1, \dots, B$, in the forest a weight $w_i^b(s)$ is assigned to each training point s_i, $b = 1, \dots, B$, accounting for the similarity between s and s_i.

Specifically, denote with $l(s)$ the leaf to which s belongs when it is dropped down the tree. Then, the weight

$$w_i^b(s) = \frac{1_{\{s_i \in l(s)\}}}{\#\{j : s_j \in l(s)\}}$$

is non zero only for s_i that are in the same leaf $l(s)$ and assign the same importance to all the s_i in $l(s)$. A total weight for each training observation s_i can consequently be obtained averaging over all the trees, i.e.,

$$w_i(s) = \frac{1}{B}\sum_{b=1}^{B} w_i^b(s),$$

and an estimate for $F_{T|S}(t \mid s)$ is finally computed as

$$\hat{F}_{T|S}(t \mid s) = \sum_{i=1}^{n} w_i(s) 1_{\{T_i \le t\}}.$$

In a similar way, an estimate $\hat{Q}_{T|S}(\tau \mid s)$ for the conditional quantile function $Q_{T|S}(\tau \mid s)$ can be easily obtained.

Estimates of the conditional distribution of T given S are thus obtained as a weighted cumulative distribution of the original $t_1, \ldots, t_n$. In this sense, quantile regression forests, when applied to atmospheric circulation patterns, can be viewed as a generalized analogues approach, since it gives more weight to days characterized by a similar sea level pressure field to the one of interest.

Quantile regression forests are shown to be a consistent statistical procedure and can even handle time-dependent data, including as covariates lagged information. For our purpose, this machine learning technique is particularly appropriate here for its ability to describe complex relationships including potentially non-linear dependencies. Furthermore, quantile regression forests are quite robust in high dimensional settings and do not suffer from the presence of multicollinearity in the predictors [9].

In this work, for a specific grid point, we focus on modeling the conditional distribution of temperature T given an atmospheric circulation pattern S at the daily scale. We use sea level pressure as a proxy for the latter, considering all its values in a squared spatial field of 60 degrees width centered on the grid point under analysis. In particular, the overall goal is to illustrate how the anomaly warm 2003 and 2018 summers, defined as the August month, would have been during a different climatic period (r_0), i.e., between 1950–1969, representing a past climate. To do that, we need to build two different models, one describing the relationship between temperature and atmospheric circulation in the reference period r_0 that we denote with $\hat{F}^{r_0}_{T|S}(t \mid s)$ and one for the 2003 and 2018 summers that we denote with $\hat{F}^{r_1}_{T|S}(t \mid s)$. While the former is trained with all the years composing r_0, for the latter we build the training dataset using all the years between 1997 and 2018, excluding 2003 and 2018.

As a final consideration, we note that the approach proposed here is still valid from a theoretical point of view even if the chosen set of covariates S would not be informative at all for the target variable T. In this case, the marginal distribution $F_T(t)$ and the conditional distribution $F_{T|S}(t \mid s)$ would coincide and the results described above are reduced to a marginal comparison of the different stochastic behaviour of T in the different reference periods.

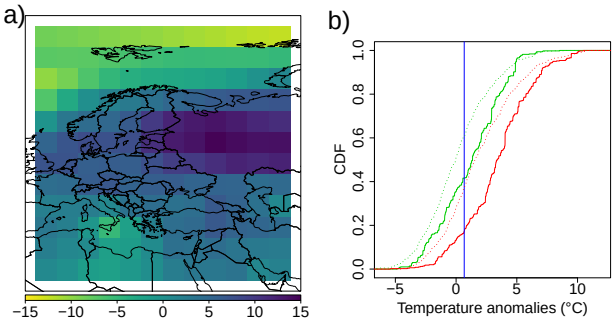

Figure 1: a) Sea level pressure anomaly field (Pa) of 2^{th} August 2003 relative to the 1950–1969 reference period for illustration. b) Temperature marginal cumulative distribution over the 1950–1969 period (dotted green line), temperature marginal cumulative distribution over the 1993–2002;2004–2013 period (dotted red line), temperature conditional cumulative distribution relative to the 1950–1969 reference period (green line), conditional cumulative temperature distribution relative to the 1993–2002;2004–2013 reference period (red line) and observed temperature anomaly (blue line). All the results are for the Minsk grid cell on the 2^{th} August 2003.

2.3 Data and experimental setup

For this work, we consider average daily temperature over Europe from the E-OBS [1] dataset and average daily land soil moisture and sea level pressure from the NCEP/NCAR dataset [5]. For each grid point, the data are anomalies relative to their climatological average in the 1950–1969 period. Temperature and soil moisture data are regridded in order to have a 2.5×2.5 degree horizontal resolution. Sea level pressure is regridded to 5×5 degree spatial resolution.

We fit the two quantile regression forest models $\hat{F}^{r_0}_{T|S}(t \mid s)$ and $\hat{F}^{r_1}_{T|S}(t \mid s)$ described above separately for each grid point to predict the daily temperature distribution based on the corresponding sea level pressure fields and subsequently aggregate results to monthly averages for 2003 and 2018 (illustrated in Section 3). Even if the results are computed only for August, we also include July and September data in the training set to increase sample size.

3 RESULTS AND DISCUSSION

3.1 An illustrative example: Changes in the marginal and conditional temperature distribution

As an illustrative example, we consider in Figure 1a the sea level pressure anomaly field of the $2th$ August 2003 for the grid point that contains Minsk. In Figure 1b we report the conditional distributions predicted from the two models $\hat{F}^{r_0}_{T|S}(t \mid s)$ and $\hat{F}^{r_1}_{T|S}(t \mid s)$ plus the two marginal temperature distributions relative to the r_0 and r_1 periods. In this case, the conditional distributions are shifted to the right (higher temperatures) with respect to the associated marginal ones, indicating a larger likelihood of warmer temperatures after accounting for sea level pressure conditions and thus a positive contribution of anticyclonic circulation anomaly (Fig. 1a)

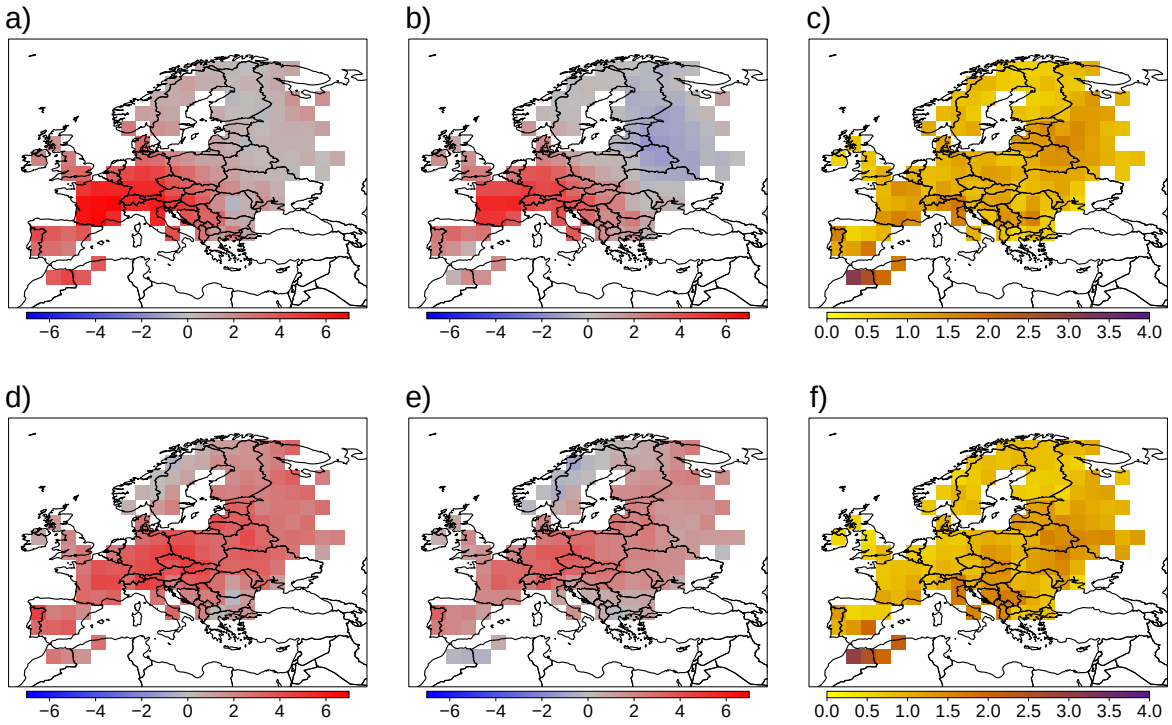

Figure 2: a) Average observed temperature anomalies (°C) for the 2003 summer, relative to the reference period 1950–1969. b) Average counterfactual temperature anomalies (°C) for the 2003 summer relative to the 1950–1969 reference period. c) Average absolute differences between the observed temperatures and their counterfactuals for the 2003 summer. d,e,f) Same results for the 2018 summer.

Table 1: Counterfactual differences (°C) for the 2003 and 2018 summers for Madrid, Hamburg and Minsk. Confidence intervals at $\alpha = 5\%$ are computed using non-parametric bootstrap.

	Madrid	Hamburg	Minsk
2003	[1.21;1.60]	[0.63;1.59]	[0.65;1.59]
2018	[0.65;1.11]	[0.15;1.10]	[0.16;1.84]

to temperature extremes. The actual event (blue line) would have been in the bulk of the conditional distribution in the early period (around 40th percentile, where green and blue lines intersect). In contrast, given the sea level pressure patterns it was a relatively cold event (around 15th percentile) in the recent period, indicating warming between the early and late period. In addition, for both reference period, the event level is relatively colder accounting for atmospheric circulation than looking at the marginal temperature distributions.

3.2 Understanding the 2003 and 2018 heatwaves conditional on atmospheric circulation

Figure 2 shows the average difference between the temperatures during August 2003 and 2018 and their counterfactuals with respect to the reference period r_0, i.e, how anomalous they would have been during 1950–1969. We define the counterfactual for a day d with observed temperature $t(d)$ and sea level pressure pattern $s(d)$ with respect to the reference period r_0 as

$$d^* = \hat{Q}^{r_0}_{T|S}\{\hat{F}^{r_1}_{T|S}[t(d) \mid s(d)] \mid s(d)\}.$$

Specifically, we thus consider as counterfactual for d the temperature in the reference period r_0 that corresponds to the same quantile level of d with respect to the recent reference period r_1. Note that both the August 2003 and 2018 would have been, on average, about by one and two degrees colder in a climate corresponding to the 1950–1969 period. However, a summer with a similar circulation pattern as the one observed in 2003 or 2018 would have still been anomalously warm one even in 1950–1969 (2b and e).

In Table 1 we focus on three specific locations and we compute their yearly average counterfactual differences as defined above. Confidence intervals at $\alpha = 5\%$ are computed through non-parametric bootstrap with $B = 100$ simulations. For all the three locations, both the August 2003 and 2018 are significantly warmer than they would have been given the same sea level pressure anomaly fields during the r_0 reference period 1950–1969. Analogously, one could compute confidence intervals for all the quantities presented here and thus perform hypothesis tests for specific research questions.

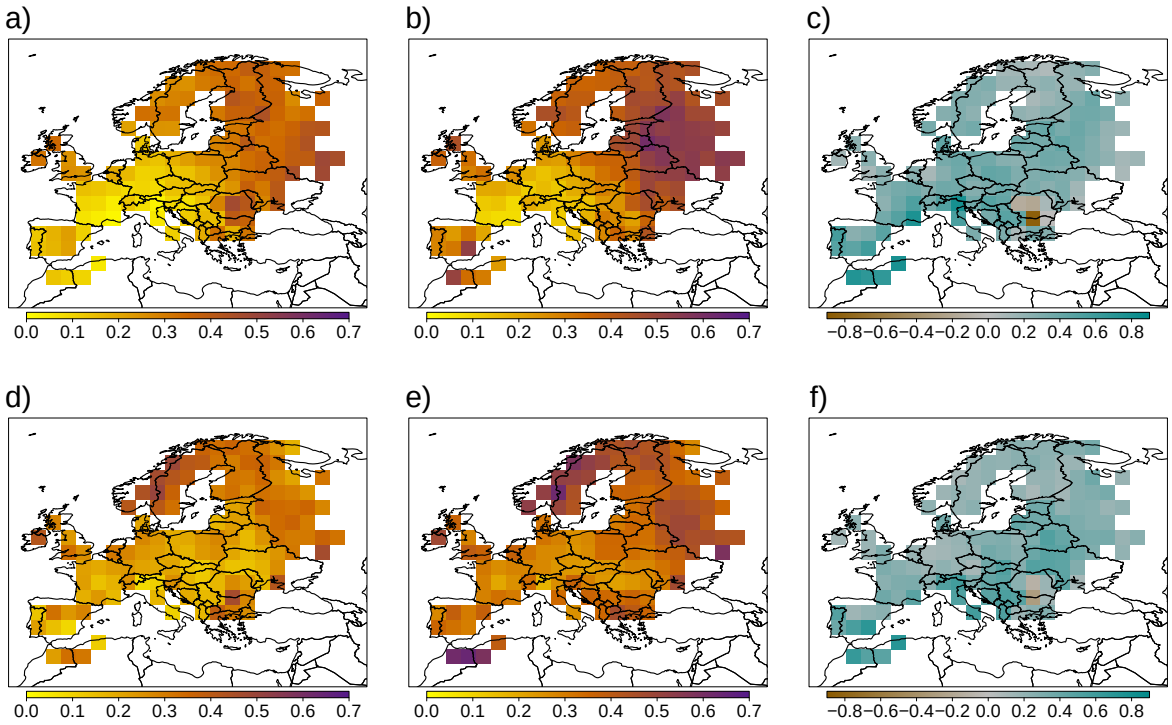

Figure 3: a,b,c) Respectively, average p_0, p_1 and fraction of attributable risk for the 2003 summer. d,e,f) Same results for the 2018 summer.

Figure 3 reports the average fraction of attributable risk [18] during the considered summers. We define the fraction of attributable risk for a specific day d as

$$\mathrm{FAR} = 1 - \frac{p_0}{p_1} = 1 - \frac{P^{r_0}_{T|S}(T > t(d) \mid s(d))}{P^{r_1}_{T|S}(T > t(d) \mid s(d))}$$
$$= 1 - \frac{1 - \hat{F}^{r_0}_{T|S}(t(d) \mid s(d))}{1 - \hat{F}^{r_1}_{T|S}(t(d) \mid s(d))}.$$

The fraction of attributable risk is a measure widely used in event attribution studies and represents the relative ratio between the likelihood p_0 of an event in a counterfactual world with weaker climate change and the likelihood p_1 of the same event in a factual world, i.e., one characterized by present-day forced climate change. Positive values, as it is the case for almost all grid points for both the 2003 and 2018 summers, indicate that such events are relatively more likely in the latter reference period than the earlier. Note than in contrast to many previous studies, FAR is here not expressed relative to pre-industrial conditions but relative to the earlier period 1950–69. In particular, regions in Eastern Europe and the Mediterranean, but also throughout Central Europe and Scandinavia, show relatively large increases in the probability, indicating warming of the conditional distribution, where FAR values hence reach relatively high values.

Atmospheric circulation is one of the main drivers of temperature anomalies. We hypothesize that other drivers or feedbacks, such as land-atmosphere interactions, particularly soil moisture feedbacks could explain the residual variability after accounting for circulation. In Figure 4 we compare the marginal Spearman correlation between soil moisture and temperature at the daily scale with the residual Spearman correlation after accounting for the influence of circulation for the 2003 and 2018 summers. The residual for a specific day d is defined as

$$e_d = \hat{F}^{r_1}_{T|S}(t(d) \mid s(d)).$$

The results are aggregated over the two considered Augusts. Grid points located in Eastern Europe and in the Mediterranean basin show negative correlation between land soil moisture and temperature for both the residual and the marginal case. Interestingly, the marginal correlation is, on average, between two and three times stronger than the residual one. The fact that the marginal correlation is systematically stronger than the conditional one points towards some joint effect of circulation and soil moisture on temperature, or some joint effect in which circulation influences temperature *and* soil moisture. In other words, if a circulation anomaly systematically induces a soil moisture anomaly, e.g. through a lack of precipitation and anomalously high evapotranspiration, dynamical adjustment can remove this effect. That is, an interpretation of the marginal correlation between temperature and soil moisture as solely a relationship between soil moisture and temperature would require some caution, as the true relationship in that case would be weaker than indicated by the marginal correlation.

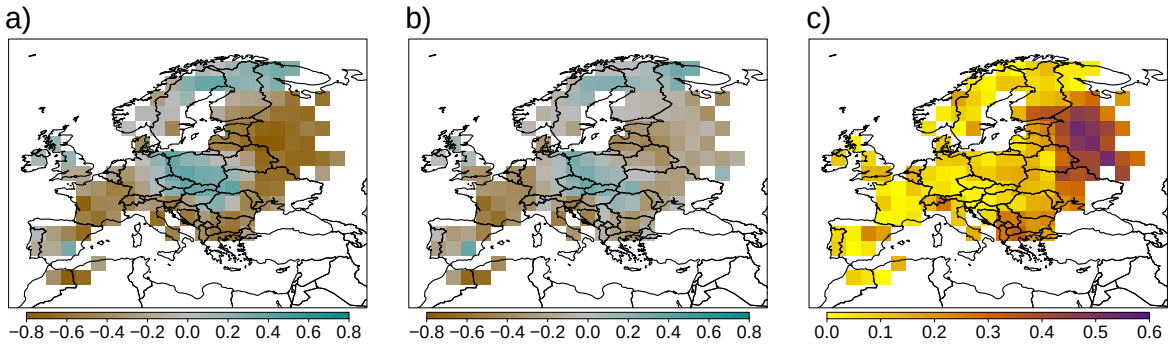

Figure 4: a) Marginal correlation between soil moisture and temperature for the 2003 and 2018 summers. b) Residual correlation between soil moisture and temperature for the 2003 and 2018 summers after accounting for the influence of atmospheric circulation. c) Absolute differences between the marginal and residual correlation for the 2003 and 2018 summers.

The goal of this work was to illustrate the potential of machine learning-based quantile regression methods for dynamically adjusting the full temperature distribution. While this yields results that will be pursued and interpreted further in future studies, and a promising methodological avenue overall, some caveats need to be mentioned. First, we have not conducted a comprehensive evaluation of the performance of quantile regression forests, neither to other, alternative methods nor to sample size. While the method in its outlined setting yields reasonable results, a comprehensive evaluation (e.g., how informative large-scale sea level pressure variations are in explaining summer temperature variability) is left to future studies. In this regard, the use of large control runs, i.e., climate model simulations in a stationary climate condition, could represent an appropriate play-field for a proper evaluation. A preliminary example can be found in the Supplementary Material. Second, the application of dynamical adjustment with sea level pressure as a proxy for atmospheric circulation has some known limitations in summer. This includes that the influence of large-scale atmospheric circulation on temperature variability is lower in summer than in winter, but also that heat lows developing on continents during extreme heat may feed back on sea level pressure and thereby affect predictions adversely. In addition, various, complex processes contribute to extreme summer heat waves, including, for instance, advection of extremely hot and dry air [13] or land surface and boundary layer feedbacks [10] that all are likely not fully captured by using an instantaneous circulation proxy. Finally, we illustrate changes in marginal and conditional temperature distributions between the mid-20th century and the early 21st century. While these changes contain the forced component of climate change, other variability such as multi-decadal internal variability that is unrelated or not captured by atmospheric circulation, may have also affected these estimates of temperature change to some degree. A further possible limitation of our proposal is given by the fact that a different model for each grid points is learned separately and thus spatial information is not fully exploited and taken into consideration. More complex statistical models, such as based on convolutional neural networks, could be developed to overcome this issue, potentially reducing the number of training years needed without lowering the accuracy of the results. Another way to reduce the time window size of the training dataset could be to build the model starting from climate model simulations instead of observational data.

4 CONCLUSIONS

In this work we propose a method useful:

- in evaluating the extent of thermodynamical change over the past 70 years;
- in a detection and attribution sense, in computing the fraction of attributable risk associated with recent anomalously warm summers while conditioned on atmospheric circulation;
- to better understand the relationships between temperature extremes and other climatological drivers, such as soil moisture;
- in explaining the relationship of large-scale atmospheric circulation to temperatures;
- for building counterfactuals of present events and evaluating how they would have been some decades ago, hence literally moving the observed record in the corresponding quantile of the conditional distribution characterizing the reference period.

We apply the methodology to observational data over Europe and show that recent warm summers were, on average, about between one and two degrees hotter than what they would have been, under the same atmospheric circulation conditions, between 1950 and 1969. Nevertheless, these two summers would have been anomaly warm also during this later reference period, indicating a positive atmospheric circulation contribution to temperature levels in these specific summers.

Different extensions of our methodology are possible. One interesting future direction could be to study if the thermodynamic forced responses under different atmospheric circulation types are of similar extent or, on the contrary, whether there are specific circulation patterns that are associated with the biggest changes, while for others the impact of climate change on temperature has been a moderate one. In the same direction, a full distribution method

as proposed here permits to investigate temperature changes that are usually not considered from classical dynamical adjustment methods, such as evaluating whether the temperature variance associated with a particular circulation pattern is wider now than in the past.

Another possible application regards the study of the contribution of atmospheric circulation to anomalously warm summers, i.e., if there were periods in the past that were particularly hot mainly because of dynamical effects. Even under a stationary climate, in fact, anomalous summers could happen due to rare circulation conditions that lead to extreme temperatures. The proposed method applied to detrended data could permit to illustrate particular atmospheric configuration under which it is more likely to observe warm summers and to evaluate in a statistically sound way the probability of heatwaves under such circumstances. Furthermore, while atmospheric circulation is certainly an important driver of temperature extremes, other factors, such that soil moisture feedbacks [10, 14] or advection of hot and dry air [13], can have an important role in characterizing heat waves intensity and duration. These factors can thus explain the residual temperature variability after accounting for atmospheric circulation. An analysis of the relationship between these factors and temperature extremes after accounting for dynamically induced effects could lead to a better understanding of their influence. Future work may aim to address limitations of the method, such as in terms of sample size needed to train the model or statistical parsimony (avoiding to train a model for each grid cell separately), or aiming towards a better understanding of limitations of dynamical adjustment in summer such as due to the complexity of processes contributing to extreme heat waves, or the development of thermal heat lows that may adversely affect dynamical adjustment results.

ACKNOWLEDGMENTS

EV acknowledges funding from the Swiss National Science Foundation (Doc.Mobility Grant 188229). FL has been supported by the Swiss National Science Foundation (Ambizione grant no. PZ00P2_174128). SS is grateful for funding provided by the Swiss Data Science Centre through the project "Data Science-informed attribution of changes in the Hydrological cycle" (DASH, ID C17-01).

REFERENCES

[1] Richard C Cornes, Gerard van der Schrier, Else JM van den Besselaar, and Philip D Jones. 2018. An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres* 123, 17 (2018), 9391–9409.

[2] Clara Deser, Adam Phillips, Vincent Bourdette, and Haiyan Teng. 2012. Uncertainty in climate change projections: the role of internal variability. *Climate dynamics* 38, 3-4 (2012), 527–546.

[3] Clara Deser, Laurent Terray, and Adam S Phillips. 2016. Forced and internal components of winter air temperature trends over North America during the past 50 years: Mechanisms and implications. *Journal of Climate* 29, 6 (2016), 2237–2258.

[4] Maurice F Huguenin, Erich M Fischer, Sven Kotlarski, Simon C Scherrer, Cornelia Schwierz, and Reto Knutti. 2020. Lack of change in the projected frequency and persistence of atmospheric circulation types over Central Europe. *Geophysical Research Letters* 47, 9 (2020), e2019GL086132.

[5] Eugenia Kalnay, Masao Kanamitsu, Robert Kistler, William Collins, Dennis Deaven, Lev Gandin, Mark Iredell, Suranjana Saha, Glenn White, John Woollen, et al. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society* 77, 3 (1996), 437–472.

[6] Jennifer E Kay, Clara Deser, A Phillips, A Mai, Cecile Hannay, Gary Strand, Julie Michelle Arblaster, SC Bates, Gokhan Danabasoglu, J Edwards, et al. 2015. The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society* 96, 8 (2015), 1333–1349.

[7] Flavio Lehner, Clara Deser, and Laurent Terray. 2017. Toward a new estimate of "time of emergence" of anthropogenic warming: Insights from dynamical adjustment and a large initial-condition model ensemble. *Journal of Climate* 30, 19 (2017), 7739–7756.

[8] Camille Li, Clio Michel, Lise S Graff, Ingo Bethke, Giuseppe Zappa, Thomas J Bracegirdle, Erich Fischer, Ben J Harvey, Trond Iversen, Martin P King, et al. 2018. Midlatitude atmospheric circulation responses under 1.5 and 2.0 degrees C warming and implications for regional impacts. *Earth System Dynamics* 9, 2 (2018), 359–382.

[9] Nicolai Meinshausen. 2006. Quantile regression forests. *Journal of Machine Learning Research* 7, Jun (2006), 983–999.

[10] Diego G Miralles, Adriaan J Teuling, Chiel C Van Heerwaarden, and Jordi Vila-Guerau De Arellano. 2014. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature geoscience* 7, 5 (2014), 345–349.

[11] Frances C Moore, Nick Obradovich, Flavio Lehner, and Patrick Baylis. 2019. Rapidly declining remarkability of temperature anomalies may obscure public perception of climate change. *Proceedings of the National Academy of Sciences* 116, 11 (2019), 4905–4910.

[12] Claudio Saffioti, Erich M Fischer, Simon C Scherrer, and Reto Knutti. 2016. Reconciling observed and modeled temperature and precipitation trends over Europe by adjusting for circulation variability. *Geophysical Research Letters* 43, 15 (2016), 8189–8198.

[13] Dominik L Schumacher, Jessica Keune, Chiel C Van Heerwaarden, Jordi Vilà-Guerau de Arellano, Adriaan J Teuling, and Diego G Miralles. 2019. Amplification of mega-heatwaves through heat torrents fuelled by upwind drought. *Nature Geoscience* 12, 9 (2019), 712–717.

[14] Sonia I Seneviratne, Daniel Lüthi, Michael Litschi, and Christoph Schär. 2006. Land–atmosphere coupling and climate change in Europe. *Nature* 443, 7108 (2006), 205–209.

[15] Theodore G Shepherd. 2014. Atmospheric circulation as a source of uncertainty in climate change projections. *Nature Geoscience* 7, 10 (2014), 703–708.

[16] Sebastian Sippel, Nicolai Meinshausen, Anna Merrifield, Flavio Lehner, Angeline G Pendergrass, Erich Fischer, and Reto Knutti. 2019. Uncovering the forced climate response from a single ensemble member using statistical learning. *Journal of Climate* 32, 17 (2019), 5677–5699.

[17] Brian V Smoliak, John M Wallace, Pu Lin, and Qiang Fu. 2015. Dynamical adjustment of the Northern Hemisphere surface air temperature field: Methodology and application to observations. *Journal of Climate* 28, 4 (2015), 1613–1629.

[18] Peter A Stott, Nikolaos Christidis, Friederike EL Otto, Ying Sun, Jean-Paul Vanderlinden, Geert Jan van Oldenborgh, Robert Vautard, Hans von Storch, Peter Walton, Pascal Yiou, et al. 2016. Attribution of extreme weather and climate-related events. *Wiley Interdisciplinary Reviews: Climate Change* 7, 1 (2016), 23–41.

[19] John M Wallace, Qiang Fu, Brian V Smoliak, Pu Lin, and Celeste M Johanson. 2012. Simulated versus observed patterns of warming over the extratropical Northern Hemisphere continents during the cold season. *Proceedings of the National Academy of Sciences* 109, 36 (2012), 14337–14342.

A SUPPLEMENTARY MATERIAL

In a context without any forcing factor, the residuals, as defined in section 3, should have standard uniform distribution. We thus consider 220 years of simulations from a CESM control run, fit a quantile regression forest for the grid point containing Paris on 20 years of randomly selected data, compute the residuals on the remaining 200 years of data and compared their distribution with the one of a standard uniform. The data have been preprocessed in the same way described in section 2C.

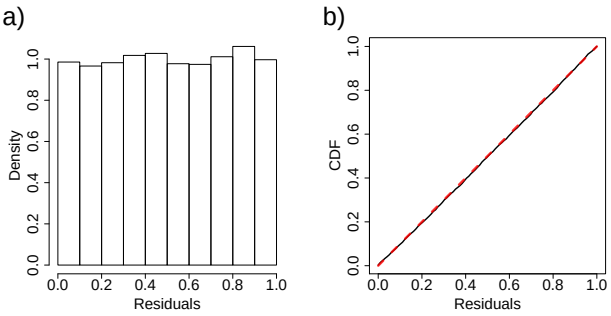

Figure 5: a) Histogram of the residuals for the grid point containing Paris for the CESM data. b) Empirical cumulative distribution function of the residuals for the grid point containing Paris for the CESM data (black line) compared to the cumulative distribution function of a standard uniform distribution (red dotted line).