



Swiss Federal Institute of Technology Zurich

Seminar for
Statistics

Department of Mathematics

Master Thesis

September 2024

Marco Froelich (22-942-601)

**Investigating year-to-year variability of hot extremes
and contributions from heat-generating mechanisms**

Submission Date: September 16th, 2024

Co-Adviser: Dr. Matthias Rothlisberger
Adviser: Prof. Dr. Nicolai Meinshausen

Preface

I would like to extend my gratitude to all the people that supported me throughout the this project. To Matthias, that from the start of my interest in this field, inspired and challenged me to create something new. You taught me many lessons that will stay with me and I value enormously the interesting conversations that stimulated a lot of excitement and creativity. To Maybritt, thank you for your support, insights, supervision and the many conversations that inspired solutions. Your encouragements from the start to the end were incredibily valuable. To Belinda, your help was essential in fortifying my physics competences and I appreciated the time you spend supporting my work. To Heini, thank you for the supervision and feedback that was instrumental in shaping a meaningful and impactful project. You all showed interest in my work and I am very grateful to have had such a competent, enthusiastic and friendly team.

This thesis has been a unique opportunity to place myself at the head of a scientific project, and it was foremost a humbling experience, as the challenges of interdisciplinary research have taught me. Throughout all the ups and down, I would like to thank my friends for the moments of laughter, for being great rubber ducks, and for all the good times. I will miss you all a lot and am eager to follow all your future adventures. Grazie mamma per essere stata sempre raggiungibile, per le poesie ispirante e le foto del grigio che mi fanno sorridere molto. Merci papa pour ces parties de ping-pong qui me sont si importantes et les bons verres de vins. To Matt and Luca, always wonderful to have you around: I cherish all the moments we have together. To my Manu, thank you for everything. Your calls and visits brought me joy and comfort, our time together is so dear to me and you are always an inspiration. I am so glad to have you by my side.

Abstract

Heat-related extremes are important meteorological phenomena that can have strong consequences on human health and the environment. Climate change is expected to exacerbate these impacts through an increase in the frequency of hot extreme occurrence and intensity. Although there exists abundant literature on the typical physical functioning of these events and their association to the variability of the climate system on different temporal scales, there lacks a global assessment of the influence of major physical processes - heat advection, adiabatic compression and diabatic heating - on the yearly variation of hot extreme magnitudes. To remediate this knowledge gap, we first propose a data-driven, systematic analysis of second-moment characteristics of yearly maxima near-surface hot extreme events and contributing heat-generating processes. Second, we apply deep-learning methods to model hot extreme Lagrangian trajectories to gain insights into important dynamical features. No physical process is globally found to dominate variability in these events and significant variance contributions exist from at least two processes, suggesting that mean-state understanding of hot extreme development may not be sufficient to explain large year-to-year differences in their magnitudes. Furthermore, this analysis reaffirms the presence of strong dependencies between the three physical mechanisms leading to a characterization of their variability by only one or two degrees of freedom in most of the world. Finally, the approach for the analysis of parcel trajectories was limited due to generally poor predictive performance, but showed that the patterns in advective, adiabatic and diabatic temperature anomaly generation follow patterns that may be predicted from their history, encouraging for future work. In addition, over oceans and many land regions we observe that adiabatic heating is minimal during the final 24h, suggesting that hot extreme primarily descend to the surface earlier than a day before, thus leading to contributions from advective and diabatic processes more likely.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | Motivation | 1 |
| 1.2 | Outline | 2 |
| 2 | Meteorological Background | 3 |
| 2.1 | Mechanisms of hot extreme development | 3 |
| 2.2 | Global quantification of physical mechanisms | 5 |
| 2.3 | Variability of hot extreme events | 5 |
| 2.4 | Methods in hot extreme literature | 6 |
| 2.5 | Temperature anomaly decomposition | 7 |
| 3 | Data | 9 |
| 3.1 | ERA5 and trajectory calculations | 9 |
| 3.2 | Cumulative total | 9 |
| 3.3 | Trajectory timeseries | 10 |
| 4 | Statistical methods | 13 |
| 4.1 | Variance decomposition of TX1day T' | 13 |
| 4.2 | Principle Component Analysis of contributors | 14 |
| 4.3 | Forecasting TX1day contributor trajectories | 16 |
| 5 | Results | 21 |
| 5.1 | Relating TX1day T' variability to component variability | 21 |
| 5.2 | Comparing contributions to TX1day T' mean and variance | 23 |
| 5.3 | Characterizing contributor dependencies | 25 |
| 5.4 | Forecasting the last 24h of TX1day trajectories | 27 |
| 6 | Conclusion | 31 |
| 6.1 | Summary | 31 |
| 6.2 | Limitations and future work | 32 |
| | Bibliography | 41 |
| A | Supplementary material on timeseries data | 43 |
| A.1 | Further examples of trajectory data | 43 |
| A.2 | Differenced timeseries data | 47 |
| B | Further background on variance | 49 |
| C | Grid-search results | 51 |

Chapter 1

Introduction

1.1 Motivation

In recent years, heat-related near-surface extreme events have globally become increasingly important. These include hot extremes and heatwaves, that we characterized by short, sudden spikes in temperatures and periods of prolonged positive temperature anomalies, respectively. Hot extremes in particular impact human health through high mortality and morbidity rates in certain demographics ([Medina-Ramón and Schwartz, 2007](#); [Oudin Åström et al., 2011](#); [Green et al., 2019](#); [Ebi et al., 2021](#)); can be damaging to agricultural and ecological systems through productivity decreases and heat stress ([Teixeira et al., 2013](#); [Gourdji et al., 2013](#); [Vogel et al., 2019](#)); and may pressure national economies by increasing energy demands ([Perera et al., 2020](#)) compounded with damage to thermal-power plant infrastructure ([Entriñen and Lordan, 2012](#)), as well as lowering labor productivity ([García-León et al., 2021](#)). Furthermore, climate change is expected to exacerbate hot extreme impacts through an increase in the frequency of their occurrence and their intensity ([Slater et al., 2021](#); [Fischer et al., 2021](#); [Thompson et al., 2022](#)). Uncovering the associations between hot extreme event magnitudes and the processes associated with the development of such events is therefore crucial for more robust understanding of future changes in extremes, leading to potential improvements in forecasting capabilities for emergency warning systems and supporting future adaptation and mitigation strategies.

Coupled with increasing mean global temperatures, recent record-shattering events highlight the importance of understanding not only the evolution and drivers of typical hot extremes, but also their variability. The Pacific North-West 2021 event is particularly illustrative, breaking the national temperature record by more than 4°C and leading to more than a thousand deaths ([White et al., 2023](#)). [Lin et al. \(2022\)](#) demonstrated that no ensemble member resolved such intense temperature 10 days before the event while analysis by [Lucarini et al. \(2023\)](#) suggested this event may become typical when conditioning on the existence of large and persistent temperature anomalies locally. Uncovering the physical differences that separate such rare events from lower-intensity hot extremes will provide insights into improving forecasting, in addition to furthering understanding of how they may evolve under climate change.

Although a good general understanding of mean-state hot extreme evolution exists and more detailed assessments have been made regionally to compare specific high-impact events, a knowledge-gap exists in quantifying the contribution of known physical processes

to hot extreme magnitude variability globally. This thesis attempts to close this gap by first proposing a systematic analysis of second-moment characteristics of hot extreme events and second applying state-of-the-art deep-learning methods to uncover patterns in trajectory features important for the development of hot extremes.

1.2 Outline

A comprehensive overview of hot extreme literature and a review of important meteorological background is presented in [Chapter 2](#). In addition, it will introduce the work that this thesis is based upon and extends. [Chapter 3](#) will describe the data used and illustrate some representative examples. [Chapter 4](#) will motivate and explain the employed methodologies, as well as provide the statistical background required to understand them. [Chapter 5](#) will present and discuss the obtained results. [Chapter 6](#) offers a broader discussion on the findings of this thesis, summarizes the work and suggests avenues for future research.

Chapter 2

Meteorological Background

It is important to recognize that the literature on heat extremes presents diverse and at times contradictory definitions of hot events. In this report, a ‘hot extreme event’ will describe an anomalously high spike in near-surface temperature occurring over the timescale of a day. A ‘heatwave’ will then describe an event of prolonged anomalously high near-surface temperature on the order of multiple days or weeks. In addition to reviewing literature on hot extremes, this section discusses heatwave research as they, although being distinct, often evolve together ([Robinson, 2001](#)).

2.1 Mechanisms of hot extreme development

The temperature of a volume of air - generally referred to as an air parcel in atmospheric sciences - can increase due to three primary mechanisms: warm **advection** involves the horizontal transport of warmer air parcels into a climatologically colder region; **adiabatic heating** is the result of descending air that compresses and thus heats; **diabatic heating** involves radiative effects including sensible heat fluxes from the Earth’s surface. These mechanisms provide a foundation to discuss atmospheric phenomena linked to hot extreme evolution and suggest a methodology, later discussed, to quantify the contribution of each physical processes to specific events.

In the mid-latitudes, anticyclonic flow situations serve as the basis for hot extreme development (e.g. [Pfahl and Wernli, 2012](#)). They are often associated with atmospheric blocks - a pattern characterized by persistent anticyclonic activity that prevents the influence of the prevailing flows ([Lupo, 2021](#)) - or amplified and quasi-stationary ridges ([Sousa et al., 2018](#)). Thereby, they develop regions of large subsidence and clear skies conducive to strong radiative heating in summer ([Lupo, 2021](#)). Indeed, atmospheric blocks co-locates with hot extremes, particularly well over the Northern ([Pfahl and Wernli, 2012](#); [Bieli et al., 2015](#); [Brunner et al., 2018](#)) and Southern ([Pezza et al., 2012](#)) mid-latitudes, although with significant spatial variability and large land-ocean contrasts. Over Europe, increased sensible heat fluxes from decreased soil moisture content and/or hydrological pre-conditioning are partly responsible for the observed land patterns, while the land-ocean contrast may be explained by a lower temperature variability due to water’s larger heat capacity.

In the polar latitudes, a diversity of atypical mechanisms may evolve due to the presence of snow and ice-sheets. Generally, an enhanced albedo effect implies lower diabatic heating and the temperature inversion created by a cold surface temperature limits adiabatic

heating. Furthermore, ice-sheets are found at relatively high altitudes (Fretwell et al., 2013) leading to adiabatic cooling from air parcels that ascend from coastal regions. Some studies have also suggested that intrusions of sub-polar air into the Arctic, such as by atmospheric rivers (AR) - corridors of rapidly ascending sub-tropical air that transport moisture and heat towards the poles - may disturb the polar climate (Ma et al., 2024; Papritz et al., 2023). This may lead to increasing downward longwave radiation that accelerates sea-ice melting (Kapsch et al., 2013; Mortin et al., 2016) and advective warming at both lower and upper levels (Papritz et al., 2023). Indeed, a case-study on the March 2022 East Antarctica heat extreme showed that intense atmospheric blocking initiated by an AR led to the weakening of the ice-sheet temperature inversion, intensifying adiabatic heating from subsiding air (Wille et al., 2024).

In the tropics, the weather is predominantly influenced by circulation anomalies (such as the El-Nino Southern Oscillation - ENSO) and local moist convective systems (although anti-cyclonic activity may still be a major synoptic driver). The limited literature on tropical hot extremes and heatwaves has found a strong association of hot extreme events with dry land conditions, indicative of strong coupling between land-surface processes and the atmosphere: anti-cyclonic activity yielding clear skies may enhance radiative heating and reduce soil-moisture (Geirinhas et al., 2018; Costa et al., 2022), in turn decreasing the potential of latent cooling - the loss of heat due to evaporation - and increasing sensible heat fluxes. Theoretical work by Byrne (2021) suggested that the onset of moist convection increases relative humidity which limits the growth of positive temperature anomalies.

Soil moisture and its complex dependence with near-surface temperature is not only important in the tropics, but also in the extra-tropics. Low soil-moisture content - dried by warm advection - is associated with increased sensible heating and decreased latent heating, which can reduce the potential for precipitation from deep convection and/or support a blocking high-pressure system, thus enhancing dry conditions (Berg et al., 2014; Wang et al., 2019; Miralles et al., 2014; Schumacher et al., 2019). Wehrli et al. (2019) showed that contributions from thermodynamic drivers such as soil moisture deficits are as important as circulation divers in 5 recent heat wave events. In the context of climate change, Vogel et al. (2019) quantified the contribution of these soil moisture and temperature feedbacks to regional increase in TX1day magnitude at more than 70% in Central Europe and Central North America, and around 50% in Amazonia, Northern Australia and Southern Africa. These mechanisms may evolve differently on small or relatively large temporal and spatial scales (Mccoll et al., 2019), further illustrating their complexity. However, a systematic and global analysis of their contribution to yearly hot extremes still lacks.

Hot extreme events over oceans have received little attention in the literature, likely due to their smaller impact on society. The absence of land-surface processes in the extra-tropics is expected to lead to a stronger influence of the large scale dynamics on hot extreme development, with adiabatic heating in the subsidence regions of the Hadley cell and advective warming in the mid-latitudes. Furthermore, we expect that regions with cooler sea surface temperatures (SSTs) than air temperature experience diabatic cooling, while in the tropics and over warm poleward currents, the opposite situation might be important with a warmer ocean surface inducing positive diabatic contributions. These processes may also involve coupling with marine heatwave, although this goes beyond the scope of this research.

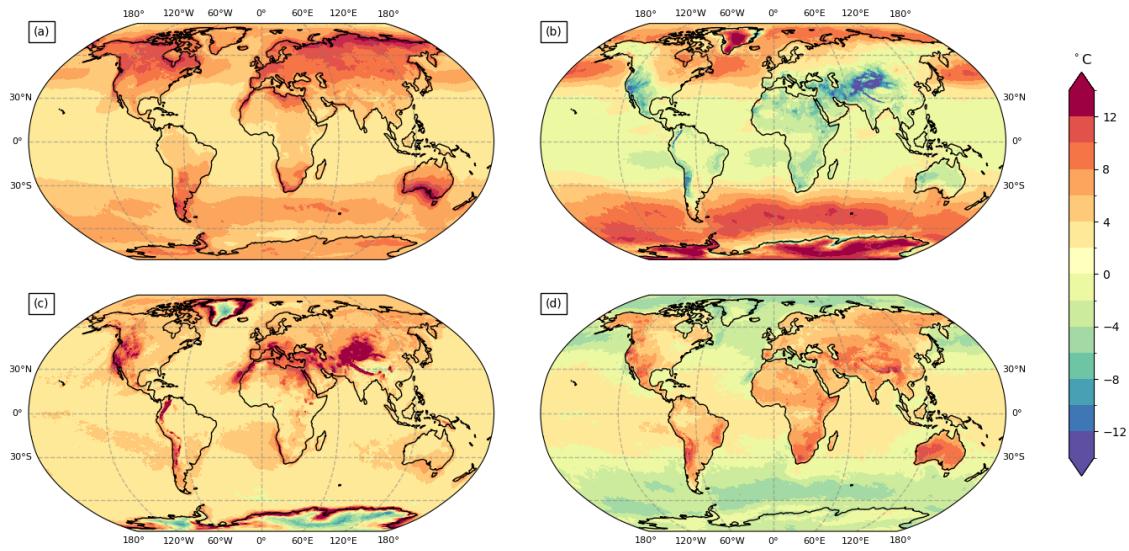


Figure 2.1 – Reproduced from Röthlisberger and Papritz (2023). The maps plot at every location the mean magnitude of (a) the temperature anomaly and the anomaly contribution from (b) advective, (c) adiabatic and (d) diabatic processes for yearly maxima TX1day events from 1980 to 2020.

2.2 Global quantification of physical mechanisms

Röthlisberger and Papritz (2023) was the first study to present a global quantification of the contribution of physical processes to hot extreme events. By considering the history of atmospheric air responsible for hot extreme events, it proposes a decomposition of the final temperature anomaly into the temperature anomaly gained through advective, adiabatic and diabatic processes. The methodology and the dataset created forms the basis of this thesis, and an in-depth description is presented in Chapter 4.1.

The analysis then quantifies the average influence of each process on hot extreme temperature anomaly magnitude (2.1). It showed that hot extreme events in higher latitudes and over land tend to have larger temperature anomalies than over the tropics and over oceans. The regions with largest temperature anomalies are in northern Russia and southern Australia. In addition, no single physical process is found to be dominant in the generation of temperature anomaly, leading to diverse geographical patterns for all three physical processes. The findings mostly agrees with the aforementioned literature, with exceptions over mountainous regions and isolated continental regions.

2.3 Variability of hot extreme events

So far, we have focused on the understanding of diverse mechanisms that yield hot extreme events across the globe. Several studies have however focused on relating the temporal variability of these processes to the resulting hot extremes variability. Here, variability is a measure of spread observed in a quantity across a temporal scale (daily, seasonally, yearly, etc.). In the European mid-latitudes, changes in variability have been attributed to an increase in anticyclonic activity and advective heating (Lenderink et al., 2007; Horton et al., 2015; Holmes et al., 2015) and to the local effects of soil-moisture (Quesada et al., 2012; Whan et al., 2015; Donat et al., 2017; Vogel et al., 2019). Wang et al. (2023) drew

similar conclusions for Asian mid-latitudes, further relating anti-cyclonic activity increases to the North Atlantic Oscillation.

The distribution of regional-to-global hot extremes is expected to both shift positively in the mean and undergo changes in variability due to climate change (Lewis and King, 2017; McKinnon et al., 2016; Slater et al., 2021; Fischer et al., 2021; Thompson et al., 2022). Simolo and Corti (2022) indeed suggested that daily temperature variability changes are more important than the observed warming trend in explaining amplifications in a majority of identified hotspot regions. In an attempt to relate these changes to dynamics, Fischer and Schär (2009) quantified the projected changes in hot extreme variability over Europe with respect to the processes associated with inter-annual, intra-seasonal and summer seasonal variability, finding the largest contribution from the former. Di Luca et al. (2020) quantified the contributions of annual and diurnal first and second-moment features to daily hot extremes over land, showing a diversity of global patterns. Suarez-Gutierrez et al. (2020) suggested instead investigating variability changes with regard to dynamical and thermodynamical sources in European mid-latitudes, and concludes that the former - encompassing effects of atmospheric blocking - is the main driver of hot extreme variability. However, these approaches lack a direct interpretation since the temporal features may be associated with different physical processes in different regions.

2.4 Methods in hot extreme literature

The weather extremes literature uses two main methodologies: sensitivity and trajectory analysis. To motivate the following work, we give a brief overview of these approaches:

Sensitivity analysis (SA) is the most widely adopted approach to study the effects of initial conditions and model parameters on outputs of a system. It involves running different experiments with simulation tools (including numerical and data-driven models) and assessing the changes in distribution of the response. Formally, sensitivity is defined as the partial derivative of the response with regard to a predictor of interest. It is then related to causality: if one assumes a linear structural model, sensitivity is a measure of the causal effect of a variable on the response. SA literature has considered both local and global measures of sensitivity. The former involves evaluating the sensitivity at a given predictor value and therefore cannot conclude general causal relationships without strong assumptions. The latter presents a difficult problem and different approaches have been proposed that include many typical statistical estimators (Razavi and Gupta, 2015).

Trajectory analysis (TA) follows the changes to air parcels along their flow within a dynamical system. This approach takes a Lagrangian perspective, as opposed to the Eulerian perspective that trackes changes to a volume of air at a fixed spatial coordinate. In the context of extreme weather, TA often involves the identification of an event within a numerical model followed by calculation of the flow backwards in time. The output is a timeseries of physical and model quantities that describe the state history of the parcel which gave rise to the extreme event. These characteristics have been used both to understand single events in case-studies (e.g. Stohl and James, 2004; Huang and Cui, 2015; Papritz and Röhlisberger, 2023) as well as to gain a general understanding of drivers of extreme weather by aggregation of multiple events (e.g. Bieli et al., 2015; Quinting and Reeder, 2017; Röhlisberger and Papritz, 2023).

This thesis uses TA to focus on process understanding of hot extreme development and will expand existing literature, introduced below.

2.5 Temperature anomaly decomposition

This investigation is based on a Lagrangian decomposition of temperature anomaly (Röthlisberger and Papritz, 2023), which quantifies the contributions of advective (adv), adiabatic (adiab) and diabatic (diab) processes in the generation of temperature anomaly (T'). Based on T' - i.e. temperature anomalies instead of absolute temperature - it leads to a more natural interpretation of the contributors, since some processes of interest - such as warming from horizontal advection - arise due to the presence of anomalously warm air within a dynamic environment.

At location \mathbf{x} and starting time of the backward trajectory t_0 , T' may be decomposed as follows:

$$\begin{aligned}
 T'(\mathbf{x}, t_X) = & - \int_{t_g}^{t_0} \frac{\partial \bar{T}}{\partial t} d\tau && \text{seas} \\
 & - \int_{t_g}^{t_0} \boldsymbol{\nu} \nabla_h \bar{T} d\tau && \text{adv} \\
 & + \int_{t_g}^{t_0} \left[\frac{\kappa T}{p} - \frac{\partial + \bar{T}}{\partial p} \right] \omega d\tau && \text{adiab} \\
 & + \int_{t_g}^{t_0} \left(\frac{p}{p_0} \right)^\kappa \frac{D\theta}{Dt} d\tau && \text{diab}
 \end{aligned} \tag{2.1}$$

where \bar{T} the temperature climatology; $t_g < t_X$ genesis time, the time at which T' was last zero; $\boldsymbol{\nu}$ the horizontal wind velocity; ∇_h the horizontal temperature gradient; p pressure; ω vertical velocity; θ potential temperature.

To study hot extremes globally and systematically, Röthlisberger and Papritz (2023) applied (2.1) to yearly maxima of 1-day temperature anomaly averages - denoted TX1day - at every grid point in ERA5 (further described in Data). Timeseries of backward trajectories - calculated with LAGRANTO (Sprenger and Wernli, 2015) - are numerically integrated to approximate the integrals. Since weather phenomenon might result from different flows, for yearly maxima event at each grid-point a total of 24 backward trajectories - corresponding to near-surface trajectory starts at 10, 30, 50 hPa and the 8 daily 3-hour timesteps - are integrated and then mean-aggregated. Each backward trajectory was back-computed until genesis $t = t_g$ up to a maximum of 120 time-steps (15 days).

Chapter 3

Data

3.1 ERA5 and trajectory calculations

Röthlisberger and Papritz (2023) applied their temperature anomaly decomposition (see [Temperature anomaly decomposition](#)) to ERA5 data from 1980 to 2020 with $0.5^\circ \times 0.5^\circ$ horizontal resolution, 137 vertical coordinates expressed as model-levels and 3-hourly temporal resolution. ERA5 data is model output that has been assimilated with observation and therefore represents the most complete global product of gridded meteorological data. Near-surface temperature is represented by 2m temperature and T' is calculated using transient climatologies of 9-year centered windows of 21-day periods, at every model-level. A complete description of ERA5 products may be found in [Hersbach et al. \(2020\)](#).

This thesis considers the dataset generated by Röthlisberger and Papritz (2023) in two ways: the first section will analyse the final decomposition of each TX1day event, while the second will analyse the timeseries that evolves when considering at each backward trajectory timestep the cumulative decomposition from genesis.

3.2 Cumulative total

The former may be considered as a budget since it represents the total T' accumulated over its trajectory. This yields an observational dataset containing - at every [latitude, longitude, year] index - the achieved T' magnitude and the three contributors corresponding to T' gained through advective, adiabatic and diabatic processes. We henceforth refer to these quantities as event T' , advective T' , adiabatic T' and diabatic T' . Numerical residuals in (2.1) - due to finite difference approximations of a continuous temporal domain - are usually small and thus ignored in future analysis. Furthermore, we assume that at a location [latitude, longitude], yearly samples are unaffected by local long-term temporal dependencies. Thus, for each location, the data consists in 41 independent draws from some multivariate distribution:

$$(T', \text{adv}, \text{adiab}, \text{diab}) \sim F$$

Initial exploration of the final decomposition shows a wide variety of marginal distributional behaviors, both for event T' and each contributor. Figure 3.1 illustrates this for three representative locations. The event T' histogram is bi-modal for the first location, normal for the second and right-skewed for the third. The bi-plots indicate mostly positive

association between event T' and each contributor, although this association is weak and even negative with advective T' for the first location. The largest TX1day T' does not necessarily develop from extremes in the contributors: for the first location the maximum event T' is achieved with near-median values of advective and adiabatic T' and only the 7th largest observed diabatic T' . The histograms of each contributor display mostly heavily skewed data and the pair-plots show mostly strong and diverse correlations between contributors as well as potentially non-linear associations.

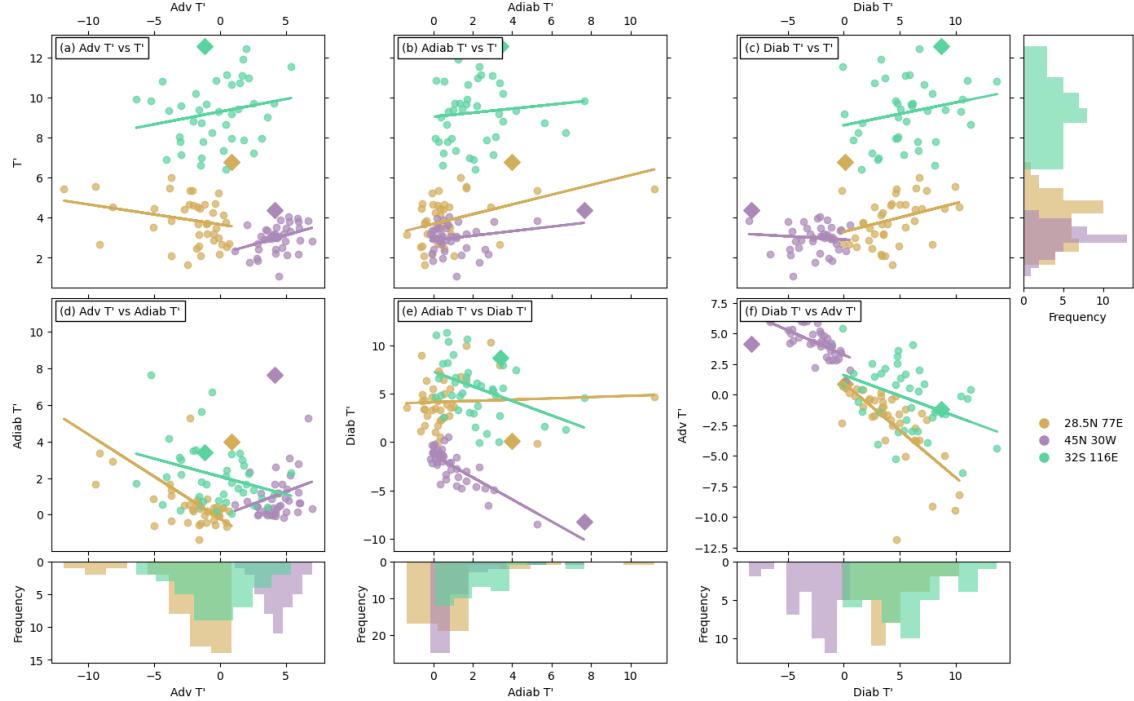


Figure 3.1 – (a-c) Plots of TX1day T' against the three contributors and (e-f) pair-plots for the contributors for years 1980 to 2020 for three locations. The T' histogram is plotted on the top-right of the figure, and the contributor histograms with 8 bins are plotted below (e-f). The TX1day event achieving largest T' over the 41 events is plotted as a diamond.

3.3 Trajectory timeseries

At every [latitude, longitude, year] index, the second dataset considers the time-cumulative evolution of the parcel's T' , decomposed into the three contributors. Since each event may have a different genesis-time, all series are padded with zeros before their genesis time, yielding at each location 41 timeseries of length 121 - a trajectory history of 15 days with 3h intervals. The data displays again diverse structures, year-to-year as well as across space. Figure 3.2 illustrates the year-to-year variability for one example location. Although some repeated structures are found in event T' and contributor T' 's across years, there are noisy periods and seemingly changing dynamics. For example, the graphs show some trajectories coupling and decoupling in time: the largest TX1day T' event exhibits advective and diabatic T' both decreasing until about 2 days before the event, then the diabatic T' steeply increases while the advective T' continues to decrease after a brief increase. The same pattern is observed between adiabatic and diabatic T' 's - although with negative correlation - up to 2 days before and then only little association is observed.

Some trajectories also exhibit sinusoidal behavior in diabatic T' with period around 12h (see a number of samples between 10 days to 3 days before the events) which is explained by the day/night variation in solar radiation.

On average, the timeseries exhibit strong auto-correlation with final T' at small time-lags, decreasing slowly until around 36h after which it cycles between 0.4 and 0.6. Strong auto-correlations are expected since the timeseries consider the accumulation of T' in time. Advective and adiabatic T' are poorly correlated with history older than 3 days, while diabatic T' remains above 0.4.

The cross-correlation between final event T' and the contributor T' 's is largest and positive for diabatic T' , slowly decreasing over 4 days with a sinusoidal behavior. The only negative correlation is with advective T' and seems to be largest at longer time-lags. Further timeseries plots for locations chosen in 3.1 are included in Appendix A.1 and show different characteristics.

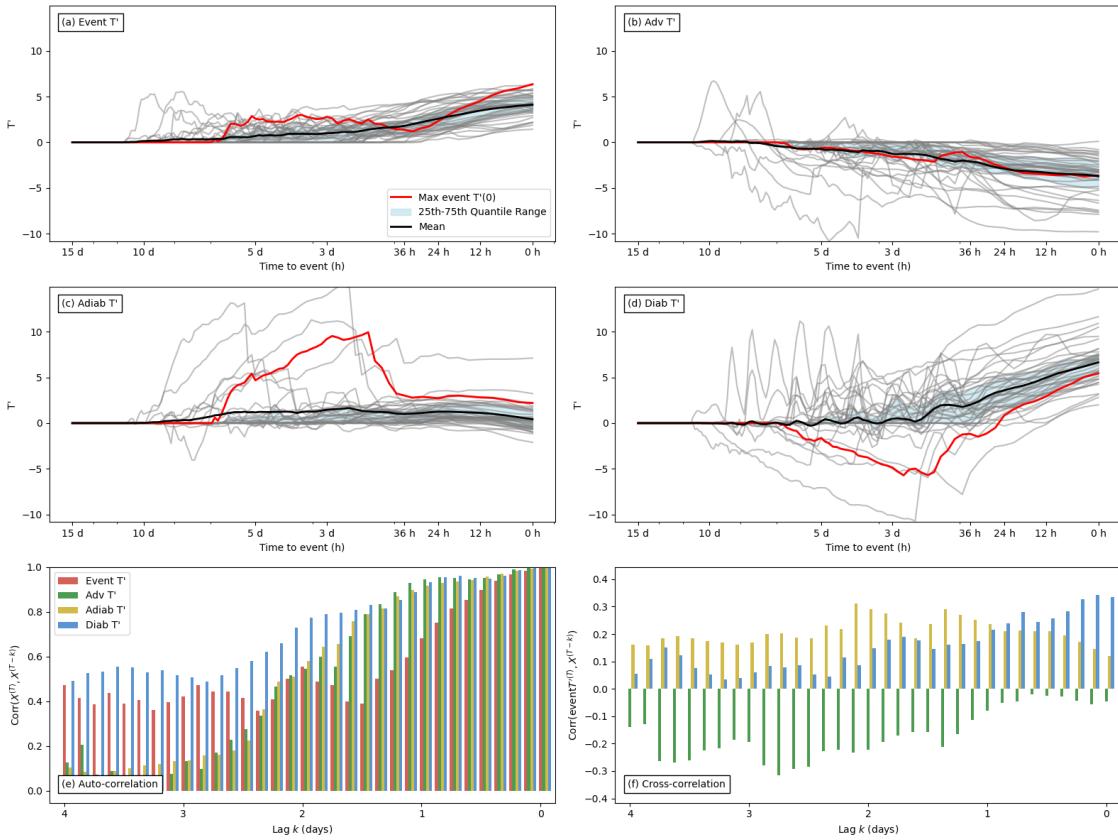


Figure 3.2 – (a-d) Plots of TX1day T' decompositions at 16°S , 48°W (vicinity of Brasilia, Brazil) for years 1980 to 2020. The x-axis is logarithmic to better represent the recent history. The trajectory yielding the largest final T' among all years is colored red, the time-mean of all trajectories in black and the inter-quantile range is highlighted in light blue. Plot (e) shows the (auto-)correlation between the final T' and the T' at different lags for the event and each contributor. Plot (f) shows the (cross-)correlation between the final event T' and the T' at different lags for each contributor. They are computed for all samples that have genesis time earlier than the considered time-lag.

A common approach to reduce auto-correlation in timeseries is compute first-order differences ($(X^{(t)} - X^{(t-1)})$ for time $t = 1, 2, \dots, T$). However, differenced data in this appli-

cation has been observed to yield noisy timeseries characterized by many local minima and maxima (see Appendix A.2) as well as considerable remaining auto-correlations. This behavior may hinder the ability of a machine learning model to learn meaningful patterns, and therefore the dataset is not differenced.

Chapter 4

Statistical methods

4.1 Variance decomposition of TX1day T'

The aim of this thesis is to gain further understanding in the influence of the three T' -generating mechanisms on TX1day events, particularly relating to the variability in yearly maxima observations. The ideal output of a statistical analysis would then be quantifying the influence of each contributor on a second-moment estimate of yearly maxima TX1day magnitude at each location.

Ranking the influence a certain predictor has on a response relative to other predictors is a common problem in applied statistics. Although R^2 is often reported to provide comparisons of model fit and thus measure the value of certain predictors in explaining the variance of the response, it does not take into account interactions between predictors. Instead, in the context of linear regression models, [Lindeman et al. \(1980\)](#) - referred to as LMG - propose to quantify predictor importance by computing the average sequential sums-of-squares over all predictor orderings. Averaging over all orderings allows this measure to capture both direct effects and secondary effects of a predictor. [Feldman \(2005\)](#) extended this approach with Proportional Marginal Variance Decomposition (PMVD) that instead considers a weighted average with weights chosen such as to enforce the exclusion criteria - a predictor with zero coefficient is allocated zero importance ([Grömping, 2007](#)). [Breiman \(2001\)](#), and extensions thereof, instead propose to measure importance of a predictor by assessing the model's loss of accuracy due to random permutations of the predictor values.

These methods have however been shown to be susceptible to data with highly correlated predictors, as is often the case in observational settings lacking causal assumptions. Furthermore, with many predictors, LMG and PMVD become computationally intensive, as they require computation of R^2 values over all possible orderings. Permutation-based methods scale linearly with the number of predictors but may require running many permutations of each column to output estimates with low variance.

[Chapter 3](#) has indeed demonstrated strong correlations between contributors that are due to inherent physical constraints. In addition, by assuming residuals to be approximately zero, the response and predictors are linearly dependent. This then limits the applicability of standard regression methods and the associated measures of predictor importance that assume them uncorrelated ([Hooker et al., 2021](#)). Addressing the constraint that the response is the sum of the three contributors, for example by transforming the contributors to a simplex space ([Greenacre, 2021](#)), is also problematic as the data does not follow the

positivity constraint: contributions to T' may be negative. Pre-processing the data to enforce positivity leads to poor representation of the domain and hinders interpretation. Furthermore, more robust scale measures - such as truncated variance - would not be appropriate as information of tail events that are often associated with the largest human impact needs to be preserved.

Then, to extend the first-moment characterization of TX1day events by [Röthlisberger and Papritz \(2023\)](#), we propose a variance decomposition of T' that quantifies the contributions of physical processes to variability. Recall that computation of (2.1) gave rise to residuals that may be assumed zero. Then the final T' of the event is related linearly to the T' generated by the three processes with a bias from the seasonality term, which is generally small. Then, for n independent events achieving temperature anomalies $\{T'\}_{1:n}$, data is described by $T'_i \approx \text{adv}_i + \text{adiab}_i + \text{diab}_i$ and therefore:

$$V(\{T'\}_{1:n}) \approx \sum_{c \in \{\text{adv, adiab, diab}\}} V(\{c\}_{1:n}) + \sum_{c, d \in \{\text{adv, adiab, diab}\}, c \neq d} \text{Cov}(\{c\}_{1:n}, \{d\}_{1:n}) \quad (4.1)$$

For further support, a brief background on variance and associated estimates is presented in Appendix B.

Applying (4.1) to every location on Earth yields a direct and systematic approach to the analysis of global patterns of hot extreme variability, while the presence of complex dependencies between physical processes is explored by the contributions from covariance terms. The linearity of the covariance operator and the fact that it is un-normalized, in comparison to correlation for instance, allows for an intuitive comparison of second-moment contributions. We argue that considering the variance of the contributors without adjusting for the influence of other contributors due to existing correlation is more informative. For example, a large variation in adiabatic T' contribution is important to report, regardless of the additional influence of advective and diabatic processes; if the sum of adiabatic and advective T' has small variance, it will be reflected in a negative covariance between the contributors. The limitation of using an un-normalized quantity is the fact that a large mean magnitude will in most cases imply a larger variance. Therefore, meaningful interpretation of second-moment estimates will also require consideration of first-moment estimates.

To summarize the spatial structure of TX1day T' variance and its decomposition, variance terms will be ranked with respect to their magnitude according to the following definition (applied to both mean and variance decomposition data): a contributor is dominant if it is at least twice as large as the second largest contributor; two contributors are dominant if they are both at least twice as large as the remaining contributor; else, no contributor is dominant. As both mean and variance decompositions are additive, this importance definition allows for a intuitive comparison. Covariance terms are omitted in these comparisons as they are observed to be generally more negative than the variance terms, except in few exceptions.

4.2 Principle Component Analysis of contributors

As discussed, strong inter-contributor correlations limit the interpretability of the variance decomposition, but are important in understanding the physical mechanisms associated

with hot extreme development and their constraints. To characterize the dependence structure between the contributors, Principal Component Analysis (PCA) is employed to assess the degrees of freedom of the heat-generating systems and provide an alternative importance definition to rank the contributors at each location.

PCA - known as Empirical Orthogonal Functions in climate sciences - is a linear transformation introduced by [Pearson \(1901\)](#) and first applied in climate sciences for weather prediction by [Lorenz \(1956\)](#). It is commonly used as a dimension reduction method - to obtain a low-dimensional representation of high-dimensional inputs - and in identifying and quantifying modes of variability such as the North-Atlantic Oscillation and the ENSO. The PCA problem consists in finding a projection $\mathbf{f}(\mathbf{x})$ that yields lower dimensional and orthogonal representations of the data that retain the most information.

Given n mean-centered samples $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{n \times p}$ of dimension p , the solution to obtain PCA transformation $\mathbf{W} = (w_{ij}) \in \mathbb{R}^{p \times l}$ for $1 \leq l \leq p$ and l -dimensional representations called PC scores $\mathbf{z}_1, \dots, \mathbf{z}_n \in \mathbb{R}^l$ is given by:

$$(\mathbf{W}, \mathbf{z}_1, \dots, \mathbf{z}_n) = \arg \min_{\mathbf{W}^\top \mathbf{W} = \mathbf{I}_l, \mathbf{z}} \sum_{i=1}^n \|\mathbf{z}_i \mathbf{W}^\top - \mathbf{x}_i\|_2^2 \quad (4.2)$$

with $f(\mathbf{x}_i) = \mathbf{z}_i = \mathbf{x}_i \mathbf{W}$. The unit column vectors $\mathbf{w}_{\bullet 1}, \dots, \mathbf{w}_{\bullet l}$ are related to the empirical covariance matrix Σ of \mathbf{X} as they are its eigenvectors corresponding to its eigenvalues $\lambda_{(1)}, \dots, \lambda_{(p)}$ ranked in decreasing magnitude:

$$\Sigma = \sum_{i=1}^p \lambda_{(i)} \mathbf{w}_{\bullet i} \mathbf{w}_{\bullet i}^\top \quad (4.3)$$

The first PC (PC1) then points in the direction with the largest variance, PC2 points in the direction orthogonal to PC1 with the largest residual variance, etc. The proportion of variance explained by a single PC can be computed using the eigenvalues:

$$\% \text{ explained variance (PC}i\text{)} = \frac{\lambda_{(i)}}{\lambda_{(1)} + \dots + \lambda_{(p)}} \quad (4.4)$$

Untruncated PCA is then applied to the advective, adiabatic and diabatic T' contributions of the 41 events at each location. To present an alternative measure of importance than the variance decomposition, each contributor is centered and scaled independently. This focuses the analysis on the contributor dependencies since it compares the variability structure without the influence of absolute magnitudes. The components are computed exactly using Singular Value Decomposition as proposed by [Halko et al. \(2011\)](#) and the analysis will consider the proportion of explained variance for each PC and relate the resulting directions of the PCs to the contributors.

Additionally, at each location we have three PC directions - i.e. the unit column vectors $\mathbf{w}_{\bullet 1}, \mathbf{w}_{\bullet 2}, \mathbf{w}_{\bullet 3}$ - that point in the orthogonal directions of largest variability. To represent these new axes at every location, we partition them in 7 categories, similarly to the dominance definition introduced in [4.1](#). Since the directions are unit-vectors in 3D space, the vector created by applying element-wise absolute-value lies on the positive octant of the surface of a unit sphere. This surface is partitioned into three corner regions (each contributor), the center (combination of all contributors) and three remaining surface divided evenly (pair-combinations), according to figure [4.1](#).

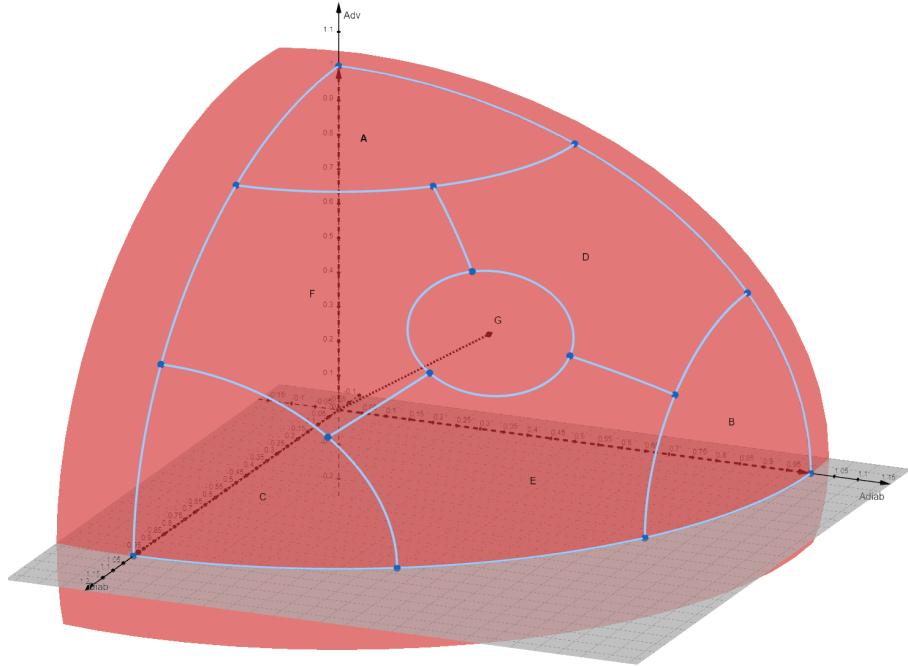


Figure 4.1 – Illustration of the partitioning of surface of PC absolute-value vectors into 7 categories: advective (A), adiabatic (B), diabatic (C), advective-adiabatic (D), adiabatic-diabatic (E), diabatic-advective (F) and all (D). The construction points all lie on planes parallel to the axes at values $\sqrt{2}/2$, 0.5 and $\sqrt{3}/2$ that correspond to 30, 45 and 60° angles between the axial planes.

Finally, PCA is preferred over machine-learning approaches because of the small sample sizes at each location, it is variance-based and therefore can be related to previous work, and has an intuitive interpretation.

4.3 Forecasting TX1day contributor trajectories

Considering final T' is useful for quantifying the cumulative contributions of physical processes to hot extreme events, as well as investigate the overall dependence patterns between them. To extend this analysis and provide a more precise assessment of the behavior of physical processes in the development of hot extreme events, we consider trajectory information. The added temporal structure allows the investigation of the underlying dynamics. For example, rapid T' growth of a certain physical process in the final day before an extreme event can be distinguished from a gradual gain that was previously reported as a single cumulative total. Understanding the timescales of advective, adiabatic and diabatic T' growth are particularly important for the development of warning systems in face of increasingly dangerous hot extreme events.

Trajectory-based analyses in hot extreme research are often carried out case-by-case. We propose leveraging the flexibility of Deep-Learning (DL) models to investigate the predictability of yearly maxima TX1day T' decomposition trajectories on a global scale. In this setting, a model is trained to forecast advective, adiabatic and diabatic T in the final period of trajectory given their past. A model that performs adequately can then be assumed to have learned meaningful patterns in the data that reflect the true dynamics of the system. Then, the model can be analysed to determine the contributors and

time-windows that most affect the prediction, thus unraveling temporal interactions of the physical mechanisms in the development of hot extreme events. Before introducing the model employed in this thesis, we give a brief overview of machine learning.

Machine learning and in particular DL architectures have recently received a lot of attention in atmospheric and climate sciences due to their performance in modelling complex high-dimensional systems and their computational efficiency in comparison to numerical weather prediction. DL models have largely been used to develop weather forecasting models (see for example Srivastava and S, 2022) and to learn parameterizations from simulated and observational datasets (Barahona et al., 2023). We first provide a brief introduction to a DL architecture that has been used extensively in timeseries applications.

Supervised DL involves the parameterization of arbitrarily complex functions - called networks - and estimation of these parameters by propagating input data through the network and evaluating the outputs' accuracy against the given truth - called the loss. By using a differentiable loss and network, the parameters can be updated. With sufficient data and learning iterations, the model parameters will be progressively updated to improve accuracy.

The modelling and processing of sequential data requires specific architectures such as Recurrent Neural Networks (RNNs) that can leverage sequential dependencies to learn representations of the underlying dynamical system. At a high level, RNN function by iteratively processing each time-step to update a set of hidden weights that retain memory - a memory cell. However, standard implementations have been shown to suffer from vanishing gradients - when parameter updates approach zero due to decaying gradient magnitudes, thus impeding learning - in applications with long-term temporal dependencies (Noh, 2021).

Long Short-Term Memory (LSTM) networks (Hochreiter and Schmidhuber, 1997) were the first extension of RNNs introduced to solve vanishing gradients. They contain a memory cell that is modulated by other units - called gates - to learn to determine when to ‘forget’ information. More formally, an LSTM memory cell is composed of hidden states H , cell states C and four gates: the input, forget, cell and output gates. Given a sequence of data, a memory cell recurrently propagates a data sequence one timestep at a time to update the gate, hidden and cell parameters. Given some initial hidden $H^{(t-1)}$ and cell states $C^{(t-1)}$, an input $x^{(t)}$ at time-step t is processed according to:

Gating:

$$\begin{aligned} i^{(t)} &= \sigma(W_{ii}x^{(t)} + b_{ii} + W_{hi}H^{(t-1)} + b_{hi}) \\ f^{(t)} &= \sigma(W_{if}x^{(t)} + b_{if} + W_{hf}H^{(t-1)} + b_{hf}) \\ g^{(t)} &= \tanh(W_{ig}x^{(t)} + b_{ig} + W_{hg}H^{(t-1)} + b_{hg}) \\ o^{(t)} &= \sigma(W_{io}x^{(t)} + b_{io} + W_{ho}H^{(t-1)} + b_{ho}) \end{aligned} \quad (4.5)$$

Memory update:

$$\begin{aligned} C^{(t)} &= f^{(t)} * C^{(t-1)} + i^{(t)} * g^{(t)} \\ H^{(t)} &= o^{(t)} * \tanh(C^{(t)}) \end{aligned} \quad (4.6)$$

where σ and \tanh are the sigma and tanh activation functions; $*$ represents element-wise

multiplication; W_{ij} and b_{ij} are matrices of learned weight and bias parameters for the input ($j = i$) and the hidden states ($j = h$); $i^{(t)}$, $f^{(t)}$, $g^{(t)}$ and $o^{(t)}$ are the forget, input and output gates at time t , respectively. A schematic representation can be found in Figure 4.2.

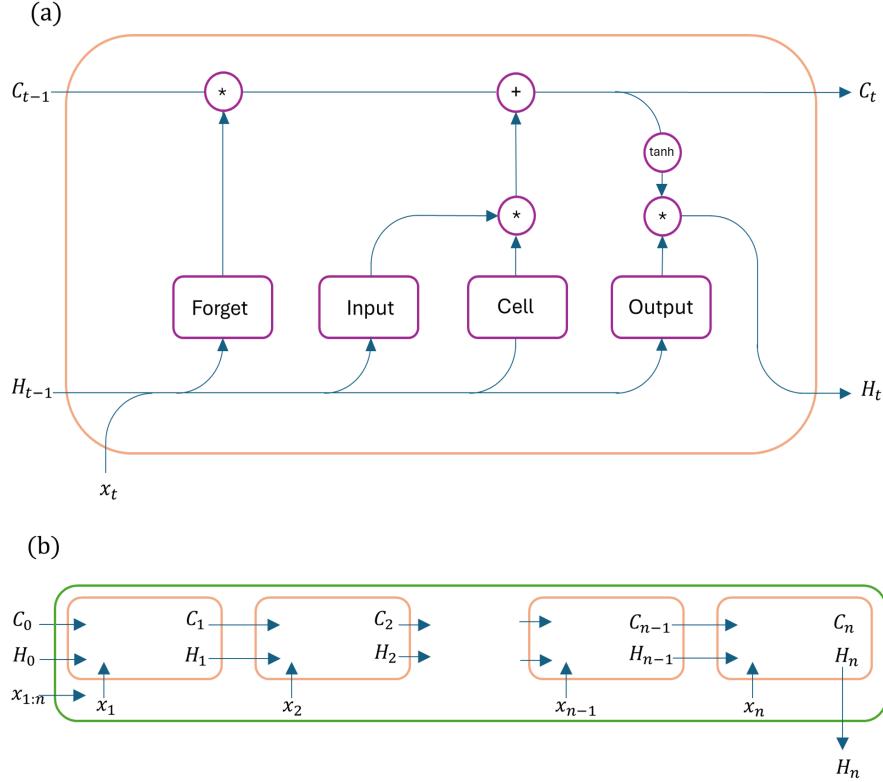


Figure 4.2 – (a) Representation of an LSTM memory cell (in orange) that, given initial hidden states $H^{(t-1)}$, cell states $C^{(t-1)}$ and input $x^{(t)}$, outputs the updated hidden $H^{(t)}$ and cell $C^{(t)}$ states. Purple shapes represent operations: * matrix multiplication, + element-wise addition, tanh the tanh activation function, and the Forget, Input and Output gates are as defined in equations 4.5. (b) The same LSTM cell but unfolded. For an input of initial hidden states $H^{(0)}$, cell states $C^{(0)}$ and data sequence $x^{(1:T)}$, it outputs final hidden states $H^{(T)}$ that can then be further processed (e.g. by a fully connected layer).

LSTM layers, composed of a number of memory cells, may be used in different ways depending on the data structure and task at hand. In timeseries forecasting applications, most models use the final hidden states $H^{(T)}$ of the last LSTM layer as lower-dimensional feature representations of every input that are then further processed by one or more fully-connected NN layers. There is a rich and evolving literature on LSTM-based architectures and LSTM variants and thus refer the reader to comprehensive review papers such as [Hu et al. \(2020\)](#) and [Han et al. \(2021\)](#).

Here, a global LSTM-based DL model is trained to forecast the final 8 timesteps (1 day) of each contributor of the TX1day event decomposition trajectories given their history. The model is composed of a single LSTM layer and a linear output layer to process the final hidden states $H^{(T)}$ into 24 outputs (3 variables x 8 timesteps). The rich diversity of timeseries behaviors and patterns as well as the occurrence of anomalous large negative and positive values motivates careful design of the optimization procedure.

First, the three contributors in the training data are centered by the global median and scaled according to the global interquartile range, as these quantities are more robust than the mean and variance and scaling has been shown to stabilize training. Second, the model is trained with stochastic gradient descent. Although this is computationally less efficient than mini-batch gradient descent for such large training datasets, initial tests showed vanishing gradients when using batches leading to constant forecasts. Third, the Huber-loss is employed to leverage the benefits of the mean-squared-error and mean-absolute-error loss: improved stability near zero and decreased sensitivity to extreme values, respectively. The Huber-loss threshold is set to its typical value 1. Fourth, generalization is improved by using L2-regularization that has been shown to be particularly beneficial for prediction tasks on multivariate time series (Zhou et al., 2019). It penalizes the magnitude of model weights to ensure that the resulting function is less sensitive to small perturbations, thus limiting over-fitting. Fifth, a dropout layer is introduced between the LSTM and fully connected layers to further promote model generalization.

Then, given network output $\hat{\mathbf{y}} = (\hat{\mathbf{y}}_{\text{adv}}, \hat{\mathbf{y}}_{\text{adiab}}, \hat{\mathbf{y}}_{\text{diab}})$ and truth $\mathbf{y} = (\mathbf{y}_{\text{adv}}, \mathbf{y}_{\text{adiab}}, \mathbf{y}_{\text{diab}})$, model parameters \mathbf{W} , the L2-regularization hyper-parameter r and Huber-loss threshold δ , the objective function is given by:

$$\begin{aligned} \text{obj}(\hat{\mathbf{y}}, \mathbf{y}) &= \frac{1}{24} \sum_{i=1}^{24} l(\hat{y}_i, y_i) + r \cdot \|\mathbf{W}\|_2 \\ l(\hat{y}, y) &= \begin{cases} 0.5(y - \hat{y})^2, & \text{if } |y - \hat{y}| < \delta \\ \delta(|y - \hat{y}| - 0.5\delta), & \text{otherwise} \end{cases} \end{aligned} \quad (4.7)$$

where $\|\cdot\|_2$ is the L2-norm and l_i the Huber-loss for one prediction.

Due to the high-resolution of ERA5 data, the dataset contains over 10 million timeseries. To select the model hyper-parameters (L2-regularization weight, LSTM hidden unit size), the globe is divided into 16 rectangular regions according to figure 4.3: 8 for training, 8 for validation. In addition, the timeseries exhibit large spatial dependencies within a given year since they may share the same flow. Therefore, to have a more faithful estimation of generalization performance, the validation regions are reduced in latitude and longitude extent by a factor proportional to the maximum of the average trajectory distance in each region. Finally, to limit computational burden and perform inference on unseen data everywhere on the globe, only years from 1980 to 1990 are used for training.

The final model parameters are obtained by randomly sampling 50'000 timeseries ($\sim 4\%$ of all training data) and training a model instance for each combination of LSTM layer hidden unit size [48, 64, 88] and L2-regularization weight $[0, 10^{-6}, 10^{-4}]$. An estimate of generalization loss is computed on half of the validation data, randomly sampled. Each instance is trained with the Adam algorithm (Kingma and Ba, 2017), a learning rate of 2×10^{-4} and for a maximum of 15 epochs, stopped early if validation loss does not improve within three epochs. LSTM hidden and cells states are re-initialized at every input with zero matrices. The parameters of the final model are chosen by considering low validation loss (see Appendix C.1 for results) relative to model complexity and expected behavior on a larger dataset. The final model is trained on the full training data and is used to forecast each test timeseries, consisting of all locations for years 1991 to 2020. To obtain a reference predictive performance, a simple constant model is created that outputs the latest input value for each contributor - i.e. the value at time 27h before the event.

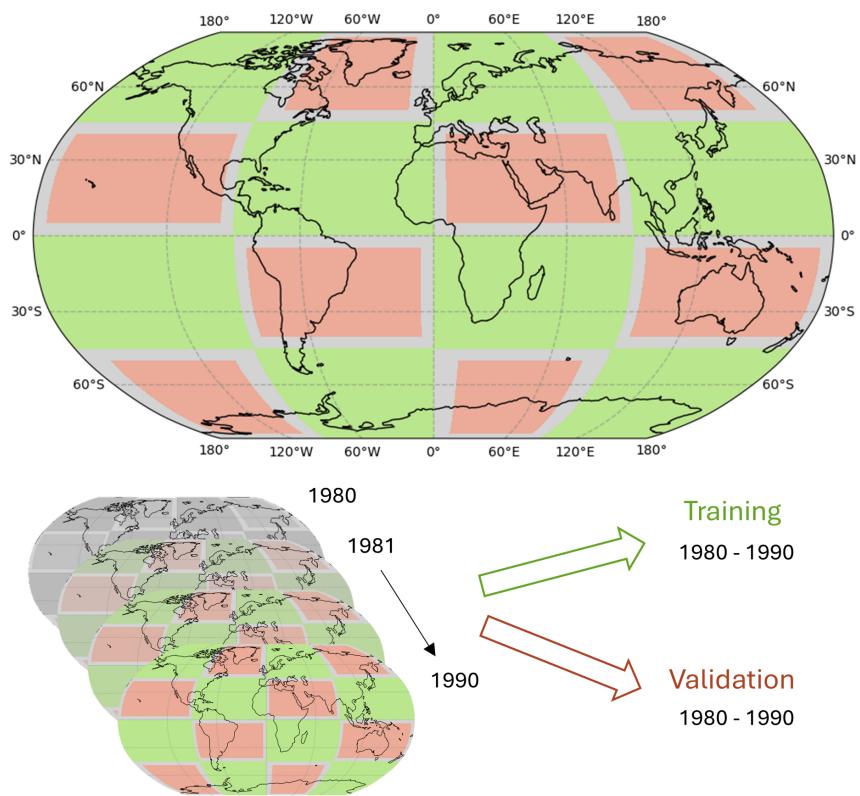


Figure 4.3 – Training and validation splitting scheme. Years between 1980 and 2000 are used for training (green shaded areas) and validation (red shaded areas). The remaining years 2001 to 2020 and the whole globe are used for inference.

Chapter 5

Results

5.1 Relating TX1day T' variability to component variability

The variance decomposition on 40 years of ERA5 data reveals large spatial diversity in year-to-year TX1day variance (Figure 5.1). As anticipated, the patterns closely resemble those of the mean of TX1day magnitudes in (Röthlisberger and Papritz, 2023), since the variance is un-normalized and therefore scales with magnitude. There is a strong ocean-land contrast, with low variances in the order of 2 K^2 over oceans and mostly above 4 K^2 over land. A general trend over land is the increase in variability further from the Equator. There exists nonetheless important exceptions: the West coasts of North America, West coasts of Southern Chile, South-western coasts of Australia and the West coasts between Mauritania and Northern Europe all exhibit large TX1day variability, while tropical land such as the Amazonia, South-east Asia and Indonesia, equatorial Africa and parts of the Eastern Saharan show low variability compared to the oceans. The largest peaks of variability occur in the bay of the Antarctic Peninsula and an oceanic hot spot centered at $70^\circ \text{ S } 160^\circ \text{ W}$, as well as relatively large hotspots in the coastal regions described above.

Individual contributors have mostly larger variance and vary spatially with no single contributor dominating globally. The advective and adiabatic T' variance have similar spatial distribution over land, while they seem to complement each other over oceans with large adiabatic T' variance in the extra-tropics and large adiabatic T' variance in the sub-tropics. Key hotspots of very large variance for both contributors include regions of high topography (particularly over Rockies, the Andes and the Himalayas) as well as the Arctic and Antarctic ice sheets, with values mostly above 32 K^2 and 64 K^2 respectively. In comparison, the variances of diabatic T' are more uniform with a less pronounced ocean-land contrast, and are particularly small over oceans east of the continents at the Equator with values between 0 and 2 K^2 . The storm track regions, especially the upstream portions of the NH, have larger advective T' variance than the surrounding oceans, which resembles the TX1day variance map.

Covariances between contributors are predominantly negative and again display diverse spatial patterns. The covariances between advective and adiabatic T' are positive only over the extra-tropical oceans; between advective and diabatic T' over small regions in the tropical oceans and the Greenland and Antarctic ice sheets; between diabatic and adiabatic T' over regions of high topography and within the African and Australian continents. Over oceans, the covariance between advective-adiabatic and advective-diabatic complement

each other with positive and negative contributions, while the covariance between diabatic and adiabatic T' is almost always negative.

Spatial patterns in mean magnitude of contributors mainly agree with those in the variance over land, but not over the oceans. The advective T' patterns are most different on the coast of Western North-America, coastal North-western Africa and the ocean region West of Australia. The diabatic T' patterns are the most different, in particular over Southern hemisphere oceans, where the mean patterns show much stronger latitudinal gradient and miss important large-variance regions on the west of the continents. The adiabatic T' patterns instead shows good agreement over the oceans, while some differences exist in the Sahara, eastern North-America and the storm tracks, both in the Northern and Southern hemispheres.

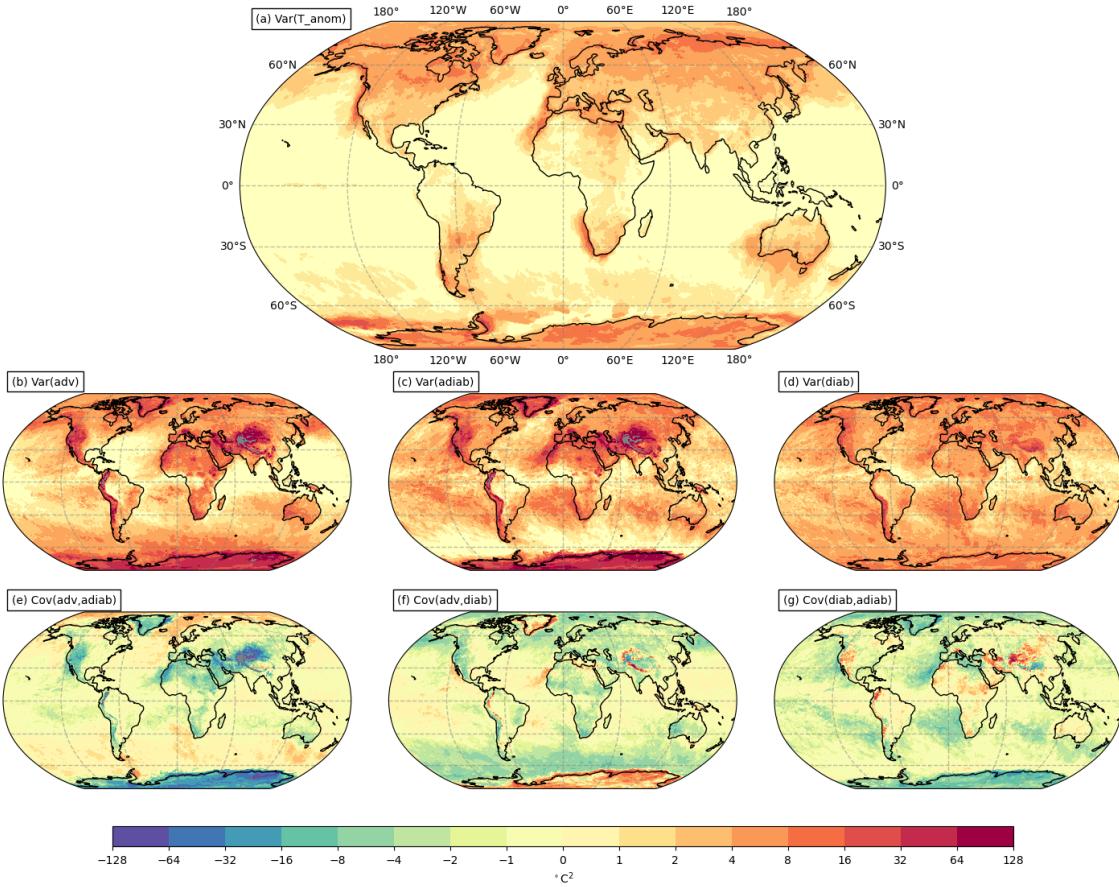


Figure 5.1 – (a) Variance of TX1day series decomposed into (b) variance of advective T' , (c) adiabatic T' , (d) diabatic T' , and the the covariances between the contributors (e,f,g). Grid points with very high variances or covariances magnitudes (below -128 and above 128, corresponding to less than 0.6% of all observations) are colored grey.

Similarly to mean-characteristics, the ocean-land contrasts in TX1day variance can be explained by water's larger heat capacity, since a slower temperature gain by solar radiation limits the short-term gain in temperature of a parcel. This is supported by a more uniform diabatic T' variance over oceans and suggests that cloud cover can influence T' growth particularly at the end of TX1day event trajectories when parcels are close to the ocean surface. Regions of larger variance over oceans - in TX1day T' and all contributors - are found near the coasts and can be attributed to trajectories that predominately evolve over

land and therefore involve more variable heat-generating mechanisms.

For example, the variability hotspot located on North-Western African coast involves the typical summer-time Azores high-pressure system inducing trade winds from North and the Saharan low-pressure system inducing the Harmattan winds that originate from the continent ([Adame et al., 2022](#)), converging on the Moroccan coast and travelling in a South-westerly direction along the coast. Yearly variability in this synoptic situation can then lead to changes in the contributions from trade winds or Harmattan winds, that evolve very differently.

The mean decomposition showed large cancellation between contributors in many regions. The generally large and negative covariances supports this fact. Furthermore, these regions tend to exhibit large variance in both/either/one of the contributing processes. The Northern and Southern Hemisphere storm tracks have large mean positive contributions from advective T' with large dampening by diabatic T' , while small but positive contributions from adiabatic T' . This study confirms that observed variability in the TX1day T' is impacted both by variability in these competing mechanisms, but additionally finds moderate contribution from adiabatic T' and the covariance between advective and adiabatic T' in the upstream portion of the storm tracks, that has previously been neglected. This observation is also noted in Eastern Brazil, continental southern Africa, Australia and the northern East coast of North America - although with diabatic T' and advective T' contributing positively and negatively in the mean, respectively. This is highly significant since changes in variability of hot extremes under climate change may not only be influenced by changes in mean-dominant physical processes, but also those that affect second-moment characteristics, supporting work by [Simolo and Corti \(2022\)](#).

Assessing the variability in these dampening mechanisms, or attributing importance of the positive or negative contributor, is difficult without considering the temporal evolution of these interactions. However, the spatial patterns in the covariance between the contributors is then a useful tool in determining regions where competing effects are present and influence the observed variability of resulting TX1day events. Furthermore, the diversity in covariance patterns is suggestive of complex and sensitive TX1day events.

5.2 Comparing contributions to TX1day T' mean and variance

Figure 5.2 maps the most important contributor/s - according to the definition introduced in [Methods](#) - for the mean ([Röhlisberger and Papritz, 2023](#)) and variance decompositions to more concisely compare their differences.

In most oceanic and subtropical land, advective and/or adiabatic T' dominates the mean while the variance is characterized additionally by significant contribution from diabatic T' . This is highlighted in the Northern and Southern Hemisphere storm tracks. In remaining regions with mean-dominating diabatic T' , the variance of TX1day T' is dominated by diabatic T' and most often advective T' - such as in Equatorial Africa, Eastern South America, Western Australia and Eastern India. In fact, only few regions have a single dominating contributor in the variance. The North-western African, South-western African and Western North-American coasts, a large part of the Mediterranean and the tropical Pacific are dominated by adiabatic processes. Small regions in the Amazonia, Western equatorial Africa and small regions south of the Northern Hemisphere storm tracks are dominated by diabatic processes. Finally, the upstream portion of the Northern

Hemisphere storm tracks and a small coastal band around the Antarctic are dominated by advective processes.

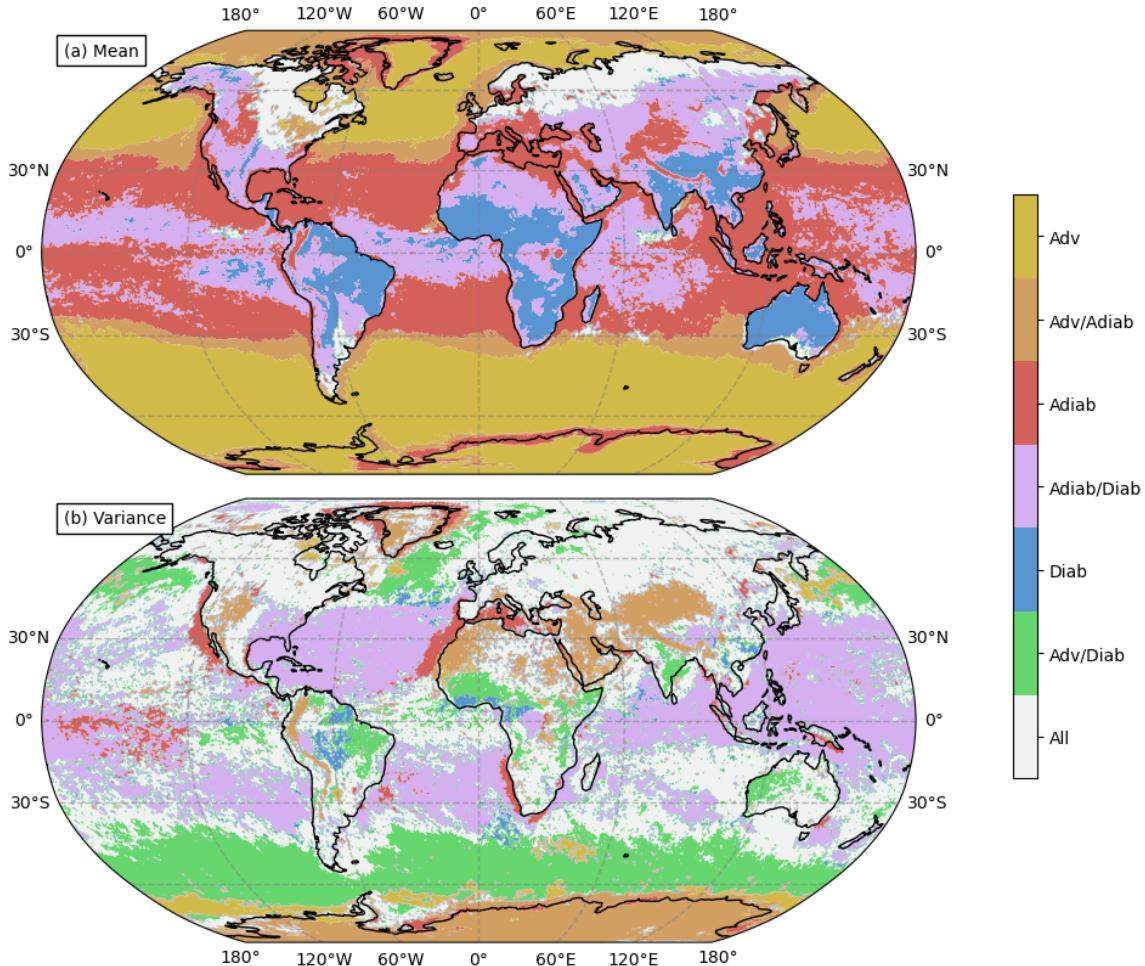


Figure 5.2 – (a) Reproduced from Röthlisberger and Papritz (2023). Maps at every location the most important contributor to average TX1day yearly maxima magnitude. **(b)** Maps at every location the most important variance term to variance of TX1day yearly maxima magnitudes. Contributor A is most important if it is 2 times larger than the second largest contributor. Contributors A and B are collectively most important if the smallest of the two is 2 times larger than the least important contributor. Else, all contributors are deemed important.

The coastal regions that showed peaks in TX1day T' variance are primarily dominated by adiabatic T' both in the variance and in the mean - the only contributor exhibiting perfect agreement globally. There is nonetheless a clear land-ocean contrast in the variance spatial patterns, as opposed to adiabatic dominating also over small land regions in the Mediterranean, the North-american east coast and the coast of Morocco. This indicates that the land-ocean boundaries not only strongly influence interdecadal variability of large-scale circulation patterns (Wake, 2014), but also local variability of hot extreme mechanisms.

Regions of high-topography mostly show adiabatic T' dominating the mean with the additional advective T' influence in the variance. This is expected given the large negative covariances between advective and adiabatic T' in these regions, suggesting that hot extreme development is modulated by mechanisms with competing processes. However,

regions with advective and adiabatic T' dominance in the variance tend to cover larger areas, particularly over the Himalayas where this regions extends from the Caspian sea to Southern Mongolia, where diabatic T' is a dominating term for the mean. Katabatic winds - high-density air flows that descend higher elevation slopes due to the force of gravity, thus represented as a positive adiabatic contribution - are highly prevalent in the Himalayas due the diabatic cooling of rising air by glacial surfaces. Under climate-change, observations have shown an increase in katabatic wind activity ([Salerno et al., 2023](#)) that then suggest the advective cooling of these flows and its dampening effect on the adiabatic parcel are more susceptible to yearly variability than the contribution from diabatic heating such as from solar radiation along the parcel trajectory and diabatic cooling due to the influence of glaciers.

Complete disagreement - i.e. no common contributors - between the maps involve only two regions. In the Southwestern Tasman Sea, advective T' mostly dominates the mean while adiabatic and diabatic T' dominate the variance. This may have important consequences for the understanding of marine heatwave development in the Tasman sea, which are attributed to both marine and atmospheric drivers, such as the positive phase of the Asymmetric Southern Annular Mode that is associated with lower wind speeds and increased downward solar radiation ([Gregory et al., 2023](#)). Since this driver does not occur every year, annual changes in diabatic T' contributions may be more important than advective T' changes in distinguishing between TX1day T' magnitudes. In regions on the North-side of Asir mountain range (Arabian Peninsula), diabatic T' mean-dominance is contrasted by advective and adiabatic T' variance-dominance. In addition to the fact figure 5.1 shows strong advective - adiabatic covariance in this region, this analysis suggests that year-to-year variability of the Indian Summer Monsoon - associated with strong subsidence and strengthening of northerly wind ([Attada et al., 2019](#)) - particularly influences hot extreme magnitude variability in the Asir mountains and in its forefront.

Although the larger influence of diabatic T' globally supports previous evidence that radiative and land surface processes dominate sub-tropical changes in summer-time near-surface temperature variability ([Holmes et al., 2015](#)), this work further highlights the importance in considering all three heat-generating mechanisms. Particularly over Australia, advective processes have been understood to be dominating the observed changes in summertime near-surface temperature variability ([Watterson et al., 2008; Holmes et al., 2015](#)). This study agrees, but further insists that for hot extremes, both diabatic - specifically in the West - and adiabatic processes cannot be ignored.

Then the variance of TX1day T' is almost entirely dominated by more processes than its mean counterpart. The implications is that process understanding of hot extreme development, particularly over land, requires more careful examination of changes in the physical processes due to climate changes. This shows that less important mechanisms for mean characterization of TX1day events become significant when quantifying year-to-year variability.

5.3 Characterizing contributor dependencies

The covariances discussed in 4.1 are useful in quantifying the importance of contributor interactions to TX1day T' variance. Since the contributors may evolve on different magnitude scales in different parts of the world, covariances may be inadequate in investigate their dependence structures. This section adjusts for the relative scales of each contributor

and uses PCA to focus on characterizing the observed variability only in the contributors.

Figure 5.3 shows that there are only few regions that have over 90% of explained variance by PC1. These are almost only over oceans and mostly in the vicinity of storm tracks - particularly between Australia and Antarctica and its coasts with over 95% of explained variance - as well as over the tropical Pacific and in the Atlantic between Brazil and the Sahel region collocating with sea-surface temperature maxima. Additionally, the Northern Atlantic and Arctic Ocean have mostly over 80% of variance explained by PC1.

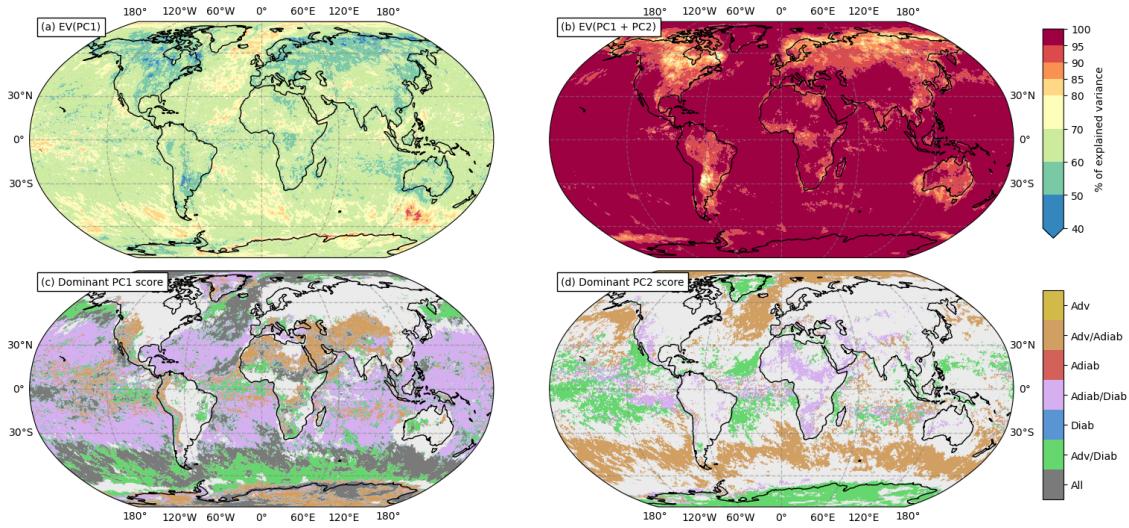


Figure 5.3 – Percentage of variance in advective, adiabatic and diabatic T' that is explained by (a) PC1, (b) PC1 and PC2. The contributor directions most aligning with PC1 and PC2 - see 5.2 - for regions with at least 95% of variance explained by PC1 and PC2 are shown in (c) and (d), respectively. Regions with less are masked in light grey.

The land regions with over 70% of explained variance are generally limited to regions in the vicinity of high topography, but also Western and equatorial Africa and Northern Colombia. The remaining land regions have between 34 and 70% of variance explained by PC1 - with the lowest values achieved in Northern Argentina, central Canada and Arctic Russia - and correspond to the areas requiring PC3 to explain more than 60% of their T' variance.

The PC scores in figure 5.3 (c-d) shows that even in regions with over 95% of the variance explained by PC1 and PC2 - although being constrained to two degrees of freedom - the PCs do not clearly align with a single or two contributors. In fact, the PC1 map supports figure 5.2 with very similar spatial patterns. All single-dominating regions in figure 5.2 are however now shown to be variable in 2 or three axes. In tropical regions, where small variances may have influenced the reliability the utilized definition of dominance, some regions with no dominating variance term now show PC1 aligning mostly with advective and either adiabatic or diabatic T' . Finally, as geometrically expected, the PC2 score aligns most with the opposing contributors.

Further interpreting the PC components is difficult without considering the temporal evolution, since the PCA is applied to the final decomposition which includes aggregated contributions. Nonetheless, this analysis demonstrates that hot extreme event trajectories involve strong dependencies between the three contributors and that they may be represented using only one or two degrees of freedom. In addition, the resulting axes of

variability do not align well with the original contributors, but instead mostly involve both advective and adiabatic T' over the oceans and both adiabatic and diabatic T' over land.

5.4 Forecasting the last 24h of TX1day trajectories

To more specifically study the dynamics between contributors of hot extreme events, we now consider the history of parcel trajectories and model them with an LSTM-based NN architecture. Recall that the final model is trained on half of all grid-points as illustrated in figure 4.3 for years between 1980 to 1990. Then, figure 5.4 shows the accuracy of this model in predicting the final 24h of advective, adiabatic and diabatic T' evolution, for the test set that consists of all grid-points for years between 1991 and 2020.

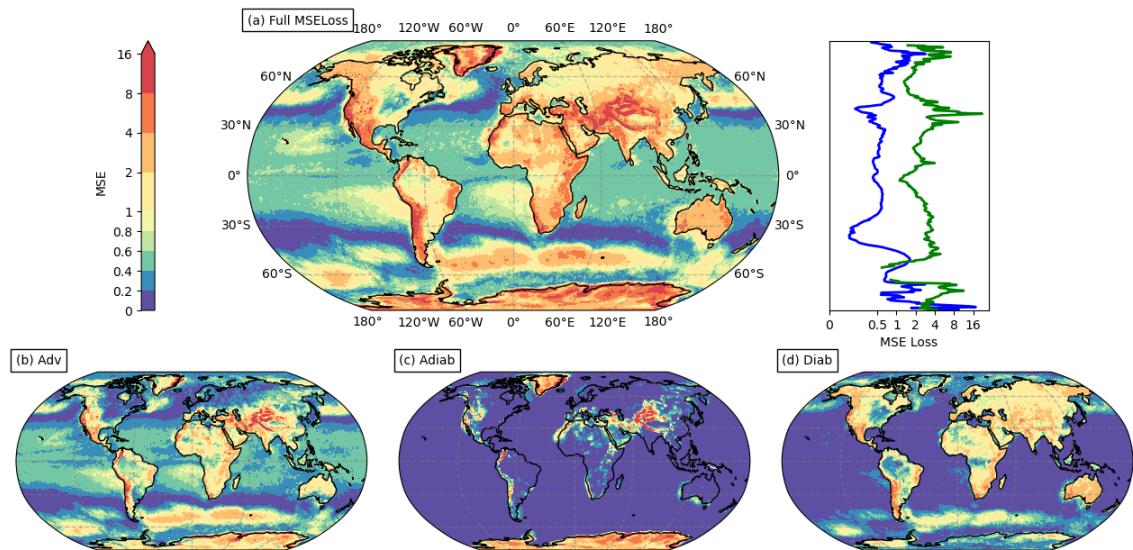


Figure 5.4 – Median mean-squared-error (MSE) loss over years 1991 to 2020 for the (a) total, (b) advective, (c) adiabatic and (d) diabatic forecasts. The line graph shows oceanic and terrestrial latitudinal averages in blue and green, respectively.

The model performs best over oceans in longitude bands between 30 and 45° latitude North and South, with MSE loss lower than 0.2 for all contributors. There is a strong land-ocean contrast evident in the full forecast as well as for each contributor, with exceptions for adiabatic T' that performs very well over many land regions. The land areas that achieve low MSE loss include eastern North-America, regions east of the Andes, northern Eurasia and east China, with all three contributors being well-forecasted. The regions exhibiting the poorest performance include regions of high topography, the Arctic and Antarctic ice sheets. In addition, hotspots that were observed to have large mean and variance in section 5.1 such as the east coasts of Northern Africa, South and North America also have peaks in MSE loss that reach above 8°C².

Generally, the oceans are forecasted accurately, with only the advective term contributing considerable MSE loss in the tropics and both Northern and Southern storm-tracks. The two regions off the coasts of Southern Africa and South America that report 0.6 to 2°C² median MSE loss collocate approximately with SST patterns associated with ENSO. During positive and negative phases of ENSO, SST gradients in the tropical to subtropical eastern Pacific strengthen significantly, coupled with increased wind stress (for e.g. Chelton et al., 2001). Poor predictability of the advective T' in this region may then be due to

more diverse trends in the evolution of advective processes and their association to the ocean surface and diabatic processes. The hotspot in the Southern Atlantic may be further explained by similar SST variability associated with the South Atlantic Dipole, as well as its year-lagged teleconnection to ENSO (Ham et al., 2021).

Figure 5.5 shows the accuracy of the baseline model: constant prediction of value 27h before the event for each contributor. This reveals a number of important insights both into the data and the limited performance of the LSTM model. A low MSE loss for tropical to mid-latitudinal oceanic regions is reasonable given that they also record the lowest average T' magnitudes, with the areas that show larger mean adiabatic T' also exhibiting larger MSE loss. The baseline outperforms the LSTM model in this areas, particularly due to advective T' that is often wrongly predicted to increase by the LSTM.

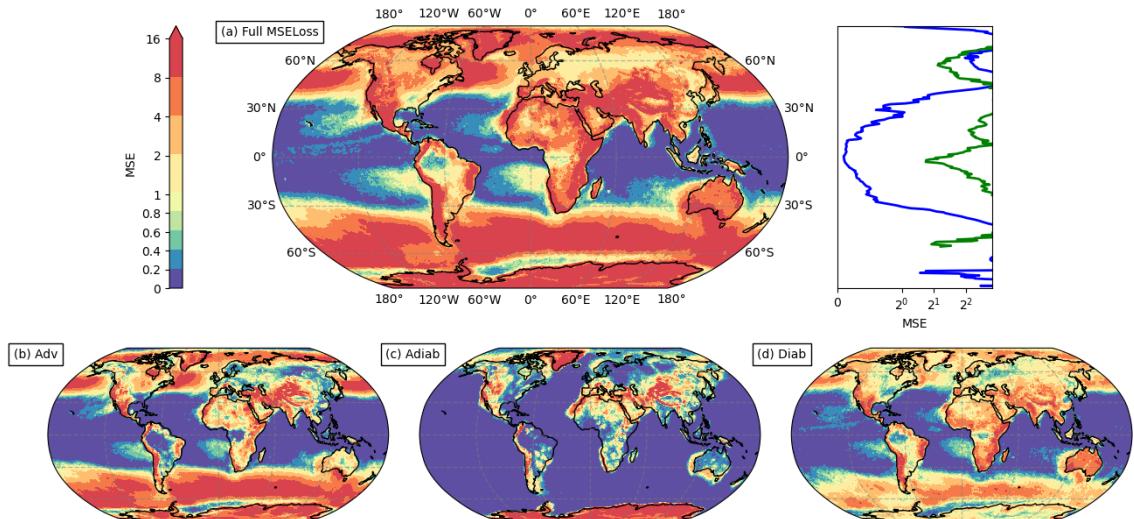


Figure 5.5 – Median mean-squared-error (MSE) loss over years 1991 to 2020 for the (a) total, (b) advective, (c) adiabatic and (d) diabatic forecasts. The line graph shows oceanic and terrestrial latitudinal averages in blue and green, respectively.

The adiabatic term shows very small MSE loss over all oceans as well as parts of South America, eastern North American, India, Northern Eurasia and parts of Australia, suggesting that globally, little to no additional adiabatic heating occurs in the 24h preceding a TX1day event. However, previous work showed large adiabatic contributions to both mean and variance of event T' . We therefore hypothesize that hot extreme events over oceans tend to evolve from parcel trajectories that descend to the surface with significant positive contributing from adiabatic T' at least 24h before the event. The final 24h period may then be dominated by diabatic T' from atmosphere interactions with the ocean surface and advective T' from prevailing circulation patterns. Exceptions occur consistently in coastal regions or in the Arctic circle, where literature supports persistent subsidence when the typical ice-sheet temperature inversion is weakened (Wille et al., 2024).

We further investigate the regions successfully modeled by the LSTM and not with the constant baseline model - most notably, over oceans in longitude bands between 30 and 45° latitude North and South. Figure 5.6 shows trajectory plots a location off the coast of Japan. Previous work reports that this region is mean-dominated by advective T' and variance-dominated by advective and diabatic T' , which agrees with the large increases in advective T' and large decreases in diabatic T' during the last 1 to 2 days. In 2014, the

LSTM model correctly predicts this behavior even though the TX1day event trajectory is just 1 timestep longer than the forecasting window, while in 2010 and 2012 it incorrectly predicts a decreasing diabatic T' trend - this is otherwise generally correct, as seen for the remaining years. In fact, the model predicts 2°C advective T' gain, approximately constant zero adiabatic T' and 0.5°C diabatic T' loss when given an input of zeros. This suggests that this is the most prevalent behavior of TX1day events with short T' trajectories, since the model most likely outputs the average for timeseries when given un-informative histories as input.

Overall, the value of the LSTM model is limited as its poor ability to correctly reproduce typical dynamics means we are unable to trust its behavior. Although outperformed over tropical oceans, it is often superior in performance against the constant baseline and this should be encouraging, indicating that in many regions, at least one of the contributor is predictable from the trajectory history. The model also generalized from the training regions to the validation regions well, as the areas of higher MSE loss do not seem to collocate with the validation locations. Then, particularly over oceans, this work suggests that forecasting future tendencies of TX1day trajectories is possible but necessitates a more complex model with more careful treatment of the inputs. We discuss potential avenues for improvement in [Chapter 6.2](#).

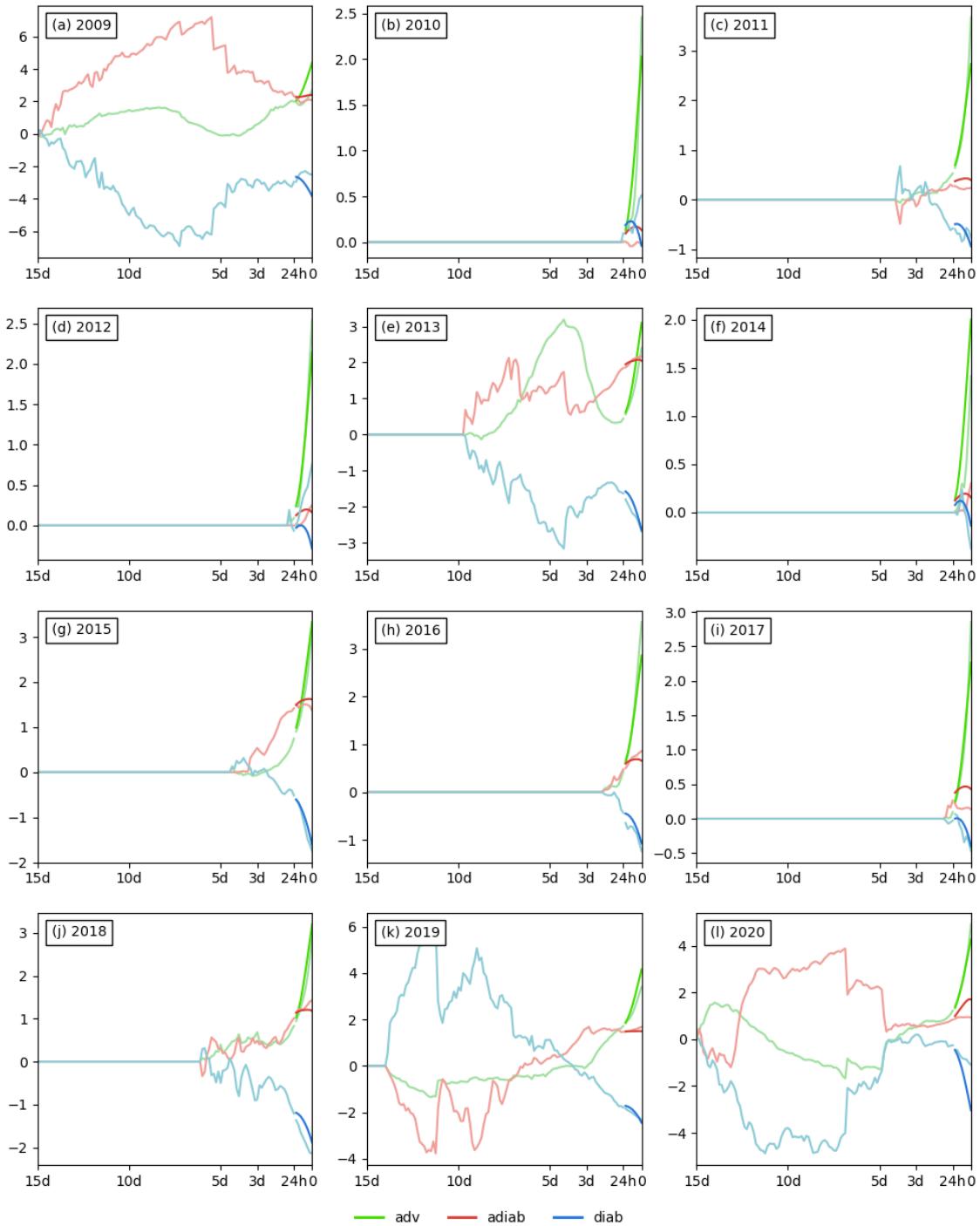


Figure 5.6 – Parcel trajectories and LSTM model predictions for location 38N 115E for the last 9 years of test data (2012-2020). The darker, more vibrant colors correspond to the predictions; the lighter colors correspond to the truth.

Chapter 6

Conclusion

6.1 Summary

Hot extremes present a threat to human health, the economy and sensitive ecosystems. Although the variability in yearly hot extreme magnitudes has been studied and regionally linked to various physical mechanisms, a global view on these events lacks. This thesis bridges this gap by exploring the spatial variability observed in the year-to-year variability of TX1day events and link it to the year-to-year variability in the contribution from advective, adiabatic and diabatic heat-generating mechanisms. By considering the Lagrangian approach proposed by [Röhlisberger and Papritz \(2023\)](#), its mean-state characterizations are extended to a comprehensive second-moment analysis.

Global patterns in TX1day T' variance are found to be similar to TX1day T' magnitudes, with predominantly larger variance over land and increasing with increasing latitude. There is however more defined land-ocean contrasts, particularly on the Western coasts of continents between 25° and 60° North and 15° and 60° South with up to 16°C² differences in variance in just 100km. No physical process is found to dominate TX1day event variability and the processes dominating the mean exhibit again similar patterns to those in the variance. Only few local regions - including the high-variability Western coasts - have a single contributor dominating the variance and these generally agree with the mean behavior. However, dominance in the mean by a contributor does not imply dominance in the variance only by this contributor, with more complete contributions from all three processes globally. Most significantly, this analysis finds that diabatic T' variance contributions are often large, despite negative contribution in the mean. This suggests that diabatic processes through diverse land-atmosphere feedbacks and cloud-cover are important in year-to-year variability of hot extremes. This supports existing literature ([Wehrli et al., 2019](#); [Miralles et al., 2014](#); [Schumacher et al., 2019](#)), although finding a larger influence in the mid- to high-latitudes than [Wehrli et al. \(2022\)](#).

Furthermore, the large and mostly negative covariances in [5.1](#) are examined through PCA, finding that most of the globe has the variability of the three contributors constrained to only one or two axes of variability. This implies that, particularly over oceans, complex mechanisms suggested by the variance decomposition and supported in the literature can be more simply characterized. In regions with variance that remains poorly explained by two axes, contributors then act more independently which may suggest the influence of a larger diversity of physical mechanisms along parcel trajectories. These areas are found in

high latitudes, such as over Northern Russia and Canada, and further suggests that Arctic amplification (Cohen et al., 2014) may be responsible for this complexity.

The approach for the analysis of parcel trajectories was limited due to poor predictive performance in many locations. Some tendencies were learned and thus out-performed the constant baseline everywhere except in tropical regions where small temperature anomalies were exaggerated by the model. We have however shown that the patterns in advective, adiabatic and diabatic temperature anomaly generation may be predicted from their history with a relative simple model, encouraging for future work. In addition, over oceans and many land regions we observe that adiabatic heating is minimal during the final 24h, suggesting that hot extreme primarily descend to the surface earlier than a day before, thus leading to contributions from advective and diabatic processes more likely.

6.2 Limitations and future work

Although providing novel insights into the global variability of physical processes contributing to hot extreme development, the methods proposed have several limitations. First of all, the variance decomposition only explores the year-to-year variability of the cumulative total T' contribution. A total estimate is meaningful, but it cannot easily make statements about the underlying dynamical structures. Nonetheless, together with PCA, these methods provide a unique perspective into the interactions involved in hot extreme development on a global scale.

This thesis was then largely successful in analyzing the patterns in the final cumulative trajectory data. However, many questions considering the timeseries history of TX1day event trajectories remain open. The approach to explore hot extreme dynamics with the use of DL for the modelling of final-day evolution was unsuccessful due to the model's poor performance. This work has however shown that some dependencies between contributors may be learned by even a simple model and that they are useful for forecasting next-day tendencies of each contributor. Given the large dataset considered and the diversity of timeseries behaviors, we suggest that future work should consider the following two changes. First, a revision of the data pre-processing: scaling by the median and inter-quartile range - although effective in dampening the influence of outlying values - did not transform each contributor to comparable ranges, which likely heavily biased the model to focus on limiting the loss due to the contributor with the largest scaled magnitude. Second, a more complex model may be required: this work focused on the contributors and therefore did not consider any additional trajectory information, such as geographical location and pressure level that are shown to play an important role. Given the large dataset, a sufficiently large neural network may learn to infer these features but to balance computational burden and model performance, these features may improve the ability of the model to capture changing dynamical mechanisms. In addition, increasing the number of LSTM hidden units may be required to obtain a richer feature-representation for each timeseries that can then be processed by the linear layer.

Another limitation of the model is that it was trained to predict only the last 24h in each trajectory. Although the model makes use of the full trajectory history and therefore learns important patterns in the evolution of each physical mechanism, it limits inference for the final 24h period and forces it to output an average behavior when trajectories are younger than the forecasting window. As was seen, some trajectories show little evolution in the final 24h - either for a single or multiple contributors - and therefore may learn to ignore

important aspects of the hot extreme development when outputting flat forecasts. Future work may then focus on training a model on diverse input windows - i.e. only using a set number of timesteps - to predict different sections of each trajectory. Exposing a model to different situations, rather than the complete history, would likely improve its ability to learn more meaningful dynamical patterns, thus improving predictive performance. Furthermore, prediction for various temporal regions would allow for a broader application of the model.

Finally, the model did not allow ‘opening the black-box’ to more precisely explore the interactions between contributors and how they relate to model predictions. This requires a model that performs well on the test-data, both in terms of low MSE loss and its ability to correctly predict general trends. Given the improvements suggested above are successful in developing a better model, future work should focus on its interpretability. Explainable AI methods have received considerable attention recently and many frameworks have been extended for multivariate timeseries data. WindowSHAP ([Nayebi et al., 2023](#)) is a post-hoc method that assigns importance to windows of input multivariate timesteps. It is particularly attractive for this problem as it is computationally more efficient than other masking based methods (see [Rojat et al., 2021](#), for a review) and provides local explanations - for an individual samples. In the context of TX1day trajectory, it then allows the identification of important time intervals and their comparison between timeseries achieving low to high final temperature anomalies. Another approach, that further reduces computational burden, is the addition of an attention mechanism to the LSTM-model presented in this thesis. Attention-mechanisms introduce learnable parameters that are optimized to give more weight to timesteps that relate to the target, thus being more discriminating than the vanilla LSTM. They have, for example, been successful in flood forecasting applications ([Ding et al., 2020](#)), where the attention weights can then be interpreted directly and compared between inputs.

Bibliography

- Adame, J. A., Notario, A., Cuevas, C. A., and Saiz-Lopez, A. (2022). Saharan air outflow variability in the 1980–2020 period. *Science of The Total Environment*, 839:156268.
- Attada, R., Dasari, H. P., Parekh, A., Chowdary, J. S., Langodan, S., Knio, O., and Hoteit, I. (2019). The role of the Indian Summer Monsoon variability on Arabian Peninsula summer climate. *Climate Dynamics*, 52(5):3389–3404.
- Barahona, D., Breen, K., Kalesse-Los, H., and Röttenbacher, J. (2023). Deep Learning Parameterization of Vertical Wind Velocity Variability via Constrained Adversarial Training. *Artificial Intelligence for the Earth Systems*, 3.
- Berg, A., Lintner, B., Findell, K., Malyshev, S., Loikith, P., and Gentine, P. (2014). Impact of Soil Moisture–Atmosphere Interactions on Surface Temperature Distribution. *Journal of Climate*, 27:7976–7993.
- Bieli, M., Pfahl, S., and Wernli, H. (2015). A Lagrangian investigation of hot and cold temperature extremes in Europe. *Quarterly Journal of the Royal Meteorological Society*, 141(686):98–108.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45:5–32.
- Brunner, L., Schaller, N., Anstey, J., Sillmann, J., and Steiner, A. K. (2018). Dependence of Present and Future European Temperature Extremes on the Location of Atmospheric Blocking. *Geophysical Research Letters*, 45(12):6311–6320.
- Byrne, M. P. (2021). Amplified warming of extreme temperatures over tropical land. *Nature Geoscience*, 14(11):837–841.
- Chelton, D. B., Esbensen, S. K., Schlax, M. G., Thum, N., Freilich, M. H., Wentz, F. J., Gentemann, C. L., McPhaden, M. J., and Schopf, P. S. (2001). Observations of Coupling between Surface Wind Stress and Sea Surface Temperature in the Eastern Tropical Pacific. *Journal of Climate*, 14(7):1479–1498.
- Cohen, J., Screen, J. A., Furtado, J. C., Barlow, M., Whittleston, D., Coumou, D., Francis, J., Dethloff, K., Entekhabi, D., Overland, J., and Jones, J. (2014). Recent Arctic amplification and extreme mid-latitude weather. *Nature Geoscience*, 7(9):627–637.
- Costa, D. F., Gomes, H. B., Silva, M. C. L., and Zhou, L. (2022). The most extreme heat waves in Amazonia happened under extreme dryness. *Climate Dynamics*, 59(1):281–295.
- Di Luca, A., de Elía, R., Bador, M., and Argüeso, D. (2020). Contribution of mean climate to hot temperature extremes for present and future climates. *Weather and Climate Extremes*, 28:100255.

- Ding, Y., Zhu, Y., Feng, J., Zhang, P., and Cheng, Z. (2020). Interpretable spatio-temporal attention LSTM model for flood forecasting. *Neurocomputing*, 403:348–359.
- Donat, M. G., Pitman, A. J., and Seneviratne, S. I. (2017). Regional warming of hot extremes accelerated by surface energy fluxes. *Geophysical Research Letters*, 44(13):7011–7019.
- Ebi, K. L., Capon, A., Berry, P., Broderick, C., Dear, R. d., Havenith, G., Honda, Y., Kovats, R. S., Ma, W., Malik, A., Morris, N. B., Nybo, L., Seneviratne, S. I., Vanos, J., and Jay, O. (2021). Hot weather and heat extremes: health risks. *The Lancet*, 398(10301):698–708.
- Entriñen, R. and Lordan, R. (2012). Impacts of extreme events on transmission and distribution systems. In *2012 IEEE Power and Energy Society General Meeting*, pages 1–10. ISSN: 1944-9925.
- Feldman, B. (2005). Relative Importance and Value. *SSRN Electronic Journal*.
- Fischer, E. M. and Schär, C. (2009). Future changes in daily summer temperature variability: driving processes and role for temperature extremes. *Climate Dynamics*, 33(7):917–935.
- Fischer, E. M., Sippel, S., and Knutti, R. (2021). Increasing probability of record-shattering climate extremes. *Nature Climate Change*, 11(8):689–695.
- Fretwell, P., Pritchard, H. D., Vaughan, D. G., Bamber, J. L., Barrand, N. E., Bell, R., Bianchi, C., Bingham, R. G., Blankenship, D. D., Casassa, G., Catania, G., Callens, D., Conway, H., Cook, A. J., Corr, H. F. J., Damaske, D., Damm, V., Ferraccioli, F., Forsberg, R., Fujita, S., Gim, Y., Gogineni, P., Griggs, J. A., Hindmarsh, R. C. A., Holmlund, P., Holt, J. W., Jacobel, R. W., Jenkins, A., Jokat, W., Jordan, T., King, E. C., Kohler, J., Krabill, W., Riger-Kusk, M., Langley, K. A., Leitchenkov, G., Leuschen, C., Luyendyk, B. P., Matsuoka, K., Mouginot, J., Nitsche, F. O., Nogi, Y., Nost, O. A., Popov, S. V., Rignot, E., Rippin, D. M., Rivera, A., Roberts, J., Ross, N., Siegert, M. J., Smith, A. M., Steinhage, D., Studinger, M., Sun, B., Tinto, B. K., Welch, B. C., Wilson, D., Young, D. A., Xiangbin, C., and Zirizzotti, A. (2013). Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. *The Cryosphere*, 7(1):375–393.
- García-León, D., Casanueva, A., Standardi, G., Burgstall, A., Flouris, A. D., and Nybo, L. (2021). Current and projected regional economic impacts of heatwaves in Europe. *Nature Communications*, 12(1):5807.
- Geirinhas, J. L., Trigo, R. M., Libonati, R., Coelho, C. A. S., and Palmeira, A. C. (2018). Climatic and synoptic characterization of heat waves in Brazil. *International Journal of Climatology*, 38(4):1760–1776.
- Gourdji, S. M., Sibley, A. M., and Lobell, D. B. (2013). Global crop exposure to critical high temperatures in the reproductive period: historical trends and future projections. *Environmental Research Letters*, 8(2):024041.
- Green, H., Bailey, J., Schwarz, L., Vanos, J., Ebi, K., and Benmarhnia, T. (2019). Impact of heat on mortality and morbidity in low and middle income countries: A review of the epidemiological evidence and considerations for future research. *Environmental Research*, 171:80–91.

- Greenacre, M. (2021). Compositional Data Analysis. *Annual Review of Statistics and Its Application*, 8(Volume 8, 2021):271–299.
- Gregory, C. H., Holbrook, N. J., Marshall, A. G., and Spillman, C. M. (2023). Atmospheric Drivers of Tasman Sea Marine Heatwaves. *Journal of Climate*, 36:5197–5214.
- Grömping, U. (2007). Estimators of Relative Importance in Linear Regression Based on Variance Decomposition. *The American Statistician*, 61(2):139–147.
- Halko, N., Martinsson, P. G., and Tropp, J. A. (2011). Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions. *SIAM Review*, 53(2):217–288.
- Ham, Y.-G., Lee, H.-J., Jo, H.-S., Lee, S.-G., Cai, W., and Rodrigues, R. R. (2021). Inter-Basin Interaction Between Variability in the South Atlantic Ocean and the El Niño/Southern Oscillation. *Geophysical Research Letters*, 48(15):e2021GL093338.
- Han, Z., Zhao, J., Leung, H., Ma, K. F., and Wang, W. (2021). A Review of Deep Learning Models for Time Series Prediction. *IEEE Sensors Journal*, 21(6):7833–7848.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellán, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.
- Hochreiter, S. and Schmidhuber, J. (1997). Long Short-term Memory. *Neural computation*, 9:1735–80.
- Holmes, C., Woollings, T., Hawkins, E., and de Vries, H. (2015). Robust Future Changes in Temperature Variability under Greenhouse Gas Forcing and the Relationship with Thermal Advection. *Journal of Climate*, 29(6):2221–2236.
- Hooker, G., Mentch, L., and Zhou, S. (2021). Unrestricted permutation forces extrapolation: variable importance requires at least one more model, or there is no free variable importance. *Statistics and Computing*, 31(6):82.
- Horton, D. E., Johnson, N. C., Singh, D., Swain, D. L., Rajaratnam, B., and Diffenbaugh, N. S. (2015). Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature*, 522(7557):465–469.
- Hu, J., Wang, X., Zhang, Y., Zhang, D., Zhang, M., and Xue, J. (2020). Time Series Prediction Method Based on Variant LSTM Recurrent Neural Network. *Neural Processing Letters*, 52(2):1485–1500.
- Huang, Y. and Cui, X. (2015). Moisture sources of an extreme precipitation event in Sichuan, China, based on the Lagrangian method. *Atmospheric Science Letters*, 16(2):177–183.
- Kapsch, M.-L., Graversen, R. G., and Tjernström, M. (2013). Springtime atmospheric energy transport and the control of Arctic summer sea-ice extent. *Nature Climate Change*, 3(8):744–748.

- Kingma, D. P. and Ba, J. (2017). Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs].
- Lenderink, G., van Ulden, A., van den Hurk, B., and van Meijgaard, E. (2007). Summer-time inter-annual temperature variability in an ensemble of regional model simulations: analysis of the surface energy budget. *Climatic Change*, 81(1):233–247.
- Lewis, S. C. and King, A. D. (2017). Evolution of mean, variance and extremes in 21st century temperatures. *Weather and Climate Extremes*, 15:1–10.
- Lin, H., Mo, R., and Vitart, F. (2022). The 2021 Western North American Heatwave and Its Subseasonal Predictions. *Geophysical Research Letters*, 49(6):e2021GL097036.
- Lindeman, R. H., Merenda, P. F., and Gold, R. Z. (1980). *Introduction to Bivariate and Multivariate Analysis*. Scott, Foresman and Company, Glenview, IL.
- Lorenz, E. N. (1956). *Empirical Orthogonal Functions and Statistical Weather Prediction*.
- Lucarini, V., Galfi, V. M., Riboldi, J., and Messori, G. (2023). Typicality of the 2021 Western North America summer heatwave. *Environmental Research Letters*, 18(1):015004.
- Lupo, A. R. (2021). Atmospheric blocking events: a review. *Annals of the New York Academy of Sciences*, 1504(1):5–24.
- Ma, W., Wang, H., Chen, G., Qian, Y., Baxter, I., Huo, Y., and Seefeldt, M. W. (2024). Wintertime extreme warming events in the high Arctic: characteristics, drivers, trends, and the role of atmospheric rivers. *Atmospheric Chemistry and Physics*, 24(7):4451–4472.
- Mccoll, K., He, Q., Lu, H., and Entekhabi, D. (2019). Short-Term and Long-Term Surface Soil Moisture Memory Time Scales Are Spatially Anticorrelated at Global Scales. *Journal of Hydrometeorology*, 20(6):1165–1182.
- McKinnon, K. A., Rhines, A., Tingley, M. P., and Huybers, P. (2016). The changing shape of Northern Hemisphere summer temperature distributions. *Journal of Geophysical Research: Atmospheres*, 121(15):8849–8868.
- Medina-Ramón, M. and Schwartz, J. (2007). Temperature, temperature extremes, and mortality: a study of acclimatisation and effect modification in 50 US cities. *Occupational and Environmental Medicine*, 64(12):827–833.
- Miralles, D. G., Teuling, A. J., van Heerwaarden, C. C., and Vilà-Guerau de Arellano, J. (2014). Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature Geoscience*, 7(5):345–349.
- Mortin, J., Svensson, G., Graversen, R. G., Kapsch, M.-L., Stroeve, J. C., and Boisvert, L. N. (2016). Melt onset over Arctic sea ice controlled by atmospheric moisture transport. *Geophysical Research Letters*, 43(12):6636–6642.
- Nayebi, A., Tipirneni, S., Reddy, C. K., Foreman, B., and Subbian, V. (2023). Window-SHAP: An efficient framework for explaining time-series classifiers based on Shapley values. *Journal of Biomedical Informatics*, 144:104438.
- Noh, S.-H. (2021). Analysis of Gradient Vanishing of RNNs and Performance Comparison. *Information*, 12(11):442.

- Oudin Åström, D., Bertil, F., and Joacim, R. (2011). Heat wave impact on morbidity and mortality in the elderly population: A review of recent studies. *Maturitas*, 69(2):99–105.
- Papritz, L., Murto, S., Röthlisberger, M., Caballero, R., Messori, G., Svensson, G., and Wernli, H. (2023). The Role of Local and Remote Processes for Wintertime Surface Energy Budget Extremes over Arctic Sea Ice. *Journal of Climate*, 36:1–40.
- Papritz, L. and Röthlisberger, M. (2023). A Novel Temperature Anomaly Source Diagnostic: Method and Application to the 2021 Heatwave in the Pacific Northwest. *Geophysical Research Letters*, 50(23):e2023GL105641.
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572.
- Perera, A. T. D., Nik, V. M., Chen, D., Scartezzini, J.-L., and Hong, T. (2020). Quantifying the impacts of climate change and extreme climate events on energy systems. *Nature Energy*, pages 150–159.
- Pezza, A. B., van Rensch, P., and Cai, W. (2012). Severe heat waves in Southern Australia: synoptic climatology and large scale connections. *Climate Dynamics*, 38(1):209–224.
- Pfahl, S. and Wernli, H. (2012). Quantifying the relevance of atmospheric blocking for co-located temperature extremes in the Northern Hemisphere on (sub-)daily time scales. *Geophysical Research Letters*, 39(12):12807.
- Quesada, B., Vautard, R., Yiou, P., Hirschi, M., and Seneviratne, S. I. (2012). Asymmetric European summer heat predictability from wet and dry southern winters and springs. *Nature Climate Change*, 2(10):736–741.
- Quinting, J. F. and Reeder, M. J. (2017). Southeastern Australian Heat Waves from a Trajectory Viewpoint. *Monthly Weather Review*, 145(10):4109–4125.
- Razavi, S. and Gupta, H. V. (2015). What do we mean by sensitivity analysis? The need for comprehensive characterization of “global” sensitivity in Earth and Environmental systems models. *Water Resources Research*, 51(5):3070–3092.
- Robinson, P. J. (2001). On the Definition of a Heat Wave. *Journal of Applied Meteorology (1988-2005)*, 40(4):762–775. Publisher: American Meteorological Society.
- Rojat, T., Puget, R., Filliat, D., Ser, J., Gelin, R., and D'iaz-Rodríguez, N. (2021). Explainable Artificial Intelligence (XAI) on TimeSeries Data: A Survey. *ArXiv*.
- Röthlisberger, M. and Papritz, L. (2023). Quantifying the physical processes leading to atmospheric hot extremes at a global scale. *Nature Geoscience*, 16:1–7.
- Salerno, F., Guyennon, N., Yang, K., Shaw, T. E., Lin, C., Colombo, N., Romano, E., Gruber, S., Bolch, T., Alessandri, A., Cristofanelli, P., Putero, D., Diolaiuti, G., Tattari, G., Verza, G., Thakuri, S., Balsamo, G., Miles, E. S., and Pellicciotti, F. (2023). Local cooling and drying induced by Himalayan glaciers under global warming. *Nature Geoscience*, 16(12):1120–1127.
- Schumacher, D. L., Keune, J., van Heerwaarden, C. C., Vilà-Guerau de Arellano, J., Teuling, A. J., and Miralles, D. G. (2019). Amplification of mega-heatwaves through heat torrents fuelled by upwind drought. *Nature Geoscience*, 12(9):712–717.

- Simolo, C. and Corti, S. (2022). Quantifying the role of variability in future intensification of heat extremes. *Nature Communications*, 13(1):7930.
- Slater, R., Freychet, N., and Hegerl, G. (2021). Substantial changes in the probability of future annual temperature extremes. *Atmospheric Science Letters*, 22(11):e1061.
- Sousa, P. M., Trigo, R. M., Barriopedro, D., Soares, P. M. M., and Santos, J. A. (2018). European temperature responses to blocking and ridge regional patterns. *Climate Dynamics*, 50(1):457–477.
- Sprenger, M. and Wernli, H. (2015). The LAGRANTO Lagrangian analysis tool – version 2.0. *Geoscientific Model Development*, 8(8):2569–2586.
- Srivastava, A. and S, A. (2022). Weather Prediction Using LSTM Neural Networks. In *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*, pages 1–4.
- Stohl, A. and James, P. (2004). A Lagrangian Analysis of the Atmospheric Branch of the Global Water Cycle. Part I: Method Description, Validation, and Demonstration for the August 2002 Flooding in Central Europe. *Journal of Hydrometeorology*, 5(4):656–678.
- Suarez-Gutierrez, L., Müller, W. A., Li, C., and Marotzke, J. (2020). Dynamical and thermodynamical drivers of variability in European summer heat extremes. *Climate Dynamics*, 54(9):4351–4366.
- Teixeira, E. I., Fischer, G., van Velthuizen, H., Walter, C., and Ewert, F. (2013). Global hot-spots of heat stress on agricultural crops due to climate change. *Agricultural and Forest Meteorology*, 170:206–215.
- Thompson, V., Kennedy-Asser, A. T., Vosper, E., Lo, Y. T. E., Huntingford, C., Andrews, O., Collins, M., Hegerl, G. C., and Mitchell, D. (2022). The 2021 western North America heat wave among the most extreme events ever recorded globally. *Science Advances*, 8(18):eabm6860.
- Vogel, E., Donat, M. G., Alexander, L. V., Meinshausen, M., Ray, D. K., Karoly, D., Meinshausen, N., and Frieler, K. (2019). The effects of climate extremes on global agricultural yields. *Environmental Research Letters*, 14(5):054010.
- Wake, B. (2014). Land–sea contrast. *Nature Climate Change*, 4(5):326–326.
- Wang, P., Zhang, Q., Yang, Y., and Tang, J. (2019). The Sensitivity to Initial Soil Moisture for Three Severe Cases of Heat Waves Over Eastern China. *Frontiers in Environmental Science*, 7.
- Wang, T., Yin, S., Hua, W., Wang, H., Luo, F., Miao, J., and Fu, Y. (2023). Decadal variability of extreme high temperature in mid- and high-latitude Asia and its associated North Atlantic air–sea interaction. *Climate Dynamics*, 61(9):4587–4601.
- Watterson, I., McGregor, J., and Nguyen, K. (2008). Changes in extreme temperatures of Australasian summer simulated by CCAM under global warming, and the roles of winds and land-sea contrasts. *Australian meteorological magazine*, 57:195–212.
- Wehrli, K., Guillod, B. P., Hauser, M., Leclair, M., and Seneviratne, S. I. (2019). Identifying Key Driving Processes of Major Recent Heat Waves. *Journal of Geophysical Research: Atmospheres*, 124(22):11746–11765.

- Wehrli, K., Luo, F., Hauser, M., Shiogama, H., Tokuda, D., Kim, H., Coumou, D., May, W., Le Sager, P., Selten, F., Martius, O., Vautard, R., and Seneviratne, S. I. (2022). The ExtremeX global climate model experiment: investigating thermodynamic and dynamic processes contributing to weather and climate extremes. *Earth System Dynamics*, 13(3):1167–1196.
- Whan, K., Zscheischler, J., Orth, R., Shongwe, M., Rahimi, M., Asare, E. O., and Seneviratne, S. I. (2015). Impact of soil moisture on extreme maximum temperatures in Europe. *Weather and Climate Extremes*, 9:57–67.
- White, R. H., Anderson, S., Booth, J. F., Braich, G., Draeger, C., Fei, C., Harley, C. D. G., Henderson, S. B., Jakob, M., Lau, C.-A., Mareshet Admasu, L., Narinesingh, V., Rodell, C., Roocroft, E., Weinberger, K. R., and West, G. (2023). The unprecedented Pacific Northwest heatwave of June 2021. *Nature Communications*, 14(1):727.
- Wille, J. D., Alexander, S. P., Amory, C., Baiman, R., Barthelemy, L., Bergstrom, D. M., Berne, A., Binder, H., Blanchet, J., Bozkurt, D., Bracegirdle, T. J., Casado, M., Choi, T., Clem, K. R., Codron, F., Datta, R., Battista, S. D., Favier, V., Francis, D., Fraser, A. D., Fourre, E., Garreaud, R. D., Genthon, C., Gorodetskaya, I. V., Gonzalez-Herrero, S., Heinrich, V. J., Hubert, G., Joos, H., Kim, S.-J., King, J. C., Kittel, C., Landais, A., Lazzara, M., Leonard, G. H., Lieser, J. L., MacLennan, M., Mikolajczyk, D., Neff, P., Ollivier, I., Picard, G., Pohl, B., Ralph, F. M., Rowe, P., Schlosser, E., Shields, C. A., Smith, I. J., Sprenger, M., Trusel, L., Udy, D., Vance, T., Vignon, E., Walker, C., Wever, N., and Zou, X. (2024). The Extraordinary March 2022 East Antarctica “Heat” Wave. Part II: Impacts on the Antarctic Ice Sheet. *Journal of Climate*, 37(3):779–799.
- Zhou, Y., Chang, F.-J., Chang, L.-C., Kao, I.-F., and Wang, Y.-S. (2019). Explore a deep learning multi-output neural network for regional multi-step-ahead air quality forecasts. *Journal of Cleaner Production*, 209:134–145.

Appendix A

Supplementary material on timeseries data

A.1 Further examples of trajectory data

Find below the trajectory timeseries visualisation plots - as figure 3.2 - for the locations in figure 3.1. Each figure contains plots (a-d) of TX1day T' decompositions for years 1980 to 2020. The x-axis is logarithmic to better represent the recent history. The trajectory yielding the largest final T' among all years is colored red, the time-mean of all trajectories in black and the inter-quantile range is highlighted in light blue. Plot (e) shows the (auto-)correlation between the final T' and the T' at different lags for the event and each contributor. Plot (f) shows the (cross-)correlation between the final event T' and the T' at different lags for each contributor. They are computed for all samples that have genesis time earlier than the considered time-lag.

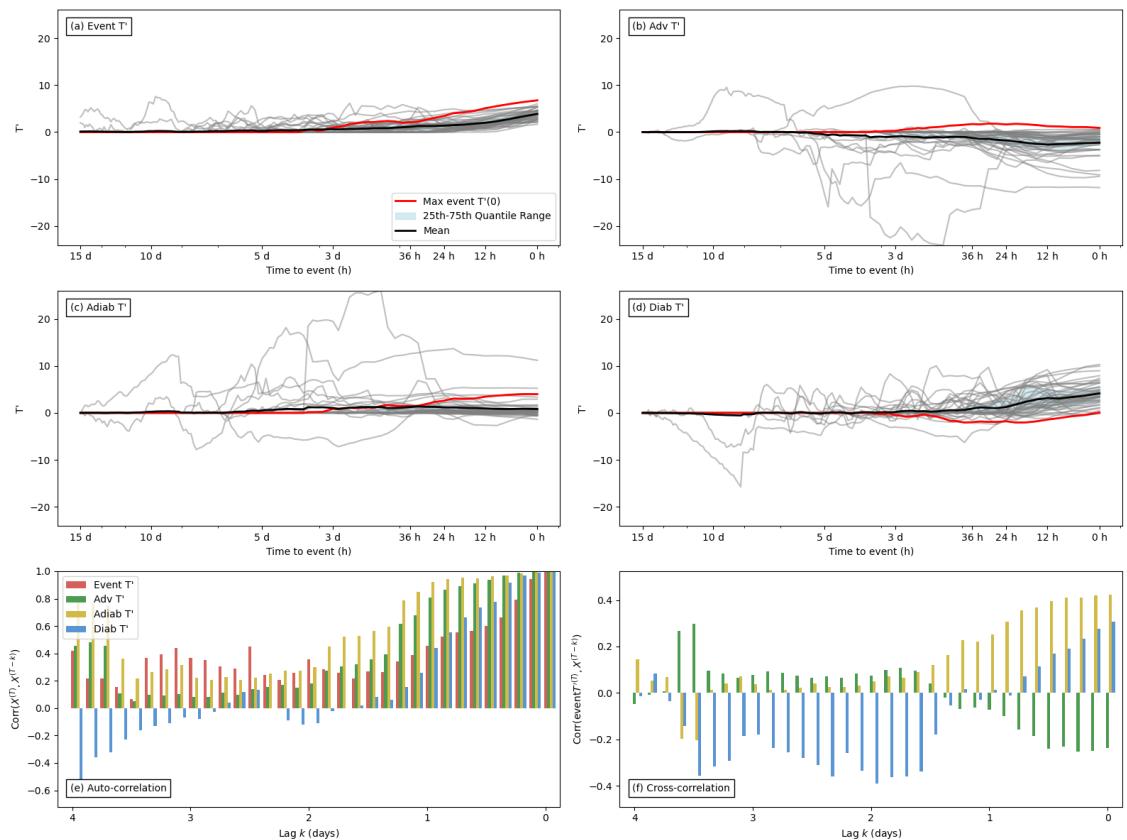


Figure A.1 – Gridpoint 28.5N 77E located in the vicinity of New Delhi (India)

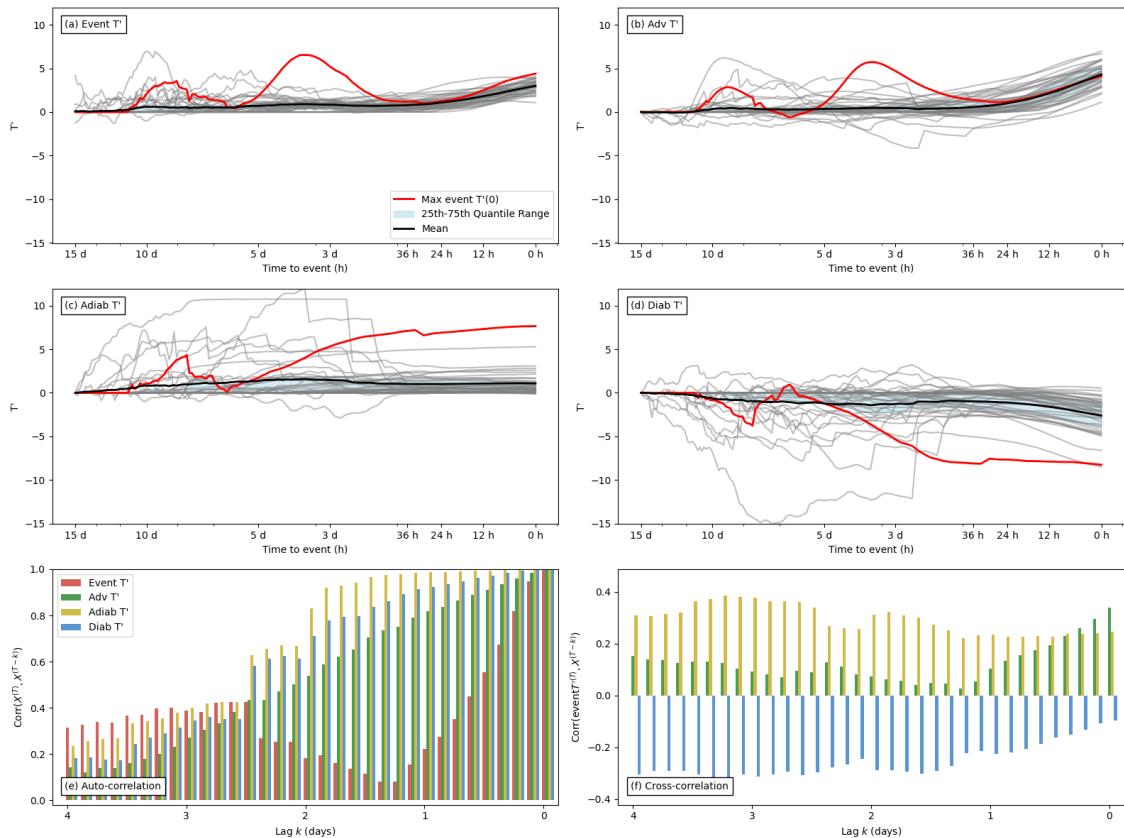


Figure A.2 – Gridpoint 45N 30W located in the middle of the Atlantic ocean.

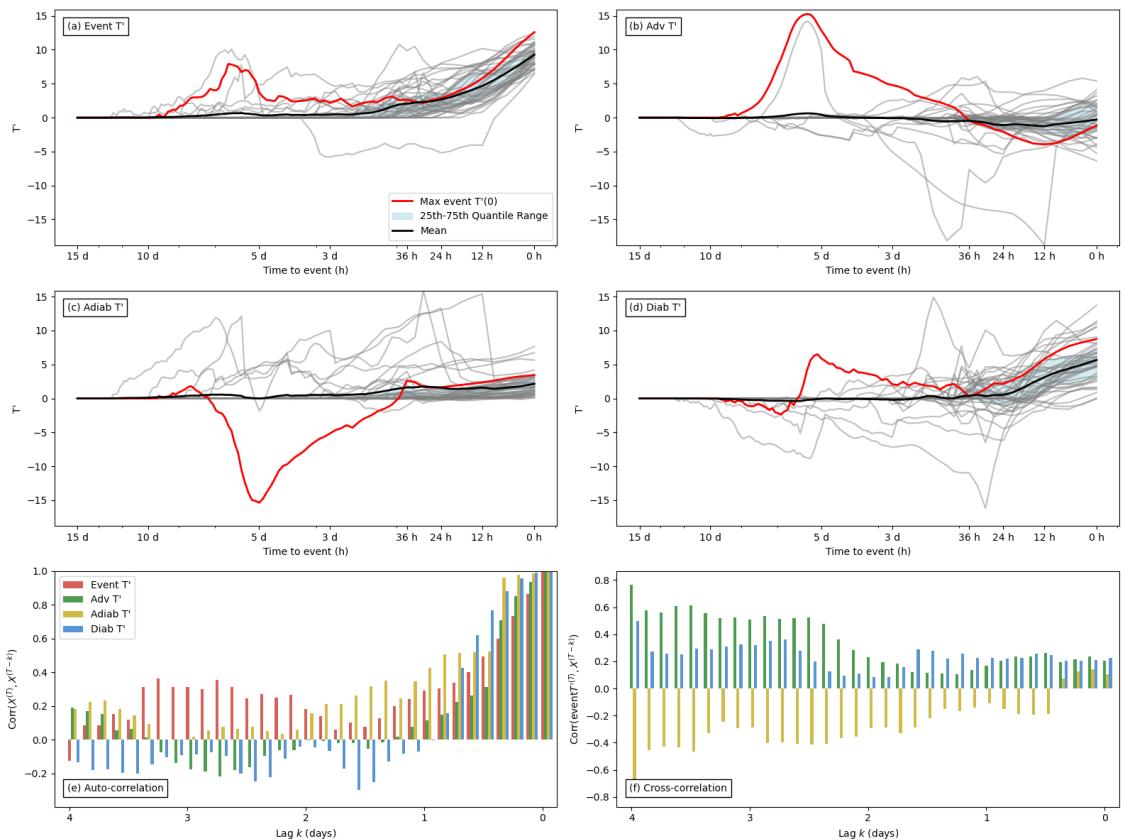


Figure A.3 – Gridpoint 32S 116E located in the vicinity of Perth (Australia).

A.2 Differenced timeseries data

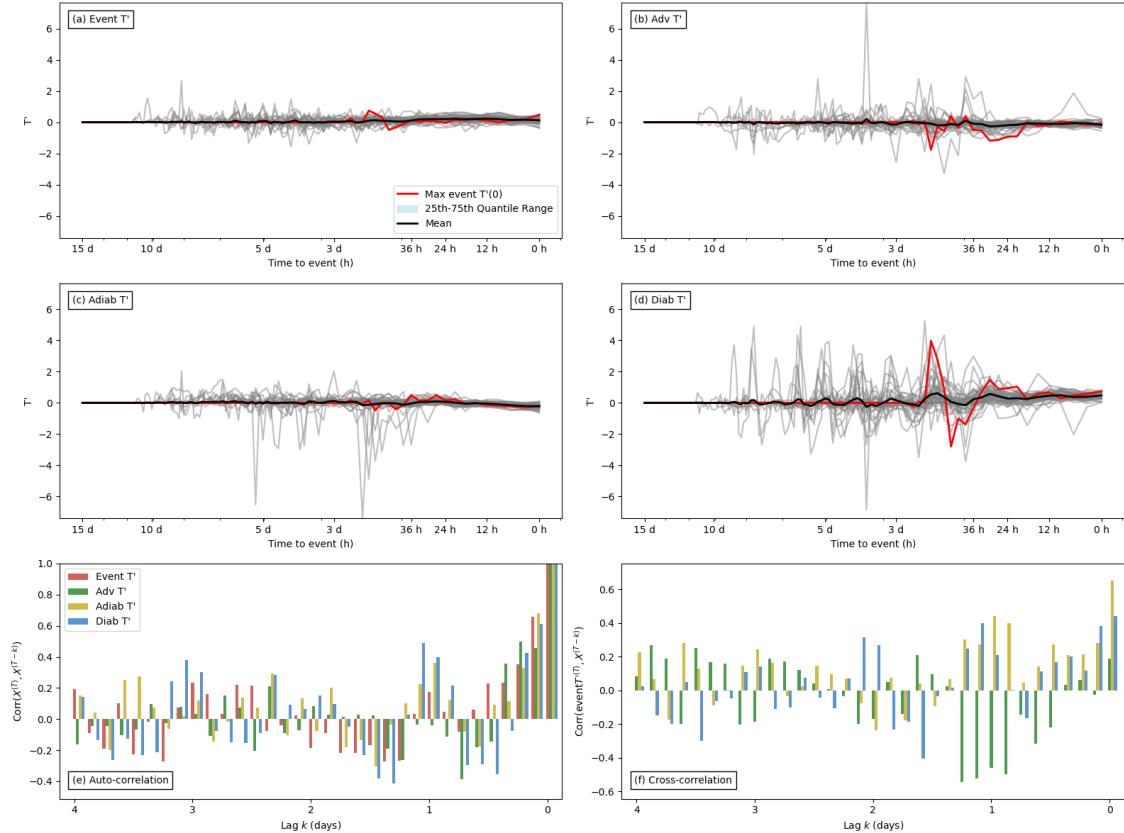


Figure A.4 – Differenced timeseries visualisation plots for location 16S 48W and years between 1980 and 2020. Explanations of each graph are the same as for Appendix A.1.

Appendix B

Further background on variance

Variance is a measure of dispersion - or spread - of a distribution and is defined as the expected squared deviation from the mean. Given random variable $X \sim \mathcal{F}$, the variance of X is given by $\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$. Conversely to the expectation, variance is a non-linear operator. Given two random variables $X \sim \mathcal{F}_X$ and $Y \sim \mathcal{F}_Y$:

$$\begin{aligned}\mathbb{V}[X + Y] &= \mathbb{E}[(X + Y - \mathbb{E}[X + Y])^2] \\ &= \mathbb{E}[(X - \mathbb{E}[X])^2] + \mathbb{E}[(Y - \mathbb{E}[Y])^2] + 2\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \quad (\text{B.1}) \\ &= \mathbb{V}[X] + \mathbb{V}[Y] + 2\text{Cov}[X, Y]\end{aligned}$$

by linearity of the expectation. The last term is called the covariance between X and Y and is a measure of linear association between two random variables. Trivially, the covariance between X and X is the variance of X .

These quantities serve as the basis for most statistical models. Normalizing covariance by the square-root of the variance of X and Y yields the Pearson correlation coefficient $\rho_{X,Y}$ that is often preferred as a summary measure of linear association since it can be compared between random variables with different scales. In regression, other common measure of association between the response Y and one or more predictors $X_{1:k}$ includes the coefficient of determination R^2 measuring the proportion of the variance of Y that is predicted by $X_{1:k}$.

In practice, one often assumed that observed data samples $\{(x_1, y_1), \dots, (x_n, y_n)\}$ are independent draws from some joint distribution. Then, assuming that the mean and covariance of the distribution are finite, the following provide the sample estimates of the variance of the marginal distributions $\hat{\mathbb{V}}$ and the covariance of the joint distribution $\hat{\text{Cov}}$:

$$\begin{aligned}\hat{\mathbb{V}}[\{x_{1:n}\}] &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \\ \hat{\text{Cov}}[\{x_{1:n}\}, \{y_{1:n}\}] &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (\text{B.2})\end{aligned}$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ is the sample mean of X . Note the normalizing factor $n - 1$ that ensures the estimate is unbiased - in repeated sampling, the average estimates equal the true quantity.

Appendix C

Grid-search results

Although the minimal loss was achieved by the model with no L2-regularization, it is expected that training on more than 30 times the number of samples will require some regularization. In addition, the difference in performance between models with 88 and 64 is small and thus prefer a less complex model. The final model then has parameters 64 hidden units and L2-regularization parameter $1e - 06$.

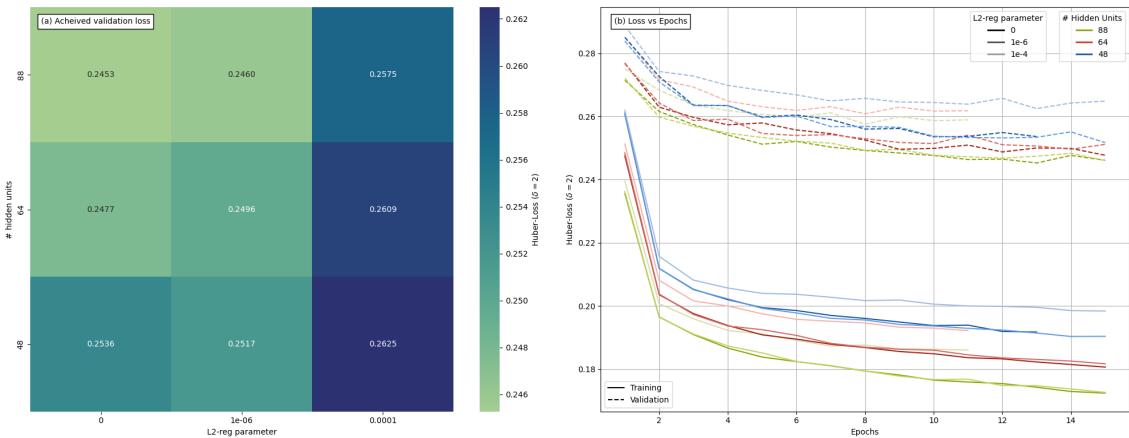


Figure C.1 – Parameter validation results showing (a) the minimally achieved loss for each model instance and (b) the training and validation loss curves over epochs for each model instance.

Declaration of originality

The signed declaration of originality is a component of every written paper or thesis authored during the course of studies. In consultation with the supervisor, one of the following three options must be selected:

- I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used no generative artificial intelligence technologies¹.
- I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used and cited generative artificial intelligence technologies².
- I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used generative artificial intelligence technologies³. In consultation with the supervisor, I did not cite them.

Title of paper or thesis:

Investigating year-to-year variability of hot extremes and contributions from heat-generating mechanisms

Authored by:

If the work was compiled in a group, the names of all authors are required.

Last name(s):

Froelich

First name(s):

Marco

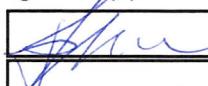
With my signature I confirm the following:

- I have adhered to the rules set out in the Citation Guide.
- I have documented all methods, data and processes truthfully and fully.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for originality.

Place, date

Zurich, 16.09.2024

Signature(s)

If the work was compiled in a group, the names of all authors are required. Through their signatures they vouch jointly for the entire content of the written work.

¹ E.g. ChatGPT, DALL E 2, Google Bard

² E.g. ChatGPT, DALL E 2, Google Bard

³ E.g. ChatGPT, DALL E 2, Google Bard