LETTER • OPEN ACCESS

Revealing trends in extreme heatwave intensity: applying the UNSEEN approach to Nordic countries

View the article online for updates and enhancements.

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

Revealing trends in extreme heatwave intensity: applying the UNSEEN approach to Nordic countries

S Berghald[1,2] , S Mayer[1,*] and P Bohlinger[3]

[1] NORCE Norwegian Research Centre, Climate, Bjerknes Centre for Climate Research, Bergen, Norway
[2] UFR PhITEM, Université Grenoble Alpes, Saint-Martin-d'Hères, France
[3] Norwegian Meteorological Institute, Bergen, Norway
[*] Author to whom any correspondence should be addressed.

E-mail: stma@norceresearch.no

Abstract

The increase in heatwave intensity, causing heat stress and crop failures in many regions is a concerning impact of global climate change. In northern Europe, significant interannual variability previously prevented robust assessments of trends in heat extremes. However, with a large-ensemble seasonal hindcasts and archived forecasts dataset covering 1981–2022 multiple realisations of weather patterns can be pooled and assessed. What are recent trends of extreme temperatures? Has the risk for a 100-year heatwave event increased in Northern Europe? We apply the UNSEEN (UNprecedented Simulated Extremes using ENsembles) approach to assess the credibility of the model ensemble and use non-stationary extreme value analysis to quantify recent trends in extreme 3-day heatwaves in late spring and early summer (May to July). We find significant non-stationarity and positive trends in annual maximum heatwave intensity. We also show that heatwave volatility, i.e. the risk of clearly outstanding heatwaves, is highest in central Scandinavia.

1. Introduction

The increase of temperature extremes is often perceived as one of the major and clearest signs of climate change (Yiou *et al* 2020, IPCC 2023, Zachariah *et al* 2023). However, this increase has not received much attention for climate impact assessment in Northern Europe. The European Environmental Agency mainly names winter storms and flood-related hazards as major areas of concern in this region, while decreases in energy demand for heating and increases in summer tourism are named as positive impacts of climatic changes (EEA 2017). Also in agriculture, temperature extremes often accompanied with heatwaves and droughts were not perceived as important risks in the past. In 2016, a report to the Norwegian government focused mainly on flooding and wet soils, while heat and drought were just perceived as potentially problematic in some areas towards the end of the century (LMD 2016). Several surveys among Norwegian farmers conducted between 2008 and 2012 also showed that while roughly 50% of the

farmers perceived drought as a potential risk for them within the following 10 years, more than 80% thought the same for heavy autumn rain and wet soils (Aasprang 2013). Higher temperatures contribute to extend the thermal growing season on one hand (Ruosteenoja *et al* 2020), however, especially the increased frequency and duration of drought and heatwaves can lead to reduced seed germination and plant growth and hence reduce crop volume and quality (Brás *et al* 2021).

The hot and dry summer in Northern Europe in 2018 came as a surprise to many, and especially to practitioners in the Scandinavian agricultural sector (Haugen *et al* 2019, Johnsson *et al* 2019, Landbruksdirektoratet 2020). Much hotter and drier conditions dominated the weather in Northern Europe from Mid–April well into autumn. May 2018 was the warmest May on record for all of the Scandinavian and Baltic countries, with temperature anomalies exceeding $+5\,^\circ\mathrm{C}$ in some areas (Blunden and Arndt 2019, Skaland *et al* 2019). More records were broken during the summer, making

2018 an exceptionally warm and dry year in this region (Blunden and Arndt 2019). This had devastating impacts especially for agriculture, with economic losses for agriculture alone of around 3 billion NOK (Norway), 4–6 billion DKK (Denmark) and 6 billion SEK (Sweden) (Johnsson *et al* 2019). For effective climate adaptation, it is therefore crucial to have a good estimate of whether the probability of such extreme events is changing (Hov 2013, Glotter and Elliott 2016, Devot *et al* 2023).

In this study, we investigate recent trends of extreme heatwave intensity in Northern Europe and if the risk for a 100-year heatwave event has increased since 1981. Some of the challenges related to high variability and short record length in observational datasets can be overcome by using large ensembles of physics-based climate models (e.g. Slater *et al* 2021). An approach that has become quite popular recently is UNSEEN (UNprecedented Simulated Extremes using ENsembles; Thompson *et al* 2017). By pooling members of seasonal predictions or climate models, UNSEEN creates a set of alternative but equally plausible realisations of past weather patterns, provided that the 'model climate' is representative of real-world climate. This approach has first been conceived by van den Brink *et al* (2005) and was standardised by Thompson *et al* (2017). Building on that, Kelder *et al* (2022) developed a standardised, easy-to-implement protocol for adapting UNSEEN and evaluating model ensembles for plausibility. In previous studies, UNSEEN has been used to assess the probability of unprecedented rainfall in the UK (Thompson *et al* 2017), heatwaves in China (Thompson *et al* 2019), exceedance probabilities of critical agrometeorological thresholds in China and the US (Kent 2019, Coughlan de Perez *et al* 2023), probability of early-spring thawing events in Siberia (Kelder *et al* 2022), critical storm-surge events in the Netherlands (van den Brink *et al* 2005), trends in extreme precipitation in Norway and Svalbard (Kelder *et al* 2020, Müller *et al* 2022), and more. We build upon and apply the UNSEEN trends approach (an extension of UNSEEN based on non-stationary extreme value theory; Kelder *et al* 2020) to detect and quantify non-stationarities in heat extremes in Northern Europe. We use the latest ECMWF seasonal prediction system SEAS5 (Johnson *et al* 2019). High-resolution large ensemble hindcasts and archived forecasts are available for the recent past (1981–2022). In this study, we identify a heatwave by assessing the 3-day averaged daily maximum temperature (TX3d) (Kew *et al* 2019) and calculate trends in the maximum TX3d in the May–July period of each year, when agriculture in Northern Europe is most vulnerable to extreme heat and drought (Johnsson *et al* 2019).
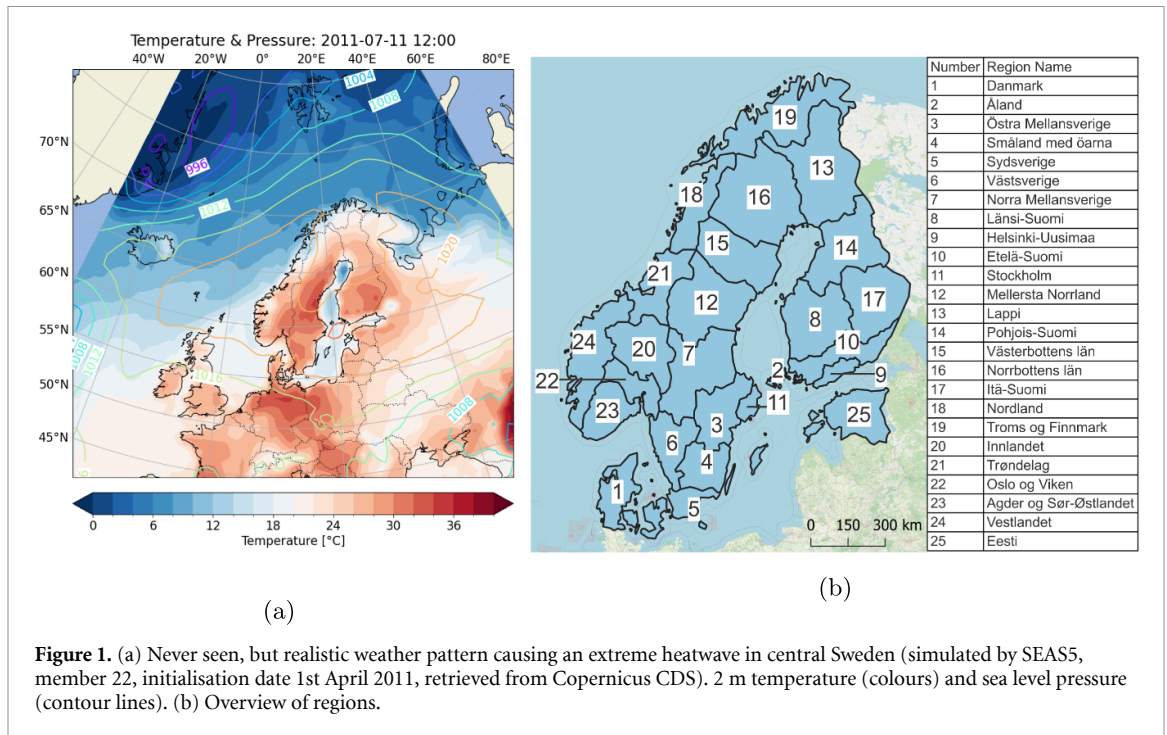
2. Materials and methods

To ensure credibility and comparability, we follow the UNSEEN workflow as presented by Kelder *et al* (2022). First, we define the event to be analysed and describe the datasets used for the analysis. Then, for each of the studied regions, we statistically evaluate the realism of the modelled ensemble time series compared to observations. Finally, we compute the trends in 1-in-100-year return values and the heatwave volatility of each region.

2.1. Heat wave definition

In this study we focus on May–July maximum 3-day averaged daily maximum temperature (TX3d). While a multi-parameter drought index accounting for both intensity and duration of a drought period might be more relevant to agriculture, we keep the event definition as simple as possible. This choice makes the analysis less sensitive to possible model shortcomings, especially in the simulation of humidity and precipitation. It also makes our results more comparable to previous studies that assessed trends in seasonal or annual daily maximum temperature (TXx; e.g. Sulikowska and Wypych 2021) or also TX3d (World Weather Attribution 2018, Philip *et al* 2023). For the warm months (March–September) in Europe, several studies found strong negative correlations between temperature and precipitation anomalies (Trenberth and Shea 2005) as well as between anomalies in temperature and cloudiness (Tomczyk *et al* 2017, Räisänen 2019). Furthermore, Schaller *et al* (2018) found a significant correlation between summer heatwave magnitudes and the number of days influenced by atmospheric blocking in Northern Europe and Western Russia. This gives us confidence to assume that extreme heat in Northern Europe is generally connected to very dry conditions, making TX3d a useful proxy for both heatwaves and drought conditions. In Northern Europe the growing season does not start before late April (Domínguez-Castro *et al* 2020) and the early onset of abnormally warm and dry conditions in May was identified as one of the main reasons for large-scale crop failure in Scandinavia in 2018 (Johnsson *et al* 2019).

2.2. Data

To assess 'alternative realities', i.e. different, but equally plausible versions of past weather, we use SEAS5, ECMWF's fifth generation seasonal forecast system (Johnson *et al* 2019). SEAS5 is a long-range physics-based forecast system initialized every month and integrated 7 months into the future. The dataset consists of hindcasts (1981–2016) containing 25 ensemble members, and archived forecasts (2017-present) containing 51 ensemble members. To

Figure 1. (a) Never seen, but realistic weather pattern causing an extreme heatwave in central Sweden (simulated by SEAS5, member 22, initialisation date 1st April 2011, retrieved from Copernicus CDS). 2 m temperature (colours) and sea level pressure (contour lines). (b) Overview of regions.

ensure consistency, we use 25 ensemble members for the whole period of 1981–2022. Despite using different initialization methods, major inhomogeneity between hindcasts and archived forecasts is not expected (Kelder *et al* 2022). By pooling lead times and ensemble members we can create a dataset containing 100 alternative realisations of possible weather outcomes for May–July of each year of the 42-year period (4 lead times, 25 members per lead time). An example of a simulated, but yet never observed weather pattern causing unprecedented heat in central Sweden is shown in figure 1(a).

In order to test credibility of the model outcomes, we need to compare them against historical observations or reanalysis products. Here we use the MET Nordic long-term consistent dataset (Norwegian Meteorological Institute 2023), which is based on the 3-km Norwegian Reanalysis NORA3 (Haakenstad *et al* 2021) hindcast. NORA3 showed very close agreement with surface observations for the 99th percentile of 2-meter surface temperature across the Norwegian mainland (Haakenstad and Breivik 2022). Preprocessing and computing of region-averaged annual maximum temperatures TX3d is done using CDO and NCO climate data operators (Zender 2008, Schulzweida 2022). We compute region-averaged TX3d by weighting grid cells according to their size (decreasing with increasing latitude) and the percentage of each grid cell within a region of interest using Python xarray (Hoyer and Hamman 2017) and Python pyscissor (Ahasan 2020). This allows us to be more precise on the boundaries of a given region than the 'black and white' masking method used for example in the 'UNSEEN-open' workflow (Kelder *et al* 2022).

We loosely base our regional analyses on the Nomenclature of Territorial Units for Statistics dataset NUTS (Eurostat 2022), which we modified to get regions of comparable size (figure 1(b)). We use QGIS geographical information system (QGIS.org 2023) for modifying the shapefiles and plotting spatial data. We choose administrative borders rather than climatic regions (e.g. Norway's temperature regions; Hanssen-Bauer and Nordli 1998) or grid cell-based analysis. As it is stated in previous studies (Bachmair *et al* 2018, Angelova and Lupio 2020), this corresponds better with the needs of risk management in societies and agriculture.

2.3. Evaluation of UNSEEN

The UNSEEN approach is based on the assumption that model ensemble members can be treated as different and equally plausible versions of past weather evolution, making it possible to increase the sample size of weather events and ultimately allowing a robust statistical analysis. Three conditions must be met (Kelder *et al* 2022): i) Ensemble member independence, because if the ensemble members are correlated, the effective sample size is much smaller than the total sample size, which results in too high confidence in the results, ii) model stability, because a drift in the model towards higher lead times may alter the realism of the ensemble; and iii) model fidelity, because only if observations and model members are statistically indistinguishable, the virtual model climate can be regarded as consistent with the 'real' climate (Anderson 1997). Based on the initial idea by van den Brink *et al* (2005), Thompson *et al* (2017) and

Thompson *et al* (2019) developed a theoretical framework to test these three conditions, which was further standardised by Kelder *et al* (2020) and Kelder *et al* (2022). Here we use the 'UNSEEN' R-package presented in the latter study (https://github.com/timokelder/UNSEEN) to test these conditions individually for each region. All tests are extensively discussed in the above-mentioned studies, thus we provide a short description for better understanding. i) Independence is tested by calculating the Spearman rank correlation coefficient (ρ) for each lead time and every distinct pair of ensemble members. The resulting distribution of Spearman ρ (one value for each pair) is compared against a theoretically expected distribution. ii) Stability is tested by separately plotting probability density functions and empirical return periods for each lead time. iii) Finally, fidelity is tested by bootstrapping 10 000 proxy model time series the same length as observational record from all years and ensemble members. Mean, standard deviation, skewness, and kurtosis are then calculated of all the proxy timeseries as well as the observations and compared to each other. If the moments of the observations fall within the 95% confidence range of the bootstrapped proxy time series, observations and ensemble are deemed statistically indistinguishable. We detect a cold bias in the mean of the ensemble data as the most common issue across all regions. This bias can be corrected by adjusting the mean. More sophisticated bias correction methods might yield better results (e.g. Maraun 2016, Sippel *et al* 2016) but are out of the scope of this study. Kelder *et al* (2022) provide an overview over potential solutions when issues with ensemble member independence, stability or fidelity are detected.

2.4. Statistical analysis

We use the obtained ensemble for two types of statistical analysis:

i) In order to show how prone a given region is to clearly outstanding extremes, we compute volatility as the mean anomaly of the 42 highest TX3d (corresponding to the top percent) of the detrended seasonal forecast dataset.

ii) We apply well-established extreme value theory (Coles 2001, AghaKouchak *et al* 2013) on the original (not detrended) TX3d time series from reanalysis (42 values) and SEAS5 (4200 values, one for each year, member and leadtime) to estimate return values and trends in extreme (100-year return period) heatwaves.

The extremal types theorem (Fisher and Tippett 1928, Coles 2001) states that for sufficiently large samples (e.g. daily temperatures for individual years), the greatest or least value of each of these samples (e.g. annual maximum temperature) follows one out of

three possible limiting distributions, which can all be summarized in the generalized extreme value (GEV) distribution. The GEV distribution is described by location ($-\infty < \mu < \infty$), scale ($\sigma > 0$), and shape ($-\infty < \xi < \infty$) parameters:

$$F(x) = \exp\left[-\left(1 + \xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right], \quad 1 + \xi\frac{x-\mu}{\sigma} > 0. \quad (1)$$

The quantiles of the distribution can then be obtained by inverting equation (1):

$$x_p = \mu - \frac{\sigma}{\xi}(1 - \{-\log(1-p)\}^{-\xi}), \quad \xi \neq 0 \quad (2)$$

where x_p corresponds to the return level associated with the return period p^{-1} (Coles 2001). For the statistical analyses in this study we use the extRemes package in R (Gilleland and Katz 2016, R Core Team 2023). We apply generalized maximum likelihood (GML) estimation method (Martins and Stedinger 2000) for parameter estimation, which is expected to be more stable than commonly used maximum likelihood (ML) estimation (El Adlouni *et al* 2007). However, where GML estimation does not converge we resort to ML (one region only). The 95% confidence intervals are calculated based on a normal distribution. Since hydroclimatic extreme value variables are often characterized by a strong skewness this is a poor assumption (Coles 2001, El Adlouni *et al* 2007). Better methods exist (e.g. parametric bootstrapping), but they are difficult to implement and much difference is not expected (Kelder *et al* 2020). However, the calculated confidence intervals are not able to take into account uncertainty due to model correctness and should therefore be interpreted with caution. Non-stationarity can be introduced by allowing one or several of the parameters to vary with a so-called covariate (AghaKouchak *et al* 2013). Following the UNSEEN Trends protocol presented by Kelder *et al* (2020) we use years as a linear covariate and redefine the parameters μ and σ as functions of time:

$$(M_{\mu,\sigma}) \begin{cases} \mu(t) = \mu_0 + \mu_1 t, \\ \ln\sigma(t) = \phi_0 + \phi_1 t \\ \xi \end{cases}. \quad (3)$$

For the scale parameter a log-link function is used to ensure $\sigma(t) > 0$. While it is theoretically possible to introduce a time varying shape parameter ξ as well, we have no reason to suspect significant changes in shape and keep ξ constant. It can also not generally be assumed that in every case a model with time-varying location and scale parameters is most appropriate, so we implement a small decision tree as suggested by Robin and Ribes (2020). We first define three submodels of the model described in

equation (3), namely $\mathcal{M}_0$ ($\mu_1 = 0, \phi_1 = 0$; stationary case), $\mathcal{M}_\mu$ ($\phi_1 = 0$), and $\mathcal{M}_{\mu,\sigma}$. We fit the three models to both the observed (from reanalysis) and simulated (from seasonal forecast) time series. We then use a likelihood ratio test as suggested by Coles (2001) to first test $\mathcal{M}_\mu$ against $\mathcal{M}_0$. If the resulting p-value is high (>0.05), $\mathcal{M}_0$ is retained (and the assumption of non-stationarity is rejected). Otherwise, the procedure is repeated testing $\mathcal{M}_{\mu,\sigma}$ against $\mathcal{M}_\mu$. This way, we assure that we do not overinterpret the data by simply assuming trends in both variation and scale parameters. In case a non-stationary model is deemed most appropriate, we can then calculate the trend in the 100-year return value between 1981 and 2022 by computing the absolute change between these two values:

$$
\frac{\Delta x_T}{\Delta t}
$$
$$
= \frac{x_T(\mu(2022), \ln\sigma(2022), \xi) - x_T(\mu(1981), \ln\sigma(1981), \xi)}{42},
$$
(4)

where x_T is defined by equation (2).

3. Results

In the following sections first an overview of the results of the UNSEEN ensemble evaluation is given, then we present the results of the statistical assessment.

3.1. Evaluation of the UNSEEN ensemble

We test the ensemble for independence, stability, and fidelity. No issues are detected for any of the regions regarding independence and stability. However, the testing for fidelity reveals partly big differences in the statistical moments of observation and ensemble distributions. A mean bias is found in all regions, with the ensemble mean being between 2 and 4.5 °C cooler than the observations. Thus, for each region we adjust the ensemble data using additive bias correction. As long as, after mean bias correction, all four moments of the observation timeseries fall within the 95% confidence intervals, we regard the fidelity test as passed. Where differences are large, but not outside the confidence intervals, this is indicated in the results table (see additional material provided on Zenodo, Berghald *et al* 2024). Only one region, namely the Åland archipelago, fails the fidelity test and is therefore excluded from further analysis.

3.2. Volatility

The results of the volatility analysis are shown in figure 2. The mean TX3d anomalies of the top percent of simulated events for each region range from 5 °C to 7 °C. We find the highest values in the central Scandinavian area, indicating a higher risk for outstanding extremes in these regions as compared to the Norwegian coast and most of Finland.
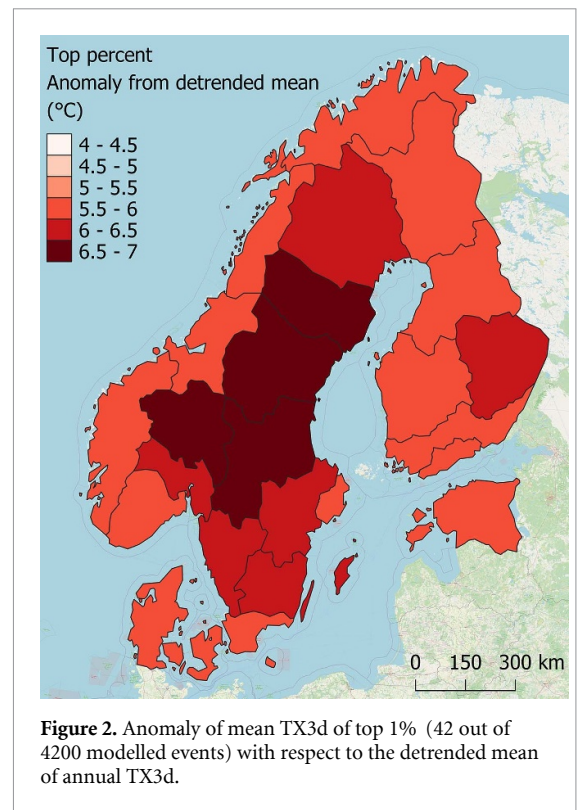
Figure 2. Anomaly of mean TX3d of top 1% (42 out of 4200 modelled events) with respect to the detrended mean of annual TX3d.

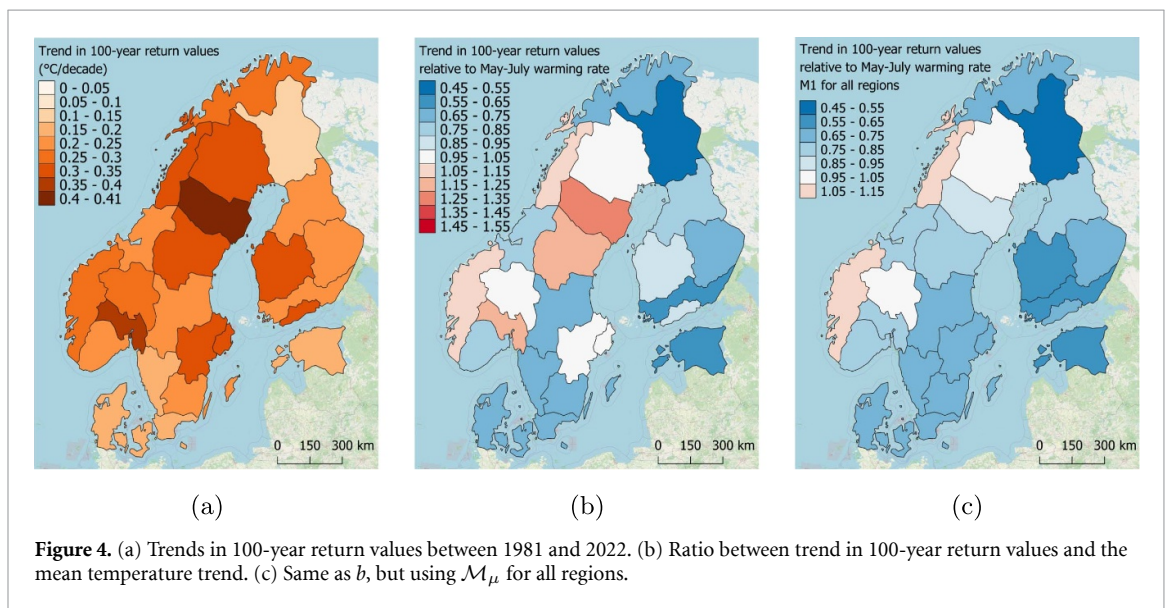### 3.3. Trends in return values and periods

3.3.1. Best model

Trends in 100-year return values of TX3d are calculated using non-stationary extreme value analysis for both reanalysis (MET Nordic long-term consistent) and ensemble data (SEAS5). We use a likelihood ratio test to confirm or reject non-stationarity in the extreme value distributions for both observed and simulated time series. Åland is the only region where non-stationarity in the observations can be confirmed at the 5% level. In contrast, non-stationarity is confirmed for all regions in the SEAS5 ensemble. Including a trend in the scale parameter ($\mathcal{M}_{\mu,\sigma}$) improves the fit in seven cases compared to variable location only ($\mathcal{M}_\mu$) as shown in figure 3(a). However, in another seven cases fitting the GEV distribution with both variable location and scale parameters fails, in these regions the model with a trend in location only is retained.

3.3.2. Changes in return periods

We compute the estimated return period in 2022 of a 100-year return value as of 1981 (figure 3(b)). The smallest change here is found in the Finnish region of Lappi (1 in 46 years) and the biggest in Nordland in Norway (1 in 20 years). While we can not find any obvious spatial patterns, it is noteworthy that the biggest changes were found in regions where using $\mathcal{M}_{\mu,\sigma}$ showed to improve the model fit (figures 3(a) and (b)).
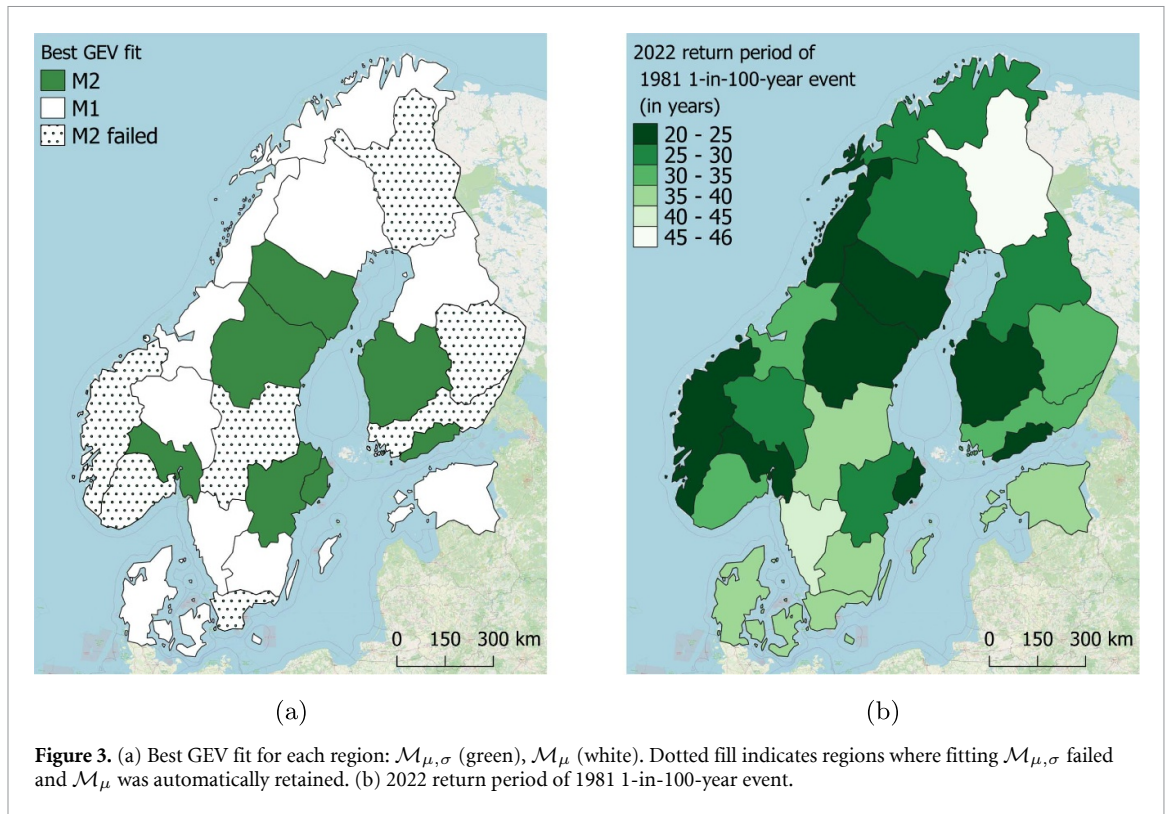
Figure 3. (a) Best GEV fit for each region: $\mathcal{M}_{\mu,\sigma}$ (green), $\mathcal{M}_{\mu}$ (white). Dotted fill indicates regions where fitting $\mathcal{M}_{\mu,\sigma}$ failed and $\mathcal{M}_{\mu}$ was automatically retained. (b) 2022 return period of 1981 1-in-100-year event.

Figure 4. (a) Trends in 100-year return values between 1981 and 2022. (b) Ratio between trend in 100-year return values and the mean temperature trend. (c) Same as *b*, but using $\mathcal{M}_{\mu}$ for all regions.

3.3.3. Trends in 100-year return values

Estimated trends from reanalysis for the 42-year period using $\mathcal{M}_{\mu}$ are between $-0.7\,°C$ and $+2.2\,°C$. However, non-stationarity could not be confirmed, which is reflected by the very large uncertainties (on average $\pm\ 3.3\,°C$). Using the much bigger SEAS5 ensemble dataset results in significantly smaller uncertainties: average 95%-confidence intervals are roughly $\pm 0.3\,°C$. We find trends in 100-year return values between $0.6\,°C$ (Lappi; corresponding to $0.14\,°C$ per decade) and $1.7\,°C$ (Västerbottens län; $0.4\,°C$ per decade) (figure 4(a)). These trends are lower than the mean temperature trends in the

same months and period along the eastern parts of the studied domain from Troms og Finnmark to Estonia as well as in southern Sweden and some regions in Norway. In central and northern Sweden as well as several regions in Norway extreme heatwave temperatures are increasing faster than the average temperature (figure 4(b)). However, if $\mathcal{M}_{\mu}$ is applied to all regions this is just the case for Vestland and Nordland at the Norwegian coast (figure 4(c)). The absolute values of the 100-year return value with covariate 2022 were found to be between $26\,°C$ along the coast of Northern Norway and $34.4\,°C$ in Eastern Middle Sweden (Östra Mellansverige), with generally

lower TX3d values along the Norwegian coast and higher values towards the southeast (supplementary material on Zenodo).

4. Discussion

In line with previous studies (World Weather Attribution 2018, Sulikowska and Wypych 2021) we do not find significant non-stationarity in observations due to high internal variability in summer weather patterns in Northern Europe. However, using the much larger SEAS5 hindcast ensemble dataset, we are able to get more robust results, revealing the until now hidden trends in heat extremes in Northern Europe. We find trends in 100-year return values (the value with 1% exceedance probability in a given year) to range from 0.1 °C–0.4 °C per decade. In most areas, these values are similar or lower than the mean warming rate in the months May–July in the same period, but in several regions the opposite is true. We also find significant changes in the probability of a given return level. For example, a heatwave event associated with a 100-year return period in 1981 is now estimated to happen once in 20–40 years. While we did not assess future trends in this study, it is very likely that these trends continue into the future, making a former extreme event the new normal (Hov 2013, Glotter and Elliott 2016).

No conclusive explanations could be found for the trend differences between regions. We find some correlation between regions with the strongest trends and where a significant trend in the scale parameter of the GEV distribution (in addition to a trend in location) is detected. Using the same statistical model for all regions leads to more similar trends. However, as especially the fidelity tests show, the statistical distributions of MJJ TX3d are quite different from one region to another, implying that the reasons for the very different trends could lie elsewhere. For example, differences in surface properties between regions or model shortcomings in their representation could cause these differences. This could be tested by repeating the same analysis using other seasonal forecast and hindcast datasets like the American Climate Forecast System (Saha *et al* 2014). However, this was out of the scope of this study. Significant (and potentially badly represented) land-use changes during the studied 42-year period, for example around the cities of Helsinki, Stockholm and Oslo are not expected, but could play a minor role.

This leads to a question that has not yet been conclusively answered in the literature. Even if the statistical distributions of models and observations do not differ significantly, it is still unclear whether the model produces the extremes of interest for the right physical reasons (e.g Vautard *et al* 2019), Philip *et al* 2020). Indeed, the computed confidence intervals are likely to be too small, since systematic model shortcomings are not accounted for (Kelder 2023). It is also by definition difficult to check how realistic simulated extremes are, especially for events far above the observed record ('unseen' events). Adapting the choice of the model to the event of interest is crucial: For regional-scale multi-day events like TX3d seasonal forecast models like SEAS5 are well-suited, while sub-daily or multi-year events cannot be captured by seasonal forecasts (Kelder 2023). Confidence in the UNSEEN ensemble can be increased by an assessment of the drivers of observed and unseen extremes, which is a time-consuming process (Kelder 2023). Here, we just checked one randomly selected 'unseen' event, which showed to be connected to dry conditions and a large-scale sea level pressure pattern typical for atmospheric blocking (figure 1(a)). Given that the statistical evaluation shows good results for all regions (section 3.1) and that the mean bias is similar between all regions we have enough confidence in the reliability of the SEAS5 dataset. An extensive discussion on the strengths and shortcomings of UNSEEN is given in Kelder (2023).

5. Conclusion

We apply the UNSEEN approach on a large ensemble seasonal hindcast dataset to robustly assess trends and changes in heat extremes in Northern Europe. Unlike previous studies based on observations or reanalysis alone, we are able to show that 100-year return values have increased over the past 42 years, thus highlighting the added value of making use of large-ensemble seasonal hindcast datasets. However, the spatial patterns of trends are difficult to interpret. While in some regions trends in 1-in-100-year extremes are higher than trends in the mean temperature, in others, they are not. Different trends in the model parameters other than linear ones could be assessed in future studies, but physical limitations of the model ensemble also need to be considered. We also find that especially central parts of Sweden and Norway have a high heatwave volatility, which means that the potential for a 'surprise'-event is highest in these areas. With continuing global climate change trends in both mean temperature and heat extremes are likely to continue. The methods applied in this study cannot be used to assess future climate change, however, the framework could potentially be adapted for the use of e.g. large ensembles of regional climate models.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.5281/zenodo.8391578. Original SEAS5 seasonal forecast data is archived and accessible for registered users at ECMWF's Meteorological Archival and Retrieval System MARS (www.ecmwf.int/en/forecasts/access-forecasts/access-archive-datasets). Description and availability of the MET Nordic

long-term consistent product can be found at MET Norways's Github pages: (https://github.com/metno/NWPdocs/wiki/MET-Nordic-dataset). Shapefiles of the NUTS dataset are available from Eurostat's webpages (https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units).

Acknowledgments

ORCID iDs

S Berghald ● https://orcid.org/0009-0005-4237-5155
S Mayer ● https://orcid.org/0000-0003-0755-6354
P Bohlinger ● https://orcid.org/0000-0002-7299-3676

References

Aasprang B 2013 Norske bønder og globale klimaendringer: En kvantitativ analyse av bønders oppfatninger av hvordan deres egne gårdsbruk vil bli påvirket av klimaendringer: Rapport 3/2013

AghaKouchak A, Easterling D R, Hsü K, Schubert S D and Sorooshian S (eds) 2013 *Extremes in a Changing Climate* (*Water Science and Technology Library*) (Springer) (https://doi.org/10.1007/978-94-007-4479-0)

Ahasan N 2020 pyscissor: a Python3 module for extracting data from netcdf file under a shapefile region (available at: https://github.com/nzahasan/pyscissor)

Anderson J L 1997 The impact of dynamical constraints on the selection of initial conditions for ensemble predictions: low-order perfect model results *Mon. Weather Rev.* **125** 2969–83

Angelova D and Lupio N B 2020 Constructing a meteorological indicator dataset for selected European NUTS 3 regions *Data Brief* **31** 105786

Bachmair S, Tanguy M, Hannaford J and Stahl K 2018 How well do meteorological indicators represent agricultural and forest drought across Europe? *Environ. Res. Lett.* **13** 034042

Berghald S, Mayer S and Bohlinger P 2024 Revealing trends in extreme heatwave intensity: applying the UNSEEN approach to Nordic countries *Zenodo* https://doi.org/10.5281/zenodo.10519316

Blunden J and Arndt D S (eds) 2019 State of the climate in 2018 *Bull. Amer. Meteor. Soc.* **100** S217–30

Brás T A, Seixas J, Carvalhais N and Jägermeyr J 2021 Severity of drought and heatwave crop losses tripled over the last five decades in Europe *Environ. Res. Lett.* **16** 065012

Coles S 2001 *An Introduction to Statistical Modeling of Extreme Values* (Springer)

Coughlan de Perez E, Ganapathi H, Masukwedza G I T, Griffin T and Kelder T 2023 Potential for surprising heat and drought events in wheat-producing regions of USA and China *npj Clim. Atmos. Sci.* **6** 56

Devot A, Royer L, Arvis B, Deryng D, Caron Giauffret E, Giraud L and Rouillard J 2023 Research for AGRI committee - the impact of extreme climate events on agriculture production in the EU *Technical Report* (European Parliament, Policy Department for Structural and Cohesion Policies)

Domínguez-Castro F, Reig F, Vicente-Serrano S M, Aguilar E, Peña-Angulo D, Noguera I and El Kenawy A M 2020 A multidecadal assessment of climate indices over Europe *Sci. Data* **7** 125

EEA. 2017 Climate change, impacts and vulnerability in Europe 2016 (Vol. EEA Report No 1/2017) (available at: www.eea.europa.eu/publications/climate-change-impacts-and-vulnerability-2016)

El Adlouni S, Ouarda T B M J, Zhang X, Roy R and Bobée B 2007 Generalized maximum likelihood estimators for the nonstationary generalized extreme value model *Water Resour. Res.* **43** W03410

Eurostat 2022 *Statistical Regions in the European Union and Partner Countries: Nuts and Statistical Regions 2021* 2022 edn (Publications Office of the European Union)

Fisher R A and Tippett L H C 1928 Limiting forms of the frequency distribution of the largest or smallest member of a sample *Math. Proc. Camb. Phil. Soc.* **24** 180–90

Gilleland E and Katz R W 2016 extRemes 2.0: an extreme value analysis package in R *J. Stat. Softw.* **72** 1–39

Glotter M and Elliott J 2016 Simulating US agriculture in a modern dust bowl drought *Nat. Plants* **3** 16193

Haakenstad H and Breivik Ø 2022 NORA3. Part II: precipitation and temperature statistics in complex terrain modeled with a nonhydrostatic model *J. Appl. Meteorol. Climatol.* **61** 1549–72

Haakenstad H, Breivik Ø, Furevik B R, Reistad M, Bohlinger P and Aarnes O J 2021 NORA3: a nonhydrostatic high-resolution hindcast of the North Sea, the Norwegian Sea and the Barents Sea *J. Appl. Meteorol. Climatol.* **60** 1443–64

Hanssen-Bauer I and Nordli P Ø 1998 *Annual and Seasonal Temperature Variations in Norway 1876-1997* Vol. DNMI Report 25/98 KLIMA

Haugen I, Hanssen-Bauer I, Skaland R G, Engh M and Svardal S 2019 *Kornproduksjon i møte med klimaendringer: Et mer klimarobust landbruk i Vestfold og Telemark* vol. TF-rapport No. 491

Hov O 2013 *Extreme Weather Events in Europe: Preparing for Climate Change Adaptation* (Norwegian Meteorological Institute)

Hoyer S and Hamman J 2017 xarray: N-D labeled arrays and datasets in Python *J. Open Res. Softw.* **5** 10

IPCC 2023 Summary for Policymakers *Climate Change 2023: Synthesis Report. A Report of the Intergovernmental Panel on Climate Change* ed H Leeand and J Romero pp 3–34 (accepted)

Johnson S J, Stockdale T N, Ferranti L, Balmaseda M A, Molteni F, Magnusson L and Monge-Sanz B M 2019 SEAS5: the new ECMWF seasonal forecast system *Geosci. Model Dev.* **12** 1087–117

Johnsson B, Søyland V, Mattsson K, Hansson P, Tunberg M, Hansson P and Ramos M A 2019 *Det Nordiska Jordbruket: Utmaningar i en Framtid präGlad av mer Extremväder* (Nordic Council of Ministers)

Kelder T 2023 *Unseen or Unrealistic? Using Ensemble Simulations to Explore Unseen Weather Extremes* (Loughborough University)

Kelder T, Marjoribanks T I, Slater L J, Prodhomme C, Wilby R L, Wagemann J and Dunstone N 2022 An open workflow to gain insights about low–likelihood high–impact weather events from initialized predictions *Meteorol. Appl.* **29** e2065

Kelder T, Müller M, Slater L J, Marjoribanks T I, Wilby R L, Prodhomme C and Nipen T 2020 Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes *npj Clim. Atmos. Sci.* **3** 47

Kent C, Pope E, Dunstone N, Scaife A A, Tian Z, Clark R and Lewis K 2019 Maize drought hazard in the Northeast farming region of China: unprecedented events in the current climate *J. Appl. Meteorol. Climatol.* **58** 2247–58

Kew S, Philip S, van Oldenborgh G J, van der Schrier G, Otto F and Vautard R 2019 The exceptional summer heat wave in Southern Europe 2017 *Bull. Am. Meteorol. Soc.* **100** S49–S53

Landbruksdirektoratet 2020 *Evaluering av iverksatte tiltak - tørkesommeren 2018* vol. Rapport No. 3/2020

LMD 2016 Landbruk og klimaendringer: Rapport fra arbeidsgruppe

Maraun D 2016 Bias correcting climate change simulations - a critical review *Curr. Clim. Change Rep.* **2** 211–20

Martins E S and Stedinger J R 2000 Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data *Water Resour. Res.* **36** 737–44

Müller M, Kelder T and Palerme C 2022 Decline of sea-ice in the Greenland Sea intensifies extreme precipitation over Svalbard *Weather Clim. Extremes* **36** 100437

Norwegian Meteorological Institute 2023 MET nordic dataset (available at: https://github.com/metno/NWPdocs/wiki/MET-Nordic-dataset)

Philip S, Kew S, van Oldenborgh G J, Otto F, Vautard R, van der Wiel K and van Aalst M 2020 A protocol for probabilistic extreme event attribution analyses *Adv. Stat. Climatol. Meteorol. Oceanogr.* **6** 177–203

Philip S, Kew S, Vautard R, Vahlberg M, Singh R, Driouech F and Otto F 2023 Extreme April heat in Spain, Portugal, Morocco & Algeria almost impossible without climate change Imperial College London

QGIS.org. 2023 QGIS geographic information system QGIS Association (available at: www.qgis.org)

R Core Team 2023 R: A language and environment for statistical computing R Foundation for Statistical Computing (available at: www.R-project.org/)

Räisänen J 2019 Energetics of interannual temperature variability *Clim. Dyn.* **52** 3139–56

Robin Y and Ribes A 2020 Nonstationary extreme value analysis for event attribution combining climate models and observations *Adv. Stat. Clim. Meteorol. Oceanogr.* **6** 205–21

Ruosteenoja K, Markkanen T and Räisänen J 2020 Thermal seasons in northern Europe in projected future climate *Int. J. Climatol.* **40** 4444–62

Saha S, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P and Becker E 2014 The NCEP climate forecast system version 2 *J. Clim.* **27** 2185–208

Schaller N, Sillmann J, Anstey J, Fischer E M, Grams C M and Russo S 2018 Influence of blocking on Northern European and Western Russian heatwaves in large climate model ensembles *Environ. Res. Lett.* **13** 054015

Schulzweida U 2022 CDO user guide Zenodo (available at: https://doi.org/10.5281/ZENODO.7112925)

Sippel S, Otto F, Forkel M, Allen M R, Guillod B P, Heimann M and Mahecha M D 2016 A novel bias correction methodology for climate impact simulations *Earth Syst. Dyn.* **7** 71–88

Skaland R G, Colleuille H, Andersen A S H, Mamen J, Grinde L, Tajet H T T and Hygen H O 2019 *Tørkesommeren 2018* vol. METinfo 14/2019

Slater L J, Anderson B, Buechel M, Dadson S, Han S, Harrigan S and Wilby R L 2021 Nonstationary weather and water extremes: a review of methods for their detection, attribution and management *Hydrol. Earth Syst. Sci.* **25** 3897–935

Sulikowska A and Wypych A 2021 Seasonal variability of trends in regional hot and warm temperature extremes in Europe *Atmosphere* **12** 612

Thompson V, Dunstone N J, Scaife A A, Smith D M, Hardiman S C, Ren H L and Belcher S E 2019 Risk and dynamics of unprecedented hot months in South East China *Clim. Dyn.* **52** 2585–96

Thompson V, Dunstone N J, Scaife A A, Smith D M, Slingo J M, Brown S and Belcher S E 2017 High risk of unprecedented UK rainfall in the current climate *Nat. Commun.* **8** 107

Tomczyk A M, Piotrowski P and Bednorz E 2017 Warm spells in Northern Europe in relation to atmospheric circulation *Theor. Appl. Climatol.* **128** 623–34

Trenberth K E and Shea D J 2005 Relationships between precipitation and surface temperature *Geophys. Res. Lett.* **32** n/a-n/a

van den Brink H W, Können G P, Opsteegh J D, van Oldenborgh G J and Burgers G 2005 Estimating return periods of extreme events from ECMWF seasonal forecast ensembles *Int. J. Climatol.* **25** 1345–54

Vautard R, Christidis N, Ciavarella A, Alvarez-Castro C, Bellprat O, Christiansen B and Yiou P 2019 Evaluation of the HadGEM3-A simulations in view of detection and attribution of human influence on extreme events in Europe *Clim. Dyn.* **52** 1187–210

World Weather Attribution 2018 Heatwave in northern Europe, summer 2018 – World Weather Attribution (available at: www.worldweatherattribution.org/attribution-of-the-2018-heat-in-northern-europe/)

Yiou P, Cattiaux J, Faranda D, Kadygrov N, Jézéquel A, Naveau P and Vrac M 2020 Analyses of the Northern European Summer Heatwave of 2018 *Bull. Am. Meteorol. Soc.* **101** S35–S40

Zachariah M, Philip S, Pinto I, Vahlberg M, Singh R, Otto F and Kimutai J 2023 *Extreme heat in North America, Europe and China in July 2023 Made Much More Likely by Climate Change (Imperial College London)* Imperial College London

Zender C S 2008 Analysis of self-describing gridded geoscience data with netCDF Operators (NCO) *Environ. Model. Softw.* **23** 1338–42