






RESEARCH ARTICLE

The regional MiKlip decadal prediction system for Europe: Hindcast skill for extremes and user-oriented variables

Julia Moemken¹  | Hendrik Feldmann¹  | Joaquim G. Pinto¹  | Benjamin Buldmann¹ | Natalie Laube¹ | Christopher Kadow^{2,3} | Andreas Paxian⁴ | Bente Tiedje⁵ | Christoph Kottmeier¹  | Jochem Marotzke⁶ 

¹Institute of Meteorology and Climate Research (IMK-TRO), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

²German Climate Computing Center (DKRZ), Hamburg, Germany

³Institute of Meteorology, Freie Universität Berlin, Berlin, Germany

⁴Deutscher Wetterdienst, Offenbach/Main, Germany

⁵Climate Service Center Germany (GERICS), Helmholtz-Zentrum Geesthacht, Hamburg, Germany

⁶Max Planck Institute for Meteorology, Hamburg, Germany

Correspondence

Julia Moemken, Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research, Wolfgang-Gaede-Str. 1, 76131 Karlsruhe, Germany.
Email: julia.moemken@kit.edu

Funding information

The study was mainly funded by the Bundesministerium für Bildung und Forschung (BMBF) under project FONa MiKlip-II; AXA Research Fund

Abstract

Regional climate predictions for the next decade are gaining importance, as this period falls within the planning horizon of politics, economy, and society. The potential predictability of climate indices or extremes at the regional scale is of particular interest. The German MiKlip project (“mid-term climate forecast”) developed the first regional decadal prediction system for Europe at 0.44° resolution, based on the regional model COSMO-CLM using global MPI-ESM simulations as boundary conditions. We analyse the skill of this regional system focussing on extremes and user-oriented variables. The considered quantities are related to temperature extremes, heavy precipitation, wind impacts, and the agronomy sector. Variables related to temperature (e.g., frost days, heat wave days) show high predictive skill (anomaly correlation up to 0.9) with very little dependence on lead-time, and the skill patterns are spatially robust. The skill patterns for precipitation-related variables (e.g., heavy precipitation days) and wind-based indices (like storm days) are less skilful and more heterogeneous, particularly for the latter. Quantities related to the agronomy sector (e.g., growing degree days) show high predictive skill, comparable to temperature. Overall, we provide evidence that decadal predictive skill can be generally found at the regional scale also for extremes and user-oriented variables, demonstrating how the utility of decadal predictions can be substantially enhanced. This is a very promising first step towards impact-related modelling at the regional scale and the development of individual user-oriented products for stakeholders.

KEYWORDS

climate services, Europe, extremes, MiKlip, regional decadal predictions, user needs

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors *International Journal of Climatology* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

1 | INTRODUCTION

Short-term climate predictions for the next 1–10 years are gaining importance in the climate science community (e.g., Meehl *et al.*, 2009). These so-called decadal predictions can also be of high value for impact modellers and decision makers in politics, economy and society (Vera *et al.*, 2010; Meehl *et al.*, 2014). Within the Coupled Model Intercomparison Project Phase 5 (CMIP5; Taylor *et al.*, 2012), a set of globally coordinated climate model experiments has been consolidated, comprising simulations for the recent past, decadal simulations and climate change projections. Based on this experience, decadal predictions also play an important role in the Decadal Climate Prediction Project (DCPP) as part of CMIP6 (Boer *et al.*, 2016; Eyring *et al.*, 2016).

Since 2007, the number of studies assessing the predictive skill of decadal systems has strongly increased (cf. Meehl *et al.*, 2014). While most studies focus on the global scale (e.g., Kim *et al.*, 2012; Müller *et al.*, 2012, 2014; Doblas-Reyes *et al.*, 2013; Bellucci *et al.*, 2015; Kadow *et al.*, 2016), some studies also investigate indices at the regional scale (e.g., Kruschke *et al.*, 2016; Moemken *et al.*, 2016). Overall, the global warming trend enables a certain predictability by itself, especially for temperature-related climate indicators. On decadal time-scales, this climate trend is of the same order of magnitude as the climate variability of global mean temperatures (Flato *et al.*, 2013; Hartmann *et al.*, 2013; Chen and Tung, 2018). On the regional scale, the variability might be even stronger and therefore could provide additional predictability. Specifically, the North Atlantic has been identified as a hot spot of decadal predictability (Sutton and Hodson, 2005; Latif and Keenlyside, 2011; Müller *et al.*, 2012; Kadow *et al.*, 2017). This predictability arises from the long-term variability pattern of the Atlantic Meridional Overturning Circulation (AMOC; Matei *et al.*, 2012) and the Atlantic Multi-decadal Variability (AMV; Zhang *et al.*, 2019). Most global decadal prediction systems provide a high hindcast skill for the AMV Index as well as North Atlantic sea surface- and 2 m-temperatures up to 9 years ahead (Doblas-Reyes *et al.*, 2013; Bellucci *et al.*, 2015). A strong potential for skilful decadal predictions has also been identified for Europe (Feldmann *et al.*, 2019; Reyers *et al.*, 2019). For example, Ghosh *et al.* (2017) provided evidence that an increased heat flux due to higher sea surface temperatures in the North Atlantic is triggered by the positive phase of the AMV. This mechanism induces a wave-like response in the sea-level pressure field and blocking-like situations downstream over Europe, thus, leading to changes in temperature and precipitation. These mechanisms are regionally and seasonally

dependent, leading to different levels of predictive skill for different climate indicators for Europe and inducing a lead-time dependency of skill.

The German research consortium MiKlip (“Mittelfristige Klimaprognosen”; www.fona-miklip.de; Marotzke *et al.*, 2016) developed a global decadal prediction system based on the Max-Planck-Institute Earth System Model (MPI-ESM) and produced several hindcast ensemble generations. In addition, MiKlip provides the first systematic efforts to establish a regional component of the decadal prediction ensemble for Europe, which culminated in the dynamical downscaling of a full global hindcast ensemble (Feldmann *et al.*, 2019). Mieruch *et al.* (2014) and Reyers *et al.* (2019) previously analysed sub-samples of this regionalised hindcast set. All three studies show added value by the downscaling approach for basic variables like surface temperature for large parts of Europe. Their results also indicate a reduction of bias and an increase of reliability in the regional system compared to the global hindcasts (cf. Feldmann *et al.*, 2019). Other efforts used statistical-dynamical downscaling for selected quantities like wind energy potentials (Moemken *et al.*, 2016) and wind gusts (Haas *et al.*, 2016). Both studies indicate predictive skill for Europe, especially for the first years (1–4) after initialisation, related to variations in the frequency of westerly weather patterns.

Mean temperature and mean precipitation are the most commonly analysed variables in decadal hindcast studies (e.g., Doblas-Reyes *et al.*, 2013; Bellucci *et al.*, 2015). The MiKlip decadal prediction website (www.fona-miklip.de/decadal-forecast/decadal-forecast-for-2019-2028/) offers operational global decadal predictions of mean temperature. However, such general products are of limited value for potential users outside the science community, like stakeholders in civil services, governments or economy (Lemos *et al.*, 2012; Hackenbruch *et al.*, 2017; Schipper *et al.*, 2019). Recent user-oriented projects (e.g., the EUPORIAS project for seasonal predictions; Buontempo *et al.*, 2018) suggest that user needs are rather specific across different sectors, and that individual prediction products are highly needed to enable applicability in the “real world.” Building on these conclusions, some efforts within MiKlip focussed on the needs of private companies and public authorities. While developing individual products for different sectors remains the aim, specific climate indices or extremes related to the basic climate variables temperature, precipitation, and wind at regional resolution are needed as an intermediate step (cf. Copernicus Climate Change Service [C3S]; Buontempo *et al.*, 2019).

In this study, we analyse the predictive skill of the regional MiKlip system focusing on extremes and climate

indices related to temperature, precipitation, wind or connected to the agronomy sector. Thus, we aim to identify potential fields of application for regional decadal predictions. Thereby, our work extends previous regional studies like Feldmann *et al.* (2019) and Reyers *et al.* (2019). Methods and datasets are described in Section 2. Section 3 focusses on the results, while a summary and discussion concludes this paper in Section 4.

2 | DATA AND METHODS

2.1 | Data and variables

We analyse the second MiKlip ensemble generation “baseline1” (Pohlmann *et al.*, 2013a), which features the Max Planck Institute for Meteorology Earth System Model in Low-Resolution mode (MPI-ESM-LR; Giorgetta *et al.*, 2013). Baseline1 combines anomaly initialisation in the ocean (using ORA-S4 ocean reanalysis [Balmaseda *et al.*, 2013]) and full-field initialisation in the atmosphere (with ERA40 [Uppala *et al.*, 2005] and ERA-Interim [Dee *et al.*, 2011]), as described in Pohlmann *et al.* (2013b). The ensemble comprises 10 decadal simulations for each annual starting year of the period 1960–2017. The individual ensemble members are generated using a 1-day-lagged initialisation (Müller *et al.*, 2012; Romanova *et al.*, 2018).

The global hindcasts are dynamically downscaled to the EURO-CORDEX domain (Giorgi *et al.*, 2009; see Figure S1) using the regional climate model COSMO-CLM (CCLM; Rockel *et al.*, 2008) at a spatial resolution of 0.44°. The downscaling is applied to all baseline1 ensemble members for all starting dates in 1960–2017, resulting in the identical number of 5,800 simulation years (58 starting dates \times 10 years \times 10 ensemble members) as for the global ensemble. A thorough evaluation of the regional decadal prediction system focussing on the basic variables temperature, precipitation, and wind speed can be found in Reyers *et al.* (2019) and Feldmann *et al.* (2019).

For the evaluation of temperature and precipitation related indices, we use the observational dataset E-OBS (V14; Haylock *et al.*, 2008) with daily temporal resolution at a regular 0.5° grid. For wind-related variables, no gridded observational dataset is available for Europe. Thus, we use a CCLM simulation with ERA40 and ERA-Interim boundary conditions as reference. For this simulation, CCLM is applied in the same setup as for the downscaling of the decadal prediction system (cf. Feldmann *et al.*, 2019).

To estimate the added value of initialisation, an ensemble of seven regional uninitialized simulations

(historicals) is used as reference. This ensemble was implemented by applying the same dynamical downscaling approach to a set of global historical runs. The global uninitialized ensemble uses the same MPI-ESM model version and setup as the global hindcasts. The runs are started from a pre-industrial control simulation and consider, among others, aerosol and greenhouse gas concentrations for the period 1850–2005 (e.g., Müller *et al.*, 2012). Note that only the first seven (out of 10) hindcast ensemble members are used to ensure a fair comparison to the available uninitialized simulations.

Based on the collection of potential user needs within MiKlip, a promising first step towards the development of individual user products is the analysis of different extremes and user-oriented variables. Most variables are defined in the “Expert Team of Climate Change Detection Indices” (ETCCDI; Van Engelen *et al.*, 2008; Zhang *et al.*, 2011). The complete set of analysed variables/indices is given in Table 1. We focus here only on a subset, as some indices are closely linked to each other and/or already discussed in other studies (e.g., Moemken *et al.*, 2016; Feldmann *et al.*, 2019).

2.2 | Data processing and skill metrics

All datasets used in this study are processed in the same way to enable comparability (Figure 1). First, yearly time series for all indices/variables are derived for hindcasts, historicals, and observations. In the next step, the raw hindcast time series are recalibrated against observations using the Decadal Climate Forecast Recalibration Strategy (DeFoReSt) developed within MiKlip by Pasternack *et al.* (2018). DeFoReSt accounts for three features characteristic for decadal datasets: lead- and start-year-dependent unconditional and conditional bias, as well as ensemble dispersion. With this aim, DeFoReSt combines the parametric drift correction by Kruschke *et al.* (2016; lead-time dependency) and the non-stationary model drift correction of Kharin *et al.* (2012; start-year dependency). The method uses first-order polynomials (linear trend) to capture start-year-dependent errors as well as third-order polynomials for lead-year dependency. In addition, Pasternack *et al.* (2018) implemented a parametric adjustment of the conditional bias and the ensemble spread by using third- and second-order polynomials, respectively. A cross-validation is included in DeFoReSt. More details can be found in Pasternack *et al.* (2018). Please note that the approach is univariate. Thus, all indices are calculated prior to recalibration.

The performance of the regional system is evaluated following the verification methods by Goddard *et al.* (2013) and Feldmann *et al.* (2019). Skill scores are calculated for

TABLE 1 List of indices, which show decadal predictive skill. Indices marked with asterisks are discussed in detail in this study

| Index | Basic variable | Definition | Reference |
|---------------------------------------|-------------------------|---|---|
| Frost days (FD)* | Minimum temperature TN | Annual count of days with $TN < 0^{\circ}\text{C}$ | Van Engelen <i>et al.</i> (2008) |
| Heatwave days (HWDS)* | Maximum temperature TX | Annual count of days with $TX > 90\text{th percentile of May-Oct}$ | Fischer and Schär (2010), Zhang <i>et al.</i> (2011) |
| Ice days (ID) | Maximum temperature TX | Annual count of days with $TX < 0^{\circ}\text{C}$ | Van Engelen <i>et al.</i> (2008) |
| Summer days (SU)* | Maximum temperature TX | Annual count of days with $TX > 25^{\circ}\text{C}$ | Van Engelen <i>et al.</i> (2008) |
| Maximum temperature summer (TASMAX)* | Maximum temperature TX | Summer (May-Oct) mean of daily TX | |
| Tropical nights (TR) | Minimum temperature TN | Annual count of days with $TN > 20^{\circ}\text{C}$ | Van Engelen <i>et al.</i> (2008) |
| Wet days (RR1) | Precipitation amount RR | Annual count of days with $RR \geq 1\text{ mm}$ | Van Engelen <i>et al.</i> (2008) |
| Heavy precipitation days (R10mm)* | Precipitation amount RR | Annual count of days with $RR \geq 10\text{ mm}$ | Van Engelen <i>et al.</i> (2008) |
| Very heavy precipitation days (R20mm) | Precipitation amount RR | Annual count of days with $RR \geq 20\text{ mm}$ | Van Engelen <i>et al.</i> (2008) |
| Flood index (RM7P80) | Precipitation amount RR | 80th percentile of 7-day running mean of RR | |
| Flood index (RM7P95)* | Precipitation amount RR | 95th percentile of 7-day running mean of RR | |
| Simple daily intensity index (SDII) | Precipitation amount RR | Annual mean precipitation amount at wet days | Van Engelen <i>et al.</i> (2008) |
| Surface wind speed Oct–Mar (SFCWIND)* | Surface wind speed V | Winter (ONDJFM) mean of daily V | |
| Storm days Oct–Mar (SFCWIND98W)* | Surface wind speed V | Winter count of days with $V > 98\text{th percentile}$ | |
| Wind energy potential (EOUT) | Surface wind speed V | $E_{\text{out}} = c_p \frac{1}{2} \rho \pi R^2 v_{80m}^3$ | Reyers <i>et al.</i> (2015), Moemken <i>et al.</i> (2016) |
| Growing degree days (GDD)* | Mean temperature TG | $GDD = \sum_{\text{Apr}}^{\text{Oct}} (TG - 10^{\circ}\text{C})$ | McMaster and Wilhelm (1997), Fraga <i>et al.</i> (2016) |
| Growing season length (GSL)* | Mean temperature TG | Number of days between first occurrence of at least 6 consecutive days with $TG > 5^{\circ}\text{C}$ and first occurrence of at least 6 consecutive days with $TG < 5^{\circ}\text{C}$ after July first | Van Engelen <i>et al.</i> (2008) |
| Heating degree days (HD) | Mean temperature TG | Annual sum of $17^{\circ}\text{C} - TG$, only for $TG < 17^{\circ}$ | Van Engelen <i>et al.</i> (2008), Feldmann <i>et al.</i> (2019) |

all 4-year periods from lead-time years 1–4 (LT1-4) to LT7-10, to account for the lead-time dependence of the decadal hindcast skill following the DCPD protocol. All lead-times cover the identical analysis period 1967–2016. This is achieved by shifting the respective starting dates for the calculation, that is, from decades 1966–2012 for LT1-4 to decades 1960–2006 for LT7-10 (see also Paxian *et al.*, 2019; their Figure 2). We apply two types of skill scores: the Anomaly Correlation Coefficient (ACC) to

estimate the overall skill, and the Ranked Probability Skill Score (RPSS; Wilks, 2011; Ferro, 2014) as measure for the reliability. The ACC is a measure for the relationship between the ensemble mean of the hindcasts/historicals and observations, ranging from -1 (perfect anti-correlation) to 1 (perfect correlation). A positive RPSS (perfect score 1) indicates that the hindcasts have a higher probability to predict an observed category than a reference dataset, and vice versa for negative RPSS. Both skill metrics use

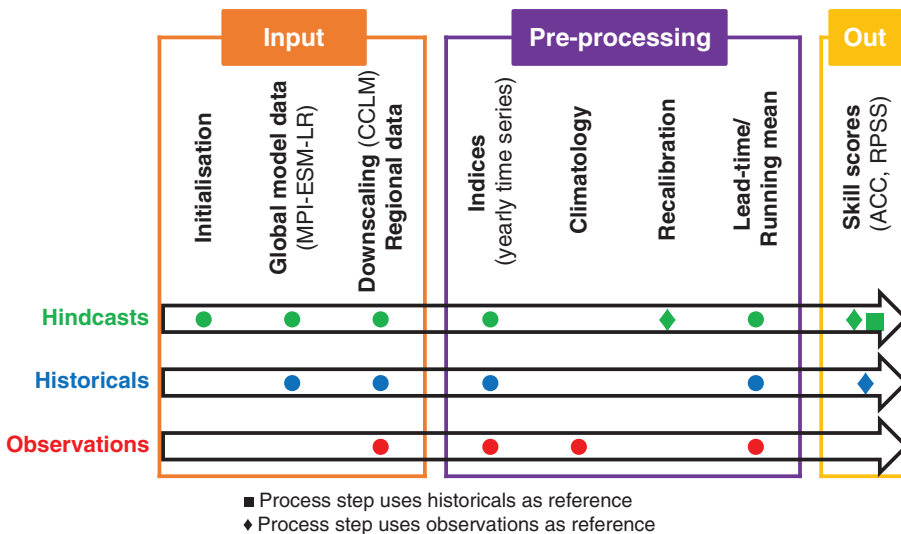


FIGURE 1 Concept for the workflow to calculate the skill metrics for regional decadal climate predictions, highlighting the different treatment of the involved datasets and their dependencies during the individual processing steps

climatology and the regional historical ensemble as reference, following the DCPD protocol (Boer *et al.*, 2016). Other widely used benchmarks for evaluating the skill of dynamical model forecasts are persistence forecasts, linear regression and statistical models (e.g., Meehl *et al.*, 2016; Suckling *et al.*, 2017). However, most results shown in our study focus on climatology as this is the most common (and comprehensible) reference for users of climate information.

The forecast and hindcast verification suite (Kadow *et al.*, 2016) of the MiKlip Central Evaluation System (CES; www-miklip.dkrz.de) provides all tools for skill assessment. The use of CES makes the results shown here comparable to most of the other MiKlip results (e.g., Pohlmann *et al.*, 2013b; Marotzke *et al.*, 2016; Feldmann *et al.*, 2019). The ACC is calculated using the MurCSS plugin (Murphy-Epstein decomposition and Continuous Ranked Probability Skill Score; Illing *et al.*, 2014), while the calculation of RPSS is performed in the ProBLEMS plugin (Probabilistic Ensemble verification for MiKlip using Specs Verification; www-miklip.dkrz.de). To remove systematic climatological biases, both tools calculate anomalies (ICPO, 2011). The significance of results at the 95% level (Goddard *et al.*, 2013) is determined using a block-bootstrapping (here 500 times) with a random re-sampling of the time series with replacement.

3 | RESULTS

3.1 | Temperature-related variables

The mean near-surface temperature shows a high predictive skill in decadal hindcasts, both on the global and the regional scale (cf. Doblas-Reyes *et al.*, 2013; Meehl

et al., 2014; Bellucci *et al.*, 2015; Feldmann *et al.*, 2019). The main reason for this high skill is the contribution given by the climate trend, which is more pronounced for temperature than for wind or precipitation. In this section, we analyse how far the skill for mean temperature extends to other temperature-related variables, with examples for warm and cold extremes. Feldmann *et al.* (2019) could show that the predictive skill of temperature in the regional MiKlip system is higher for the warm season than for the cold season (their Figures 9 and 10). Therefore, we chose the daily maximum summer temperature (TASMAX; Table 1) as a first example. TASMAX is an indicator for summerly heat conditions relevant for, for example, heat stress (Honda *et al.*, 2007). Figures 2a, b depict spatial ACC and RPSS plots for TASMAX for LT2-5, using the climatology as reference. For most of Europe, the correlation is high and significant, reaching values close to one for the coastal Mediterranean regions. However, correlations drop below 0.5 for Scandinavia, the British Isles and Greece. The RPSS exhibits similar spatial patterns compared to the ACC, but with lower values. Nevertheless, RPSS is positive everywhere except for Greece. In general, the skill scores and the fraction of significant grid points exhibit nearly no lead-time dependence (Figure S2). While the skill of TASMAX is already high and significant in the raw data, the recalibration leads to a further improvement (Figure S3). Skill scores are much lower and less significant when using the regional historicals as reference (Figure S4–S8). While no added value of initialisation is detected for the British Isles and Greece, the hindcast initialisation improves the predictive skill for the rest of Europe (particularly for RPSS with values up to 0.6).

The results for the so-called summer days (SU, maximum temperature above 25°C; Table 1) are shown in Figures 2c, d (ACC, RPSS). Skill scores are positive and

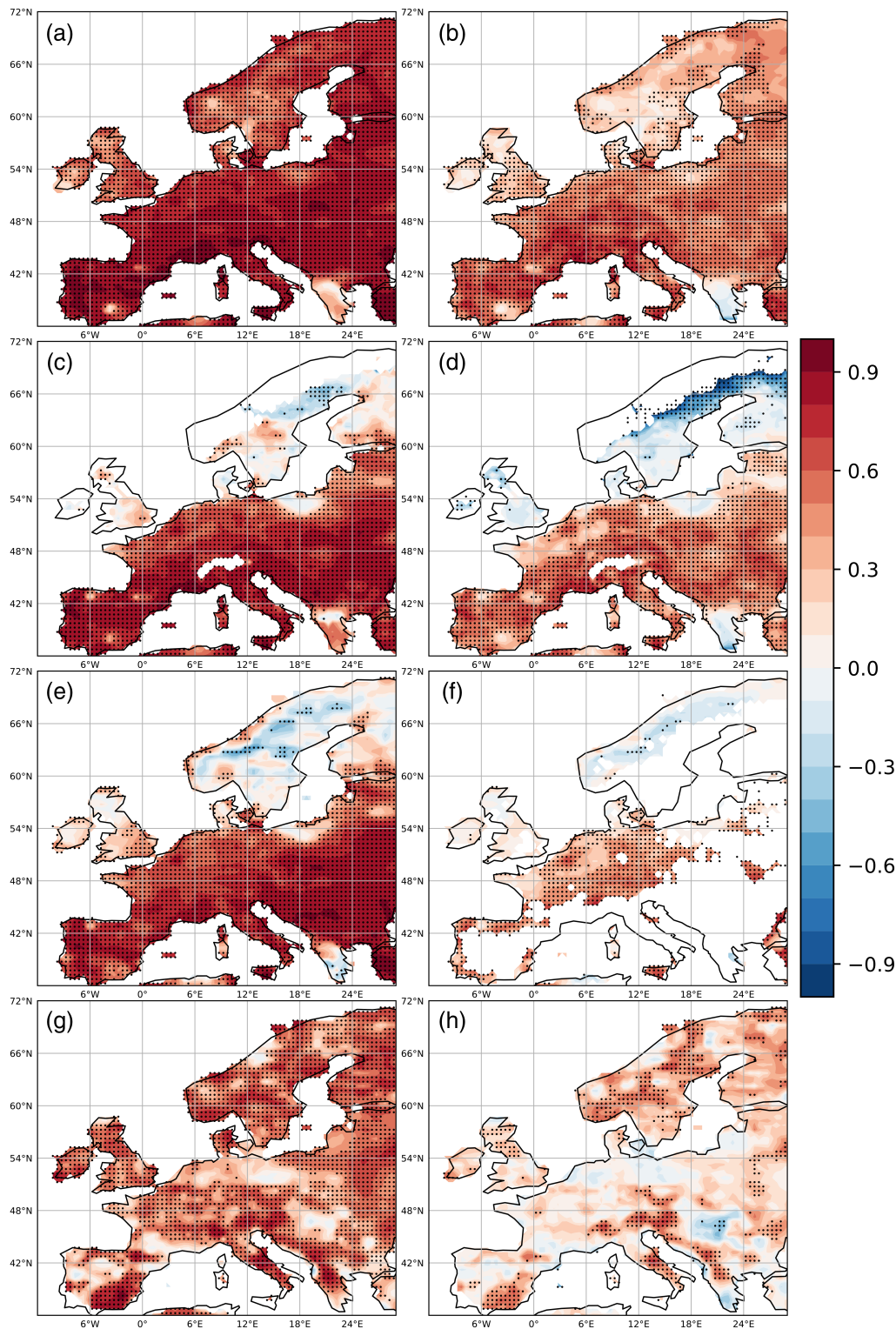


FIGURE 2 Spatial distribution of ACC (first column) and RPSS (second column) for lead year 2–5 for daily maximum summer temperature [TASMAX; (a), (b)], summer days [SU; (c), (d)], heatwave days [HWDS; (e), (f)], and frost days [FD; (g), (h)]. Reference is the climatology. Black dots indicate significant skill at the 95% level

significant for large parts of Southern, Central and Eastern Europe (ACC for LT2-5 up to 0.8), while less skill is found for Scandinavia and the British Isles. As for

TASMAX, the skill scores vary only slightly with the choice of lead-time (Figure S2). The chosen index has some obvious limitations: It is not meaningful in regions

where the temperatures seldom exceed the fixed threshold (e.g., Scandinavia) and grid points are often marked by missing values after recalibration. In these areas, percentile-based thresholds and indices may be more appropriate. Nevertheless, the recalibration is able to increase the overall predictive skill of SU (Figure S3), mainly by adjusting the hindcast trend to the observed one (Figures S9–S10).

A more complex phenomenon related to warm temperature extremes is heat waves. So far, we analysed the predictive skill for the number of heat wave days per year (HWDS; Table 1). The skill pattern of ACC (Figure 2e) shows a North–South-gradient, similar to the spatial patterns of SU: Correlations are positive, significant for Southern, Central and Eastern Europe, and negative for Scandinavia and the British Isles. In general, skill scores for HWDS are lower compared to SU or TASMAY.

Finally, frost days (FD, minimum temperature below 0°C; Figures 2g, h) were also analysed as an example related to cold temperature extremes, revealing skill for large parts of Europe. Compared to warm temperature extremes, skill scores are lower (ACC of up to 0.7) for Southern, Central, and Eastern Europe, but higher (and positive) for Northern Europe. While the ACC is significant for large parts of Europe, the RPSS shows significant skill only for Scandinavia and parts of Italy and Spain.

The results hint at the potential of the regional decadal system to predict not only mean temperatures but also the likelihood of both cold and warm temperature extremes, especially when considering the

climatology as reference. Skill scores using the historical ensemble as reference are generally lower and less significant (Figures S4–S8). Nevertheless, large parts of Europe exhibit an added value of initialisation, particularly for warm temperature extremes. For all indices, the skill scores vary only slightly with the choice of lead-time (-Figure S2), leading to the conclusion that the skill is mostly attributed to the forcing and not the initialisation. Results for tropical nights (TR) and ice days (ID) are comparable to those of SU and FD, but with an increased fraction of missing values due to the higher thresholds (not shown).

3.2 | Precipitation-related variables

Regarding precipitation-related variables, expectations are generally lower than for temperature-derived parameters (cp. Mieruch *et al.*, 2014 or Reyers *et al.*, 2019 for mean precipitation). In this section, we focus on heavy precipitation events, as extremes are more relevant for stakeholders than mean values. Figure 3a shows the ACC for heavy precipitation days (R10mm; Table 1) for LT2-5. The skill pattern is more heterogeneous compared to temperature-related variables. Correlations are positive (up to 0.8) and significant for Scandinavia and parts of Central and Southern Europe. Again, spatial patterns of the RPSS look similar, but with slightly lower values (Figure 3b). Similar results can be found for a percentile-based precipitation index aiming at the identification of

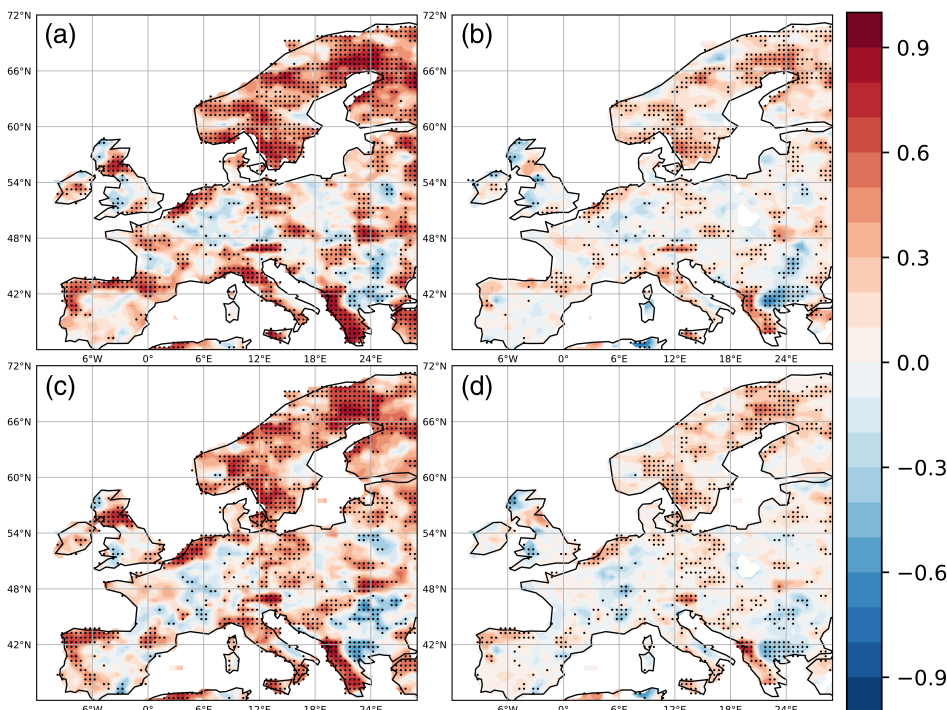


FIGURE 3 Spatial distribution of ACC (first column) and RPSS (second column) for lead year 2–5 for heavy precipitation days [R10mm; (a), (b)] and flood index [RM7P95; (c), (d)]. Reference is the climatology. Black dots indicate significant skill at the 95% level

precipitation events potentially causing floods in major European river catchments (RM7P95; Table 1). Both the ACC and the RPSS (Figures 3c, d) are comparable to those of R10mm.

The regional decadal prediction system is less skilful for variables related to precipitation than for temperature-based indices. However, results are promising for several European regions (e.g., Scandinavia or the Mediterranean), where the hindcasts beat not only the climatology, but also the uninitialized ensemble (Figures S4–S8). The recalibration improves the predictive skill and reduces the lead-time dependence for both R10mm and RM7P95 (Figures S2 and S3). Skill scores for other precipitation-related quantities (Table 1) like very heavy precipitation days (R20mm) or simple daily intensity index (SDII) are similar to the results presented above (not shown).

3.3 | Wind-related variables

For wind, only a few studies focus on decadal predictability. For example, Haas *et al.* (2016) and Moemken *et al.* (2016) found some predictive for wind gust and wind energy potentials over Central Europe, while Kruschke *et al.* (2016) found some skill for Northern hemispheric cyclones. Here, we extend these studies and analyse further potentially user-relevant wind-related variables. The first example is the mean surface wind speed in the extended winter season (ONDJFM). This

period is of interest as, for example, the probability of windstorms affecting Central Europe is higher than in summer (e.g., Donat *et al.*, 2010). The ACC for LT2-5 (Figure 4a) is positive (up to 0.6) for the Mediterranean region, the North and the Baltic Sea, and parts of Central Europe. However, the skill is not significant for most of the grid points. The RPSS shows larger regions with negative skill (Figure 4b). For both skill metrics, the skill averaged over Europe decreases for longer lead-times in spite of the recalibration (e.g., LT6-9; Figure S2).

We also estimated the predictive skill of a simplified storm severity index (e.g., Pinto *et al.*, 2012) for the extended winter season (SFCWIND98W; Table 1). Both ACC and RPSS (Figures 4c, d) show positive skill scores for the Mediterranean region, Eastern Europe and parts of North and Baltic Sea, with highest values (ACC up to 0.8) for the Adriatic Sea. However, only the ACC values are significant in certain regions.

In general, variables related to extreme wind speeds are less skilful than temperature or precipitation-based indices when using the climatology as reference. Nevertheless, some European regions show predictive skill. In addition, wind-related variables exhibit a stronger lead-time dependence (Figure S2). The applied recalibration method is able to enhance the predictive skill (Figure S3) and even turns the on average negative ACC values into slightly positive ones. Results look different if the historical ensemble is considered. Here, both indices show a significant added value of initialisation for several European

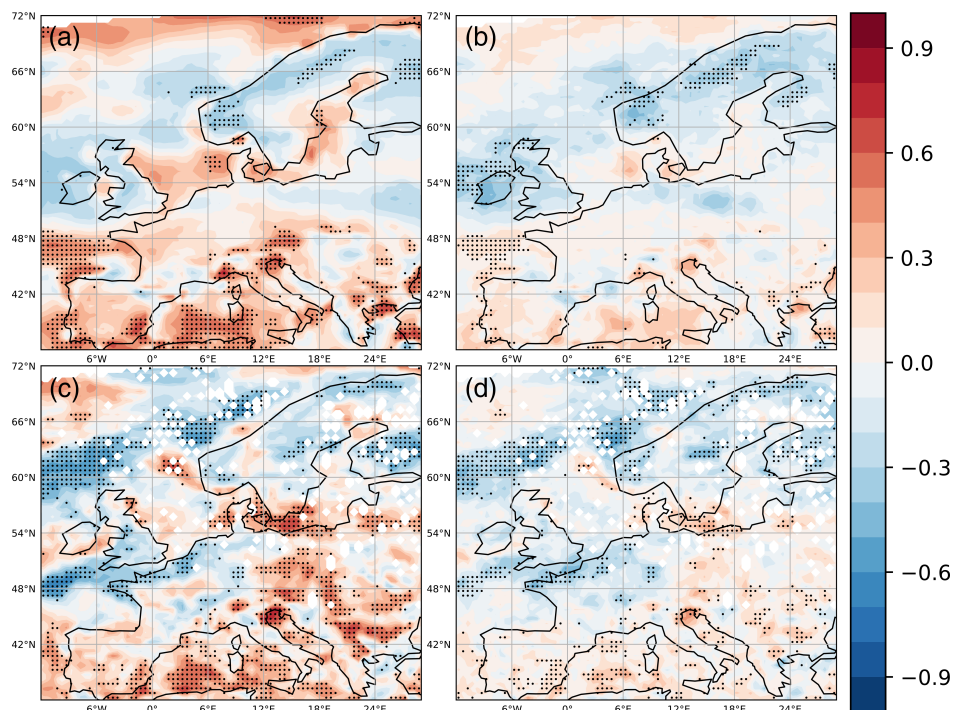


FIGURE 4 Spatial distribution of ACC (first column) and RPSS (second column) for lead year 2–5 for winter wind speed [SFCWIND; (a), (b)] and winter storm days [SFCWIND98W; (c), (d)]. Reference is the climatology. Black dots indicate significant skill at the 95% level

regions (like Mid-Europe or the British Isles), especially for SFCWIND (Figures S4–S8).

3.4 | Indices related to the agronomy sector

Finally, we analysed two indices relevant for the agronomy sector, namely, the length of the growing season (GSL) and the growing degree days (GDD; Table 1). Both parameters are based on temperature. Therefore, a certain predictive skill can be expected. Nevertheless, they focus on different aspects compared to the variables analysed in Section 3.1. GSL is often used to determine which crops can be grown in a specific region. With this climate indicator, we analyse if shifts in the timing and length of the growing season show predictive skill on decadal timescales. Figure 5a depicts the ACC of GSL for LT2-5. Correlations are positive (up to 0.6) and mostly significant for Northern, Central, and Eastern Europe. GSL shows no skill for parts of France, Spain, and the British Isles. RPSS values are lower and less significant, with the negative skill extending towards Eastern and Northern Europe (Figure 5b).

GDD addresses the integrated temperature over the summer half-year and is a frequently used measure for describing and predicting the growth and development processes of various crops. In general, the predictive skill is much higher for GDD than for GSL. Both ACC and RPSS (Figures 5c, d) show high positive (up to 0.9 for

ACC) and significant values for Southern, Central, and Eastern Europe. Except for Scandinavia, the spatial pattern is comparable to TASMAY. This region exhibits partly negative skill scores, especially for RPSS and longer lead-times.

The regional decadal prediction system provides promising results for parameters relevant for the agronomy sector (see also heating degree days [HD] in Feldmann *et al.*, 2019). Skill scores are comparable to both mean and extreme temperatures, irrespective of the applied reference dataset (Figures S4–S8). Additionally, the skill scores show practically no lead-time dependence when averaged over Europe (Figure S2). In contrast to other presented variables in this study, both GSL and GDD do not seem to benefit from the applied recalibration (Figure S3).

4 | SUMMARY AND DISCUSSION

In this study, we assess the potential for decadal predictability for user-oriented variables and extremes using the regional MiKlip decadal prediction system. This effort extends previous work, which focussed on basic variables (Feldmann *et al.*, 2019; Reyers *et al.*, 2019). The main conclusions are (see also Figure 6):

- All temperature-related indices show high-predictive skill (mean ACC for Europe up to 0.8) for all lead-times, indicating that skill extends beyond mean

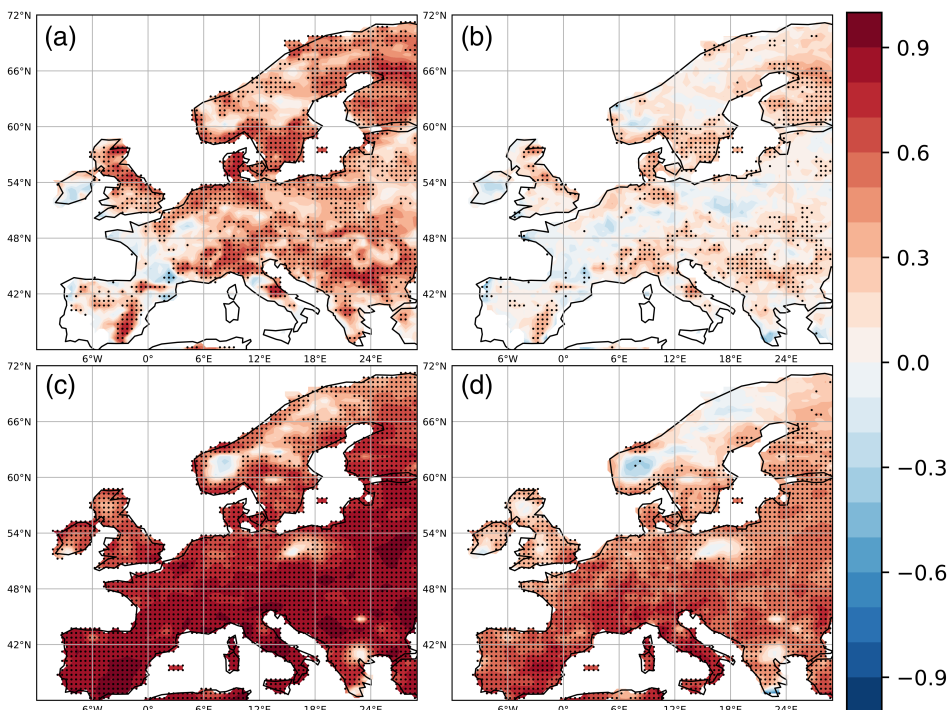
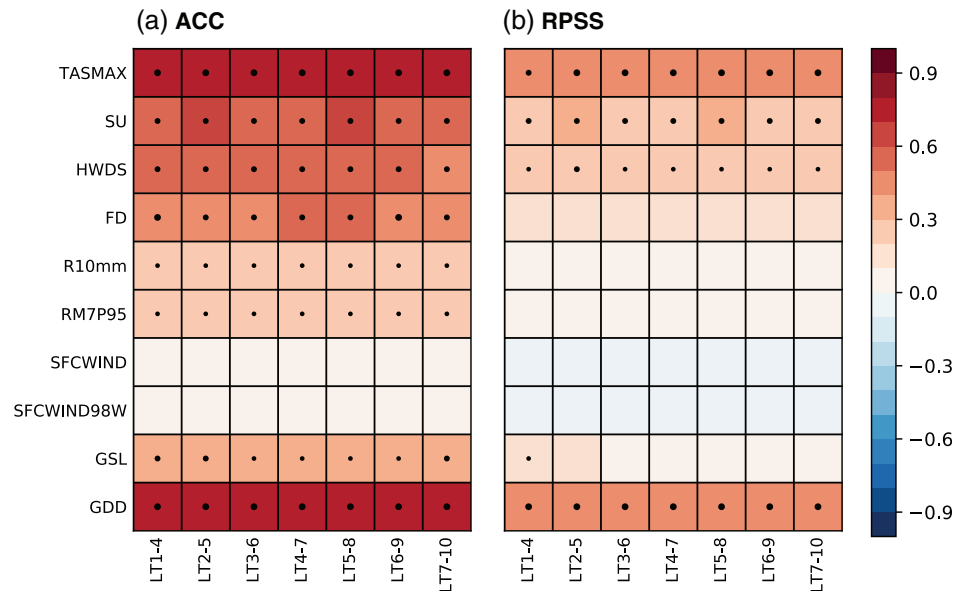


FIGURE 5 Spatial distribution of ACC (first column) and RPSS (second column) for lead year 2–5 for growing season length [GSL; (a), (b)] and growing degree days [GDD; (c), (d)]. Reference is the climatology. Black dots indicate significant skill at the 95% level

FIGURE 6 Skill scores averaged over Europe divided by index and lead-time: (a) ACC, and (b) RPSS. Reference is the climatology. Black dots indicate the area fraction of significant grid points: No marker if less than 25% are significant, small dots if 25–50% are significant, medium-sized dots if 50–75% are significant and large dots if more than 75% are significant



temperature. The skill patterns are spatially robust and often significant (Figure 6). In general, the predictive skill is higher for variables related to summer and high temperatures (in line with Mieruch *et al.*, 2014; Feldmann *et al.*, 2019). The predictive skill of TASMEX is comparable to that of daily mean temperature (Feldmann *et al.*, 2019). The variables SU and HWDS address more heat-related aspects, and show high predictive skill from Central to Southern Europe, where the MiKlip system might provide useful information on the expected heat stress for a given period. Notably, HWDS, which indicates the likely duration of heat waves, hints at the potential for valuable climate information for users. The skill of SU and HWDS is lower for the British Isles and Scandinavia, where a lower number of hot days and thus lower heat stress is generally found, making this type of climate information less valuable. For cold extremes, FD shows a lower but still significant predictive skill for most of Europe (Figure 6). This information might be valuable, for example, for de-icing activities (Schipper *et al.*, 2019).

- As expected, results are more mixed for precipitation-based parameters: The spatial skill patterns are more heterogeneous and the skill scores are generally lower (mean ACC for Europe up to 0.3; Figure 6). This is in line with findings by Mieruch *et al.* (2014) and Reyers *et al.* (2019) for mean precipitation in a sub-sample of the full-regional MiKlip ensemble. The selected indices (R10mm and RM7P95) address more extreme precipitation events and might thus be useful to users, as such events may, for example, trigger floods in large river catchments.
- Skill patterns are spatially diverse and often not significant for the analysed wind-related indices. Skill scores

are highest for Southern and parts of Eastern Europe, but generally lower compared to both temperature and precipitation (mean ACC for Europe up to 0.1; Figure 6). Reyers *et al.* (2019) found similar results (with lower values) for mean wind speed in a sub-sample of the regional system. This is in partial contradiction with previous studies on wind energy potentials and wind gusts (Haas *et al.*, 2016; Moemken *et al.*, 2016), which had found higher predictive skill for Central Europe with statistical-dynamical approaches and for shorter lead-times (LT1-4). However, their results exhibited a pronounced lead-time dependence.

- Results look promising for indices connected to the agronomy sector, as skill scores are high and often significant for large parts of Europe (mean ACC up to 0.75; Figure 6). The skill patterns are comparable to those of temperature extremes. Indicators like GDD are calculated as integrals over longer periods and might therefore compensate potential timing errors in the annual cycle. Such integrated indices might thus provide more robust information than threshold-based indices.
- The majority of the discussed indices show a negligible lead-time dependence in their predictive skill (Figure 6), especially after recalibration. Feldmann *et al.* (2019) concluded that in some regions there is a superposition of skill due to (a) the initialisation in the first years and (b) the long-term climate trend in later years (their Figure 5). Particularly for temperature-based indices, the high skill originates mainly from the strong trend and neither from the represented variability nor from the initialisation (see Figures S9–S10).

- The applied recalibration approach is generally able to improve the skill of the regional decadal prediction system. The level of improvement depends on the variable and region. This is in line with results by Feldmann *et al.* (2019) for the surface temperature at the regional scale. On the global scale, DeFoReSt improves the prediction skill for the surface temperature (Pasternack *et al.*, 2018) as well as the GPCC drought index (Paxian *et al.*, 2019). Regarding the different components of the recalibration method, the hindcasts seem to benefit mostly from the adjustment of their trend towards the observed one (Figures S9–S10).
- The focus on the climatology as reference is the most common choice in climate services. In general, skill scores are lower and less significant when considering historical simulations as reference (Figures S4–S8). Nevertheless, several indices analysed here (e.g., SFCWIND or RM7P95) show a clear added value of initialisation for various regions in Europe. Future work could extend the current study by using other widely used references as, for example, persistence forecasts or statistical models (e.g., Suckling *et al.*, 2017).

We conclude that the regional MiKlip decadal prediction system has potential for several user-oriented variables and climate extremes (Figure 6). The predictive skill depends on both the variable and the region. While the temperature-based indices are skilful for most of Europe, significant predictive skill for precipitation- and wind-related variables is limited to certain European regions. These regions include Scandinavia and large parts of the European coastal areas for precipitation-based indices, and the Mediterranean and parts of Eastern Europe for wind extremes. The regional differences affect the applicability of the decadal prediction system. Nevertheless, keeping these limitations in mind, the MiKlip system can provide valuable information for users outside the science community, for example, decision makers in politics, economy or society. Potential areas of application for temperature extremes are the health sector (action plans to reduce mortality), the humanitarian risk reduction (forecast-based financing to reduce effects of events) or the IT sector (heat dependent server outages, cooling systems). For precipitation extremes, fields of application may be the hydrology sector (water management) or inland shipping (river water levels), while the predictability of wind extremes might be important for forestry (removal of storm damages) or insurance companies (wind damages to buildings).

Our results illustrate in general terms how the utility of decadal predictions can be enhanced substantially with

little additional effort. It is often not necessary to drive an impact model with climate-model output to establish the predictive skill for a climate-impact variable. Instead, it suffices to establish the skill for a climate index that is empirically known to relate to important applications (Table 1). This strategy may well require defining climate indices that are not a priori obvious from a climate-physics standpoint. But once an index is established, investigating its predictive skill from an existing suite of climate predictions is straightforward.

The development of custom-tailored (user-driven) indices that show decadal predictive skill and the development of more specific user products in close cooperation with individual stakeholders is strongly needed and has already started within MiKlip and other projects (e.g., Falloon *et al.*, 2018; Paxian *et al.*, 2019). However, it has to be differentiated between analysing the predictive skill of user-oriented variables and the actual applicability of information from decadal predictions. In order to create valuable information based on decadal predictions, that is, the information is integrated in the established workflow of stakeholders, their features need to be specified from the users' (and not the scientific) point of view (Buontempo *et al.*, 2019). These features include beneficial skill measures, prediction types (ensemble mean, probabilistic predictions), temporal or spatial requirements (lead-time, reference period, temporal aggregation, area of interest), reference datasets, and information formats (numerical, graphical, text). To succeed in identifying what drives the interest of potential users in the rather unknown field of decadal predictions, work must focus on cooperative communication, mutual understanding, and trust building between users and scientists. Nevertheless, the present work is an indispensable first step towards pushing the science of decadal predictability to actual "real world" applicability.

ACKNOWLEDGEMENTS

The research programme MiKlip is funded by the German ministry of education and research (BMBF, contract numbers 01LP1518A/D, 01LP1519A-D). The MPI ensemble simulations were provided by the MPI Hamburg, Germany. All simulations were performed at the German Climate Computing Center (DKRZ). We thank the developers of the MiKlip CES for their efforts and support. We acknowledge the E-OBS data set from the EU-FP6 project ENSEMBLES (<http://ensembles-eu.metoffice.com>) and the data providers in the ECA&D project (<http://www.ecad.eu>). We thank the anonymous reviewers for their helpful comments. Pinto thanks the AXA Research Fund for support. Open access funding enabled and organized by Projekt DEAL.


AUTHOR CONTRIBUTIONS

Moemken, Feldmann, and Pinto conceived and designed the study, and wrote the original paper draft. Moemken, Feldmann, Buldmann, and Laube performed the data analysis. Kadow contributed with software and analysis tools. Paxian and Tiedje organized the user workshops and gathered information on the user requirements for tailored products. All authors discussed the results and contributed with manuscript revisions.

ORCID

Julia Moemken  <https://orcid.org/0000-0002-9432-0202>

Hendrik Feldmann  <https://orcid.org/0000-0001-6987-7351>

Joaquim G. Pinto  <https://orcid.org/0000-0002-8865-1769>

Christoph Kottmeier  <https://orcid.org/0000-0002-2196-2052>

Jochem Marotzke  <https://orcid.org/0000-0001-9857-9900>

REFERENCES

- Balmaseda, M.A., Mogensen, K. and Weaver, A.T. (2013) Evaluation of the ECMWF Ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, 139, 1132–1161. <https://doi.org/10.1002/qj.2063>.
- Bellucci, A., Haarsma, R., Gualdi, S., Athanasiadis, P.J., Caian, M., Cassou, C., Fernandez, E., Germe, A., Jungclaus, J., Kröger, J., Matei, D., Müller, W., Pohlmann, H., Salas y Melia, D., Sanchez, E., Smith, D., Terray, L., Wyser, K. and Yang, S. (2015) An assessment of a multi-model ensemble of decadal climate predictions. *Climate Dynamics*, 44, 2787–2806. <https://doi.org/10.1007/s00382-014-2164-y>.
- Boer G.J., Smith D.M., Cassou C., Doblas-Reyes F., Danabasoglu G., Kirtman B., Kushnir Y., Kimoto M., Meehl G. A., Msadek R., Mueller W.A., Taylor K.E., Zwiers F., Rixen M., Ruprich-Robert Y., Eade R. (2016) The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. *Geoscientific Model Development*, 9, 3751–3777. <http://dx.doi.org/10.5194/gmd-9-3751-2016>.
- Buontempo, C., Hanlon, H.M., Bruno Soares, M., Christel, I., Soubeyroux, J.M., Viel, C., Calmanti, S., Bosi, L., Falloon, P., Palin, E.J., Vanvyve, E., Torralba, V., Gonzalez-Reviriego, N., Doblas-Reyes, F., Pope, E.C.D., Newton, P. and Liggins, F. (2018) What have we learnt from EUPORIAS climate service prototypes? *Climate Services*, 9, 21–32. <https://doi.org/10.1016/j.cliser.2017.06.003>.
- Buontempo, C., Hutjes, R., Beavis, P., Berckmans, J., Cagnazzo, C., Vamborg, F., Thépaut, J.N., Bergeron, C., Almond, S., Amici, A., Ramasamy, S. and Dee, D. (2019) Fostering the development of climate services through Copernicus climate change service (C3S) for agricultural applications. *Weather and Climate Extremes*, 27, 100226. <https://doi.org/10.1016/j.wace.2019.100226>.
- Chen, X. and Tung, K. (2018) Global-mean surface temperature variability: space-time perspective from rotated EOFs. *Climate Dynamics*, 51, 1719–1732. <https://doi.org/10.1007/s00382-017-3979-0>.
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N. and Vitart, F. (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597. <https://doi.org/10.1002/qj.828>.
- Doblas-Reyes, F.J., Andreu-Burillo, I., Chikamoto, Y., Garcia-Serrano, J., Guemas, V., Kimoto, M., Mochizuki, T., Rodrigues, L.R.L. and van Oldenborgh, G.J. (2013) Initialized near-term regional climate change prediction. *Nature Communications*, 4, 1715. <https://doi.org/10.1038/ncomms2704>.
- Donat, M.G., Leckebusch, G.C., Pinto, J.G. and Ulbrich, U. (2010) Examination of wind storms over Central Europe with respect to circulation weather types and NAO phases. *International Journal of Climatology*, 30, 1289–1300. <https://doi.org/10.1002/joc.1982>.
- Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J. and Taylor, K.E. (2016) Overview of the coupled model Intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9, 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>.
- Falloon, P., Soares, M.B., Manzanar, R., San-Martin, D., Liggins, F., Taylor, I., Kahana, R., Wilding, J., Jones, C., Comer, R., de Vreede, E., Som de Cerff, W., Buontempo, C., Brookshaw, A., Stanley, S., Middleham, R., Pittams, D., Lawrence, E., Bate, E., Peter, H., Uzell, K. and Richards, M. (2018) The land management tool: developing a climate service in Southwest UK. *Climate Services*, 9, 86–100. <https://doi.org/10.1016/j.cliser.2017.08.002>.
- Feldmann, H., Pinto, J.G., Laube, N., Uhlig, M., Moemken, J., Früh, B., Pohlmann, H., Pasternack, A. and Kottmeier, C. (2019) Skill and added value of the MiKlip regional decadal prediction system for temperature over Europe. *Tellus A*, 71, 1618678. <https://doi.org/10.1080/16000870.2019.1618678>.
- Ferro, C.A.T. (2014) Fair scores for ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, 140, 1917–1923. <https://doi.org/10.1002/qj.2270>.
- Fischer, E.M. and Schär, C. (2010) Consistent geographical patterns of changes in high-impact European heatwaves. *Nature Geoscience*, 3, 398–403. <https://doi.org/10.1038/NNGEO866>.
- Flato, G., et al. (2013) Evaluation of climate models. In: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (Eds.) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press.
- Fraga, H., Santos, J.A., Malheiro, A.C., Oliveira, A.A., Moutinho-Pereira, J. and Jones, G.V. (2016) Climatic suitability of Portuguese grapevine varieties and climate change adaptation. *International Journal of Climatology*, 36, 1–12. <https://doi.org/10.1002/joc.4325>.

- Ghosh, R., Müller, W.A., Baehr, J. and Bader, J. (2017) Impact of observed North Atlantic multidecadal variations to European summer climate: a linear baroclinic response to surface heating. *Climate Dynamics*, 48, 3547–3563. <https://doi.org/10.1007/s00382-016-3283-4>.
- Giorgetta, M.A., Jungclaus, J., Reick, C.H., Legutke, S., Bader, J., Böttlinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K., Glushak, K., Gayler, V., Haak, H., Hollweg, H.D., Ilyina, T., Kinne, S., Kornbluh, L., Matei, D., Mauritsen, T., Mikolajewicz, U., Mueller, W., Notz, D., Pithan, F., Raddatz, T., Rast, S., Redler, R., Roeckner, E., Schmidt, H., Schnur, R., Segschneider, J., Six, K.D., Stockhause, M., Timmreck, C., Wegner, J., Widmann, H., Wieners, K.H., Claussen, M., Marotzke, J. and Stevens, B. (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the coupled model Intercomparison project phase 5. *Journal of Advances in Modeling Earth Systems*, 5, 572–597. <https://doi.org/10.1002/jame.20038>.
- Giorgi, F., Jones, C. and Asrar, G.R. (2009) Addressing climate information needs at the regional level: the CORDEX framework. *Bulletin of the World Meteorological Organization*, 58, 175–183.
- Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., Kharin, V., Merryfield, W., Deser, C., Mason, S.J., Kirtman, B.P., Msadek, R., Sutton, R., Hawkins, E., Fricker, T., Hegerl, G., Ferro, C.A.T., Stephenson, D.B., Meehl, G.A., Stockdale, T., Burgman, R., Greene, A.M., Kushnir, Y., Newman, M., Carton, J., Fukumori, I. and Delworth, T. (2013) A verification framework for interannual-to-decadal prediction experiments. *Climate Dynamics*, 40, 245–272. <https://doi.org/10.1007/s00382-012-1481-2>.
- Haas, R., Reyers, M. and Pinto, J.G. (2016) Decadal predictability of regional-scale peak winds over Europe using the earth system model of the max-Planck-Institute for Meteorology. *Meteorologische Zeitschrift*, 25, 739–752. <https://doi.org/10.1127/metz/2015/0583>.
- Hackenbruch, J., Kunz-Plapp, T., Müller, S. and Schipper, J.W. (2017) Tailoring climate parameters to information needs for local adaptation to climate change. *Climate*, 5, 25. <https://doi.org/10.3390/cli5020025>.
- Hartmann, D.L., et al. (2013) Observations: Atmosphere and Surface. In: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P.M. (Eds.) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY: Cambridge University Press.
- Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D. and New, M. (2008) A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research*, 113, D20119. <https://doi.org/10.1029/2008JD010201>.
- Honda, Y., Kabuto, M., Ono, M. and Uchiyama, I. (2007) Determination of optimum daily maximum temperature using climate data. *Environmental Health and Preventive Medicine*, 12, 209–216. <https://doi.org/10.1265/ehpm.12.209>.
- ICPO. (2011) Data and bias correction for decadal climate predictions. *CLIVAR Publication Series*, 150, 6pp.
- Illing, S., Kadow, C., Kunst, O. and Cubasch, U. (2014) MurCSS: a tool for standardized evaluation of decadal Hindcast systems. *Journal of Open Research Software*, 2 1–3. <https://doi.org/10.5334/jors.bf>.
- Kadow, C., Illing, S., Kröner, I., Ulbrich, U. and Cubasch, U. (2017) Decadal climate predictions improved by ocean ensemble dispersion filtering. *Journal of Advances in Modeling Earth Systems*, 9, 1138–1149. <https://doi.org/10.1002/2016MS000787>.
- Kadow, C., Illing, S., Kunst, O., Rust, H.W., Pohlmann, H., Müller, W.A. and Cubasch, U. (2016) Evaluation of forecasts by accuracy and spread in the MiKlip decadal climate prediction system. *Meteorologische Zeitschrift*, 25, 631–643. <https://doi.org/10.1127/metz/2015/0639>.
- Kharin, V.V., Boer, G.J., Merryfield, W.J., Scinocca, J.F. and Lee, W.S. (2012) Statistical adjustment of decadal predictions in a changing climate. *Geophysical Research Letters*, 39, L19705. <https://doi.org/10.1029/2012GL052647>.
- Kim, H.M., Webster, P.J. and Curry, J.A. (2012) Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophysical Research Letters*, 39, L10701. <https://doi.org/10.1029/2012GL051644>.
- Kruschke, T., Rust, H.W., Kadow, C., Müller, W.A., Pohlmann, H., Leckebusch, G.C. and Ulbrich, U. (2016) Probabilistic evaluation of decadal predictions of northern hemisphere winter storms. *Meteorologische Zeitschrift*, 25, 721–738. <https://doi.org/10.1127/metz/2015/0641>.
- Latif, M. and Keenlyside, N.S. (2011) A perspective on decadal climate variability and predictability. *Deep-Sea Research II*, 58 (17–18), 1880–1894. <https://doi.org/10.1016/j.dsr2.2010.10.066>.
- Lemos, M.C., Kirchhoff, C.J. and Ramprasad, V. (2012) Narrowing the climate information usability gap. *Nature Climate Change*, 2, 789–794. <https://doi.org/10.1038/NCLIMATE1614>.
- Marotzke, J., Müller, W.A., Vamborg, F.S.E., Becker, P., Cubasch, U., Feldmann, H., Kaspar, F., Kottmeier, C., Marini, C., Polkova, I., Prömmel, K., Rust, H.W., Stammer, D., Ulbrich, U., Kadow, C., Köhl, A., Kröger, J., Kruschke, T., Pinto, J.G., Pohlmann, H., Reyers, M., Schröder, M., Sienz, F., Timmreck, C. and Ziese, M. (2016) MiKlip: a National Research Project on decadal climate prediction. *Bulletin of the American Meteorological Society*, 97, 2379–2394. <https://doi.org/10.1175/BAMS-D-15-00184.1>.
- Matei, D., Pohlmann, H., Jungclaus, J., Müller, W., Haak, H. and Marotzke, J. (2012) Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model. *Journal of Climate*, 25, 8502–8523. <https://doi.org/10.1175/JCLI-D-11-00633.1>.
- McMaster, G.S. and Wilhelm, W.W. (1997) Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, 87, 291–300. [https://doi.org/10.1016/S0168-1923\(97\)00027-0](https://doi.org/10.1016/S0168-1923(97)00027-0).
- Meehl, G.A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas-Reyes, F., Hawkins, E., Karspeck, A., Kimoto, M., Kumar, A., Matei, D., Mignot, J., Msadek, R., Navarra, A., Pohlmann, H., Rienecker, M., Rosati, T., Schneider, E., Smith, D., Sutton, R., Teng, H., van Oldenborgh, G.J., Vecchi, G. and Yeager, S. (2014) Decadal climate prediction: an update from the trenches. *Bulletin of the American Meteorological Society*, 95, 243–267. <https://doi.org/10.1175/BAMS-D-12-00241.1>.
- Meehl, G.A., Goddard, L., Murphy, J., Stouffer, R.J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M.A., Greene, A.M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D.,

- Stammer, D. and Stockdale, T. (2009) Decadal prediction – can it be skilful? *Bulletin of the American Meteorological Society*, 90, 1467–1485. <https://doi.org/10.1175/2009BAMS2778.1>.
- Meehl, G.A., Hu, A. and Teng, H. (2016) Initialized decadal prediction for transition to positive phase of the Interdecadal Pacific oscillation. *Nature Communications*, 7, 11718. <https://doi.org/10.1038/ncomms11718>.
- Mieruch, S., Feldmann, H., Schädler, G., Lenz, C.J., Kothe, S. and Kottmeier, C. (2014) The regional MiKlip decadal forecast Ensemble for Europe: the added value of downscaling. *Geoscience Model Development*, 7, 2983–2999. <https://doi.org/10.5194/gmd-7-2983-2014>.
- Moemken, J., Reyers, M., Buldmann, B. and Pinto, J.G. (2016) Decadal predictability of regional scale wind speed and wind energy potentials over Central Europe. *Tellus A*, 68, 29199. <https://doi.org/10.3402/tellusa.v68.29199>.
- Müller, W.A., Baehr, J., Haak, H., Jungclaus, J.H., Kröger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J.S. and Marotzke, J. (2012) Forecast skill of multi-year seasonal means in the decadal prediction system of the max Planck Institute for Meteorology. *Geophysical Research Letters*, 39, L22707. <https://doi.org/10.1029/2012GL053326>.
- Müller, W.A., Pohlmann, H., Sienz, F. and Smith, D. (2014) Decadal climate predictions for the period 1901–2010 with a coupled climate model. *Geophysical Research Letters*, 41, 2100–2107. <https://doi.org/10.1002/2014GL059259>.
- Pasternack, A., Bhend, J., Liniger, M.A., Rust, H.W., Müller, W.A. and Ulbrich, U. (2018) Parametric decadal climate forecast recalibration (DeFoReSt 1.0). *Geoscientific Model Development*, 11, 351–368. <https://doi.org/10.5194/gmd-11-351-2018>.
- Paxian, A., Ziese, M., Kreienkamp, F., Pankatz, K., Brand, S., Pasternack, A., Pohlmann, H., Modali, K. and Früh, B. (2019) User-oriented global predictions of the GPCP drought index for the next decade. *Meteorologische Zeitschrift*, 28, 3–21. <https://doi.org/10.1127/metz/2018/0912>.
- Pinto, J.G., Karremann, M.K., Born, K., Della-Marta, P.M. and Klawa, M. (2012) Loss potentials associated with European windstorms under future climate conditions. *Climate Research*, 54, 1–20. <https://doi.org/10.3354/cr01111>.
- Pohlmann, H., Müller, W.A., Kulkarni, K., Kameswarrao, M., Matei, D., Vamborg, F.S.E., Kadow, C., Illing, S. and Marotzke, J. (2013b) Improved forecast skill in the tropics in the new MiKlip decadal climate predictions. *Geophysical Research Letters*, 40, 5798–5802. <https://doi.org/10.1002/2013GL058051>.
- Pohlmann, H., Smith, D.M., Balmaseda, M.A., Keenlyside, N.S., Masina, S., Matei, D., Müller, W.A. and Rogel, P. (2013a) Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system. *Climate Dynamics*, 41, 775–785. <https://doi.org/10.1007/s00382-013-1663-6>.
- Reyers, M., Feldmann, H., Mieruch, S., Pinto, J.G., Uhlig, M., Ahrens, B., Früh, B., Modali, K., Laube, N., Moemken, J., Müller, W., Schädler, G. and Kottmeier, C. (2019) Development and prospects of the regional MiKlip decadal prediction system over Europe: predictive skill, added value of regionalization, and ensemble size dependency. *Earth System Dynamics*, 10, 171–187. <https://doi.org/10.5194/esd-10-171-2019>.
- Reyers, M., Pinto, J.G. and Moemken, J. (2015) Statistical-dynamical downscaling for wind energy potentials: evaluation and applications to decadal hindcasts and climate change projections. *International Journal of Climatology*, 35, 229–244. <https://doi.org/10.1002/joc.3975>.
- Rockel, B., Will, A. and Hense, A. (2008) Special issue: regional climate modelling with COSMO-CLM (CCLM). *Meteorologische Zeitschrift*, 17, 347–348. <https://doi.org/10.1127/0941-2948/2008/0309>.
- Romanova, V., Hense, A., Wahl, S., Brune, S. and Baehr, J. (2018) Skill assessment of different ensemble generation schemes for retrospective predictions of surface freshwater fluxes on inter and multi-annual timescales. *Meteorologische Zeitschrift*, 27, 111–124. <https://doi.org/10.1127/metz/2017/0790>.
- Schipper, J.W., Hackenbruch, J., Lentink, H.S. and Sedlmeier, K. (2019) Integrating adaptation expertise into regional climate simulations through tailored climate parameters. *Meteorologische Zeitschrift*, 28, 41–57. <https://doi.org/10.1127/metz/2019/0878>.
- Suckling, E.B., van Oldenborgh, G.J., Eden, J.M. and Hawkins, E. (2017) An empirical model for probabilistic decadal prediction: global attribution and regional hindcasts. *Climate Dynamics*, 48, 3115–3138. <https://doi.org/10.1007/s00382-016-3255-8>.
- Sutton, R.T. and Hodson, D.L.R. (2005) Atlantic Ocean forcing of north American and European summer climate. *Science*, 309, 115–118. <https://doi.org/10.1126/science.1109496>.
- Taylor, K.E., Stouffer, R.J. and Meehl, G.A. (2012) An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93, 485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D.C., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A.C.M., Berg, L.V.D., Bidlot, J., Bormann, N., Cairnes, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, I., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.F., Morcrette, J.J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P. and Woollen, J. (2005) The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131, 2961–3012. <https://doi.org/10.1256/qj.04.176>.
- Van Engelen, A., Klein Tank, A., van der Schrier, G. and Klok, L. (2008) *European Climate Assessment & Dataset (ECA&D), Report 2008*. De Bilt, Netherlands: KNMI. https://www.ecad.eu/documents/ECAD_report_2008.pdf.
- Vera, C., Barange, M., Dube, O.P., Goddard, L., Griggs, D., Kobysheva, N., Odada, E., Parey, S., Polovina, J., Poveda, G., Seguin, B. and Trenberth, K. (2010) Needs assessment for climate information on decadal time scales and longer. *Procedia Environmental Sciences*, 1, 275–286. <https://doi.org/10.1016/j.proenv.2010.09.017>.
- Wilks, D.S. (2011) *Statistical Methods in the Atmospheric Sciences*, 3rd revised edition. Oxford, UK: Academic Press.
- Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y.O., Marsh, R., Yeager, S.G., Amrhein, D.E. and Little, C.M. (2019) A review of the role of the Atlantic meridional overturning circulation in Atlantic multidecadal variability and associated climate impacts. *Reviews of Geophysics*, 57, 316–375. <https://doi.org/10.1029/2019RG000644>.
- Zhang, X., Alexander, L., Hegerl, G.C., Jones, P., Klein Tank, A., Peterson, T.C., Trewin, B. and Zwiers, F.W. (2011) Indices for monitoring changes in extremes based on daily temperature

and precipitation data. *WIREs Climate Change*, 2, 851–870.
<https://doi.org/10.1002/wcc.147>.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Moemken J, Feldmann H, Pinto JG, *et al.* The regional MiKlip decadal prediction system for Europe: Hindcast skill for extremes and user-oriented variables. *Int J Climatol*. 2021;41 (Suppl. 1):E1944–E1958.
<https://doi.org/10.1002/joc.6824>