https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Weather and Climate Dynamics Discussions
Open Access
EGU

An unsupervised learning approach to identifying blocking events: the case of European summer

Carl Thomas[1], Apostolos Voulgarakis[1,2], Gerald Lim[3], Joanna Haigh[1,4], and Peer Nowack[1,4,5,6]

[1]Department of Physics, Imperial College London, South Kensington Campus, London, SW7 2BW, UK
[2]School of Environmental Engineering, Technical University of Crete, Chania, Crete, 73100, Greece
[3]Centre for Climate Research Singapore, 36 Kim Chuan Road, 537054, Singapore
[4]Grantham Institute, Imperial College London, SW7 2AZ, UK
[5]Climatic Research Unit, School of Environmental Sciences, Norwich, NR4 7TJ, UK
[6]Data Science Institute, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK

Correspondence: Carl Thomas (c.thomas18@imperial.ac.uk)

Abstract. Atmospheric blocking events are mid-latitude weather patterns, which obstruct the usual path of the polar jet streams. They are often associated with heat waves in summer and cold snaps in winter. Despite being central features of mid-latitude synoptic-scale weather, there is no well-defined historical dataset of blocking events. Various blocking indices (BIs) have thus been suggested for automatically identifying blocking events and are frequently used to study their occurrence historically as well as in climate model simulations. However, BIs can show significant regional and seasonal differences and therefore several indices are typically applied in parallel to test scientific robustness. Here, we introduce a new blocking index using self-organizing maps (SOMs), an unsupervised machine learning approach, and compare its detection skill to some of the most widely applied BIs. To enable this intercomparison, we first create a new ground truth time series classification of European blocking based on expert judgement. We then demonstrate that our method (SOM-BI) has several key advantages over previous BIs because it exploits all of the spatial information provided in the input data and avoids the need for arbitrary thresholds. Using ERA5 reanalysis data (1979-2019), we find that the SOM-BI identifies blocking events with a higher precision and recall than other BIs. We present case studies of the 2003 and 2019 European heat waves and highlight that well-defined groups of SOM nodes can be an effective tool to reliably and accurately diagnose such weather events. We further compare the skill at detecting historic blocking events by applying our new SOM-BI to several meteorological variables that are associated with the study of blocking, including geopotential height, sea level pressure and four variables related to potential vorticity. The 500 hPa geopotential height anomaly field is the variable that most effectively supports the identification of blocking events with our new approach. Finally, we evaluate the SOM-BI performance on around 100 years of climate model data from a pre-industrial simulation with the new UK Earth System Model (UK-ESM1). For the model data, all blocking detection methods have lower skill than for the ERA5 reanalysis, but SOM-BI performs noticeably better than the conventional indices. SOM-BI performs well using at least 20 years of training data, which suggests that observational records are sufficiently long to train our new method effectively. Overall, our results demonstrate the significant potential for unsupervised learning to complement the study of blocking events in both reanalysis and climate modelling contexts.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Weather and Climate Dynamics Discussions
Open Access
EGU

1 Introduction

Atmospheric blocking events are large-scale mid-latitude anticyclones that can persist for several days, which obstruct the typical westerly flow pattern (Rex, 1950). Blocking systems are often associated with regional extreme weather events, particularly heat waves in summer and cold snaps in winter. For example, the 2003 summer heat wave and 2009/10 winter cold events in Europe were both associated with atmospheric blocking (Black et al., 2004; Cattiaux et al., 2010). The evolution of atmospheric blocking itself is nonlinear (Palmer, 1999) and the underlying complex physical mechanisms are not yet fully understood (Nakamura and Huang, 2018; Woollings et al., 2018). There is a large seasonal, inter-annual and decadal variability in the occurrence of blocking (Kennedy et al., 2016; Brunner et al., 2017), which compounds the problem of separating externally forced changes from internal variability (Barnes et al., 2014; Shepherd, 2014). As a result, the influence of climate change on blocking remains an open question (Francis and Vavrus, 2012; Barnes, 2013; Hassanzadeh et al., 2014; Barnes and Polvani, 2015; Barnes and Screen, 2015; Francis and Vavrus, 2015; Coumou et al., 2018; Mann et al., 2018).

In order to better understand blocking and to investigate the influence of climate change, there have been significant efforts to develop methods that can automatically detect blocking in long meteorological records. Since "any attempt to identify blocked cases with certainty from an inspection of the longer available record of surface analyses would require a prohibitive expenditure of time" (Rex, 1950), blocking indices (BIs) have been developed to objectively identify blocked events (Lejenäs and Økland, 1983; Dole and Gordon, 1983; Tibaldi and Molteni, 1990; Pelly and Hoskins, 2003). However, the multiplicity of these BIs, with a variety of thresholds for defining the area, persistence and magnitude of blocked features on different atmospheric dynamical variables, means that these methods necessarily carry the burden of somewhat subjective definitions. Notably, while previous intercomparisons of BIs show similar global climatologies, and while all indices capture many of the basic features of atmospheric blocking within their definitions, there are known regional and seasonal differences (Croci-Maspoli et al., 2007; Barriopedro et al., 2010; Pinheiro et al., 2019). In addition, whilst spatial climatologies obtained from these BIs have been compared extensively, to the best of our knowledge there has been no direct time series comparison of the BIs beyond case study analyses such as those in Scherrer et al. (2006) and Pinheiro et al. (2019).

Other frequently used methods to study the climatology of, and changes in, blocking include a dynamical metric referred to as "sinuosity" (Cattiaux et al., 2016; Blackport and Screen, 2020), as well as an unsupervised machine learning approach called self-organizing maps (SOMs), a form of cluster analysis (Skific and Francis, 2012; Horton et al., 2015; Mioduszewski et al., 2016; Gibson et al., 2017a). Owing to the incomplete understanding of blocking paired with large internal dynamical variability, it has been highlighted that consistency across various methods in detecting long-term changes is a fundamental requirement to identify trends with certainty (Woollings et al., 2018; Barnes et al., 2014). To the best of our knowledge, there has been no previous study that directly compared an SOM approach to the BIs or other dynamical indices.

With the advent of modern computational methods and better instrumentation and observations, extensive study of the available record of surface analyses to identify blocking events no longer requires a prohibitive expenditure of time. Here, we

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Weather and Climate Dynamics Discussions
Open Access
EGU

therefore define a new binary ground truth dataset (GTD) of European blocking events across June–July–August (JJA) 1979-2019, based on a five day threshold, historical information and expert judgement. Our understanding of blocking events has been informed by the BIs and the various objective definitions that have been proposed, but we do not rely on any BIs for our study. This provides an objective ground for a time series comparison with the BIs. We also study the sinuosity metric as an alternative means to describe the mid-latitude atmosphere, and present case studies of the prominent 2003 and 2019 European heat waves, where we show how well sinuosity, the BIs and SOMs describe the blocking events and how they relate to surface temperature extremes.

We then use SOMs to develop a new blocking index (SOM-BI). This SOM-BI method has advantages over previous BIs because it exploits all of the spatial information provided in the input data and avoids the need for arbitrary thresholds. It also provides a new way of studying blocking events that can more intuitively distinguish between different regimes and locations of blocking events, which the other indices are lacking. We find that through comparison with three BIs used in a recent inter-comparison study (Pinheiro et al., 2019), the SOM-BI method has an improved skill at detecting regional blocking events. Since the SOM-BI method is not bound to a specific meteorological variable, we also quantify how its detection skill varies with the variable used, from geopotential height anomaly fields to potential vorticity maps. While there have been theoretical discussions on the importance of the meteorological variable used to define and identify blocking (Pelly and Hoskins, 2003; Chen et al., 2015), the dependence of skill of blocking detection methods has not been quantified before. Finally, we evaluate the performance of the SOM-BI on 101 years of a pre-industrial control run carried out with the UK Earth System model (UKESM1-0-LL), and identify a significant improvement in skill over the BIs as for the historical blocking events. A key advantage is that the longer climate model simulation allow us to test the robustness of our method compared to other BIs over longer timescales, and to study the dependence of the SOM-BI detection skill on the number of years included in the algorithm's training dataset.

Our paper is structured as follows. In section 2 and its subsections, we introduce the meteorological reanalysis and climate model data, the new GTD, the BIs, sinuosity, SOMs, and our new SOM-BI. In section 3, we present the main results of our analysis. We first compare the various blocking identification methods by means of the 2003 and 2019 European heat wave case studies (section 3.1), followed by an evaluation and intercomparison of the methods on ERA5 reanalysis and UK-ESM1-0-LL climate model data (sections 3.2 and 3.3). In section 3.4, we discuss how the performance of our new SOM-BI depends on the length of the data record used to train the algorithm. In section 3.5, we test the feasibility to train SOM-BI on ERA5 data to then reliably identify blocking in climate model data, and vice versa. In section 3.6 we briefly discuss the effect of other hyperparameters on the SOM-BI skill. In section 4, we summarise and discuss our key results, and propose avenues for future work, especially concerning the detection of blocking in climate change simulations.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Weather and Climate Dynamics Discussions
Open Access
EGU

2 Methods

2.1 Meteorological data

As a proxy for observed dynamical states over Europe, we used ERA5 reanalysis data from the European Centre for Medium Range Weather Forecasts (ECMWF, Hersbach et al., 2020). The pre-industrial climate model data was obtained from simulations carried out with the UK Earth System Model (UKESM1-0-LL) as part of Coupled Model Intercomparison Project Phase 6 (CMIP6, Sellar et al., 2019; Eyring et al., 2016). For ERA5, we used gridded data at a spatial resolution of 1° x 1° across 1979-2019, and created daily averages derived from 3-hourly intervals. In UKESM1-0-LL, we used 101 years of daily data from the pre-industrial run of the r1i1p1f2 ensemble member, across the arbitrarily defined 1960-2060 period. We used the UKESM1-0-LL (hereafter UKESM) data at the native resolution of 1.25° x 1.875° to develop the GTD plots and regridded to a 2° x 2° grid for the SOM analysis. When training and testing between the ERA5 and UKESM data (section 3.5), we also regridded the ERA5 data to a 2° x 2° grid.

For both types of datasets, we used the following common meteorological variables to characterize the dynamical state of the atmosphere at any given time: geopotential height anomalies at 500 hPa (Z_{500}), mean sea level pressure (MSLP), vorticity anomalies at 500 hPa (ζ_{500}). For ERA5, we also used vertically integrated potential vorticity across 150-500 hPa (VPV), isentropic potential vorticity on 350 K and 330 K (IPV_{350} and IPV_{330}) and potential temperature on the PV= 2 PVU surface (θ-PV). These PV-based variables have all been used in the context of understanding atmospheric blocking (Hoskins et al., 1985; Crum and Stevens, 1988; Pelly and Hoskins, 2003) but are not available from the CMIP6 archive. The 350 K and 330 K isentropes were chosen because these intersect with the tropopause in the mid-latitude summer, as shown in Fig. 1 of Liniger and Davies (2004), and therefore represent upper-level dynamics. For the case study analyses we have also used the surface horizontal wind fields and surface temperature (T_{surf}).

Following Pinheiro et al. (2019) and Grotjahn and Zhang (2017), we define the anomaly fields that we study by subtracting a long-term daily mean (LTDM) from the data instead of subtracting the daily average. This is a smoothed function of the 365-day seasonal cycle across Z_{500}, VPV and T_{surf} using the first six harmonics of their Fourier series. The purpose of this is to smooth out the daily mean fields, which can otherwise show excessive variation between neighbouring days across the seasonal cycle.

Where we have studied the T_{surf} and Z_{500} anomaly in ERA5, we have subtracted thermodynamic warming by detrending across time. Following Jézéquel et al. (2017) we subtract a spatially uniform trend over the domain being considered, so that the horizontal gradients of the field are not altered. We depart from the Jézéquel et al. (2017) method by subtracting a linear Z_{500} anomaly trend instead of a cubic spline interpolation, since we assume that in the 1979-2019 time period the thermodynamic dilation of the troposphere can be approximated as linear, so removing nonlinear trends could risk removing the dynamical changes in the atmosphere that we are interested in. We also apply the same detrending approach to the pre-industrial UKESM data, to remove any minor remaining trends in the data, e.g. due to the finite spin-up time of the control simulations (Gregory et al., 2004; Nowack et al., 2017; Mansfield et al., 2020).

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

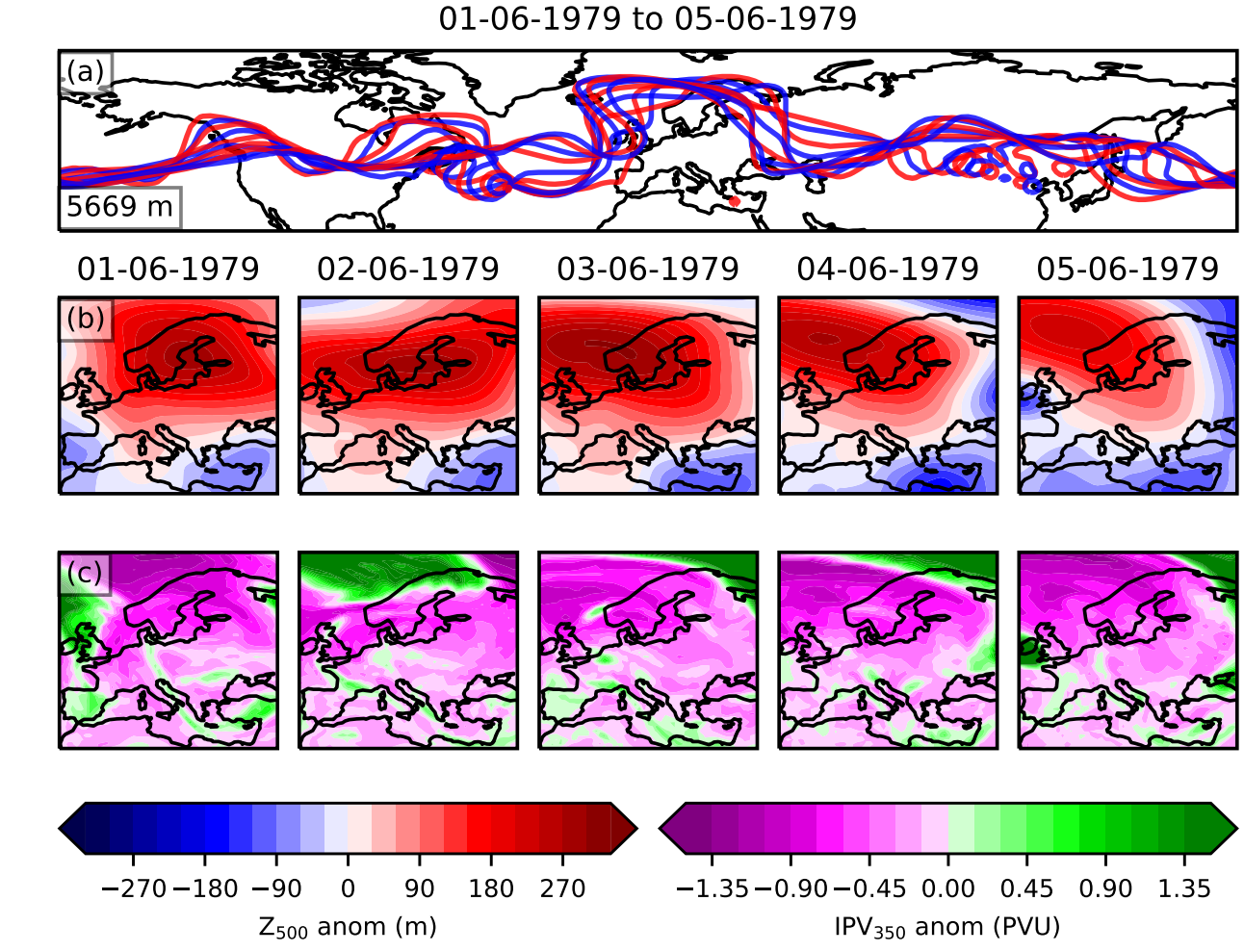

Figure 1. The information used to classify blocks in the ERA5 ground truth dataset (GTD). (a) shows the Z_{500} contour for the averaged value across 30-70 °N, indicated in the bottom left of the panel. The red and blue colours highlight the contours at midnight and midday respectively. (b) and (c) show the Z_{500} time detrended anomaly and IPV_{350} anomaly for each day.

2.2 Creating the ground truth dataset (GTD)

In order to objectively compare the blocking indices, we develop a ground truth dataset (GTD) of blocking events in JJA Europe, here defined as 30–75 °N, 10W–40 °E, following IPCC AR5 definitions (Stocker et al., 2013). The northern latitude is extended to 76 °N when using data on a 2° x 2° grid. The GTD has been derived by studying every successive five day period from 28 May to 4 September 1979-2019, and identifying whether or not a blocking high persisted across that period. By including the last four days at the end of May and the first four days of September, we ensure that we capture all blocking

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

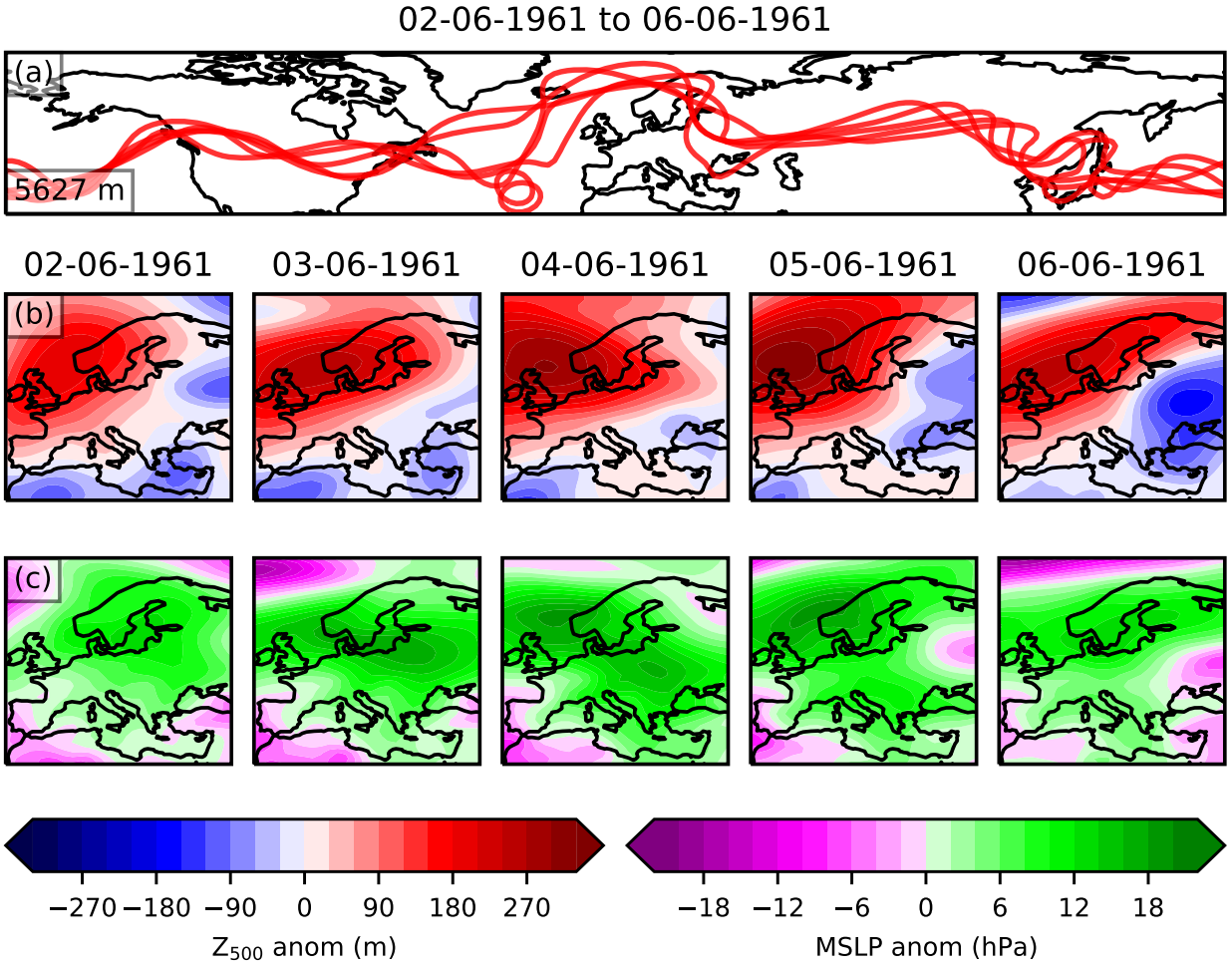

Figure 2. The information used to classify blocks in the UKESM GTD. (a) shows the daily Z_{500} contour for the averaged value across 30-70 °N, indicated in the bottom left of the panel. (b) and (c) show the Z_{500} time detrended anomaly and MSLP anomaly for each day.

events within the JJA period. Five days was chosen since this a typical persistence threshold for blocking indices (Verdecchia et al., 1996; Schwierz et al., 2004; Scherrer et al., 2006; Pinheiro et al., 2019), although a persistence of 7-10 days with weaker BI thresholds for amplitude and area has also been used (Rex, 1950; Lejenäs and Økland, 1983). We highlight that the SOM-BI defined below does not require definitions of such thresholds.

A diagram of the type of information analysed to label each individual day is shown in Fig. 1, for the example period 1-5 June 1979. This period was labelled as blocked, since Fig. 1a clearly shows a continuous North shift in the Z_{500} contours over Europe and Fig. 1b shows a substantial positive Z_{500} anomaly which persists across Northern Europe. The IPV_{350} maps in Fig. 1c highlight filaments and regions where there is fast moving air. Once the total set of all 4001 5-day periods starting on

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

consecutive days within JJA 1979-2019 has been classified (labelled 0 for non-blocked and 1 for blocked periods), blocking events are reconstructed to form a time series where each day is labelled as blocked or not.

A similar approach was adopted to classify 9494 five-day periods from 101 years of JJA data in the UKESM pre-industrial control run, with an example blocked period shown in Fig. 2. As in Fig. 1, there is a clear quasi-stationary high centered on a region slightly north of the UK. This is indicated by the Z_{500} contours which show a significant northward protrusion over this region, and by the substantial Z_{500} anomaly across all panels in Fig. 1b. Since PV is not available in CMIP6 data and the physical variables used to derive PV are not available at sufficiently high vertical resolution, we instead show the MSLP anomaly field in Fig. 2c, which also indicates a high pressure region consistent with Figs. 2a and 2b.

2.3 Blocking Indices (BIs)

One way of describing atmospheric flow and investigating trends in atmospheric dynamics is by using proxy indices such as those used to classify if a blocking event is occurring. There are many blocking indices (BIs) that have been used to create a blocking climatology, and these have been rigorously compared (Barriopedro et al., 2010; Pinheiro et al., 2019). Some BIs are based on measuring persistent anomalies of a relevant pressure field in a particular location. This builds on the pioneering work of Elliot and Smith (1949), who identified events of persistent sea level pressure (SLP) anomalies above a particular threshold. This approach was extended by Dole and Gordon (1983) who investigated persistent anomalies in the Z_{500} field. A similar approach was taken by Schwierz et al. (2004) who identified anomalies in the vertically averaged potential vorticity field (VPV), averaged over 150-500 hPa. This approach was inspired by the work of Pelly and Hoskins (2003), who defined blocking as the negative latitudinal potential temperature gradient on the dynamical tropopause. By taking a vertical average of the potential vorticity field from the mid-troposphere to the lower stratosphere, Schwierz et al. (2004) formulate a 3-D dynamically based index.

Another common approach to studying blocking trends is to use the absolute gradient of Z_{500} across fixed latitudes. This was first developed in Lejenäs and Økland (1983) and refined in a commonly applied form by Tibaldi and Molteni (1990). This definition focuses on blocking events as persistent anticyclones that reverse the Z_{500} gradient. The method has been adopted widely, refined (Diao et al., 2006; Barriopedro et al., 2010), and extended to a range of latitudes (Scherrer et al., 2006).

All of these methods have been further developed by Pinheiro et al. (2019) who applied four thresholds for each blocking index: the magnitude of the anomaly, the persistence of the blocking event, a minimum area over which the anomaly takes place and an overlap criterion which measures if there is continuity across the blocked region between different days. We adopt their thresholds and as such study the three indices compared in Pinheiro et al. (2019) including their modifications:

- **AGP** - the geopotential height gradient method, which is the Tibaldi and Molteni (1990) index as adapted by Scherrer et al. (2006) to construct a two-dimensional field of geopotential height gradients
- **DG83** - the Dole and Gordon (1983) method of investigating positive geopotential height anomalies
- **S04** - the Schwierz et al. (2004) method of identifying persistent anomalies in the potential vorticity field (VPV) averaged over 150-500 hPa (VPV).

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Weather and Climate Dynamics Discussions
Open Access
EGU

However, our analysis differs from the methodology outlined by Pinheiro et al. (2019) in two ways, reflecting the fact that our study is regional instead of global. Firstly, we apply all thresholds defined by Pinheiro et al. (2019) only to those grid cells within the region of study so that we exclude events that are on the edges of the domain. Such events would be considered blocking events if the domain studied was extended. Additionally, Pinheiro et al. (2019) applied a spatial smoothing to their global threshold field, which defines the minimum threshold for each grid cell to be blocked. Although we have applied the LTDM smoothing of the seasonal cycle (which we subtract from variables to calculate field anomalies, section 2.1) and we also use a spatially varying threshold field, we have not applied this spatial smoothing to our threshold field. We found that the resulting blocking climatologies shown in Fig. A3 are broadly consistent with those presented in Fig. 6 of Pinheiro et al. (2019), underlining that this regional use of the BIs is still valid.

We note that more indices have been proposed, including hybrid approaches combining the AGP and DG83 indices (Barriopedro et al., 2010; Dunn-Sigouin et al., 2013; Woollings et al., 2018), the PV-θ approach developed by Pelly and Hoskins (2003) and the finite amplitude wave activity (FAWA) method (Huang and Nakamura, 2015). However, with the three methods included here, as well as the SOM method and sinuosity metric described below, we expect to see results that are sufficiently representative of the range of blocking detection methods available, and to be able to highlight their most important similarities and discrepancies.

2.4 Sinuosity

Another way of investigating changes in mid-latitude atmospheric dynamics is by developing a metric that can describe the variability of the zonal flow in form of a continuous function of time. Cattiaux et al. (2016) characterised the trajectory of the mid-latitude flow by measuring the length of the 500 hPa isohypse (line of constant geopotential height) at a value averaged across 30-70 °N, such that the atmospheric flow at approximately 50 °N is described for each day. Such isohypses are shown in Fig. 1a. The length of each isohypse across is then normalized by the length of the 50 °N Earth circle to obtain a measure of the sinuosity of the atmospheric flow.

2.5 Self-organizing map (SOM)

The fifth method we used to investigate trends in atmospheric circulation regimes in European summer is self-organizing map cluster analysis (SOM; Kohonen, 1982). This is an increasingly popular unsupervised machine learning technique in synoptic meteorology to learn representative patterns of weather regimes and to investigate their trends (Hewitson and Crane, 2002; Liu and Weisberg, 2005; Huth et al., 2008; Sheridan and Lee, 2011; Johnson, 2013; Horton et al., 2015; Xu et al., 2016; Singh et al., 2016; Diffenbaugh et al., 2017; Sánchez-Benítez et al., 2019). In our context here, the SOM algorithm is trained with daily spatial maps of dynamical states of the atmosphere above Europe, as for example characterized by maps of geopotential height anomalies (Fig. 1b). By iteratively cycling through all samples of such meteorological maps, the algorithm learns representative patterns of atmospheric dynamical states, which are referred to as “SOM nodes”.

First, the number of nodes is specified and the SOM is initialised either with random values or with principal component analysis patterns. Then for each day from the input field, the Euclidean distance between that meteorological pattern and each

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

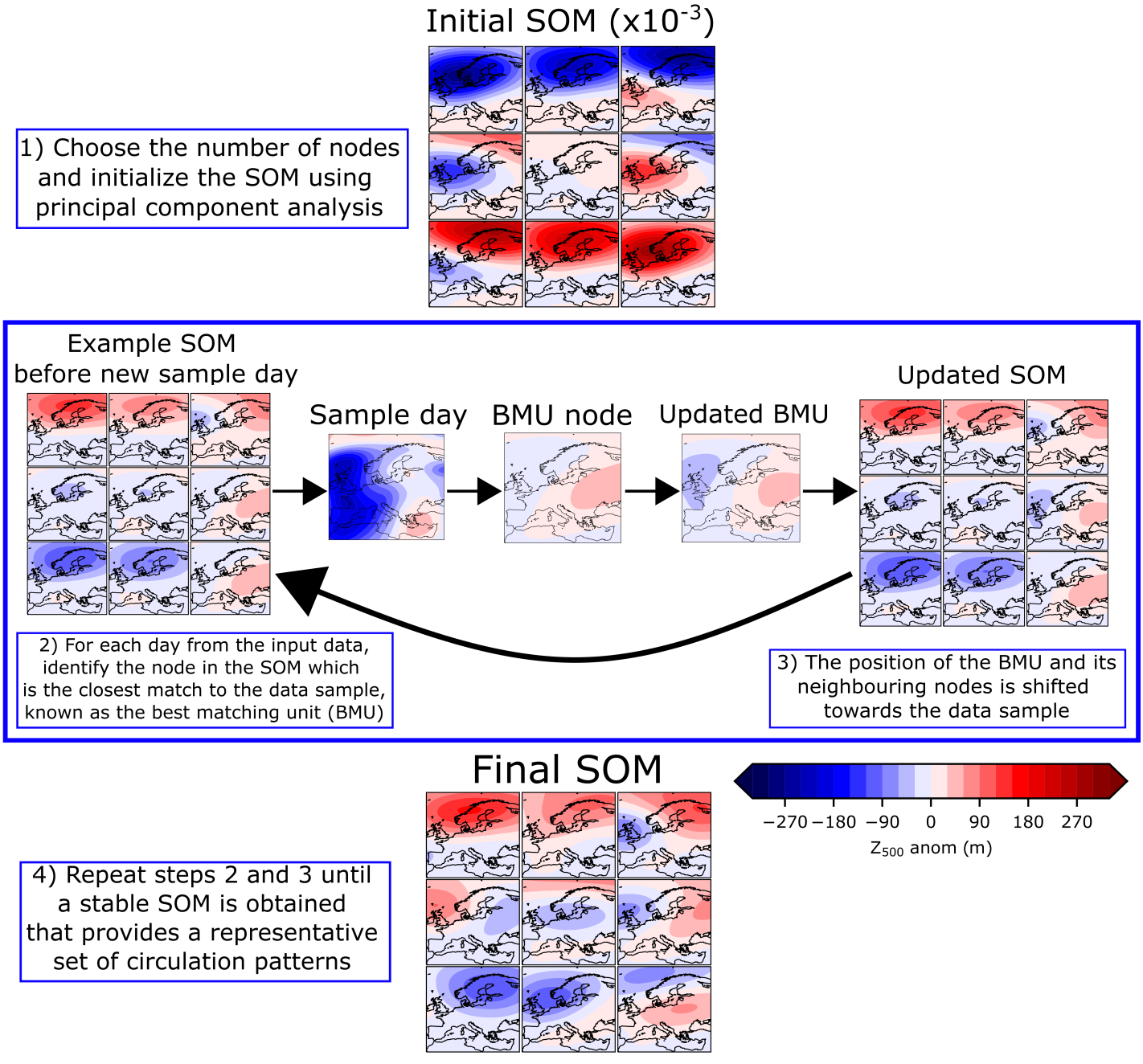

Figure 3. The self-organizing map algorithm. Shown using a 3x3 node SOM with ERA5 Z_{500} JJA 1979-2019. The example SOM is identical to the initial SOM in the first iteration.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

node pattern is calculated. The node with the smallest Euclidean distance to the sample day is known as the best matching unit (BMU) for that day. Then the BMU pattern is updated to shift towards the pattern of the sample day. The neighbouring SOM nodes are also updated to shift towards the sample day according to a Gaussian neighbourhood function. For each iteration the updates tend to become smaller as the SOM nodes converge towards a representative pattern of atmospheric dynamical states. A decay function on the updates is applied, which ensures convergence. Finally, a stable SOM is obtained with a set of nodes that each provide a representative composite of circulation patterns, arranged according to their similarity on a row-column grid (i.e. the map). A diagram of the training procedure is shown in Fig. 3. The number of nodes to be learned by the algorithm, or in other words the number of representative weather patterns one aims to learn for a particular meteorological problem, chosen by the user. In section 3.3, we show how the SOM-BI performance depends on the number of nodes and how this provides an objective criterion to select this number.

SOMs are of particular relevance in atmospheric science because they maintain the topological properties of the input space. Once optimised, each node pattern represents a possible state of the atmosphere, and the nodes are arranged in order of similarity, thus representing a continuum of atmospheric states. This contrasts with other methods of dimension reduction such as principal component analysis, where the identified patterns are orthogonal. Such purely mathematical representations are typically less meaningful from a physical point of view, whilst each SOM node maintains physical significance as it can closely resemble actual atmospheric states found in meteorological data, with the nodes on the row-column grid representing smooth transitions across those possible atmospheric states (see the similarity of neighbouring nodes in the final SOM grid in Fig. 3). We have implemented the SOM algorithm using the somoclu package (Wittek et al., 2017).

2.6 The self-organizing map blocking index (SOM-BI)

Once we have created the GTD, this can be used as a means of developing a blocking index using a SOM approach. For a given variable from the ERA5 dataset, we can specify a node number and arrangement of nodes (number of rows and columns, Fig. 3) and then learn the corresponding SOM nodes from that data. Figure 4a shows the trained pattern for Z_{500} anomalies in ERA5 28 May-4 Sep 1979-2019 for 9 nodes arranged in a 3x3 grid. Since each day in the dataset has been matched to a BMU, we can identify which nodes are associated with blocked days according to our GTD. Figure 4b compares the histograms of those nodes which are and are not associated with the GTD blocking events. As expected, the three nodes with large positive Z_{500} anomalies (nodes 1, 2 and 3) are most closely associated with blocking events, and the nodes with large negative Z_{500} anomalies (nodes 7 and 8) are rarely associated with blocking events. However, nodes 1, 2 and 3 still occur on 15% of non-blocked days, and 28% of the blocked days are also matched with one of the other six nodes, including 3% of blocked days matched with nodes 7 and 8. This tells us that while the SOM nodes can indicate the occurrence of blocked events, there is no node or single combination of nodes that can be consistently identified with blocking events with high skill.

However, from every five day period within the GTD, we can identify an associated "group" of nodes. For example, a five day period can be associated with nodes 1 and 4 (any arrangement of nodes 1 and 4 across five days), and this would mean that [1,4] is the associated group of nodes for that five day period. Since each five day period has been classified either as blocked or not blocked, it raises the possibility that a set of such groups can be more specifically associated with blocking. We identify

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

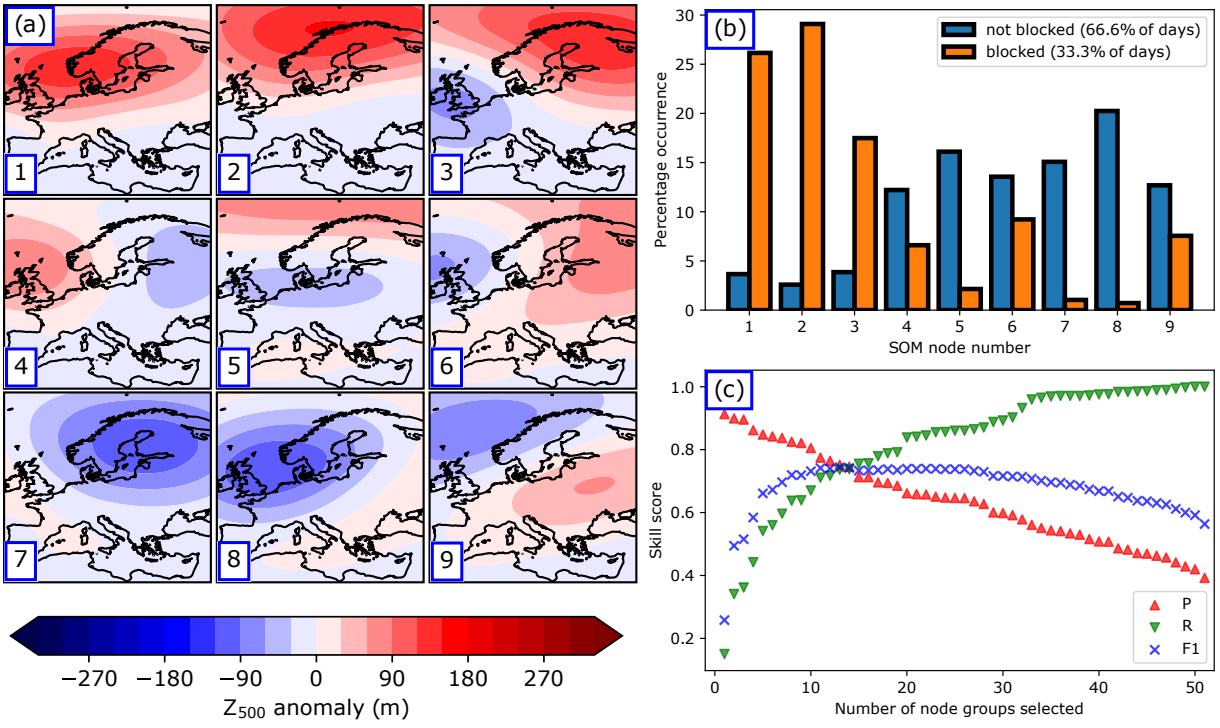

Figure 4. The SOM blocking index (SOM-BI). (a) The trained 3x3 SOM for Z_{500} time detrended anomaly. (b) Normalised histograms showing the distributions of occurrence of BMUs for the days identified as blocked or non-blocked within the GTD. (c) The SOM-BI optimisation of the set of node groups against three different skill scores (precision (P), recall (R) and F1 score) that are associated with the GTD blocking events.

the optimal set of node groups associated with blocking by ordering the list of all possible node groups (e.g. [1,2,3], [1,4], [1], [1,2,3,4,6] etc) from the node groups that have the highest to lowest precision (P) at identifying blocking events.

2.7 Classification skill measures

Fig. 4 (c) shows the binary classification skill according to the measures of precision, recall and F1 score when applying the 9-node SOM-BI to ERA5 data. The three skill measures are shown for consecutive cases where we successively add node groups as described above in order from highest to lower precision to the set of groups that we associate with blocking. In other words, once a new group has been added to the set of groups, this new group will define a series of blocked periods within our SOM-BI approach. Precision (P) is defined as the ratio of true positives to total detected positives. For example, a precision of 0.8 indicates that 80% of the events identified by a method are true positives and the remaining 20% are false positives. Recall (R) is the number of true positives divided by the total number of actual events. A recall of 0.8 indicates that 80% of all total blocking events are captured by the classification method, but 20% are false negatives. A higher recall

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

is typically associated with a loss in precision, as identifying more events also means that one typically identifies more false positive events. Therefore, a careful balance between precision and recall is usually sought after. One widely used skill metric to achieve this balance is the F1 score, which is the harmonic mean of precision and recall:

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (1)$$

which can vary between 0 (worst case, low detection skill) and 1 (best score). If either P or R are low, the F1 score tends towards 0, thus indicating low detection skill in at least one of the two measures. For example, if there is a small number of node groups selected in the SOM-BI, the precision is very high but the recall is very low - a small number of blocking events is well described but many blocking events are missed by the classification. When a larger number of node groups with a decreasing precision is included, then precision decreases and recall increases; more events are described but there is also a higher proportion of false positives. For the 3x3 SOM learned from ERA5 data, the node group with the highest precision is [1], with P = 0.91 and R = 0.15, followed by [2] with P = 0.89 and R = 0.19 and [1, 2, 6] with P = 0.87 and R = 0.03. Initially with only one node group in the set there is a high P and low R, but as more node groups are added to the set of node groups, P decreases but R increases. We identify the optimal set of node groups by the value which maximises the F1 score (Fig. 4c). We perform this classification for a range of node numbers and meteorological variables to identify an optimal performance in section 3.3.

3 Results

3.1 Case study analyses

We compare the blocking identification methods (i.e. SOMs/SOM-BI, the three conventional BIs, and sinuosity) for two examples of well-known 2003 and 2019 European heat waves that were linked to blocking states of the atmosphere (Figs. 5 and 6). In addition, we study two blocking events from UKESM, to investigate how blocking events are be described in the climate model. From the 101 years investigated in the pre-industrial control run we have found the largest extent of heat extremes to occur in an extended heat wave shown in Appendix Fig. A1. This is contrasted with Fig. A2, which shows the end of a blocking event and a weaker transitory anticyclone. Both UKESM events are discussed further in Appendix A.

The 2003 European heat wave was a record-breaking heat wave that had significant societal impacts (Robine et al., 2008) and was shown to have been made at least twice as likely due to anthropogenic climate change (Stott et al., 2004). According to climate change projections, such heat waves will become commonplace by the 2040s irrespective of future emissions scenarios (Christidis et al., 2014). The most extreme temperatures during this heat wave were recorded from the 6-12 August, where the peak temperature recorded was in Southern France at 41 °C. Black et al. (2004) reports that atmospheric flow anomalies were recorded in early August, although there was a relatively weak signature of blocking. The 2003 heat wave remained the European temperature record until 2019, when surface temperatures of 46 °C were observed in central France. The 2019 heat wave was concurrent with persistent hot air that originated in North Africa (the so-called “Saharan heat bubble”), which was sustained by an omega block centered on Western Europe (Mitchell et al., 2019).

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

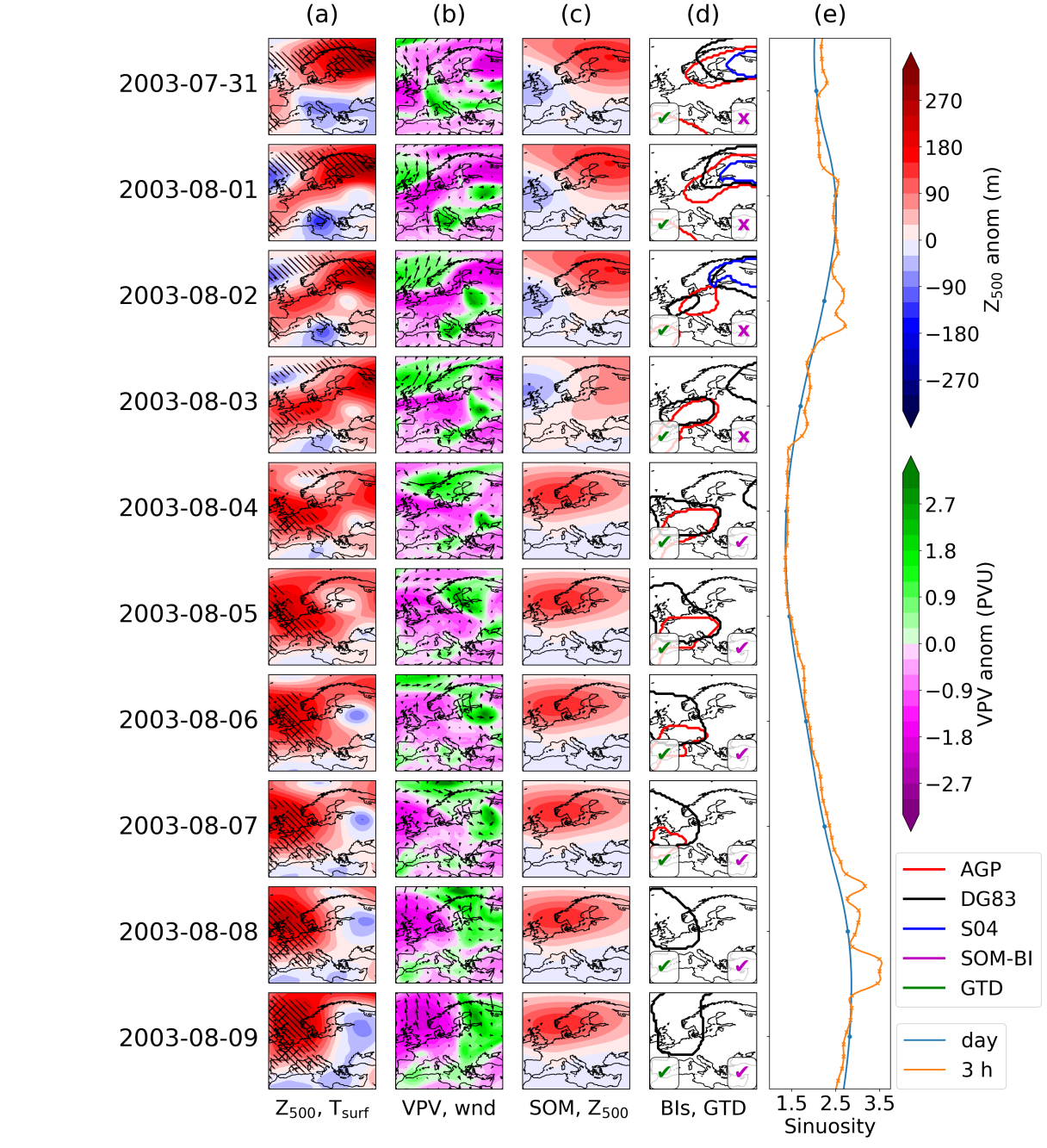

Figure 5. The 2003 European heat wave. (a) shows the detrended 500 hPa geopotential height anomaly for each day. Left (right) hatching indicates where the local surface temperature exceeds the 90th (99th) percentile for the detrended 2 m temperature. (b) shows the potential vorticity anomaly vertically averaged across 150-500 hPa, with arrows showing the 10-m wind field. (c) shows the corresponding SOM pattern for detrended 500 hPa geopotential height anomalies from 9 nodes. (d) shows the contours identified as blocked in this region in the AGP, DG83 and S04 indices. A green (magenta) tick or cross indicates if the GTD (SOM-BI) identifies the day as blocked or not. (e) shows the time series for sinuosity of the 500 hPa gph field across the same period. The blue line shows the daily values and orange showing the 3 h snapshots of sinuosity.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

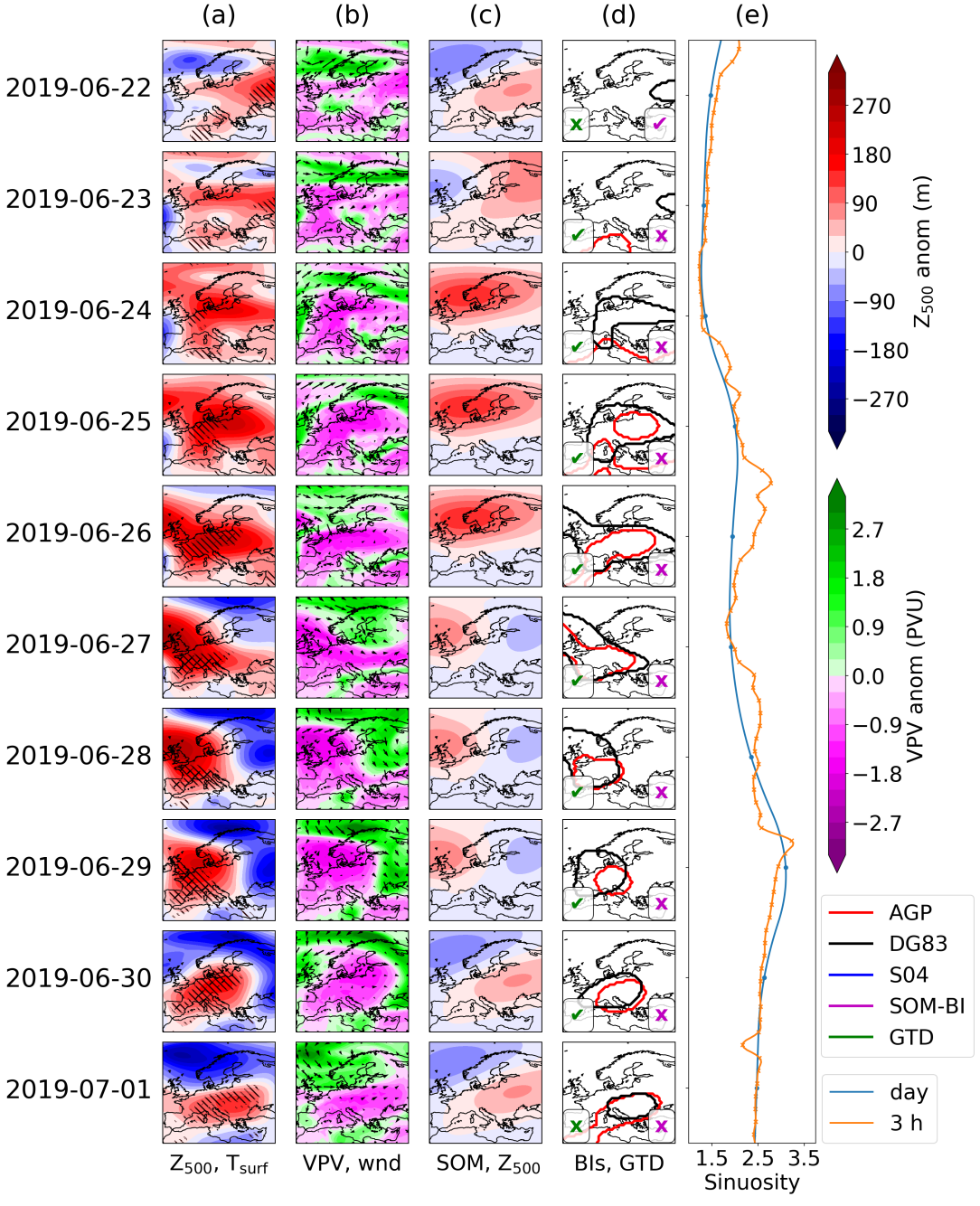

Figure 6. As in Fig. 5 but for the 2019 European heat wave.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Figs. 5a and 6a show daily maps of detrended Z_{500} anomalies for the two events, the field used by the DG83 index to identify blocking events. The hatching indicates detrended surface temperature extremes at the 90th and 99th percentile. It can be seen that across all cases there are significant positive Z_{500} anomalies which are associated with temperature extremes. Figures 5b and 6b show the vertically averaged potential vorticity (VPV) field, used by the S04 index to identify blocking, and also the 10 m winds. The VPV field is consistently anti-correlated with the Z_{500} field, and significant negative anomalies in the VPV field tend to be associated with stationary surface winds, particularly across 26-29 June 2019. Figures 5c and 6c show the BMU SOM pattern for the case of 9 nodes for detrended Z_{500} anomaly fields. Whilst the SOM nodes clearly track the features shown in the Z_{500} maps, a range of BMUs are identified in both case studies even though there is a consistent extreme weather event across these time periods. In the 2003 case study in ERA5, three SOM BMUs and four transitions between BMUs are shown in Fig. 5c. These all show positive Z_{500} anomalies in the Northern part of the domain, even though the meteorological situation varies meridionally more than zonally, particularly across 4-9 August 2003. An even greater variety of BMUs is observed in the 2019 case, where four nodes and four transitions between SOM nodes are shown in Fig. 6c.

This creates a difficulty of interpretation - whilst the SOM can identify the best matching spatial pattern of Z_{500} anomalies, these particular SOM patterns do not correspond to circulation regimes as conventionally understood, since even minor shifts in the domain (such as the change from the 2-3 August 2003) can cause the corresponding pattern to shift. The frequency of these shifts and sensitivity of the SOM is dependent both on the number of nodes chosen and the domain size. Smaller domains with fewer SOMs show more consistency in the synoptic weather patterns, but when these are sufficiently reduced (such as for four SOMs over the Mediterranean), the SOMs become less distinguishable and lose even more of their explanatory power to represent pattern variations across the domain (not shown). Overall, the fact that several SOM nodes occur during the case study blocking events shows that individual SOM patterns will not be able to consistently identify blocking events with high precision or recall, contrary to how SOMs are typically used in many applications in the literature. However, well-defined groups of nodes, as we will show below, can indeed achieve this task and can thus be used for the purpose of our new SOM-BI.

Figs. 5d and 6d show the contours demarking blocked regions as identified by the three different BIs. A tick or cross in the bottom left and right corners indicates whether the day was identified as blocked or not in the GTD or the SOM-BI. For the SOM-BI labelling, Z_{500} 20 SOM nodes are chosen on the basis of the optimisation of the SOM-BI in Fig. 7a. Across all case studies the AGP and DG83 indices clearly track regions of positive Z_{500} anomalies, and often show overlapping contours. However, the AGP index frequently identifies blocking in North Africa. This feature of the AGP index has been identified in climatologies that investigate Northern Hemisphere JJA (Pinheiro et al., 2019), and is associated with Z_{500} reversal due to the subtropical high (Davini et al., 2012). These features are physically very different to the mid-latitude blocking events that are associated with flow reversal, so are considered an anomalous feature of this index. The S04 index also tracks the same feature in the 2003 heat wave until 3 August 2003, but does not identify any blocking associated with the 2019 heat wave. The SOM-BI describes the 2003 heat wave well, but we note that it also does not capture the 2019 blocking pattern coincident with the 2019 heat wave. This is because the SOM nodes are too variable over the 2019 event such that the set of nodes which best match the Z_{500} anomaly fields are not more generally associated with blocking at high precision. For example, the SOM nodes across 27-30 June 2019 indicate mixed patterns which do not obviously correspond to blocking over a consistent area (the positive

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Weather and Climate Dynamics Discussions
Open Access
EGU

maxima shift from the British Isles to Eastern Europe within a day). This lack of pattern consistency is mostly the result of an unfortunate balance between the positive and negative Z_{500} anomalies on these days, where the latter play a major role in the allocation of the BMU during this period. We discuss the possibility of ignoring negative anomalies in our assignments of the BMUs in section 3.6, but found that this modification does not improve the SOM-BI performance overall. In summary, there are blocking events such as these that will also not be described well by the new SOM-BI, but as we show below the SOM-BI performs as good as or even better than many conventional BIs in most cases.

The daily sinuosity time series is shown in Figs. 5e and 6e. We also show 3 h snapshots for the ERA5 data here. The values of sinuosity are high where there are substantial perturbations in the flow across the domain, particularly across 8-9 August 2003 and 29-30 June 2003. The values are small where the Z_{500} are more consistent across the domain, such as across 4-6 August 2003, even though there are significant positive anomalies across this period. We conclude that, although high sinuosity is associated with blocking on a larger domain such as the Euro-Atlantic region (Cattiaux et al., 2016), high sinuosity does not necessarily correspond to blocking on a smaller domain. We note that similar features as those described above are shown in the UKESM case studies in Appendix Figs. A1 and A2.

3.2 Blocking index comparison in ERA5 and UKESM with GTD

A climatological comparison of the BIs over JJA Europe confirms what has been discussed in the case study analyses above, and is consistent with the results of Pinheiro et al. (2019), which are broadly consistent with other BI climatologies. We show the spatial distribution of blocking climatologies according to three conventional blocking indices in Fig. A3. Where our analysis substantially differs from the literature is in our regional approach and consideration of direct time series comparisons among the BIs as well as to our new SOM-BI. We do not explicitly consider the time-averaged climatological distributions of blocking events over Europe (as shown in Fig. A3). For this, we first apply all BIs to the historical ERA5 data, and each index will label each five day period either as blocked or not blocked. This binary dataset we then compare to our manually labelled GTD.

Table 1 compares the precision, recall and F1 scores of these BIs and our new SOM-BI against the GTD for this domain-based comparison in both ERA5 and UKESM. We further compare the time series of blocking classifications among the BIs themselves to quantify how consistent the BIs are with each other. The key results are underlined. In both ERA5 and UKESM, the best blocking index is the SOM-BI, with a F1 score of 0.74 in ERA5 and 0.71 in UKESM. All of the indices consistently perform worse in UKESM than in ERA5. This is because blocking is less frequent in the model and several of those blocking patterns identified in UKESM are less distinct (Fig. A2). This is probably associated with mean biases in the representation of Z_{500} that have been observed across several climate models (Scaife et al., 2010; Schaller et al., 2018). The DG83 index performs almost as well as the SOM-BI in ERA5 with an F1 score of 0.73, but there is a significant reduction in performance to 0.60 when applied to UKESM data. The AGP index in turn shows an even weaker skill than DG83 in both reanalysis and model, with a larger drop in skill to 0.60 and 0.28 in ERA5 and UKESM, respectively. The fact that SOM-BI still shows a relatively good score for UKESM suggests that the SOM-BI can be particularly useful in studying regional blocking in climate models. In particular, since a model ensemble may exhibit a variety of intensities of blocking, the SOM-BI would be able to

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Dataset	Method	Days blocked	Precision	Recall	F1	F1 wrt AGP	F1 wrt DG83	F1 wrt S04	F1 wrt SOM-BI
ERA5	GTD	33.4%	1	1	1	**0.53 (0.58)**	**0.73**	**0.19**	**0.74**
ERA5	AGP	91.7% (65.4%)	0.40 (0.56)	0.78 (0.61)	0.53 (0.58)	1 (0.72)	0.55 (0.61)	0.11 (0.15)	0.54 (0.53)
ERA5	DG83	34.3%	0.72	0.75	0.73	0.55	1	0.19	0.69
ERA5	S04	5.3%	0.69	0.11	0.19	0.11	0.19	1	0.15
ERA5	SOM-BI	35.0%	0.73	0.76	0.74	0.54	0.69	0.15	1
ERA5	BLO	100%	0.33	1	0.50	0.79	0.51	0.10	0.52
ERA5	RND	33.4%	0.33	0.33	0.33	0.44	0.34	0.09	0.34
UKESM	GTD	29.0%	1	1	1	**0.28 (0.31)**	**0.60**	-	**0.71**
UKESM	AGP	20.8% (17.0%)	0.27 (0.41)	0.29 (0.24)	0.28 (0.31)	1 (0.71)	0.23 (0.29)	-	0.26 (0.26)
UKESM	DG83	21.6%	0.90	0.45	0.60	0.23	1	-	0.55
UKESM	SOM-BI	29.6%	0.70	0.72	0.71	0.26	0.55	-	1
UKESM	BLO	100%	0.29	1	0.45	0.28	0.23	-	0.71
UKESM	RND	29.0%	0.29	0.29	0.29	0.30	0.19	-	0.29

Table 1. A comparison of skill scores of the original BIs and the new SOM-BI against the GTD for ERA5 1979-2018 and UKESM for JJA over the European domain. Where not indicated the skill scores are measured with respect to the relevant GTD. "BLO" indicates the skill score for the trivial case of every day labelled as blocked, and "RND" where a random allocation of blocked days has occurred with the same proportion of blocked days as the GTD. The values in brackets indicate the AGP index where latitudes 30-40 °N are excluded to remove the anomalous identification related to the subtropical high.

overcome the limitations of BIs, where (particularly in the case of AGP) thresholds are defined with respect to the observational record. Since the anomalous flow patterns associated with blocking will be more consistent across datasets, the SOM-BI can identify blocking events across a model ensemble with greater accuracy. The consistent skill of the SOM-BI across both ERA5 and UKESM has been further verified by swapping the training and test datasets between each dataset, as described in section 3.5.

Numbers in brackets next to the AGP score indicate the performance of the AGP index as a binary labelling of regional blocking where the grid cells beneath 40 °N have been excluded. This is to remove unintended features arising from subtropical highs, as found in Figs. A3a and A3d and discussed in section 3.1. Removing this feature increases the precision of the index, but decreases the recall, and in ERA5 leads to a significantly improved F1 score. In the UKESM case similar changes to precision and recall occur but the overall F1 score remains low.

A case where every day in Europe is labelled as blocked ("BLO") is also compared, which represents the case of perfect recall (=1) but a low precision. This case gives an F1 score of 0.53 for the GTD for ERA5 and 0.45 for the GTD of UKESM,

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

and provides a useful benchmark of basic performance. Surprisingly, the AGP index only performs marginally better than BLO for ERA5, and performs worse in the UKESM case. Whilst S04 has a higher precision than BLO, because the recall is so low the total F1 score is much lower (0.19). Finally, we compare a random labelling of blocked and non-blocked days, where the proportion of blocked days is equal to that of the GTD ("RND"). This gives an equal precision and recall because the number of true positives is equal to the number of false negatives. The F1 score of RND still exceeds that of S04, with 0.33 for ERA5 and 0.29 for UKESM, and is comparable to the F1 score of AGP in UKESM.

3.3 SOM-BI skill dependence on the choice of SOM node number and the meteorological variable

The key hyperparameter in the SOM-BI is the number of nodes (k), which here is similar to identifying the optimal number of circulation patterns required to effectively classify European summer weather regimes. In addition, there are a number of meteorological variables from which we could learn the SOM patterns, which in turn will also influence the skill of our SOM-BI method. The dependency of the skill of our BI on these two factors is quantified in the following. Figures 7 and 8 show how precision (P), recall (R) and F1 score depend on k and the meteorological variable in ERA5 and UKESM, respectively. Specifically, we compare the skill metrics for cases where we learn the SOM nodes from Z_{500}, MSLP and ζ_{500} anomalies. For ERA5, we additionally consider four PV-related variables (VPV, θ-PV, IPV_{350} and IPV_{330}) shown in Fig. 7d-g.

Another hyperparameter related to the number of nodes is the row x column (nxm) arrangement of nodes. For example, 16 nodes can be arranged as 16x1, 4x4, 8x2, 2x8 or 1x16. These different arrangements affect the topology of the SOM, the initialization of the nodes and which nodes are counted in the neighbourhood of other nodes during the update process of the SOM (Fig. 3). For each k in Figs. 7-9 we have used the arrangement of nodes that maximises the average number of nearest neighbours between each node (e.g. using 4x4 nodes for k = 16). This approach maximally exploits the SOM topology. We have also used $n \geqslant m$ (for example using a 9x2 arrangement instead of a 2x9 arrangement of nodes for k = 18) to preferentially arrange the SOM topology zonally across the domain rather than meridionally. This is done because there is greater variability in the occurrence of blocking patterns zonally than meridionally across Europe (Fig. A3).

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

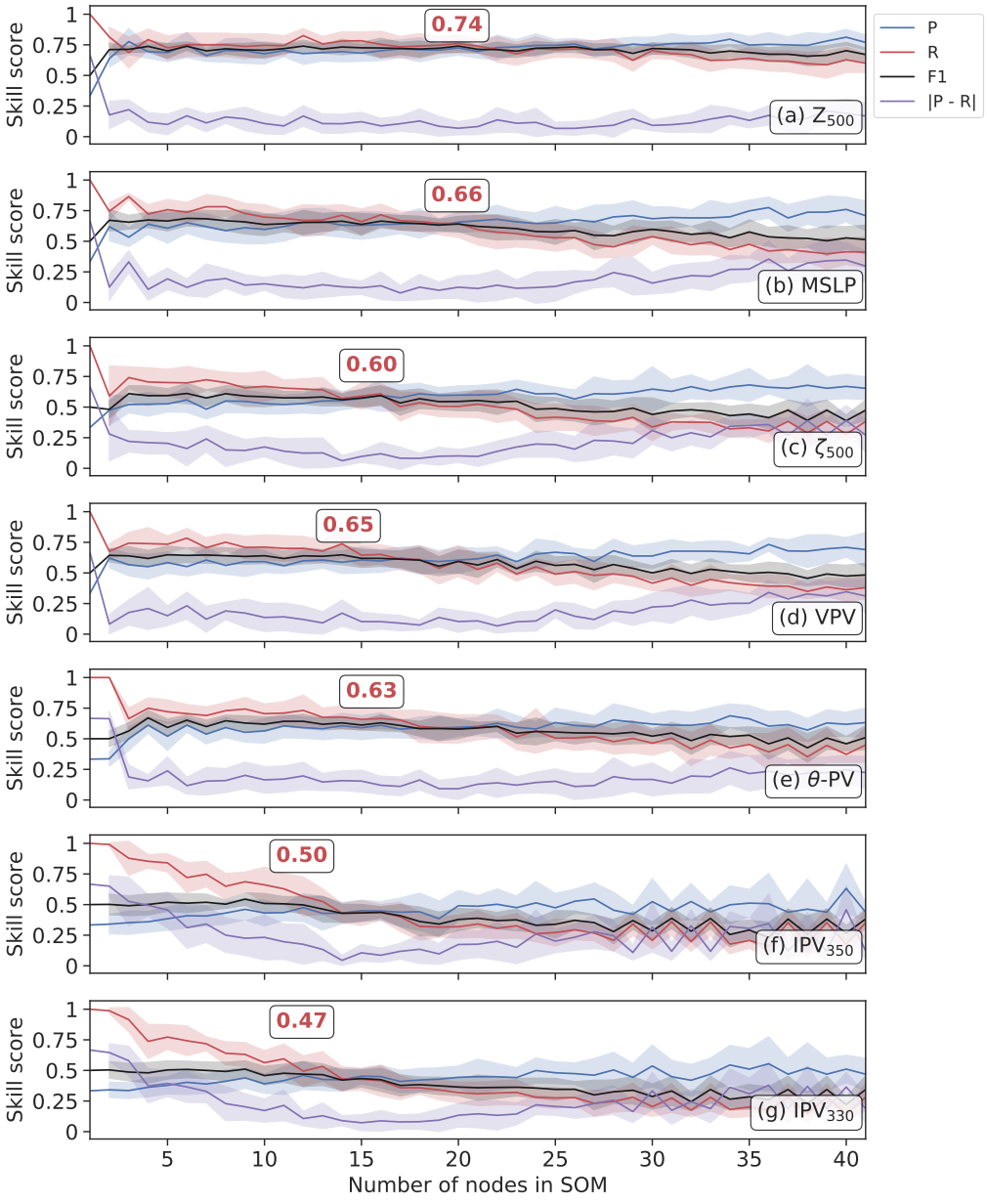

Figure 7. A comparison of the performance of the SOM blocking index for seven variables in the ERA5 1979-2019 historical period with a varying number of nodes in the SOM. Precision (P), recall (R) and F1 scores are calculated, and the absolute difference between precision and recall is also shown. Error bands show the standard deviation ($\pm 1\sigma$) for 10-fold cross-validation. The red number inset into each panel shows the optimal F1 score and the position of the box indicates the corresponding optimal node number. The optimal value is defined by the node number where the F1 score is close to its maximum value and the difference between precision and recall is close to the minimum value.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

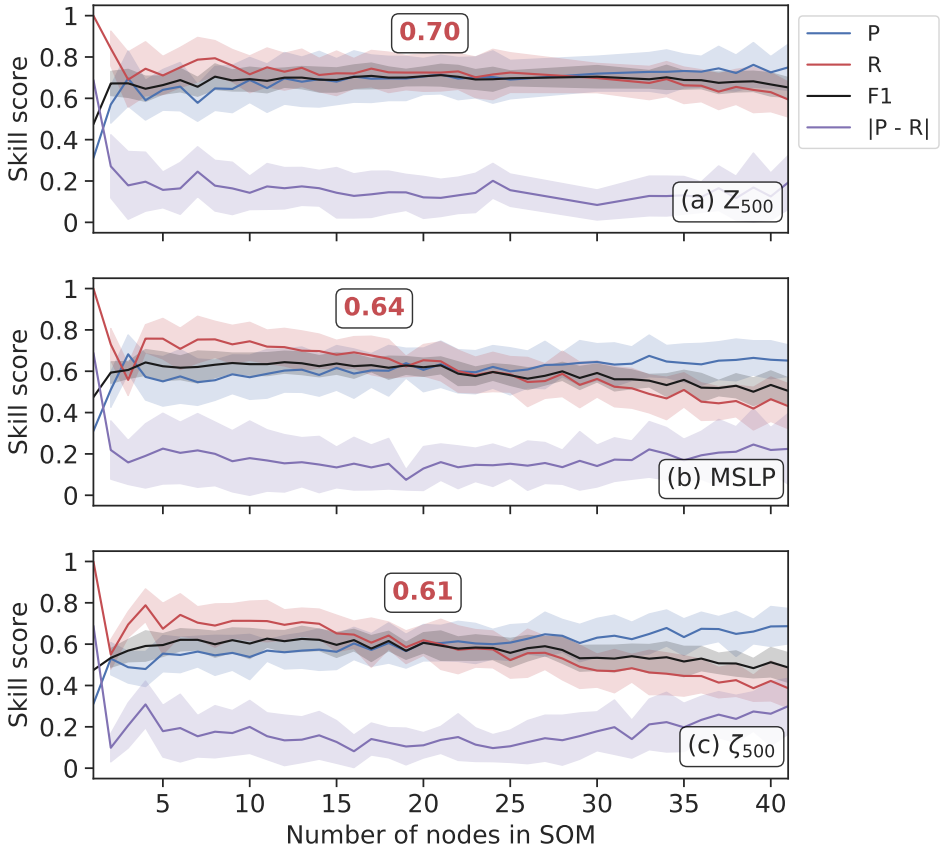

Figure 8. A comparison of the SOM blocking index performance for three variables in 101 years from the UKESM pre-industrial control period with a varying number of nodes in the SOM. Precision (P), recall (R) and F1 scores are calculated, and the absolute difference between precision and recall is also shown. Error bands show the standard deviation ($\pm 1\sigma$). The red number inset into each panel shows the optimal F1 score and the position of the box indicates the corresponding node number. As above, the largest F1 score is for Z_{500}, indicating that Z_{500} is the best variable tested for analysing blocking patterns using the SOM-BI in UKESM.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

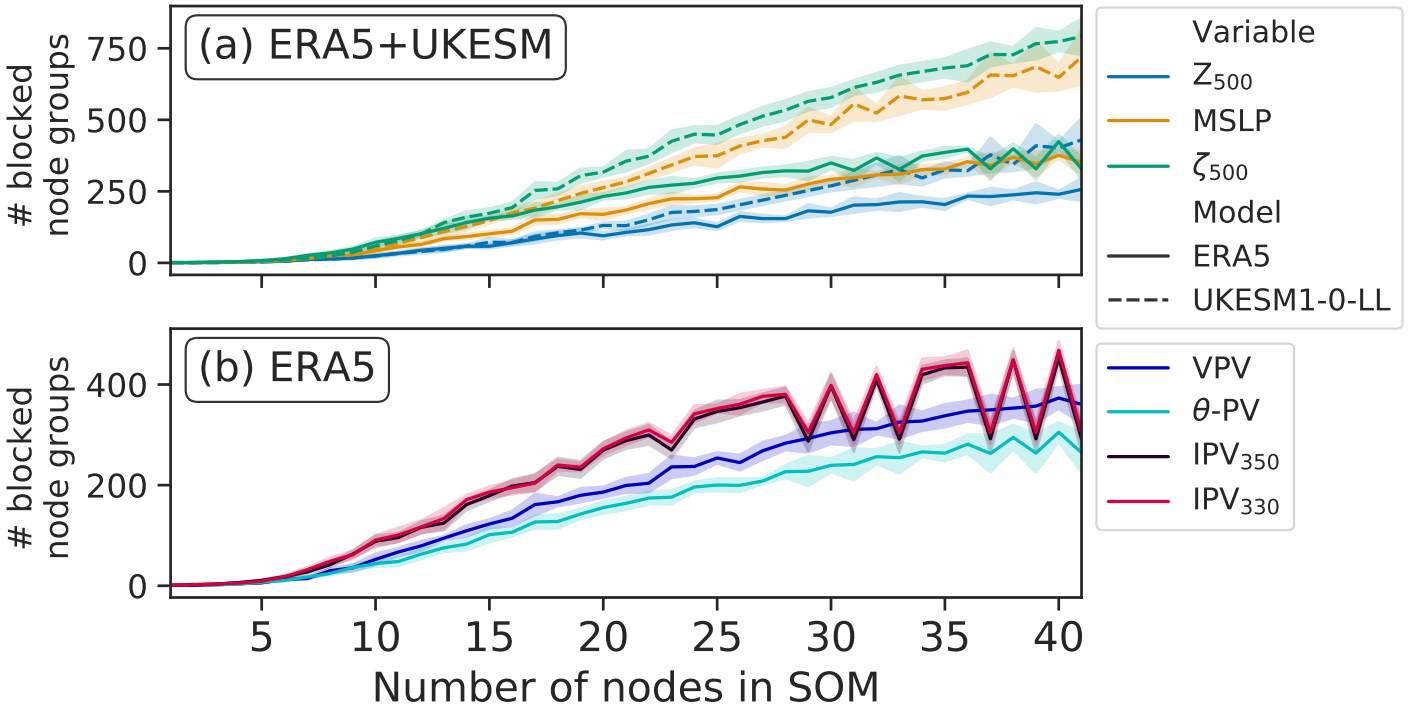

Figure 9. The number of node groups that are identified as blocked in the SOM-BI for ERA5 and UKESM for a range of node numbers and variables. The panels separate the variables available in both models from those only available in ERA5. Error bands show the standard deviation ($\pm 1\sigma$).

The results are shown for $1 \leqslant k \leqslant 41$. To measure out-of-sample skill, we used 10-fold cross-validation, where the GTD was split into 10 separate sections for testing the SOM-BI. The SOM-BI is trained on nine of the ten data sections and its skill is evaluated on the remaining section. The skill scores shown only indicate how well the SOM-BI is able to predict the test period in question, which was not used for training. This ensures that the SOM-BI has not been tuned to the data we measure our skill against, which could give it an unfair advantage compared to the other BIs. For ERA5 we used 4 year periods (1980-1984, ... , 2015-2019 inclusive) to test on and trained on the remaining 37 years, with each 4-year period once serving as the independent test set. In UKESM 10-year periods (1960-1959, ... , 2050-2059 inclusive) were used for testing the SOM-BI and it was trained on the remaining 91 year period. This 10-fold cross-validation procedure produces a range of precision, recall and F1 scores for each node number. Figures 7 and 8 show the mean values for precision, recall, F1 and the absolute difference between precision and recall. Figure 9 compares the number of groups of nodes identified as blocked for each variable. Error bands indicate the standard deviation of each skill metric ($\pm 1\sigma$).

Common features are observed for each variable for a very small or large number of SOM nodes. For small k the SOM-BI identifies more days as blocked, such that R » P. This indicates that the SOM is under-fitting the data for European circulation

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

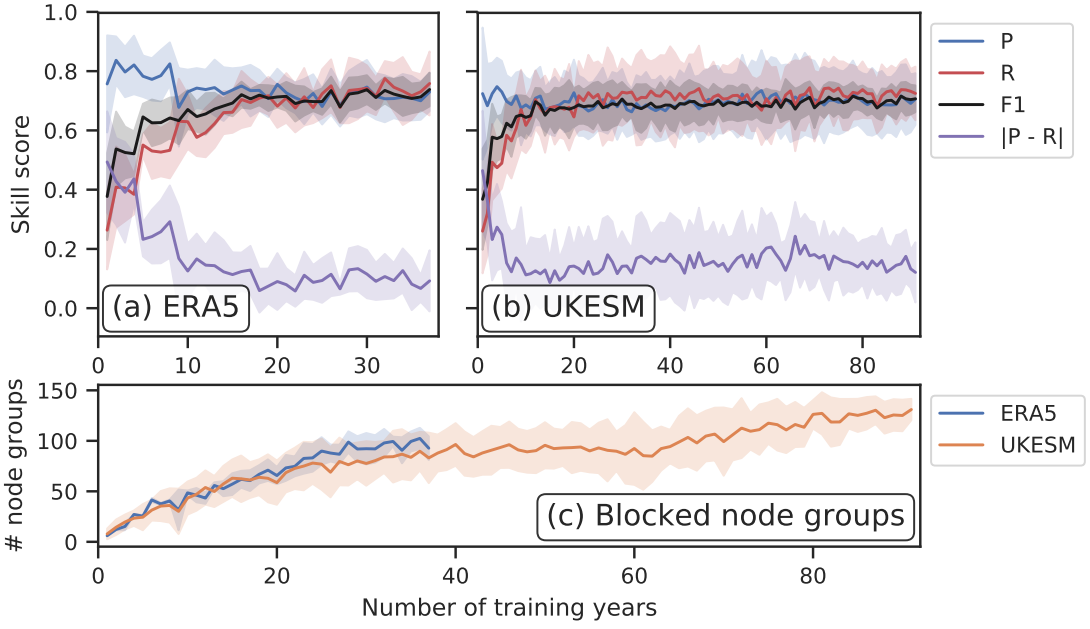

Figure 10. SOM-BI skill depending on the number of training years. (a) and (b) show the skill scores for ERA5 and UKESM, and (c) shows how the number of node groups associated with blocking varies with the length of the training record. 10-fold cross-validation is used, with 4 and 10 years used to test the SOM-BI for ERA5 and UKESM respectively. In both models Z_{500} is the variable tested with 20 (5x4) nodes in the SOM. Error bands indicate standard deviation ($\pm 1\sigma$) in the skill scores depending on the training/test set combination.

patterns across the domain and so the algorithm lacks a precise delineation of blocking events. In other words, it could be beneficial to increase k to be able to represent a larger number of dynamical states and thus to detect and describe blocking events more precisely. For large k, R « P, showing that the SOM-BI is trending towards overfitting the training data. We deduce that the optimal k occurs when the difference between P and R is small and the F1 score is close to its maximum value.

From Figs. 7 and 8 we find that for both ERA5 and UKESM the Z_{500} anomaly is the best variable to use with the SOM-BI, with a mean F1 score of 0.74 and 0.71 for $k = 20$ and 21 in ERA5 and UKESM respectively. From Fig. 9a, Z_{500} also shows the lowest number of blocked node groups for a given k, which shows that the blocked node groups are physically more explanatory in Z_{500} than the blocked node groups associated with other for other variables, making the SOM-BI results easier to interpret physically. MSLP is the second most effective variable, with an optimum F1 score of 0.66 and 0.64 for ERA5 and UKESM respectively. This lower peak performance is because the MSLP field has a lower signal-to-noise ratio as it is influenced by effects within the boundary layer such as heat lows. The PV-related variables exhibit a variety of lower skills, where the VPV field performing at a similar level to MSLP, since the vertical integration of the VPV variable enables it to capture the pattern of blocking better than other PV-based variables (Schwierz et al., 2004).

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

3.4 SOM-BI skill dependence on number of training years

One important verification for the SOM-BI is to ensure its robustness over long timescales. Contrary to the other BIs, the SOMs learn from training data. Therefore, the SOM-BI skill on test data will also be a function of how representative the training samples are of general states of the atmosphere. Here we investigate if the observational record, for example, is long enough to indeed ensure the same performance of the SOM-BI described above over longer timescales. For this purpose, we train the SOM-BI algorithm on a range of different numbers of training years, while keeping the number of years to test the algorithm performance consistent. Importantly, there is no overlap between the training and test data to ensure that the skill evaluation is truly independent, following the idea of statistical cross-validation (see e.g. Nowack et al., 2018; Mansfield et al., 2020). Figure 10 shows the results of this analysis for Z_{500} and 20 nodes across both (a) ERA5 and (b) UKESM datasets, which is the best performing case according to our analysis above. Since the datasets have different lengths (41 years vs. 101 years), we tested the model on 4 and 10 years for each dataset respectively. For a small number of years, the algorithm only sees a few blocking events and so only identifies the particular node groups that are associated with these blocking events rather than identifying node groups that are in general associated with blocking events. This leads to a high precision for a small number of training years, particularly in the ERA5 data, since the SOM-BI is effectively over-fitting on a few events, but the recall and overall F1 score are low. This behaviour is confirmed by Fig. 9 (c), which shows that there is a small set of node groups associated with blocking for a small number of training years.

Figs. 9 (a) and (b) both show that the recall and F1 scores increase asymptotically for a larger number of training years, and the precision decreases asymptotically. These variations become very small after 20 years for both ERA5 and UKESM, which indicates that for around 20 years the SOM-BI seems to have achieved an optimal performance. Figure 9 (c) shows that the number of node groups associated with blocking continues to increase in both ERA5 and UKESM even after this point, with 120 node groups identified with blocking for UKESM over 91 years compared to 95 node groups over 37 years. However, these extra node groups occur rarely in the blocking datasets since they do not significantly affect either the precision or recall of the algorithm and are therefore not physically meaningful.

3.5 Cross-comparison of SOM-BI skill

For the SOM-BI to be effectively applied to understand future trends in atmospheric blocking, we need to verify that the training of the SOM-BI on the observational record is consistent with CMIP6 models. This step is necessary to ensure that the SOM-BI can identify blocking patterns in the models. If it is possible for the SOM-BI to identify blocking patterns in a CMIP6 model from training on the observations, then this shows potential for the SOM-BI to be applied consistently across a model ensemble. Furthermore, if the SOM-BI can be trained on a CMIP6 model and tested on the observations, differences in the skill of the SOM-BI would highlight limitations in that model's ability to represent blocking patterns. This could be applied across a model ensemble to compare the skill of different models at representing blocking patterns.

To investigate the feasibility of such studies, we test the skill of the SOM-BI algorithm by training Z_{500} data on the 41 years from the ERA5 dataset and testing on the UKESM and vice versa. Table 2 shows the differences in the optimal performance

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Training dataset	Test dataset	F1 score	Number of nodes	Number of blocked node groups
ERA5	ERA5	0.74	20	95
UKESM	ERA5	0.74	21	134
UKESM	UKESM	0.71	20	131
ERA5	UKESM	0.71	19	99

Table 2. A comparison of the optimal F1 score for when Z_{500} ERA5 and UKESM datasets are trained and tested on themselves and each other respectively. The corresponding node number and number of blocked node groups is shown. When the dataset is tested on itself, 10-fold cross-validation is used and the mean value is shown. The optimal F1 score is identified by finding the node number with the smallest difference between precision and recall whilst maintaining a relatively high F1 score.

for Z_{500} across the different datasets. In all cases several node numbers were tested, and we identified an optimal node number of 20 or 21 across all the configurations of training and testing data. There was also a good performance of the SOM-BI for other node numbers that is consistent with Fig. 7a (not shown). The stable performance of the SOM-BI shows that there is a consistent range of synoptic weather patterns between the ERA5 and UKESM for European summer. It also indicates a consistency between the labeling that occurred in the GTD across ERA5 and UKESM, despite the reduced performance in the blocking indices to label the GTD. This further shows that UKESM describes blocking patterns in a similar enough way to the historical observations for useful study of blocking events, which in turn reinforces the validity of studies in blocking trends from the CMIP6 archive (Davini and D'Andrea, 2020). Finally, this underscores the potential for the SOM-BI to be used in understanding future trends and diagnosing model skill across the CMIP model ensemble.

3.6 Dependence of SOM-BI skill on other parameters

Apart from the SOM node number, number of years trained over and training dataset, there are several other parameters that could be modulated within the SOM-BI framework. First, we investigated 5-fold cross-validation on the ERA5 dataset, which involves testing the SOM-BI on 8 years of data five times. This was found to have a marginally lower performance than 10-fold cross-validation. Furthermore, we tested an alternative approach to identifying the corresponding best matching unit for the SOM pattern, where we only used positive anomalies to define the BMU. Since we are only interested in positive anomalies it is possible that such an approach would increase the skill score, particularly for events such as the 2019 European heat wave (Fig. 6). However, this modification was found to have a negligible effect on the overall SOM-BI skill.

4 Discussion and Conclusions

Using self-organizing maps (SOMs), we have developed a new blocking index (SOM-BI). This has involved the creation of a new time series dataset (GTD) to describe when blocking events have historically occurred over a region. By studying the case of European summer, we have identified a similar or better skill score for SOM-BI compared to several other blocking

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

indices (BIs) using ERA5 reanalysis data from 1979-2019. We further applied our new approach to a pre-industrial control run from UK-ESM1, and find that our method shows consistent skill for this model dataset, whereas the other BIs substantially lose performance in this case. Whilst no individual SOM node directly corresponds to a weather regime such as blocking, with an optimal node number we can develop a set of node groups which are associated with blocking. We have also found that 20 years are needed to train the SOM-BI, which underlines that the SOM-BI has a robust level of performance across the reanalysis period and in typical lengths of climate model simulations. The performance of the SOM-BI is also robust to the dataset used to train it, since it shows good performance when it is trained on the ERA5 data and tested on UKESM and vice versa. These results show that unsupervised learning can be usefully applied to understand regional blocking events, both historically and in the future.

We have confirmed that individual SOM nodes do not represent weather patterns perfectly so that care needs to be taken in using SOM patterns as a means of diagnosing weather patterns (Gibson et al., 2017b). If individual SOM nodes were used to create a blocking index, or if a small node number was used (3-6 nodes) there would be a high recall and low precision in detecting blocking using this approach, which would be the equivalent to some of the approaches applied elsewhere (e.g. Horton et al., 2015). If a higher node number (12+) was used and only one node was associated with blocking, then there would be a high precision and low recall, and overall a lower F1 score than for a low node number. However, by using a large number of nodes and studying groups of nodes across periods of five days, we have developed an algorithm that can regionally identify blocking patterns with optimal precision and recall, and which outperforms several conventional blocking indices for this task.

Using this algorithm has involved the creation of a GTD, a binary dataset that identifies regional blocking events. There are several limitations to this approach. Firstly, the choice of domain is arbitrary, and events which are on the edges of the domain are excluded, even though a large region within the domain could be considered blocked. In addition, the task of assigning a binary label to each day can be further complicated, since there is subjectivity in assigning a binary label to the onset and decay of blocking events. However, by focusing on events which are centered within the domain, a broad agreement with the identification of blocked events was achieved, despite the subjective nature of this approach. The fact that the SOM-BI exhibits consistently good skill across ERA5 and UKESM even when the SOM is trained on the opposite dataset underscores the validity of our labelling applied to both model and reanalysis data.

The use of SOMs as a blocking index is also limited by the regional approach. Whilst the BIs can be used to provide climatologies that show occurrence as in Fig. A3, the SOM-BI does not produce such a gridded global climatology. However, by studying the association of blocked events with different nodes in the SOM and identifying prominent node groups and their trends overtime, we expect that the SOM-BI can provide information about the changing nature of blocked events over a region which would be difficult to obtain from a BI. Characteristics such as the stationarity, amplitude and location of blocked events emerge from the study of the SOM nodes and their association with blocked events and other weather regimes over time.

In particular, we intend to apply this method to future trends across CMIP5 and CMIP6 models to better understand the patterns of blocking in models, diagnose model skill at reproducing the historic patterns of European circulation regimes and compare projections of future changes in blocking patterns. Since the SOM-BI explicitly identifies blocking events from unique node groups, there is great potential to study trends in these node groups, which will indicate changes both in the frequency

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

of future blocking events and their characteristics, such as stationarity, spatial extent and location. We also make our GTDs available for both ERA5 and UKESM, which have wider application in understanding historic blocking events, how they interact with other meteorological phenomena (such as heat waves and droughts) and comparing blocking patterns between reanalyses and CMIP6 models (cf. process-based climate model evaluation, Nowack et al. (2020)). We consequently encourage similar ground truth datasets to be created for other world regions and seasons, and highlight that our method could then be trained for and applied to those regions.

Code and data availability. The scripts used for the self-organizing map blocking index, the plots for this paper and the ground truth datasets for labelling of blocking events in JJA Europe (in both ERA5 1979-2019 and UKESM pre-industrial control 1960-2060) can be accessed in github.com/carlmagnusthomas/SOM-BI. ERA5 data is available from confluence.ecmwf.int and UKESM data is available from esgf-node.llnl.gov

Appendix A: UKESM case studies

In the UKESM pre-industrial run, we show in Fig. A1 part of a heat wave in the (arbitrary) year 2014. This year shows the largest spatial extent of heat extremes, where the number of grid cells exceeding the 90th (99th) temperature percentile peaks at 66% (24%) on 19 (20) July 2014. To complement this extreme case, we also show in Fig. A2 a period from the 2030 summer, which shows the edge of a blocking pattern in Eastern Europe on the 19 July and an anticyclone shifting across Europe over 20-27 July.

Since VPV is not available as a variable, the S04 blocking index cannot be calculated, and we have instead shown MSLP in Figs. A1b and A2b. Furthermore, since Z_{500} is not available at a sub-daily resolution, we only show the daily sinuosity in Figs. A1e and A2e.

Many of the same features are observed. Extreme heat is associated with persistent high pressure and stationary surface winds. The MSLP field is broadly correlated with the Z_{500} anomalies, but frequently the Z_{500} anomaly doesn't represent the MSLP anomalies well, such as on the 25th July 2014 shown in Fig. A1, where low MSLP is contrasted with high Z_{500}. The AGP index in general performs worse than in ERA5, since the zonal Z_{500} gradients are not as prominent. The DG83 index is still able to describe blocking patterns from the Z_{500} anomalies. The SOM-BI labelling is generally consistent with the ground truth dataset in both cases.

The MSLP field is broadly correlated with the Z_{500} anomalies, but frequently the Z_{500} anomaly does not represent the MSLP anomalies well, such as on the 25th July 2014 shown in Fig. A1, where low MSLP is contrasted with high Z_{500}. The surface wind fields in UKESM similarly show the easterly wind direction associated with high pressure and vice versa, particularly when the MSLP anomalies are also strong such as on the 20-21 July 2030 in Fig. A2.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

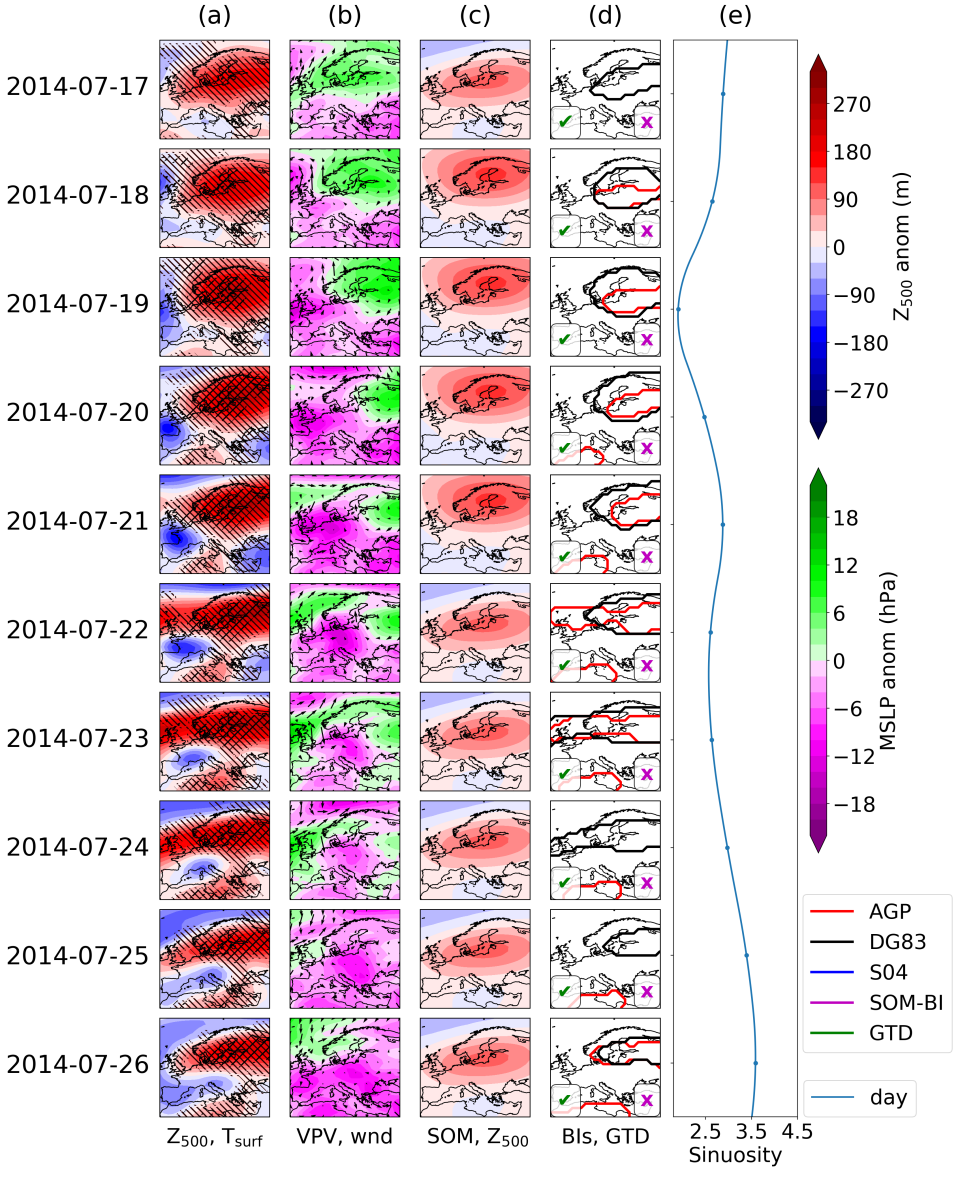

Figure A1. As with the case studies shown in figures 5 and 6, but for a heat wave in UKESM.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

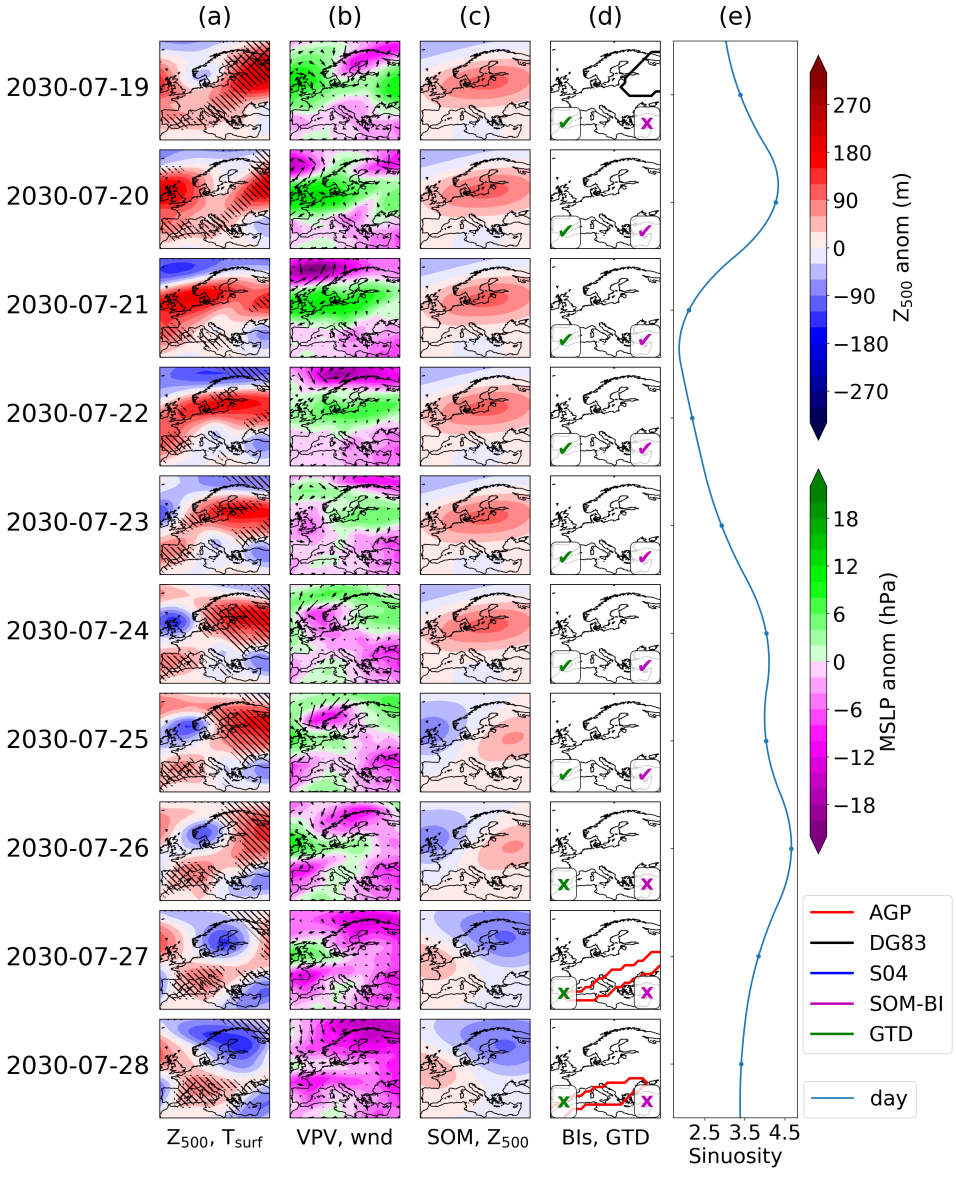

Figure A2. As above, but for a transient period in UKESM.

Author contributions. CT developed the SOM-BI method and labelled the GTDs under the supervision of PN and AV. GL developed the code for the SOMs and JH provided input into the discussion of the use of SOMs. CT wrote the initial draft of the paper and all authors contributed to its review and refinement. PN suggested the study.

Competing interests. No competing interests are present.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

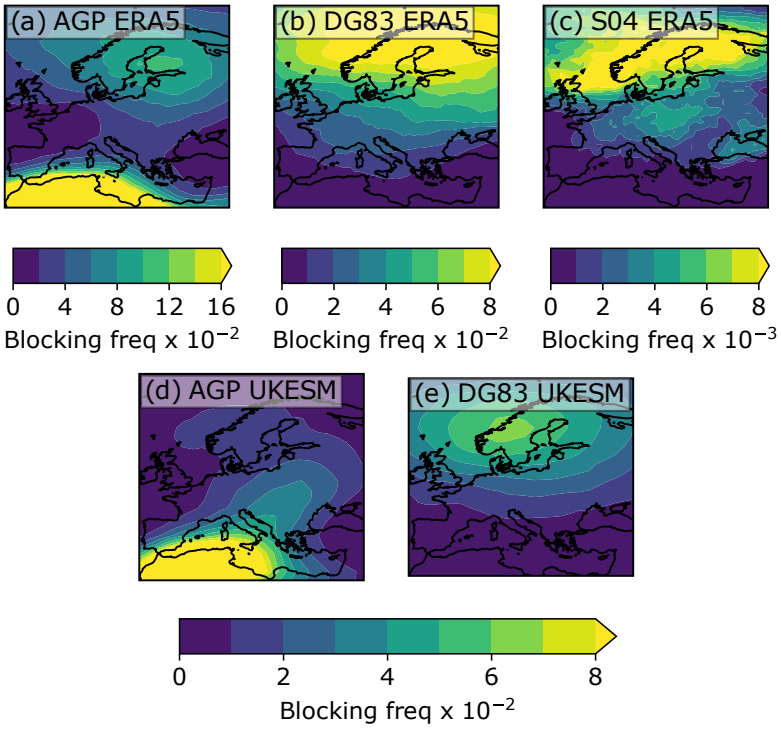

Figure A3. Occurrence of blocking events per grid cell across JJA Europe for three BIs in ERA5 1979-2019 and two BIs in UKESM JJA 1960-2060.

Acknowledgements. PN is supported through an Imperial College Research Fellowship. AV is partially funded by the Leverhulme Centre for Wildfires, Environment and Society through the Leverhulme Trust, grant RC-2018-023. CT is funded by the Natural Environment Research Council SSCP DTP, grant NE/L002515/1.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

References

Barnes, E. A.: Revisiting the evidence linking Arctic amplification to extreme weather in midlatitudes, Geophysical Research Letters, 40, 4734–4739, https://doi.org/10.1002/grl.50880, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/grl.50880, 2013.

Barnes, E. A. and Polvani, L. M.: CMIP5 Projections of Arctic Amplification, of the North American/North Atlantic Circulation, and of Their Relationship, Journal of Climate, 28, 5254–5271, https://doi.org/10.1175/JCLI-D-14-00589.1, https://doi.org/10.1175/JCLI-D-14-00589.1, 2015.

Barnes, E. A. and Screen, J. A.: The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it?, WIREs Climate Change, 6, 277–286, https://doi.org/10.1002/wcc.337, https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.337, 2015.

Barnes, E. A., Dunn-Sigouin, E., Masato, G., and Woollings, T.: Exploring recent trends in Northern Hemisphere blocking, Geophysical Research Letters, 41, 638–644, https://doi.org/10.1002/2013GL058745, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2013GL058745, 2014.

Barriopedro, D., García-Herrera, R., and Trigo, R.: Application of blocking diagnosis methods to General Circulation Models. Part I: A novel detection scheme, Climate Dynamics, 35, 1373–1391, https://doi.org/10.1007/s00382-010-0767-5, 2010.

Black, E., Blackburn, M., Harrison, G., Hoskins, B., and Methven, J.: Factors contributing to the summer 2003 European heatwave, Weather, 59, 217–223, https://doi.org/10.1256/wea.74.04, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/wea.74.04, 2004.

Blackport, R. and Screen, J. A.: Insignificant effect of Arctic amplification on the amplitude of midlatitude atmospheric waves, Science Advances, 6, https://doi.org/10.1126/sciadv.aay2880, https://advances.sciencemag.org/content/6/8/eaay2880, 2020.

Brunner, L., Hegerl, G. C., and Steiner, A. K.: Connecting Atmospheric Blocking to European Temperature Extremes in Spring, Journal of Climate, 30, 585–594, https://doi.org/10.1175/JCLI-D-16-0518.1, https://doi.org/10.1175/JCLI-D-16-0518.1, 2017.

Cattiaux, J., Vautard, R., Cassou, C., Yiou, P., Masson-Delmotte, V., and Codron, F.: Winter 2010 in Europe: A cold extreme in a warming climate, Geophysical Research Letters, 37, https://doi.org/10.1029/2010GL044613, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010GL044613, 2010.

Cattiaux, J., Peings, Y., Saint-Martin, D., Trou-Kechout, N., and Vavrus, S. J.: Sinuosity of midlatitude atmospheric flow in a warming world, Geophysical Research Letters, 43, 8259–8268, https://doi.org/10.1002/2016GL070309, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL070309, 2016.

Chen, G., Lu, J., Burrows, D. A., and Leung, L. R.: Local finite-amplitude wave activity as an objective diagnostic of midlatitude extreme weather, Geophysical Research Letters, 42, 10,952–10,960, https://doi.org/10.1002/2015GL066959, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GL066959, 2015.

Christidis, N., Jones, G., and Stott, P.: Dramatically increasing chance of extremely hot summers since the 2003 European heatwave, Nature Climate Change, 5, 46–50, https://doi.org/10.1038/nclimate2468, 2014.

Coumou, D., Di Capua, G., Vavrus, S., Wang, L., and Wang, S.: The influence of Arctic amplification on mid-latitude summer circulation, Nature Communications, 9, 2959, https://doi.org/10.1038/s41467-018-05256-8, https://doi.org/10.1038/s41467-018-05256-8, 2018.

Croci-Maspoli, M., Schwierz, C., and Davies, H. C.: A Multifaceted Climatology of Atmospheric Blocking and Its Recent Linear Trend, Journal of Climate, 20, 633–649, https://doi.org/10.1175/JCLI4029.1, https://doi.org/10.1175/JCLI4029.1, 2007.

Crum, F. X. and Stevens, D. F.: A Case Study of Atmospheric Blocking Using Isentropic Analysis, Monthly Weather Review, 116, 223–241, https://doi.org/10.1175/1520-0493(1988)116<0223:ACSOAB>2.0.CO;2, https://doi.org/10.1175/1520-0493(1988)116<0223:ACSOAB>2.0.CO;2, 1988.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Davini, P. and D'Andrea, F.: From CMIP3 to CMIP6: Northern Hemisphere Atmospheric Blocking Simulation in Present and Future Climate, Journal of Climate, 33, 10 021–10 038, https://doi.org/10.1175/JCLI-D-19-0862.1, https://doi.org/10.1175/JCLI-D-19-0862.1, 2020.

Davini, P., Cagnazzo, C., Gualdi, S., and Navarra, A.: Bidimensional Diagnostics, Variability, and Trends of Northern Hemisphere Blocking, Journal of Climate, 25, 6496–6509, https://doi.org/10.1175/JCLI-D-12-00032.1, https://doi.org/10.1175/JCLI-D-12-00032.1, 2012.

Diao, Y., Li, J., and Luo, D.: A New Blocking Index and Its Application: Blocking Action in the Northern Hemisphere, Journal of Climate, 19, 4819–4839, https://doi.org/10.1175/JCLI3886.1, https://journals.ametsoc.org/doi/abs/10.1175/JCLI3886.1, 2006.

Diffenbaugh, N. S., Singh, D., Mankin, J. S., Horton, D. E., Swain, D. L., Touma, D., Charland, A., Liu, Y., Haugen, M., Tsiang, M., and Rajaratnam, B.: Quantifying the influence of global warming on unprecedented extreme climate events, Proceedings of the National Academy of Sciences, 114, 4881–4886, https://doi.org/10.1073/pnas.1618082114, https://www.pnas.org/content/114/19/4881, 2017.

Dole, R. M. and Gordon, N. D.: Persistent Anomalies of the Extratropical Northern Hemisphere Wintertime Circulation: Geographical Distribution and Regional Persistence Characteristics, Monthly Weather Review, 111, 1567–1586, 1983.

Dunn-Sigouin, E., Son, S.-W., and Lin, H.: Evaluation of Northern Hemisphere Blocking Climatology in the Global Environment Multiscale Model, Monthly Weather Review, 141, 707–727, https://doi.org/10.1175/MWR-D-12-00134.1, https://doi.org/10.1175/MWR-D-12-00134.1, 2013.

Elliot, R. and Smith, T.: A study of the effect of large blocking highs on the general circulation in the northern hemisphere westerlies, J Meteor, 6, 67–85, 1949.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: 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, https://www.geosci-model-dev.net/9/1937/2016/, 2016.

Francis and Vavrus: Evidence for a wavier jet stream in response to rapid Arctic warming, Environmental Research Letters, 10, 014 005, https://doi.org/10.1088/1748-9326/10/1/014005, https://doi.org/10.1088%2F1748-9326%2F10%2F1%2F014005, 2015.

Francis, J. A. and Vavrus, S. J.: Evidence linking Arctic amplification to extreme weather in mid-latitudes, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL051000, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL051000, 2012.

Gibson, P., Pitman, A., Lorenz, R., and Perkins-Kirkpatrick, S.: The Role of Circulation and Land Surface Conditions in Current and Future Australian Heat Waves, Journal of Climate, 30, https://doi.org/10.1175/JCLI-D-17-0265.1, 2017a.

Gibson, P. B., Pitman, A. J., Lorenz, R., and Perkins-Kirkpatrick, S. E.: The Role of Circulation and Land Surface Conditions in Current and Future Australian Heat Waves, Journal of Climate, 30, 9933–9948, https://doi.org/10.1175/JCLI-D-17-0265.1, https://doi.org/10.1175/JCLI-D-17-0265.1, 2017b.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivity, Geophysical Research Letters, 31, https://doi.org/https://doi.org/10.1029/2003GL018747, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003GL018747, 2004.

Grotjahn, R. and Zhang, R.: Synoptic Analysis of Cold Air Outbreaks over the California Central Valley, Journal of Climate, 30, 9417–9433, https://doi.org/10.1175/JCLI-D-17-0167.1, https://doi.org/10.1175/JCLI-D-17-0167.1, 2017.

Hassanzadeh, P., Kuang, Z., and Farrell, B. F.: Responses of midlatitude blocks and wave amplitude to changes in the meridional temperature gradient in an idealized dry GCM, Geophysical Research Letters, 41, 5223–5232, https://doi.org/https://doi.org/10.1002/2014GL060764, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL060764, 2014.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren,

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803, 2020.

Hewitson, B. and Crane, R.: Self-Organizing Maps: Applications to synoptic climatology, Climate Research, 22, 13–26, https://doi.org/10.3354/cr022013, 2002.

Horton, D. E., Johnson, N. C., Singh, D., Swain, D. L., Rajaratnam, B., and Diffenbaugh, N. S.: Contribution of changes in atmospheric circulation patterns to extreme temperature trends, Nature, 522, 465 – 468, https://doi.org/10.1038/nature14550, 2015.

Hoskins, B. J., McIntyre, M. E., and Robertson, A. W.: On the use and significance of isentropic potential vorticity maps, Quarterly Journal of the Royal Meteorological Society, 111, 877–946, https://doi.org/10.1002/qj.49711147002, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711147002, 1985.

Huang, C. S. Y. and Nakamura, N.: Local Finite-Amplitude Wave Activity as a Diagnostic of Anomalous Weather Events, Journal of the Atmospheric Sciences, 73, 211–229, https://doi.org/10.1175/JAS-D-15-0194.1, https://doi.org/10.1175/JAS-D-15-0194.1, 2015.

Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynová, M., Kyselý, J., and Tveito, O. E.: Classifications of atmospheric circulation patterns: recent advances and applications., Annals of the New York Academy of Sciences, 1146, 105–52, 2008.

Johnson, N.: How many ENSO flavors can we distinguish?, Journal of Climate, 26, 4816–4827, https://doi.org/10.1175/JCLI-D-12-00649.1, 2013.

Jézéquel, A., Yiou, P., and Radanovics, S.: Role of circulation in European heatwaves using flow analogues, Climate Dynamics, https://doi.org/10.1007/s00382-017-3667-0, 2017.

Kennedy, D., Parker, T., Woollings, T., Harvey, B., and Shaffrey, L.: The response of high-impact blocking weather systems to climate change, Geophysical Research Letters, 43, 7250–7258, 2016.

Kohonen, T.: Self-organized formation of topologically correct feature maps, Biological Cybernetics, 43, 59–69, https://doi.org/10.1007/BF00337288, https://doi.org/10.1007/BF00337288, 1982.

Lejenäs, H. and Økland, H.: Characteristics of northern hemisphere blocking as determined from a long time series of observational data, Tellus A, 35A, 350–362, https://doi.org/10.1111/j.1600-0870.1983.tb00210.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0870.1983.tb00210.x, 1983.

Liniger, M. A. and Davies, H. C.: Seasonal differences in extratropical potential vorticity variability at tropopause levels, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/https://doi.org/10.1029/2004JD004639, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JD004639, 2004.

Liu, Y. and Weisberg, R. H.: Patterns of ocean current variability on the West Florida Shelf using the self-organizing map, Journal of Geophysical Research: Oceans, 110, https://doi.org/10.1029/2004JC002786, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004JC002786, 2005.

Mann, M. E., Rahmstorf, S., Kornhuber, K., Steinman, B. A., Miller, S. K., Petri, S., and Coumou, D.: Projected changes in persistent extreme summer weather events: The role of quasi-resonant amplification, Science Advances, 4, https://doi.org/10.1126/sciadv.aat3272, 2018.

Mansfield, L. A., Nowack, P. J., Kasoar, M., Everitt, R. G., Collins, W. J., and Voulgarakis, A.: Predicting global patterns of long-term climate change from short-term simulations using machine learning, npj Climate and Atmospheric Science, 3, 44, https://doi.org/10.1038/s41612-020-00148-5, https://doi.org/10.1038/s41612-020-00148-5, 2020.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Mioduszewski, J. R., Rennermalm, A. K., Hammann, A., Tedesco, M., Noble, E. U., Stroeve, J. C., and Mote, T. L.: Atmospheric drivers of Greenland surface melt revealed by self-organizing maps, Journal of Geophysical Research: Atmospheres, 121, 5095–5114, https://doi.org/https://doi.org/10.1002/2015JD024550, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015JD024550, 2016.

Mitchell, D., Kornhuber, K., Huntingford, C., and Uhe, P.: The day the 2003 European heatwave record was broken, The Lancet Planetary Health, 3, e290–e292, https://doi.org/10.1016/s2542-5196(19)30106-8, https://doi.org/10.1016/s2542-5196(19)30106-8, 2019.

Nakamura, N. and Huang, C. S. Y.: Atmospheric blocking as a traffic jam in the jet stream, Science, 361, 42–47, https://doi.org/10.1126/science.aat0721, https://science.sciencemag.org/content/361/6397/42, 2018.

Nowack, P., Braesicke, P., Haigh, J., Abraham, N., Pyle, J., and Voulgarakis, A.: Using machine learning to build temperature-based ozone parameterizations for climate sensitivity simulations, Environmental Research Letters, 13, 104 016, https://doi.org/10.1088/1748-9326/aae2be, https://iopscience.iop.org/article/10.1088/1748-9326/aae2be, 2018.

Nowack, P., Runge, J., Eyring, V., and Haigh, J. D.: Causal networks for climate model evaluation and constrained projections, Nature Communications, 11, 1415, https://doi.org/10.1038/s41467-020-15195-y, https://doi.org/10.1038/s41467-020-15195-y, 2020.

Nowack, P. J., Braesicke, P., Luke Abraham, N., and Pyle, J. A.: On the role of ozone feedback in the ENSO amplitude response under global warming, Geophysical Research Letters, 44, 3858–3866, https://doi.org/https://doi.org/10.1002/2016GL072418, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016GL072418, 2017.

Palmer, T. N.: A Nonlinear Dynamical Perspective on Climate Prediction, Journal of Climate, 12, 575–591, https://doi.org/10.1175/1520-0442(1999)012<0575:ANDPOC>2.0.CO;2, https://doi.org/10.1175/1520-0442(1999)012<0575:ANDPOC>2.0.CO;2, 1999.

Pelly, J. L. and Hoskins, B. J.: A New Perspective on Blocking, Journal of the Atmospheric Sciences, 60, 743 – 755, https://doi.org/10.1175/1520-0469(2003)060<0743:ANPOB>2.0.CO;2, https://journals.ametsoc.org/view/journals/atsc/60/5/1520-0469_2003_060_0743_anpob_2.0.co_2.xml, 2003.

Pinheiro, M. C., Ullrich, P. A., and Grotjahn, R.: Atmospheric blocking and intercomparison of objective detection methods: flow field characteristics, Climate Dynamics, https://doi.org/10.1007/s00382-019-04782-5, https://doi.org/10.1007/s00382-019-04782-5, 2019.

Rex, D. F.: Blocking Action in the Middle Troposphere and its Effect upon Regional Climate, Tellus, 2, 275–301, https://doi.org/10.1111/j.2153-3490.1950.tb00339.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2153-3490.1950.tb00339.x, 1950.

Robine, J.-M., Cheung, S. L. K., Roy, S. L., Oyen, H. V., Griffiths, C., Michel, J.-P., and Herrmann, F. R.: Death toll exceeded 70,000 in Europe during the summer of 2003, Comptes Rendus Biologies, 331, 171 – 178, https://doi.org/https://doi.org/10.1016/j.crvi.2007.12.001, http://www.sciencedirect.com/science/article/pii/S1631069107003770, dossier : Nouveautés en cancérogenèse / New developments in carcinogenesis, 2008.

Sánchez-Benítez, A., Barriopedro, D., and García-Herrera, R.: Tracking Iberian heatwaves from a new perspective, Weather and Climate Extremes, 28, 100 238, https://doi.org/10.1016/j.wace.2019.100238, 2019.

Scaife, A. A., Woollings, T., Knight, J., Martin, G., and Hinton, T.: Atmospheric Blocking and Mean Biases in Climate Models, Journal of Climate, 23, 6143 – 6152, https://doi.org/10.1175/2010JCLI3728.1, https://journals.ametsoc.org/view/journals/clim/23/23/2010jcli3728.1.xml, 2010.

Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M., and Russo, S.: Influence of blocking on Northern European and Western Russian heatwaves in large climate model ensembles, Environmental Research Letters, 13, 054 015, https://doi.org/10.1088/1748-9326/aaba55, https://doi.org/10.1088/1748-9326/aaba55, 2018.

https://doi.org/10.5194/wcd-2021-1
Preprint. Discussion started: 14 January 2021

Scherrer, S., Croci-Maspoli, M., Schwierz, C., and Appenzeller, C.: Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region, International Journal of Climatology, 26, 233 – 249, https://doi.org/10.1002/joc.1250, 2006.

Schwierz, C., Croci-Maspoli, M., and Davies, H. C.: Perspicacious indicators of atmospheric blocking, Geophysical Research Letters, 31, https://doi.org/10.1029/2003GL019341, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003GL019341, 2004.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, Journal of Advances in Modeling Earth Systems, 11, 4513–4558, https://doi.org/10.1029/2019MS001739, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001739, 2019.

Shepherd, T. G.: Atmospheric circulation as a source of uncertainty in climate change projections, Nature Geoscience, 7, 703 EP –, https://doi.org/10.1038/ngeo2253, perspective, 2014.

Sheridan, S. C. and Lee, C. C.: The self-organizing map in synoptic climatological research, Progress in Physical Geography: Earth and Environment, 35, 109–119, https://doi.org/10.1177/0309133310397582, https://doi.org/10.1177/0309133310397582, 2011.

Singh, D., Swain, D. L., Mankin, J. S., Horton, D. E., Thomas, L. N., Rajaratnam, B., and Diffenbaugh, N. S.: Recent amplification of the North American winter temperature dipole, Journal of Geophysical Research: Atmospheres, 121, 9911–9928, https://doi.org/10.1002/2016JD025116, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025116, 2016.

Skific, N. and Francis, J.: Self-Organizing Maps: A Powerful Tool for the Atmospheric Sciences, https://doi.org/10.5772/54299, 2012.

Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., eds.: IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change., 2013.

Stott, P. A., Stone, D. A., and Allen, M. R.: Human contribution to the European heatwave of 2003, Nature, 432, 610–614, https://doi.org/10.1038/nature03089, https://doi.org/10.1038/nature03089, 2004.

Tibaldi, S. and Molteni, F.: On the operational predictability of blocking, Tellus A, 42, 343–365, https://doi.org/https://doi.org/10.1034/j.1600-0870.1990.t01-2-00003.x, https://www.onlinelibrary.wiley.com/doi/abs/10.1034/j.1600-0870.1990.t01-2-00003.x, 1990.

Verdecchia, M., Visconti, G., D'Andrea, F., and Tibaldi, S.: A Neural Network Approach for blocking recognition, Geophysical Research Letters, 23, 2081–2084, https://doi.org/10.1029/96GL01810, 1996.

Wittek, P., Gao, S. C., Lim, I. S., and Zhao, L.: somoclu: An Efficient Parallel Library for Self-Organizing Maps, Journal of Statistical Software, 78, https://doi.org/10.18637/jss.v078.i09, http://dx.doi.org/10.18637/jss.v078.i09, 2017.

Woollings, T., Barriopedro, D., Methven, J., Son, S.-W., Martius, O., Harvey, B., Sillmann, J., Lupo, A. R., and Seneviratne, S.: Blocking and its Response to Climate Change, Current Climate Change Reports, 4, 287–300, https://doi.org/10.1007/s40641-018-0108-z, https://doi.org/10.1007/s40641-018-0108-z, 2018.

Xu, G., Osborn, T. J., Matthews, A. J., and Joshi, M. M.: Different atmospheric moisture divergence responses to extreme and moderate El Niños, Climate Dynamics, 47, 393–410, https://doi.org/10.1007/s00382-015-2844-2, https://doi.org/10.1007/s00382-015-2844-2, 2016.