AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2022MS003038

Key Points:

- We show how to make climate model results easier to use by reducing large ensembles to just the ensemble mean and a single pattern
- We apply the methods to European rainfall in two climate model ensembles, and gain new insights into the structure of likely changes
- Comparing ways to define the representative deviation, we find that simple compositing and a linear method give similar results

Correspondence to:
S. Jewson,
stephen.jewson@gmail.com

Citation:
Jewson, S., Messori, G., Barbato, G., Mercogliano, P., Mysiak, J., & Sassi, M. (2023). Developing representative impact scenarios from climate projection ensembles, with application to UKCP18 and EURO-CORDEX precipitation. *Journal of Advances in Modeling Earth Systems*, *15*, e2022MS003038. https://doi.org/10.1029/2022MS003038

Received 9 FEB 2022
Accepted 19 DEC 2022

Author Contributions:
Conceptualization: Stephen Jewson
Data curation: Giuliana Barbato, Paola Mercogliano, Jaroslav Mysiak
Formal analysis: Stephen Jewson
Investigation: Stephen Jewson
Methodology: Stephen Jewson
Project Administration: Stephen Jewson
Resources: Stephen Jewson
Software: Stephen Jewson, Maximiliano Sassi
Validation: Stephen Jewson
Visualization: Stephen Jewson

Developing Representative Impact Scenarios From Climate Projection Ensembles, With Application to UKCP18 and EURO-CORDEX Precipitation

Stephen Jewson[1], Gabriele Messori[2,3], Giuliana Barbato[4], Paola Mercogliano[4], Jaroslav Mysiak[4], and Maximiliano Sassi[5]

[1]Lambda Climate Research Ltd., London, UK, [2]Department of Earth Sciences and Centre of Natural Hazards and Disaster Science (CNDS), Uppsala University, Uppsala, Sweden, [3]Department of Meteorology and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden, [4]Euro-Mediterranean Center on Climate Change (CMCC) Foundation, Lecce, Italy, [5]Risk Management Solutions Ltd., London, UK

Abstract Calculating impacts from climate projection ensembles can be challenging. A simple approach might consider just the ensemble mean, but this ignores much of the information in the ensemble and does not explore the range of possible impacts. A more thorough approach would consider every ensemble member, but may be computationally unfeasible for many impact models. We investigate the compromise in which we represent the ensemble by the mean and a single deviation from the mean. The deviation from the mean would ideally be representative both of variability in the ensemble, and have a significant impact, according to some impact metric. We compare methods for calculating the deviation from the mean, based on traditional compositing and a statistical method known as Directional Component Analysis (DCA). DCA is based on linearizing the impact metric around the ensemble mean. We illustrate the methods with synthetic examples, and derive new mathematical results that clarify the interpretation of DCA. We then use the methods to derive scenarios from the UKCP18 and EURO-CORDEX projections of future precipitation in Europe. We find that the worst ensemble member is not robust, but that deviations from the ensemble mean calculated using compositing and DCA are robust. They thus give robust insight into the patterns of change in the ensemble. We conclude that mean and representative deviation methods may be suitable for climate projection users who wish to explore the implications of the uncertainty around the ensemble mean without having to calculate the impacts of every ensemble member.

Plain Language Summary Using the results from climate models to force models that calculate the impacts of climate change can be challenging when ensembles are large. For detailed impacts models, it may not be computationally feasible to use all the data. We investigate ways to reduce the ensemble of different results down to just an ensemble of two patterns of change: one pattern for the mean, and one pattern for variations around the mean. Using these two patterns allows us to estimate the typical impact and the possible range of impacts. The pattern for variations around the mean we call a "representative deviation pattern." An ideal representative deviation pattern would capture the most relevant differences from the mean, where what defines relevance may vary according to the application. We compare simple and traditional methods for deriving the representative deviation pattern with a recently derived statistical method called Directional Component Analysis. We apply all the methods to help better understand the changes in rainfall expected in Europe due to climate change. We find robust results that show that the deviation from the mean with the greatest total rainfall consists of a pattern with more rainfall in most of Europe but less rainfall in the far North.

1. Introduction

There is an extensive literature that describes how users of climate information might process ensemble climate projections, covering both Single Model Ensembles (SMEs) and Multi-Model Ensembles (MMEs). The methods described include methods for bias correction (Chen et al., 2019; Knutti et al., 2017; Maraun, 2016), methods for identifying the forced response (Barnes et al., 2019; Sippel et al., 2019; Wills et al., 2020), analysis of the potential benefits of spatial smoothing (Masson & Knutti, 2011; Räisänen & Ylhäisi, 2010), methods for assigning weights to the different models in an MME (Abramowitz et al., 2019; Knutti et al., 2017; Sanderson et al., 2017), methods for understanding uncertainty (Deser et al., 2010; Frankcombe et al., 2015; Hawkins & Sutton, 2009;

Writing – original draft: Stephen Jewson
Writing – review & editing: Stephen Jewson, Gabriele Messori, Giuliana Barbato, Paola Mercogliano

Thompson et al., 2015), methods for reducing uncertainty (Hall et al., 2019; Jewson et al., 2021), and methods for creating smaller and hence more manageable ensembles (Evans et al., 2013; Herger et al., 2018).

In this article, we will revisit the broad question of how to make effective use of the information in climate model ensembles, with a focus on creating climate impact scenarios. We imagine a hypothetical user of climate projections who currently considers only the ensemble mean from an SME such as the UKCP18 Regional Projection Ensemble (Lowe et al., 2018) or from an MME such as the CMIP5 project (Taylor et al., 2012), the CMIP6 project (Eyring et al., 2016) or the EURO-CORDEX project (Jacob et al., 2014, 2020). They use an impact model, perhaps for calculating total flood or wind damages across a large region, or for calculating the impact of climate change on GDP. Their impact model already incorporates simulations of weather variability, as is common, for instance, in the impact models used in the insurance industry. The reason for using climate models is to understand possible changes in statistics of weather variability under climate change. They look at impacts of the ensemble mean for multiple cases based on the four main RCPs (Moss et al., 2010) or the five SSPs (Riahi et al., 2017), for different points in time and for different adaptation strategies. Their impact models are complex, and the use of the ensemble mean already entails significant computational costs. They would like to extend their analysis to consider the impact of uncertainty around the ensemble mean, and to estimate the range of possible impacts, but running the impact model for every ensemble member for every case is not feasible. Is there a framework that would allow them to understand the impacts of the uncertainty around the ensemble mean, including the range of possible impacts, without having to use the whole ensemble?

To bridge the gap between the two extremes of analyzing the impacts of only the ensemble mean and analyzing the impacts of every ensemble member, one might consider using only a subset of the members of the full ensemble, either chosen randomly, or according to some kind of algorithm (Evans et al., 2013; Herger et al., 2018). The smallest subset, apart from using just the ensemble mean, would be to use the ensemble mean and one additional pattern, so that the full impact model only has to be run twice. This approach has recently been studied by Scher et al. (2021) (henceforth SJM) in the context of a 50-member ensemble of medium range forecasts of temperature. To determine which single additional pattern to use, SJM considered a linear function that gives a measure of which ensemble members are most important in terms of impact, where impact is considered undesirable. They then compared four different methods that could be used to find a single reasonable worst-case deviation from the ensemble mean, given the linear impact function. "Reasonable" in the phrase reasonable worst-case means that the probability of this pattern should not be too low, so that the pattern is representative of the ensemble, and "worst" means that the pattern should have a reasonably large positive value of the linear impact function, similar to the more extreme members of the ensemble. The four methods they considered were: taking the worst single member from the ensemble, creating a composite pattern by averaging the worst 10% of the members of the ensemble, and two methods based on Directional Component Analysis (DCA, Jewson (2020)). DCA is a method for deriving patterns of variability from space-time data sets such that the patterns have both a large linear impact and a high probability density. They compared these four methods and found the two DCA methods to be the most robust (i.e., the results varied less when the methodology was subjected to small changes).

We build upon the ideas in Jewson (2020) and SJM, but in a different context and with a different goal. Unlike SJM, we do not try to find a reasonable worst-case, which is by definition fairly extreme. We instead seek to find a pattern that represents both the likely variability of the ensemble around the mean, and captures the possible impact of the variability around the mean. Such a pattern is not necessarily extreme, and we will call this pattern a representative deviation, where "representative" implies both that the pattern should be likely in terms of probability of occurrence, and significant in terms of impact. We can then calculate a typical impact from the ensemble mean pattern, and estimate the range of possible impacts using the representative deviation pattern. We extend the previous work in four ways. First, we provide more context for DCA by discussing in detail how the linear impact function used in DCA can be considered as a linearization of a non-linear impact function, and how the method can be generalized to higher order approximations. Second, we provide illustrative examples of four methods that can be used to identify a representative deviation, using ensembles of 10 members. The methods are: the single ensemble member with the largest impact (W1), the mean of the two ensemble members with the largest impact (W2), the mean of the four ensemble members with the largest impact (W4), and DCA. For these examples we imagine applying the methods to climate projections of precipitation and include cases in which the definition of linear impact is based on increased precipitation in all regions, and a case in which it is based on increased precipitation in some regions, and decreased precipitation in others. This latter case is motivated by the situation in Europe where parts of northern Europe are projected to experience an increase in precipitation

because of climate change, while parts of southern Europe are projected to experience a decrease in precipitation (European Environment Agency, 2017). In this situation, a Europe-wide large impact case might consist of greater than the ensemble mean precipitation in northern Europe and less than the ensemble mean precipitation in southern Europe. Third, we derive various additional mathematical properties of DCA, that help to clarify the interpretation of DCA and its relationship to W2, W4, and compositing methods in general. Fourth, we use our framework to derive impact scenarios from state-of-the-art high-resolution climate simulations from the UKCP18 and EURO-CORDEX projects.

In Section 2 we describe the data we will use. In Section 3 we review the DCA methodology and discuss how to generalize DCA to higher order impact functions. In Section 4 we illustrate our four methods for finding representative deviation patterns. In Section 4 we present new mathematical properties of DCA that help clarify the interpretation. In Section 6 we apply the representative deviation methodologies to results from the UKCP18 and EURO-CORDEX climate model ensembles. In Section 7 we draw some conclusions. Appendices A, B, C, D, and E include a number of related mathematical derivations.

2. Data Sources

2.1. Synthetic Data

In Section 4 we analyze synthetic ensemble data to illustrate the representative deviation methods in various ways. The synthetic data was created by sampling from a bivariate normal distribution. The two dimensions are intended to represent changes in future time-mean precipitation values at two different but nearby locations. Each ensemble member in the synthetic data is intended to represent output from a different member from a climate model ensemble. We simulate ensembles of both 10 and 1,000 members, and we also simulate 500 repeats of the 10 member ensemble to understand the robustness of the results to the statistical sampling involved in creating the ensembles. Thousand members is larger than most climate ensembles, but using 1,000 member ensembles is helpful for illustrating how the methods work.

2.2. UKCP18 Data

The first set of climate model output we use in Section 6 was produced by the United Kingdom Meteorological Office (UKMO) as part of the 2018 iteration of the United Kingdom Climate Projections (UKCP18) project (Lowe et al., 2018). The output comprises an SME consisting of 12 limited area atmospheric model runs at 12 km horizontal resolution, produced by the UKMO Unified Model Global Atmosphere GA7, and driven by perturbed variants of the UKMO global climate model HadGEM3-GC3.05, all for RCP8.5. These runs cover Europe for the period 1981–2080. The boundary conditions for the climate model simulations were observed and projected forcings including, in particular, increasing green-house gases. The climate model runs were started from different initial conditions in 1900, and use random numbers to perturb the physical parametrizations (see the explanation in Lowe et al. (2018)). We will consider the changes in precipitation climate between the periods 1981–2010 and 2011–2040, which we will refer to as precipitation changes. As a result of the initialization, the model states of the different ensemble members during the periods that we are studying would be expected to be randomly out of phase with respect to each other in terms of decadal and interdecadal climate variability, which we require for our study. We only use 10 of the 12 members for our analysis, in order to better compare with our EURO-CORDEX ensemble, which only has 10 members. We have also calculated the UKCP18 results using 12 members, and they were not materially different from the results calculated using 10 members (not shown).

2.3. EURO-CORDEX Data

The second set of climate model data we use in Section 6 is precipitation data extracted from the EURO-CORDEX ensemble projections of future climate (Benestad et al., 2017; Jacob et al., 2014, 2020). EURO-CORDEX is an MME, and we use data from 10 separate projections, simulated by 10 different combinations of global models and regional models: a list of the models used is given in Table 1. The model output is at 0.11-degree resolution (roughly 12 km). Once again, we will consider the changes in precipitation climate between the periods 1981–2010 and 2011–2040, for RCP8.5.

Table 1
The EURO-CORDEX Climate Models Used in This Study

Model	Driving GCM	GCM member	RCM
M1	CNRM-CM5	r1i1p1	ALADIN53
M2	IPSL-CM5A-MR	r1i1p1	RCA4
M3	CNRM-CM5	r1i1p1	RCA4
M4	CNRM-CM5	r1i1p1	CCLM4-8-17
M5	EC-EARTH	r12i1p1	CCLM4-8-17
M6	EC-EARTH	r12i1p1	RACMO22E
M7	EC-EARTH	r12i1p1	RCA4
M8	EC-EARTH	r1i1p1	RACMO22E
M9	EC-EARTH	r3i1p1	HIRHAM5
M10	IPSL-CM5A-MR	r1i1p1	WRF331F

3. Mathematical Definitions and Basic Properties

For all the ensembles we consider, the ensemble members are converted to anomalies from the ensemble mean and written as vectors x, each with M components corresponding to the number of spatial points multiplied by the number of variables being considered. In general, for the mathematical methods we discuss, the ensemble members could have zero, one, two or three spatial dimensions, and a time dimension, and could include multiple variables, but in all cases the data would be reduced to vectors with M components in this way. In our examples we consider patterns of precipitation with two horizontal spatial dimensions, representing time average precipitation at the surface. Given an ensemble of N members we write the N vectors in a matrix X. This matrix has dimensions M by N. The empirical spatial covariance matrix of the vectors, which gives the empirical covariance between every possible pair of spatial points, is given by $C = XX^T/N$, which is an M by M matrix.

We define the impact of a climate state y as $a = a(y)$. We make the simplifying assumption that a is a scalar, such as the total flood or wind damage, or gross domestic product. Using a scalar allows us to consider minimizing the impact. In general, the impact of a climate state is likely to be a complex non-linear function: consider, for instance, the relationships between changes in precipitation and flood damage, or changes in wind-speed and wind damage, which are highly non-linear.

We write the ensemble mean as x_0, so that $y = x_0 + x$. We can then expand the function a around the ensemble mean as:

$$a(y) = a(x_0) + x^T \nabla a(x_0) + \frac{1}{2} x^T H(x_0)\, x + \ldots \tag{1}$$

where $a(x_0)$ is the impact of the ensemble mean, $\nabla a(x_0)$ indicates a vector of the M derivatives of a, evaluated at the ensemble mean, and $H(x_0)$ indicates an M by M square matrix of second derivatives of a, known as a Hessian matrix, also evaluated at the ensemble mean. For a general non-linear function, the impact of the ensemble mean $a(x_0)$ is unlikely to be equal to the mean impact of the ensemble. For a linear function, however, they would be equal.

If we write $r = \nabla a(x_0)$ and neglect higher order terms, this equation becomes a linear equation for impact:

$$a(y) = a(x_0) + x^T r \tag{2}$$

If we write $a_0(x_0) = a(x_0)$, and $a'(x) = x^T r$ then this becomes

$$a(y) = a_0(x_0) + a'(x) \tag{3}$$

The linear term $a'(x) = x^T r$ is just a weighted sum of values in the anomaly pattern x, where the vector r gives the weights on each location. Equivalently r can be interpreted as the direction of the gradient of the full non-linear impact, in the M dimensional space of x, pointing toward values of highest impact. This linear approximation only has to be a good approximation across the range of variability in the ensemble being considered, and not across the entire range of variability of y and x.

One of simplest possible examples of impact that can be expressed using $a = a_0 + a'$ is the total precipitation change (change relative to the previous climate) across a precipitation field. This is vastly simpler, as an impact model, than models for total flood damage, wind damage, or change in GDP, but is useful as an example. If x is the spatial field of precipitation anomalies, a_0 is the total precipitation change of the ensemble mean and r is a vector with components that account for the area represented by each data location, then a is the total precipitation change due to the ensemble mean plus the anomaly pattern x. A more complex example might involve using the components of the vector r to apply weighting as a function of spatial variation in population density.

In the rest of this study we will assume that the impact is linear, as given by Equation 2, unless stated otherwise. It may be linear because the impact is genuinely linear, or it may be the linearization of a non-linear impact

function. We provide some discussion of the possible extensions that include non-linear terms, to provide some additional context for the linear analysis, in Sections 3.2.3 and 5.2.

3.1. Definitions of W1, W2, and W4

The W1 method consists of just taking the member of the ensemble with the highest linear impact. The W2 and W4 methods are composites that consist of taking the average of the two and four members with the highest linear impacts, respectively.

These simple methods could be used, without any further analysis, as methods for defining representative deviations from the ensemble mean. However, it is also worth exploring the use of DCA, following SJM, for two reasons. First, SJM showed that in their example DCA gave more robust estimates than using a composite average of the worst 10% of the ensemble, and second, the theoretical framework on which DCA is based, and the various mathematical results concerning the properties of DCA, can help with interpretation of the patterns derived.

3.2. Definition of DCA

The derivation of DCA in Jewson (2020) is based on assuming that the variability in the ensemble can be modeled as a MultiVariate Normal (MVN) distribution, and trying to find the value of the vector x that maximizes probability density for a given linear impact. The derivation then involves solving a Lagrangian maximization problem with the Lagrangian L given by:

$$L = -x^T C^{-1} x + 2\lambda(x^T r - a_c) \tag{4}$$

where λ is a Lagrange multiplier and a_c is a given level of the linear impact. This has the solution $x \propto Cr$, which is an M component vector in the direction of the vector Cr. This vector can be interpreted as a spatial pattern. The direction of this vector does not depend on the value of the Lagrange multiplier λ or the level of impact a_c. We write the first DCA vector as g, and define the direction of g as the direction of Cr, and so

$$g \propto Cr \tag{5}$$

Since $C = XX^T/N$, we see that $g \propto (XX^T/N)r \propto X(X^T r)$, which can also be written as

$$g \propto Xv \tag{6}$$

where $v = X^T r$ is a vector that gives the linear impacts of the individual members of the ensemble. Based on these expressions we see that the direction of the first DCA vector g can be calculated either using the covariance matrix ($g \propto Cr$) or by calculating the linear-impact-weighted average of the ensemble members ($g \propto Xv$). The covariance-based expression makes DCA look like a parametric method, while the weighted average expression shows that DCA can also be interpreted as a non-parametric method that consists of an impact weighted composite average.

Equations 5 and 6 are proportional relationships and define the direction of the first DCA vector g (i.e., the shape of the spatial pattern described by g), but not the magnitude of the vector (i.e., not the amplitude of the spatial pattern described by g). To make the definition of the first DCA vector unique, Jewson (2020) defines the first DCA pattern as a unit vector, that is, a vector for which $g^T g = 1$. The length of the unit vector DCA pattern can then be scaled to an appropriate value, depending on the application. Subsequent DCA patterns after the first can also be derived (see Jewson (2020)) to create a basis of orthogonal spatial patterns (patterns for which ${x_1}^T x_2 = 0$ for all combinations of patterns x_1 and x_2) and a corresponding method for matrix factorization, although in this study we will only consider the first pattern, which we will refer to simply as *the* DCA pattern or DCA vector.

3.2.1. Mathematical Properties of DCA

The mathematical properties of DCA depend on the statistical distribution of the ensemble. Jewson (2020) shows that if the ensemble is distributed as an MVN, which allows us to calculate a probability density for every possible pattern vector x, then DCA is the anomaly spatial pattern that maximizes the probability density among all patterns that correspond to a given value of the linear impact function (*property* 1). Part of the utility of DCA

arises from the fact that the direction of the DCA vector is then independent of the level of linear impact used in this definition, and depends only on the direction of the gradient of linear impact. In other words, whatever level of impact is of interest, DCA is the pattern that maximizes the probability density. Conversely, the DCA pattern is also the anomaly spatial pattern that maximizes the linear impact for a given value of the probability density (*property 2*). These are the mathematical properties that justify the use of the DCA pattern as a pattern that is representative of deviations from the ensemble mean (because it maximizes probability density in property 1), and that captures representative changes in the impact (because it maximizes impact in property 2). We will present some further properties that can be used to define DCA in an even simpler way in Section 5 below.

If the ensemble is not MVN, but is still elliptically distributed (for example, is distributed as a multivariate *t* distribution) then W1, W2, W4, and DCA all give patterns that would be plausible as ensemble members, and these two properties of DCA may still hold. More precise details are given in Appendix A.

If the ensemble is not elliptically distributed it is useful to distinguish between what we will call *smoothable* and *non-smoothable* ensembles. We define smoothable ensembles as those that have the property that weighted linear combinations of events in the ensemble are plausible alternative members of the ensemble, while non-smoothable ensembles are those which do not have this property. Climate projections of seasonal mean temperature might be considered smoothable, since there are no obvious constraints on what such a field might look like, and averaging together ensemble members in some way will likely create a pattern that looks like it could be another ensemble member. On the other hand, short-term weather forecasts of localized convective precipitation would probably not be considered smoothable, because averaging together convective events from different ensemble members does not create a realistic field of convective events. For ensembles that are not elliptically distributed but are still smoothable, W1, W2, W4, and DCA can be considered as non-parametric methods for creating representative scenarios, but DCA no longer possesses optimality properties related to impact and probability density. For ensembles that are non-smoothable, W1, W2, W4, and DCA could all be calculated, but only W1 would give a plausible ensemble member (since it is an ensemble member). W2, W4, DCA, and indeed the ensemble mean are not plausible as possible members of the ensemble. An example of the ensemble mean, various composite patterns and the DCA pattern for a non-smoothable ensemble, in which they clearly cannot be considered as possible ensemble members, is given in SJM, Appendix A.

3.2.2. Scaling of DCA

In several of the examples in this article we will scale the length of the DCA vector so that it gives a reasonable representation of the variability in the ensemble, as follows. First, we determine the DCA direction g using Equation 5 above and normalize g to be a unit vector. Second, we project all the members of the ensemble onto the unit vector g to create a series of projected values. Third, we calculate the standard deviation of this series of projections and set the length of the DCA vector to a value of 1.5 standard deviations. This gives a pattern which is 1.5 standard deviations of ensemble variability away from the ensemble mean in the direction of the DCA pattern, and hence represents a reasonable level of both ensemble variability and linear impact relative to the ensemble members. Within this method for scaling, the choice of 1.5 standard deviations is arbitrary, and could be replaced by any other number of standard deviations, or percentiles of the distribution of the projected values.

3.2.3. Generalizations of DCA to Non-Linear Impact Functions

DCA, as defined above, uses a linear impact function. Real-world impact is likely, in most cases, to be non-linear, at least to some degree. The linear definition of impact is the simplest possible formulation that allows us to explore the connections between the impact of a pattern and its probability density, and to determine from that a straightforward methodology for defining representative deviation patterns. In some cases, the errors introduced by using a linear impact function may be smaller than the errors introduced by other uncertainties in the analysis, and hence the linear approximation may be acceptable. Alternatively, the impact function can be extended to include quadratic terms, as follows. Revisiting Equation 1, if we retain both the linear and quadratic terms, we obtain

$$a = a_0 + x^T r + \frac{1}{2} x^T H\, x \tag{7}$$

Once again, this approximation only has to be a good approximation across the range of variability in the single ensemble being considered, and not across the entire range of variability of x and y.

In this expression the impact of the full climate state y is made up of three parts: the impact due to the ensemble mean a_0, the linear impact $x^T r$, and the quadratic impact $\frac{1}{2}x^T H\, x$. The quadratic impact can then itself be split into two parts: the local impact relating to the diagonal elements of H and the non-local impact relating to the off-diagonal elements of H, which contains quadratic cross-terms which model the impact as depending on the product of climate variables at different locations. Writing $H = H^{\text{diag}} + H^{\text{offdiag}}$, Equation 7 becomes

$$a = a_0 + x^T r + \frac{1}{2}x^T H^{\text{diag}}\, x + \frac{1}{2}x^T H^{\text{offdiag}}\, x \tag{8}$$

We can once again consider finding the vector x that maximizes probability density for a given impact, leading to the following Lagrangian:

$$L = -x^T C^{-1}\, x + \lambda\left(a_0 + x^T r + \frac{1}{2}x^T H x - a_c\right) \tag{9}$$

Which has the solution

$$x \propto (2I - \lambda C H)^{-1}\, Cr = A\, Cr \tag{10}$$

Where we have written $A = (2I - \lambda C H)^{-1}$. We will refer to this as a quadratic DCA pattern. Equation 10 is an expression that could be readily evaluated given an ensemble, the linear impact vector r and the quadratic impact matrix H. The pattern x is no longer a weighted sum of the ensemble members, but is the weighted sum $S = Xv$ transformed by the matrix A:

$$x \propto ACr \propto AXX^T r = AX\left(X^T r\right) = AXv = AS \tag{11}$$

The pattern x is also no longer constant in direction as a function of the Lagrange multiplier, as it is in the linear case. In Equation 10, different values of the Lagrange multiplier λ lead to solutions with different values for the probability density, lengths of the vector and directions.

We can simplify Equation 10 in two stages. First, we might imagine a situation in which the quadratic impact is just the weighted sum of the impacts at each location, with no cross-terms, that is, $H^{\text{offdiag}} = 0$. Under this first simplification the matrix H is then diagonal, $H = H^{\text{diag}}$, and represents the quadratic curvature of the impact at each location. This could be appropriate for wind damage as a function of wind speed, since wind damage is generally local. It would be less appropriate for flood damage as a function of precipitation, since hydrological processes may make the relationship between precipitation and flood damage non-local. Second, we might imagine a situation in which there is no quadratic impact at all, in which case the matrix H is zero, and the solution reverts to the linear DCA solution.

A similar, although slightly more complicated, analytical solution also exists for a cubic approximation to the impact, based on the analytical method for solving vector quadratic equations. If a cubic approximation is not sufficient then it may be necessary to run the full impact model with every member of the ensemble separately.

3.3. Interpretation of SMEs and MMEs

In Section 6 we apply representative deviation analysis to an SME and an MME. SMEs and MMEs have different statistical properties, and in some situations the resulting representative deviation patterns need to be interpreted differently. We now discuss the issues related to interpretations of SMEs and MMEs.

For both SMEs and MMEs, quantities calculated using the ensemble, such as means, variances and covariances, can be used to give insight into the structure of the ensemble. Someone who wishes to understand as much as possible about the structure of the ensemble they are using may appreciate the information provided by these quantities. Also, for both SMEs and MMEs, parameters calculated from the ensemble can be used to fit distributions to the ensemble, which could then be used as subjective predictive distributions. This distribution fitting would typically come after appropriate post-processing, including bias correction and weighting schemes. This is discussed in more detail below.

The main distinguishing property of SMEs, relative to MMEs, is that for SMEs it is clear how one could generate more members of the ensemble, ad infinitum, and one could imagine creating an infinite population of climate model runs in this way. There is therefore a clearly defined infinite population that we can consider the individual model runs in an SME as being sampled from. As a result, it is easy to justify applying statistical results from frequentist statistics that relate to infinite populations, and that relate to the sampling properties of finite samples from those populations. Many of the quantities estimated based on an SME ensemble, such as means, variances and covariances, are therefore not just useful summaries of the structure of the ensemble, and not just a way to fit predictive distributions, but can also be considered as estimates of the corresponding population quantities, and have well understood sampling properties. As the size of the ensemble increases, the sample estimates would converge toward their population analogs. These statistical properties of SMEs are advantageous for certain types of analysis. The main disadvantage of SMEs is that they do not sample model uncertainty in a completely satisfactory way, since the development of any single climate model involves somewhat arbitrary decisions about how to implement the parametrizations. Attempts have been made to overcome this problem using stochastic parameters within the parametrizations of the models used in SMEs (and stochastic parametrizations are used in the UKCP18 ensemble that we analyze below) but this approach still works within the framework of a single specific implementation, and cannot be expected to account for all the model uncertainty related to possible alternative implementations. If, for instance, the convection parametrization in the model used to create an SME is significantly unrealistic, then even with stochastic parameters no members of the ensemble will be realistic, and the ensemble may not encompass reality. As a result of model uncertainty, even though the relationship between ensemble quantities in an SME and the corresponding population quantities is well understood, the relationship between ensemble quantities and the actual future outcome is in general not well understood. It is the relationship with the actual future outcome that is critical for understanding whether the ensemble is likely to be useful as a basis for making good predictions of the future. Climate modelers may nevertheless make the pragmatic choice to use quantities derived from an SME, perhaps adjusted via bias correction and weighting, to create predictive distributions. Given the number of subjective decisions involved in making such distributions, such predictions are best interpreted as representing the modelers' subjective assessment of the probabilities of future events, according to the subjective interpretation of probability as used, for instance, in subjective Bayesian statistics. Since such subjective assessments are the only way to make probabilistic predictions of future climate, and probabilistic predictions are needed to make decisions, these subjective predictions are extremely important for quantitative applications of climate information.

Relative to SMEs, multi-model ensembles (MMEs) have the advantage that they include different models in which parameterizations have been implemented differently, and so hopefully give a better representation of model uncertainty. Continuing with the convection parametrization example: it would be hoped that at least one of the models in the ensemble has a reasonably realistic convective parametrization. The disadvantage of MMEs is that they do not fit as easily into the population-and-sample framework of frequentist statistics. One could perhaps consider the models in an MME as samples from the distribution of all possible models, but the meaning of "distribution of all possible models" is rather hard to make concrete. As a result it is more difficult to define what the population is, and in this article we will not consider MMEs as providing samples from a population of models. With this interpretation of MMEs, quantities such as means, variances, and covariances can still be calculated, and may again be useful summaries of the structure of the ensemble, and as a way to fit predictive distributions. However, they should not be considered as estimates of population quantities, as they would be for SMEs, because of the lack of a clearly defined population. As with SMEs, the relationship between these quantities and the actual outcome is not in general well understood, and once again it is this relationship that is the critical factor that determines whether the ensemble is likely to make good predictions of the future. As with SMEs, climate modelers may make the pragmatic choice to use quantities calculated from an MME, most likely adjusted using weighting and bias correction, to create a predictive distribution, representing their subjective assessment of the probabilities of future events.

With our two example ensembles, it is impossible to know which is likely to give the better predictions. As an SME, the UKCP18 ensemble may underestimate model uncertainties, while the EURO-CORDEX ensemble, as an MME, may sample model uncertainty better but it is hard to know how to weight the different members of the ensemble (for simplicity, we have chosen equal weighting). We therefore consider it useful to consider both, as recommend by Tucker et al. (2021) in their analysis of the UKCP18 ensemble.

The implications of this discussion for our representative deviation methods are as follows. For both ensembles, all the representative deviation methods we calculate can be considered as potentially giving useful insights into the structure of the ensembles. If the ensembles are used to create predictive probabilities, using multivariate normal or elliptic distributions, then in addition DCA has a number of optimality properties as described above, which apply in the context of the fitted distribution. For the W2, W4, and DCA methods derived from an SME, one can go further and define corresponding population quantities, and if the ensemble size were increased, the sample estimates would gradually converge onto these population quantities (assuming that as the ensemble size increases W2 and W4 are defined as the means of the ensemble members with the largest 20% and 40% of impact, respectively, rather than just as the top 2 and 4 members). W1, defined as the member with the largest impact, would not, however, converge as the ensemble size increases.

4. Illustrations and Examples

We now show a number of illustrations and examples of DCA applied to ensembles, extending the illustrations and examples in Jewson (2020) and SJM. This section is intended to help develop an intuitive understanding of the various methods we are considering, while Section 5 explores some of the mathematical properties of the methods. We consider a linear impact function. Linear impact functions may arise because the impact is genuinely linear, or they may arise as approximations to non-linear impact functions. In Section 4.1 we show conceptual illustrations that relate directly to the definitions of DCA given in Section 3.2. The definitions are based on probability density and impact, and the conceptual illustrations show how DCA arises from the interaction of contours of constant probability density and impact. In Section 4.2, we give more concrete illustrations in which we calculate DCA patterns using simulated data. These show how the shape of a cloud of simulated data points interacts with the gradient of impact to create the DCA pattern. In Section 4.3, we then show some comparisons between W1, W2, W4, and DCA. We will assume the ensemble is MVN.

4.1. Illustrations

Figure 1 illustrates DCA in four different situations, in two dimensions. We will assume that the two dimensions represent changes in time mean precipitation at two locations. In each case the ensemble is represented by a single elliptical contour of constant probability density, with higher densities inside the ellipse. The ensemble mean is shown by the red dot at the center of the ellipse. Figure 1a illustrates a situation in which the impact is simply the sum of the precipitation changes at the two locations, and so the impact vector $\boldsymbol{r}$ (illustrated by the black arrow) points toward the top right corner of the diagram where the total precipitation change is greatest. A line of constant impact, which is perpendicular to the impact vector $\boldsymbol{r}$, is shown by the straight diagonal line. The DCA pattern is shown by the blue dot at the point at which the straight line is tangent to the ellipse. As discussed above in Section 3.2.2 the DCA pattern can be scaled in different ways according to the application, and in this case the DCA pattern is scaled to lie on the contour of probability density shown. The DCA pattern does not align with the impact vector because it also takes account of the covariance structure of the ensemble, that is, the elliptical shape of the distribution. This diagram relates to the two properties of DCA given in Section 3.2.1 above as follows. Considering property 1: we can see that of all points on the straight line, which have the same impact (in this case, the same total precipitation change), the point that touches the ellipse is the point with the highest probability density (since all the other points on the straight line lie outside the ellipse and are hence at lower probability densities). Considering property 2: we can see that of all the points on the ellipse, which are at the same probability density, the point at which the straight line touches the ellipse is the point with highest impact (since it is the point which is furthest from the ensemble mean in the direction of the impact vector). Furthermore, we can see that if we move away from the DCA pattern in any direction then either the impact or the probability density will reduce: moving into the ellipse or along the ellipse always reduces impact, while moving out of the ellipse always reduces probability density. A proof of this property that as we move away from the DCA pattern in any direction then either the impact or the probability density will reduce is given in Section 5.1.2 below.

Figure 1b illustrates the same ensemble (i.e., the same ellipse) as Figure 1a but a situation in which the definition of impact puts twice as much weight on precipitation change at the second location (the vertical axis) as on the precipitation change at the first location (the horizontal axis). As a result, $\boldsymbol{r}$ now points more vertically than in Figure 1a, and the line of constant impact is more horizontal. This changes the DCA vector, which now has more precipitation at location 2 than location 1. Figure 1c illustrates a different ensemble, but for the same impact

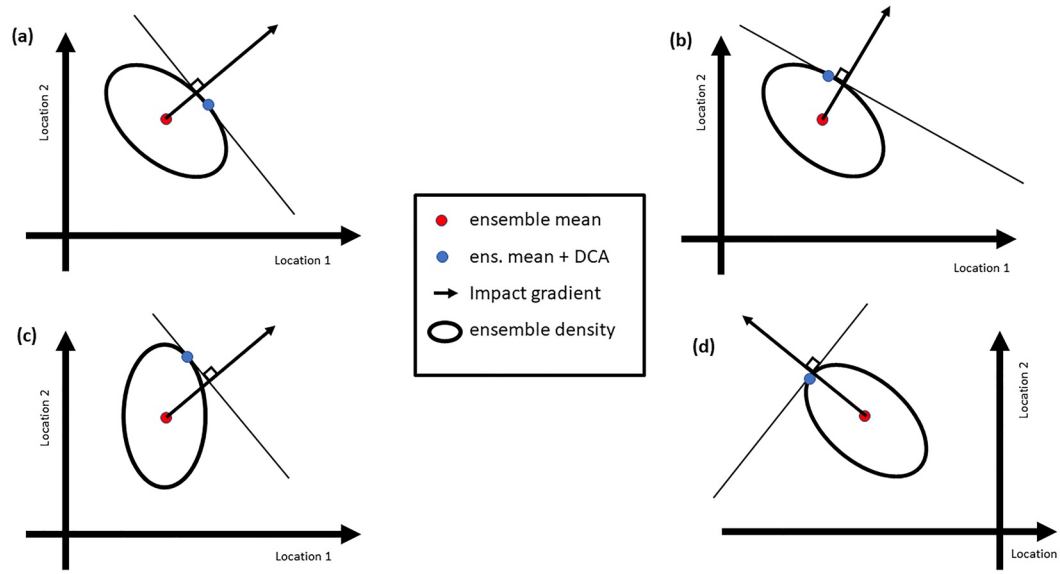

Figure 1. Four illustrations of Directional Component Analysis (DCA), applied to climate variables at locations 1 and 2. The ellipse in each case illustrates a contour of probability density for an ensemble prediction for these two locations. The ensemble is assumed to be bivariate normally distributed. The red dot in each case illustrates the ensemble mean, the arrow indicates the gradient of increasing impact, the straight line, which is perpendicular to the arrow, indicates a line of constant impact and the blue dot indicates the resulting DCA pattern. Examples (a and b) show the same ensemble, but different directions of impact, and hence different DCA patterns. Examples (a and c) show different ensembles, but the same direction of impact, and hence different DCA patterns. Example (d) shows a different ensemble and a new direction of impact.

vector as in Figure 1a. This also changes the DCA vector relative to Figure 1a. Finally, Figure 1d illustrates a more complex situation using a different ensemble and a situation in which the impact increases as precipitation changes increase at location 2 but impact *decreases* as precipitation changes increase at location 1. This might occur in the situation discussed in the introduction in which at location 2 the main concern with respect to climate change and future precipitation is related to increases in precipitation (e.g., such as concerns about increased flooding in the UK) while at location 1 the main concern is with respect to decreases in precipitation (e.g., such as concerns about increased drought in some parts of Southern Europe). In this case the vector r points to the top left corner of the diagram, and the lines of constant impact are correspondingly different, and run from the lower left to the upper right. The DCA direction is also correspondingly very different, and selects a pattern that consists of increased precipitation at location 1 and decreased precipitation at location 2, relative to the ensemble mean, as the representative deviation.

4.2. DCA Examples

We now give some examples of DCA patterns calculated using synthetic ensemble data (Section 2.1). Our first set of examples is based on synthetic ensembles with 1,000 members and is shown in Figure 2. These examples correspond to the diagrams shown in Figure 1. Figure 2a shows one ensemble with 1,000 ensemble members, generated using means of 3 (location 1) and 4 (location 2), variances of 1 (location 1) and 0.84 (location 2), and a correlation of −0.4. The figure shows the ensemble mean (red dot) and the direction of the impact vector r (black arrow), which puts equal weight on both precipitation variables (following Figure 1a). It also shows the DCA pattern scaled to 1.5 standard deviations of impact, as described in Section 3.2.2 (blue dot). The DCA pattern is calculated using Equation 5. Figure 2b shows the same ensemble, with the same ensemble members, but now with a different impact vector r which puts twice as much weight on the precipitation change on the vertical axis (following Figure 1b). This shifts the DCA pattern. Figure 2c shows a different ensemble, generated using means of 3 and 4, variances of 0.25 and 1, and a correlation of 0. However, it uses the same impact vector as Figure 2a (following Figure 1c). This leads to a different DCA pattern relative to Figure 2a. Figure 2d shows a different ensemble again, generated using means of −4 and 3, variances of 1 and 0.84, and a correlation of −0.4, and an impact vector which puts equal weight on negative precipitation change on the horizontal axis and positive

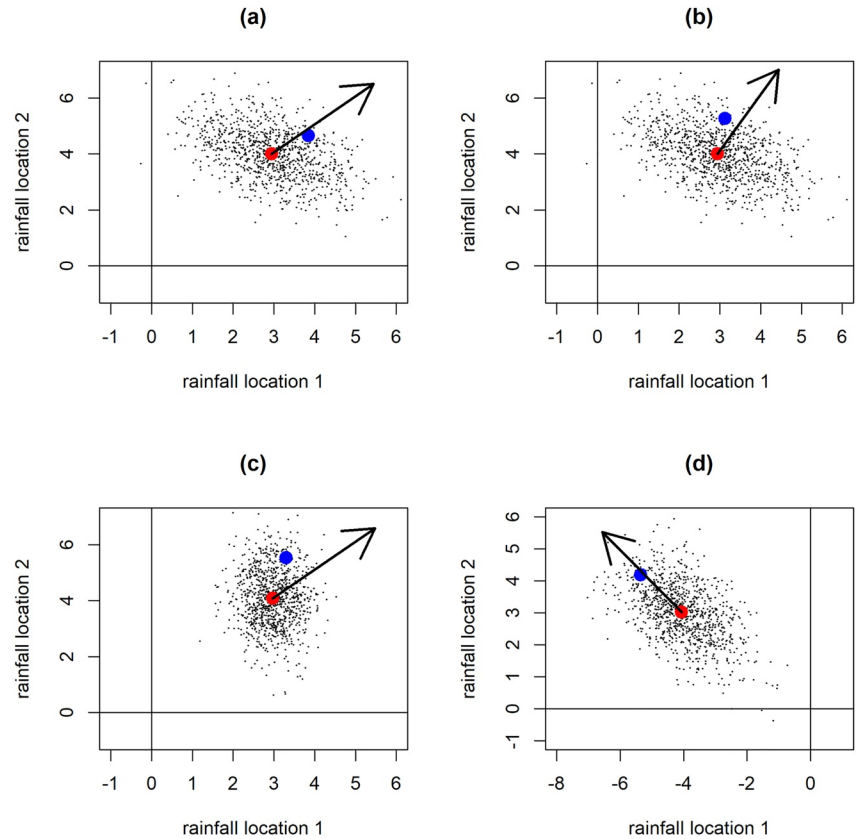

Figure 2. Four randomly generated 1,000 member examples and related quantities. In each case the ensemble members are shown by black dots, the ensemble mean is shown by a red dot, the gradient of increasing impact is shown by the arrow and the ensemble mean plus the Directional Component Analysis pattern, scaled to 1.5 standard deviations of impact, is shown by a blue dot. Examples (a and b) use the same ensemble, while examples (a and c) use the same impact vector. These examples correspond to the illustrations in Figure 1.

precipitation change on the vertical axis (following Figure 1d). This leads to a DCA pattern which also has negative precipitation change on the horizontal axis and positive precipitation change on the vertical axis.

Overall, the above examples illustrate the way in which the DCA pattern is determined by both the impact function and the shape of the probability density function of the ensemble.

4.3. W1, W2, W4, and DCA Examples

We now give some examples of applying W1, W2, W4, and DCA to simulated data. These examples are based on ensembles with just 10 members, to help us understand the behavior of the representative deviation methods on small ensembles, of the same size as the UKCP18 and EURO-CORDEX ensembles we analyze in Section 6. Figures 3a and 3b show the results of the four different methods applied to two simulated ensembles. These ensembles were generated using the same parameters as the ensembles in Figures 2a and 2b, but with only 10 members. The elliptical structure of the ensemble is now much less clear from so few members.

Figures 3c–3f explore the robustness of the four methods for estimating a representative deviation for ensembles of size 10. They show results from 500 ensembles, each of size 10, from the same underlying mean, variances, and correlation as were used in Figures 3a and 3b. For each of the 500 ensembles, we calculate W1, W2, W4, and DCA, and these are illustrated in Figures 3c–3f. The mean of the 500 ensemble means is given by the red dots, and the individual ensemble members are not shown. The 500 values of W1 show a large spread, even though the parameters of the underlying ensemble generation process are the same for each of the 500 values. This is because W1 relies on just a single ensemble member. The spread reduces as we move from W1–W2 to W4–the DCA pattern. This shows that the DCA pattern is more precisely estimated than any of the other methods. This

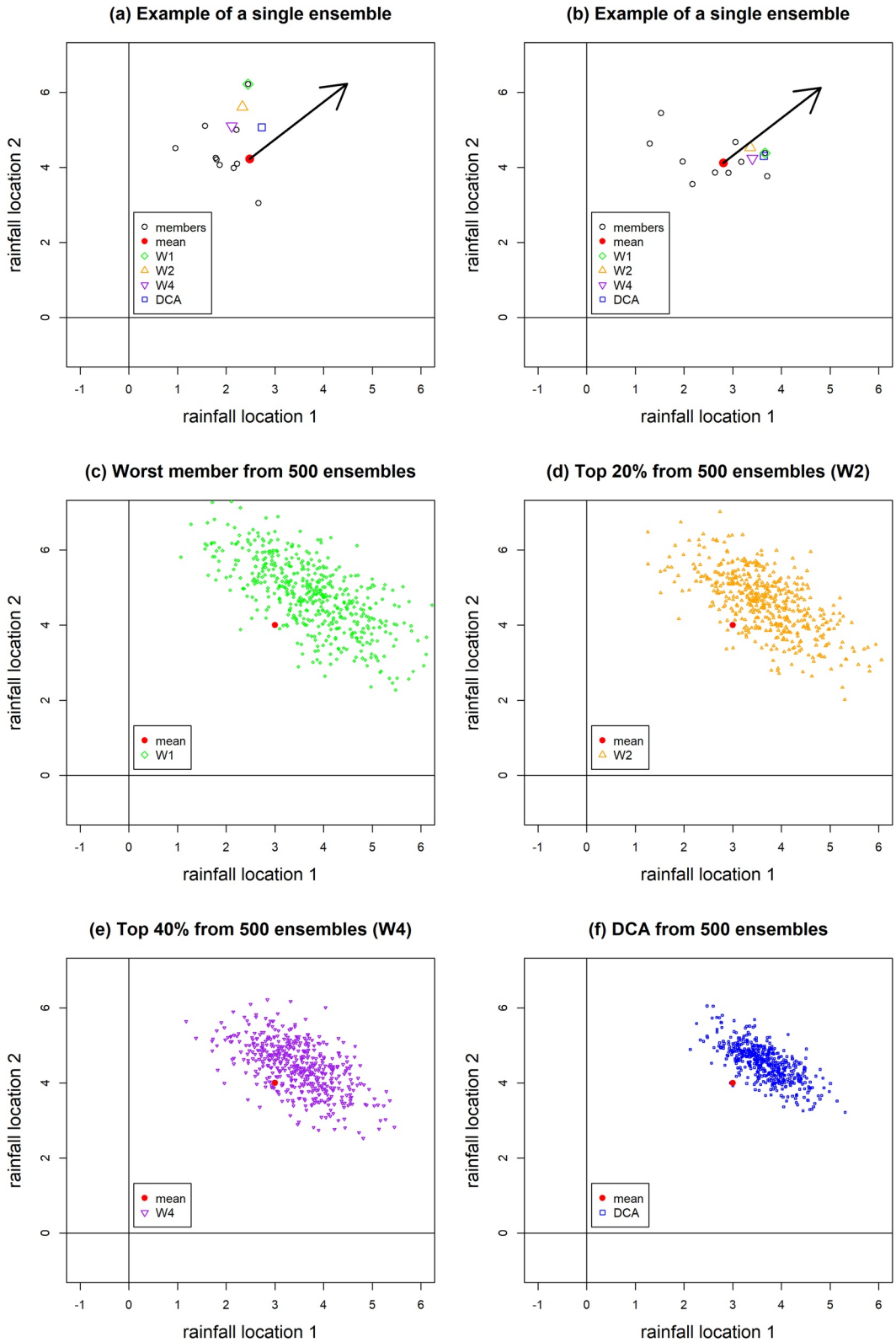

Figure 3.

is because the DCA pattern is based on more information: from Equations 5 and 6 we see that it uses information from the entire ensemble, rather than from just a subset of the ensemble.

SJM studied the robustness of DCA in detail, comparing with the robustness of the most extreme ensemble member and the average of the five most extreme ensemble members, for a 50 member ensemble of ECMWF medium range temperature forecasts. They considered statistical robustness, that is, robustness to the random sampling that occurs in the creation of the ensemble, robustness to small changes in the physical domain, and robustness to the use of normal distributions. In all tests they found that DCA was more robust than the two alternative methods they considered.

The situation in which using more data leads to less variable estimates is common in statistics. The downside of using more data is that it may smooth over important features. In our case, for an ensemble that has a distribution very far from an elliptical distribution, it may be that only W1 gives a pattern which is a plausible sample from the ensemble. For ensembles with more regular distributions, W2 and W4 may give patterns that are plausible samples from the ensemble, but DCA, which smooths the data the most, may not. For ensembles that are close to elliptically distributed, all the methods will give patterns which are plausible samples from the ensemble. In practice, for high dimensional ensembles with small ensemble sizes like the climate model ensembles we use in Section 6, it is very difficult to ascertain mathematically which methods will give plausible samples, and some judgment is required. We will see in Section 6 that comparing the patterns from the four methods can be useful in this respect.

5. Additional Mathematical Properties of DCA

In Section 5.1 below we present a number of additional mathematical properties of DCA for the MVN case, that help with the interpretation. Those readers less interested in the mathematical properties may wish to skip to Section 6. In this section we show that: (a) DCA can be defined simply as a pattern that maximizes two easy to understand metrics, avoiding the Lagrangian framework used in the original definitions; (b) DCA can also be defined simply as a tail-conditional expectation, which demonstrates that there is a mathematical connection between DCA, W2, and W4, and also shows that DCA can be considered as a generalization of the ensemble mean; and (c) DCA can also be defined as a weighted tail-conditional expectation, which connects DCA with both unweighted and weighted composites.

The idea that one pattern has multiple mathematical properties is a familiar one, since the ensemble mean itself has multiple mathematical properties, each of which helps with the interpretation of the ensemble mean in different ways and in different situations (for example, the ensemble mean is the pattern that minimizes the in-sample mean squared error across the ensemble, and if the ensemble is symmetrically distributed then the ensemble mean is also the ensemble median). For the derivations below we will follow the standard approach of assuming that the covariance of the MVN is known. We then define DCA using Equation 5, evaluated using this known covariance matrix. DCA patterns calculated using samples from this MVN distribution could be considered as sample estimates of the population DCA pattern, following the usual framework of frequentist statistics. We will illustrate these additional properties in Figure 4 with a simple two-dimensional example, and then in Section 5.2 we describe the calculation of a quadratic DCA pattern, which we also illustrate, for comparison with the results related to linear DCA.

5.1. Additional Mathematical Properties of DCA

The optimality properties of DCA discussed in Section 3.2.1 above (property 1 and property 2) are of the form "the DCA vector maximizes X given Y." In Sections 5.1.1 and 5.1.2 below we will show that DCA also satisfies two statements of the simpler form "the DCA vector maximizes Z." Furthermore, Equation 6 above shows that

Figure 3. Panels (a and b) show ensembles generated using the same parameters as Figure 2a, but now for just 10 members in each ensemble. They also show representative deviations, calculated using W1 (most extreme member, green diamond), W2 (average of the two most extreme members, orange triangle), W4 (average of the four most extreme members, purple inverted triangle), and Directional Component Analysis (DCA) (blue square). Panels (c–f) show representative deviations generated from 500 ensembles created using the same parameters as the ensembles shown in panels (a and b). The red dots show the mean of the ensemble means. Panel (a) shows 500 examples of W1, one for each ensemble. Panel (b) shows 500 examples of W2. Panel (c) shows 500 examples of W4, and panel (d) shows 500 DCA patterns.

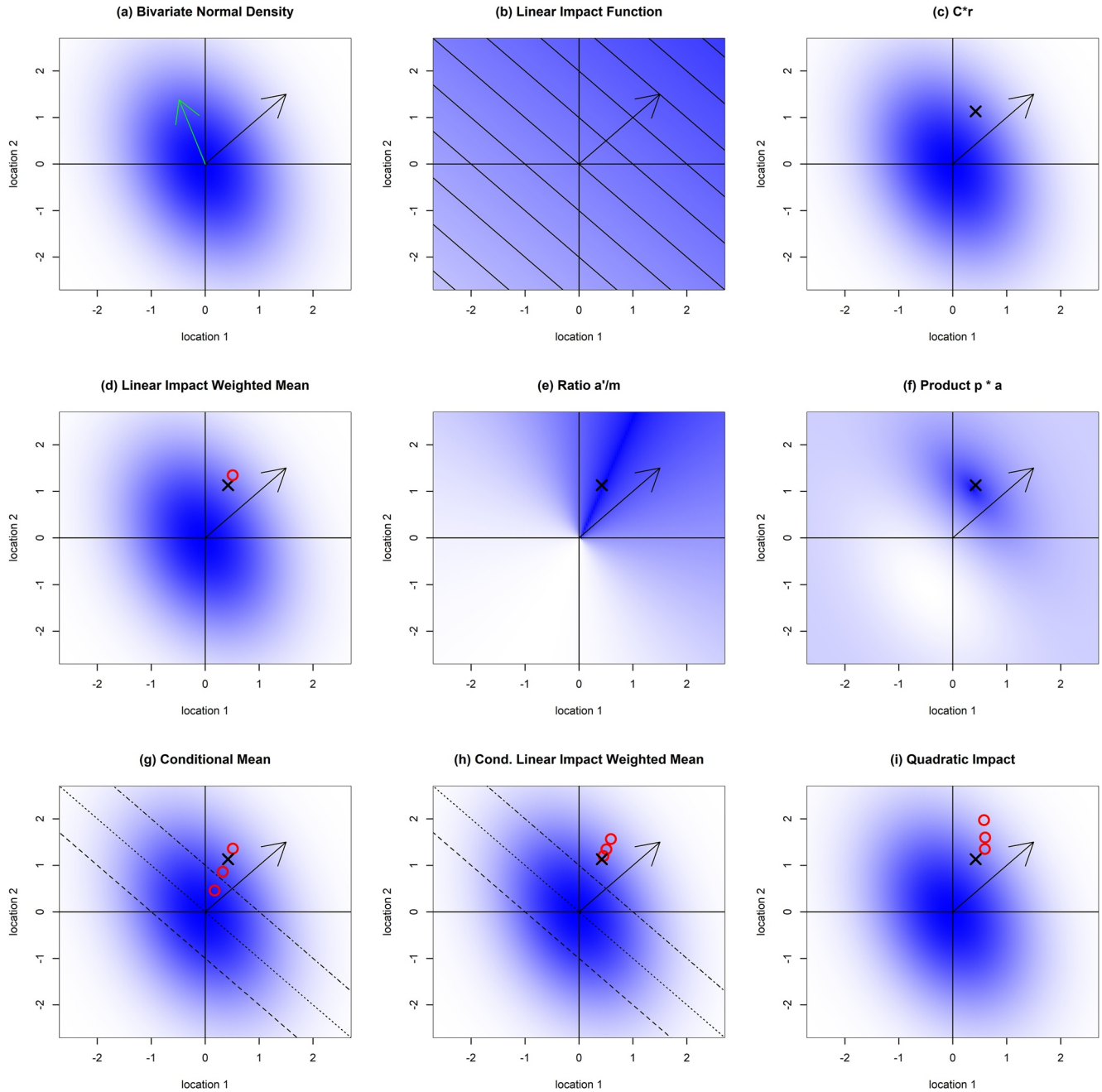

Figure 4. Illustrations of mathematical properties of Directional Component Analysis (DCA). In each panel, the gradient of increasing impact is shown by the black arrow that points toward the top right hand corner. Panel (a) shows probability densities p for a bivariate normal, with darker shading indicating higher densities, and lower values of the Mahalanobis distance m from the origin. The density is elliptical, and the principal axis of the ellipse is given by the green arrow. Panel (b) shows values of a linear impact function a increasing toward the top right-hand corner, with darker shading indicating higher impact. Panel (c) shows the DCA pattern as a black cross, calculated using Equation 5 with r normalized to unit length. This black cross is repeated in all subsequent panels. Panel (d) shows the DCA pattern as a red circle, calculated using Equation 6. Panel (e) shows values of the ratio a'/m, with darker shading indicating higher values. Panel (f) shows values of the product pa', with darker shading showing higher values. Panel (g) shows three examples of the conditional mean pattern, conditional on exceeding a certain level of impact, as red circles. The impact thresholds are given by the three diagonal lines. Panel (h) shows three examples of the impact-weighted conditional mean pattern as red circles, conditional on exceeding a certain level of impact. The impact thresholds are given by the diagonal lines. Panel (i) shows three examples of quadratic DCA patterns, as red circles.

the DCA pattern can be considered as the linear-impact-weighted average of the members of the ensemble, which suggests that DCA may have properties related to the expectation over possible patterns in the MVN. In Section 5.1.3 below we explore this idea and present two ways in which the DCA pattern can be considered as an expectation. This allows us to make theoretical connections between DCA, W2, and W4.

The shading in Figure 4a shows values of the probability density p for a bivariate normal distribution with mean zero, variances of 1 (location 1) and 2 (location 2), and a correlation of −0.4, representing the probability densities for an ensemble. Darker shading indicates higher probability density, and the black arrow (in this and subsequent panels in Figure 4) shows the direction of the impact vector r, which is in the direction (1, 1). The probability density has an elliptical shape, and the green arrow shows the principle axis of the ellipse, which is also the first eigenvector of the covariance matrix. The green arrow is scaled to a length equal to the square root of the corresponding eigenvalue.

The shading in Figure 4a can also be interpreted as showing a quantity known as the Mahalanobis distance, defined by $m = \sqrt{x^T C^{-1} x}$. Mahalanobis distance is a multivariate generalization of the univariate concept of the number of standard deviations from the mean. For the MVN the Mahalanobis distance is proportional to minus the log of the probability density, and so (since numeric values are not specified for the color scale) the shading in Figure 4a can also be interpreted as showing the Mahalanobis distance, with *lighter* shading showing a greater distance. The shading in Figure 4b shows values of the linear impact function, which is $a' = 30(x_1 + x_2)$, where x_1 and x_2 are the distances in the horizontal and vertical directions. The diagonal lines show constant values of this linear impact function. Figure 4c, and all subsequent panels of Figure 4, show the DCA pattern, given by a black cross, calculated by scaling r to be a unit vector and then using Equation 5. Figure 4d shows the DCA pattern, given by a red circle, calculated using Equation 6, without further scaling. The red circle in Figure 4d can be seen to represent the same spatial pattern as shown by the black cross (i.e., has the same direction from the origin) but with a slightly different length scaling.

5.1.1. The DCA Pattern Maximizes the Ratio of Linear Impact to Mahalanobis Distance

The first new property of DCA that we describe (*property* 3) is that the linear DCA vector maximizes the ratio of the linear impact a' to the Mahalanobis distance. We write this ratio as

$$c(x) = \frac{a'(x)}{m(x)} \tag{12}$$

This property of DCA emphasizes that in the MVN case DCA finds patterns that have both a high linear impact (a large value of a') and at the same time have a high probability density, which corresponds to a low value of m. The proof that maximizing Equation 12 leads to the same expression for the DCA pattern as Equation 5 above is given in Appendix B.

Figure 4e shows values of the ratio $c = a'/m$, with darker shading showing larger values. We can see that there is a direction along which this ratio is maximized. This direction aligns with the direction of the DCA pattern given by the black cross.

5.1.2. The DCA Pattern Maximizes the Product of Linear Impact and Weighted Probability Density

The second new property of DCA that we describe (*property* 4) is that any linear DCA pattern, of any length scaling, can be described as the pattern that maximizes the product of the linear impact a' and the probability density p to some positive power, that is, the function

$$f(x) = p(x)^n\, a'(x) \tag{13}$$

In Appendix C we show that maximizing Equation 13 leads to a pattern which is a scaling of the DCA pattern given by Equation 5 and that different values of n give all possible scalings. Choosing $n = 1$, this means that the DCA direction can be described as the direction that maximizes the product of probability density and impact, $p(x)a'(x)$, and this is the simplest and most easy to understand definition of DCA. One could also attempt to find patterns that maximize this function for probability distributions other than the MVN, although the solutions may not be DCA patterns, according to the definition of DCA given in Section 3.2.1 above.

An implication of this property is that when $n = 1$ there is no other pattern which has both a higher probability density and a higher linear impact, and that any change to the DCA pattern would lead to either the probability density or the impact reducing. This is also clear from the examples illustrated in Figure 1.

Figure 4f shows values of $f(x)$ for $n = 1$, with darker shading showing higher values. We see that there is indeed a maximum in the direction of the DCA vector.

5.1.3. The DCA Vector Is Parallel to the Tail Conditional Expectation

The third new property of DCA that we describe (*property* 5) is that, for the MVN, the direction of the linear DCA vector is the same as the direction of the vector defined as the expectation over all possible vectors, conditional on exceeding a certain value for the linear impact. We call this the tail conditional expectation. In Figure 4b, this expectation is the expectation over the region above and to the right of one of the diagonal lines. This property is shown in Appendix D below.

Given this property one could estimate DCA as the ensemble mean, conditional on exceeding a certain threshold value for the linear impact. This estimator would use less information from the ensemble than Equation 5 or Equation 6. Whether or not it would give a more representative pattern depends on the structure of the ensemble, as discussed above in Section 4.3.

Figure 4g shows patterns estimated in this way, for three different levels of linear impact, given by the three diagonal lines, which have linear impacts of −30, 0, and 30. Linear impact values give the impact relative to the impact of the ensemble mean, and so a negative value means less impact than the ensemble mean, but not necessarily a negative total impact. We see that the patterns do indeed lie along the DCA direction, given by the black cross.

This property helps explain why SJM found that using the mean of the five most extreme members of the ensemble as a definition of reasonable worst case gave similar results to DCA: the mean of the five most extreme members is in fact an alternative (less precise) estimator for DCA. It also shows that W2 and W4 can be considered as estimates of the DCA pattern in the MVN case.

There is an extension of this property to *weighted* tail conditional expectations (*property* 6). This is illustrated in Figure 4h and shown in Appendix E. A special case of the weighted property is that the expectation over *all* patterns, weighted by the linear impact, is parallel to DCA (*property* 7). This is the expectation version of Equation 6.

5.1.4. Summary of Properties

Given all the above, we can summarize the key properties of DCA, as applied to MVN distributed data, as follows:

1. When scaled appropriately, the DCA pattern gives the unique spatial pattern that maximizes the probability density, for a given level of linear impact, for any level of linear impact (Section 3.2.1, property 1).
2. When scaled appropriately, the DCA pattern gives the unique spatial pattern that maximizes the linear impact, for a given level of probability density, for any level of probability density (Section 3.2.1, property 2).
3. All scalings of the DCA pattern maximize the ratio of linear impact to Mahalanobis distance. All scalings give the same value for this ratio (Section 5.1.1, property 3).
4. For any given scaling, the DCA pattern is the unique spatial pattern that maximizes the linear impact multiplied by a positive power of the probability density (Section 5.1.2, property 4).
5. When scaled appropriately, the DCA pattern is the unique spatial pattern that equals the expectation of all possible patterns, conditional on exceeding a certain linear impact (Section 5.1.3, property 5).
6. When scaled appropriately, the DCA pattern is the unique spatial pattern that equals the *weighted* expectation of all possible patterns, conditional on exceeding a certain linear impact, and weighted by the linear impact (Section 5.1.3, property 6).
7. When scaled appropriately, the DCA pattern equals the linear impact-weighted expectation of *all* possible patterns (Section 5.1.3, property 7).

5.2. Quadratic Impact Example

For comparison, we also illustrate a quadratic DCA pattern. Figure 4i shows patterns generated from Equation 10, using a quadratic impact function. These patterns correspond to the same linear impact vector as used in the previous panels, along with a quadratic impact matrix H given by the identity matrix, with values of λ of −2, 0, and 2. We see that the patterns for different values of λ are no longer parallel, as expected.

6. Climate Model Results

We now show representative deviation patterns derived from the UKCP18 and EURO-CORDEX ensembles described in Sections 2.2 and 2.3, using the W1, W2, W4, and DCA methods. We use two different ensembles for the same variable, RCP, and time period so that we can assess the robustness of the results to the choice of ensemble. Figures 5 and 6 show results based on annual precipitation, and a linear impact metric defined as total precipitation anomaly. Figure 7 shows results based on annual precipitation, and a more complex impact function. Figure 8 shows results based on seasonal precipitation.

The UKCP18 ensemble mean shows increasing precipitation in the north of the domain, decreasing precipitation in the south of the domain, and a mixture of increases and decreases in between. The ensemble standard deviation shows spatial variations, but with less obvious large-scale spatial structure than the ensemble mean. Of the four candidate patterns for representative deviation, the one that is least similar to the other three is the W1 pattern. The two patterns that are the most similar are the W4 and DCA pattern. That these patterns are similar is an indication of robustness for both, and is an indication that neither is too influenced by the potential weaknesses of the methods used to derive them. It suggests that the W4 pattern is not strongly affected by sampling variability, and that the DCA pattern is not strongly affected by possible non-elliptical behavior in the ensemble distribution. Either pattern would therefore be a reasonable candidate for a representative deviation from the ensemble mean. Both patterns show negative precipitation anomalies in Norway, and positive precipitation anomalies in much of the rest of Europe, except Spain. The interpretation based on the properties of DCA, subject to the assumption that the ensemble is approximately elliptically distributed, is that this is the most likely pattern to cause increases in the total precipitation relative to the ensemble mean. One might imagine that the most likely pattern to cause increases in the total precipitation would consist of more precipitation everywhere, but this analysis shows that is not the case. A pattern consisting of more precipitation everywhere would contain more total precipitation than the pattern given by DCA but would be less likely to occur. The spatial structures in the W4 and DCA patterns arise because of spatial correlations of patterns in the ensemble, and show that there is a consistent tendency across the model runs in the ensemble for lower than ensemble mean precipitation in Norway to be associated with greater than ensemble mean precipitation further south. One might speculate that this is caused by a southward shift in the storm track in some of the model runs, moving precipitation from Norway further south. If we were to reverse the sign of the impact function, then the DCA pattern would reverse. In other words, the most likely pattern to cause *decreases* in the total precipitation amount would involve positive precipitation in Norway and negative precipitation anomalies further south. The W4 pattern for the reversed sign impact function may not be the exact reverse of the W4 pattern shown, but would likely be similar, since the sum of all the ensemble members is zero by definition.

Figure 6 shows corresponding results, but for the EURO-CORDEX ensemble. Strikingly, the ensemble mean is now very different, with most of Europe showing increases in precipitation, and only the South showing slight decreases. The differences between the two ensemble means highlights the issue of model uncertainty, as discussed in Section 2.7, and are also likely to be due to the small size of the two ensembles. Understanding these differences in more detail, for instance in terms of which times of year are driving the differences, to what extent the ensembles overlap and whether it would make sense to combine the two ensembles into a super-ensemble in some way, would be very informative, but is beyond the scope of the present study. The standard deviation pattern of the EURO-CORDEX ensemble is similar to that for UKCP18.

Considering the four candidates for representative deviation derived from EURO-CORDEX: once again the most unique pattern of the four is the W1 pattern, and the two most similar patterns are the W4 and DCA patterns. The W1 patterns from UKCP18 and EURO-CORDEX are not similar. This is to be expected, because the most extreme pattern in any ensemble is always highly random. The W4 and DCA patterns from UKCP18 and EURO-CORDEX are, however, somewhat similar, in that all four patterns show the north-south dipole with decreases in precipitation in the north and increases further south. That the representative deviation pattern for precipitation consists of this North-South dipole therefore seems to be a reasonably robust result: robust to using either W4 or DCA, and robust to using a different ensemble.

Figure 7 shows representative deviation patterns calculated using W4 and DCA, with the same data as used for Figures 5 and 6, but now using a different impact function, where positive impact arises from positive anomalies where the ensemble mean is positive, but from negative anomalies where the ensemble mean is negative. This

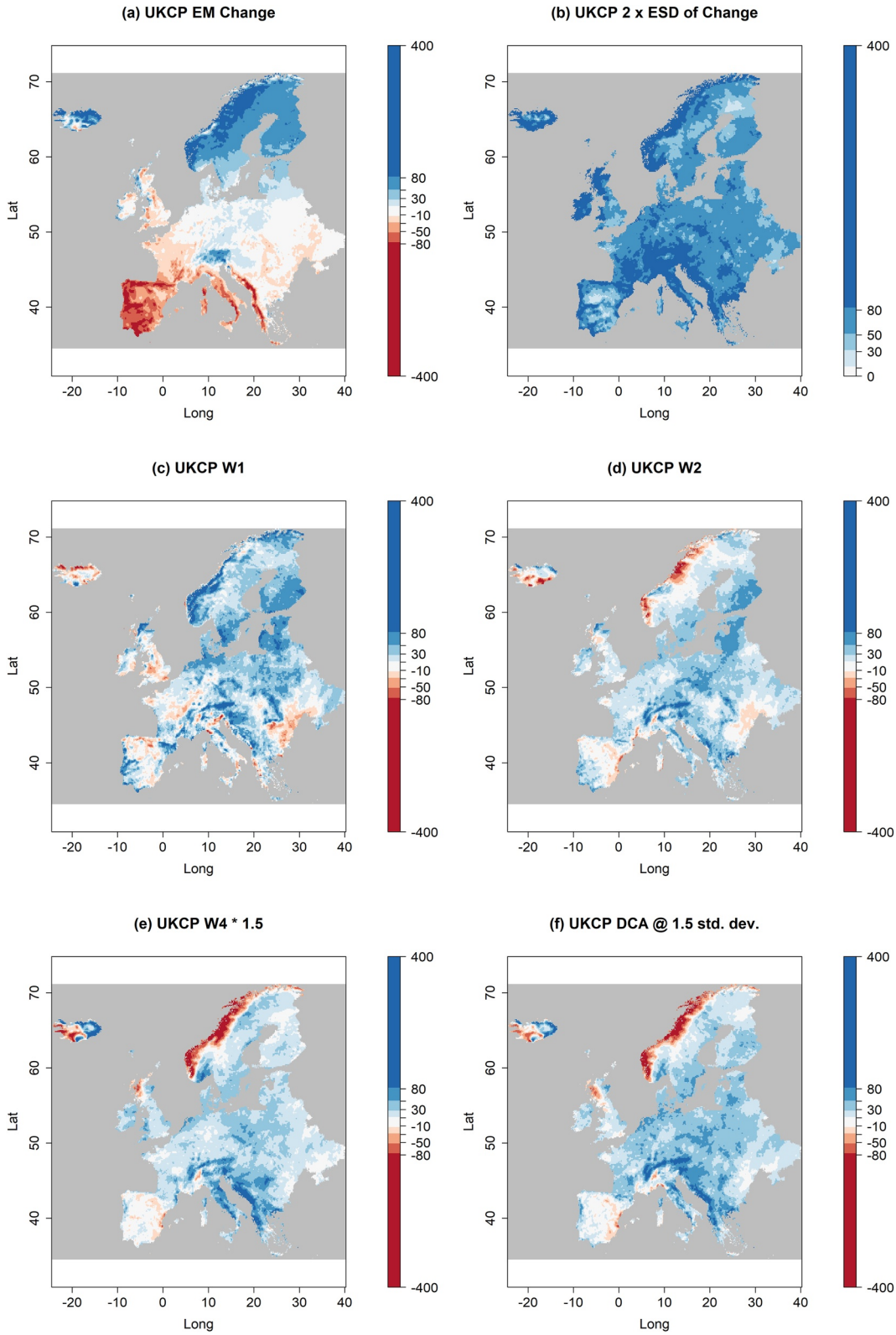

Figure 5.

kind of impact function was previously discussed in the introduction and in Section 4.1 and 4.2. The two UKCP patterns are broadly similar, with regional differences. They are very different from the representative deviation patterns shown in Figure 5 based on an impact metric of total precipitation. The two EURO-CORDEX patterns are also similar, but this time are also similar to the representative deviation patterns shown in Figure 6 based on an impact metric of total precipitation. The UKCP and EURO-CORDEX patterns shown in Figure 7 are very different. This difference in the behavior of the representative deviation patterns between the two ensembles arises because the patterns shown in Figure 7 are defined in a way that takes the sign of the ensemble mean change into account, and the ensemble mean is different for the two ensembles. The patterns in Figure 7 arise from a complicated interaction between the sign of the ensemble mean, the covariance structure of the ensembles and the variability in the ensembles.

The above examples are all based on annual data. For many applications, analyzing higher temporal resolution information would be beneficial. As an example, Figure 8 shows DCA patterns for the EURO-CORDEX data, now calculated season by season, using an impact function defined as the total rainfall within each season. The patterns are somewhat different from the annual pattern shown in Figure 6f. The winter pattern has the largest amplitude, with a dipole of decreasing precipitation in Norway and increasing precipitation in the Iberian peninsula. The spring pattern has the second largest amplitude of the four. The summer and autumn patterns have smaller amplitude, suggesting that the numerical models in the ensemble agree to a greater extent on the changes in precipitation in those seasons.

The real impacts of precipitation variability depend in a detailed way on the temporal and spatial aspects of the precipitation, down to sub-daily timescales and very fine spatial scales. The impacts of climate change then depend on how such variability changes. The use of annual or seasonal fields, and the representative deviation approach, are rather drastic approximations, that throw away much of the spatial and temporal detail. In some cases, such approximations may be useful, while in other cases a more detailed approach may be required.

7. Conclusion

Ensemble climate projections present a large amount of information and ideally, given infinite time and resources, users of climate projections would consider the possible impact of every ensemble member in their analysis. However, this may not always be feasible since some impact models are simply too computationally intensive to run for every ensemble member, every RCP or SSP, every socio-economic scenario (such as migration scenarios), every adaptation scenario (such as flood defense building scenarios), and every time point required. In these situations, an alternative to considering the whole ensemble is to use a smaller ensemble, and an extreme case of using a smaller ensemble is to use the ensemble mean and just one other pattern. We explore the idea of selecting this pattern carefully to be representative of both the variability in the ensemble and possible increases in the impact relative to the ensemble mean. We refer to this as the mean and representative deviation approach. It is very similar to the mean and worst-case approach described in Scher et al. (2021) but is less focused on extremes of impact. Given the ensemble mean and a representative deviation pattern we can calculate both a typical level of impact, using the ensemble mean, and estimate the likely range of impacts, using the representative deviation. One way to create a single representative deviation pattern is by adding (or subtracting) two standard deviations locally. However, the spatial pattern that is created in this way is not, in most situations, a reasonable pattern of uncertainty since it assumes unrealistically high correlations of the uncertainty between locations. We consider four alternatives, all of which rely on a linear definition of the impact of deviations from the ensemble mean. The first method (W1) uses just the most extreme member of the ensemble, in terms of linear impact, as the representative deviation. However, the most extreme member is highly affected by the randomness of the ensemble generation process, and hence is not robust. In a multi-model ensemble, in which some models may be less realistic than others, there is also a risk that the model that determines the most extreme member may be one of the less realistic models. The second and third methods we consider (W2 and W4) consist of averaging the most extreme 2 and 4 members of the ensemble (for an ensemble of 10 members). The fourth method uses the

Figure 5. Annual mean precipitation changes from the UKCP18 ensemble described in the text, for the difference between the periods 2011–2040 and 1981–2010, in mm/yr. EM is ensemble mean and ESD is ensemble standard deviation. Panel (a) shows the ensemble mean change, panel (b) shows twice the standard deviation of change. Panels (c–f) show candidates for representative deviations from the ensemble mean, for an impact metric consisting of domain averaged precipitation. Panel (c) shows the W1 pattern, (d) shows the W2 pattern, (e) shows the W3 pattern, multiplied by 1.5 so that it can be plotted with the same color scale, and (f) shows the Directional Component Analysis pattern at 1.5 standard deviations of impact from the ensemble mean.

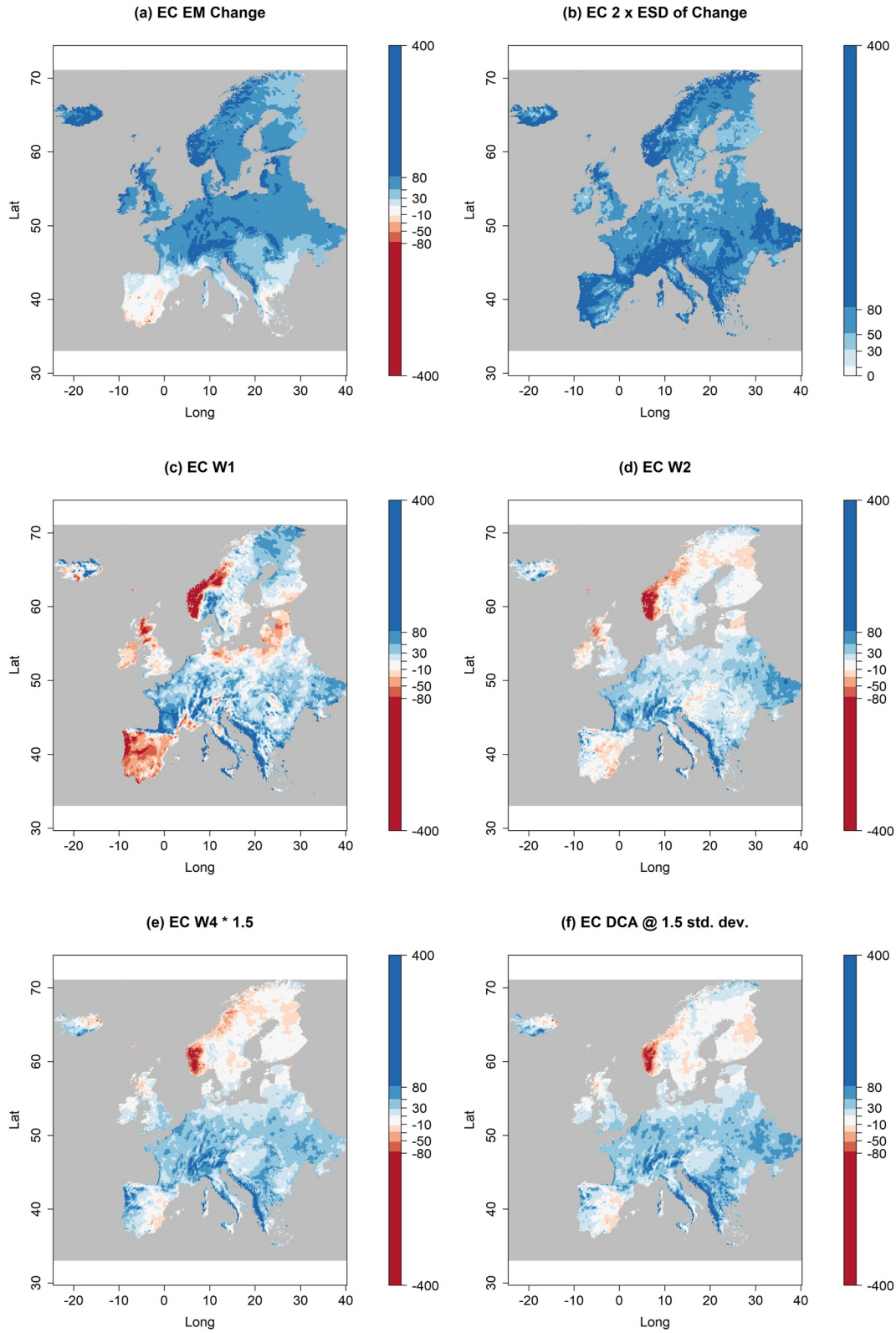

Figure 6. As Figure 5, but now for the EURO-CORDEX ensemble described in the text.

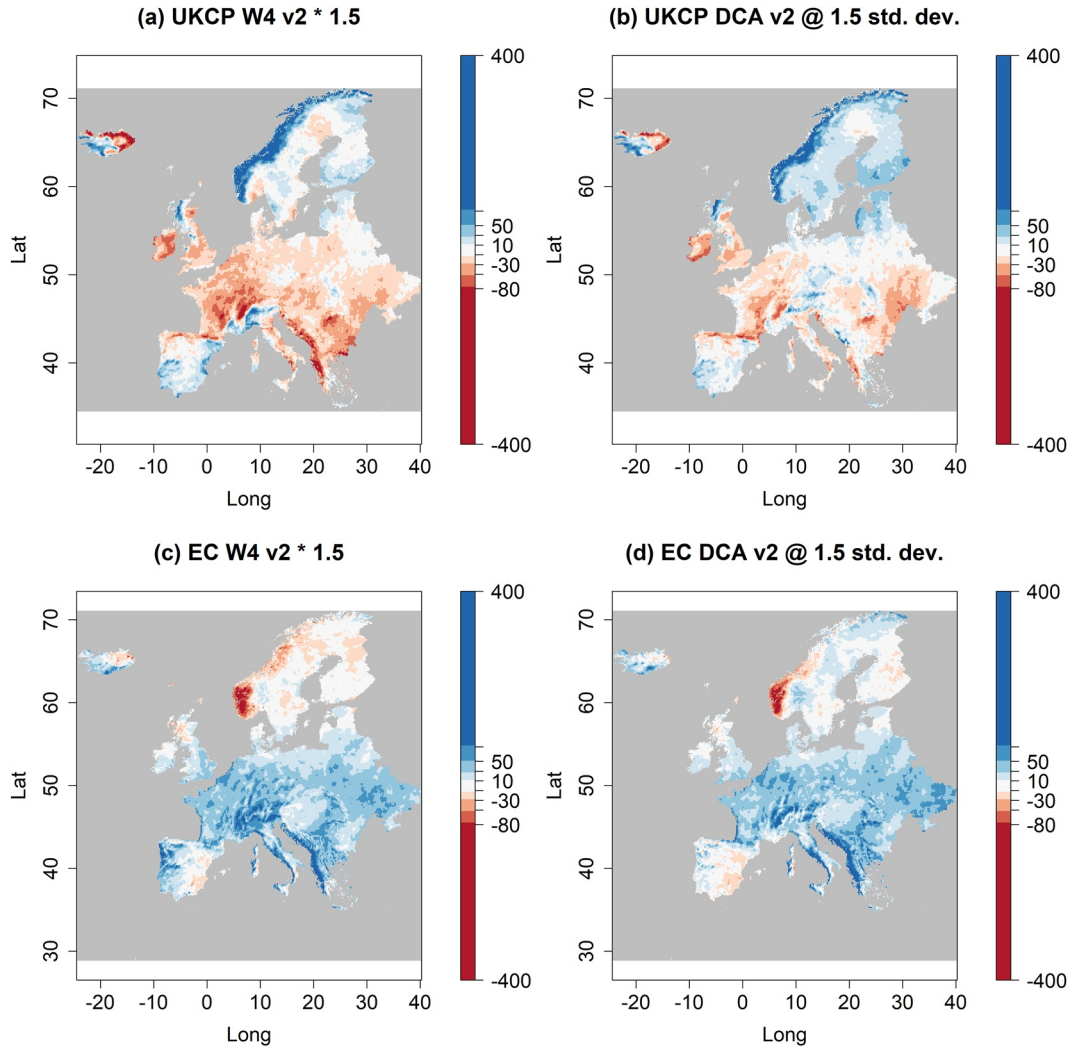

Figure 7. Representative deviations from the ensemble mean, for an impact metric consisting of domain averaged precipitation, multiplied by the local sign of the ensemble mean, so that positive impact is created by positive precipitation anomalies where the ensemble mean is positive, and negative precipitation anomalies where the ensemble mean is negative. Panels (a and b) show results from UKCP18, and panels (c and d) show results from EURO-CORDEX. Panels (a and c) show W4 patterns, multiplied by 1.5 to match the general amplitude of the Directional Component Analysis (DCA) patterns, and panels (b and d) show DCA patterns, normalized to 1.5 standard deviations.

statistical method of DCA (Jewson (2020)), which identifies patterns that maximize a combination of probability density and impact, under certain assumptions.

We have extended the work of Jewson (2020) and Scher et al. (2021) in a number of ways. We have discussed how the linear impact function used in DCA can be derived as an approximation to a more general non-linear impact function, and shown that a higher order quadratic version of DCA is possible, and is scarcely more complex to apply than linear DCA, but is harder to use and interpret because the pattern varies with the level of impact. We have given additional illustrations and simulated examples of W1, W2, W4, and linear DCA, including a case where the highest linear impacts are defined by increasing precipitation in some locations and decreasing precipitation in others. We have presented several new properties of DCA for multivariate normally distributed data, including (a) optimality properties that avoid the use of Lagrangian calculus and show that DCA maximizes certain simple functions related to probability and linear impact and (b) properties that show that DCA can be interpreted as an expectation in several different ways, including as a generalization of the ensemble mean from the expectation over all patterns to the expectation over all patterns with linear impact over a threshold. These new properties give additional insight, help with interpretation, and make a connection between W2, W4, and

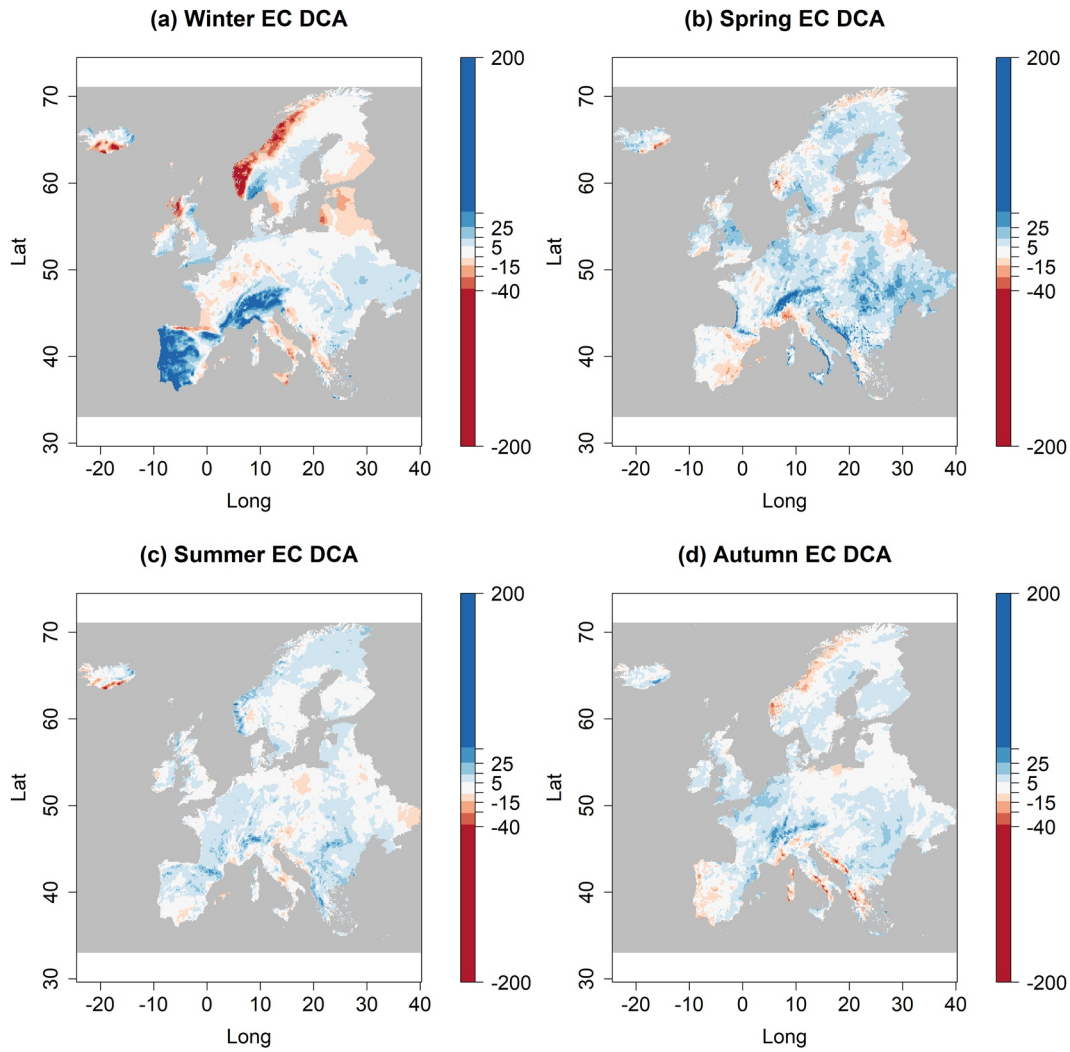

Figure 8. Representative deviations from the ensemble mean, calculated using Directional Component Analysis, for the EURO-CORDEX data used in Figure 6, but now for four seasons.

DCA. It turns out that W2, W4, and DCA would converge as the ensemble size increases in an SME, if the SME is multivariate normally distributed. Finally, we have applied our representative deviation methods to precipitation changes from high-resolution UKCP18 and EURO-CORDEX climate projection ensembles. Analyzing annual mean precipitation using an impact function defined as total precipitation, we find in both ensembles that the most extreme member of the ensemble gives different results to the other methods, and conclude that it is best considered as random and not representative of the variability in the ensemble as a whole. The W2, W4, and DCA methods give somewhat similar results, with the greatest similarities between the W4 and DCA results. These three methods all give patterns that are very different from simply adding two standard deviations from the mean, and are strongly affected by spatial patterns and spatial correlations within the models. The W4 and DCA results from both ensembles consist of a dipole, with decreasing precipitation in Norway and increasing precipitation further south, that may be related to different north-south shifts of the storm track in the different ensemble members. Based on the properties of DCA, this pattern can be interpreted as the most likely pattern to cause an increase in total precipitation relative to the ensemble mean. We have also analyzed the UKCP18 and EURO-CORDEX ensembles using a different impact function, and have analyzed the seasonal dependence of the DCA pattern in the EURO-CORDEX ensemble.

Using the ensemble mean and one additional pattern does not give a full picture of the uncertainty around the ensemble mean, nor of the range of possible impacts. That can only be achieved by a deeper analysis, involving all

the ensemble members. In cases where the impact is non-linear, but the methods we have considered generate a valid ensemble pattern via linearization or approximation of the impact function, then the impact calculated from that pattern using the full non-linear impact function is still a valid impact scenario. In cases where the impact is highly non-linear, using the ensemble mean and one additional pattern may not work at all. For instance, one could imagine a case in which there are critical thresholds in the response that are only exceeded by the most extreme ensemble members, but in which the additional pattern does not exceed those thresholds. One can also imagine cases in which the entire ensemble itself no longer gives reasonable estimates of the impact distribution. Also, in cases where there is no single scalar impact function that summarizes the climate impacts in a useful way, there may be no justifiable way to derive a single additional pattern. For instance, if the goal is to estimate impact everywhere as well as possible, no single scalar impact function would be definable.

However, if there are practical reasons for why it is only possible to consider the ensemble mean and one additional pattern, then using an average of a number of the most extreme impact patterns, or DCA, are both good candidates for deriving an impactful, representative deviation pattern. If the two methods agree, that gives an increased level of robustness. Averaging the most extreme impact patterns has the advantage that it is less likely to be misleading for ensembles with distributions that are very far from elliptical, while the DCA pattern has the advantage that it possesses a number of optimality properties, as discussed in this article, and is more robust because it is calculated using more data.

Appendix A: Mathematical Properties of DCA for Elliptical Distributions

Elliptical distributions are those multivariate distributions for which the probability density for anomalies x from the mean can be written as

$$p(x) = k\, f\left(x^T C^{-1} x\right) \tag{A1}$$

where k is a normalizing constant. Examples are the multivariate normal distribution and the multivariate t distribution. If the function f is strictly decreasing (as it is for the multivariate normal and multivariate t) then maximizing $p(x)$ is equivalent to minimizing $x^T C^{-1} x$. The derivations of property 1 and property 2 (given in Jewson (2020)) are therefore still valid. Alternatively, we can rederive Directional Component Analysis (DCA) from Equation A1 as follows.

In Jewson (2020) DCA is derived using the Lagrange function

$$L = -x^T C^{-1} x + 2\lambda\left(x^T r - a_c\right) \tag{A2}$$

Using Equation A1 the Lagrange function becomes:

$$L = -k\, f\left(x^T C^{-1} x\right) + 2\lambda\left(x^T r - a_c\right) \tag{A3}$$

which has the maximum if

$$k\, f'\left(x^T C^{-1} x\right) C^{-1} x = \lambda r \tag{A4}$$

If f is strictly decreasing then f' is never zero, and this equation the same solution as Equation 5 ($x \propto Cr$).

To cover all elliptical distributions, including those for which f is not strictly decreasing, property 1 has to be amended to become "DCA is the spatial pattern that minimizes the *Mahalanobis distance* for a given value of the linear impact function".

Appendix B: The DCA Pattern Maximizes the Ratio of Linear Impact to Mahalanobis Ratio

We will now show that the DCA pattern provides a solution to the question: given an ensemble with covariance matrix C, what is the spatial pattern x that maximizes the ratio of the linear impact $a' = x^T r$ to the Mahalanobis distance from the ensemble mean $m(x) = \sqrt{x^T C^{-1} x}$. The Mahalanobis distance represents the distance between

the pattern x and the origin, measured using a metric that takes into account the covariance structure: it can be loosely described as being proportional to the number of equally spaced contour lines of probability density between the origin and the pattern x. In one dimension it is the number of standard deviations from the mean. This derivation applies to any probability distribution, although it is most meaningful for elliptically distributed ensembles.

We will consider the function

$$c(x) = \frac{a'(x)}{m(x)} = \frac{x^T r}{\sqrt{x^T C^{-1} x}} \tag{B1}$$

The length of the vector x does not affect the ratio c because it cancels between the numerator and the denominator.

Differentiating Equation B1 by the vector x gives

$$\frac{dc}{dx} = \left(m\frac{da'}{dx} - a'\frac{dm}{dx}\right)/m^2 = \left(mr - \frac{a'C^{-1}x}{m}\right)/m^2 = \left(m^2 r - a'C^{-1}x\right)/m^3 \tag{B2}$$

Setting this equation to zero gives

$$m^2 r = a'C^{-1}x \tag{B3}$$

And hence

$$m^2 Cr = a'x \tag{B4}$$

This is a vector equality, and the two sides of this equation must be parallel, from which we see that x must be parallel to Cr, which means it is parallel to the DCA pattern given by Equation 5. Since the length of the vectors on both sides of Equation B4 are proportional to the length of x squared, this equation says nothing about the length of x: any length of x is a solution. This is confirmed by the example shown in Figure 4e.

In addition, we note that the ratio in Equation 12 can also be written using the fact that $m^2 = \ln(p_0) - \ln(p)$, where p is the probability density of x and p_0 is the probability density for the ensemble mean, x_0, giving $c = a'/\sqrt{\ln(p_0) - \ln(p)}$. This expression also illustrates that DCA captures the idea of finding directions with large values of both a' and p.

Appendix C: The DCA Pattern Maximizes the Product of Linear Impact and Weighted Probability Density

We now show that, for MultiVariate Normal (MVN) ensembles, DCA provides a solution to the question: what pattern maximizes the product of the linear impact and the weighted probability density? This property emphasizes directly that DCA finds patterns that have both a high linear impact and a high probability density. We consider the function:

$$f(x) = p^n(x)a'(x) \tag{C1}$$

where n is positive. Changing the value of n varies the weight on the probability density, relative to the impact.

Differentiating with respect to x, and using the result that for the MVN, $\frac{dp}{dx} = -pC^{-1}x$, gives

$$\frac{df}{dx} = \frac{d(p^n a')}{dx} = p^n\frac{da'}{dx} + na'p^{n-1}\frac{dp}{dx} = p^n r - nx^T r p^n C^{-1}x = p^n\left(r - nx^T rC^{-1}x\right) \tag{C2}$$

Setting this equation equal to zero gives

$$nx^T rx = Cr \tag{C3}$$

where the left hand side is the vector x scaled by the scalar $nx^T r$.

This is a vector equality, and the two sides of this equation must be parallel, from which we see that x must be parallel to Cr, which means it is parallel to the DCA pattern given by Equation 5. The lengths of the vectors on both sides must also be equal which gives

$$x = \frac{Cr}{\sqrt{nr^T Cr}} \tag{C4}$$

This is the unique solution to Equation C1 and is a scaled version of DCA. Large values of n make the solution shorter. Every scaling of DCA is a solution of Equation C1 for a different value of n.

Appendix D: The DCA Vector Is Parallel to the Tail Conditional Expectation

To derive this property, we first need to show that the DCA pattern is parallel to the impact conditional expectation, that is, the vector derived as the expectation over all possible vectors, conditional on a certain single value for the linear impact. In Figure 4b, this expectation is the expectation along one of the diagonal lines. In greater than two dimensions, lines of constant linear impact become planes and hyperplanes, and this expectation is then the expectation over one of these planes or hyperplanes.

This initial property can be shown as follows. Property 1 from Section 2.3.1 above states that DCA maximizes probability density for a given level of linear impact. For an MVN in two or more dimensions, the probability density for a given level of linear impact is also an MVN, in a space with one fewer dimension. For instance, in two dimensions the subspace defined by a given level of linear impact is a line (see the lines in Figure 4b) and the probability density of points on this line is a univariate normal. In three dimensions the subspace defined by a given level of linear impact is a plane, and the probability density of patterns on this plane is a bivariate normal.

To demonstrate this, consider a two-dimensional case, with dimensions x and y, and consider a line of constant impact $x + y = \text{const}$. We can parametrize distance along this line using the value of x and so the probability density, conditional on being on this line, can be written as $p(x|x + y)$. Using Bayes' theorem we can factorize this into

$$p(x|x + y) = \frac{p(x + y|x)p(x)}{p(x + y)} \tag{D1}$$

In this equation, $p(x)$ is a normal distribution by definition, $p(x + y|x)$ is a normal distribution because shifting a normal distribution by a constant gives another normal distribution, and $p(x + y)$ is a normal distribution because the sum of normal distributions is a normal distribution. The terms on the RHS are therefore all normal distributions, and they combine to form a single normal distribution. As a result, the LHS must also be a normal distribution. This demonstration can be extended to other linear impact functions, and to multivariate distributions. It does not work analogously for the multivariate t distribution, since the distributions on the RHS of Equation D1 would not combine to be a multivariate t distribution.

The point with maximum probability density of an MVN is also the mean of that MVN (just as the mode and mean coincide for a univariate normal distribution) and so the DCA pattern is the mean of the reduced dimension MVN.

We write this property for a pattern x, for any value of a', as the proportional relationship:

$$E\left(x|x^T r = a'\right) \propto Cr \tag{D2}$$

We will integrate this expression over a range of levels of linear impact a', and since the integral of DCA patterns is a DCA pattern, we find that DCA is also the expectation conditional on a range of linear impacts. Integrating Equation D2 over a range of linear impacts gives

$$\int_u^v E\left(x|x^T r = a'\right) da' \propto \int_u^v Cr\, da' \tag{D3}$$

The LHS is equal to $E\left(x|x^T r \in (u, v)\right)$. Since Cr does not depend on a' the RHS simplifies to Cr multiplied by a constant factor, giving

$$E\left(x|x^T r \in (u,v)\right) \propto Cr \tag{D4}$$

For instance, we could integrate over the region between two of the diagonal lines in Figure 4b.

Setting $v = \infty$ gives the special case

$$E\left(x|x^T r \in (u,\infty)\right) \propto Cr \tag{D5}$$

in which the range extends to infinity on one side, which we refer to as "tail conditioning," showing that DCA is the tail conditional expectation. If the range extends to infinity on both sides, then the expectation becomes the ensemble mean, which corresponds to the DCA pattern with a length of zero. In this sense the DCA pattern is a generalization of the ensemble mean.

Appendix E: The DCA Vector Is Parallel to the Weighted Tail Conditional Expectation

The fourth new property of DCA that we describe (*property* 6) is that, for MVN data, the direction of the linear DCA vector is the same as the direction of the vector defined as the expectation over all spatial patterns, weighted by the linear impact at each point, and conditional on exceeding a certain value for the linear impact. Relative to property 5, the only difference is the weighting, and we call this the *weighted* tail conditional expectation. In Figure 4b, the weighted tail conditional expectation is the expectation over the region above and to the right of one of the diagonal lines, weighted at each point by the linear impact. This is shown in Appendix E below.

Figure 4h shows patterns estimated in this way, for three different levels of linear impact, given by the three diagonal lines, which have linear impacts of −30, 0, and 30. The patterns calculated as weighted tail conditional expectations do indeed lie on the DCA direction given by the black cross.

This property suggests that an alternative way to derive composites would be to weight the composites using the level of impact, and this may have advantages over more traditional unweighted composites, although we do not explore weighted composites in this article.

This property also follows from Equation D2. In words, the derivation proceeds as follows: since DCA is the expectation of the spatial pattern, conditional on a certain value for the linear impact, it is also the expectation of the spatial pattern weighted by the linear impact, conditional on a certain value for the linear impact. Integrating over a range of impacts then shows that DCA is the expectation of the spatial pattern, weighted by the linear impact at each point, conditional on exceeding a certain value for the linear impact.

Mathematically, the derivation starts with Equation D2, and by multiplying both sides by any integrable function of a', $h(a')$ and integrating over a range of values of a', giving:

$$\int_u^v h\left(a'\right) E\left(x|x^T r = a'\right) da' \propto \int_u^v h\left(a'\right) Cr\, da' \tag{E1}$$

Considering the LHS first: writing the expectation as an integral:

$$\text{LHS} = \int_u^v h\left(a'\right) \int_{x \in V_1} xp(x) dV_1\, da' = \int_u^v \int_{x \in V_1} h\left(a'\right) xp(x)\, dV_1\, da' \tag{E2}$$

where V_1 is the volume defined by the values of x that satisfy $x^T r = a'$, dV_1 is a volume element in this volume, and $h(a')$ can be taken inside the dV_1 integral because it is a constant within this volume.

The two integrals can then be combined into one, giving:

$$\text{LHS} = \int_{x \in V_2} h\left(a'\right) xp(x)\, dV_2 = E\left(h\left(a'\right) x | x \in V_2\right) \tag{E3}$$

where V_2 is the volume defined by the values of x that satisfy $a' = x^T r \in (u,v)$, dV_2 is a volume element in this volume, and $dV_2 = dV_1 da'$. Considering the RHS of Equation E1, and integrating in the same way

$$\text{RHS} = \int_u^v h(a')Cr\,da' = Cr\int_u^v h(a')\,da' \propto Cr \tag{E4}$$

Combining Equations E3 and E4 gives:

$$E\left(h(a')x|x\in V_2\right) \propto Cr \tag{E5}$$

Setting $v = \infty$ gives the special case

$$E\left(h(a')x|x\in V_3\right) \propto Cr \tag{E6}$$

where V_3 is the volume defined by the values of x that satisfy $x^T r\in(u,\infty)$. This is what we set out to prove: the weighted tail conditional expectation over all possible patterns is parallel to DCA.

Setting both $u = -\infty$ and $v = \infty$ gives the special case

$$E\left(h(a')x\right) \propto Cr \tag{E7}$$

which says that the expectation over all patterns, weighted by the linear impact, is parallel to DCA. This is the expectation version of Equation 6.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The climate model data used in this study is freely available from https://euro-cordex.net and https://catalogue.ceda.ac.uk. The computer code used to perform the calculations and generate the figures is available at Jewson (2022).

Acknowledgments
The authors would like to thank Francesco Repola from CMCC for assisting with data extraction and Casper Christophersen, Marie Scholer, and Luisa Mazzotta from EIOPA for catalyzing the collaboration. They would also like to thank the reviewers for helpful comments that have led to various improvements in the manuscript. G. Messori received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant 948309).

References

Abramowitz, G., Herger, N., Gutmann, E., Hammerling, D., Knutti, R., Leduc, M., et al. (2019). Model dependence in multi-model climate. *Earth System Dynamics*, *10*(1), 91–105. https://doi.org/10.5194/esd-10-91-2019

Barnes, E. A., Hurrell, J. W., Ebert-Uphoff, I., Anderson, C., & Anderson, D. (2019). Viewing forced climate patterns through an AI lens. *Geophysical Research Letters*, *46*(22), 13389–13398. https://doi.org/10.1029/2019gl084944

Benestad, R., Haensler, A., Hennemuth, B., Ill, T., Jacob, D., Keup-Thiel, E., et al. (2017). Guidance for EURO-CORDEX. [Online]. Retrieved from https://www.euro-cordex.net/imperia/md/content/csc/cordex/euro-cordex-guidelines-version1.0-2017.08.pdf

Chen, J., Brissette, F. P., Zhang, X. J., Chen, H., Guo, S., & Zhao, Y. (2019). Bias correcting climate model multi-member ensembles to assess climate change impacts on hydrology. *Climatic Change*, *153*(3), 361–377. https://doi.org/10.1007/s10584-019-02393-x

Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2010). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, *38*(3), 527–546. https://doi.org/10.1007/s00382-010-0977-x

European Environment Agency. (2017). Indicator assessment: Mean precipitation. [Online]. Retrieved from https://www.eea.europa.eu/data-and-maps/indicators/european-precipitation-2/assessment

Evans, J. P., Ji, F., Abramowitz, G., & Ekström, M. (2013). Optimally choosing small ensemble members to produce robust climate simulations. *Environmental Research Letters*, *8*(4), 044050. https://doi.org/10.1088/1748-9326/8/4/044050

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Frankcombe, L. M., England, M. H., Mann, M. E., & Steinman, B. A. (2015). Separating internal variability from the externally forced climate response. *Journal of Climate*, *28*(20), 8184–8202. https://doi.org/10.1175/jcli-d-15-0069.1

Hall, A., Cox, P., Huntingford, C., & Klein, S. (2019). Progressing emergent constraints on future climate change. *Nature Climate Change*, *9*(4), 269–278. https://doi.org/10.1038/s41558-019-0436-6

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, *90*(8), 1095–1108. https://doi.org/10.1175/2009bams2607.1

Herger, N., Abramowitz, G., Knutti, R., Angélil, O., Lehmann, K., & Sanderson, B. M. (2018). Selecting a climate model subset to optimise key ensemble properties. *Earth System Dynamics*, *9*(1), 135–151. https://doi.org/10.5194/esd-9-135-2018

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., et al. (2014). EURO-CORDEX: New high-resolution climate change projections for European impact research. *Regional Environmental Change*, *14*(2), 563–578. https://doi.org/10.1007/s10113-013-0499-2

Jacob, D., Teichmann, C., Sobolowski, S., Katragkou, E., Anders, I., Belda, M., et al. (2020). Regional climate downscaling over Europe: Perspectives from the EURO-CORDEX community. *Regional Environmental Change*, *20*(2), 1–20. https://doi.org/10.1007/s10113-020-01606-9

Jewson, S. (2020). An alternative to PCA for estimating dominant patterns of climate variability and extremes, with application to US and China seasonal rainfall. *Atmosphere*, *11*(4), 354. https://doi.org/10.3390/atmos11040354

Jewson, S. (2022). Computer code related to the publication: "Developing representative impact scenarios from climate projection ensembles, with application to UKCP18 and EURO-CORDEX precipitation" [Online]. *Zenodo*. https://doi.org/10.5281/zenodo.5095229

Jewson, S., Barbato, G., Mercogliano, P., Mysiak, J., & Sassi, M. (2021). Improving the potential accuracy and usability of EURO-CORDEX estimates of future rainfall climate using frequentist model averaging. *Nonlinear Processes in Geophysics*, *28*(3), 329–346. https://doi.org/10.5194/npg-28-329-2021

Knutti, R., Sedláček, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., & Eyring, V. (2017). A climate model projection weighting scheme accounting for performance and interdependence. *Geophysical Research Letters*, *44*(4), 1909–1918. https://doi.org/10.1002/2016gl072012

Lowe, J., Bernie, D., Bett, P., Bricheno, L., Brown, S., Calvert, D., et al. (2018). UKCP Science overview report [Online]. Retrieved from https://www.metoffice.gov.uk/research/collaboration/ukcp/guidance-science-reports

Maraun, D. (2016). Bias correcting climate change simulations-a critical review. *Current Climate Change Reports*, *2*(4), 211–220. https://doi.org/10.1007/s40641-016-0050-x

Masson, D., & Knutti, R. (2011). Spatial-scale dependence of climate model performance in the CMIP3 ensemble. *Journal of Climate*, *24*(11), 2680–2692. https://doi.org/10.1175/2011jcli3513.1

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., et al. (2010). The next generation of scenarios for climate change research and assessment. *Nature*, *463*(7282), 747–756. https://doi.org/10.1038/nature08823

Räisänen, J., & Ylhäisi, J. S. (2010). How much should climate model output be smoothed in space? *Journal of Climate*, *24*(3), 867–880. https://doi.org/10.1175/2010jcli3872.1

Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., et al. (2017). The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, *42*, 153–168. https://doi.org/10.1016/j.gloenvcha.2016.05.009

Sanderson, B. M., Wehner, M., & Knutti, R. (2017). Skill and independence weighting for multi-model assessments. *Geoscientific Model Development*, *10*(6), 2379–2395. https://doi.org/10.5194/gmd-10-2379-2017

Scher, S., Jewson, S., & Messori, G. (2021). Robust worst-case scenarios from ensemble forecasts. *Weather and Forecasting*, *36*(4), 1357–1373. https://doi.org/10.1175/waf-d-20-0219.1

Sippel, S., Meinshausen, N., Merrifield, A., Lehner, F., Pendergrass, A. G., Fischer, E., & Knutti, R. (2019). Uncovering the forced climate response from a single ensemble member using statistical learning. *Journal of Climate*, *32*(17), 5677–5699. https://doi.org/10.1175/jcli-d-18-0882.1

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*(4), 485–498. https://doi.org/10.1175/bams-d-11-00094.1

Thompson, D. W., Barnes, E. A., Deser, C., Foust, W. E., & Phillips, A. S. (2015). Quantifying the role of internal climate variability in future climate trends. *Journal of Climate*, *28*(16), 6443–6456. https://doi.org/10.1175/jcli-d-14-00830.1

Tucker, S. O., Kendon, E. J., Bellouin, N., Buonomo, E., Johnson, B., & Murphy, J. M. (2021). Evaluation of a new 12 km regional perturbed parameter ensemble over Europe. *Climate Dynamics*, *58*(3), 879–903. https://doi.org/10.1007/s00382-021-05941-3

Wills, R. C., Battisti, D. S., Armour, K. C., Schneider, T., & Deser, C. (2020). Pattern recognition methods to separate forced responses from internal variability in climate model ensembles and observations. *Journal of Climate*, *33*(20), 8693–8719. https://doi.org/10.1175/jcli-d-19-0855.1