

# Multi-variate factorisation of numerical simulations

Daniel J. Lunt<sup>1</sup>, Deepak Chandan<sup>2</sup>, Harry J. Dowsett<sup>3</sup>, Alan M. Haywood<sup>4</sup>, George M. Lunt<sup>5</sup>, Jonathan C. Rougier<sup>6</sup>, Ulrich Salzmann<sup>7</sup>, Gavin A. Schmidt<sup>8</sup>, and Paul J. Valdes<sup>1</sup>

<sup>1</sup>School of Geographical Sciences, University of Bristol, UK

<sup>2</sup>Department of Physics, University of Toronto, Canada

<sup>3</sup>U.S. Geological Survey, USA

<sup>4</sup>School of Earth and Environment, University of Leeds, UK

<sup>5</sup>AECOM, UK

<sup>6</sup>School of Mathematics, University of Bristol, UK

<sup>7</sup>Geography and Environmental Sciences, Northumbria University

<sup>8</sup>NASA Goddard Institute for Space Studies, USA

**Correspondence:** Daniel J. Lunt (d.j.lunt@bristol.ac.uk)

**Abstract.** Factorisation (also known as ‘factor separation’) is widely used in the analysis of numerical simulations. It allows changes in properties of a system to be attributed to changes in multiple variables associated with that system. There are many possible factorisation methods; here we discuss three previously-proposed factorisations that have been applied in the field of climate modelling: the linear factorisation, the Stein and Alpert (1993) factorisation, and the Lunt et al. (2012) factorisation. We

5 show that, when more than two variables are being considered, none of these three methods possess all three-four properties of ‘uniqueness’, ‘symmetry’, and ‘completeness’ and ‘purity’. Here, we extend each of these factorisations so that they do possess these properties for any number of variables, resulting in three factorisations – the ‘linear-sum’ factorisation, the ‘shared-interaction’ factorisation, and the ‘sealed-total-scaled-residual’ factorisation. We show that the linear-sum factorisation and the shared-interaction factorisation reduce to be identical in the case of four or fewer variables, and we conjecture that  
10 this holds for any number of variables. We present the results of the factorisations in the context of three past studies that used the previously-proposed factorisations. This reveals that only the linear-sum/shared-interaction factorisation possesses a fourth property—‘boundedness’, and as such we recommend the use of this factorisation in applications for which these properties are desirable.

*Copyright statement.* TEXT

## 15 1 Introduction

Factorisation (also known as ‘factor separation’) consists of attributing the total change of some property of a system to multiple components, each component being associated with a change to an internal variable of the system. Multiple tests can be carried out to inform this factorisation, with each test (or simulation in the case of numerical applications) consisting of different combinations of variables. Factorisation experiments are used in many disciplines, with early applications being

20 in agricultural field experiments (Fisher, 1926), and widespread application in industrial and engineering design (Box et al., 2005) ~~and other fields such as medicine (e.g. Smucker et al., 2019)~~. The experiments that underpin such analysis are called  
21 'factorial experiments'. In some cases, in particular when there are a large number of variables, not all combinations of all  
22 variables are tested (usually due to practical or computational limitations), and some previous work has focused on optimising  
23 the experimental design of such 'fractional factorial' experiments (e.g. Domagni et al., 2021). Furthermore, each test often has  
24 an associated error or uncertainty, and may be carried out multiple times. Analysis of such experimental designs is typically  
25 carried out using analysis of variance (ANOVA), in which the total change is represented as a model consisting of a series of  
'main effects', one for each factor, and 'interaction effects' between the factors (Montgomery, 2013).

26 In this paper, we focus on factorisation of numerical model simulations of the climate system; in this case, the factorisation  
27 typically consists of attributing a fundamental property of the climate system to multiple internal model parameters and/or  
28 boundary conditions ~~(Stein and Alpert, 1993; Lunt et al., 2012)~~. In common with previously proposed factorisation methods  
29 in this field (Stein and Alpert, 1993; Lunt et al., 2012), we limit our analysis to the case where there are two possible values for  
30 each variable, and where all combinations of all variables have been simulated; such an experimental design is called a  $2^k$  (or  
31 two-level) full factorial experiment (Montgomery, 2013). Also in common with these studies, we assume that there is zero (or  
32 negligible) uncertainty in each simulation, which is consistent with the deterministic nature of most climate models. Factorisa-  
33 tion has been ~~used applied~~ extensively in the climate literature; some ~~key~~ examples include Claussen et al. (2001), Hogrefe et al.  
34 (2004), van den Heever et al. (2006), and Schmidt et al. (2010); see also the collected studies in Alpert and Sholokhman (2011).  
35 The factorisation ~~method~~ proposed by Stein and Alpert (1993) has ~~currently~~ been cited more than ~~250~~ ~~280~~ times according to  
Web of Science.

## 2 Previous factorisation methods

40 In order to introduce and discuss previous factorisation methods, we use an example case study from the field of climate science.  
41 We turn to the ~~Last Glacial Maximum (LGM), 21~~ Pliocene, ~~~3 million years ago (Haywood et al., 2016, 2020; Dowsett et al., 2016)~~  
42 ~~, 000 years ago~~, the most recent time ~~that the Earth has experienced a large-scale ice age. The LGM was 4–6 of~~ prolonged  
43 ~~natural global warmth relative to pre-industrial (Burke et al., 2018). The Pliocene oceans were on average about 2.5–3.5 °C~~  
44 ~~colder warmer~~ than pre-industrial ~~(Annan and Hargreaves, 2013; Snyder, 2016)~~ (McClymont et al., 2020); for this example, we  
45 would like to know how much of this ~~cooling warmth~~ was due to ~~a decrease an increase~~ in atmospheric CO<sub>2</sub> concentration and  
46 how much ~~was~~ due to the ~~presence reduction in extent and volume~~ of large ice sheets. In this case we would use a climate model  
47 to carry out simulations with ~~various~~ combinations of high and low CO<sub>2</sub> concentrations, and with ~~and without~~ ~~two different~~  
48 ~~configurations of~~ ice sheets. In general there are interactions between the variables so that the contributions from them do not  
49 ~~add sum~~ linearly.

50 It is worth at this stage introducing some notation. Here, we restrict ourselves to the case where there are two possible values  
51 for each variable, denoted '0' and '1'; having more than two values increases the computational cost of a factorisation, and  
52 can reduce the number of factors that can be assessed in a fixed computing budget. We name the fundamental property of the

**Figure 1.** ~~3~~Three different factorisation methods of temperature,  $T$ , for ~~2~~two variables (CO<sub>2</sub> and ice sheets). (a) Linear factorisation, (b) Stein and Alpert (1993) factorisation, (c) Lunt et al. (2012) factorisation. The temperature,  $T$  can be considered as a surface in a third dimension sitting above the 2-dimensional plane of CO<sub>2</sub> and ice sheets. In Equations 1,3, and 4,  $\Delta T_1 = \Delta T_{CO2}$  and  $\Delta T_2 = \Delta T_{ice}$ .

climate system that we are factorising as  $T$ . If there are  $N$  variables, then the results of all possible simulations can be uniquely identified by  $T$  followed by  $N$  subscripts of either 0 or 1, with each subscript representing the value of a variable, with the variables in some predefined order. For our LGM-Pliocene example with two variables ( $N = 2$ ), we have CO<sub>2</sub> (variable 1) and ice (variable 2) contributing to a global mean temperature ( $T$ ); in this case there are 4 possible model simulations: a control (pre-industrial) simulation with pre-industrial CO<sub>2</sub> and pre-industrial ice ( $T_{00}$ ), a second simulation with LGM-Pliocene CO<sub>2</sub> and pre-industrial ice ( $T_{10}$ ), a third simulation with pre-industrial CO<sub>2</sub> and LGM-Pliocene ice ( $T_{01}$ ), and an LGM simulation with LGM a Pliocene simulation with Pliocene CO<sub>2</sub> and LGM-Pliocene ice ( $T_{11}$ ) (see Figure 1a).

## 60 2.1 The linear factorisation

The simplest factorisation that can be carried out is a linear one. For the LGM-Pliocene example with 2 factors, 3 simulations are carried out in which variables are changed consecutively; for example,  $T_{00}$ ,  $T_{10}$ , and  $T_{11}$ . The factorisation of the total change,  $\Delta T$ , between contributions due to CO<sub>2</sub> ( $\Delta T_1$ ) and ice ( $\Delta T_2$ ) would then be:

$$\begin{aligned} \Delta T_1 &= T_{10} - T_{00} \\ 65 \quad \Delta T_2 &= T_{11} - T_{10}. \end{aligned} \quad (1)$$

This factorisation is illustrated graphically in Figure 1(a). However, an equally valid linear factorisation would be

$$\begin{aligned} \Delta T_1 &= T_{11} - T_{01} \\ \Delta T_2 &= T_{01} - T_{00}, \end{aligned} \quad (2)$$

and in a non-linear system this would in general give a different answer to Equation 1. In this sense, the linear factorisation method is not ‘unique’. However, it is ‘complete’ in the sense that the individual factors sum to the total change,  $\Delta T$  exactly, i.e.  $\Delta T_1 + \Delta T_2 = T_{11} - T_{00}$ . Considering the linear factorisation as a ‘path’ starting at  $T_{00}$  and ending at  $T_{11}$ , it is also ‘symmetric’, in that if we instead started from  $T_{11}$  we would retrieve the same numerical values for the two linear factorisations (differing just by a minus sign for the numerical value of each factor). It is also ‘pure’ in that it does not need additional interaction terms (see Section 2.2 and Section 2.3) in order to make it complete.

## 75 2.2 The Stein and Alpert (1993) factorisation

Stein and Alpert (1993) proposed an alternative factorisation method, illustrated in Figure 1(b). In this, for the LGM-Pliocene case, all four possible simulations are carried out, and the factorisation performed relative to the preindustrial pre-industrial case ( $T_{00}$ ) for all variables. The non-linear terms are then all grouped together in a term which is named an interaction term

**Figure 2.** Simulations and linear factorisations in an  $N = 3$  factorisation. Edges that represent changes in CO<sub>2</sub> are in red, changes in ice are in blue, and edges in vegetation are in green. The paths associated with all three possible linear factorisations are shown with dotted lines.

(sometimes called the ‘synergy’),  $S$ :

$$\begin{aligned} 80 \quad \Delta T_1 &= T_{10} - T_{00} \\ \Delta T_2 &= T_{01} - T_{00} \\ S &= T_{11} - T_{10} - T_{01} + T_{00}. \end{aligned} \quad (3)$$

In contrast to the linear factorisation, the Stein and Alpert (1993) factorisation is unique. It is also complete because  $\Delta T_1 + \Delta T_2 + S = T_{11} - T_{00}$  (in fact,  $S$  is defined such that the factorisation is complete). ~~However, As a result of the interaction term,  $S$ , it is not ‘pure’.~~ ~~In addition, it is not~~ symmetric; if we instead performed the factorisation relative to  $T_{11}$ , we would in general obtain a different numerical value of the factorisation (i.e.  $\Delta T_1 = T_{01} - T_{11}$ ).

### 2.3 The Lunt et al (2012) factorisation

Lunt et al. (2012) proposed another factorisation, in which the factorisation for a particular variable is defined as the mean of the difference between each pair of simulations that differ by just that variable. This is illustrated in Figure 1(c); the factorisation of ice is represented by the mean of the two blue lines and the factorisation of CO<sub>2</sub> is represented by the mean of the two red lines:

$$\begin{aligned} \Delta T_1 &= \frac{1}{2} \{ (T_{10} - T_{00}) + (T_{11} - T_{01}) \} \\ \Delta T_2 &= \frac{1}{2} \{ (T_{01} - T_{00}) + (T_{11} - T_{10}) \}. \end{aligned} \quad (4)$$

The  $N = 2$  factorisation in Equation 4 is unique, complete, ~~and symmetric~~<sup>symmetric, and pure</sup>. It is worth noting that Equation 4 can be interpreted in multiple ways – either (i) as described above, the factorisation averages all the possible pairs of simulations that differ solely by a change in that variable, i.e. for a particular variable it is the mean of either the horizontal or vertical edges of the square in Figure 1(c); or (ii) it is the average of the two possible linear factorisations in Equations 1 and 2; or (iii) it is the average of the two possible Stein and Alpert (1993) factorisations obtained by swapping the ~~LGM~~<sup>Pliocene</sup> and pre-industrial (in which case the ~~synergy~~<sup>interaction</sup> terms cancel); or (iv) it is the Stein and Alpert (1993) factorisation but with the ~~synergy~~<sup>interaction</sup> term,  $S$ , shared equally between the two factors.

In extending to  $N = 4$  variables, Lunt et al. (2012) assumed that the first of these interpretations would still hold for any number of variables. However, consider the  $N = 3$  case illustrated in Figure 2, in which we have added vegetation as a third variable to contribute to ~~LGM~~<sup>cooling</sup>~~Pliocene~~<sup>warming</sup>. Averaging the edges (interpretation (i) above) would result in a

factorisation:

$$\begin{aligned} 105 \quad \Delta T'_1 &= \frac{1}{4} \{ (T_{100} - T_{000}) + (T_{110} - T_{010}) + (T_{101} - T_{001}) + (T_{111} - T_{011}) \} \\ \Delta T'_2 &= \frac{1}{4} \{ (T_{010} - T_{000}) + (T_{110} - T_{100}) + (T_{011} - T_{001}) + (T_{111} - T_{101}) \} \\ \Delta T'_3 &= \frac{1}{4} \{ (T_{001} - T_{000}) + (T_{101} - T_{100}) + (T_{011} - T_{010}) + (T_{111} - T_{110}) \} \end{aligned} \quad (5)$$

Although this is unique and symmetric, it is not complete, because  $\Delta T'_1 + \Delta T'_2 + \Delta T'_3 \neq T_{111} - T_{000}$  (this is immediately apparent by considering the  $T_{111}$  term, for which terms; the three lines in Equation 5 each include a term equal to  $\frac{1}{4}T_{111}$ , which sum to  $\frac{3}{4}T_{111}$ , whereas they are required to sum to  $T_{111}$  for a complete factorisation). As such, an additional synergy term,  $S$ , interaction term, in the sense of  $S$  in Equation 3, would be required for the factorisation to be complete in  $N = 3$  dimensions (in which case it would no longer be pure). Note that the Lunt et al. (2012) factorisation is complete for  $N = 2$  without such an interaction term, but this is a case specific to  $N = 2$  as a result of cancellation of terms in Equation 4.

## 115 2.4 Summary of previous factorisations

As shown above, neither the linear, or the Stein and Alpert (1993), or the Lunt et al. (2012) factorisation methods possess all three properties of uniqueness, symmetry, purity, and completeness in  $N > 2$  dimensions. These properties are often desirable in a factorisation, because any factorisation that lacks one of these properties is less easy to interpret. For example, for the LGM-Pliocene example above, uniqueness means that we can have a single answer to the question “why is the LGM colder than the preindustrial Pliocene warmer than the pre-industrial”. Symmetry means that we obtain the same answer to the question “why is the LGM colder than the preindustrial Pliocene warmer than the pre-industrial” as to the question “why is the pre-industrial warmer than the LGM colder than the Pliocene”. Completeness means that the answer to the question “how much warmer is the Pliocene than the pre-industrial” is equal to the sum of the individual factors (plus an interaction term if one exists). “Purity” means that we can answer the question “why is the LGM colder Pliocene warmer than the pre-industrial” by referring solely to contributions from our fundamental factors  $\text{CO}_2$ , ice, and vegetation, i.e. without including additional synergistic interaction terms that are not attributed to a single factor. These synergistic interaction terms are important and interesting, but there are cases where it can be useful or essential to only include attributable terms in the factorisation.

## 3 Extensions to the previous factorisations

Here we discuss possible extensions to the three previous factorisations discussed above, that are unique, symmetric, pure, and complete in  $N$  dimensions.

### 130 3.1 Extension to the linear factorisation: The linear-sum factorisation

The linear-sum factorisation arises from a generalisation to  $N > 2$  dimensions of the second interpretation of Equation 4; i.e. it arises from averaging all the possible linear factorisations. This will result in a complete and pure factorisation because each

individual linear factorisation is itself complete and pure. For three dimensions, this is illustrated by the dotted lines in Figure

135 2.

Each possible linear factorisation can be represented as a non-returning ‘path’ from the vertex  $T_{000}$  to the opposite vertex  $T_{111}$ , traversing edges along the way (dotted lines in Figure 2). When considering the sum of all possible paths, some edges are traversed more than others. In general, those edges near the initial or final vertices are traversed more times than edges that are further away from these vertices. As such, when we average the possible linear factorisations, different edges (corresponding to different terms in the factorisation) will have different weightings. This is in contrast to Equation 5 where each term (i.e. edge of the cube) has the same weighting. For three dimensions, Figure 2 shows that the 6 edges adjacent to the initial and final vertex are traversed twice, whereas the 6 other edges are traversed only once. Therefore, the factorisation is :

$$\begin{aligned}\Delta T_1 &= \frac{1}{6} \{2(T_{100} - T_{000}) + (T_{110} - T_{010}) + (T_{101} - T_{001}) + 2(T_{111} - T_{011})\} \\ \Delta T_2 &= \frac{1}{6} \{2(T_{010} - T_{000}) + (T_{110} - T_{100}) + (T_{011} - T_{001}) + 2(T_{111} - T_{101})\} \\ 145 \quad \Delta T_3 &= \frac{1}{6} \{2(T_{001} - T_{000}) + (T_{101} - T_{100}) + (T_{011} - T_{010}) + 2(T_{111} - T_{110})\}. \end{aligned} \quad (6)$$

This factorisation is complete ( $\Delta T_1 + \Delta T_2 + \Delta T_3 = T_{111} - T_{000}$ ), unique, ~~and symmetric~~symmetric, and pure.

To generalise to  $N$  dimensions, consider an  $N$ -dimensional cube, which has a total of  $2^N$  vertices and  $N \times 2^{N-1}$  edges. There are  $2^{N-1}$  edges in each dimension. There are  $N!$  paths from the initial vertex of the cube to the final opposite vertex, each of which consists of a traverse along  $N$  edges. Therefore, in each dimension there are a total of  $N!$  edges traversed for all paths combined.

As for the 3-dimensional case above, let us label each vertex,  $V$ , of this  $N$ -dimensional cube as  $V_{a_1 \dots a_N}$ , where each  $a_i$  is either 0 or 1. A value  $a_i = 0$  represents the first value for variable  $i$ , and  $a_i = 1$  represents the second value for variable  $i$ . Each vertex is also associated with a system value, denoted  $T_{a_1 \dots a_N}$  (see Figure 2 for the case  $N = 3$ ).

All factorisations consist of partitioning the total change,  $\Delta T = T_{1\dots 1} - T_{0\dots 0}$  between  $N$  factors. Each factor is associated 155 with a dimension,  $i$ , in the  $N$ -dimensional cube. The factorisation for dimension  $i$  is  $\Delta T_i$ .

For the linear-sum factorisation, all paths that we consider start at the origin vertex,  $0 \dots 0$ , and end at the opposite vertex  $1 \dots 1$ , and are made up of a series of edges. For all edges on the  $N$ -dimensional cube, let us define  $X$  as the set of all possible starting vertices, for a given  $N$ . For example, for  $N = 3$ ,  $X = \{000, 001, 010, 011, 100, 101, 110\}$ . Let us define  $X_i$  as the set of all possible starting vertices for an edge that is oriented in the  $i$ th dimension, i.e. all those vertices that have a 0 in the 160  $i$ th subscript. For example, for  $N = 3$  and  $i = 2$ ,  $X_2 = \{000, 001, 100, 101\}$ . Let us define  $Y_i$  as the set of all possible ending vertices for an edge that is oriented in the  $i$ th dimension, so that  $Y_i$  is related to  $X_i$  by changing the  $i$ th subscript of each element from 0 to 1. For example, for  $N = 3$  and  $i = 2$ ,  $Y_2 = \{010, 011, 110, 111\}$ . Order  $X_i$  and  $Y_i$  so that their elements correspond. Then we write  $X_i^j$  to indicate the  $j$ th element of  $X_i$ , and  $Y_i^j$  as indicating the  $j$ th element of  $Y_i$ . For example, for the  $X_2$  defined above,  $X_2^3 = 100$ .

$$\Delta T'_i = \frac{1}{2^{N-1}} \sum_{j=1}^{2^{N-1}} (T_{Y_i^j} - T_{X_i^j}) \quad (7)$$

For the linear-sum factorisation, we instead carry out a weighted average, with the weight for each edge in dimension  $i$  given by how many times it is traversed in all  $N!$  paths. Consider all the paths that traverse an edge which starts at a vertex defined by  $k$  subscripts of '1' and  $N - k$  subscripts of '0'. There are  $k!$  possible paths to the start of this edge, and  $(N - k - 1)!$  paths 170 from the end of this edge to the final corner (defined by  $N$  subscripts of '1'). Therefore, there are  $k! \times (N - k - 1)!$  paths that use this edge. As such, we can write the linear-sum factorisation as:

$$\Delta T_i = \frac{1}{N!} \sum_{j=1}^{2^{N-1}} \left\{ k_i^j! (N - 1 - k_i^j)! (T_{Y_i^j} - T_{X_i^j}) \right\}, \quad (8)$$

where  $k_i^j$  is the number of subscripts of '1' in  $X_i^j$ .

For example, for  $N = 4$  and  $i = 1$ , we have  $N! = 24$  edges traversed in this dimension, and  $2^{N-1} = 8$  edges. 175  $X_1 = \{0000, 0001, 0010, 0100, 0011, 0101, 0110, 0111\}$ , and  $Y_1 = \{1000, 1001, 1010, 1100, 1011, 1101, 1110, 1111\}$ . For those edges with a starting subscript with  $k = 0$  subscripts of '1' (i.e. 0000), the weighting  $k! (N - 1 - k)! = 0! (4 - 1 - 0)! = 6$ . For those edges with a starting subscript with  $k = 1$  subscripts of '1' (i.e. 0001, 0010, 0100), the weighting  $k! (N - 1 - k)! = 1! (4 - 1 - 1)! = 2$ . For those edges with a starting subscript with  $k = 2$  subscripts of '1' (i.e. 0011, 0101, 0110), the weighting  $k! (N - 1 - k)! = 2! (4 - 1 - 2)! = 2$ . For those edges with a starting subscript with  $k = 3$  subscripts of '1' (i.e. 0111), the 180 weighting  $k! (N - 1 - k)! = 3! (4 - 1 - 3)! = 6$ . Therefore, for  $N = 4$  and  $i = 1$ , we have:

$$\begin{aligned} \Delta T_1 = & \frac{1}{24} \{ 6(T_{1000} - T_{0000}) + 2(T_{1001} - T_{0001}) + 2(T_{1010} - T_{0010}) + 2(T_{1100} - T_{0100}) + \\ & 2(T_{1011} - T_{0011}) + 2(T_{1101} - T_{0101}) + 2(T_{1110} - T_{0110}) + 6(T_{1111} - T_{0111}) \}. \end{aligned} \quad (9)$$

### 3.2 Extension to the Stein and Alpert (1993) factorisation: the shared-interactions factorisation

As stated in Section 2.3, the Lunt et al. (2012) factorisation for  $N = 2$  can be interpreted as being identical to the Stein and Alpert 185 (1993) factorisation but with the synergy interaction term shared between the two factors (thereby removing the interaction term, resulting in a pure factorisation). Here we explore what happens when this interpretation is generalised to  $N > 2$  dimensions. For consistency, we use the same notation as (Stein and Alpert, 1993) Stein and Alpert (1993). In their notation,  $\hat{f}_1$  represents the difference between a simulation in which only factor  $i$  is modified with a simulation in which no factors are modified, and  $\hat{f}_{ijk\dots}$  represents interaction terms between the different factors. For example, for our original  $N = 2$  example 190 illustrated in Figure 1 and given in Equation 3,  $\Delta T_1 \equiv \hat{f}_1$ ,  $\Delta T_2 \equiv \hat{f}_2$ , and  $S \equiv \hat{f}_{12}$ .

For our LGM-Pliocene example for  $N = 3$ ,  $\hat{f}_{12}$  is the interaction term (i.e. the synergy) between factors 1 and 2 ( $\text{CO}_2$  and ice),  $\hat{f}_{13}$  is the interaction term between factors 1 and 3 ( $\text{CO}_2$  and vegetation),  $\hat{f}_{23}$  is the interaction term between factors 2 and

**Figure 3.** (a) Visual representation of the shared-interaction factorisation for  $N = 3$ , as given by Equation 10. The straight dotted lines represent the sharing of the interactions according to Equation 11. (b) Visual representation of the shared-interaction factorisation for  $N = 4$ . The straight dotted lines represent the sharing of the interactions according to Equation 13.

3 (ice and vegetation), and  $\hat{f}_{123}$  is the interaction **term** between all three factors. In this case, Stein and Alpert (1993) give that

$$\begin{aligned}
 \Delta T &= \hat{f}_1 + \hat{f}_2 + \hat{f}_3 + \hat{f}_{12} + \hat{f}_{13} + \hat{f}_{23} + \hat{f}_{123} \\
 195 \quad \hat{f}_1 &= T_{100} - T_{000} \\
 \hat{f}_2 &= T_{010} - T_{000} \\
 \hat{f}_3 &= T_{001} - T_{000} \\
 \hat{f}_{12} &= T_{110} - (T_{100} + T_{010}) + T_{000} \\
 \hat{f}_{13} &= T_{101} - (T_{100} + T_{001}) + T_{000} \\
 200 \quad \hat{f}_{23} &= T_{011} - (T_{010} + T_{001}) + T_{000} \\
 \hat{f}_{123} &= T_{111} - (T_{110} + T_{101} + T_{011}) + (T_{100} + T_{010} + T_{001}) - T_{000}. \tag{10}
 \end{aligned}$$

As discussed in Section 2.2, this factorisation is not symmetric or unique (e.g. we could define  $\hat{f}_1 = T_{011} - T_{111}$ ), **and it is only or pure, but it is** complete if we include all the interaction terms, which are not attributed to any particular factor. By extending the interpretation of **shared synergy the shared interaction term** in 2 dimensions discussed in Section 2.3, we can

205 choose to share the interaction terms equally between their contributing factors, an approach applied by Schmidt et al. (2010) (although they carried out a **partial fractional** factorisation in which not all combinations of all variables were included). This results in a factorisation that is complete **and pure** (because we are just re-partitioning the interaction terms). It turns out that it is also symmetric. For example for  $\text{CO}_2$ ,

$$\Delta T_1 = \hat{f}_1 + \frac{1}{2}\hat{f}_{12} + \frac{1}{2}\hat{f}_{13} + \frac{1}{3}\hat{f}_{123}. \tag{11}$$

210 This factorisation for  $N = 3$  is represented visually in Figure 3(a). Equations 10 and 11 give that, for  $\text{CO}_2$ ,

$$\Delta T_1 = \frac{1}{6}\{2(T_{100} - T_{000}) + (T_{110} - T_{010}) + (T_{101} - T_{001}) + 2(T_{111} - T_{011})\}. \tag{12}$$

This is identical to the equivalent term in Equation 6, indicating that the shared-interaction and linear-sum interpretations are identical for  $N = 3$ , and that therefore for  $N = 3$  the shared-interaction factorisation is unique, symmetric, **pure**, and complete.

Stein and Alpert (1993) give the generalisation of their factorisation to  $N$  factors (their Equations 11-16). For  $N = 4$ , the 215 interaction terms are shared so that, for example for  $\text{CO}_2$ ,

$$\Delta T_1 = \hat{f}_1 + \frac{1}{2}(\hat{f}_{12} + \hat{f}_{13} + \hat{f}_{14}) + \frac{1}{3}(\hat{f}_{123} + \hat{f}_{124} + \hat{f}_{134}) + \frac{1}{4}\hat{f}_{1234}. \tag{13}$$

This factorisation for  $N = 4$  is represented visually in Figure 3(b). Again, for  $N = 4$  this is **the same as the identical to the** linear-sum interpretation (Equation 9). We conjecture that for any  $N$  these two interpretations will give identical results.

In the sealed-total-scaled-residual factorisation, the Lunt et al. (2012) factorisation is modified so that it is complete (and remains pure). This is achieved by taking the total residual term required for completeness (the ‘synergy’,  $S$  in the sense of (Stein and Alpert, 1993)), and sharing this between the factors in proportion to the sign and magnitude of their Lunt et al. (2012) factorisation. For the  $N = 3$  example of the LGM Pliocene, we have that the synergy, residual term,  $R$ , is defined such that

$$\Delta T'_1 + \Delta T'_2 + \Delta T'_3 + SR = T_{111} - T_{000}, \quad (14)$$

where the  $\Delta T'_i$  are defined in Equation 5. We then share the synergy this residual proportionally across the three factors, such that

$$\begin{aligned} \Delta T_1 &= \Delta T'_1 + \frac{S\Delta T'_1}{\Delta T'_1 + \Delta T'_2 + \Delta T'_3} \frac{R|\Delta T'_1|}{|\Delta T'_1| + |\Delta T'_2| + |\Delta T'_3|} \\ 230 \quad \Delta T_2 &= \Delta T'_2 + \frac{S\Delta T'_2}{\Delta T'_1 + \Delta T'_2 + \Delta T'_3} \frac{R|\Delta T'_2|}{|\Delta T'_1| + |\Delta T'_2| + |\Delta T'_3|} \\ \Delta T_3 &= \Delta T'_3 + \frac{S\Delta T'_3}{\Delta T'_1 + \Delta T'_2 + \Delta T'_3} \frac{R|\Delta T'_3|}{|\Delta T'_1| + |\Delta T'_2| + |\Delta T'_3|} \end{aligned} \quad (15)$$

Equations 14 and ?? reduce to:

$$\begin{aligned} \underline{\Delta T_1} &\equiv \underline{\Delta T'_1} \frac{T_{111} - T_{000}}{\Delta T'_1 + \Delta T'_2 + \Delta T'_3} \\ \underline{\Delta T_2} &\equiv \underline{\Delta T'_2} \frac{T_{111} - T_{000}}{\Delta T'_1 + \Delta T'_2 + \Delta T'_3} \\ 235 \quad \underline{\Delta T_3} &\equiv \underline{\Delta T'_3} \frac{T_{111} - T_{000}}{\Delta T'_1 + \Delta T'_2 + \Delta T'_3}. \end{aligned}$$

This shows that this factorisation can also be interpreted as simply sealing the Lunt et al. (2012) factorisation so that the sum of the factors equals  $T_{111} - T_{000}$ . In  $N$  dimensions, this generalises to:

$$\Delta T_i = \Delta T'_i \frac{T_{1\dots i} - T_{0\dots 0}}{\sum_{j=1}^N \Delta T'_i} + \frac{R|\Delta T'_i|}{\sum_{j=1}^N |\Delta T'_i|} \quad (16)$$

where  $\Delta T'_i$  is defined in Equation 7, and

$$240 \quad R = T_{1\dots 1} - T_{0\dots 0} - \sum_{j=1}^N \Delta T'_j. \quad (17)$$

For example, for  $N = 4$  and  $i = 1$  we have:

$$\begin{aligned}
 \Delta T'_1 &= \frac{1}{8} \{ (T_{1000} - T_{0000}) + (T_{1001} - T_{0001}) + (T_{1010} - T_{0010}) + (T_{1100} - T_{0100}) + \\
 &\quad (T_{1011} - T_{0011}) + (T_{1101} - T_{0101}) + (T_{1110} - T_{0110}) + (T_{1111} - T_{0111}) \} \\
 \underline{R} &\equiv \underline{T_{1111} - T_{0000}} - (\Delta T'_1 + \Delta T'_2 + \Delta T'_3 + \Delta T'_4) \\
 245 \quad \underline{\Delta T_1} &= \Delta T'_1 \frac{\underline{T_{1111} - T_{0000}}}{\underline{\Delta T'_1 + \Delta T'_2 + \Delta T'_3 + \Delta T'_4}} + \frac{R|\Delta T'_1|}{|\Delta T'_1| + |\Delta T'_2| + |\Delta T'_3| + |\Delta T'_4|}; \tag{18}
 \end{aligned}$$

and similarly for  $\Delta T'_2$ ,  $\Delta T'_3$ , and  $\Delta T'_4$ .

## 4 Implications for previous published work

Here we discuss three examples of papers in which the Lunt et al. (2012) factorisation has been used. For each, we show how using our factorisations would affect the results in that paper.

### 250 4.1 Implications for Lunt et al (2012)

Lunt et al. (2012) presented a factorisation of global mean temperature change in the Pliocene (~~3 million years ago, the most recent time of prolonged natural global warmth relative to pre-industrial~~) into four variables: CO<sub>2</sub>, orography, ice, and vegetation. As described in Section 2.3, in extending to  $N = 4$  variables, the Lunt et al. (2012) factorisation is unique~~and symmetric, symmetric, and pure~~, but not complete. Using their notation, their factorisation for the CO<sub>2</sub> variable is (equivalent to Equation 255 9 in their paper):

$$\begin{aligned}
 dT'_{CO_2} &= \frac{1}{8} \{ (T_c - T) + (T_{oc} - T_o) + (T_{ic} - T_i) + (T_{vc} - T_v) + \\
 &\quad (T_{ocv} - T_{ov}) + (T_{oci} - T_{oi}) + (T_{civ} - T_{iv}) + (T_{ociv} - T_{oiv}) \}. \tag{19}
 \end{aligned}$$

The equivalent linear-sum/shared-interaction factorisation is given by Equation 9, which in the notation of Lunt et al. (2012) is:

$$\begin{aligned}
 260 \quad dT_{CO_2} &= \frac{1}{24} \{ 6(T_c - T) + 2(T_{oc} - T_o) + 2(T_{ic} - T_i) + 2(T_{vc} - T_v) + \\
 &\quad 2(T_{ocv} - T_{ov}) + 2(T_{oci} - T_{oi}) + 2(T_{civ} - T_{iv}) + 6(T_{ociv} - T_{oiv}) \}, \tag{20}
 \end{aligned}$$

and similarly for the other three variables.

The equivalent ~~sealed-total scaled-residual~~ factorisation is given by Equation 18, which in the notation of Lunt et al. (2012) is:

$$\begin{aligned}
 265 \quad \underline{R} &\equiv \underline{T_{ociv} - T} - (dT'_{CO_2} + \underline{dT'_{orog}} + \underline{dT'_{ice}} + \underline{dT'_{veg}}) \\
 \underline{dT_{CO_2}} &= dT'_{CO_2} \frac{\underline{T_{ociv} - T}}{\underline{dT'_{CO_2} + dT'_{orog} + dT'_{ice} + dT'_{veg}}} + \frac{R|dT'_{CO_2}|}{|dT'_{CO_2}| + |dT'_{orog}| + |dT'_{ice}| + |dT'_{veg}|} \tag{21}
 \end{aligned}$$

where  $dT'_{CO_2}$  is given in Equation 19; and similarly for the other three variables.

In Lunt et al. (2012), although Equation 19 (Equation 9 in their paper) was presented, the four variables were actually factorised by two  $N = 2$  factorisations for all the analysis in that paper (Equation 13 in their paper). Because for  $N = 2$  dimensions the Lunt et al. (2012), linear-sum/shared-interaction, and ~~sealed-total scaled-residual~~ factorisations are identical, the actual results related to Pliocene temperature change presented in Lunt et al. (2012) would not be affected by using our proposed new factorisations.

## 4.2 Implications for Haywood et al (2016)

Haywood et al. (2016), in the context of the experimental design for model simulations of the Pliocene in the PlioMIP project, 275 presented a 3-variable factorisation of Pliocene warming into components due to  $CO_2$ , topography, and ice, based on the Lunt et al. (2012) factorisation (presented in their Section 3.2).

An alternative, using the linear-sum/shared-interaction factorisation that is complete, is obtained from Equation 6, which in their notation is, for  $CO_2$  (and analogously for the other two components):

$$dT'_{CO_2} = \frac{1}{6} \{ 2(E^{400} - E^{280}) + (Eo^{400} - Eo^{280}) + (Ei^{400} - Ei^{280}) + 2(Eoi^{400} - Eoi^{280}) \} \quad 280 \quad (22)$$

Another alternative, using the ~~sealed-total scaled-residual~~ factorisation that is complete, is obtained from [Equation 24](#) [Equations 14 and 15](#), which in their notation is, for  $CO_2$  (and analogously for the other two components):

$$\begin{aligned} dT'_{CO_2} &= \frac{1}{4} \{ (E^{400} - E^{280}) + (Eo^{400} - Eo^{280}) + (Ei^{400} - Ei^{280}) + (Eoi^{400} - Eoi^{280}) \} \\ R &\equiv \underline{Eoi^{400} - E^{280} - (dT'_{CO_2} + dT'_{orog} + dT'_{ice})} \\ 285 \quad dT'_{CO_2} &= dT'_{CO_2} \frac{Eoi^{400} - E^{280}}{dT'_{CO_2} + dT'_{orog} + dT'_{ice}} + \frac{R|dT'_{CO_2}|}{|dT'_{CO_2}| + |dT'_{orog}| + |dT'_{ice}|}. \end{aligned} \quad (23)$$

## 4.3 Implications for Chandan and Peltier (2018)

Chandan and Peltier (2018) applied the  $N = 3$  factorisation of Lunt et al. (2012) (Equation 5), as also given by Haywood et al. (2016) (first line of Equation 23), to their suite of Pliocene simulations. The factorisation was applied to each gridcell in the model, resulting in  $192 \times 288 = 55,296$  factorisations over the globe. The two-dimensional mid-Pliocene minus ~~preindustrial temperature anomaly~~ [pre-industrial temperature anomaly, reproduced here in Figure 4\(a\)](#), was factorised into contributions originating from a change in  $CO_2$ , orography and ice sheets, ~~and is reproduced here in Figure 4(a)~~. Figure 4(b–d) shows the results of the original factorisation and is identical to those presented in Figure 7 of Chandan and Peltier (2018). Figure 4(f–h) shows the factorisation of the same anomaly using the linear-sum/shared-interaction method (Equations 22) while Figure 4(j–m) shows the results of employing the ~~sealed-total scaled-residual~~ method (Equations 23). [The first thing to note is that the three factorisations all have very similar results; visually it is difficult to tell them apart on a regional scale, and they result in global means for each factor that differ by less than 10%. This is because, in this example, the non-linearities \(i.e. the](#)

**Figure 4.** Comparison of various factorisation methods. (a) The mid-Pliocene minus ~~preindustrial~~ pre-industrial anomaly ~~modeled~~ ( $T_{111} - T_{000}$ ) ~~modelled~~ by Chandan and Peltier (2017). (b–m) The top three rows present factorisations of the total anomaly into contributions arising from changes to CO<sub>2</sub> (upper, (b,f,j)), orography (middle, (c,g,k)) and ice sheets (lower, (d,h,l)), while the bottom row shows the residual ~~required for completeness~~ (‘synergy’),  $T_{111} - T_{000} - (\Delta T_{CO_2} + \Delta T_{orog} + \Delta T_{ice})$ , (e,i,m). Note that the residual term,  $R$ , for panel (e) is given by Equation 14, and is equal to  $T_{111} - T_{000} - (\Delta T'_{CO_2} + \Delta T'_{orog} + \Delta T'_{ice})$ . The first column (b,c,d,e) shows results using the methodology of Lunt et al. (2012) and is identical to results reported in Figure 7 of Chandan and Peltier (2018). The second column (f,g,h,i) shows results from the linear-sum/shared-interaction factorisation (Eq. [Equation 6](#)) and the third column (j,k,l,m) shows results of the ~~scaled-total scaled-residual~~ factorisation (Eq. [24](#)[Equation 15](#)). The ~~pink circles in the factorized results shown in the rightmost column highlight regions where the scaled-total factorisation has very large negative or positive compensating values for at the three factors, due to the very small values top-right of denominator term appearing each panel give the global mean, in Equants of °C, 24 at those locations.~~

~~interaction terms) are relatively small. As such, the main conclusions of the Chandan and Peltier (2018) study are robust to a change in factorisation methodology.~~

The bottom row in Figure 4 shows, for the case of each method, the residual difference between the sum of all the factors and the total change (i.e. the ~~synergy interaction/synergy terms~~ in the sense of Stein and Alpert (1993)). The Lunt et al. (2012) method yields spatially coherent structures in the residual whose magnitude can be comparable to those of the factorized components, whereas the residuals for the other two methods are zero by definition, because they are ~~complete pure~~ (in the Figures they are very close to zero – essentially numerical noise due to round-off error). The non-linearity (indicated by the magnitude of the ~~synergy residual term~~ associated with the Lunt et al. (2012) factorisation) is greatest in the North Atlantic (Figure 4d), and is likely associated with changes in the sea-ice margin that are non-linearly influenced by all three boundary conditions (CO<sub>2</sub>, orography, and ice sheets).

~~Figure 4(j–l) reveals a problem with the scaled-total method. In these panels, the pink circles show regions where the scaled-total factorisation has very large negative or positive values for the three factors. At these locations the denominator in Eq. 24 is very small, resulting in very large magnitude positive or negative results for each factorised components, which sum to a much smaller number. This is clearly not a meaningful result (because the values in the same region in Figure 4(a) are not unusual), and although in this analysis these issues are found to occur only at isolated locations, in other cases there is potential for the problem to be more widespread. In response to this, we introduce an a fourth property of factorisations – ‘boundedness’. A factorisation is bounded if the factorisation for a particular variable (e.g.  $\Delta T_1$ ) is bounded by the minimum and maximum of all the possible single-factor factorisations for that variable. For example, for four dimensions, a factorisation is bounded if  $\Delta T_1$  has a value that is not greater than the largest, or smaller than the smallest, term in Equation 18. The linear-sum/shared-interaction factorisation is by definition bounded, because it consists of a weighted average of those very terms. In contrast, the scaled-total factorisation is not bounded, and as such it should only be used with caution. It is worth noting that if absolute weightings were used in Equation ??, such that the scaled-total factorisation became (e.g. for CO<sub>2</sub>) We also explored using a version of the scaled-residual factorisation in which the residual terms was shared, not by the absolute~~

320 magnitude of the individual factors, but by their relative values, so that Equation 15 became:

$$\Delta T_1 = \Delta T'_1 + \frac{S|\Delta T'_1|}{|\Delta T'_1| + |\Delta T'_2| + |\Delta T'_3|} \frac{R\Delta T'_1}{\Delta T'_1 + \Delta T'_2 + \Delta T'_3}, \quad (24)$$

325 then the factorisation would not result in spuriously large values (because the denominator could never approach zero) and similarly for the other variables. However, the factorisation would still not be bounded in our definition. For example, if  $\Delta T'_1$  were negative and consisted of all negative terms in the first line of Equation 5,  $\Delta T_1$  could still be positive if  $S$  were sufficiently large at a small number of gridcells this produced highly divergent results, when the denominator in Equation 24 became very close to zero (see Supp Info, Figure 1).

## 5 Conclusions

In this paper, we have reviewed three previously-proposed factorisations, and extended them to produce factorisations that are unique, symmetric, pure, and complete. We have presented them for 3 dimensions (i.e. 3 factors), and generalised to  $N$  dimensions. The first factorisation, ‘linear-sum’ (Equation 8), averages all the possible linear factorisations on the  $N$ -dimensional cube. The second factorisation, ‘shared-interaction’, shares the interaction terms between each corresponding factor equally. The linear-sum and shared-interaction factorisations are shown to reduce to be identical for  $N \leq 4$ , and we conjecture that this holds for any  $N$ . The third factorisation, ‘sealed-total scaled-residual’ (Equation 16), averages all the contributions associated with the edges of the  $N$ -dimensional cube, and scales them by the total change in the property being factorised takes the residual term required for completeness in the Lunt et al. (2012) factorisation, and shares it between the factors, in proportion to their magnitude. We have presented results of these extended factorisations in the context of previous work carried out by Lunt et al. (2012), Haywood et al. (2016), and Chandan and Peltier (2018) in the context of Pliocene climate change. This reveals that the sealed-total factorisation is not bounded, and therefore can lead to anomalous results that are hard to interpret. Therefore we We recommend the use of the linear-sum/shared-interaction factorisation or the scaled-residual factorisation for cases where the properties of uniqueness, symmetry, purity, and completeness, and boundedness are desirable. In some cases, the interaction terms may, of course, be of great interest, and in such cases a non-pure factorisation (e.g. Stein and Alpert, 1993) can be very informative. Also, it is worth noting that if the interaction terms are zero, i.e. we have a completely linear system, then all the factorisation methods reduce to be identical. The properties of all the factorisations discussed in this paper are shown in Table ?? for 2, 3, 4, and  $N$  dimensions.

345 The methods that we present here will be of particular use in the analysis of systems with multiple variables, and have application beyond solely climate science.

*Code and data availability.* The model fields underlying Figure 4 are available from the University of Toronto Dataverse in netcdf format: <https://doi.org/10.5683/SP2/QGK5B0>. The code used to calculate the factorisations illustrated in Figure 4 is available in the Supplement, in both python and ncl. The ncl version is also available on github, here: [https://github.com/danlunt1976/factor\\_separation/blob/master/factor\\_separation.ncl](https://github.com/danlunt1976/factor_separation/blob/master/factor_separation.ncl)

350 *Author contributions.* DJL led the study and wrote the first draft of the paper. GAS made Figure 3 and DC made Figure 4. DJL devised the linear-sum factorisation, GAS devised the shared-interaction factorisation, and GML devised the scaled-residual factorisation. JCR provided the derivation of Equation 8. HJD, US, PJV, and AMH developed the boundary conditions and early Pliocene modelling that underpin the Pliocene simulations discussed. All authors contributed to writing the final version of the paper.

*Competing interests.* The authors declare no competing interests.

355 *Acknowledgements.* DJLacknowledges, JCR, PJV, and US acknowledge support from NE/P01903X/1. HJD states that "Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government".

## References

Alpert, P. and Sholokhman, T., eds.: Factor Separation in the Atmosphere: Applications and Future Prospects, Cambridge University Press, <https://doi.org/10.1017/CBO9780511921414>, 2011.

360 Annan, J. D. and Hargreaves, J. C.: A new global reconstruction of temperature changes at the Last Glacial Maximum, *Climate of the Past*, 9, 367–376, <https://doi.org/10.5194/cp-9-367-2013>, 2013.

Box, G. E. P., Hunter, J. S., and Hunter, W. G.: Statistics for Experimenters, John Wiley and Sons, 2 edn., 2005.

Burke, K. D., Williams, J. W., Chandler, M. A., Haywood, A. M., Lunt, D. J., and Otto-Bliesner, B. L.: Pliocene and 365 Eocene provide best analogs for near-future climates, *Proceedings of the National Academy of Sciences*, 115, 13 288–13 293, <https://doi.org/10.1073/pnas.1809600115>, <https://www.pnas.org/content/115/52/13288>, 2018.

Chandan, D. and Peltier, W. R.: Regional and global climate for the mid-Pliocene using the University of Toronto version of CCSM4 and PlioMIP2 boundary conditions, *Climate of the Past*, 13, 919–942, <https://doi.org/10.5194/cp-13-919-2017>, 2017.

Chandan, D. and Peltier, W. R.: On the mechanisms of warming the mid-Pliocene and the inference of a hierarchy of climate sensitivities with relevance to the understanding of climate futures, *Climate of the Past*, 14, 825–856, <https://doi.org/10.5194/cp-14-825-2018>, 2018.

370 Claussen, M., Brovkin, V., and Ganopolski, A.: Biogeophysical versus biogeochemical feedbacks of large-scale land cover change, *Geophysical Research Letters*, 28, 1011–1014, <https://doi.org/10.1029/2000GL012471>, 2001.

Domagni, K. F., Hedayat, A., and Sinha, B. K.: On the nature of saturated 2k- factorial designs for unbiased estimation of non-negligible parameters, *Journal of Statistical Planning and Inference*, 211, 33 – 40, <https://doi.org/https://doi.org/10.1016/j.jspi.2020.05.006>, <http://www.sciencedirect.com/science/article/pii/S0378375820300641>, 2021.

375 Dowsett, H., Dolan, A., Rowley, D., Moucha, R., Forte, A. M., Mitrovica, J. X., Pound, M., Salzmann, U., Robinson, M., Chandler, M., Foley, K., and Haywood, A.: The PRISM4 (mid-Piacenzian) paleoenvironmental reconstruction, *Climate of the Past*, 12, 1519–1538, <https://doi.org/10.5194/cp-12-1519-2016>, <https://cp.copernicus.org/articles/12/1519/2016/>, 2016.

Fisher, R. A.: The arrangement of field experiments, *Journal of the Ministry of Agriculture of Great Britain*, 33, 503–513, 1926.

Haywood, A. M., Dowsett, H. J., Dolan, A. M., Rowley, D., Abe-Ouchi, A., Otto-Bliesner, B., Chandler, M. A., Hunter, S. J., Lunt, D. J., 380 Pound, M., and Salzmann, U.: The Pliocene Model Intercomparison Project (PlioMIP) Phase 2: Scientific objectives and experimental design, *Climate of the Past*, 12, 663–675, <https://doi.org/10.5194/cp-12-663-2016>, 2016.

Haywood, A. M., Tindall, J. C., Dowsett, H. J., Dolan, A. M., Foley, K. M., Hunter, S. J., Hill, D. J., Chan, W.-L., Abe-Ouchi, A., Stepanek, C., Lohmann, G., Chandan, D., Peltier, W. R., Tan, N., Contoux, C., Ramstein, G., Li, X., Zhang, Z., Guo, C., Nisancioglu, K. H., Zhang, Q., Li, Q., Kamae, Y., Chandler, M. A., Sohl, L. E., Otto-Bliesner, B. L., Feng, R., Brady, E. C., von der Heydt, A. S., Baatsen, M. L. J., 385 and Lunt, D. J.: The Pliocene Model Intercomparison Project Phase 2: large-scale climate features and climate sensitivity, *Climate of the Past*, 16, 2095–2123, <https://doi.org/10.5194/cp-16-2095-2020>, <https://cp.copernicus.org/articles/16/2095/2020/>, 2020.

Hogrefe, C., Lynn, B., Civerolo, K., Ku, J.-Y., Rosenthal, J., Rosenzweig, C., Goldberg, R., Gaffin, S., Knowlton, K., and Kinney, P. L.: Simulating changes in regional air pollution over the eastern United States due to changes in global and regional climate and emissions, *Journal of Geophysical Research: Atmospheres*, 109, <https://doi.org/10.1029/2004JD004690>, 2004.

390 Lunt, D. J., Haywood, A. M., Schmidt, G. A., Salzmann, U., Valdes, P. J., Dowsett, H. J., and Loptson, C. A.: On the causes of mid-Pliocene warmth and polar amplification, *Earth and Planetary Science Letters*, 321-322, 128 – 138, <https://doi.org/10.1016/j.epsl.2011.12.042>, 2012.

McClumont, E. L., Ford, H. L., Ho, S. L., Tindall, J. C., Haywood, A. M., Alonso-Garcia, M., Bailey, I., Berke, M. A., Littler, K., Patterson, M. O., Petrick, B., Peterse, F., Ravelo, A. C., Risebrobakken, B., De Schepper, S., Swann, G. E. A., Thirumalai, K., Tierney, J. E., van der Weijst, C., White, S., Abe-Ouchi, A., Baatsen, M. L. J., Brady, E. C., Chan, W.-L., Chandan, D., Feng, R., Guo, C., von der Heydt, A. S., Hunter, S., Li, X., Lohmann, G., Nisancioglu, K. H., Otto-Bliesner, B. L., Peltier, W. R., Stepanek, C., and Zhang, Z.: Lessons from a high-CO<sub>2</sub> world: an ocean view from  $\sim$  3 million years ago, *Climate of the Past*, 16, 1599–1615, <https://doi.org/10.5194/cp-16-1599-2020>, <https://cp.copernicus.org/articles/16/1599/2020/>, 2020.

395 Montgomery, D. C.: *Design and Analysis of Experiments*, Wiley, 2013.

400 Schmidt, G. A., Ruedy, R., Miller, R. L., and Lacis, A. A.: The attribution of the present-day total greenhouse effect, *J. Geophys. Res.*, 115, D20106, <https://doi.org/10.1029/2010JD014287>, 2010.

Smucker, B., Krzywinski, M., and Altman, N.: Two-level factorial experiments, *Nature Methods*, 16, 211–212, <https://doi.org/https://doi.org/10.1038/s41592-019-0335-9>, 2019.

Snyder, C. W.: Evolution of global temperature over the past two million years, *Nature*, 538, 226–228, 2016.

405 Stein, U. and Alpert, P.: Factor separation in numerical simulations, *J. Atmos. Sci.*, 50, 2107–2115, 1993.

van den Heever, S. C., Carrió, G. G., Cotton, W. R., DeMott, P. J., and Prenni, A. J.: Impacts of Nucleating Aerosol on Florida Storms. Part I: Mesoscale Simulations, *Journal of the Atmospheric Sciences*, 63, 1752–1775, <https://doi.org/10.1175/JAS3713.1>, 2006.