



# Estimation of the local response to a forcing in a high dimensional system using the fluctuation-dissipation theorem

F. C. Cooper<sup>1</sup>, J. G. Esler<sup>2</sup>, and P. H. Haynes<sup>3</sup>

<sup>1</sup>Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, UK

<sup>2</sup>Department of Mathematics, University College, London, UK

<sup>3</sup>Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK

Correspondence to: F. C. Cooper (f.c.cooper@littlestick.com)

Received: 1 August 2012 – Revised: 22 February 2013 – Accepted: 28 February 2013 – Published: 26 April 2013

**Abstract.** The fluctuation-dissipation theorem (FDT) has been proposed as a method of calculating the response of the earth's atmosphere to a forcing. For this problem the high dimensionality of the relevant data sets makes truncation necessary. Here we propose a method of truncation based upon the assumption that the response to a localised forcing is spatially localised, as an alternative to the standard method of choosing a number of the leading empirical orthogonal functions. For systems where this assumption holds, the response to any sufficiently small non-localised forcing may be estimated using a set of truncations that are chosen algorithmically. We test our algorithm using 36 and 72 variable versions of a stochastic Lorenz 95 system of ordinary differential equations. We find that, for long integrations, the bias in the response estimated by the FDT is reduced from  $\sim 75\%$  of the true response to  $\sim 30\%$ .

## 1 Introduction

An important problem in atmospheric sciences is how the climate responds to a forcing. For example how does the mean local temperature respond to an increase in carbon dioxide or a change in the incident solar radiation. Explicit simulation can estimate the effects of a particular forcing; however, simulation can be prohibitively expensive if the response to a large set of possible forcings is required. One method of reducing the cost of a local estimate is to truncate the calculation to a local simulation at a high resolution forced by boundary conditions given by a low resolution global simulation. For estimating the response to a forcing applied over a short time this regional approach seems effective, (see for

example the UK Met office numerical weather prediction model Davies et al., 2005); however, it remains to be seen if it can be extended to the long time climate response. The hypothesis explored in this paper is that a statistical approximation based upon the fluctuation-dissipation theorem (FDT), using only data on the unforced system, will have some skill in predicting the response to a forcing. The form of the FDT that we consider predicts only the linear component of the response, i.e. the response of a (fully non-linear) system to a sufficiently small forcing, and much in the same way that simulation is expensive in computational effort, the FDT is expensive in data. Some sort of truncation is necessary to reduce the quantity of data required. In this paper we describe a method of truncating a data set locally for the purposes of applying the FDT. This method, inspired by the technique of local simulation at high resolution, is a practical alternative to the globally distributed truncation to a particular number of empirical orthogonal functions (EOFs); see Joliffe (2002) chapter 6. The linear nature of the FDT then allows for the responses calculated with local truncation to many small localised forcings to be added together to give the global response to a forcing that is not localised.

The expected response  $\delta\mathbf{x}(t)$  of a system to a forcing as a function of time  $t$  is denoted by

$$\langle\delta\mathbf{x}(t)\rangle = \langle\mathbf{x}_f(t)\rangle - \langle\mathbf{x}(t)\rangle,$$

where  $\mathbf{x}(t)$  represents the state vector of the system without forcing applied,  $\mathbf{x}_f(t)$  represents the state vector of the system with the forcing applied and the angled brackets  $\langle\dots\rangle$  denote the ensemble average over a large number of independent realisations. Assuming that the system is sufficiently random and that its probability density function (PDF)  $\rho(\mathbf{x})$

is differentiable, the linear component of  $\langle \delta \mathbf{x}(t) \rangle$  is given by

$$\langle \delta \mathbf{x}(t) \rangle = - \int_0^t \left\langle \mathbf{x}(\tau) \left[ \frac{\nabla_{\mathbf{x}} \rho(\mathbf{x}(0))}{\rho(\mathbf{x}(0))} \right]^T \right\rangle \delta \mathbf{f}(t - \tau) d\tau, \quad (1)$$

where  $\delta \mathbf{f}(t)$  represents a forcing vector at time  $t$  and  $T$  denotes the vector transpose; see for example Marconi et al. (2008) for a recent review. Throughout this paper we make the assumption that the mean of  $\mathbf{x}$  has been subtracted and therefore  $\langle \mathbf{x}(t) \rangle = 0$ .

For practical evaluation of (1) knowledge of  $\rho(\mathbf{x})$  is required. If it is assumed that  $\rho(\mathbf{x})$  is well approximated by a Gaussian, (1) reduces to the Gaussian FDT (sometimes called the quasi-Gaussian FDT)

$$\langle \delta \mathbf{x}(t) \rangle \approx \int_0^t \mathbf{C}(\tau) \mathbf{C}(0)^{-1} \delta \mathbf{f}(t - \tau) d\tau, \quad (2)$$

where  $\mathbf{C}(\tau)$  is the lag  $\tau$  covariance matrix. Gritsun and Branstator (2007) amongst others have applied (2) to atmospheric general circulation model simulations with promising results. The Gaussian FDT requires the inversion of a covariance matrix which can be difficult when considering a large state vector  $\mathbf{x}$ . In order to evaluate (2) Gritsun and Branstator reduced the size of their state vector by first selecting particular variables and truncating their system from 18 352 components to 1800 EOFs. Majda et al. (2010) suggests some strategies for truncation in EOF space, however, a general quantitative theory for such truncation, e.g. predicting any biases introduced, has not yet been developed.

In the context of global climate modelling aspects of the dynamical system, such as separation of timescales associated with hydrostatic and geostrophic balance, can be exploited to facilitate truncation. For example, Ring and Plumb (2008) use the solution of an ‘‘Eliassen problem’’ based on geostrophic balance to truncate an idealised general circulation model by expressing all variables in terms of the zonal wind. Their truncation is then used, for example, to convert a particular forcing of global temperature into an equivalent forcing of the zonal wind, that can then be used in a truncated version of (1).

One problem with (2) is that the assumption of Gaussianity where it is not appropriate may cause unacceptably large errors in a response prediction (Cooper and Haynes, 2013). An alternative approach, appropriate for short time responses, is to use an ensemble adjoint technique (Eyink et al., 2004). Equation (1) is recast into a different form that depends upon a tangent linear model. This technique overcomes the problem of an unknown PDF, but is valid only up to a maximum lag  $\tau$ , and requires the availability of the tangent linear model (and hence cannot be used with data derived from measurements). Abramov and Majda (2007) introduce a ‘‘blended response algorithm’’ that extends the Eyink et al. (2004) approach to longer lags by assuming that (2) is accurate for

sufficiently large  $\tau$ . An alternative suggested by Cooper and Haynes (2011) is to use the methods of non-parametric statistics to approximate  $\rho(\mathbf{x})$ . Unfortunately, the non-parametric FDT they suggest is susceptible to problems due to a high dimensional state vector and some form of truncation appears helpful.

To explore possible truncation strategies a simple toy model is an appropriate first step. We consider as our test bed a stochastic version of the Lorenz 95 system (Lorenz, 1995) whose state vector  $\mathbf{x} = (x_1, x_2, x_3, \dots, x_d)$  contains  $d > 3$  variables and is governed by

$$\frac{dx_j}{dt} = x_{j-1}(x_{j+1} - x_{j-2}) - x_j + f + \delta f_j + a\xi, \quad (3)$$

where  $f$  is a constant equal to 8 in this paper. This choice of  $f$  makes the system chaotic.  $j = 1, 2, 3, \dots, d$ ,  $\delta f_j$  represents an additional small forcing,  $\xi$  represents a Gaussian white noise term with unit variance with  $a$  a constant controlling its amplitude. The state vector is periodic, so that in (3)  $x_0 = x_d$ ,  $x_{-1} = x_{d-1}$  and  $x_{d+1} = x_1$ . The precise details of our integrations are detailed in Appendix A. (The Lorenz 95 system is sometimes referred to as the Lorenz 96 system in the literature due to a confusion regarding the dates of first publication.) This particular system is appropriate to this investigation because the dimensionality of its phase space is easy to vary and can be made large enough to allow exploration of the effectiveness of truncation, it is symmetric in the sense that each point has the same relation to its neighbours, and its solution with an appropriate choice of  $f$ , is chaotic. In addition, the system is local in the sense that each element of the state vector only directly effects its close neighbours. More distant elements of the state vector are only affected indirectly allowing for the possibility of correlation decaying with distance as in many physical systems. Lucarini and Sarno (2011) demonstrate that the Lorenz 95 system possesses many of the properties required for investigation using some form of linear response theory and Abramov and Majda (2007) tested the various forms of their blended response algorithm (that requires a tangent linear model) to predict the response of the Lorenz 95 system. They reported predictions with similar accuracy to those of this paper. We must bear in mind that the Lorenz 95 system may have a very different underlying nature to that of the true climate system, so the results reported here may or may not be applicable to that case.

The response after a long time has elapsed of each element of  $\mathbf{x}$  to a continuous unit forcing applied to a single point  $x_k$  of the Lorenz 95 system (3) is plotted in Fig. 1. The response is estimated from two long integrations of (3), with  $\delta f_k = 0$  and 1, respectively. An initial period, longer than any correlation times, is discarded and the experiment is repeated ten times for robustness; see Appendix A for details. The final element of  $\mathbf{x}$  in each plot is linked by (3) to the first. The response to the left of the applied force is relatively small and almost completely restricted to two points. The response to

the right is larger in places with the amplitude reducing in a non-monotonic manner with distance. The overall impression is that the response is somewhat local to the forcing and does not strongly depend upon the dimensionality of  $\mathbf{x}$ . In effect, the extra points in the  $d = 72$  system do not make a significant contribution to the response and we should be able to simply ignore them when making a response estimate. The experiment repeated for  $d = 18$  and  $d = 144$  does not alter this conclusion. Figure 1 suggests that the response will only be well captured if the system is truncated locally rather than in a non-local Fourier or EOF space. For the purpose of applying (1), we are interested in the linear component of the response. For a forcing sufficiently small to be in the linear regime, the response to forcing more than one element of  $\mathbf{x}$  may be estimated as the sum of the responses due to forcing the single elements separately. In the case of the Lorenz 95 system, all grid points are equivalent and we can calculate the response to any linear forcing by considering the response of a single grid point. In other words, the linear response to any combination of  $\delta f_j$  may be estimated using Fig. 1. Therefore, we restrict ourselves to forcing  $\delta f$  only in the direction of  $x_k$ , so that  $\delta f_k = 1$  and  $\delta f_j = 0$  for  $j \neq k$ .

To investigate possible truncations, we first split the state vector into three components

$$\mathbf{x} = \begin{pmatrix} x_k \\ \mathbf{x}_d \\ \mathbf{x}_i \end{pmatrix},$$

where  $x_k$  is the element to which a forcing is applied,  $\mathbf{x}_d$  is a vector of the components of  $\mathbf{x}$  that are closely related to  $x_k$ , and  $\mathbf{x}_i$  is a vector of the components of  $\mathbf{x}$  that are in some sense approximately independent of  $x_k$  during an integration. We now introduce a localised response approximation (LRA) that the PDF  $\rho_{k|(d,i)}$  of  $x_k$  given a particular value of all of the other state vector elements is equal to the PDF  $\rho_{k|d}$  of  $x_k$  given only the values of  $\mathbf{x}_d$ ,

$$\rho_{k|(d,i)}(x_k|\mathbf{x}_d, \mathbf{x}_i) \approx \rho_{k|d}(x_k|\mathbf{x}_d). \quad (4)$$

This assumption is justified by the fact that changes to  $\rho_{k|(d,i)}$  are negligible given sufficiently small variations in  $\mathbf{x}_i$ , i.e. typical random or chaotic variations in  $\mathbf{x}_i$  are not large enough to influence the value of  $x_k$ . The utility of the LRA therefore depends upon the relative dimension of  $\mathbf{x}_d$  and  $\mathbf{x}_i$ . We use the LRA of (4) and the fact that

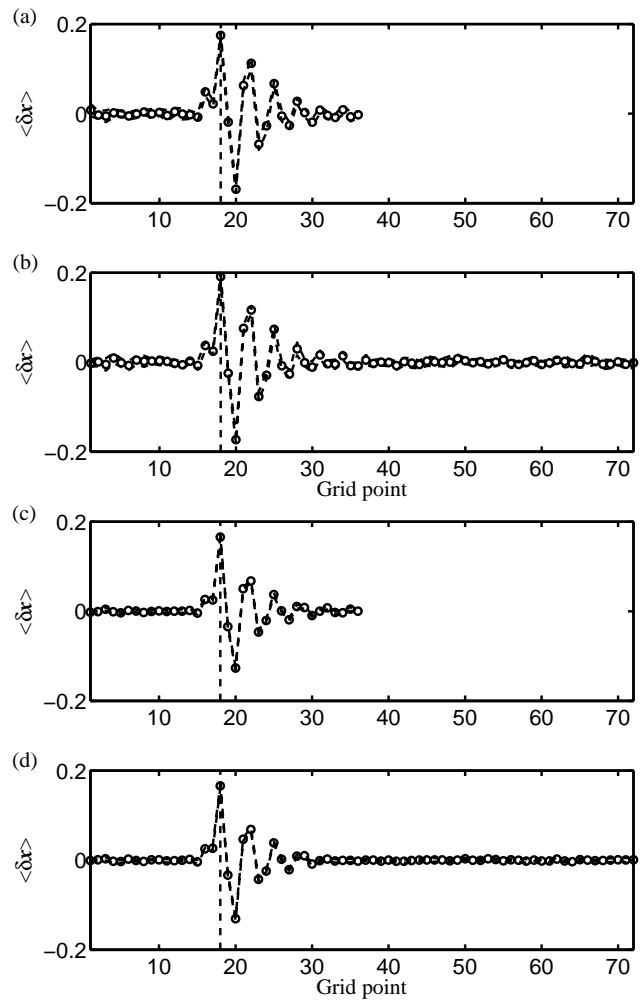
$$\rho_{k|(d,i)}(x_k|\mathbf{x}_d, \mathbf{x}_i) = \frac{\rho(\mathbf{x})}{\int \rho(\mathbf{x}) dx_k} \quad (5)$$

and

$$\rho_{k|d}(x_k|\mathbf{x}_d) = \frac{\int \rho(\mathbf{x}) dx_i}{\iint \rho(\mathbf{x}) dx_i dx_k} \quad (6)$$

to write

$$\rho(\mathbf{x}) = \frac{\rho_{k,d}(x_k, \mathbf{x}_d) \rho_{d,i}(\mathbf{x}_d, \mathbf{x}_i)}{\rho_d(\mathbf{x}_d)}, \quad (7)$$



**Fig. 1.** The response  $\langle \delta x \rangle$  of the Lorenz 95 system to a forcing of  $\delta f_{18} = 1.0$  with the location of  $\delta f_{18}$  indicated by the vertical dashed line. To compare with typical fluctuations during an integration, the standard deviation of a single element of  $\mathbf{x}$  for  $\delta f = 0$  is estimated as  $\sim 3.6$  for  $a = 0$ . The number of elements in the state vector for (a) and (c) is  $d = 36$  and for (b) and (d) is  $d = 72$ . Plots (a) and (b) correspond to the deterministic system with  $a = 0$  and (c) and (d) correspond to the same system with additional stochastic noise,  $a = 2$ . The integrations are sufficiently long to obtain an accurate response, (2 standard deviations divided by the square root of the number of integrations is smaller than graphical accuracy), which is the same as that reported by Abramov and Majda (2007); see their Fig. 2. The difference between the first 36 points of (a) and (b) or (c) and (d) is not significantly different from zero with this set of integrations. Note that the response to linearly forcing multiple grid points simultaneously is simply the sum of the responses to the decomposed single grid point forcings.

where

$$\rho_{d,i}(\mathbf{x}_d, \mathbf{x}_i) = \int \rho(\mathbf{x}) dx_k,$$

$$\rho_{k,d}(x_k, \mathbf{x}_d) = \int \rho(\mathbf{x}) dx_i$$

and

$$\rho_d(\mathbf{x}_d) = \int \int \rho(\mathbf{x}) d\mathbf{x}_i dx_k.$$

Substituting (7) into (1) and restricting forcing to the direction of  $x_k$ , we obtain

$$\langle \delta \mathbf{x}(t) \rangle \approx - \int_0^t \left\langle \mathbf{x}(\tau) \frac{\partial \rho_{k,d}(x_k(0), \mathbf{x}_d(0))}{\partial x_k} \frac{1}{\rho_{k,d}(x_k(0), \mathbf{x}_d(0))} \right\rangle \delta f_k(t - \tau) d\tau. \quad (8)$$

Note that the response in the direction of  $\mathbf{x}_i$  due to  $\delta f_k$  is zero:

$$\langle \delta \mathbf{x}_i(t) \rangle \approx 0. \quad (9)$$

The PDF  $\rho_{k,d}$  in (8) is a function only of  $x_k$  and  $\mathbf{x}_d$ . Dependence upon  $\mathbf{x}_i$  has been eliminated. If we now assume Gaussianity, we can obtain (2) for the system that involves smaller covariance matrices obtained using only  $x_k$  and  $\mathbf{x}_d$ . Now suppose we are interested in the response to a forcing  $\delta f$  that is not restricted to  $\delta f_k$ , we may use the fact that we are in the linear regime to compute the linear response to forcing a single element  $\delta f_j$  for each  $j$  and sum the results.

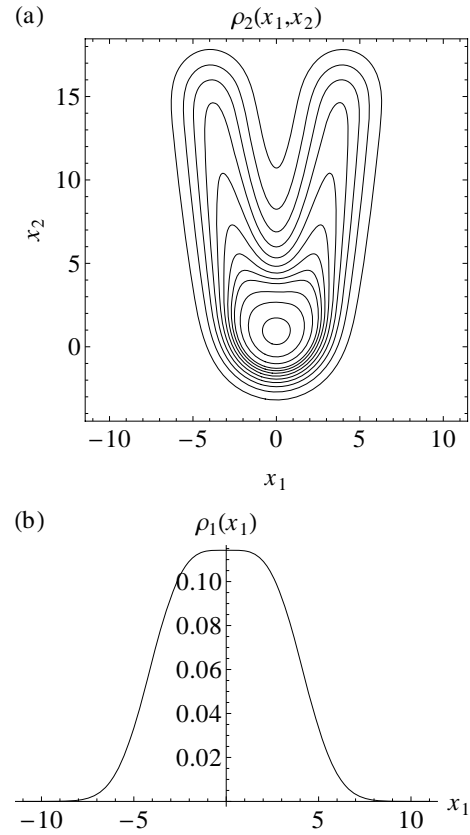
## 2 A truncation algorithm

An algorithm that finds the appropriate truncation of phase space according to (8) may now be developed. We define

$$\mathbf{x}_r = \begin{pmatrix} x_k \\ \mathbf{x}_d \end{pmatrix}.$$

Now we do not know in advance which elements of  $\mathbf{x}$  to include in  $\mathbf{x}_d$  and therefore truncate iteratively towards numerical convergence of our estimate of the response  $\langle \delta \mathbf{x}(t) \rangle$ . Our starting point will be the maximum truncation possible, i.e. setting  $\mathbf{x}_r$  to have only the element to which the forcing is applied,  $\mathbf{x}_r(t) = (x_k(t))$ . We then obtain an estimate for  $\langle \delta \mathbf{x}(t) \rangle$  using a non-parametric FDT algorithm (A1) which we denote by  $\widehat{\delta \mathbf{x}}_r(t)$ ; see Appendix A for details. Next we try a truncation that in addition includes another grid point and see if the estimated response of grid point  $k$  changes. If it does change, then using two grid points gives us a more accurate estimate of  $\rho(\mathbf{x})$  present in (1) and approximated in (8), see Fig. 2. On the other hand if it does not change then perhaps point  $k$  and the tested point are independent. We repeat this process, checking point  $k$  with all other grid points and pick the grid point that contributed the largest change in  $\widehat{\delta \mathbf{x}}_r(t)$  to include in our truncated state vector  $\mathbf{x}_r$ . For example, if grid point  $k+3$  caused the estimated response to change the most, then for the next round of testing we choose

$$\mathbf{x}_r = \begin{pmatrix} x_k \\ x_{k+3} \end{pmatrix}.$$



**Fig. 2.** The contours of a two dimensional marginal PDF  $\rho_2(x_1, x_2)$  (a). Integrating  $\rho_2(x_1, x_2)$  in the  $x_2$  direction gives the one dimensional marginal PDF  $\rho_1(x_1)$  (b). In this case the plots illustrate that the quantities  $\partial \rho_2 / \partial x_1$  and  $\partial \rho_1 / \partial x_1$ , which may be used in (8) are different.

The entire process is then repeated to find the next points to add to  $\mathbf{x}_r$ , this time looking at the largest change in the response at points  $k$  and  $k+3$ , (estimated using the normalised root mean squared difference, RMS)

$$\text{RMS distance} = \sqrt{\frac{\sum_{j=1}^{d'} (\widehat{\delta x}_{r,j}(t) - \widehat{\delta y}_{r,j}(t))^2}{\sum_{j=1}^{d'} (\widehat{\delta y}_{r,j}(t))^2}},$$

where  $\widehat{\delta y}_{r,j}(t)$  is the response estimated from the previous round of tests and  $d'$  indicates the current dimensionality of  $\mathbf{x}_r$ . We continue to successively add more and more grid points to  $\mathbf{x}_r$  until the estimated response no longer changes. At this point we assume that we have a good approximation of the optimal truncation so that (8) is a good approximation of (1).

One problem with our truncation algorithm is that we are comparing a response estimated using  $d'$  grid points with a response estimated using  $d'+1$  grid points. Our estimate  $\widehat{\delta \mathbf{x}}_r(t)$  is biased by an amount depending somehow upon the dimensionality of phase space; see Eq. (21) of Cooper and Haynes (2011). We must account for this or otherwise any

change in the estimated response may be due to bias rather than a genuine interaction of variables. We do so by adding a variable of Gaussian distributed random numbers (with approximately the same variance as that of  $x_k$ ) to the truncated state vector  $\mathbf{x}_d$ . So in the example above we would at some point compare the response given the vectors

$$\begin{pmatrix} x_k(t) \\ x_G(t) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x_k(t) \\ x_{k+3}(t) \end{pmatrix},$$

where  $x_G$  represents the Gaussian random numbers.

### 2.1 Further truncation

It is necessary to estimate a subset of the domain to use as a starting point for the above calculation. This is because it is computationally expensive to apply the FDT to every single grid point, several times, when the number of grid points is large. One choice of region in an atmospheric context could be the set  $\mathcal{D}$  of all points where the integral of absolute positive lagged correlations is greater than some threshold  $c_{\text{abs,min}}$ , i.e. the set

$$\mathcal{D} = \{ \mathbf{x} : c_{\text{abs},j}(\mathbf{x}) > c_{\text{abs,min}} \},$$

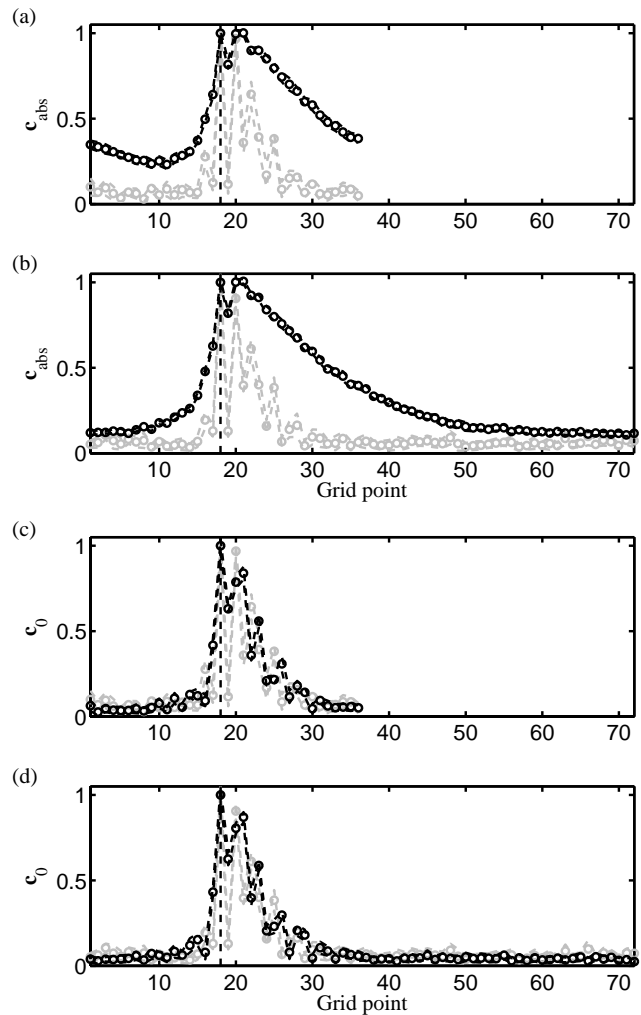
for all  $j = 1, 2, 3, \dots, d$ , where

$$c_{\text{abs}}(\mathbf{x}) = \int_0^\infty | \langle \mathbf{x}(\tau) x_k(0) \rangle | d\tau. \tag{10}$$

Then the elements of  $\mathbf{c}_{\text{abs}}$  that are greater than  $c_{\text{abs,min}}$  would be the only points considered when looking for an appropriate truncation. The zero in the lower limit indicates that we are looking for the response to a forcing, rather than all variables that have some relation to  $x_k$ . For the Lorenz 95 system,  $\mathbf{c}_{\text{abs}}$  is plotted in Fig. 3 which shows that the (normalised) absolute value of the response lies under the (normalised) values of  $\mathbf{c}_{\text{abs}}$ . Note that in any suitably random system, elements of  $\mathbf{c}_{\text{abs}}$  cannot be zero because we take the absolute value of the correlation estimate which, in turn, is subject to statistical uncertainty. Therefore, if we are able to estimate the contribution of this uncertainty, for example from two independent integrations, we can pick an appropriate value for our threshold  $c_{\text{abs,min}}$ . For the Lorenz 95 integrations considered here the normalised value of  $c_{\text{abs,min}} \sim 0.12$  and from Fig. 3 we would be able to eliminate about one third of the points from the 72-element system and none of the points from the 36-element system.

Equation (10) specifies a rather conservative estimate of the region to consider and as such may not solve the problem given limited computational resources. For this investigation, rather than  $\mathbf{c}_{\text{abs}}$ , a more useful truncation region is based upon  $c_0$  which is given by

$$c_0 = \left| \int_0^\infty \langle \mathbf{x}(\tau) x_k(0) \rangle d\tau \right| \tag{11}$$



**Fig. 3.** Evaluation of (10), plots (a) and (b), and (11), plots (c) and (d), for the deterministic ( $a = 0$ ) Lorenz 95 system is indicated by the black lines. The grey lines indicate the absolute value of the response,  $|\langle \delta x \rangle|$ . The number of elements in the state vector for (a) and (c) is  $d = 36$  and for (b) and (d) is  $d = 72$ . The vertical scale has been normalised so that  $c_{\text{abs},k} = c_{0,k} = |\delta x_k| = 1$  where  $k = 18$ . The location of the forcing  $\delta f_{18}$  is indicated by the vertical dashed line and statistical error, 2 standard deviations divided by the square root of the number of integrations, is smaller than graphical accuracy.

with a threshold  $c_{0,\text{min}}$  estimated in the same way. This is in fact the absolute value of row  $k$  of the matrix  $\int_0^\infty \mathbf{C}(\tau) d\tau$  that appears in the expression for the Gaussian FDT (2). It is related to the spectral density via Parseval's theorem. The motivation for not simply using (2) to estimate our truncation region is that given a large state vector and limited data, which is typical in a climate context, it can be difficult or even impossible to accurately estimate  $\mathbf{C}(0)^{-1}$ . In any case where (11) is zero, (2) gives zero response anyway. Figure 3 shows that in our case the region estimated by (11) is a much tighter fit to the response region and the number of points to

consider is reduced considerably. However, to find an appropriate threshold value  $c_{0,\min}$ , accurate evaluation of (11) is necessary. This may not always be possible and is not such a problem for (10). Another problem with (11) is that it is only a good approximation for systems with close to Gaussian statistics. If a system is extremely non-Gaussian, (11) may be a poor approximation and we must resort to using (10) instead.

## 2.2 Application of the algorithm

By invoking the LRA (4), we now have an algorithm that we can actually apply to high dimensional data sets provided the response to a localised forcing is sufficiently local. Using Fig. 3c and d, we consider only the region between points  $k-7$  to  $k+16$  inclusive and apply the FDT for each truncation. The RMS difference between the true response at all grid points, (found by integrating (3)), and the response predicted by our estimation of (8) and (9) is plotted as a function of the number of elements included in  $\mathbf{x}_T$  in Fig. 4 (black circles).

Figure 4 demonstrates that for the Lorenz 95 system, predictions of the response using our estimator of (8) have some skill. This figure illustrates that for a truncation to 10 grid points of the 36-element system, with the noise term  $a=0$ , the RMS error squared (i.e. the mean squared difference), is made up by variance due to a finite run length  $\sim 5\%$ , missing responses due to truncation  $\sim 29\%$  and bias in the FDT algorithm  $\sim 66\%$ . For the 36- and 72-element systems, the results have similarities which indicates that for the Lorenz 95 system we may be close to the point where the predictive skill of our algorithm is independent of the dimensionality of the full state vector  $\mathbf{x}$ . Similar results (not shown here) were obtained with an 18-element system.

An interesting point to make about Fig. 4 is that there appears to be an optimal truncation at around  $d'=9$ . Unfortunately, we did not predict the precise location of this optimum in advance. Initially the skill increases non-monotonically as more grid points are added before reaching the optimum truncation. The RMS error then increases slowly up to its value when all grid points are included (without any truncation). We can hypothesise as to the reasons behind this behaviour. Firstly, we have no reason to believe that high skill at predicting the response of the first point on its own would translate to other dynamical systems in some general case. Secondly, the reduction in RMS error with more than four grid points could be due to the FDT calculation including a better approximation of the underlying PDF in the manner described in Fig. 2. Thirdly, the increase in RMS error beyond the optimal truncation is at least partly due to the bias present in our estimator of (8). An intuitive explanation of the cause of this bias could be simply that complex shapes present in a PDF become harder and harder to approximate accurately as their dimensionality increases. In our case we are approximating the underlying PDF by a convolution of the data (from integrations) with a kernel function; see Ap-

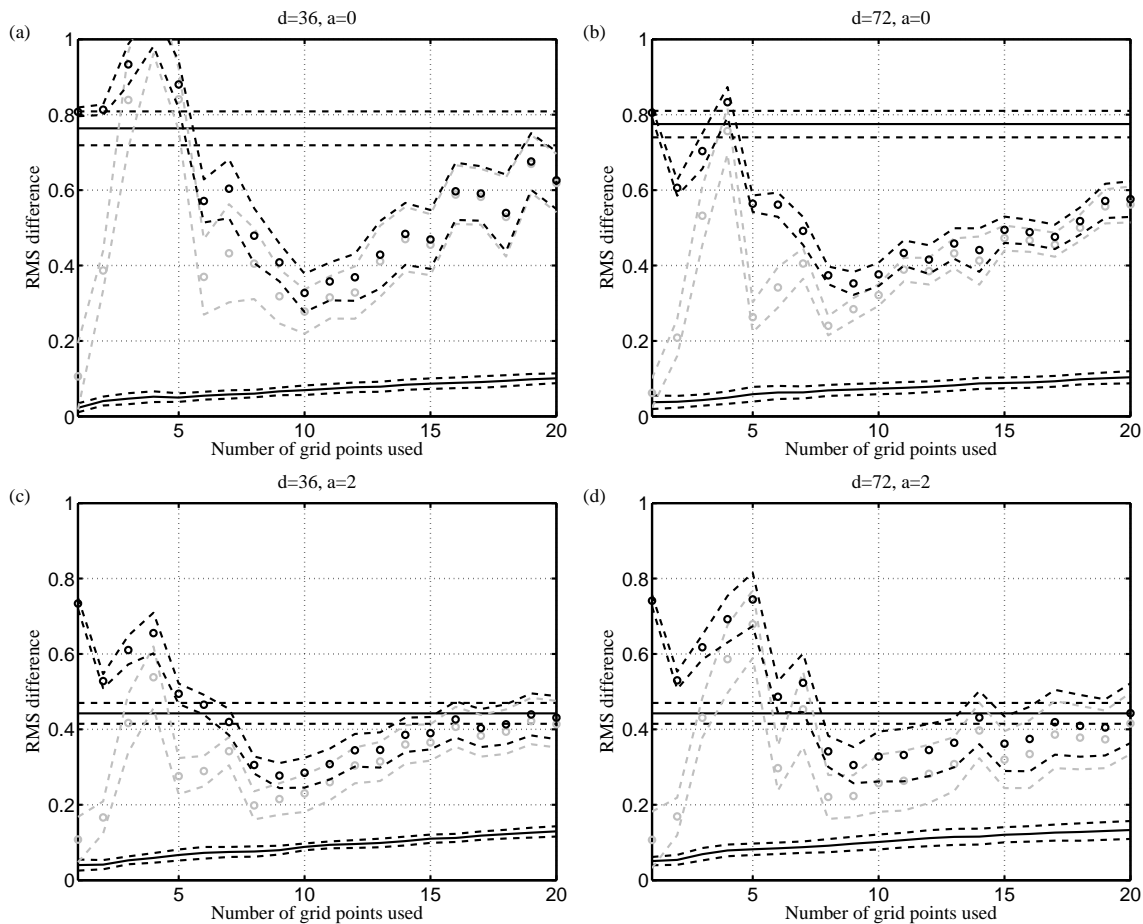
pendix A. As the dimensionality  $d'$  of phase space increases, the integral over all  $\mathbf{x}_T$  of the kernel function must remain constant. Fixing an acceptable level of variance in the predictions of the FDT and increasing the dimensionality has the effect of requiring a broader kernel to approximate the PDF with. Eventually, the kernel is so broad that one may as well approximate the convolution of the data and the kernel with a single Gaussian blob with non-isotropic covariance. If this is true for our system then the response predicted by the Gaussian form of the FDT (2) and Eq. (1) using our estimator, should be approximately the same; see Fig. 5.

The fact that our estimator approximates a smoothing convolution of the PDF leads us to suspect that it has skill when applied to stochastic systems where the white noise term leads to a genuinely smooth underlying PDF. Although it may appear slight, the FDT (A1) has more skill for the stochastic case where  $a=2$  than the deterministic case where  $a=0$ ; see Fig. 4. Given sufficient smoothing (by increasing the value of  $a$ ), we eventually get to the situation where the non-Gaussian aspects of the underlying PDF are small corrections to a  $d'$  dimensional Gaussian. In this case the Gaussian approximation (2) is a good one.

## 3 Conclusions

If we wish to apply the fluctuation-dissipation theorem (FDT) to predict the response to forcing in general circulation models, reanalysis data sets, and measurements of the climate system, truncation of the huge state vector present in these systems seems necessary. We have argued that truncation of the state vector to include only the locality of the forcing (according to variations in predictions of the FDT) may be appropriate to some chaotic high dimensional systems and have tested this hypothesis using a 36- and 72-element Lorenz 95 system. The practical implications of this localised response approximation (LRA) are, for example, that if we consider the atmospheric response to a heating over one small part of the planet, perhaps there are regions that contribute very little to the response. If these regions exist then they may be safely ignored in a FDT calculation.

One method of truncation that has previously been applied (see for example Gritsun and Branstator, 2007) is to use only a given number of leading EOFs. Each EOF is described by a fixed global pattern and the associated principal component time series describes its evolution in time. For systems where the response to a local forcing is spatially localised, a localised truncation seems preferable. To illustrate this we note that the responses of the 36- and 72-element Lorenz 95 systems to a good level of accuracy are the same, i.e. the extra points in the 72-element system can be ignored. Truncation in EOF space does not take advantage of this fact but localised truncations do. In addition, if the response to a local forcing is sufficiently localised, truncation to these points allows us to apply the FDT to data sets of any dimensionality.

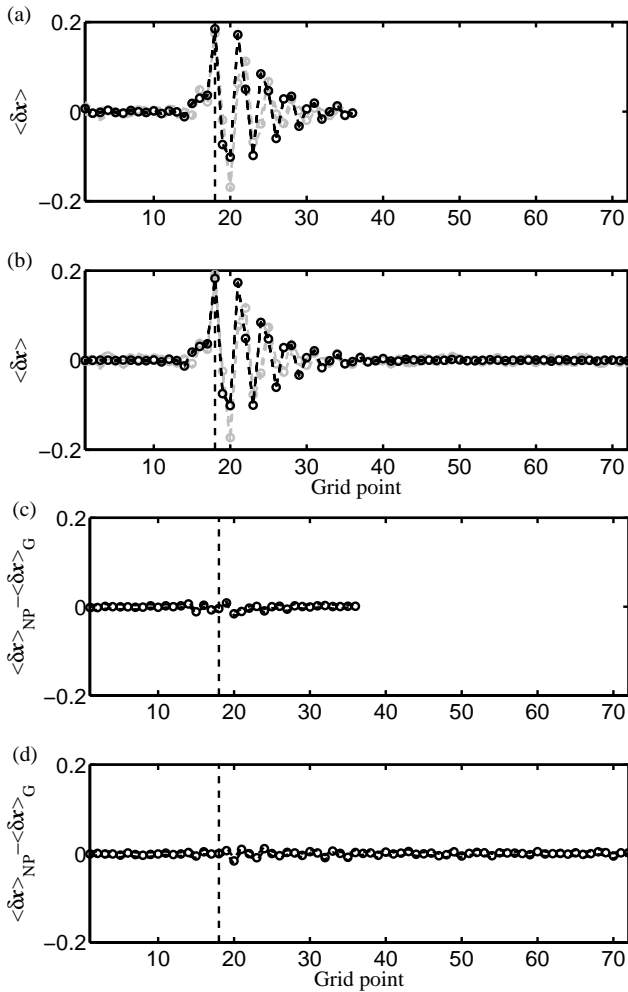


**Fig. 4.** RMS difference between the response of the Lorenz 95 system estimated by integrating (3) and using the non-parametric FDT (A1) where the number of elements in the full state vector for (a) and (c) is  $d = 36$  and for (b) and (d) is  $d = 72$ . Plots (a) and (b) correspond to the deterministic system with  $a = 0$  and plots (c) and (d) correspond to the same system with additional stochastic noise,  $a = 2$ . The black circles indicate the RMS difference of the entire state vector using (8) and (9). The grey circles indicate the RMS difference of only the elements to which the calculation is applied using (8). The solid line at the bottom of each plot indicates the minimum possible RMS difference of the grey circles with this data set, i.e. the difference between the responses estimated using two independent sets of data found by integrating (3). The straight horizontal line is the RMS difference between integrating (3) and the response estimated using the Gaussian form of the FDT (2) for the 36-element system; see Fig. (5) for a plot of this response. The dashed lines, indicating uncertainty, represent two standard deviations divided by the square root of the number of independent integrations. An estimate with zero skill, (given by (9) with the entire state vector being independent of the forcing), has a RMS difference of one using this normalisation. It is possible to do worse than this by overestimating the response by a factor of greater than 2.

The FDT applied here gives only the linear component of the response of the non-linear system considered. This is a good approximation for the full response if the forcing is sufficiently small. Being linear, the sum of the responses estimated separately for two different forcings is identical to the single response estimated for the sum of these two forcings. We may therefore estimate the entire set of possible linear responses by appropriate summing of the responses to forcing each grid point individually. Similarly, quantities such as the response in the mean, or any other combination of several variables, may be estimated from the appropriate combination of responses to separate local forcings. Forcings sufficiently large to make the response strongly non-linear are in

general not susceptible to any simple linear decomposition into localised forcings. In addition (1) only applies to the linear component of any response calculation and large forcings require a generalisation (e.g. Boffetta et al., 2003). In some situations, even with simple systems similar to that considered here, the response may be entirely non-linear (Lacorata and Vulpiani, 2007).

The non-parametric FDT algorithm requires that we use an approximation of a PDF using a convolution of the data and a kernel function. The kernel function fills the gaps between data points. Given more data these gaps are smaller and the kernel function can also be narrower leading to a better PDF approximation. However, increasing dimensionality



**Fig. 5.** Response of the deterministic ( $a = 0$ ) Lorenz 95 system estimated using the Gaussian FDT (2), plots (a) and (b), and the difference between the Gaussian FDT (2) and the non-parametric FDT (A1) as an estimator of (1) without truncation, plots (c) and (d). The number of elements in the state vector for (a) and (c) is  $d = 36$  and for (b) and (d) is  $d = 72$ . The grey lines indicate the true response plotted in Fig. 1 and the RMS difference between the grey and black points is  $\sim 75\%$  as indicated by the uppermost straight horizontal lines in Fig. 4. The location of the forcing  $\delta f_{18}$  is indicated by the vertical dashed line and statistical error, 2 standard deviations divided by the square root of the number of integrations, is smaller than graphical accuracy.

necessarily increases the width of the kernel function relative to the distance between data points leading to a worse PDF approximation. In a high dimensional space, the kernel functions significantly overlap and although we have not shown it in any analytical way, the non-parametric FDT seems to approach the Gaussian FDT (Fig. 5). If the response to a local forcing is spread over a large number of grid points (and an accurate transformation to a smaller number of grid points is not available), then the dimensionality of phase space in which the non-parametric FDT must approximate a PDF is

also large. If the PDF has a complex shape then its accurate approximation is extremely difficult and is likely to require more data than is possible to obtain. Deterministic systems may even have a fractal PDF that makes use of (1) difficult to justify. Other expressions such as the linear response relation of Ruelle (1998) may have to be considered.

In addition to using truncation, there are several improvements that could be made to the non-parametric FDT algorithm. It may be possible to obtain insight into its bias by considering different kernels with the same data set, or to reduce the bias via adaptive density estimation, see Silverman (1986) and Cooper and Haynes (2011). If restrictions such as conservation laws can be placed upon a system's PDF, perhaps by analytical consideration of the underlying equations, it may also be easier to approximate. Another approach may be to adopt the tangent linear algorithm of Eyink et al. (2004) with the “blended response” modification of Abramov and Majda (2007) which replaces estimation of the PDF by estimation (or calculation if the underlying equations are available) of a tangent linear system.

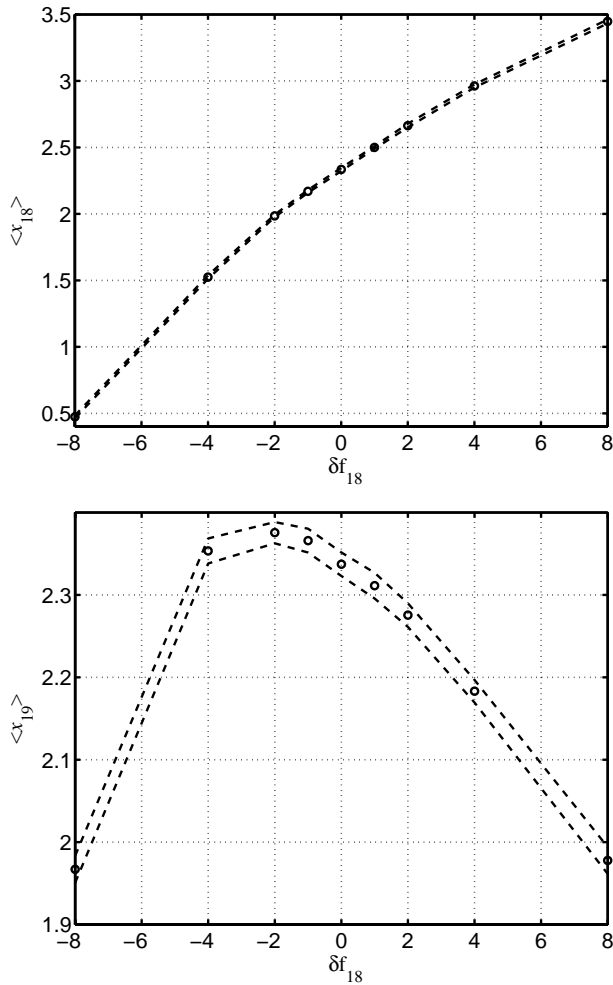
## Appendix A

### Experiment specification

We test the FDT (1) using the stochastic Lorenz 95 system introduced above where we choose the constant  $f = 8$ . It is approximated by applying a fourth-order Runge–Kutta method (Press et al., 2002) to the deterministic part of (3) and the method of Rümelin (Rümelin, 1982) to the stochastic part using the Mersenne Twister algorithm (Matsumoto and Nishimura, 1998) for random numbers. We choose state vectors of size  $d = 36$  and  $d = 72$  and starting with random initial conditions we use a time step of  $\Delta t = 10^{-3}$  and after discarding  $10^5$  time steps iterate from  $t = 0$  to  $t = 10^6$ . This may be compared with  $\sim 1.74$ , the value of the largest Lyapunov exponent for the deterministic system (Abramov and Majda, 2007). We perform 10 identically specified integrations with different initial conditions for each experiment in order to get a measure of the statistical uncertainty. Equation (8) is only valid in the linear regime, i.e. for a sufficiently small forcing. In order to discover the range over which the linear approximation is valid, we perform several integrations of (3) with  $-8 \leq \delta f_k \leq 8$  and  $\delta f_j = 0$  for  $j \neq k$  and  $k = 18$ . The average value of elements 18 and 19 of the state vector of the  $d = 36$ ,  $a = 0$  integrations averaged over  $10^{10}$  time steps is plotted in Fig. A1. For a forcing of  $\delta f_{18} = \pm 1$  the non-linear component of the response is small in comparison with the size of the total response. A similar range and accuracy of linearity is observed for all elements of the state vector, and if the integrations are carried out with  $d = 72$ .

Taking care that the relevant integrals have sufficiently converged, (10) is used to approximate infinity in the upper limit of the integral, the Gaussian FDT (2) is evaluated using





**Fig. A1.** The mean value of state vector element  $x_{18}$  and  $x_{19}$  of a 36-element Lorenz 95 system as a function of forcing applied to  $x_{18}$  to illustrate the range of linearity of the response. The dashed lines indicate two standard deviations divided by the square root of the number of independent integrations (of which there are 10). Plots of the response of other vector elements or of the elements of a 72-element or stochastic Lorenz 95 system yield similar ranges of linearity with different gradients near the origin.

the standard unbiased estimator for the covariance  $\mathbf{C}(\tau)$ . Following Cooper and Haynes (2011), we approximate (8) by the non-parametric method

$$\hat{\mathbf{A}}(s\Delta t) = \sum_{j=1}^m X_{j+s} \left[ \frac{\sum_{i=1}^n (X_j - X_i)^T E(X_j; X_i, h)}{h^2 \sum_{i=1}^n E(X_j; X_i, h)} \right]$$

and

$$\hat{\delta \mathbf{x}}(h) = \sum_{s=0}^r \mu_s \hat{\mathbf{A}}(0)^{-1} \hat{\mathbf{A}}(s\Delta t) \delta f, \quad (\text{A1})$$

where  $\hat{\delta \mathbf{x}}(h)$  is the estimated response as a function of the free parameter  $h$  and  $\mathbf{X}$  is the appropriate segment of the

state vector recorded at a particular time. Here a unit of the indices  $i$  and  $j$  correspond to a time between recordings of  $\mathbf{X}$  of 10 and a unit of the index  $s$  corresponds to a time between recordings of  $10^{-3}$ . So for example if  $i = 5$ ,  $j = 7$  and  $s = 137$  then the values of  $t$  that  $\mathbf{X}_i$ ,  $\mathbf{X}_j$  and  $\mathbf{X}_{j+s}$  correspond to are 50, 70 and 70.137, respectively. To evaluate (A1) we use  $r = 10^4$  (representing 10 time units to approximate infinity as the upper limit of the integral in (8)),  $m = 10^5$ ,  $n = 10^5$ ,  $\mu_0 = 0.5\Delta t$  and  $\mu_s = \Delta t$  for  $s > 0$  with our choice being motivated by a trade off between accuracy and available computational power; see Cooper and Haynes (2011). In Cooper and Haynes (2011), the function  $E$  was represented by  $N$ , a  $d$  dimensional Gaussian with isotropic covariance  $h$ . Here for computational efficiency we use the multi-dimensional Epanechnikov kernel (Silverman, 1986) without normalisation

$$E(\mathbf{x}; \mathbf{y}, h) = \begin{cases} h^2 - (\mathbf{x} - \mathbf{y})^T (\mathbf{x} - \mathbf{y}), & |\mathbf{x} - \mathbf{y}| < h \\ 0, & \text{otherwise.} \end{cases} \quad (\text{A2})$$

To estimate the response using (A1), a value of the free parameter  $h$  must be chosen. We choose the smallest guess of  $h$  that corresponds to the standard deviation of the estimated response of element  $x_{18}$  being no larger than five percent of the estimate of its mean. Our choice of  $h$  is not well optimised because this is computationally expensive and the benefits of improving upon a simple guess appear small; see Cooper and Haynes (2011) for a discussion.

*Acknowledgements.* Thanks go to H. Salman for useful discussions. This research was supported by the Natural Environment Research Council Grant NE/G003122/1 to J. G. Esler and by the EU funded SHIVA project.

Edited by: O. Talagrand

Reviewed by: two anonymous referees

## References

- Abramov, R. V. and Majda, A. J.: Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems, *Nonlinearity*, 20, 2793–2821, doi:10.1088/0951-7715/20/12/004, 2007.
- Boffetta, G., Lacorata, G., Musacchio, S., and Vulpiani, A.: Relaxation of finite perturbations: Beyond the fluctuation-response relation, *Chaos*, 13, 806–811, doi:10.1063/1.1579643, 2003.
- Cooper, F. C. and Haynes, P. H.: Climate Sensitivity via a Non-parametric Fluctuation-Dissipation Theorem, *J. Atmos. Sci.*, 68, 937–953, doi:10.1175/2010JAS3633.1, 2011.
- Cooper, F. C. and Haynes, P. H.: Assessment of the fluctuation-dissipation theorem as an estimator of the response of the tropospheric response to forcing, *Q. J. Roy. Meteor. Soc.*, submitted, 2013.

- Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A., and Wood, N.: A new dynamical core for the Met Office's global and regional modelling of the atmosphere, *Q. J. Roy. Meteor. Soc.*, 131, 1759–1782, doi:10.1256/qj.04.101, 2005.
- Eyink, G. L., Haine, T. W. N., and Lea, D. J.: Ruelle's linear response formula, ensemble adjoint schemes and Lévy flights, *Nonlinearity*, 17, 1867–1889, doi:10.1088/0951-7715/17/5/016, 2004.
- Gritsun, A. and Branstator, G.: Climate Response Using a Three-Dimensional Operator Based on the Fluctuation-Dissipation Theorem, *J. Atmos. Sci.*, 64, 2558–2575, doi:10.1175/JAS3943.1, 2007.
- Joliffe, I. T.: *Principal Component Analysis*, Springer Series in Statistics, 2nd Edn., 2002.
- Lacorata, G. and Vulpiani, A.: Fluctuation-Response Relation and modeling in systems with fast and slow dynamics, *Nonlin. Processes Geophys.*, 14, 681–694, doi:10.5194/npg-14-681-2007, 2007.
- Lorenz, E. N.: Predictability: a problem partly solved, *Proceedings from the ECMWF Seminar on Predictability*, volume Vol. I, ECMWF, Reading, UK, 1995.
- Lucarini, V. and Sarno, S.: A statistical mechanical approach for the computation of the climatic response to general forcings, *Nonlin. Processes Geophys.*, 18, 7–28, doi:10.5194/npg-18-7-2011, 2011.
- Majda, A. J., Gershgorin, B., and Yuan, Y.: Low-Frequency Climate Response and Fluctuation-Dissipation Theorems: Theory and Practice, *J. Atmos. Sci.*, 67, 1186–1201, doi:10.1175/2009JAS3264.1, 2010.
- Marconi, U. M. B., Puglisi, A., Rondoni, L., and Vulpiani, A.: Fluctuation-dissipation: Response theory in statistical physics, *Physics Reports*, 461, 111–195, doi:10.1016/j.physrep.2008.02.002, 2008.
- Matsumoto, M. and Nishimura, T.: Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator, *ACM T. Model. Comput. S.*, 8, 3–30, 1998.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: *Numerical Recipes in C++: The Art of Scientific Computing*, Cambridge University Press, UK, 2002.
- Ring, M. and Plumb, R. A.: The Response of a Simplified GCM to Axisymmetric Forcings: Applicability of the Fluctuation-Dissipation Theorem, *J. Atmos. Sci.*, 65, 3880–3898, doi:10.1175/2008JAS2773.1, 2008.
- Ruelle, D.: General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium, *Phys. Lett. A*, 245, 220–224, doi:10.1016/S0375-9601(98)00419-8, 1998.
- Rümelin, W.: Numerical Treatment of Stochastic Differential Equations, *SIAM J. Numer. Anal.*, 19, 604–613, 1982.
- Silverman, B. W.: *Density Estimation for Statistics and Data Analysis*, *Mg. Stat. Pro.*, Vol. 26, Chapman and Hall, 1986.