Nonlinear Processes
in Geophysics

Empirical correction techniques: analysis and applications to chaotically driven low-order atmospheric models

I. Trpevski[1], **L. Basnarkov**[1,2], **D. Smilkov**[1,3], and **L. Kocarev**[1,2,4]

[1]Macedonian Academy of Sciences and Arts, Skopje, Macedonia
[2]SS. Cyril and Methodius University, Faculty of Computer Science and Engineering, P.O. Box 393, Skopje, Macedonia
[3]The MIT Media Lab, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA
[4]BioCircuits Institute, University of California, San Diego, California, USA

Correspondence to: I. Trpevski (itrpevski@manu.edu.mk)

Abstract. Contemporary tools for reducing model error in weather and climate forecasting models include empirical correction techniques. In this paper we explore the use of such techniques on low-order atmospheric models. We first present an iterative linear regression method for model correction that works efficiently when the reference truth is sampled at large time intervals, which is typical for real world applications. Furthermore we investigate two recently proposed empirical correction techniques on Lorenz models with constant forcing while the reference truth is given by a Lorenz system driven with chaotic forcing. Both methods indicate that the largest increase in predictability comes from correction terms that are close to the average value of the chaotic forcing.

1 Introduction

In forecasting the state of the atmosphere, the occurrence of two fundamental types of errors is inevitable (Palmer, 2000). The first type, often called *internal error growth*, results from the amplification of initial condition uncertainties due to atmospheric instabilities (Kalnay, 2002). The present tools for tackling internal error growth include data assimilation techniques and ensemble forecasting (Anderson, 2001; Hunt et al., 2004; Merkova et al., 2011). The second type is called external or model error and comes from the fact that atmosphere has larger complexity and resolution than its representations by models (Orrell et al., 2001; Judd and Smith, 2004). With mathematical language, the atmosphere has (many) more degrees of freedom, or variables, and the

equations of its evolution (if they exist) have more complex structure than those of the atmospheric models. Furthermore, subgrid-scale atmospheric processes such as cloud formation are not resolved by the models. Instead, they are approximated by different parametrization schemes, the choice of which mainly depends on the preference of the investigators at a weather forecasting center. The uncertainties coming from model errors limit our ability to make useful predictions with any individual model, and current tools that account for model error are typically called empirical correction techniques (Leith, 1978; DelSole and Hou, 1999; Danforth et al., 2007; Allgaier et al., 2012; Basnarkov and Kocarev, 2012).

The simplest form of model error occurs in the perfect model scenario where the model and the truth belong to the same class. The model structure is correct but the true parameter values are unknown, and the goal of eliminating model error is the same as that of parameter estimation (Smith et al., 2010). One of the difficulties in reducing model error in this setting comes from infrequent observation of the truth. For tackling this problem we provide a new iterative linear regression technique that works efficiently even when the sampling interval between two successive observations of the truth is very large.

When the model is structurally different from the truth, one can modify the dynamical equations of the model based on the statistics of the forecast errors between the model and the truth. The question arises as to how different methods compare to each other in this approach, and this paper attempts to address it by investigating two promising techniques on a chaotically driven toy climate model. Both of these methods aim to optimize the average model tendency

error (i.e., the difference between the true change of the variables and the model tendency over a time series of observations) in terms of some linear parameters. The first method was originally proposed by Leith (1978) with the aim to correct the model tendency equations by adding a state-independent or model bias term, and a state-dependent linear operator. A version of Leith's correction method has been applied successfully on a low-dimensional atmospheric model with $O(10^2)$ degrees of freedom (DelSole and Hou, 1999), while a modification of the technique has given promising results on atmospheric models with as much as $O(10^5)$ degrees of freedom (Danforth et al., 2007). Recently, it has been shown for a toy climate system that the increase in the average forecasting accuracy comes at the cost of substantial qualitative differences between the dynamics of a corrected model and its reference truth system (Allgaier et al., 2012).

The second method compared in this paper stems from a more recent idea that one can combine the advantages of different imperfect models to produce an improved forecast of the truth (Wiegerinck et al., 2011; van den Berge et al., 2011). In particular this approach constructs a model whose tendencies are given by a weighted combination of the tendencies of the imperfect models. Akin to the Leith's method, the weights are determined so as to minimize the average model tendency error. This method has not been tested as extensively as Leith's method on larger models, but a recent study shows that it offers improved forecast on a low-dimensional Lorenz '96 system (Basnarkov and Kocarev, 2012).

The experimental test bed on which we test these two techniques is given by a chaotically driven Lorenz 63 system (Lorenz, 1963) – essentially capturing the fact that at a more fundamental level the atmosphere is an open system subject to a variable forcing in time. The external forcing is given by a chaotic drive signal obtained from the variables of another Lorenz 63. We assume that we cannot resolve the driving process and model it using a standard parametrization scheme, namely by adding constants in the Lorenz 63 model.

We note that this type of on-line empirical correction methods considered in this paper differs from a posteriori correction methods (offline), such as subtracting the bias from the forecast itself, in that they modify the dynamical equations of the model. Danforth and Kalnay (2008) recently pointed out that online correction reduces not only the growth of the bias but also the nonlinear growth of nonconstant (state-dependent and random) forecast errors during the model integration.

The rest of the paper proceeds as follows. In Sect. 2 we first look at the new method for estimating the parameters of a perfect model in the case of infrequent observational data. We then describe the two empirical correction techniques in Sect. 3. Their forecasting properties on the chaotically driven Lorenz 63 system are analyzed in Sect. 4. We provide a discussion and conclusion in Sect. 5.

2   An iterative linear regression method for model correction

Forecast errors originate from both initial condition errors and model errors, and it is often assumed that one dominates over the other, although in general this need not be true. Shadowing time (Orrell, 2001) for example depends on the interplay between the two types of errors. Here we assume that one can get a good *analysis* of the true state in terms of the model variables with current state-of-the-art data assimilation methods so that model error dominates short-term forecast errors. This allows us to focus on improving the model by reducing its model error. In the following we denote the analysis of the truth with x_a and the model state vector with x. Also, assume that the model tendency is given by

$$\dot{x} = T(x, p),\qquad(1)$$

where the parameters appear linearly in the equations.

We present an iterative linear regression (ILR) method that can effectively estimate all the parameters appearing in the model equations (1) when the sampling interval τ between two successive observations of the truth is large. The efficiency of the technique is demonstrated in the perfect model scenario where "true" or optimal values of the parameters of a system exist but are unknown. In the imperfect model scenario, one needs to define a goal (such as best forecast lead time) in terms of which one hopes to find the best parameter values (Du and Smith, 2012). Standard linear regression techniques work by minimizing the Mean Square Tendency Error (MSTE) with respect to the parameters p:

$$E_1(p) = \langle |(\tilde{\dot{x}}_a - T(x_a, p))|^2 \rangle,\qquad(2)$$

where $\langle \rangle$ denotes time average over N observations $\{x_a(t_0), x_a(t_0 + \tau), \ldots, x_a(t_0 + (N-1)\tau)\}$ sampled with frequency τ. The estimated tendency of the truth $\tilde{\dot{x}}_a$ at time t can be obtained using standard interpolation techniques or finite differences from the observations.

We now modify method (2) to be used as a model correction method:

$$E_2(p) = \langle |(\tilde{\dot{x}}_a - \tilde{\dot{x}} - c(x_a, p))|^2 \rangle,\qquad(3)$$

where $c(x, p)$ is the correction of the initial model (1) and $\tilde{\dot{x}}(t)$ is estimated from time series observation $\{x(t_0), x(t_0 + h) \ldots\}$ of frequency h obtained by integrating the initial model Eq. (1). A question that arises naturally is why to estimate the model tendencies from time series instead of directly evaluating the model equations. The reason is to reduce the effect of the error that occurs, because we observe the truth at infrequent discrete times. Suppose the sampling interval τ is not small enough; then $\tilde{\dot{x}}_a(t)$ will be an incorrect estimation of $\dot{x}_a(t)$. If $\dot{x}(t)$ is used instead of $\tilde{\dot{x}}(t)$, then there is going to be a non-zero correction $c(x, p)$ even

though the initial model is equivalent to the truth, because $\dot{x}(t) = \dot{x}_a(t) \neq \tilde{\dot{x}}_a(t)$ due to the error introduced by any estimation scheme of the tendency. However, if $\tilde{\dot{x}}(t)$ is used instead and estimated by matching $h = \tau$ (i.e., we observe our current model the same way we observe the truth (with frequency τ)), then $\tilde{\dot{x}}(t) = \tilde{\dot{x}}_a(t)$ will hold and no correction is needed.

Once we obtain the correction $c(x, p)$, we can apply it to our initial model (1) and run a new iteration of the procedure with the corrected model. The convergence criteria can be decided depending on the model and application. In the following experiment we stop the iteration when the change in any of the estimated coefficients between two iterations is smaller than 10^{-4}.

For testing the method we use the Lorenz system with standard parameter values as a reference truth. The time series from the reference truth $\{x_a(t)\}$ is obtained by running fourth-order Runge–Kutta with timestep 0.01 from $t = 0$ to $t = 20$ and then taking every seventh observation – thus $\tau = 0.07$. We start the iteration procedure with a null model $\dot{x} = 0$, and the correction term has the general quadratic form

$$c_j(x, p) = p_{0,j} + \sum_{i=1}^{3} p_{i,j} x_i + \sum_{i=1}^{3} \sum_{k \geq i} p_{ik,j} x_i x_k \qquad (4)$$

for $j = 1, 2, 3$. The estimation of the tendency of the truth $\tilde{\dot{x}}_a(t)$ is done by using spline interpolation (Press et al., 2007) on the time series, because for the Lorenz system it gives better derivative estimation than finite differences, although for higher-order models this need not be true. We estimate $\tilde{\dot{x}}(t)$ by first generating time series $\{x(t - kh) \ldots x(t - h), x(t), x(t + h) \ldots x(t + kh)\}$ from the initial model by running backward and forward Runge–Kutta for k steps starting with the initial state $x_a(t) \equiv x(t)$. We found that a local time series with $k = 6$ is sufficient for good estimation $\tilde{\dot{x}}(t)$. When backward integration is not applicable due to instabilities, one can use a concentrated finite differencing scheme (Kravtsov et al., 2005):

$$\tilde{\dot{x}}(t) = \frac{1}{2h} \big(-3x(t) + 4x(t + h) - x(t + 2h) \big). \qquad (5)$$

On model time series we use spline interpolation and estimate $\tilde{\dot{x}}(t)$ at time t. Using the differences $\tilde{\dot{x}}_a - \tilde{\dot{x}}$ as target variables, we finally solve Eq. (3) using least squares to obtain the correction term. We can then recursively construct a new model $T_{n+1}(x, p) = T_n(x, p) + c(x, p)$. The results from the 1st, 2nd and last iteration of the procedure are summarized in Table 1. For the first iteration there are significant errors in the estimation of the bias terms and modest errors for the parameters ρ and β. Note that the first iteration is equivalent to applying standard multiple polynomial regression. In the second iteration we can see significant improvement in the estimation of the three Lorenz parameters, but the errors in the bias (constant) terms remain large. The

procedure converges after 13 iterations, and the largest error in any of the final estimated parameters is less than 1 %.

The iterative correction method cannot be applied straightforward to the imperfect model scenario since the imperfections in general cannot be resolved. One direction that is probably worth exploring is the use of stochastic parametrizations to account for the model imperfections. A recent approach that iteratively constructs inverse stochastic models with multilevel stochastic forcing has been effectively applied in explaining climate variability (Kravtsov et al., 2005; Kondrashov et al., 2005). This approach and the one presented in this section are algorithmically similar in that they apply least-squares iteratively to determine better model coefficients.

3 Empirical correction techniques for imperfect models

3.1 Leith's correction method

Leith (1978) proposed a technique that reduces model tendency error by adding both a state-independent (constant) term and a state-dependent (linear) term to the model tendency equations. He pointed out that for short forecast times the model errors can neither interact nor compound nonlinearly so a linear equation that approximates the model error might be useful. The state-independent term accounts for the difference in the time average between the model and the truth, while the state-dependent term accounts for the difference between the variances of the model and truth (Allgaier et al., 2012). For a general dynamical system given by Eq. (1), the proposed improved model has the following form:

$$\dot{x} = T(x, p) + \mathbf{L}x + b. \qquad (6)$$

The operator $\mathbf{L}$ and the constant forcing term b are obtained by least squares optimization of the mean square tendency error $\langle e^T e \rangle$ where the model tendency error e at a particular time t is given by

$$e = \tilde{\dot{x}}_a(t) - T(x_a(t)) - \mathbf{L}x_a(t) - b. \qquad (7)$$

We follow (Allgaier et al., 2012) and first calculate the state-independent correction:

$$b = \langle \Delta x \rangle / \tau, \qquad (8)$$

where $\langle \Delta x \rangle$ denotes the average of forecast errors over the entire training period. Note that the average forecast error has to be divided by the length of the sampling interval to obtain the instantaneous model correction term. This average provides an estimate for the systematic model error generated during the observation window, and adding it to the model equations should produce a model with less bias:

$$T^+(x) = T(x) + b. \qquad (9)$$

Table 1. Estimated coefficients in the quadratic form (4) for the 1st, 2nd and last iteration of the procedure.

$p(1)$	1	x	y	z	x^2	y^2	z^2	xy	xz	yz
$\dot{x}$	−0.6286	−11.1669	10.8096	0.0858	0.0168	−0.0004	−0.0031	−0.0114	0.0362	−0.0274
$\dot{y}$	11.3944	25.5828	−0.0911	−1.3503	−0.1735	−0.0598	0.0375	0.1930	−0.9332	−0.0182
$\dot{z}$	−8.4870	−0.9195	0.7317	−1.3070	0.3116	0.0690	−0.0471	0.6929	0.0262	−0.0233

$p(2)$	1	x	y	z	x^2	xy	xz	y^2	yz	z^2
$\dot{x}$	−0.1927	−10.2821	10.1960	0.0347	0.0077	−0.0011	−0.0013	−0.0040	0.0086	−0.0064
$\dot{y}$	4.8727	27.6313	−0.8928	−0.5571	−0.0734	−0.0293	0.0155	0.0867	−0.9898	−0.0017
$\dot{z}$	−3.6866	−0.3729	0.3118	−2.1688	0.0949	0.0273	−0.0159	0.9007	0.0105	−0.0095

$p(13)$	1	x	y	z	x^2	xy	xz	y^2	yz	z^2
$\dot{x}$	0.0058	−10.0008	10.0002	−0.0008	0.0000	0.0000	0.0000	−0.0000	0.0000	−0.0000
$\dot{y}$	−0.0063	28.0035	−1.0019	−0.0000	−0.0003	−0.0001	0.0000	0.0003	−1.0001	0.0001
$\dot{z}$	−0.0043	−0.0061	0.0038	−2.6668	0.0000	0.0001	0.0000	0.9999	0.0002	−0.0001

To calculate Leith's operator $\mathbf{L}$, we recompute the forecast error time series Δx with the bias-corrected model (9) and then produce two time series of anomalies:

$$x'_a(t) = x_a(t) - \langle x_a \rangle;$$
$$\Delta x'(t) = \Delta x(t) - \langle \Delta x \rangle. \tag{10}$$

Looking for an improved model of the form

$$T^{++}(x) = T^+(x) + \mathbf{L}(x - \langle x_a \rangle), \tag{11}$$

we arrive at the following expressions for Leith's operator $\mathbf{L}$:

$$\mathbf{L} = \mathbf{C}_{\Delta x' x'_a} \mathbf{C}_{x'_a x'_a}. \tag{12}$$

on the right-hand side is the lagged cross covariance matrix obtained by taking the outer product $\Delta x(t + \tau)' x_a(t)'^T$ over the entire training period, while the second matrix is the inverse of the cross covariance matrix of the truth $\mathbf{C}_{x'_a x'_a} = \langle x_a(t)'^T x_a(t)' \rangle^{-1}$.

3.2 Weighted combination of imperfect models

A more recent approach is to produce an improved model using weighted linear combination of the tendencies of several imperfect models. The imperfect models can be any number of different models expressing the change in the same atmospheric quantities. In general circulation models (GCMs), the differences are typically in the parametrization of subgrid-scale processes and forcing parameters. More formally, assume that we are given M models and that the tendency for the i-th component of model μ is given by

$$\dot{x}_i^\mu = T_i^\mu(x). \tag{13}$$

This technique constructs an improved model whose tendencies are given by a weighted combination of the tendencies of the different models. The equation for the i-th component

of the improved model is given by

$$\dot{x}_i^W = \sum_\mu w_i^\mu T_i^\mu = \boldsymbol{w}_i^T \boldsymbol{T}_i, \tag{14}$$

where $\boldsymbol{w}_i = \left[w_i^1, w_i^2, \ldots, w_i^M\right]^T$ is the vector of weights and $\boldsymbol{T}_i = \left[T_i^1, T_i^2, \ldots, T_i^M\right]^T$ is the vector of tendencies of different models. We are left with the task of searching for weights w_μ^i that will give a model with improved forecast when compared to the individual models. The search can be done with respect to any cost function, but the nature of the problem allows us to make a fast least square fit of the weights with respect to MSTE. In general, there are different basis functions T^i for every model component x^i, so we have to solve a separate regression problem for each one. The mean square tendency error for the i-th model component is

$$E\left(\boldsymbol{w}^i\right) = \langle |\dot{x}_a^i - \boldsymbol{w}_i^T \boldsymbol{T}_i(x)|^2 \rangle, \tag{15}$$

where the angle brackets denote time average. The gradient of the error function takes the form

$$\nabla E\left(\boldsymbol{w}^i\right) = \langle \left(\dot{x}_a^i - \boldsymbol{w}_i^T \boldsymbol{T}_i\right) \boldsymbol{T}_i^T \rangle, \tag{16}$$

and after setting it to zero we obtain the following linear system for the weights:

$$\mathbf{C}_i \boldsymbol{w}_i = \boldsymbol{c}_i. \tag{17}$$

Note that the covariance matrices between the tendencies of different models are given by $\mathbf{C}_i = \langle \boldsymbol{T}_i \boldsymbol{T}_i^T \rangle$. Also the covariance vectors between the true tendency and those of the different models are $\boldsymbol{c}_i = \langle \dot{x}_a^i \boldsymbol{T}_i \rangle$. To avoid over-fitting of the weights to the training set, a regularization term can be added to the error function (Bishop, 2006):

$$E(\boldsymbol{w}_i) + \frac{\lambda}{2} \boldsymbol{w}_i^T \boldsymbol{w}_i. \tag{18}$$

The regularization coefficient λ controls the relative importance of the data-dependent error $E(\boldsymbol{w}_i)$ and the regularization term $\frac{1}{2}\boldsymbol{w}_i^T\boldsymbol{w}_i$. Also, the regularization term is a quadratic function of the weights $\boldsymbol{w}$ allowing the solution of the problem to be expressed in closed form:

$$\boldsymbol{w}_i = (\lambda\mathbf{I} + \mathbf{C}_i)^{-1}\boldsymbol{c}_i. \tag{19}$$

The choice of the regularization parameter λ can influence the results, and typically we have to run different simulations for different values of λ in order to produce a good result. A clever ansatz from machine learning tells us that we can start with a small value for λ of 0.01 and increase it by roughly three times its value in each simulation. In our simulations we explored the parameter space for the values $[0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100, 300, 1000]$ and found that a value of $\lambda = 3$ gives the best results.

4 Results on the chaotically driven Lorenz 63 system

We now compare both empirical correction techniques in a realistic imperfect model scenario where the forecast models have different dimension and parametrization from the truth. Because atmospheric models are nonlinear and have chaotic behavior, we took the paradigmatic Lorenz 63 system as a model. As truth we considered a Lorenz system that is driven externally by some nontrivial functions. The role of this forcing can be given to the variables of another Lorenz 63 system that is independent. The whole six-dimensional system is given by

$$\dot{x}_1 = \sigma(y_1 - x_1) + \epsilon z_2,$$
$$\dot{y}_1 = x_1(\rho - z_1) - y_1,$$
$$\dot{z}_1 = x_1 y_1 - \beta z_1 + \delta(x_2 + \eta),$$
$$\dot{x}_2 = \sigma(y_2 - x_2),$$
$$\dot{y}_2 = x_2(\rho - z_2) - y_2,$$
$$\dot{z}_2 = x_2 y_2 - \beta z_2. \tag{20}$$

As can be seen the subsystem 2 is the ordinary Lorenz 63 system, which drives the subsystem 1. The latter is thus a non-autonomous dynamical system with variable forcing. The parameters ϵ and δ determine the magnitude of the driving, while η is some constant drift[1]. For simplicity we have taken the standard parameter values in both subsystems $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$. The values used in the simulations for the other parameters are $\epsilon = 1$, $\delta = 1$ and $\eta = 2$. Furthermore, we assume that only the subsystem 1 is accessible for observations, and our aim is to model its time series $x_1(t)$, $y_1(t)$ and $z_1(t)$. In Fig. 1 we show the projection of

[1]Inserting the drift η seems artificial, but our aim in this toy case is to have a constant as a parameter that will represent external forces. Since x_2 has both positive and negative values, we have added drift. On the other hand, z_2 is always positive and modeling its influence in $\dot{x}_1$ with a constant comes naturally.
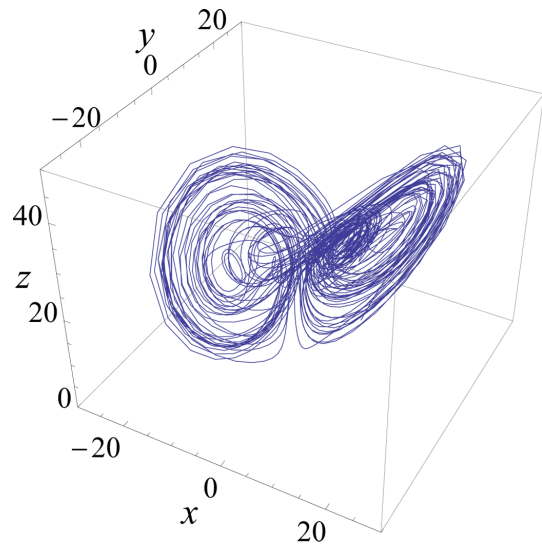

Fig. 1. Projection of the observations in the model space. The trajectory resembles those obtained from the ordinary Lorenz 63 system.

the full attractor on the three-dimensional space (x_1, y_1, z_1) where it is clear that it looks like the familiar Lorenz attractor.

The forecast model is the Lorenz 63 system with constant "forcing" terms added in the equations for the variables x and z. In most advanced atmospheric models, the additive constants are used to model the external forcing, which is due to some unresolved processes. The models developed at different meteorological centers differ in choice of the unresolved processes that are parameterized and also in the parameter values. In this simple example, these different choices can be mimicked by making two models that parameterize the external driving at the tendencies of x_1 and z_1 variables separately. It means that in the first case only the unresolved process appearing at the tendency of variable x is parameterized with some parameter α:

$$\dot{x}_3 = \sigma(y_3 - x_3) + \alpha;$$
$$\dot{y}_3 = x_3(\rho - z_3) - y_3;$$
$$\dot{z}_3 = x_3 y_3 - \beta z_3. \tag{21}$$

The value of α should be determined by fitting the model's output with the observations of the reality – subsystem 1 in Eq. (20). This optimization procedure models the parameter fitting that is performed by the meteorologists.

Similarly, one can build another model equation that will have a parametrization of the process that drives the variable z:

$$\dot{x}_4 = \sigma(y_4 - x_4),$$
$$\dot{y}_4 = x_4(\rho - z_4) - y_4,$$
$$\dot{z}_4 = x_4 y_4 - \beta z_4 + \gamma, \tag{22}$$

with the parameter γ. Thus the Eqs. (21) and (22) can be regarded as two imperfect models developed at different

meteorological centers and serve as a test bed for comparing the empirical correction techniques. We assumed values for α and γ of 30 and 4, respectively.

4.1 Training and testing

For the purpose of training of each of the techniques, we used a set of observations comprised of 100 time units of the variables of subsystem 1. Both the simulated truth and the model are integrated with fourth-order Runge–Kutta with a time step of 0.01. The sampling interval is assumed to be 1 model time steps, so the resulting training set is 10 000 points.

To compare the methods we generated a run of 1 million points, from which we chose randomly a set of 1000 starting points for which trial forecasts were performed. For each forecast, anomaly correlation (AC) is computed with respect to the reference truth and then averaged over all trials. Anomaly correlation is typically used to determine the length of time for which a model forecast is useful, and it is given by

$$\text{AC} = \frac{(\boldsymbol{x}_a - \langle \boldsymbol{x}_a \rangle)^T \cdot (\boldsymbol{x} - \langle \boldsymbol{x}_a \rangle)}{\sqrt{|\boldsymbol{x}_a - \langle \boldsymbol{x}_a \rangle|^2 |\boldsymbol{x} - \langle \boldsymbol{x}_a \rangle|^2}} \tag{23}$$

for a particular time. AC is essentially the dot product of the anomalous model state with the anomalous true state. A forecast is typically considered useful for as long as its AC remains above 0.6. The AC is calculated for each imperfect model and both of the corrected models and averaged over the 1000 random trials. The results of the forecast performance in terms of the AC metric are shown in Fig. 2. As expected, both correction techniques offer better performance when compared to the imperfect models. The first peculiar observation is that the parametrization in the first equation is more important since it gives better forecast. On the other hand, both correction techniques show very similar performance, with Leith's method edging out the weighted combination model by a small margin. A closer look at the estimated coefficients by the two correction methods reveals why this is the case. The coefficients in Leith's operator along with the estimated bias term are shown in Table 2. Obviously there is little if any correction appearing before the linear terms, while the bias values are strikingly close to the average values of the chaotic forcing in the truth (approximately 23.5 and 2). Similar results are obtained for the weighted combination where the estimated forcings have values in the equations x and z of 20.2 and 0.76, respectively. One can notice from Table 3 that the coefficients in front of the linear and quadratic terms differ from those of the original Lorenz by 15 % in the equation x and by 5 % in the equation z, because the weights in front of the different models do not add up to one. This can be readily removed by solving Eq. (15) with a quadratic programming technique where a linear constraint can be put on the weights so that they sum up to one. In particular, we have used a subspace trust-region technique
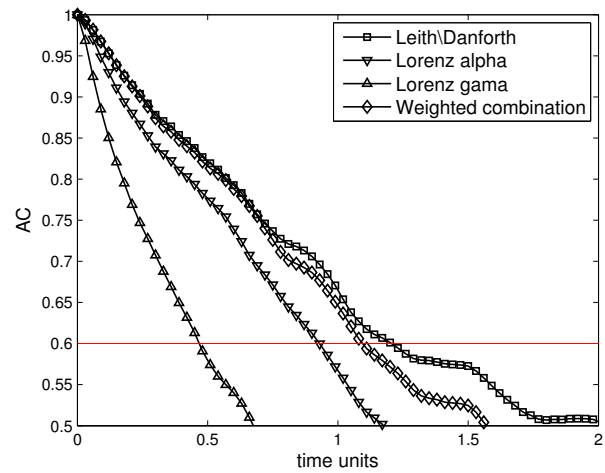

Fig. 2. Anomaly correlation calculated for the two imperfect models and the models obtained by the two model correction techniques.

Table 2. The estimated coefficients of Leith's operator (12) and the bias term (8).

	$\dot{x}$	$\dot{y}$	$\dot{z}$
Lx	0.0907	0.0024	−0.0459
Ly	−0.0450	−0.0009	0.0159
Lz	−0.0350	−0.0010	0.0200
b	22.7862	0.0958	1.3730

for solving quadratic programming problems based on the interior-reflective Newton method (Coleman and Li, 1996).

This indication from both methods shows that the biggest improvement in predictability comes from estimating the average value of the chaotic driving in the variable x, which in general is not straightforward since the driving comes from some unresolved processes. When we calculated the AC metric for a Lorenz system with parameterizations $\alpha = 23.5$ and $\gamma = 2$, there was virtually no difference with Leith's method. This conclusion also has implications for models with stochastic parametrization. The parameters of the stochastic forcing term should be estimated so as to align its average value with that of the chaotic drive signal. Capturing higher-order statistics of the chaotic drive signal with the stochastic parametrization might provide additional increase in predictability.

5 Conclusions

In this work we have analyzed techniques for reducing model error by empirically correcting the dynamical equations of the models. For the case when the observations of the truth are infrequent, we proposed an iterative linear regression method that effectively estimates all the parameters in the perfect model scenario. The key insight into this method is to match the observation time of the model to the observation time of the truth so that the error resulting from infrequent observations is reduced. An open question is how to apply

Table 3. The estimated weights of the weighted combination model (14) appearing in front of the corresponding imperfect model tendencies.

	w_1	w_2	w_3
Model 3	0.6737	1.0000	0.7586
Model 4	0.1859	0	0.1940

this method in the context of an imperfect model scenario where model error is always present and each iteration in the procedure needs not produce a better model than the previous. One possible extension of the method is in the direction of models with stochastic parametrization of the imperfections.

We have also tested and analyzed two promising techniques for model correction on a realistic, imperfect model scenario where the truth was given by a chaotically driven Lorenz system and the models are given by the standard Lorenz system with a constant forcing. Both techniques indicate that the biggest performance gain comes from estimating a value for the constant forcing that is close to the average of the chaotic drive forcing. Nevertheless, this may not be true for higher-order climate models, and applying the techniques to correct parameters in front of linear and quadratic terms in operational models is worth exploring because of the importance of weather and climate forecasting. Currently, in operational weather and climate models, Leith's method has been successfully applied only for the state-independent terms (DelSole et al., 2008; Yang et al., 2008) due to the fact that calculating the full linear term is computationally prohibitive. The weighted combination of models has not been applied in any form on operational models, although work is underway in applying the technique to state-of-the-art European climate models (Shen and Keenlyside, 2012).

References

Allgaier, N., Harris, K., and Danforth, C.: Empirical correction of a toy climate model, Phys. Rev. E, 85, 026201, doi:10.1103/PhysRevE.85.026201, 2012.

Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, Mon. Weather Rev., 129, 2884–2903, 2001.

Basnarkov, L. and Kocarev, L.: Forecast improvement in Lorenz 96 system, Nonlin. Processes Geophys., 19, 569–575, doi:10.5194/npg-19-569-2012, 2012.

Bishop, C.: Pattern recognition and machine learning, Springer New York, 2006.

Coleman, T. F. and Li, Y.: A reflective Newton method for minimizing a quadratic function subject to bounds on some of the variables, SIAM J. Optimiz., 6, 1040–1058, 1996.

Danforth, C. and Kalnay, E.: Impact of online empirical model correction on nonlinear error growth, Geophys. Res. Lett., 35, L24805, doi:10.1029/2008GL036239, 2008.

Danforth, C., Kalnay, E., and Miyoshi, T.: Estimating and correcting global weather model error, Mon. Weather Rev., 135, 281–299, 2007.

DelSole, T. and Hou, A.: Empirical correction of a dynamical model. Part I: Fundamental issues, Mon. Weather Rev., 127, 2533–2545, 1999.

DelSole, T., Zhao, M., Dirmeyer, P., and Kirtman, B.: Empirical correction of a coupled land-atmosphere model, Mon. Weather Rev., 136, 4063–4076, 2008.

Du, H. and Smith, L.: Parameter estimation through ignorance, Phys. Rev. E, 86, 016213, doi:10.1103/PhysRevE.86.016213, 2012.

Hunt, B., Kalnay, E., Kostelich, E., Ott, E., Patil, D., Sauer, T., Szunyogh, I., Yorke, J., and Zimin, A.: Four-dimensional ensemble Kalman filtering, Tellus A, 56, 273–277, 2004.

Judd, K. and Smith, L.: Indistinguishable states II: The imperfect model scenario, Physica D, 196, 224–242, 2004.

Kalnay, E.: Atmospheric modeling, data assimilation and predictability, Cambridge University Press, Cambridge, UK, 2002.

Kondrashov, D., Kravtsov, S., Robertson, A., and Ghil, M.: A hierarchy of data-based ENSO models, J. Climate, 18, 4425–4444, 2005.

Kravtsov, S., Kondrashov, D., and Ghil, M.: Multilevel regression modeling of nonlinear processes: Derivation and applications to climatic variability, J. Climate, 18, 4404–4424, 2005.

Leith, C.: Objective methods for weather prediction, Annu. Rev. Fluid Mech., 10, 107–128, 1978.

Lorenz, E.: Deterministic nonperiodic flow, J. Atmos. Sci., 20, 130–141, 1963.

Merkova, D., Szunyogh, I., and Ott, E.: Strategies for coupling global and limited-area ensemble Kalman filter assimilation, Nonlin. Processes Geophys., 18, 415–430, doi:10.5194/npg-18-415-2011, 2011.

Orrell, D.: Modelling nonlinear dynamical systems: chaos, error, and uncertainty, Ph.D. thesis, Oxford University, Oxford, UK, 2001.

Orrell, D., Smith, L., Barkmeijer, J., and Palmer, T. N.: Model error in weather forecasting, Nonlin. Processes Geophys., 8, 357–371, doi:10.5194/npg-8-357-2001, 2001.

Palmer, T.: Predicting uncertainty in forecasts of weather and climate, Rep. Prog. Phys., 63, 71–116, 2000.

Press, W., Teukolsky, S., Vetterling, W., and Flannery, B.: Numerical recipes, 3rd Edn., The art of scientific computing, Cambridge University Press, Cambridge, UK, 2007.

Shen, M.-L. and Keenlyside, N.: Report on climate super model after learning, Tech. rep., 1–27, available at: http://www.sumoproject.eu/downloads/SUMO_WP2_yr2_deliverable2_2.pdf, 2012.

Smith, L., Cuéllar, M., Du, H., and Judd, K.: Exploiting dynamical coherence: A geometric approach to parameter estimation in nonlinear models, Phys Lett. A, 374, 2618–2623, 2010.

van den Berge, L. A., Selten, F. M., Wiegerinck, W., and Duane, G. S.: A multi-model ensemble method that combines imper-

fect models through learning, Earth Syst. Dynam., 2, 161–177, doi:10.5194/esd-2-161-2011, 2011.

Wiegerinck, W., Burgers, W., and Selten, F.: Report low/intermediate complexity super models, Tech. rep., 1–21, available at: http://www.sumoproject.eu/downloads/SUMO_WP5_yr2_deliverable5_2.pdf, 2011.

Yang, X., DelSole, T., and Pan, H.: Empirical correction of the NCEP global forecast system, Mon. Weather Rev., 136, 5224–5233, 2008.