

A State-Space Approach to Analyze Structural Uncertainty in Physical Models

Azam Moosavi^a and Adrian Sandu^b

^{a,b}Computational Science Laboratory, Department of Computer Science,
Virginia Polytechnic Institute and State University, Blacksburg, VA 24060.

^aazmosavi@vt.edu, ^bsandu@cs.vt.edu

Abstract. Physics-based computer models, such as fluid flow simulations, seek to approximate the behavior of a real system based on the physical equations that govern the evolution of that system. The model approximation of reality, however, is imperfect because it is subject to uncertainties coming from different sources: finite model resolution, uncertainty in model parameter values, uncertainty in input data such as external forcings, and uncertainty in the structure of the model itself. Many studies to date have considered the effects of parameter and data uncertainty on model outputs, and have offered solutions to obtain the best fitted parameter values for a model. However, much less effort has been devoted to the study of structural uncertainty, which is caused by our incomplete knowledge about the true physical processes, and manifests itself as missing dynamics in the model. This paper seeks to understand structural uncertainty by studying the observable errors, i.e., the discrepancies between the model solutions and measurements of the physical system. The dynamics of these errors is modeled using a state-space approach, which enables to identify the source of uncertainty and to recognize the missing dynamics inside model. Furthermore, the model solution can be improved by correcting it with the error predicted by the state-space approach. The proposed methodology is applied to two test problems, Lorenz-96 and a stratospheric chemistry model.

Keywords: uncertainty, state-space models, structural uncertainty.

1. Introduction

First-principles models encapsulate our knowledge about the physical processes that govern the evolution of a natural system. The solutions of such models approximate the evolution of the physical system under consideration. However, models are imperfect since they are subject to multiple uncertainties associated with the input data, external forcings, parameters, and the structure of the model. The process of uncertainty quantification includes the identification, characterization, propagation, analysis, and finally the reduction of these uncertainties [1]. The identification of sources of uncertainty considers parameter uncertainty, structural uncertainty, algorithmic uncertainty, experimental uncertainty and interpolation uncertainty [2]. In forward uncertainty propagation various uncertainties are run through the model to assess the total uncertainty in the outputs. In inverse uncertainty propagation the values

of uncertain parameters are estimated using the discrepancy between the model and observations [3].

Structural uncertainty, also named model-form uncertainty, is due to insufficient knowledge of the true physics and therefore to an incomplete representation of reality by the model [4, 5]. While uncertainties due to data and parameter values have been studied extensively, comparatively little work has been devoted to date to studying structural uncertainties. Two approaches have been employed to resolve the structural uncertainty: model averaging and discrepancy modeling (or model calibration). Model averaging considers weighted averages of the outputs of different models, with the weights obtained from a metric of model adequacy, e.g., likelihood with respect to the data. The limitations of this approach are that it requires the use of multiple models, and that the models may not be independent. Discrepancy modeling assesses the model based on the discrepancy between model outputs and the real data [6].

In [7] a Gaussian Process is used to model the discrepancy between the model outputs and reality, and then a Markov Chain Monte Carlo (MCMC) is employed to sample the posterior distribution of the discrepancy. In [2] a Bayesian approach is taken to quantify different aspects of model uncertainty such as uncertainty in parameters, parametric variability, residual variability, observation error, code uncertainty, and model inadequacy. The model inadequacy is defined as the difference between the true mean value of the physical quantity of interest and the model output for the true values of the inputs. The posterior distribution over the parameters is sampled by MCMC to obtain the calibrated values of parameters. In [6] a distinction is drawn between external discrepancies that correspond to the entire model inadequacies, and internal discrepancies corresponding to individual sub-functions that compose the model. Internal discrepancies are defined with respect to the output of each sub-function when given the true input values. A joint distribution over internal discrepancies and the input data is specified. The sensitivity of the model output to each of the internal discrepancies is determined in order to assess the relative importance of the structural uncertainty within each sub-model. If the output is insensitive to all internal discrepancies then the full model is considered sufficiently accurate. However, large impacts of internal discrepancies mean that there is considerable prediction uncertainty arising from uncertain model structure, that this uncertainty should be reduced. The Integrated Bayesian Uncertainty Estimator (IBUNE) discussed in [8] considers uncertainties coming from parameters, inputs, and model structure. IBUNE combines Bayesian model averaging (BMA) to reduce the structural uncertainties of individual sub-models with the Shuffled Complex Evolution Metropolis (SCEM) algorithm for probabilistic parameter estimation to the input and parameter uncertainties.

This study seeks to identify the missing dynamics in models that leads to the discrepancy between the model outputs and the measurements of the physical world. Our approach is to represent the dynamics of the structural uncertainty using linear state-space models, and to identify the parameters of these models from the known model states and model-observations discrepancies. The predicted structural

uncertainties can be used to correct the model solutions, and therefore to improve the model. Previous work that has applied the theory of linear dynamical systems to structural uncertainty includes the use of an extended Kalman filter for system identification of seismic model errors [9], and the use of an ensemble Kalman filter to manage the structural uncertainties in reservoir models [10].

The reminder of this paper is organized as follows, Section 2 describes the proposed methodology for analysis of structural uncertainty. Numerical results are presented in Section 3. Section 4 summarizes the findings of this work and draws the conclusions.

2. Structural uncertainty

Consider a physical process with internal state $\mathbf{v}_t$ at the discrete time $t \in \mathbb{N}$, evolving according to the dynamic equation:

$$\mathbf{v}_t = \mathcal{P}(\mathbf{v}_{t-1}) + \eta_t, \quad t = 1, \dots, T. \quad (1)$$

The exact physics $\mathcal{P}$ and the physical state $\mathbf{v}_t$ are not known exactly. We take measurements of a finite number of observables $\mathbf{y}_t \in \mathbb{R}^m$ of the physical process:

$$\mathbf{y}_t = h(\mathbf{v}_t) + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \mathbf{R}_t), \quad t = 1, \dots, T, \quad (2)$$

where ϵ_t is the observation error, assumed to be normally distributed.

A computer model $\mathcal{M}$ evolves the model state $\mathbf{x}_t \in \mathbb{R}^n$ from one discrete time (t) to the next ($t + 1$):

$$\mathbf{x}_t = \mathcal{M}(\mathbf{x}_{t-1}, \hat{\theta}), \quad t = 1, \dots, T. \quad (3a)$$

Without loss of generality we assume that the parameters of the model $\theta \in \mathbb{R}^\ell$ take the best-fitting values $\hat{\theta}$ and we don't have any parameter uncertainty. The observation operator $\mathcal{H}$ maps the model state onto the observation space; the model-predicted values $\mathbf{z}_t \in \mathbb{R}^m$ of the observations (2) are:

$$\mathbf{z}_t = \mathcal{H}(\mathbf{x}_t), \quad t = 1, \dots, T. \quad (3b)$$

The computer model $\mathcal{M}$ seeks to approximate the true physical process; specifically, the model state approximates the physical state,

$$\mathbf{x}_t \approx \xi(\mathbf{v}_t), \quad t = 1, \dots, T, \quad (4)$$

where the operator ξ can represent, for example, the sampling of a continuous temperature field onto a finite dimensional computational grid. The model dynamics (3a) approximates the dynamics of the physical system (1). If we initialize the model at time t with an idealized value (a projection of the real state), the model prediction at time $t + 1$ will differ from the reality:

$$\xi(\mathbf{v}_t) = \mathcal{M}(\xi(\mathbf{v}_{t-1}), \hat{\theta}) + \delta_t(\mathbf{v}_{t-1}), \quad t = 1, \dots, T. \quad (5)$$

The discrepancy δ_t between model prediction and the reality, projected onto the model space, is the structural uncertainty of concern here. In the discussion that follows we will make the following simplification. We assume that the physical system is finite dimensional, $\mathbf{v}_t \in \mathbb{R}^n$, and that the model state lives in the same space as reality, i.e., $\mathbf{x}_t \approx \mathbf{v}_t$ and $\xi(\cdot) \equiv id$ is the identity operator in (4), and $\mathcal{H}(\cdot) \equiv h(\cdot)$ in (2) and (3b). This assumption means that the discretization errors are very small, and that the main source of error is the missing representation of some physical processes. With these assumptions the evolution equations for the physical system (1) and the physical observations equation (2) become, respectively:

$$\mathbf{v}_t = \mathcal{P}(\mathbf{v}_{t-1}) + \eta_t \quad (6a)$$

$$= \mathcal{M}(\mathbf{v}_{t-1}, \hat{\theta}) + \delta_t(\mathbf{v}_{t-1}) + \eta_t, \quad t = 0, \dots, T-1,$$

$$\mathbf{y}_t = h(\mathbf{v}_t) + \epsilon_t. \quad (6b)$$

We seek to understand the structure of the model-form error (5) by comparing the observations (6b) of the real system against the model predicted values of these observables (3b), i.e., by considering the discrepancies in the observable quantities:

$$\Delta_t = \mathbf{z}_t - \mathbf{y}_t \in \mathbb{R}^m, \quad t = 1, \dots, T. \quad (7)$$

Specifically, our goal is to model the state error δ_t in order to gain understanding of the missing dynamics, i.e., of the physical processes that are not captured by the model $\mathcal{M}$. Moreover, good estimates of the discrepancy δ_t allow to improve model predictions by applying the correction (6a) to model results:

$$\mathbf{v}_t \approx \mathbf{x}_t + \delta_t. \quad (8)$$

We consider two techniques to quantify the structure of the uncertainty in the models. The first is the global structural uncertainty approach, which models the discrepancy between the model state and physical state over the entire state space. The second is the local uncertainty approach, which only takes account the variables in the vicinity of the source of uncertainty and not all variables in the system. We now give a detailed description of these approaches.

2.1. Global models for structural uncertainty

Rooted in control engineering, state-space models (SSMs) are dynamical systems that describe the probabilistic dependence between the latent variables and the observed measurements [11]. A state-space vector evolves forward in time under the control of external inputs. An output equation captures the relationship between the system state, the input, and the output. A linear time-invariant (LTI) discrete-time SSM has the form:

$$\zeta_t = \mathbf{A} \cdot \zeta_{t-1} + \mathbf{B} \cdot \mathbf{u}_{t-1} + \mathbf{K} \cdot \nu_{t-1}, \quad t = 1, \dots, T; \quad \zeta_0 = \text{given}, \quad (9a)$$

$$\gamma_t = \mathbf{C} \cdot \zeta_t + \mathbf{D} \cdot \mathbf{u}_t + \epsilon_t, \quad t = 1, \dots, T, \quad (9b)$$

where (9a) is the state equation, and (9b) is the output equation. Here $\boldsymbol{\zeta}_t \in \mathbb{R}^{n_{\text{ssm}}}$ is the state vector, $\mathbf{u}_t \in \mathbb{R}^{p_{\text{ssm}}}$ is the input or control variable, $\boldsymbol{\gamma}_t \in \mathbb{R}^{m_{\text{ssm}}}$ is output of the system (observable quantities), $\boldsymbol{\nu}_t \in \mathbb{R}^{q_{\text{ssm}}}$ is the disturbance of the model, and $\boldsymbol{\epsilon}_t \in \mathbb{R}^{m_{\text{ssm}}}$ is the observation error. The evolution of the system is captured by the dynamics or state matrix $\mathbf{A} \in \mathbb{R}^{n_{\text{ssm}} \times n_{\text{ssm}}}$, the input matrix $\mathbf{B} \in \mathbb{R}^{n_{\text{ssm}} \times p_{\text{ssm}}}$, the output or sensor matrix $\mathbf{C} \in \mathbb{R}^{m_{\text{ssm}} \times n_{\text{ssm}}}$, the feed-through or feed forward matrix $\mathbf{D} \in \mathbb{R}^{m_{\text{ssm}} \times p_{\text{ssm}}}$, and the disturbance matrix $\mathbf{K} \in \mathbb{R}^{n_{\text{ssm}} \times q_{\text{ssm}}}$.

In order to understand the structure of uncertainty in this work we employ SSMs (9) to describe the evolving discrepancy between the model and the true physics, as follows:

- The state vector $\boldsymbol{\zeta}_t$ consists of hidden states that represent the dynamics of the structural model error, and are driven by the error dynamics of the model.
- The hidden error $\boldsymbol{\zeta}_t$ depends on the model solution at every time step and on the discrepancy between the model outputs and the physical observations. The inputs of our SSM are therefore the model predictions (3a) :

$$\hat{\mathbf{x}}_t = \mathcal{M}(\mathbf{x}_{t-1}, \hat{\theta}), \quad t = 1, \dots, T. \quad (10)$$

- The outputs of the SSM are the observed discrepancies (7) between the model forecast (10) and reality:

$$\hat{\Delta}_t = \hat{\mathbf{z}}_t - \mathbf{y}_t = h(\hat{\mathbf{x}}_t) - \mathbf{y}_t, \quad t = 1, \dots, T. \quad (11)$$

- The outputs of the SSM predict the total stat-space discrepancy $\hat{\Delta}_t$.

The SSM that models the global error dynamics is:

$$\boldsymbol{\zeta}_t = \mathbf{A} \cdot \boldsymbol{\zeta}_{t-1} + \mathbf{B} \cdot \hat{\mathbf{x}}_t + \mathbf{K} \cdot \boldsymbol{\eta}_t, \quad \boldsymbol{\zeta}_0 = \text{given}, \quad (12a)$$

$$\hat{\Delta}_t = \mathbf{C} \cdot \boldsymbol{\zeta}_t + \mathbf{D} \cdot \hat{\mathbf{x}}_t + \boldsymbol{\epsilon}_t. \quad (12b)$$

From the sequence of known inputs $\hat{\mathbf{x}}_t$ and the known outputs $\hat{\Delta}_t$, $t = 1, \dots, T$ one can perform a system identification and fit the matrices $\mathbf{A}, \mathbf{B}, \mathbf{K}, \mathbf{C}, \mathbf{D}$. The matrix entries contain information about the structure and source of uncertainty in models; this aspect will be discussed in detail in the numerical experiments Section 3.

It is important to note that in the global approach, the SSM (12) approximates the dynamics of the model error over the entire model space. Since the dimension of the hidden error vector $\boldsymbol{\zeta}_t \in \mathbb{R}^{n_{\text{ssm}}}$ can be much smaller than the dimension of the model state-space error vector $\boldsymbol{\delta}_t \in \mathbb{R}^n$, $n_{\text{ssm}} < n$, the hidden error vector predicted by (12) needs to be lifted up to the model space:

$$\boldsymbol{\delta}_t = \mathbf{F} \cdot \boldsymbol{\zeta}_t + \mathbf{G} \cdot \hat{\mathbf{x}}_t. \quad (13)$$

In order to define the projection matrices $\mathbf{F} \in \mathbb{R}^{n \times n_{\text{ssm}}}$ and $\mathbf{G} \in \mathbb{R}^{n \times n}$ and one needs to use additional information about the structure of the model and the observation operator. Some possible definitions are discussed below.

- The global error dynamics model (12) can be defined on the entire state-space, i.e., $n_{\text{ssm}} = n$. In this case the hidden error approximates the entire model error vector $\zeta_t \approx \delta_t$, and therefore $\mathbf{F} = \mathbf{I}$ and $\mathbf{G} = \mathbf{0}$ in (13).
- The hidden error dynamics vector (12) can represent the projection of the full model error onto a subspace spanned by the orthonormal basis $\Phi \in \mathbb{R}^{n \times n_{\text{ssm}}}$ with $\Phi^T \cdot \Phi = \mathbf{I} \in \mathbb{R}^{n_{\text{ssm}} \times n_{\text{ssm}}}$. In this case we have:

$$\zeta_t \approx \Phi^T \cdot \delta_t, \quad \delta_t \approx \Phi \Phi^T \cdot \zeta_t \quad \Rightarrow \quad \mathbf{F} = \Phi \Phi^T \quad \text{and} \quad \mathbf{G} = \mathbf{0}.$$

The reduced order basis $\Phi \in \mathbb{R}^{n \times n_{\text{ssm}}}$ can be obtained from snapshots of a more complex model trajectories by proper orthogonal decomposition.

- The lifting operation (13) can be constructed using the observation operator. Specifically, assuming a small model error we have:

$$\hat{\Delta}_t = h(\mathbf{x}_t) - h(\mathbf{v}_t) \approx \frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}_t) (\mathbf{x}_t - \mathbf{v}_t) = -\mathbf{H}_t \delta_t, \quad \text{where} \quad \mathbf{H}_t = \frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}_t). \quad (14)$$

Let $\mathbf{H}_t^+$ be the pseudo-inverse of $\mathbf{H}$. From (14) and (12b) we have that:

$$\mathbf{x}_t - \mathbf{v}_t = -\delta_t \approx \mathbf{H}_t^+ \hat{\Delta}_t \approx \mathbf{H}_t^+ \mathbf{C} \zeta_t + \mathbf{H}_t^+ \mathbf{D} \hat{\mathbf{x}}_t \quad \Rightarrow \quad \mathbf{F} = \mathbf{H}_t^+ \mathbf{C} \quad \text{and} \quad \mathbf{G} = \mathbf{H}_t^+ \mathbf{D}. \quad (15)$$

Before the next time step the model forecast (10) is corrected with the predicted discrepancy (13) in order to obtain a better approximation of the true physical state:

$$\mathbf{x}_t = \hat{\mathbf{x}}_t + \delta_t. \quad (16)$$

2.2. Local models for structural uncertainty

In order to capture the global error dynamics the SSM (12) needs to be large. Finding the matrix coefficients from the inputs and outputs of this model is highly challenging. Moreover, many models have a dynamics driven by local dependencies among variables. The model structural errors due to local inaccuracies in the representation of physics have local impacts.

In order to address these points we consider local error models that seek to capture the structural uncertainty associated with only subsets of variables of the model. The assumption is that structural errors are associated with certain parts of the model, and that the effects of one type of structural error is observed locally in the evolution of “nearby” model variables and quantities of interest. By nearby we mean variables in physical proximity (e.g., grid points located at a short distance from each other) or variables that are tightly coupled to one another. Some prior knowledge about the source of structural errors may be required in order to define the local subset of variables affected by it. The interaction between local and remote variables is assumed to be sufficiently weak such that we can capture the evolution of errors using only local models: the errors depend on local variable states as inputs, and provide the

discrepancies between the model forecasts and true physics only on for those observable quantities that measure local variables.

The large SSM (12) modeling the global error dynamics is split into a set of L small-dimensional “local” SSMs:

$$\zeta_t^\ell = \mathbf{A}^\ell \cdot \zeta_{t-1}^\ell + \mathbf{B}^\ell \cdot \hat{\mathbf{x}}_t^\ell + \mathbf{K}^\ell \cdot \eta_t^\ell, \quad \zeta_0^\ell = \text{given}, \quad (17a)$$

$$\hat{\Delta}_t^\ell = \mathbf{C}^\ell \cdot \zeta_t^\ell + \mathbf{D}^\ell \cdot \hat{\mathbf{x}}_t^\ell + \epsilon_t^\ell, \quad (17b)$$

$$\delta_t^\ell = \mathbf{F}^\ell \cdot \zeta_t^\ell + \mathbf{G}^\ell \cdot \hat{\mathbf{x}}_t^\ell, \quad (17c)$$

$$\mathbf{x}_t^\ell := \hat{\mathbf{x}}_t^\ell + \delta_t^\ell, \quad \ell = 1, \dots, L. \quad (17d)$$

Here $\mathbf{x}^\ell$ consists of a local subset of model variables (e.g., associated with a certain geographic region), ζ_t^ℓ are the hidden states associated with the local model errors, and $\hat{\Delta}_t^\ell$ are the observations associated with that particular geographic region. In the local approach the large model (12) is replaced by a set of small error models (17), one for each local set of variables that are affected by structural errors.

In absence of any prior knowledge regarding the location of potential sources of uncertainty one can iterate over subsets of variables of the model, assume the uncertainty is associated with those variables, and construct the local error SSM (17) using only the variables in the locality of the source of uncertainty. The state is corrected according to (17d). The difference between the observables associated with the corrected solution (17d) and the real data allows to rank the different possible sources of error, and to select the ones that lead to the smallest differences. Algorithm 1 summarizes this approach.

Algorithm 1 Identifying the subsets of variables most affected by model uncertainty using local structural uncertainty modeling.

UncertaintyVars=Subset of variables in the model concern of having uncertainty.

For variable= UncertaintyVars[0]+L: UncertaintyVars[end]-L

$\hat{\mathbf{x}}^\ell = \mathbf{x}(\text{variable} - \frac{L}{2} : \text{variable} + \frac{L}{2})$

 Obtain $\mathbf{x}^\ell$ using (17)

 rank(variable) = RMSE ($y, \mathbf{x}^\ell$)

End For

[min,index]=minimum(rank)

return $\mathbf{x}^\ell = \mathbf{x}(\text{index} - \frac{L}{2} : \text{index} + \frac{L}{2})$

3. Numerical experiments

In order to illustrate the proposed modeling of structural errors we employ two test systems. The first one is the Lorenz-96 model with 40 variables [12], and the second model is the stratospheric chemistry model [13]. For our experiments we use the SSM library in Matlab [14] to obtain the state-space equations of the error models, i.e., the $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ matrices. We use the identity observation operator, i.e., we observe the errors in all model states. From (15) with $\mathbf{H}_t^+ = \mathbf{I}$ we have $\delta_t \approx \hat{\Delta}_t$.

3.1. Lorenz-96 model

The one dimensional Lorenz model is given by [12]:

$$\frac{dX_k}{dt} = -X_{k-2}X_{k-1} + X_{k-1}X_{k+1} - X_k + F, \quad k = 1, 2, \dots, K, \quad (18)$$

with $K = 40$ variables, periodic boundary conditions, and a forcing term $F = 8$. The variables of Lorenz model are periodic meaning that, values of X_{k-K} and X_{k+K} is equal to X_k .

A more complex Lorenz model includes two distinct time scales. It couples each variable X_k of the slow model (18) with a fast oscillating system described by the variables $\{Y_{j,k}\}_{k=1,\dots,K,j=1,\dots,J}$. The two-scale Lorenz model (19) adds five fast variables ($J = 5$) to each slow variable in the basic Lorenz (18):

The equations are:

$$\begin{aligned} \frac{dX_k}{dt} &= X_{k-1}(X_{k+1} - X_{k-2}) - X_k - (hc/b) \sum_{j=1}^J Y_{j,k} \\ \frac{dY_{j,k}}{dt} &= -cbY_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - cY_{j,k} + (hc/b)X_k, \\ k &= 1, 2, \dots, K, \quad j = 1, 2, \dots, J. \end{aligned} \quad (19)$$

System (19) has 200 variables. We select the coupling coefficient value $h = 0.2$. The parameters $c = 1$ and $b = 10$ are chosen such that the convective (fast) scales oscillate ten times faster than the large (slow) scales.

We consider (19) to be the real physical system, and (18) to be the computer model approximation. The mapping of variables is $\mathbf{v}_t = \{X_k, Y_{j,k}\}_{1 \leq k \leq K, 1 \leq j \leq J}$, $\mathbf{x}_t = \{X_k\}_{1 \leq k \leq K}$, and $\xi(\mathbf{v}_t) = \{X_k\}_{1 \leq k \leq K}$. For simplicity, we use the identity observation operator $\mathcal{H}(\mathbf{x}_t) = \mathbf{x}_t$. Using the discrepancy between the model solution and the true physics we employ the global and the local modeling approaches to analyze the structure of uncertainty in the model.

3.1.1. Global structural uncertainty modeling results For the global approach at each time step we observe the discrepancy between the model solution and the real data for all model variables (7). The model (18) is integrated forward in time starting from real data to obtain the model forecast $\hat{\mathbf{x}}_t$. The SSM states ζ_t are predicted from past values, the observed discrepancies, and the model forecast. Before the next step is taken the model solution is corrected using the predicted global error according to (12a). This procedure is summarized by Algorithm 2.

In the first experiment we have perturbed the 12th and 30th variables in the Lorenz system, and kept all other slow variables without the additional forcing from the fast scales not captured by the model. This means that the structural uncertainty affects only two variables in the model. The system will be integrated forward in time for 1000 time steps, and the number of error hidden states was empirically chosen equal to eight. Figure 1 shows the corrected model solution using with reduced uncertainty.

Algorithm 2 Reducing the uncertainty in the Lorenz model using the global approach.

$$\Delta_t = \hat{\mathbf{x}}_t - \mathbf{v}_t, \quad t \in \{\text{training times}\}.$$

$$[\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \zeta_0] = SSM(\{\Delta_t, \hat{\mathbf{x}}_t\}_{t \in \{\text{training times}\}}).$$

$$\mathbf{x}_0 = y_1$$

For $t = 1 : T$

$$\hat{\mathbf{x}}_t = \mathcal{M}(\mathbf{x}_{t-1}, \hat{\theta})$$

$$\zeta_t = \mathbf{A} \zeta_{t-1} + \mathbf{B} \hat{\mathbf{x}}_t$$

$$\hat{\Delta}_t = \mathbf{C} \zeta_t + \mathbf{D} \hat{\mathbf{x}}_t$$

$$\mathbf{x}_t = \hat{\mathbf{x}}_t + \hat{\Delta}_t$$

End For

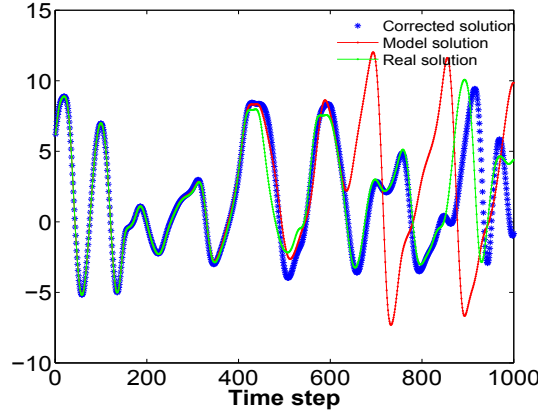


Figure 1. The evolution of a Lorenz variable in time for model solution, real solution and corrected model solution. The corrected model solution of the Lorenz system is obtained using the predicted global model error. The 12th and 30th variables in the model are perturbed, meaning that they are affected by the structural uncertainty.

The SSM coefficient matrices are shown in Figure 2 when perturbing only the 12th and 30th variables in the Lorenz system. The output matrix $\mathbf{C}$ identifies well the source of uncertainty: the larger entries correspond to the 12th and 30th variables. The state matrix $\mathbf{A}$ has the largest entries near the diagonal, indicating that global error correlations are spatially localized. The input matrix $\mathbf{B}$ shows no special structure that reveals the uncertainty location. The feed forward matrix $\mathbf{D}$ also captures the perturbed states.

In the second experiment the structural uncertainty affects the 8th, 16th, 25th, and 33rd variables in the Lorenz system (18). Figure 3 shows the SSM coefficient matrices. The sensor matrix $\mathbf{C}$ again identifies well the perturbed variables, with entry peaks corresponding to the perturbed variables. The state matrix $\mathbf{A}$ has the largest entries near the diagonal, indicating that global error correlations are spatially localized. The input matrix $\mathbf{B}$ does not have any special structure. The feed forward matrix $\mathbf{D}$ also captures the perturbed states.

Furthermore, the length of the the time series used in SSM for tuning the parameters, affects the prediction power. The more data used, the parameters of SSM

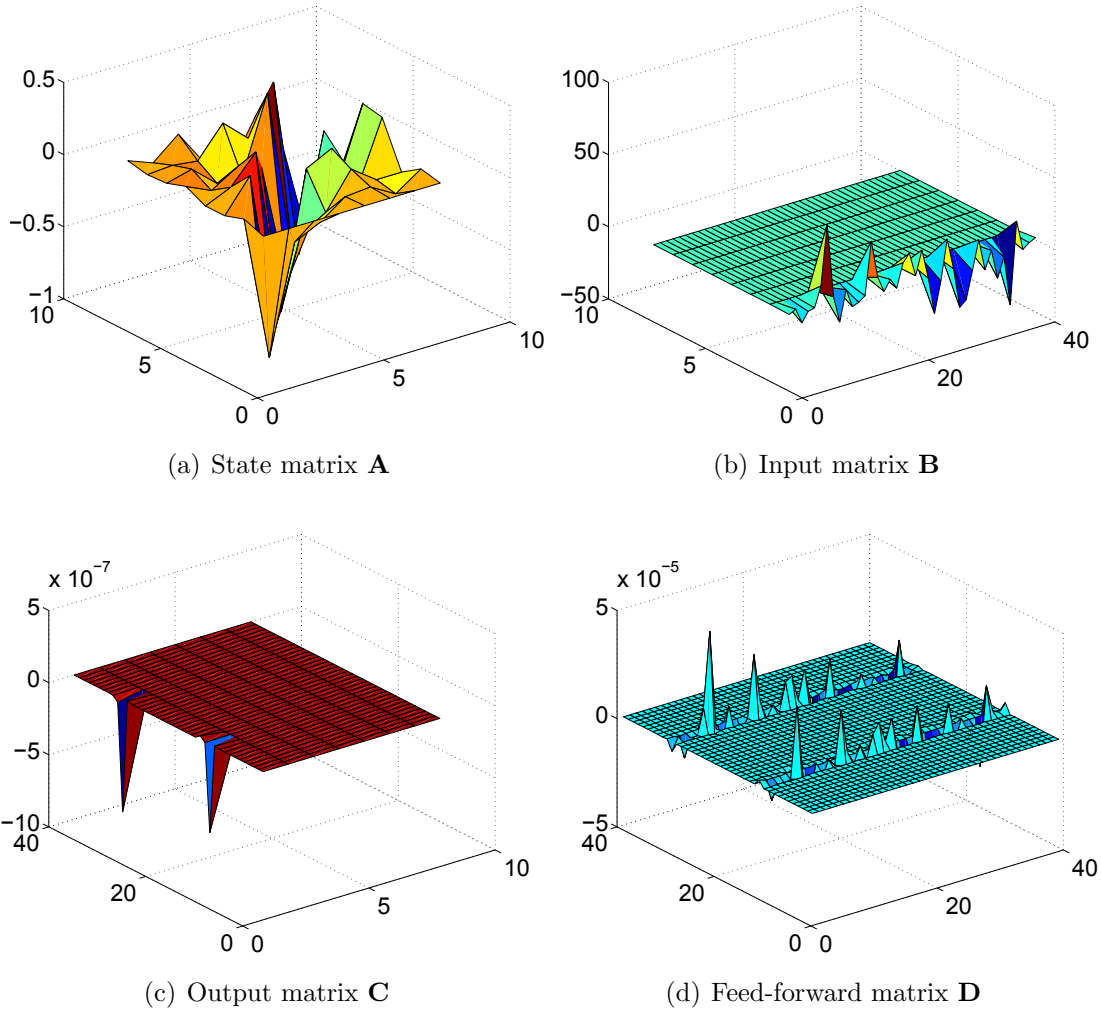


Figure 2. The structure of the SSM coefficient matrices for the global error model. The structural uncertainty affects the 12th and 30th variables in the Lorenz model (18).

gets better tuned and the hidden dynamics of the error will be identified better. Table (1.a) illustrates the predictability power of the SSM using different length of time series. The variables we use also impacts the performance of SSM to great extent. Table (1.b) also shows the predictability power of the SSM using different number of variables.

(a) Effect of length of time-series on SSM performance		(b) Effect of using different variable numbers on SSM performance	
Length of time-series	RMSE	Number of variables	RMSE
1000	2.2449	40	2.2449
800	3.0075	20	3.5834
400	4.1406	10	5.2142

Table 1. RMSE between corrected model solution and real data

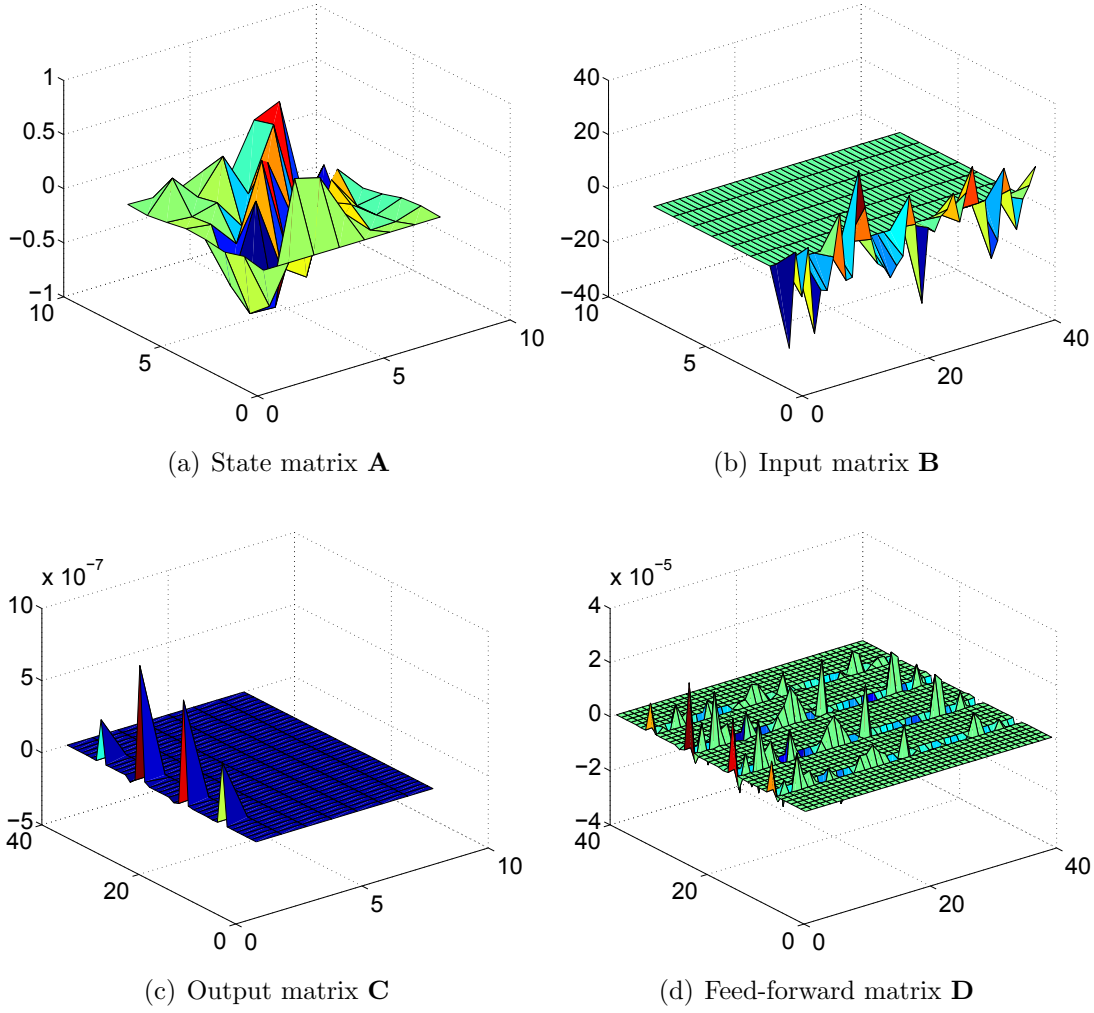


Figure 3. The structure of the SSM coefficient matrices for the global error model. The structural uncertainty affects the 8th, 16th, 25th and 33rd variables in the Lorenz model (18).

3.1.2. Local structural uncertainty modeling results In the local approach all the assumptions made for the global approach also hold. The difference is that we construct the SSM representing the uncertainty in variable i using only a subset of variables in the physical vicinity of i , and do not include all variables in the system. Therefore, for local error modeling one needs to determine the size of the local neighborhood where errors are correlated with the error in variable i . In case of the Lorenz model (18) the physical distance between variables equal the difference between their indices, modulo 20 (due to the periodic boundary conditions). Therefore one needs to determine the decorrelation distance L such that the local error model for variable i uses information from variables indexed $i - L$ to $i + L$.

For the first set of experiments we set the localization radius to $L = 1$. The 22nd variable is perturbed, and we use the model predictions $\hat{\mathbf{x}}$ of variables 21, 22, 23 to build the local error SSM. The corrected model solution, obtained according to (17), is shown

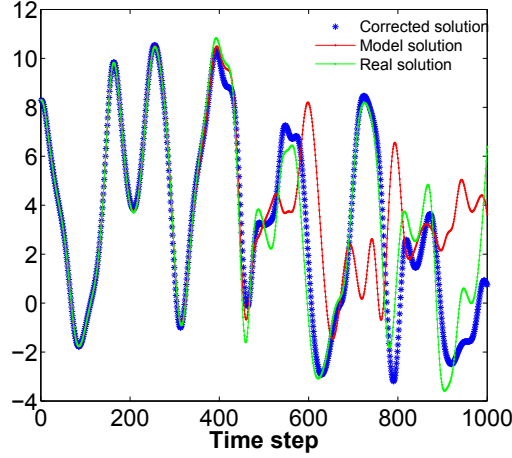


Figure 4. The evolution of a Lorenz variable in time for model solution, real solution and corrected model solution. The corrected model solution is obtained using the predicted local error. The structural uncertainty affects the 22nd variable in the model, and the localization radius is one.

in Figure 4. Figure 5 shows the matrix coefficients for the local error SSM. Since we only included three variables in the SSM the dimension of the output matrix $\mathbf{C}$ is 3×3 ; the largest entry corresponds to the perturbed variable. The state matrix $\mathbf{A}$ is nearly the identity matrix, the input matrix $\mathbf{B}$ does not have any special structure, and the feed-forward matrix $\mathbf{D}$ is zero. However, the corrected solution is closer to the “truth”, as shown in Figure 4.

In the next experiment we set the localization radius to $L = 5$. Figure 6 shows the corrected solution using a local error SSM model with 11 variables. The corrected solution is much closer to the truth than the one obtained for $L = 1$, see also Figure 4. This implies that an accurate corrected model solution depends on a good choice of the localization radius; in this experiment including more variables is beneficial. Figure 7 shows the SSM matrix coefficients for this experiment. The sensor matrix $\mathbf{C}$ has a peak in the middle variable corresponding to the perturbed variable. The structure of the rest of the matrices are very similar to the experiment where the localization radius is three.

However, in practice, the real source of uncertainty is unknown and we might not have sufficient knowledge about the states of the model that are affected by large structural uncertainties. One can find the most uncertain variables by iterating over subsets of model states, as summarized in Algorithm 1. At each iteration we assume that the source of uncertainty affects one particular subset of variables. A local error SSM is constructed, the corrected solution is computed, and the root mean square of the difference between the corrected solution and the real trajectory is found. The subsets of variables corresponding to the minimum root mean square differences are considered the most likely to be affected by structural uncertainty, since local corrections applied to those variables results in most improvements of the overall model predictions.

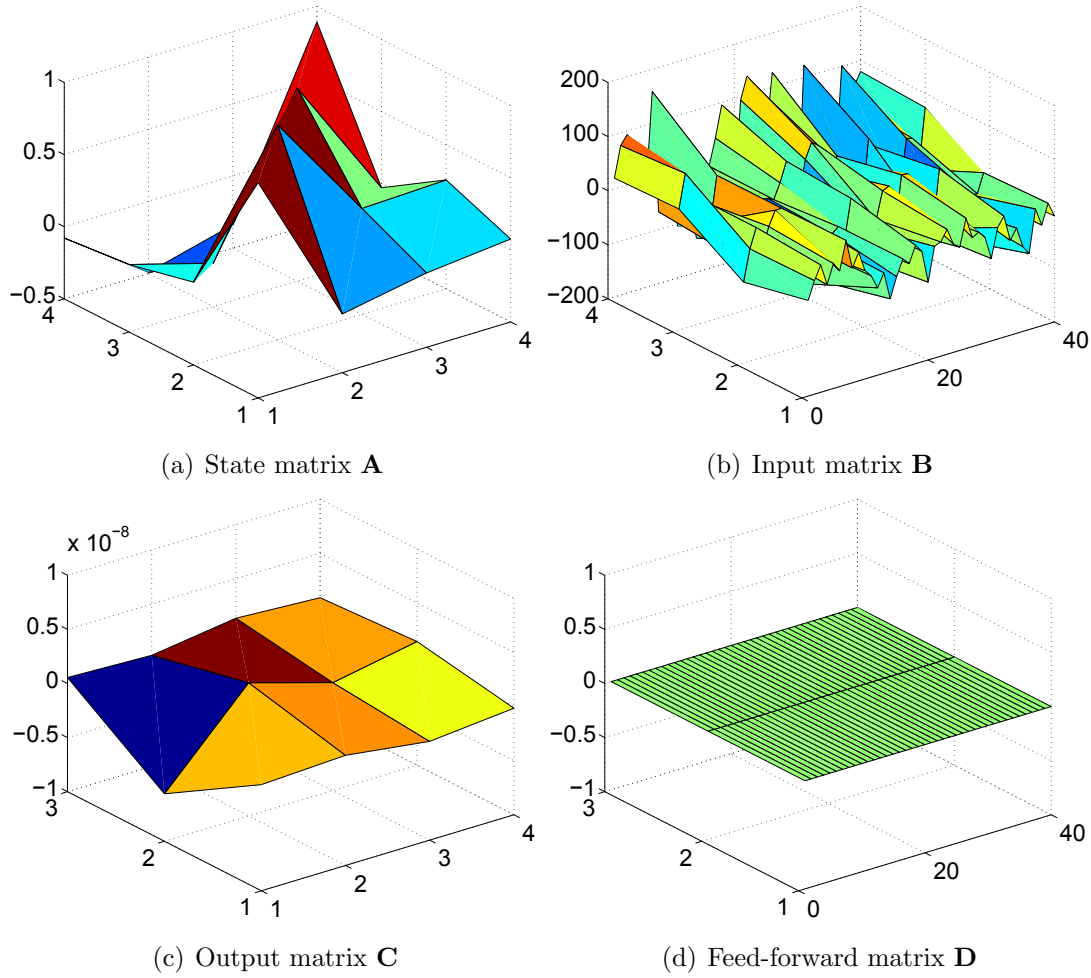


Figure 5. The local error SSM coefficients for the Lorenz system (18). The structural uncertainty affects the 22nd variable in the model, and the localization radius is one.

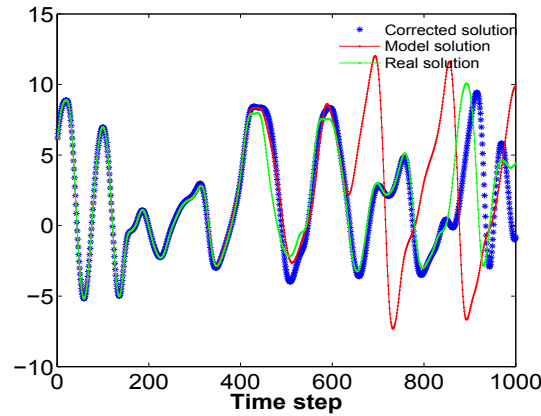


Figure 6. The evolution of a Lorenz variable in time for model solution, real solution and corrected model solution. The corrected model solution is obtained through applying the local model discrepancy. The structural uncertainty affects the 22nd variable in the model, and the localization radius is five.

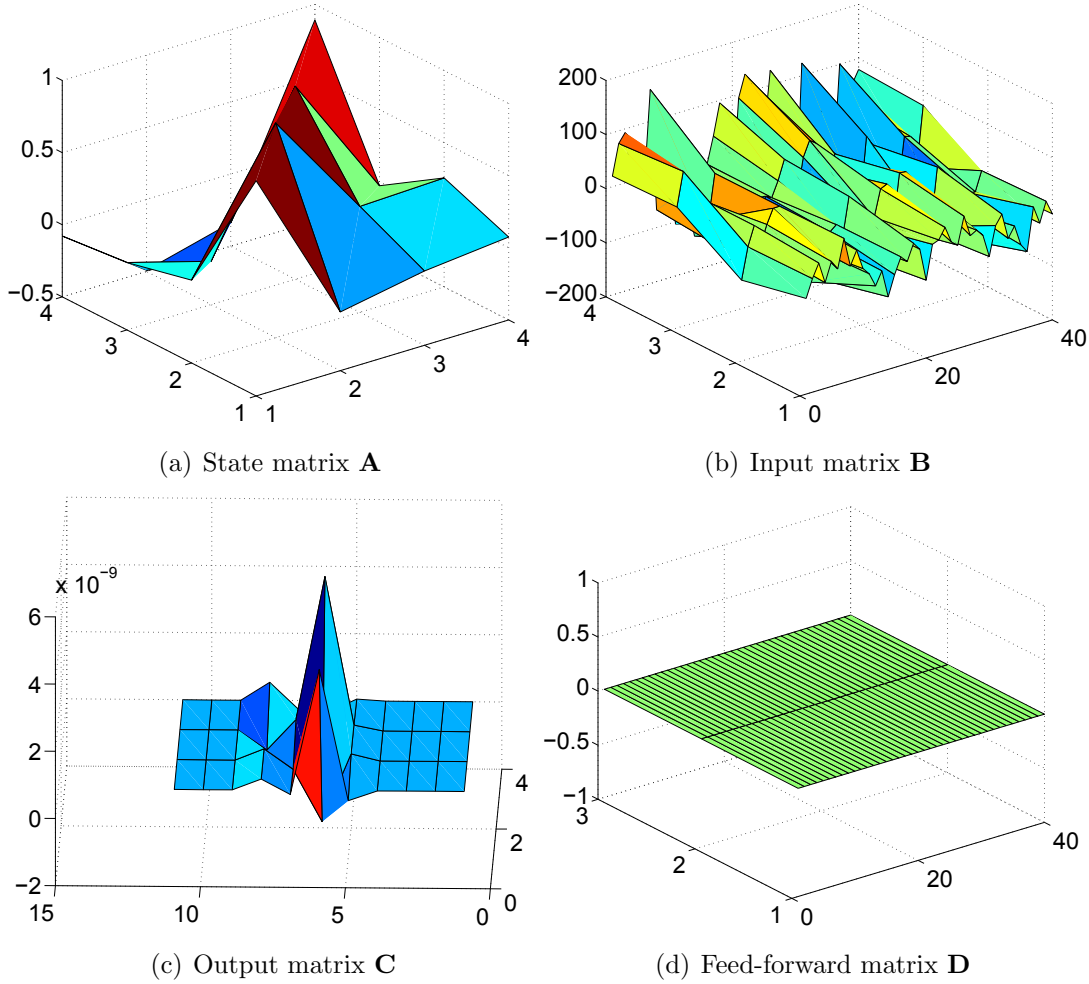


Figure 7. The local error SSM coefficients for the Lorenz system (18). The structural uncertainty affects the 22nd variable in the model, and the localization radius is five.

In the third experiment we perturb one variable in the Lorenz system. We iterate over all variables in the Lorenz system, and for each we apply the local approach with a given localization radius. The differences between the corrected solution and the real trajectory are obtained. We consider the root mean square errors between the vectors of solutions at each time moment along the trajectory. The histogram of minima of the root mean square differences will show for which variable corrections result in the minimum error. This helps identify the potential source of uncertainty in the system. Figure 8 shows the histogram of root mean square differences between corrected solution and real data after perturbing the 30th variable. We iterate over all variables in the Lorenz system, and for each we apply the local approach with a localization radius of three. The frequency of the minimum of the discrepancy between model solution and real data for the 30th variable is higher than for other variables. In another experiment we perturb the 14th variable in the Lorenz system with the localization radius of five. Figure 8 shows the histogram of discrepancies perturbing the 14th variable; again, the frequency of the minimum of the root mean square differences between model solution and real

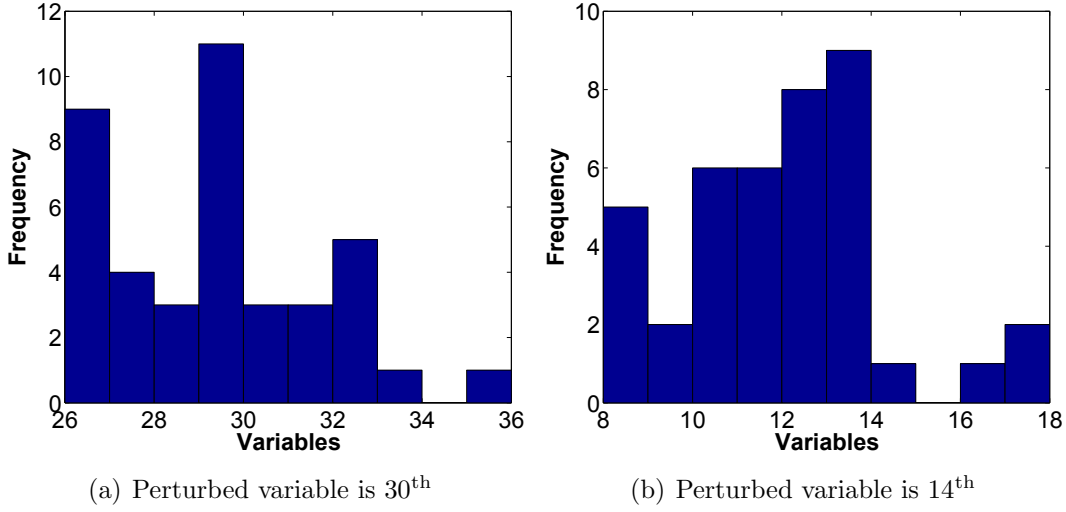
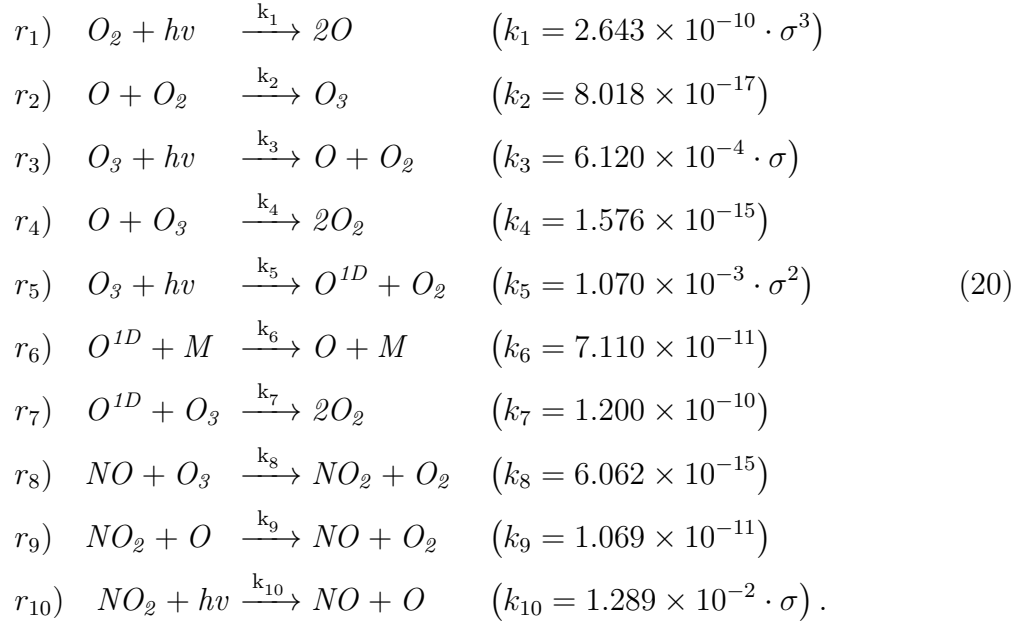


Figure 8. Histogram of minimum root mean square discrepancies between the corrected solution and the exact trajectory using the local error modeling approach.

data for the 14th variable is higher than for other variables, which indicates that the source of uncertainty is well identified.

3.2. Stratospheric chemistry model

A basic stratospheric chemistry mechanism [13] is given by the following set of reactions:



Here $M = 8.120E + 16 \text{ molecules/cm}^3$ is the atmospheric number density. The rate coefficients are scaled for time expressed in seconds, and $\sigma(t) \in [0, 1]$ represents the normalized sunlight intensity during a diurnal cycle [13].

In our experiments we consider the full set of equations (20) as describing the true physical system (6a). Our model (3a) is the set of reactions $r_1 \dots r_7$; not captured by

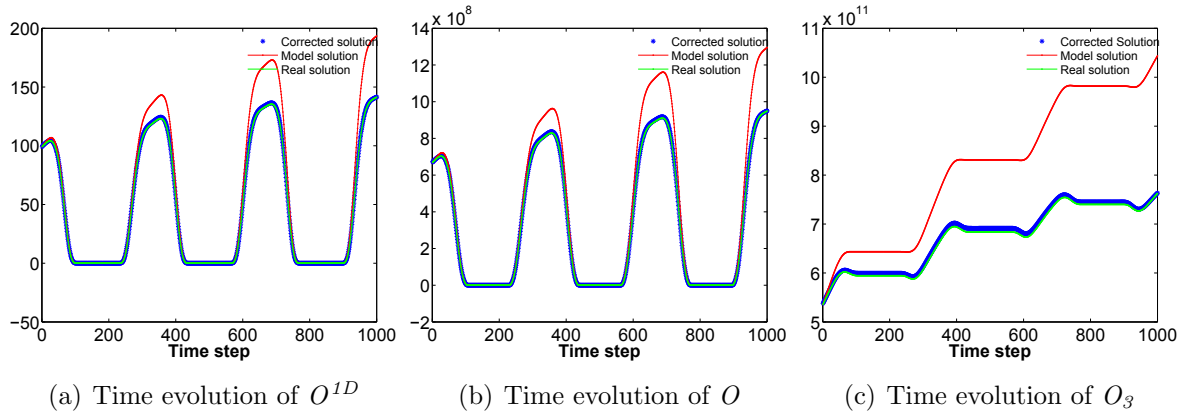


Figure 9. Model forecasts and corrected model solutions for the stratospheric chemical system with missing reactions. A global model of structural uncertainty is used.

the model are those reactions r_8, r_9, r_{10} that involve nitrous oxides (NO, NO_2). The three missing reactions correspond to missing dynamics in the model, i.e., physical phenomena that exist in reality but are not captured by the computer model. Our goal is to understand what is the missing dynamics in the chemical system that leads to systematic differences between the model and reality. In order to tackle this we construct a global error model (12) that estimates the global discrepancy between the system of full reactions and the system with three missing reactions. The assumptions made for the Lorenz model also hold here, i.e. we don't have observation error and disturbance, and we use the identity observation operator, therefore according to (15) we have $\delta_t \approx \hat{\Delta}_t$.

Figure 9 shows the solutions of the reference system, as well as the original and the corrected solutions of the model with missing physics. Correcting the solutions of the model with missing dynamics allows to recover very well the trajectories of the physical (reference) model. Figure 10 shows the coefficient matrices of the global SSM for the stratospheric model. There is no particular structure associated with these matrices. While the linear SSM dynamics can mimic the evolution of structural errors in the nonlinear chemical system (20), it cannot explain the real source of discrepancies as coming from three missing reactions.

Having the corrected solution, one is interested to learn about the number of missing chemical reactions as well as their corresponding reaction rates; the missing chemical reactions represent the missing dynamics in the model. Assuming we know what the missing reactions are, the corresponding reaction rates can be obtained through an optimization process that adjusts their values such as to minimize the model-observation discrepancies. In this experiment we carry out the optimization using a genetic algorithm (GA). The reason for using the GA approach rather than a traditional gradient-based approach is that computing the Jacobian of this system is very expensive. GAs are a class of evolutionary algorithms that is used both for constrained and unconstrained optimization problems. A population of candidate solutions is generated randomly in

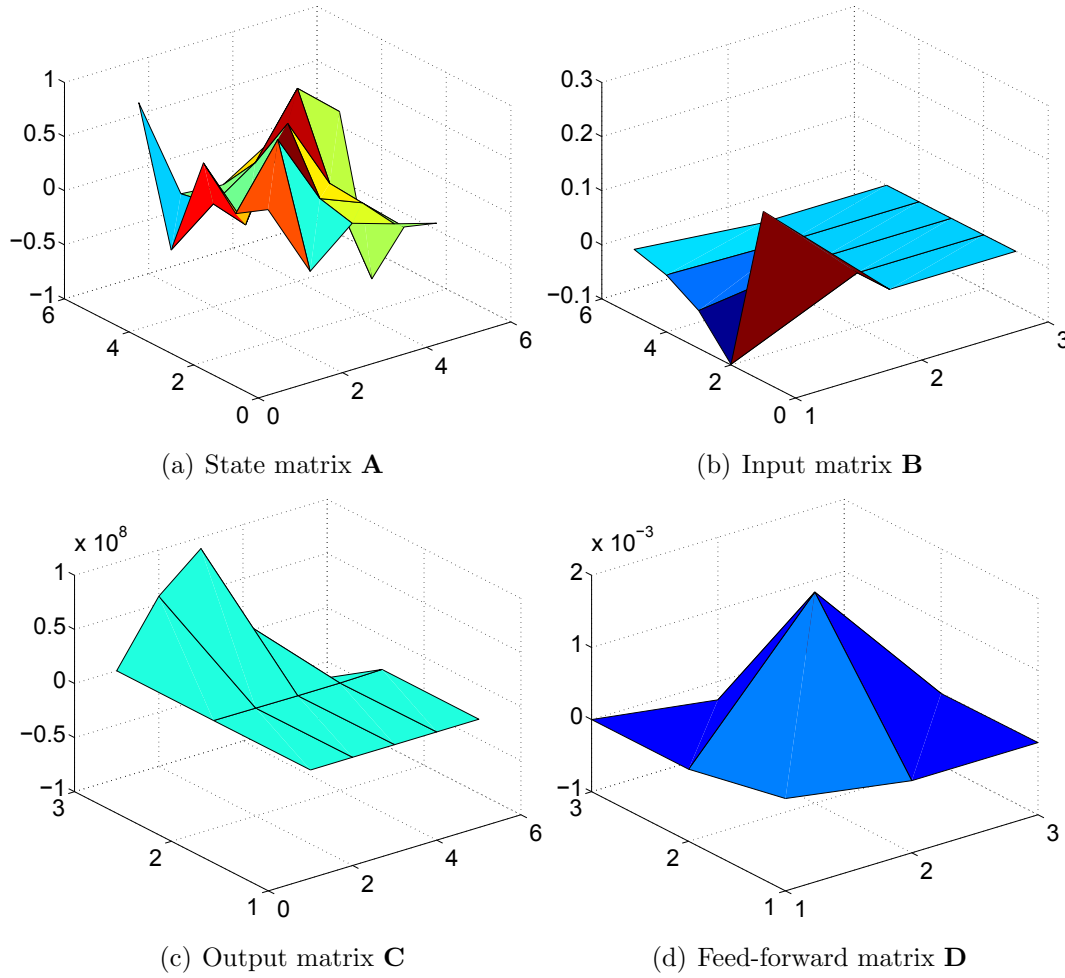


Figure 10. The structure of the matrix coefficients of the global-error SSM for the stratospheric chemical system.

the domain of search space and the fitness of each individual solution is evaluated using the value of the objective function in the optimization problem. In an iterative process the population of candidate solutions then evolves toward the best solution in the search space. At each iteration the properties of the candidate solutions are changed randomly by cross-over and mutation (operations inspired by biological evolution) to ensure the dispersion of the possible solutions throughout the search space. The best fitted solutions will be selected to form the next generation of feasible solutions. The algorithm is stopped when the maximum number of iterations is reached, or the highest fitness value was obtained, or successive iterations do not improve the result any further [15].

For this experiment we use Matlab's GA toolbox with a population size of 30, 50 iterations, cross over rate of 0.85, and a mutation rate of 0.05. The fitness function is the root mean square difference between the true physics (solution of the stratospheric model with all reactions) and the solution of the model (stratospheric model with three reactions having uncertain reaction coefficients). With tight upper bounds and

lower bounds for the reaction rates of the missing reactions, the optimization finds the reaction rate coefficients for r_8, r_9, r_{10} to have the values 7.781×10^{-15} , 1.937×10^{-11} , and 3.4621×10^{-2} , respectively. The absolute differences between the approximated values and the true values of the rate coefficients are small, 1.719×10^{-15} , 8.687×10^{-12} , and 0.021, respectively. This experiment illustrates the fact that any information we have about the source of structural uncertainty (e.g., three missing chemical reactions) can and should be used to obtain a more accurate, physically based description of the associated model errors.

4. Conclusions

Physics-based computer models of the real world are imperfect since not all physical processes that drive reality are fully captured by the model. This work considers the structural model uncertainty, i.e., the model inadequacy due to missing aspects of the dynamics of the true physical system. We study structural uncertainty based on the information provided by the discrepancy between the model solution and the true state of the physical system, as measured by the available observations.

The proposed approach is to approximate the dynamics of structural errors using linear state-space models. The parameter matrices of these models are obtained from fitting their response to match the mapping of model states as inputs to the observed model-reality differences as outputs.

Two different strategies to model the error dynamics are discussed: the global approach and the local approach. The global approach seeks to build one state-space model that approximates the entire structural uncertainty over the model’s entire state-space. The local approach uses the ansatz that structural errors are only correlated locally, and that the correlation decreases with increasing distance between model components. Low-dimensional local state-space models are constructed to approximate the dynamics of the local errors, and they use only information from subsets of variables and data.

Numerical experiments are carried out with two test problems, the Lorenz-96 system and a stratospheric chemistry model. These experiments reveal that the state-space discrepancy models provide intuition about the variables affected most by the missing dynamics. Global error models can identify the sources of uncertainty inside the physics-based models. Local error models capture the evolution of uncertainty over subsets of variables and are considerably less expensive to construct. When there is insufficient knowledge about the source of structural uncertainty the local approach allows to locate the subsets of variables that are most affected by it. The structural errors estimated with both global and local approaches can be used to correct the model solution and obtain improved estimates of the true physics. However, even if the state-space models can reproduce well the dynamics of the error, they may do this using an internal dynamics that is unrelated to the model dynamics or the true physics. In particular, reproducing the error dynamics does not necessarily explain the missing physics, as it was illustrated

in the experiments with the stratospheric chemistry system. Consequently, in order to fully understand structural errors it is important to incorporate in the analysis all available information about the nature of the missing physical processes.

Acknowledgment

This work was funded in part by the AFOSR DDDAS 15RT1037 award and by the Computational Science Laboratory at Virginia Tech.

References

- [1] Pilch M, Trucano T G and Helton J C 2006 Unlimited Release SAND2006-5001, Sandia National Laboratory, Albuquerque, New Mexico **87185** 2
- [2] Kennedy M C and O'Hagan A 2001 Journal of the Royal Statistical Society: Series B (Statistical Methodology) **63** 425–464
- [3] Chantrasmi T and Iaccarino G 2012 International Journal for Uncertainty Quantification **2**
- [4] Refsgaard J C, Van der Sluijs J P, Brown J and Van der Keur P 2006 Advances in Water Resources **29** 1586–1597
- [5] Briggs A H, Weinstein M C, Fenwick E A, Karnon J, Sculpher M J and Paltiel A D 2012 Medical Decision Making **32** 722–732
- [6] Strong M and Oakley J E 2014 SIAM/ASA Journal on Uncertainty Quantification **2** 106–125
- [7] Higdon D, Kennedy M, Cavendish J C, Cafo J A and Ryne R D 2004 SIAM Journal on Scientific Computing **26** 448–466
- [8] Ajami N K, Duan Q and Sorooshian S 2007 Water Resources Research **43**
- [9] Hoshiya M and Saito E 1984 Journal of Engineering Mechanics **110** 1757–1770
- [10] Seiler A, Rivenæs J C, Aanonsen S I and Evensen G 2009 Structural uncertainty modelling and updating by production data integration SPE/EAGE Reservoir Characterization & Simulation Conference
- [11] Rowell D 2002 URL: <http://web.mit.edu/2.14/www/Handouts/StateSpace.pdf>
- [12] Lorenz E N 1996 Predictability: A problem partly solved Proc. Seminar on predictability vol 1
- [13] Sandu A 2002 IMA VOLUMES IN MATHEMATICS AND ITS APPLICATIONS **130** 21–38
- [14] Van Overschee P and De Moor B 2012 Subspace identification for linear systems: Theory—Implementation—Applications (Springer Science & Business Media)
- [15] Melanie M 1999 Cambridge, Massachusetts London, England, Fifth printing **3** 62–75