Error Estimation for Reduced Order Models of Dynamical Systems

Linda Petzold and Chris Homescu
University of California Santa Barbara

Radu Serban
Lawrence Livermore National Laboratory

Objective and Approach

OBJECTIVE

- To judge the quality of the reduced model of a dynamical system by estimating its error and region of validity.

APPROACH

- The overall approach is general. Here we concentrate on Proper Orthogonal Decomposition (POD).

- Estimates and bounds of the reduced model errors are obtained using a combination of small sample statistical condition estimation and error estimation using the adjoint method.

- This approach allows the assessment of regions of validity, i.e., ranges of perturbations in the original system over which the reduced model is still appropriate.

POD for Dynamical Systems

- The *PROPER ORTHOGONAL DECOMPOSITION* (POD) reduction is the most efficient choice among linear decompositions in the sense that it retains, on average, the greatest possible kinetic energy.

- It provides the best approximating affine subspace to a given set of time snapshots of the solution, collected into an observation matrix

$$\boldsymbol{Y} = \left[\boldsymbol{y}(\boldsymbol{t}_1) - \bar{\boldsymbol{y}}, \ldots, \boldsymbol{y}(\boldsymbol{t}_m) - \bar{\boldsymbol{y}}\right]$$

where $\bar{\boldsymbol{y}}$ is the mean of these observations.

- POD seeks a subspace S and the corresponding projection matrix P, so that the total square distance ||Y – PY|| is minimized.

POD for Dynamical Systems (continued)

- Using the singular value decomposition (SVD) of the observation matrix, the projection matrix corresponding to the optimal POD subspace S is obtained as $P = \rho\rho^T \in R^{n\times n}$, where ρ is the matrix of projection onto S, the subspace spanned by the reduced basis obtained from the SVD.

- The matrix $\rho \in R^{n\times k}$ consists of the columns V_i (i=1,...,k), the singular vectors corresponding to the k largest singular values.

- The error of the projection is given by

$$\min_S \|Y - PY\|^2 = \sum_{j=k+1}^{n} \lambda_j$$

POD for Dynamical Systems (continued)

- A POD-based reduced model for

$$\frac{dy}{dt} = f(t, y, p), \quad y(t_0) = y_0$$

can be constructed by projecting onto S the vector field f(s,t) at each point $s \in S$.

- The reduced model in subspace coordinates is:

$$\frac{dy^S}{dt} = \rho^T f(t, \rho y^S + \bar{y}, p), \quad y^S(t_0) = \rho^T (y_0 - \bar{y})$$

- The reduced model in full space coordinates is:

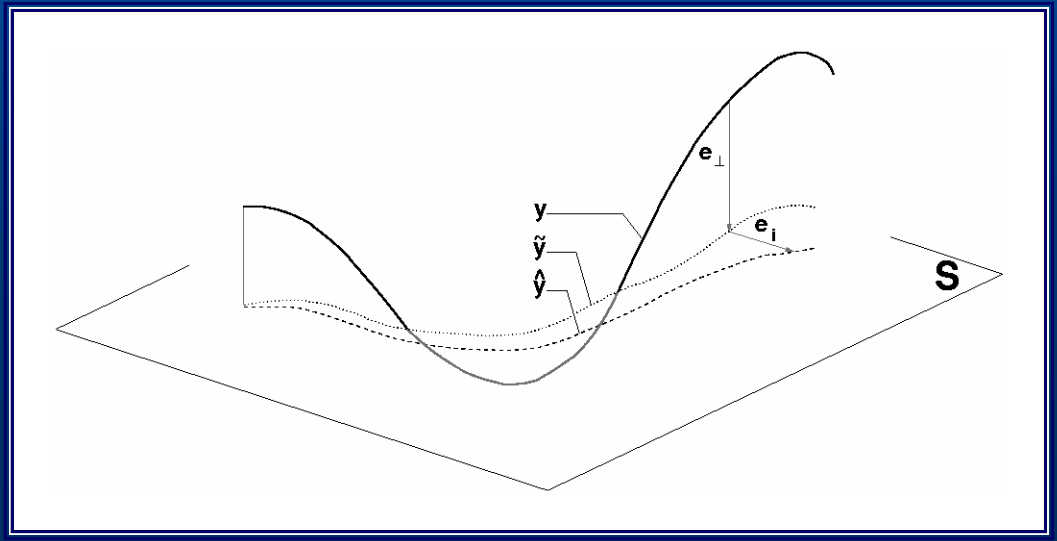
$$\frac{d\hat{y}}{dt} = Pf(t, \hat{y}, p), \quad \hat{y}(t_0) = P(y_0 - \bar{y}) + \bar{y}$$

Error of the POD Approximation

Let $\hat{y}(t)$ be the solution of the POD-reduced model, and $\tilde{y}(t)$ the projection onto S of the solution $y(t)$ of the original problem.

The total approximation error can be split into

- subspace approximation error $\quad e_{\perp}(t) = \tilde{y}(t) - y(t)$
- integration error in the subspace S $\quad e_{i}(t) = \hat{y}(t) - \tilde{y}(t)$



Small Sample Statistical Method for Condition Estimation (SCE)

- For any vector $v \in R^n$, if z is selected uniformly and randomly from the unit sphere S_{n-1}, the expected value is $E\left(\left|z^T v\right|\right) = W_n \|v\|$
- We estimate the norm ||v|| using the expression

$$\xi = \frac{\left|z^T v\right|}{W_n}, \text{ where } W_n \approx \sqrt{\frac{2}{\pi(n-1/2)}} \text{ are the Wallis factors.}$$

- For additional random vectors $z_2,\ldots,z_q$, the estimate

$$\xi_q = \frac{W_q}{W_n}\sqrt{\left|z_1^T v\right|^2 + \left|z_2^T v\right|^2 + \ldots + \left|z_q^T v\right|^2}$$

satisfies

$$\Pr\left(\frac{\|v\|}{\gamma} \le \xi_2 \le \gamma\|v\|\right) \approx 1 - \frac{\pi}{4\gamma^2}$$

$$\Pr\left(\frac{\|v\|}{\gamma} \le \xi_3 \le \gamma\|v\|\right) \approx 1 - \frac{32}{3\pi^2\gamma^3}$$

...

SCE for Estimation of Errors Due to Model Reduction

- For the norm $||e(t_f)||$ of the error vector, the quantities $z^T_j e(t_f)$ (for some random vector z_j selected uniformly from S_{n-1}) are computed using an adjoint model.

- While for one given ODE system the forward model is most efficient for estimating the norm of the error, the combination of the adjoint approach and the SCE can be used to estimate the region of validity of a reduced model, using the concept of “condition number” for the error equation corresponding to each perturbation.

- Although these estimates provide only approximate upper bounds for the norms of the errors, they have the advantage of allowing *a-priori* estimates of the errors induced by perturbations.

Estimation of the Total Approximation Error

- To first order approximation, the total error satisfies

$$\frac{de}{dt} = J(\hat{y},t)e - Qf(\hat{y},t), \quad e(t_0) = -Q(y_0 - \bar{y})$$

where **J** is the Jacobian of **f**, and **Q = I - P**. Using

$$z^T e(t_f) = -\int_{t_0}^{t_f} \lambda^T(s) Qf(\hat{y}(s),s)\,ds - \lambda^T(t_0) Q(y_0 - \bar{y})$$

where z is a random vector uniformly selected from the unit sphere $S_{n\text{-}1}$ and λ is the solution of the adjoint system

$$\frac{d\lambda}{dt} = -J^T(\hat{y},t)\lambda, \quad \lambda(t_f) = z$$

we obtain the SCE for the norm of the total error

$$\|e(t_f)\| \approx \frac{W_q}{W_n} \sqrt{\sum_{j=1}^{q} \left| \int_{t_0}^{t_f} \lambda^T(s) Qf(\hat{y}(s),s)\,ds + \lambda^T(t_0) Q(y_0 - \bar{y}) \right|^2}$$

Estimation of the Subspace Integration Error

- In the S coordinate system, where $e_i^S = \rho^T e_i$ and $e_i = \rho\, e_i^S$, the subspace integration error satisfies to first order

$$\frac{d\boldsymbol{e}_i^S}{dt} = \rho^T \boldsymbol{J}(\hat{\boldsymbol{y}},t)\rho \boldsymbol{e}_i^S + \boldsymbol{J}(\hat{\boldsymbol{y}},t)\boldsymbol{e}_\perp, \quad \boldsymbol{e}_i^S(t_0) = 0$$

- For a random vector z^S uniformly selected from the unit sphere $S_{k\text{-}1}$

$$\left(\boldsymbol{z}^S\right)^T \boldsymbol{e}_i^S(t_f) = \int_{t_0}^{t_f} \mu_{\hat{y}}^T(s)\rho^T \boldsymbol{J}(\hat{\boldsymbol{y}}(s),s)\boldsymbol{e}_\perp(s)\,ds$$

where $\mu_{\hat{y}}$ solves the adjoint system $\dfrac{d\mu_{\hat{y}}}{dt} = -\rho^T \boldsymbol{J}^T(\hat{\boldsymbol{y}},t)\rho\, \mu_{\hat{y}}, \quad \mu_{\hat{y}}(t_f) = \boldsymbol{z}^S$

- The SCE estimate for the norm of the subspace integration error is

$$\left\|\boldsymbol{e}_i(t_f)\right\| \approx \frac{W_q}{W_n}\sqrt{\sum_{j=1}^{q}\left|\int_{t_0}^{t_f} \mu_{\hat{y}}^T(s)\rho^T \boldsymbol{J}(\hat{\boldsymbol{y}}(s),s)\boldsymbol{e}_\perp(s)\,ds\right|^2}$$

Condition Number for the Subspace Integration Error

- For a unit vector z_j^S we define $\theta^j(s) = J^T(\hat{y}(s), s)\rho\mu_{\hat{y}}(s)$

  and $\kappa_j(e_i) = \left\|\theta^j\right\|_{L_1} = \int_{t_0}^{t_f} \sum_{i=1}^{n} \left|\theta_i^j(s)\right| ds$

- We have $\left\|e_i(t_f)\right\| \le \kappa(e_i) \cdot \left\|e_\perp\right\|_{L_\infty}$

  where $\kappa(e_i) = \frac{W_q}{W_n} \sqrt{\sum_{j=1}^{q} \kappa_j^2(e_i)}$ is the “condition number”

  for the subspace integration error

- For the norm of the projection error $e_\perp$ we have

$$\left\|e_\perp\right\|_{L_\infty} \le \left\|e_\perp\right\|_{L_2} = \sqrt{\sum_{j=k+1}^{n} \lambda_j} \equiv ErrPOD$$

Estimation of Regions of Validity

- Consider perturbations to initial conditions. Other perturbations are treated similarly.
- Let $\boldsymbol{Y}$ be the solution of the ODE obtained by applying an IC perturbation and $\hat{\boldsymbol{Y}}$ the solution of the corresponding POD-reduced model, with P based on y. Defining $\boldsymbol{E}_1 = \hat{\boldsymbol{Y}} - \boldsymbol{Y}$ and $\Delta(\boldsymbol{t}) = \boldsymbol{E}_1(\boldsymbol{t}) - \boldsymbol{e}(\boldsymbol{t})$,

  we have $\left| \|\boldsymbol{e}(t_f)\| - \|\Delta(t_f)\| \right| \le \|\boldsymbol{E}_1(t_f)\| \le \|\boldsymbol{e}(t_f)\| + \|\Delta(t_f)\|$
- The error Δ is split into $\Delta_\perp$ (orthogonal to **S**) and Δ_i (parallel to S).

  An SCE estimate of the norm of $\Delta_\perp(t_f)$ is given by

$$\|\Delta_\perp(t_f)\| \approx \frac{W_q}{W_n}\sqrt{\sum_{j=1}^{q}\left|z_j'^T \Delta_\perp(t_f)\right|^2} = \frac{W_q}{W_n}\sqrt{\sum_{j=1}^{q}\left|\lambda_{\hat{y}}^T(t_0)\,\delta y_0\right|^2} \le \kappa(\Delta_\perp)\cdot\|\delta y_0\|$$

where $\lambda_{\hat{y}}$ satisfies $\quad \dfrac{d\lambda_{\hat{y}}}{dt} = -J^T(\hat{y},t)\,\lambda_{\hat{y}}\,, \quad \lambda_{\hat{y}}(t_f) = Qz_j'$

Estimation of Regions of Validity (continued)

- The "condition number" $\kappa(\Delta_\perp)$ is defined as

$$\kappa(\Delta_\perp) = \frac{W_q}{W_n}\sqrt{\sum_{j=1}^{q}\kappa_j^2(\Delta_\perp)}, \quad \kappa_j(\Delta_\perp) = \left\|\lambda_{\hat{y}}(t_0)\right\|_2$$

- To first order, $\Delta_i(t) = 0$, i.e., a perturbation to the initial conditions of the original ODE does not introduce additional subspace integration errors. Thus

$$\left|\left\|e(t_f)\right\| - \left\|\Delta_\perp(t_f)\right\|\right| \le \left\|E_1(t_f)\right\| \le \left\|e(t_f)\right\| + \kappa(\Delta_\perp)\cdot\left\|\delta y_0\right\|$$

Numerical Results

- The estimates (and bounds), obtained with q=1,2,3, where q is the number of orthogonal vectors used by the SCE, are shown as follows:
  - Total approximation error: $\left\| \boldsymbol{e}(\boldsymbol{t}_f) \right\| = \left\| \hat{\boldsymbol{y}}(\boldsymbol{t}_f) - \boldsymbol{y}(\boldsymbol{t}_f) \right\|$
  - Error bound for $\left\| \boldsymbol{E}_1^{ic}(\boldsymbol{t}_f) \right\|$, as predicted by the condition number $\kappa(\Delta_\perp^{ic})$
- Examples: linear advection-diffusion and pollution model

Linear Advection-Diffusion Model

$$u_t = p_1 u_{xx} + p_2 u_x$$

$$B.C. \quad u(0,t) = u(2,t) = 0.0$$

$$I.C. \quad u(x,0) = u_0(x) = x(2-x)e^{2x}$$

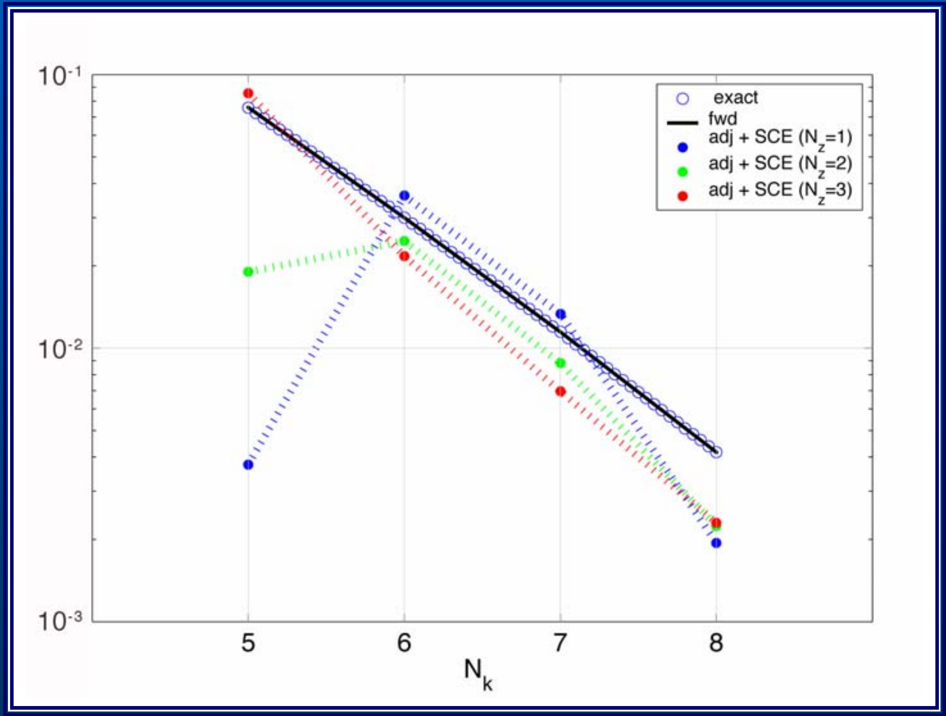
- With $y_i(t) = u(x_i,t)$, central differencing, and eliminating boundary values, we obtain n ODEs

$$\frac{dy_i}{dt} = p_1 \frac{y_{i+1} - 2y_i + y_{i-1}}{\Delta x^2} + p_2 \frac{y_{i+1} - y_{i-1}}{2\Delta x}, \quad y_i(0) = u_0(x_i)$$

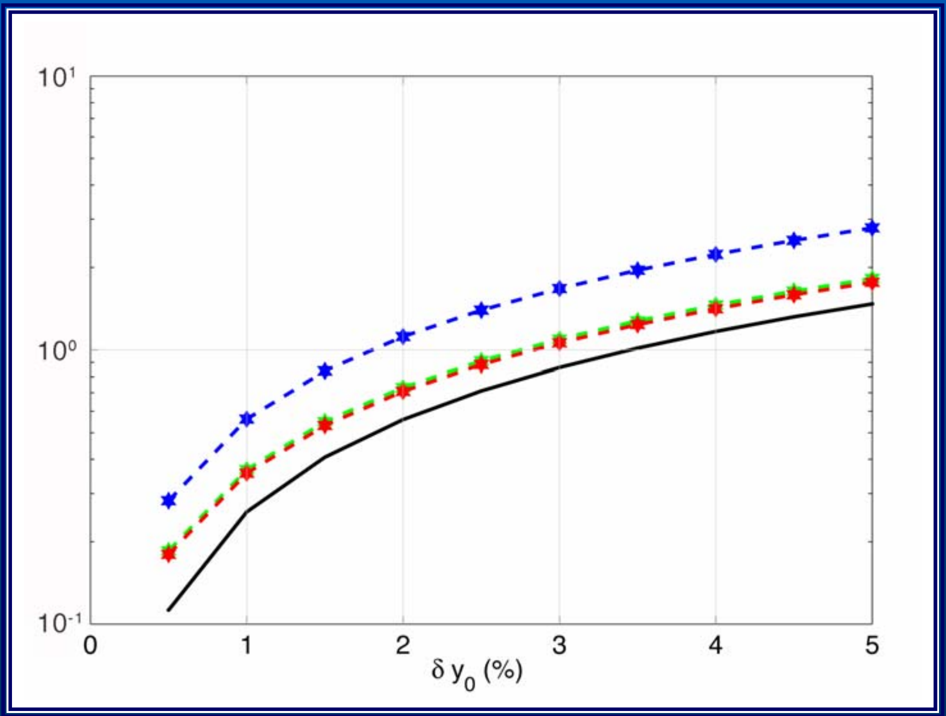
- The problem parameters were p_1=0.5, p_2=1.0, and n=100
- The POD projection matrices were based on 100 data points equally spaced in the interval $[t_0, t_f] = [0.0, 0.3]$.

Advection-Diffusion Model Approximation Errors

Total Error



Approximation error E_1 as a function of the IC perturbation k=5



The solid (black) lines represent the corresponding norms computed by the forward integration of the error equations.
The dotted (colored) lines describe SCE estimates.
The dashed (colored) lines represent the bounds of the errors.
The (blue) line made of circles represents the norm of the "exact error," $\boldsymbol{e}(t_f) = \hat{\boldsymbol{y}}(t_f) - \boldsymbol{y}(t_f)$

Pollution Model

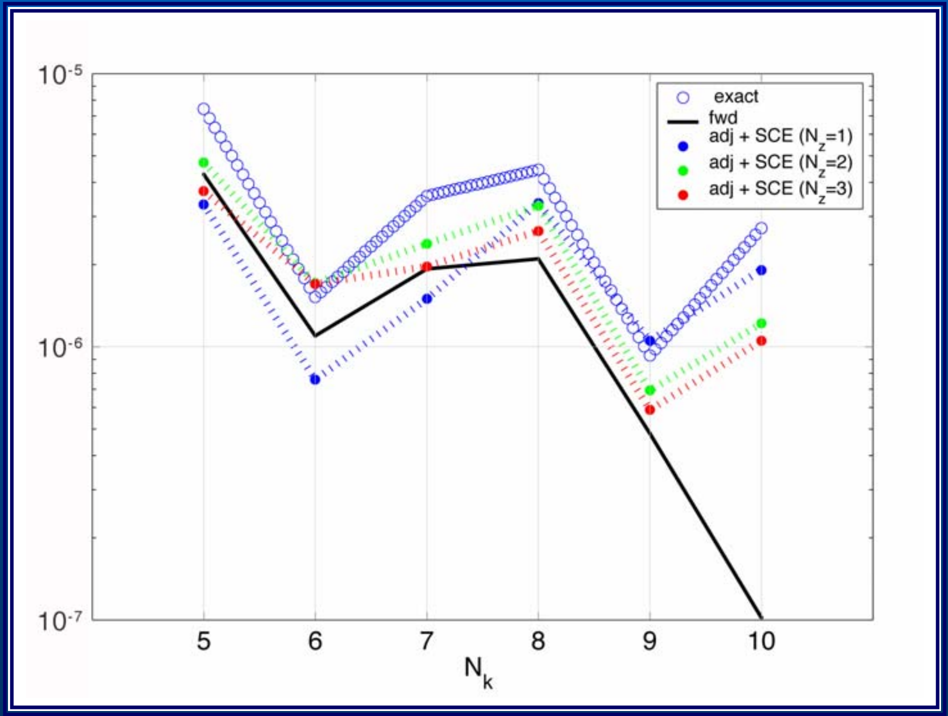
- This is a highly stiff ODE system consisting of 25 reactions and 20 chemical species

$$\frac{dy}{dt} = f(y, p), \quad y(0) = y_0$$

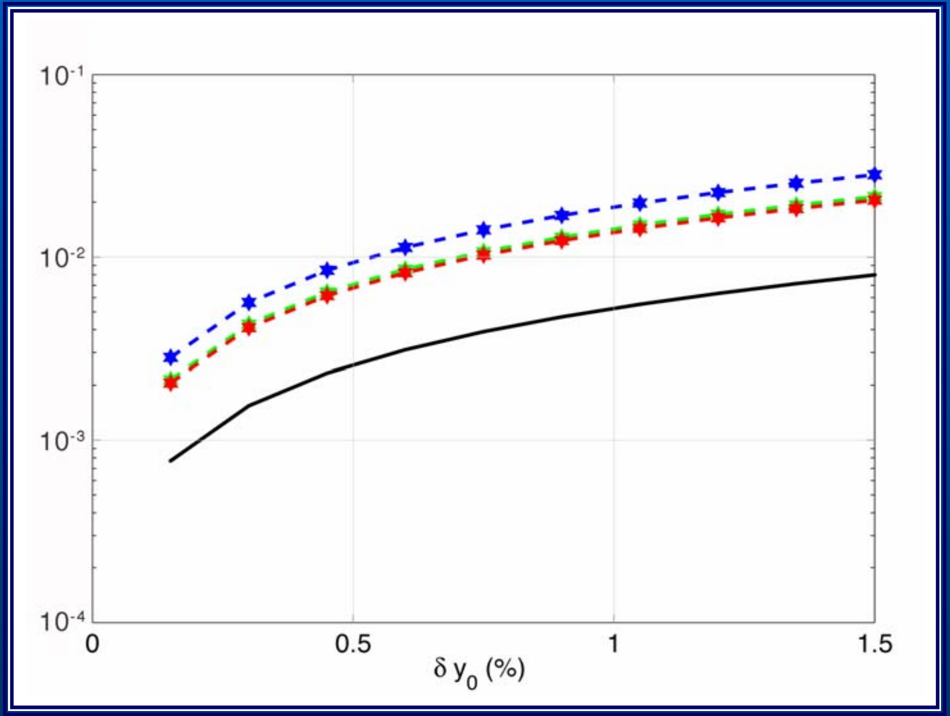
- The POD projection matrix was based on 1000 data points equally spaced in the interval $[t_0, t_f] = [0.0, 1.0]$.

Pollution Model Approximation Errors

Total Error



Approximation error E_1 as a function of the IC perturbation k=5



The solid (black) lines represent the corresponding norms computed by the forward integration of the error equations.
The dotted (colored) lines describe SCE estimates.
The dashed (colored) lines represent the bounds of the errors.
The (blue) line made of circles represents the norm of the “exact error,” $\boldsymbol{e}(t_f) = \hat{\boldsymbol{y}}(t_f) - \boldsymbol{y}(t_f)$

Discussion

ERROR BOUNDS

- do not rely on the solution of the perturbed system and therefore provide *a-priori* assessments of validity.
- are based on the continuous error equation and therefore independent of the integration method.
- can be obtained via similar procedure for perturbations in right hand side, rather than in initial conditions
- can be extended to projections other than POD.
  - good results for reduced order models of chemical kinetics