Predictive Scientific Simulations for Complex Systems

EMIL M CONSTANTINESCU
Mathematics and Computer Science
Argonne National Laboratory

Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.

Funding: U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research under contract DE-AC02-06CH11357

September 17, 2018
ASCAC Meeting
Washington DC

Overview

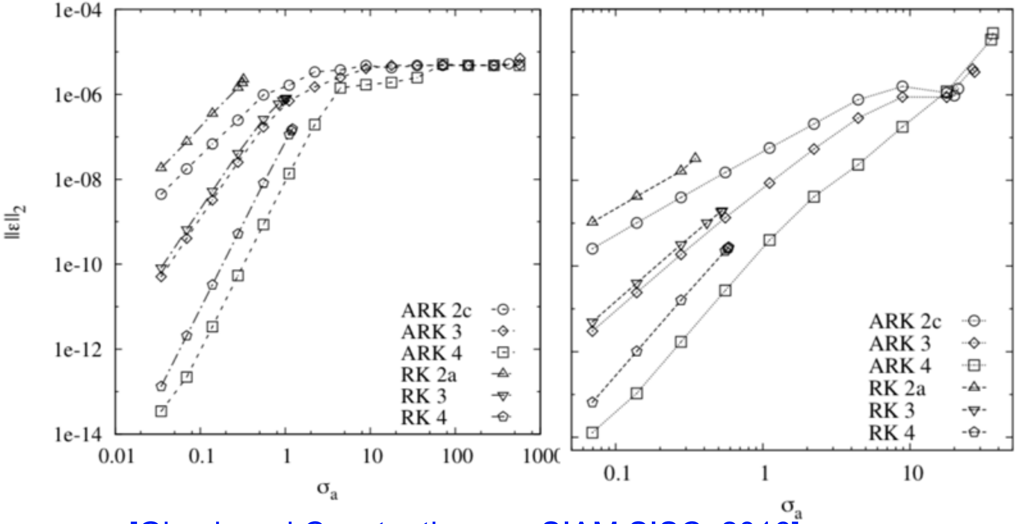
Making complex dynamic simulations predictive by using data

- Complex systems and their predictability
- Part I: Numerical errors
- Part II: Estimating model form errors
- Part III: Probabilistic predictions

Accuracy Challenges in Complex Systems

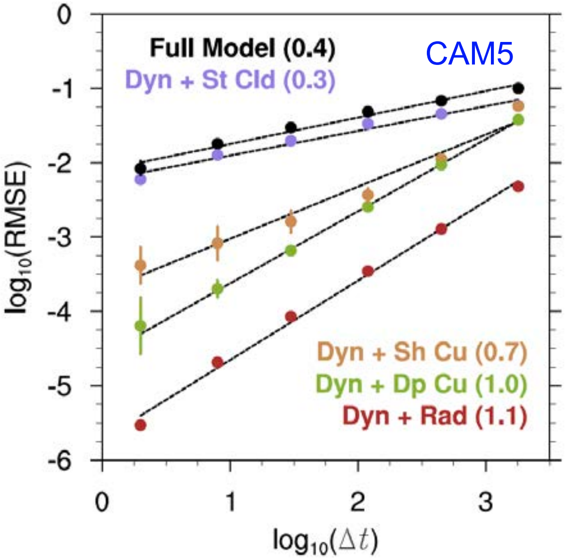
- Convergence is well understood and controlled for individual components
- This is what we think of the simulations

E.g, simulation of 3D Euler equations



[Ghosh and Constantinescu, SIAM SISC, 2016]

- Short simulation of CAM, part of CESM
- Convergence is degrading as we add "physics"
- This is what simulations of complex systems are behaving like



[Hui Wan, Philip Rasch, Mark Taylor, and Christiane Jablonowski, 2015]

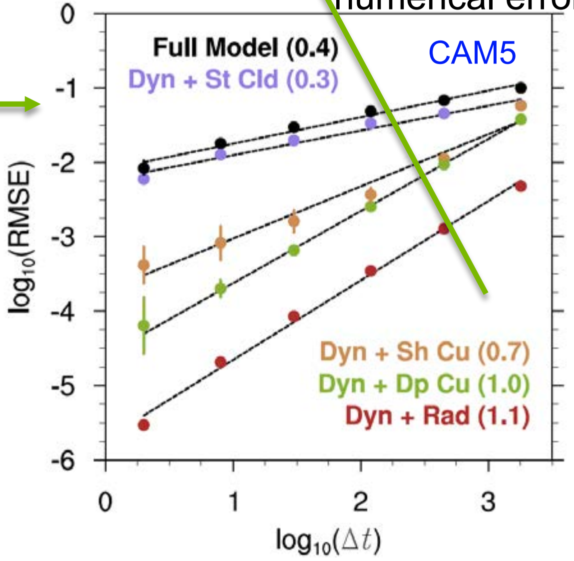
Accuracy Challenges in Complex Systems

- Short simulation of CAM, part of CESM
- Convergence is degrading by adding "physics"
- This is what simulations of complex systems are behaving like

Complexity emerges as the number of components and their interactions increase

1. Numerical error was assumed to be much smaller than model form errors
2. Sophisticated models can be predictive
3. Numerical error estimation and control is critical in assessing the quality of the solution

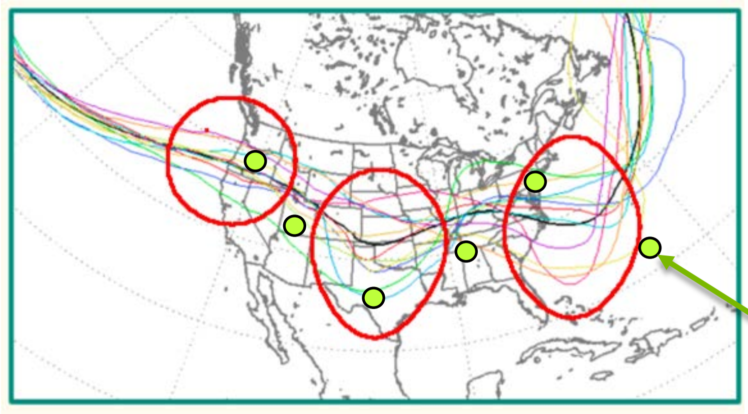
Increase in complexity brings and increase in numerical errors



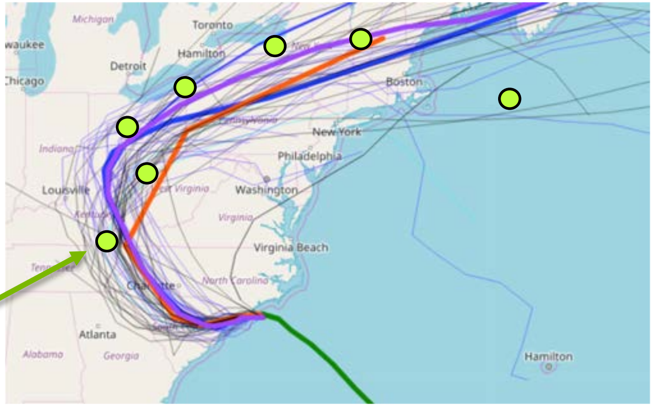
[Hui Wan, Philip Rasch, Mark Taylor, and Christiane Jablonowski, 2015]

Assessing Predictive Simulations

- Simple models: numerical analysis: $\varepsilon = ||u_{\text{simulation}} - u_{\text{exact solution}}||$ (verification)
- Simple models with data: $\varepsilon = ||u_{\text{simulation}} - u_{\text{observations}}||$ (validation)
- Complex models: set a Quantity of Interest (QoI); e.g., $T_{\text{average}} = \int_{\mathcal{D}} T(x)\mathrm{d}x$
- Stochastic models: try to match moments, marginals,…
- Measuring predictability is driven by the scientific question and becomes a complex task in itself



assume we have observations

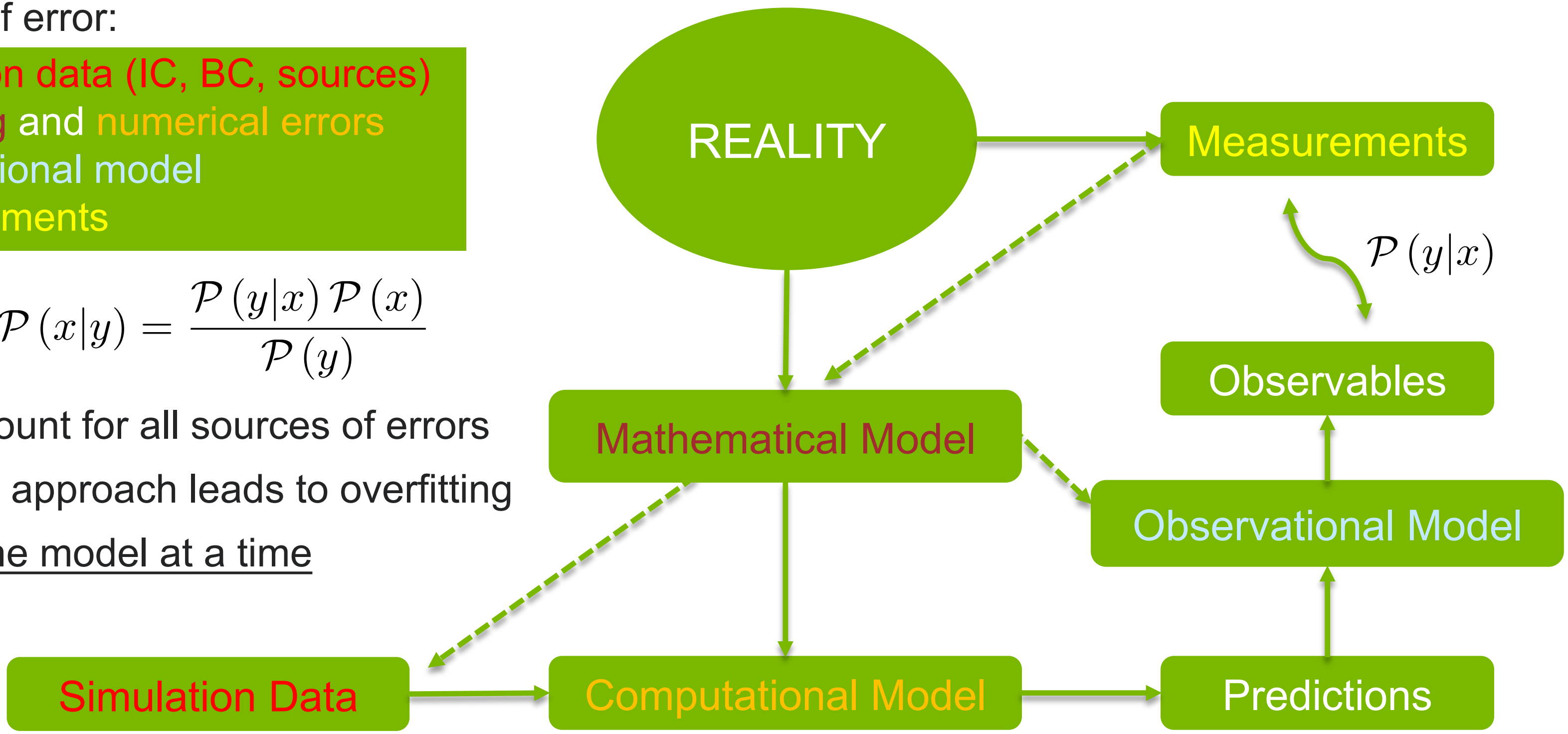


Sources of Error that Hinder Predictions When Using Simulations (Physics) and Data (Observations)

- Sources of error:
  - simulation data (IC, BC, sources)
  - modeling and numerical errors
  - observational model
  - measurements

Use Bayes: $\mathcal{P}(x|y) = \frac{\mathcal{P}(y|x)\,\mathcal{P}(x)}{\mathcal{P}(y)}$

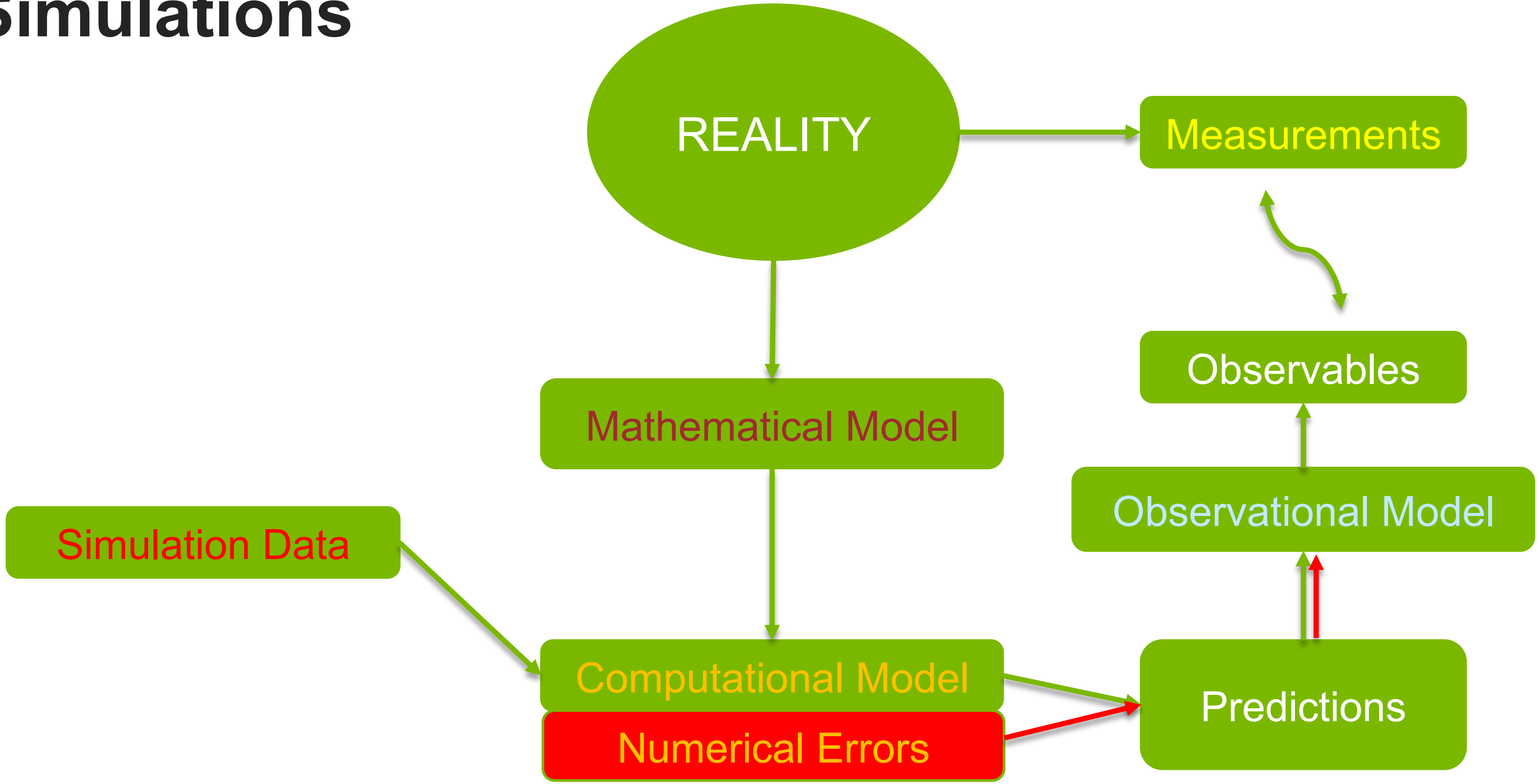
- Need to account for all sources of errors
- Reductionist approach leads to overfitting
- Cannot fit one model at a time



PART I: NUMERICAL ERRORS

initial/boundary conditions; sources → Computational Model → solution, what about errors?

Accounting for Numerical Errors in Predictive Simulations



Making Inferences and Predictions with Model Errors

- Shallow water equation with errors in a "physics" component

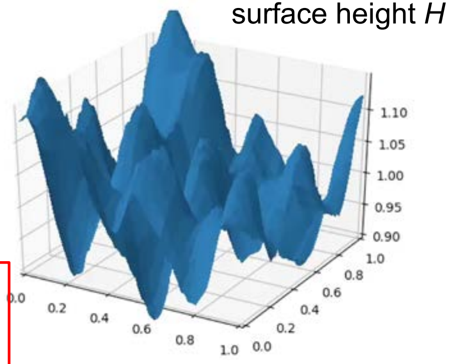
$$\frac{\partial H}{\partial t} = -\frac{\partial U}{\partial x} - \frac{\partial V}{\partial y} + \nu_H \nabla^2 H$$

$$\frac{\partial U}{\partial t} = -\frac{\partial U^2/H}{\partial x} - \frac{\partial UV/H}{\partial y} - \frac{1}{2}\frac{\partial g H^2}{\partial x} + \nu_U \nabla^2 U$$

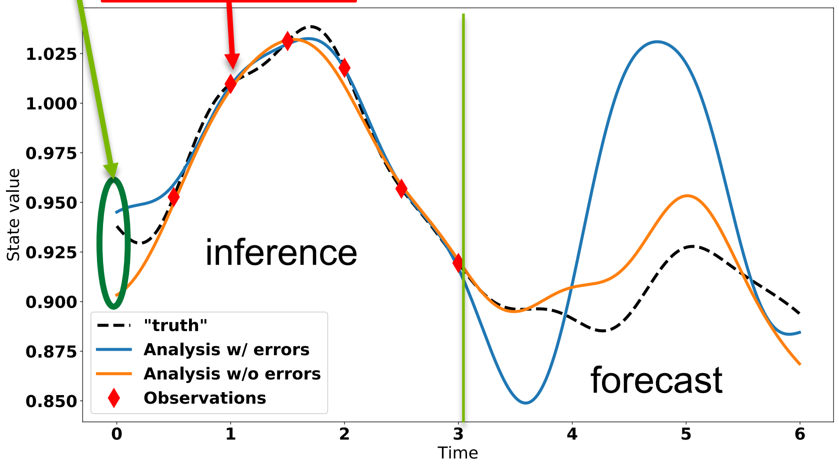
$$\frac{\partial V}{\partial t} = -\frac{\partial UV/H}{\partial x} - \frac{\partial V^2/H}{\partial y} - \frac{1}{2}\frac{\partial g H^2}{\partial y} + \nu_V \nabla^2 V$$

$$g = F(x, y) + \varepsilon$$

surface height H



Variable H at a location in space over time



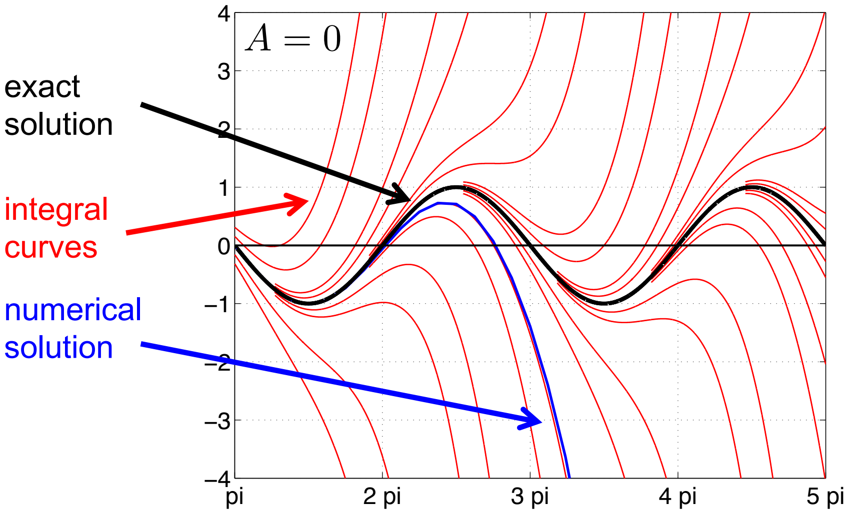
- Both resolved and under-resolved models perform well during the inference, have different behavior during the forecast

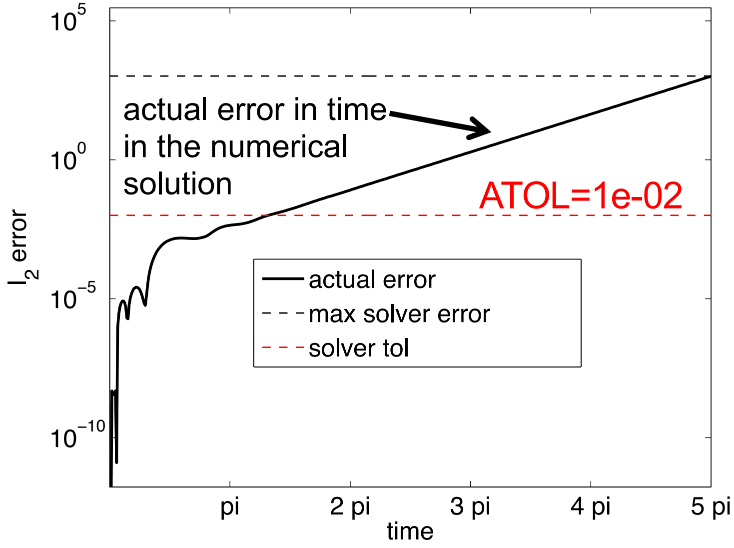
Canonical Example Illustrating the Compounding Effect of Numerical Errors

- Problem: $y' = y - \sin(t) + \cos(t)$ $\qquad y(t) = A\exp(t) + \sin(t)$

  $y(0) = A, \quad A = 0$

- Any solver with local error control with ATOL = 1e-02





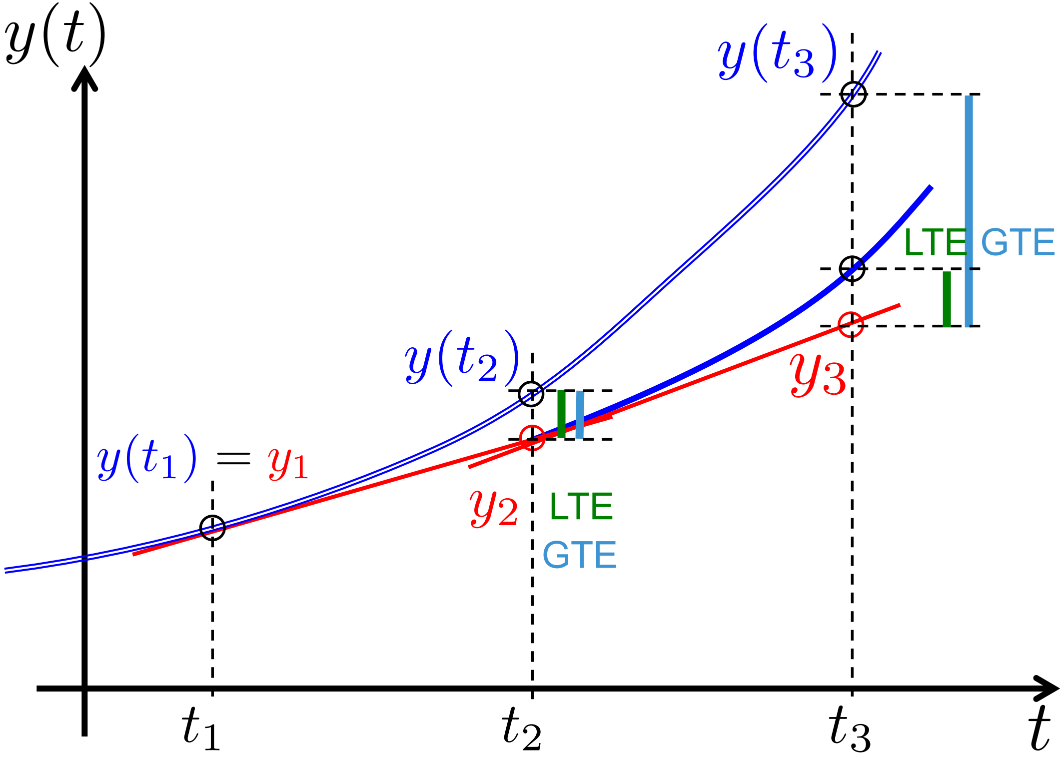
- No warnings if errors are really really large

ODE Dictionary: Local and Global Truncation Errors

- ODE (semi-discrete PDE) system:

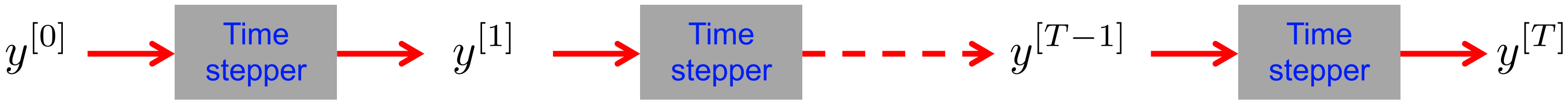
$$y' = f(t, y)\,,\;\; y(t_1) = y_1$$

- Local truncation error (LTE) is used to control the integration error
- *Actual error* or global truncation error (GTE) or a posteriori error is not used in solvers

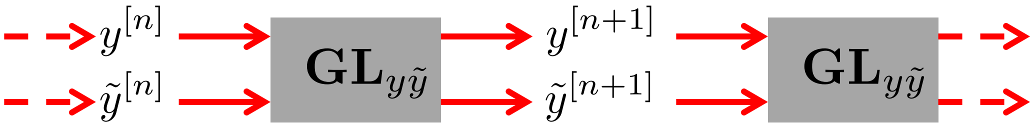


Estimating Numerical Errors

- Traditional time stepping w/o error estimation; only the solution is propagated



- Many ways of estimating a posteriori errors; typically expensive, and thus not used
- We have demonstrated that:
  1. Two solutions need to be propagated to estimate errors solutions
  2. Introduced a general framework – most previous methods are particular cases of it
  3. Leads to the **most efficient possible algorithm within general linear methods**
  4. Can be used for estimation as well as control



Emil Constantinescu, "Generalizing global error estimation for ordinary differential equations by using coupled time-stepping methods." Journal of Computational and Applied Mathematics, 2018.

New Framework for Error Estimation Based on GLEE

- The general strategy leads to very efficient methods; need to propagate two solutions

Problem: ODE
$y' = f(y)$

- Example: an explicit Runge-Kutta can be modified to propagate the numerical errors

- A general result: extrapolation, solving for the correction, Zadunaisky procedure,… can be expressed in the same framework

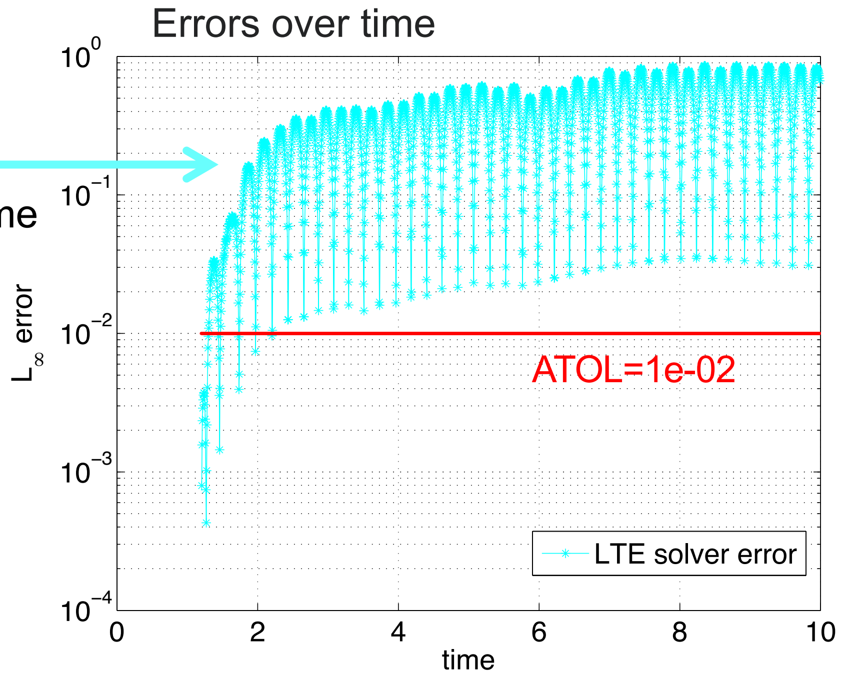
$$
\begin{aligned}
Y_1 &= y^{[n-1]} \\
Y_2 &= -9y^{[n-1]} + 10\tilde{y}^{[n-1]} + \Delta t f(Y_1) \\
Y_3 &= 2y^{[n-1]} - \tilde{y}^{[n-1]} + \Delta t \left(\frac{1}{4} f(Y_1) + \frac{1}{4} f(Y_2) \right) \\
y^{[n]} &= y^{[n-1]} + \Delta t \left(\frac{1}{12} f(Y_1) + \frac{1}{12} f(Y_2) + \frac{5}{6} f(Y_3) \right) \\
\tilde{y}^{[n]} &= \tilde{y}^{[n-1]} + \Delta t \left(\frac{1}{6} f(Y_1) + \frac{1}{6} f(Y_2) + \frac{2}{3} f(Y_3) \right)
\end{aligned}
$$

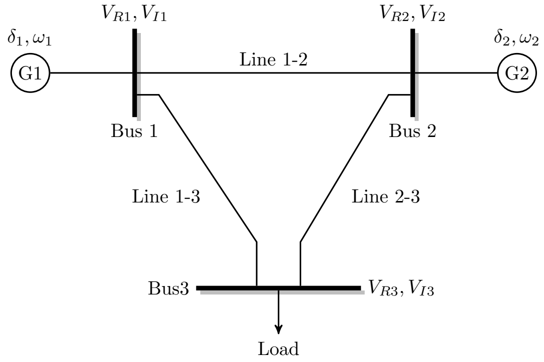
- Error is easily estimated: $\varepsilon^{[n]} = \tilde{y}^{[n]} - y^{[n]}$

Transient Stability Analysis in a 3-Bus Power Grid

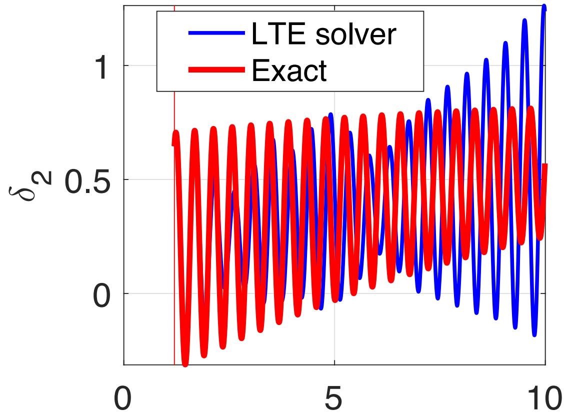
- Simple power grid model, introduce a fault, study dynamic behavior, set error ATOL = 1e-02

Global error over time





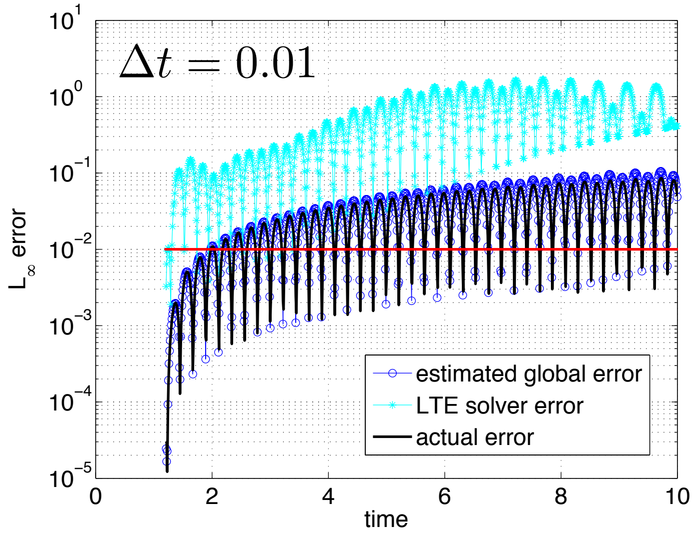
Solution component – evolution over time

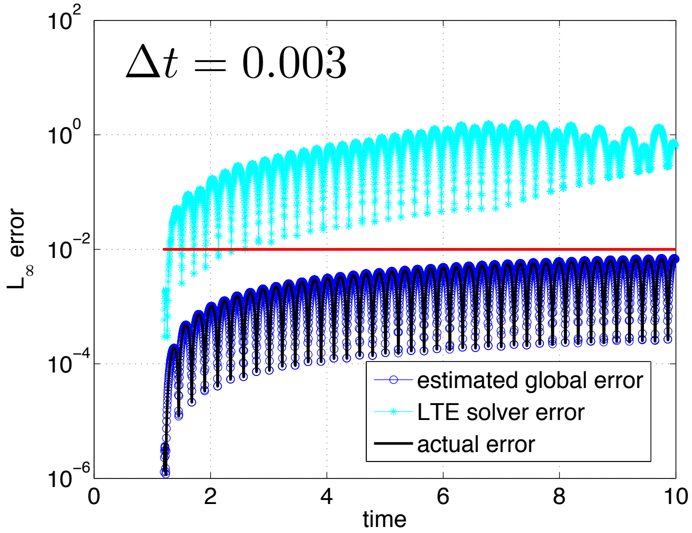


Error Control Using a 2nd Order Method

1. Strategy for estimation: estimation is automatic

2. Strategy for control: two passes, to achieve ATOL = 0.01 a time step of 0.0030823 should have been used instead of 0.01

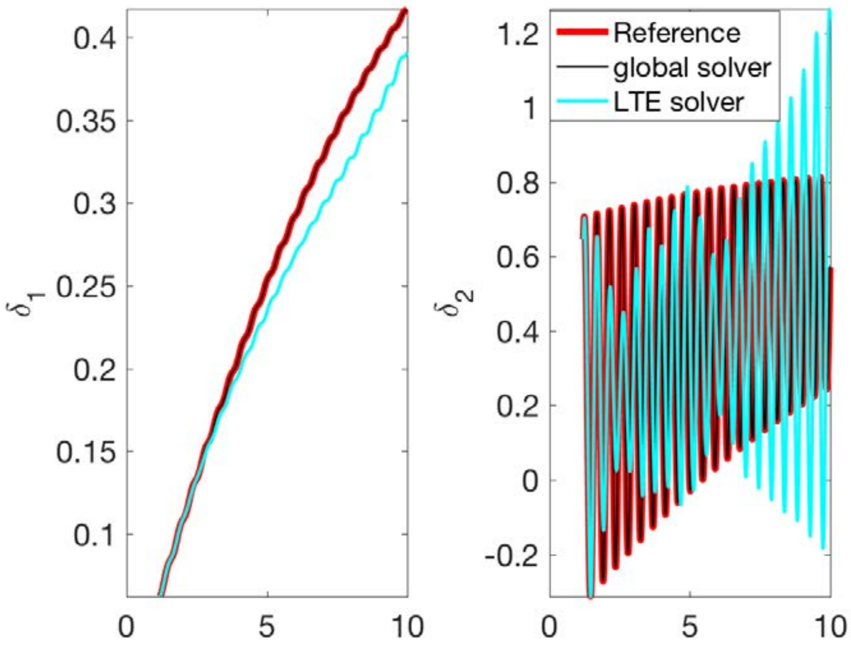




- Proportionality error controller tolerance can also be used

Corrected Solution and Solution Obtained by Local Error Estimation

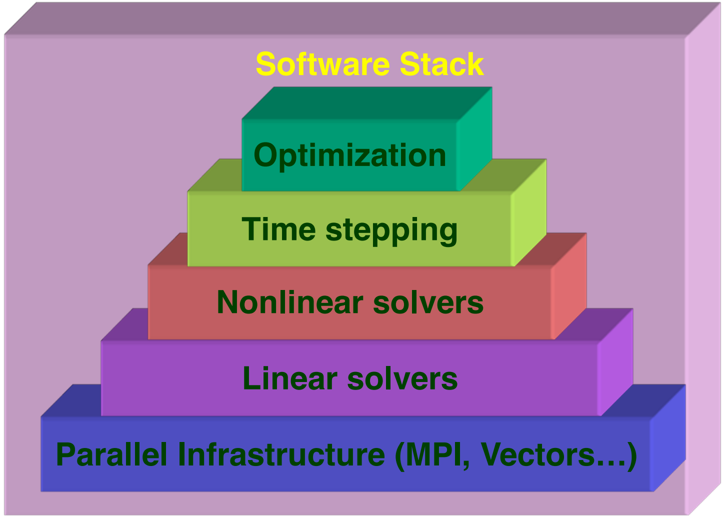
- Solution obtained with ATOL = 1e-02 can lead to wrong interpretation
- Solution computed with estimated global error and corrected looks okay quantitatively and qualitatively
- Note that this is an expensive process



Implementation of Error Estimation in PETSc: Portable Extensible Toolkit for Scientific Computation

- Open-source numerical library for large-scale parallel computation
- Portability
  - Unix, Linux, MacOS, Windows
  - 32/64 bit, real/complex, ...
  - C, C++, Fortran, Python
- Extensibility
  - ParMetis, SuperLU, SuperLU_Dist, MUMPS, hypre, UMFPACK, Sundials, Elemental, ScaLAPACK, UMFPack, …
- Toolkit
  - Iterative solvers and preconditioners
  - Parallel nonlinear solvers
  - **Time stepping (ODE and DAE) solvers**
  - **Adjoint sensitivity analysis**
  - **Event support**
  - **Support for network data structures**
  - **Optimization**

1000s of users in academia and industry



Scalable Solver Suite for ODEs/DAEs/PDEs

$$G(t, y, y') = F(t, y)$$

All PETSc time integrators use a unified interface

implicit(ODE/DAE) "+" explicit "=" IMEX

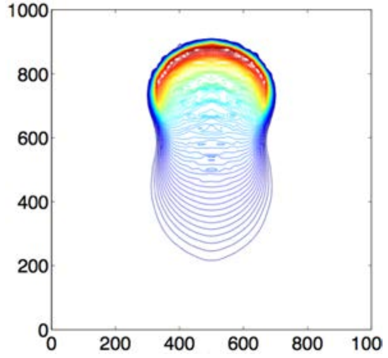
$g(y, y') = 0$ $y' = f(y)$ $y' = f(y) + g(y)$

mpirun –np 4 burgers.exe -ts_type rk

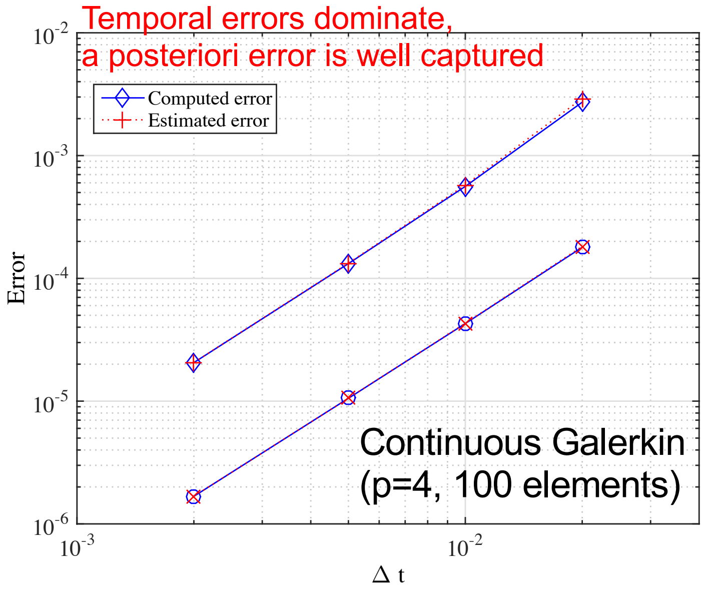
mpirun –np 4 burgers.exe -ts_type glee

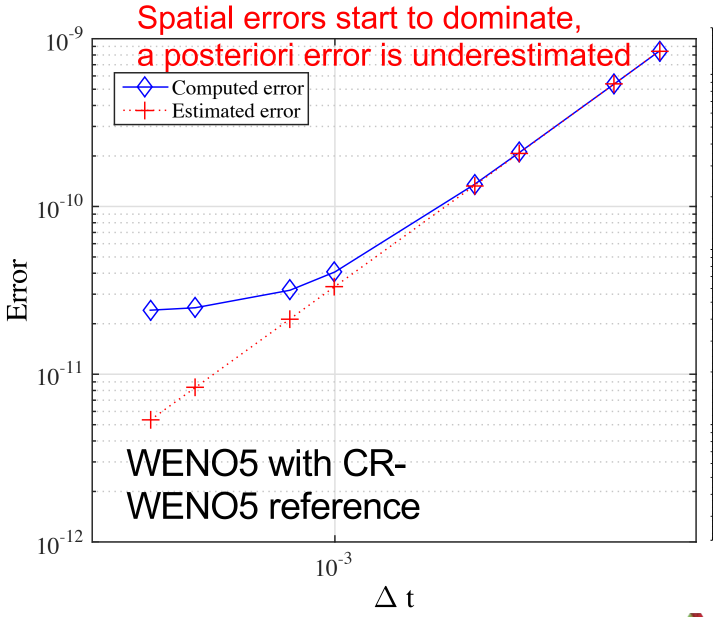
PETSc name	Name	Class	Type	Order
euler	forward Euler	one-step	explicit	1
ssp	multistage SSP	Runge-Kutta	explicit	≤ 4
rk	explicit RK	Runge-Kutta	explicit	≥ 1
theta,beuler,cn	theta-method	one-step	implicit	1, 2
alpha	alpha-method	one-step	implicit	2
glle	general linear	multi/step/stage	implicit	≤ 3
eimex	extrapolated IMEX	one-step	IMEX	≥ 1, adapt.
ARK-IMEX	additive RK	IMEX RK	IMEX	$1 - 5$
RosW	Rosenbrock	Rosenbrock-W	lin. implicit	$1 - 4$
glee	general linear with global error est.	general linear	explicit or implicit	≥ 1

Global Error Estimation for PDEs



- Solve 2D rising thermal bubble: compressible Euler/gravity field
- Final times 1-400 sec.
- Spectral elements: [Giraldo, Kelley, Constantinescu, JCP 2013]
- Finite difference/volume grids: [Ghosh, Constantinescu, SISC 2016]

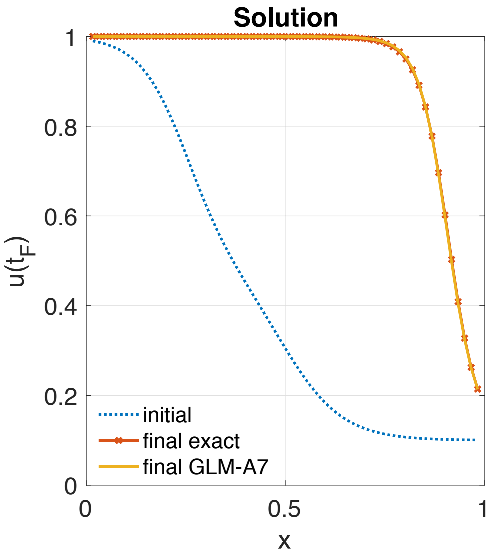




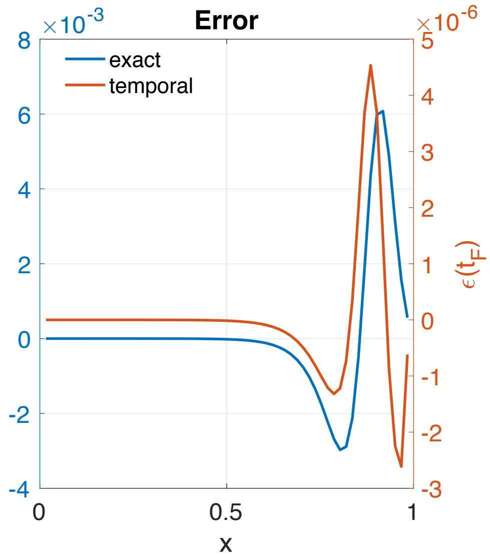
Deterministic Error Estimation for PDEs

- Burgers PDE equation with exact solution: $$\frac{\partial u}{\partial t} = \nu \frac{\partial^2 u}{\partial x^2} - u\frac{\partial u}{\partial x}\,, \quad (x,t) \in [0,1] \times [0,1]$$
- Inject errors in the time stepping method at every step from a spatial error estimator

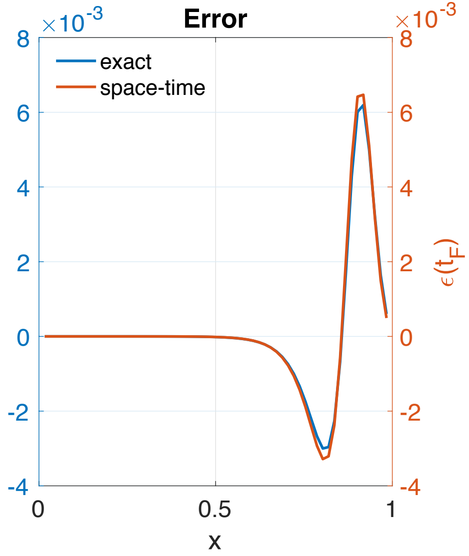
Solution



Ignoring spatial errors



Considering spatial errors



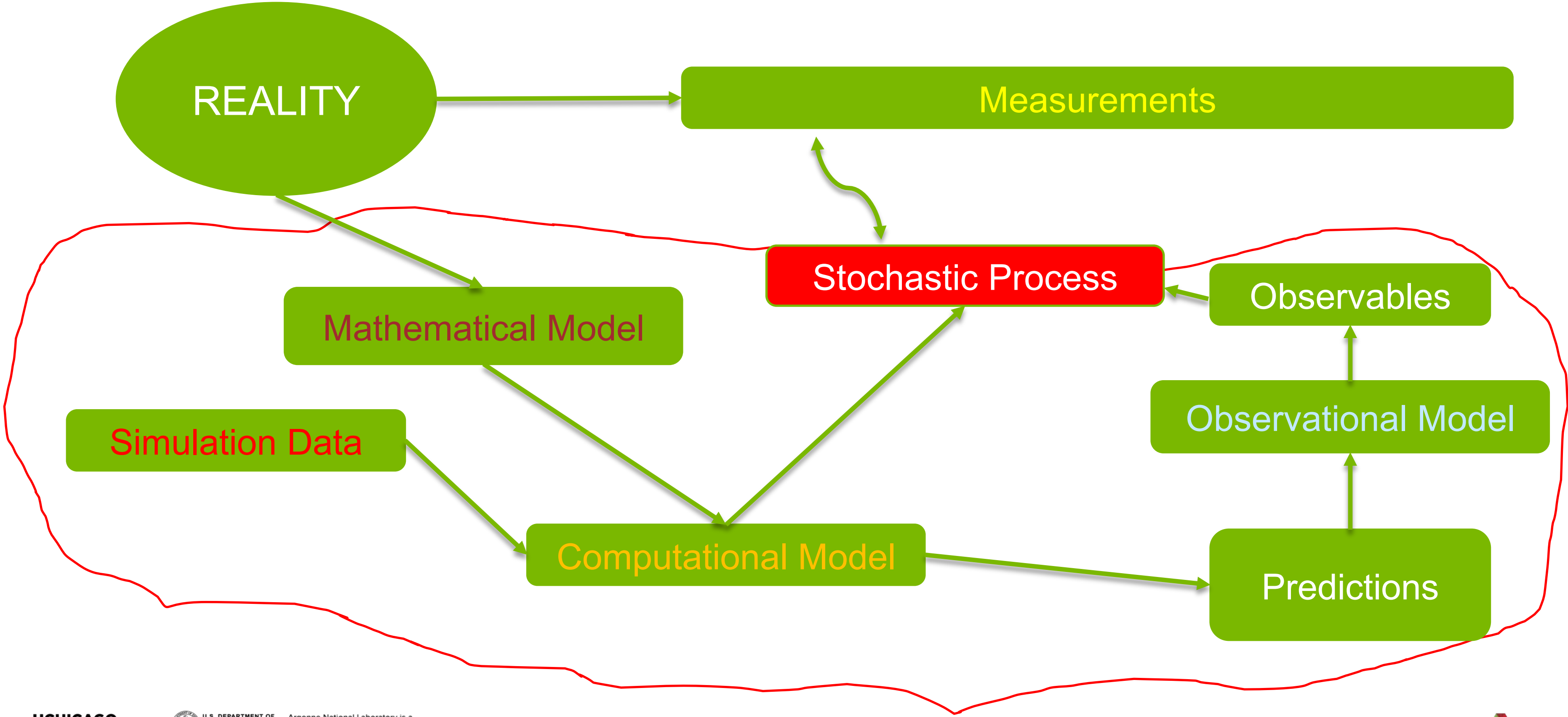
Summary on Numerical Errors

- Numerical errors may dominate your computations even when using error control
- We can control and estimate numerical errors efficiently
- Not accounting for numerical errors may lead to spurious predictions and inferences
- Implementation in PETSc can be used for estimating errors during time integration, RHS (e.g., spatial) error estimation must be provided

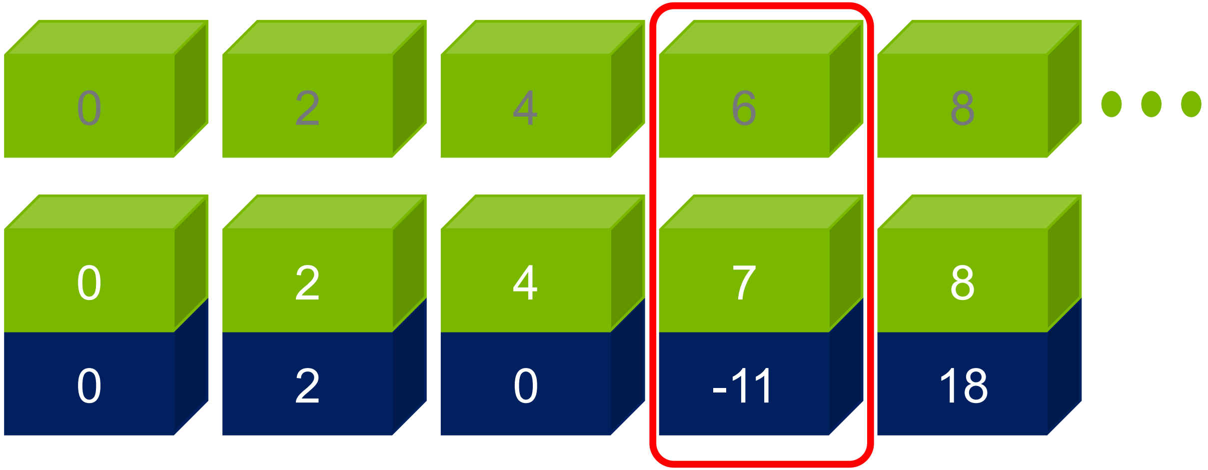
Emil Constantinescu, "Generalizing global error estimation for ordinary differential equations by using coupled time-stepping methods." Journal of Computational and Applied Mathematics, 2018.

PART II: MODEL FORM ERRORS

Framework for Predictive Modeling with Model-Form Errors



Thought Experiment



Proposed numerical model

$$\widehat{y}_n = \widehat{y}_{n-1} + 2$$

True process:

$$\widehat{y}_n = \frac{5}{4}\widehat{y}_{n-1} + 2 - \frac{1}{4}\tilde{y}_{n-1}$$

$$\tilde{y}_n = -\frac{13}{4}\tilde{y}_{n-1} + 2 + \frac{9}{4}\widehat{y}_{n-1}$$

$$\begin{cases} \widehat{y}_n = a\widehat{y}_{n-1} + b + c\tilde{y}_{n-1} \\ \tilde{y}_n = f\tilde{y}_{n-1} + h + g\widehat{y}_{n-1} \end{cases} \Rightarrow \widehat{y}_n = a\widehat{y}_{n-1} + b + c\left\{ f^n\tilde{y}_0 + \sum_{j=1}^{n} f^{n-j}(g\widehat{y}_{j-1} + h) \right\}$$

$$= a\widehat{y}_{n-1} + b + cN(\tilde{y}_0, \widehat{y}_0) + cG(\widehat{y}_1, \ldots, \widehat{y}_n)$$

Can just compute $\widehat{y}_n$ from and estimate $\tilde{y}_0$ to have a complete solution for $\widehat{y}_{0,1,\ldots,n}$

Nonlinear Variation of Constants and Mori-Zwanzig Formalism

- One way to express missing dynamics from a simulation is to use the (nonlinear) variation of constants for differential equations, or Duhamel formula, or Mori-Zwanzig

- Difference model:

$$\widehat{y}_n = a\widehat{y}_{n-1} + b + cN(\tilde{y}_0, \widehat{y}_0) + cG(\widehat{y}_1, \ldots, \widehat{y}_n)$$

- Mori-Zwanzig formalism:

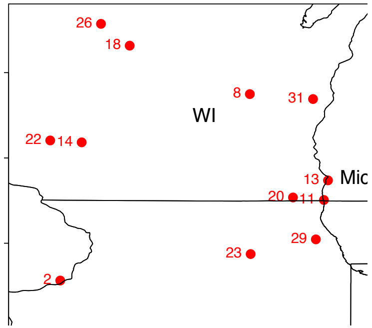
$$\frac{\mathrm{d}\widehat{y}(t)}{\mathrm{d}t} = S(t, \widehat{y}(t)) + N(t, \widehat{y}(t_0), \tilde{y}(t_0)) + \int_0^t K(t-s)F(\widehat{y}(s))\, ds$$

- In principle, the real dynamics can be approximated by a stochastic term and a functional of the computed solution

Practical Application: Develop an Augmented Wind Forecast

- Use numerical weather simulations and observational data to correct the forecast and estimate an error model

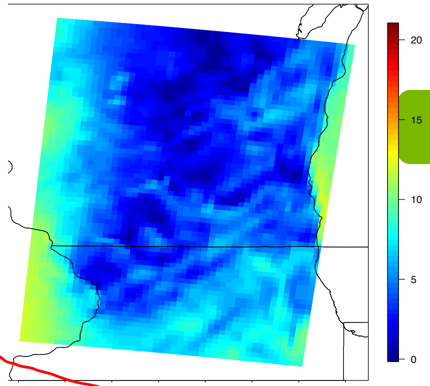
Ground observations of the wind velocity



Measurements

Prediction

Stochastic Process

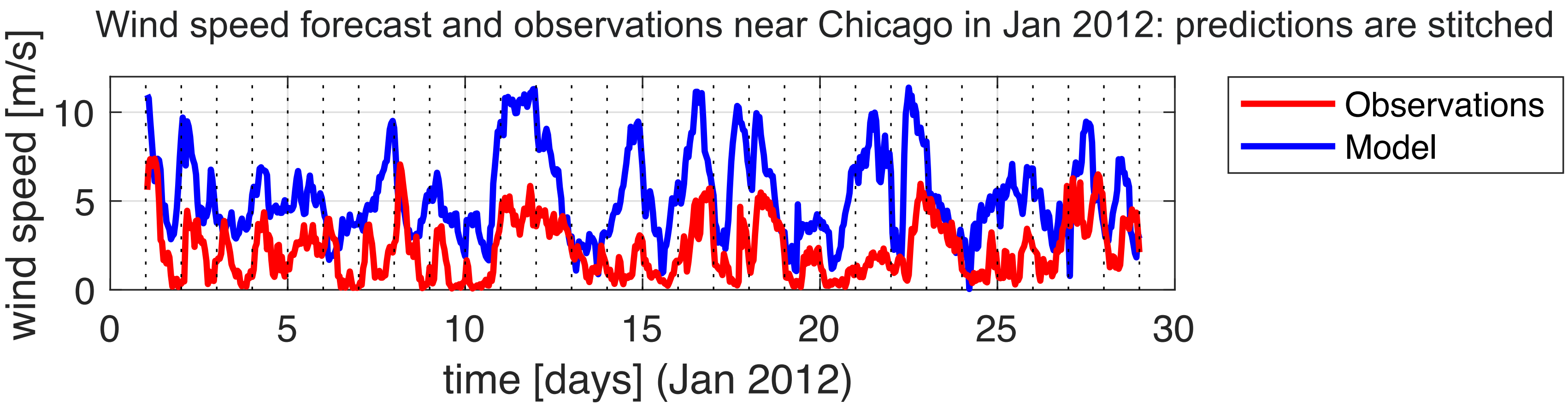


Observables

Wind field obtained from numerical weather predictions

Prediction and Measurement Data for One Measured Location

- The prediction is issued every 24 hours for 24 hours ahead
- Use 12 independent measurements around Chicago (median RMSE shown)
- Simulations generally overpredicts the wind fields



Probabilistic Formulation

- Express the posterior distribution of the observations given simulations of the future and past and observations (in the past)

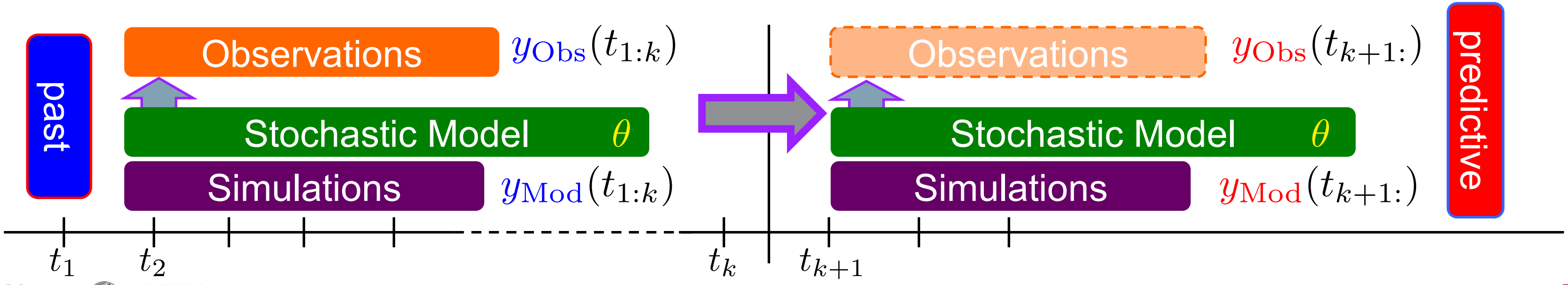
$$p\left(y_{\mathrm{Obs}}(t_{k+1:}) \middle| y_{\mathrm{Mod}}(t_{k+1:}), y_{\mathrm{Obs}}(t_{1:k}), y_{\mathrm{Mod}}(t_{1:k})\right)$$

$$= \int p\left(y_{\mathrm{Obs}}(t_{k+1:}), \theta \middle| y_{\mathrm{Mod}}(t_{k+1:}), y_{\mathrm{Obs}}(t_{1:k}), y_{\mathrm{Mod}}(t_{1:k})\right) \, \mathrm{d}\theta$$

$$\approx \int p\left(y_{\mathrm{Obs}}(t_{k+1:}) \middle| \theta, y_{\mathrm{Mod}}(t_{k+1:})\right) p\left(\theta \middle| y_{\mathrm{Obs}}(t_{1:k}), y_{\mathrm{Mod}}(t_{1:k})\right) \, \mathrm{d}\theta$$

$$\approx p\left(y_{\mathrm{Obs}}(t_{k+1:}) \middle| \theta^*, y_{\mathrm{Mod}}(t_{k+1:})\right) \qquad \theta^* = \mathrm{argmax}_\theta \; p\left(\theta \middle| y_{\mathrm{Obs}}(t_{1:k}), y_{\mathrm{Mod}}(t_{1:k})\right)$$

Key:
future
past



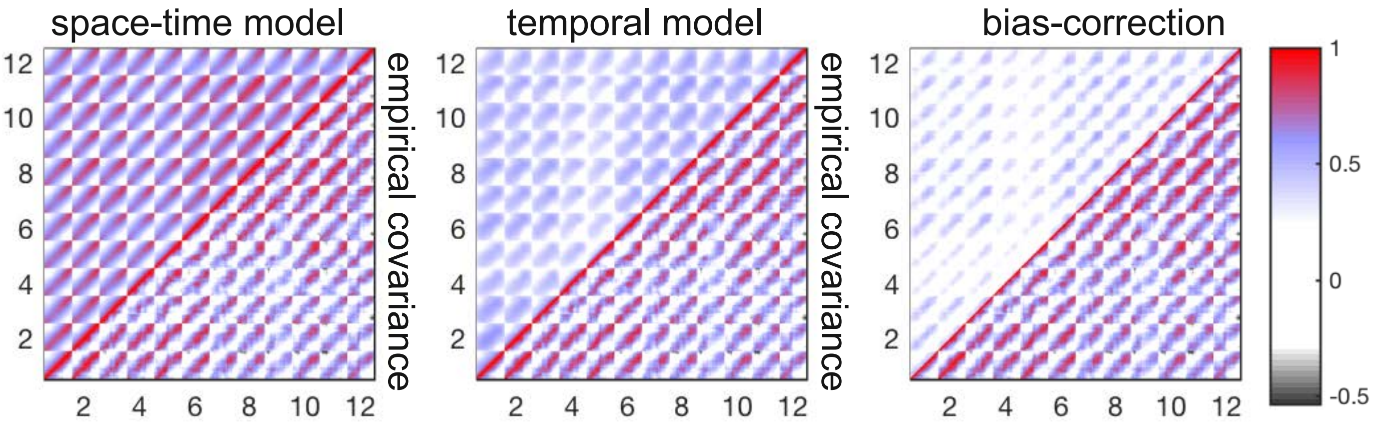
Stochastic Modeling: Use Gaussian Processes

- Model simulation and observations as a joint Gaussian process

hierarchical space-time covariance structure

$$\begin{pmatrix} Y_{\text{Obs}} \\ Y_{\text{Model}} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_{\text{Obs}}(\theta) \\ \mu_{\text{Model}}(\theta)\end{pmatrix}, \begin{pmatrix} \Sigma_{\text{Obs}}(\theta) & \Sigma_{\text{Obs}|\text{Model}}(\theta) \\ \Sigma^T_{\text{Obs}|\text{Model}}(\theta) & \Sigma_{\text{Model}}(\theta)\end{pmatrix}\right)$$

- Correlation structure of the GP with various models

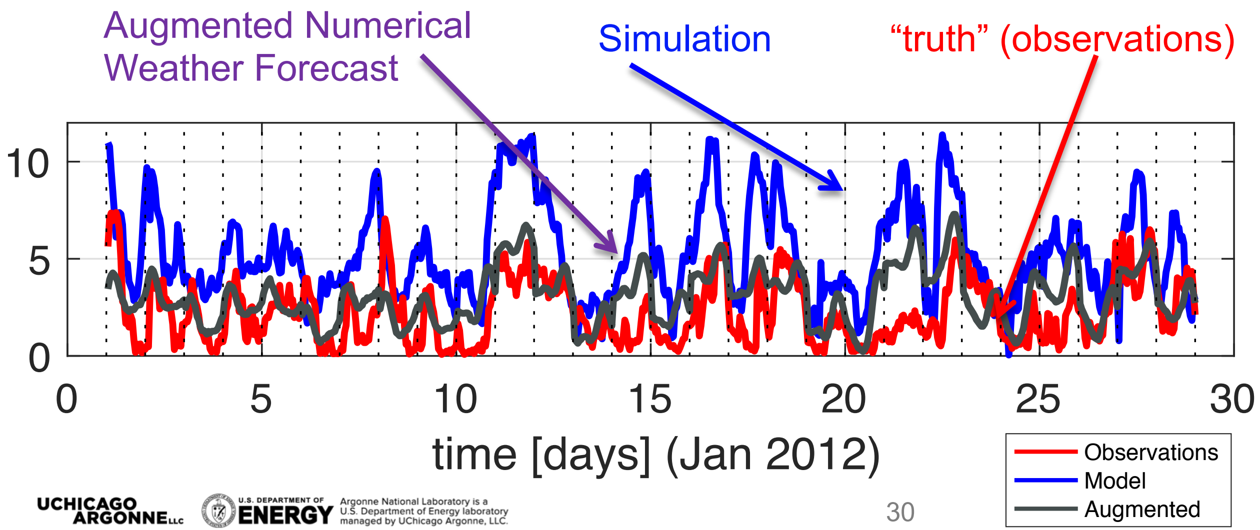


J. Bessac, E.M. Constantinescu, and M. Anitescu, Stochastic simulation of predictive space-time scenarios of wind speed using observations and physical models, Annals of Applied Statistics, 2018

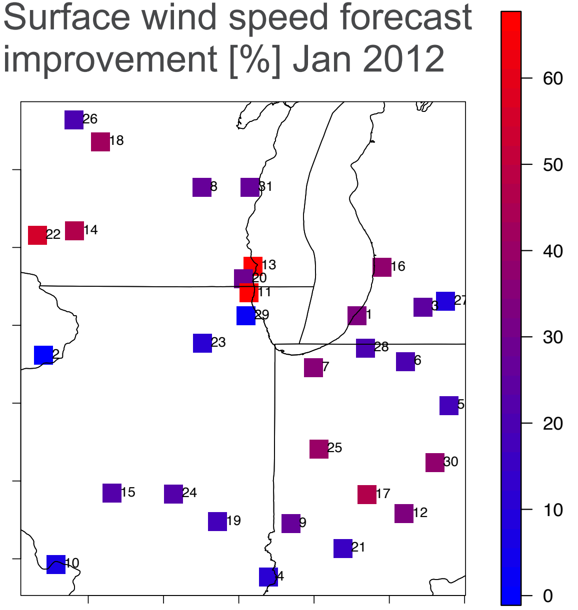
J. Hart, J. Bessac, and E.M. Constantinescu, Global sensitivity analysis for statistical model parameters, under review, 2018

Augmenting the Forecast

- Use the GP to compute the posterior distribution of the wind at the observation locations conditioned on the forecast
- GP is calibrated once every 10 days, use ~90 space-time variables, reduced to ~70 via global sensitivity analysis
- The location with median improvement

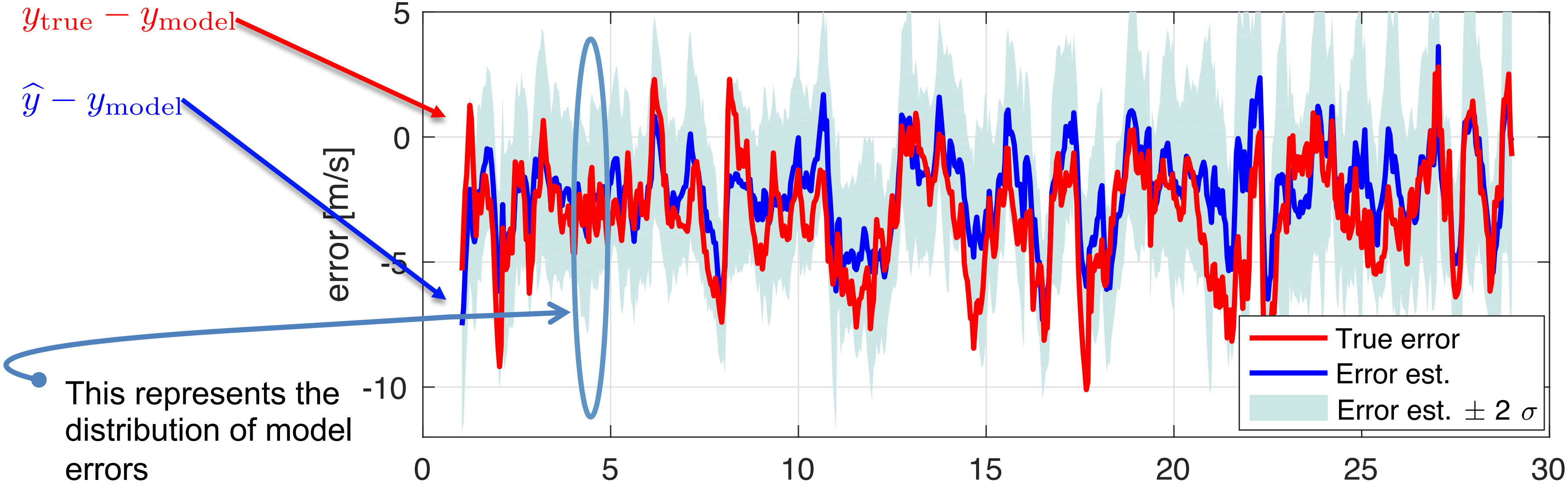


Surface wind speed forecast improvement [%] Jan 2012



Forecast of the Model-Form Errors at One Point in Space over a Month by Using GPs

- Wind speed errors over a month: difference between model and reality
- Stochastic forecast for the errors and confidence intervals tracks true errors well



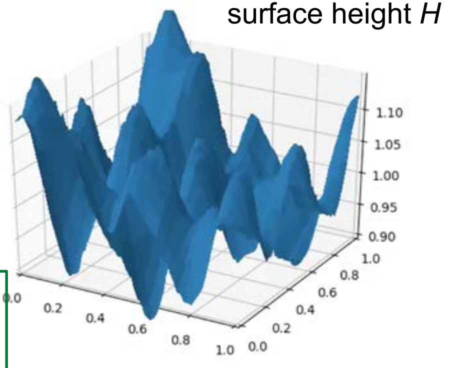
Making Inferences and Predictions with Model Errors

- Including model errors in the inference process: 4D-Var, a smoother estimated at the maximum a posteriori estimate

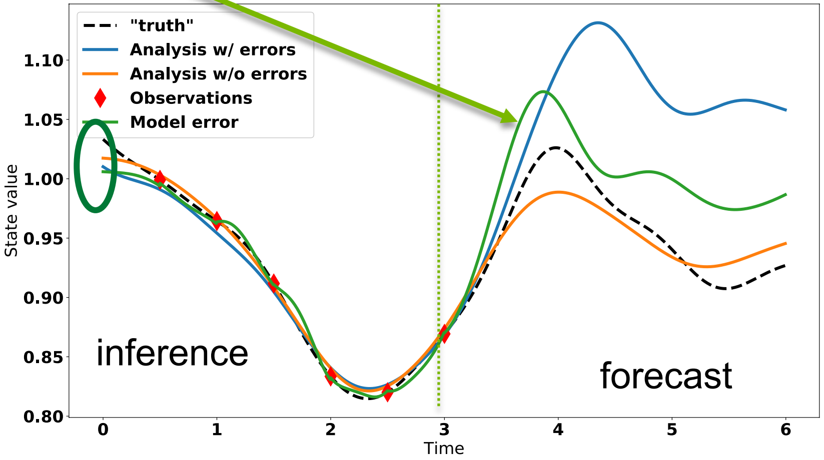
$$\mathcal{J}(y_{t_0}, y_{t_1}, \ldots, y_{t_N}) = \frac{1}{2}||y_{t_0} - y_B||_{Q_B^{-1}} + \frac{1}{2}\sum_{k=0}^{N}||H_k(y_{t_k}) - y_k^o)||_{R_k^{-1}} + \frac{1}{2}\sum_{k=0}^{N-1}||y_{t_{k+1}} - \mathcal{M}_k(y_{t_k})||_{Q_k^{-1}}$$

$$[y_{t_0}^*, y_{t_1}^*, \ldots, y_{t_N}^*)]^T = \arg\max \mathcal{J}(\cdots)$$

surface height H



Accounting for model errors



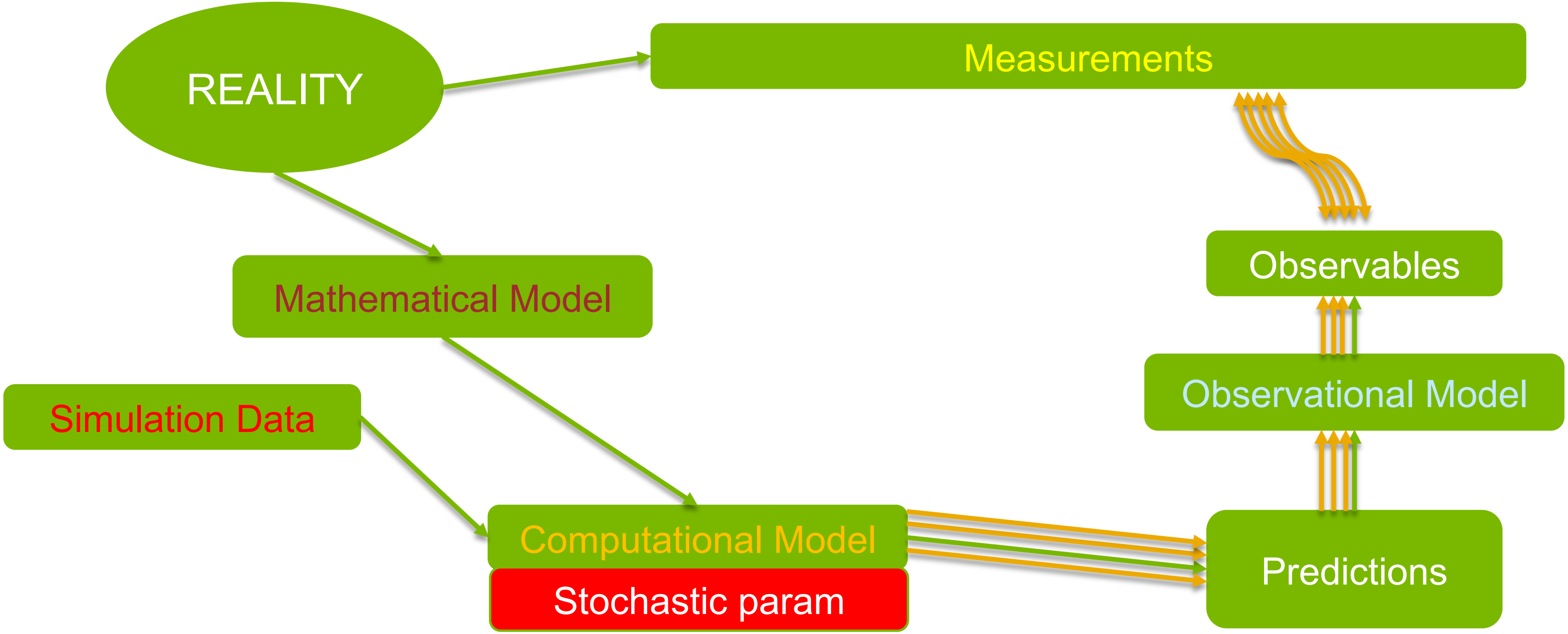
Summary on Model-Form Errors

- Model-form errors are typically the dominant obstacle for complex predictions
- These are difficult to estimate and are application dependent
- Statistical model-form errors can be included in the inference process or augment predictions

J. Bessac, E.M. Constantinescu, and M. Anitescu, Stochastic simulation of predictive space-time scenarios of wind speed using observations and physical models, Annals of Applied Statistics, 2018
J. Hart, J. Bessac, and E.M. Constantinescu, Global sensitivity analysis for statistical model parameters, under review, 2018

PART III: PROBABILISTIC PREDICTIONS

Framework for Stochastic Predictions

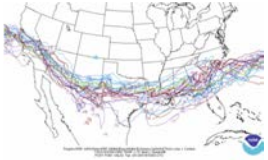


Calibrate Simulations with Stochastic Predictions

- Observations (real process): $\boldsymbol{d}_T = \boldsymbol{d}_{\text{true}} + \varepsilon_{\boldsymbol{d}}$
- Model with stochastic parameters: $F(x, m; \xi) = 0$

x – states, m – parameters, ξ – noise

E.g., geophysical applications



E.g., power grid

$x = [y, z]^T$

$$\begin{cases} \dot{y} &= g(y, z; m) \\ 0 &= h(y, z; \xi) \end{cases}$$

E.g., subsurface flow

$-\nabla \cdot (e^{\mathrm{m}} \nabla x) = \xi$

- What is the best m s.t. $\boldsymbol{d}$ agrees with $\boldsymbol{d}_T$?
- Define cost/loss function: $S(\boldsymbol{d}_T, \boldsymbol{d}; m)$
- Seek a solution: $m^* = \arg\min_m S(\boldsymbol{d}_T, \boldsymbol{d}; m)$

How Do We Measure the Distance Between Prediction Distributions and (Point) Observations?

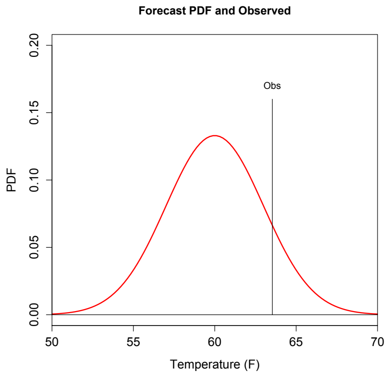
- Need qualitative and quantitative measures for distances from distributions to realizations
  - Calibration: statistical similarity of forecasts and observations; i.e., P → Y

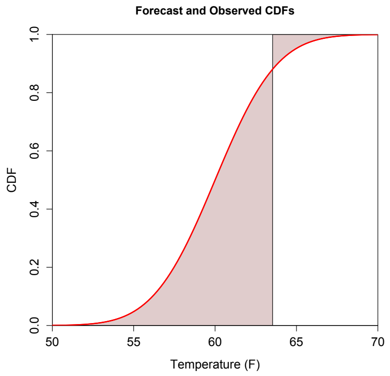
  - Sharpness: the concentration of the predictive distributions; e.g., ↘var(P)

Measuring the Distance Between Prediction Distributions and (Point) Observations

- Statistics: forecast verification, use proper metrics (proper scoring rules)
  - Loss function
  - Performance metric

- CRPS
- Energy score (ES):
- Variogram score (VS):





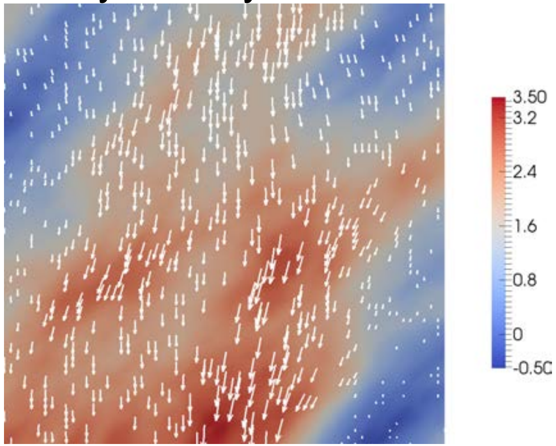
Subsurface (Groundwater, Porous Media) Model

- Elliptic PDE modeling porous media with random forcing -> SPDE

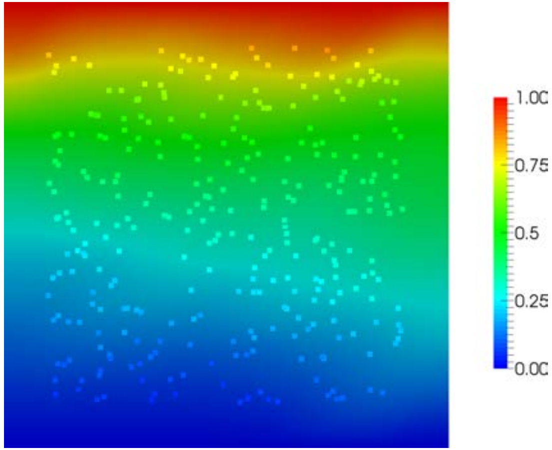
$$\begin{aligned} -\nabla \cdot (e^{\mathrm{m}} \nabla u) &= f \quad \text{in } \mathcal{D}, \\ u &= g \quad \text{on } \Gamma_D, \\ e^{\mathrm{m}} \nabla u \cdot \boldsymbol{n} &= h \quad \text{on } \Gamma_N, \end{aligned}$$

$$f \sim \pi_\xi, \ \ \pi_\xi \text{ known distribution}$$

Darcy velocity: $q = -e^{\mathrm{m}} \nabla u$

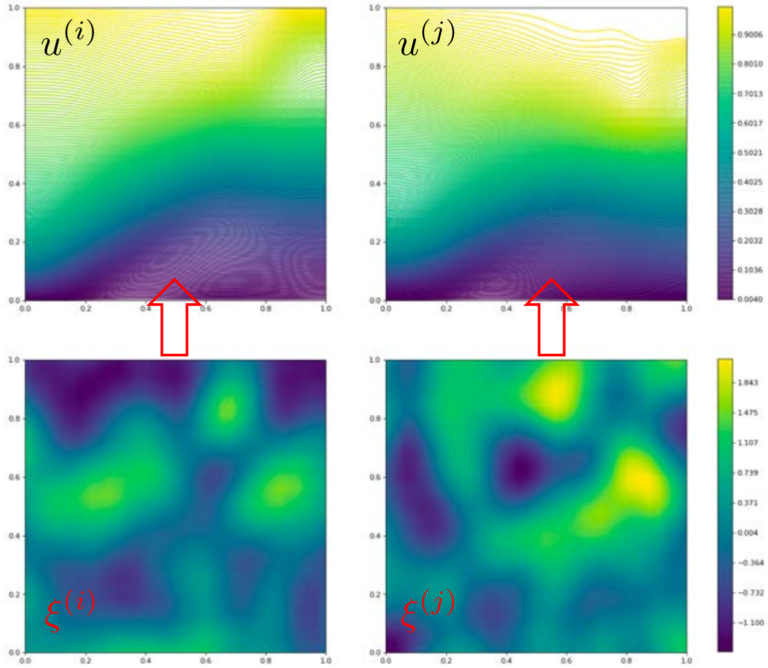


state & observations



Subsurface (Groundwater, Porous Media) Model with Stochastic Forcing

- Elliptic PDE, with random forcing: $-\nabla \cdot (e^{\mathrm{m}} \nabla u) = f$
- Use hIPPYlib -> FEniCS -> PETSc for simulation



with N. Petra, J. Bessac, C.G. Petra, Statistical Treatment of Inverse Problems Constrained by Differential Equations-based Models with Stochastic Terms, in prep, 2018

Subsurface Model – Problem Setup

- Formulation with prior (Bayesian, regularization):

$$m^* = \arg\min_m S(\boldsymbol{d}_T, \boldsymbol{d}; m) + m^T [\gamma \nabla^2 + \delta]^{-1} m \,, \quad \text{s.t. } F(u, m; \xi) = 0$$

- 300 observations, 32x32 grid, $N_s = 1, 4, 8, 32, 64; \; \xi \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \exp\left(-\frac{d^2(x, x', y, y')}{\ell^2_{\Delta x, \Delta y}}\right) + \delta I\right)$

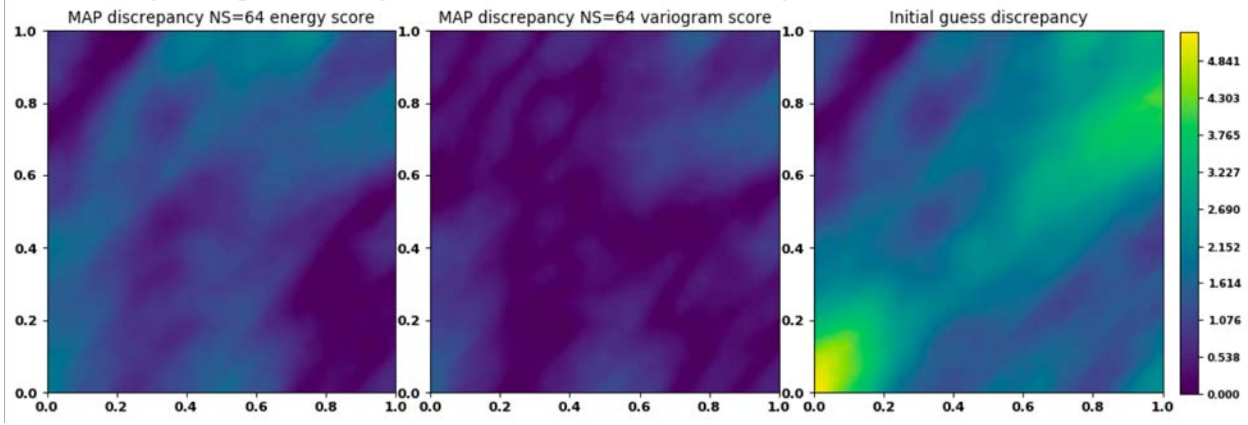
- BFGS (q-Newton), ~60-130 iterations

- Use adjoint-based gradient computation

- Outcomes: prediction of the control (parameter) field and predictive samples (ensemble)

Subsurface Model – Optimization Results

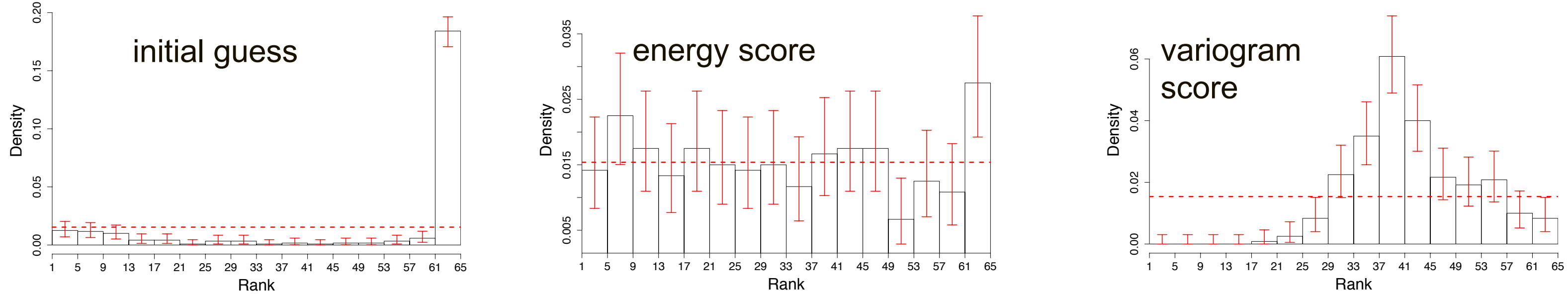
- Discrepancy of the parameter at the MAP point



RMSE of the parameter at the MAP point

Initial guess: **2.182**		
N_s	ES	Var-sc
1	1.137	**1.124**
8	1.108	**0.660**
64	1.176	**0.612**

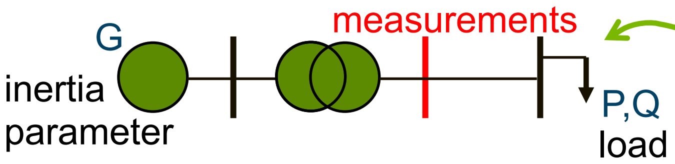
- Results for 64 samples: rank histograms (qualitative)

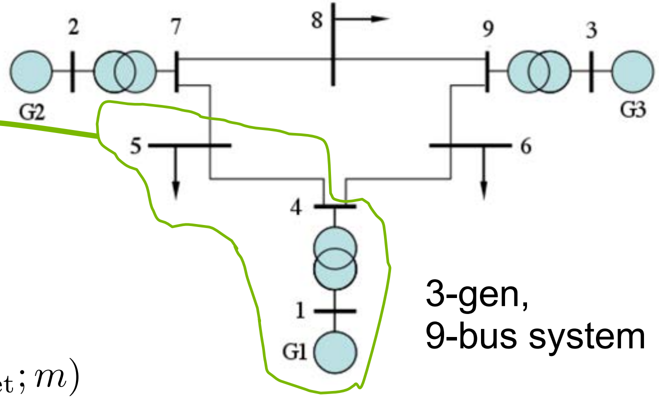


- ES gives better agreement with the forecast, whereas VS gives a better fit of the parameter

Power Grid Model – Problem Formulation

Estimate generator inertia from unobserved fluctuations in the load





3-gen,
9-bus system

- Generator model with exciter (index-1 DAE):
  7 differential and 8 algebraic equations (=>SDAE)

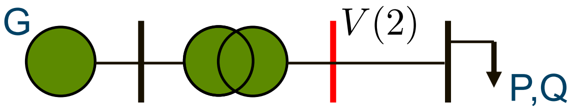
$$\frac{\partial X_{\text{Gen}}}{\partial t} = f(X_{\text{Gen}}, Y_{\text{Gen}}, Y_{\text{Net}}; m)$$

$$0 = g(X_{\text{Gen}}, Y_{\text{Gen}}, Y_{\text{Net}}; \underbrace{[P, Q]}_{\xi})$$

- Known the probability law of the load perturbations [P,Q]
- Inertia parameter -> voltage dynamics
- Measure voltage at one of the busses

$$S(\boldsymbol{d}_T, \boldsymbol{d}; m)$$

Power Grid Model – Problem Setup



- Observational operator:
  real & imaginary V at bus 2

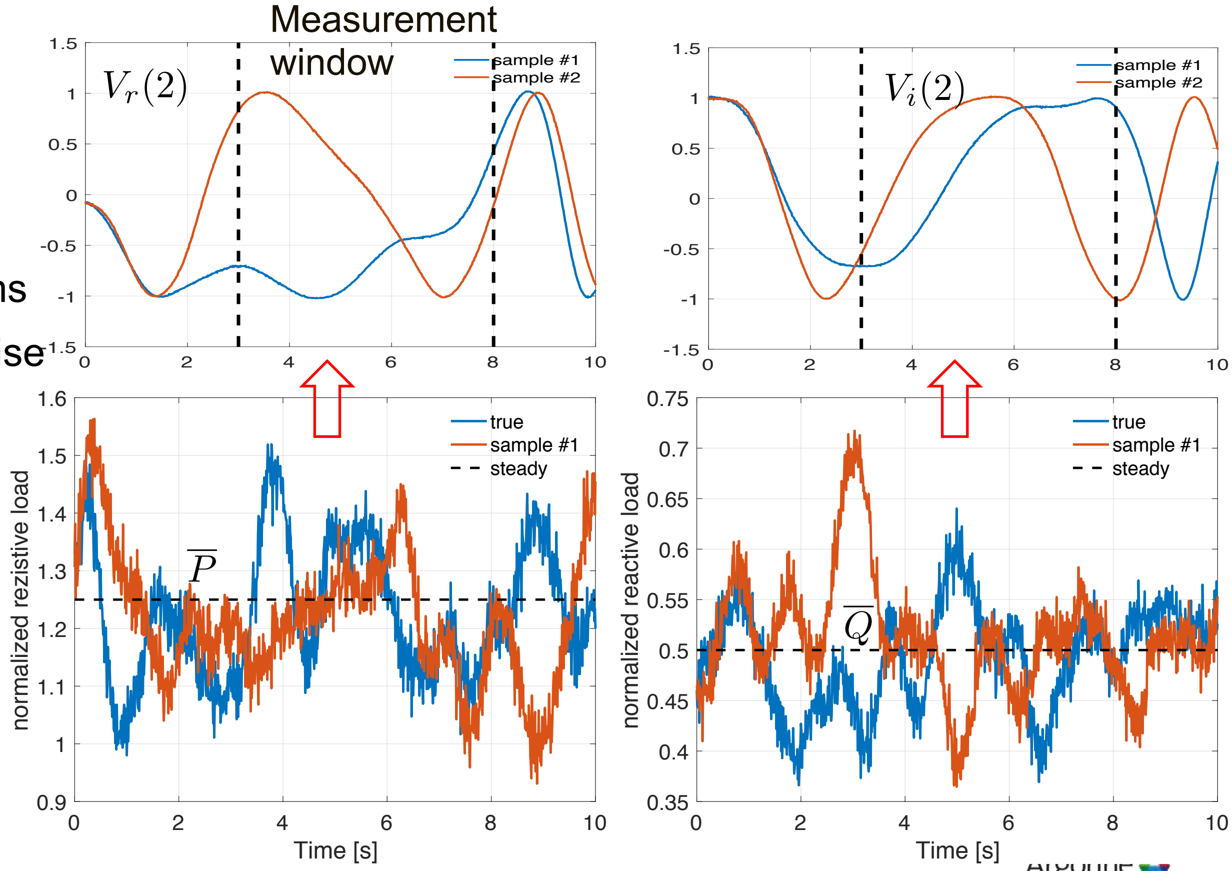
- Time window 10s, time step 10ms
- 500 samples, no observation noise

- Stochastic load (known PDF):

$$\xi = [P, Q]$$

$$P(t) \sim \mathcal{N}\left(\overline{P}, 0.1^2\, e^{-\frac{(t-t')^2}{2\times 0.001}} + 0.1\right)$$

$$Q(t) \sim \mathcal{N}\left(\overline{Q}, 0.05^2\, e^{-\frac{(t-t')^2}{2\times 0.001}} + 0.1\right)$$

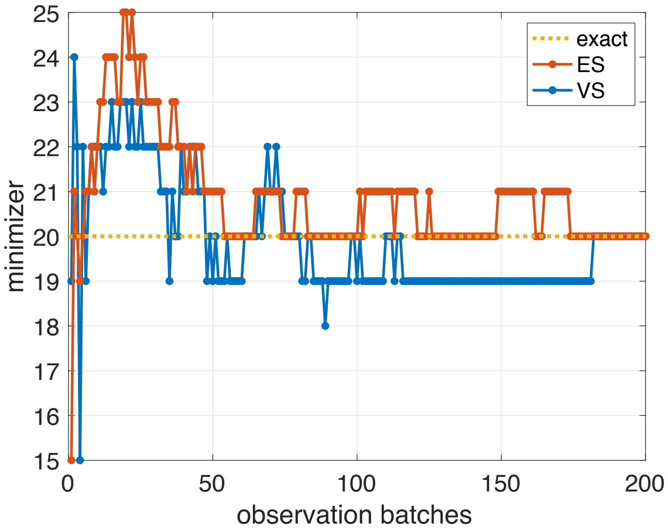


Power Grid Model – Convergence in Score

- Score variability under increasing number of observations, expected score:

$$S_n(\boldsymbol{d}_T, \boldsymbol{d}) = \frac{1}{n} \sum_k S(\boldsymbol{d}_T^{(k)}, \boldsymbol{d})$$

- Grid search: parameter value 10,11,12,…,30 (true=20); observation batches: 1,2, …, 100

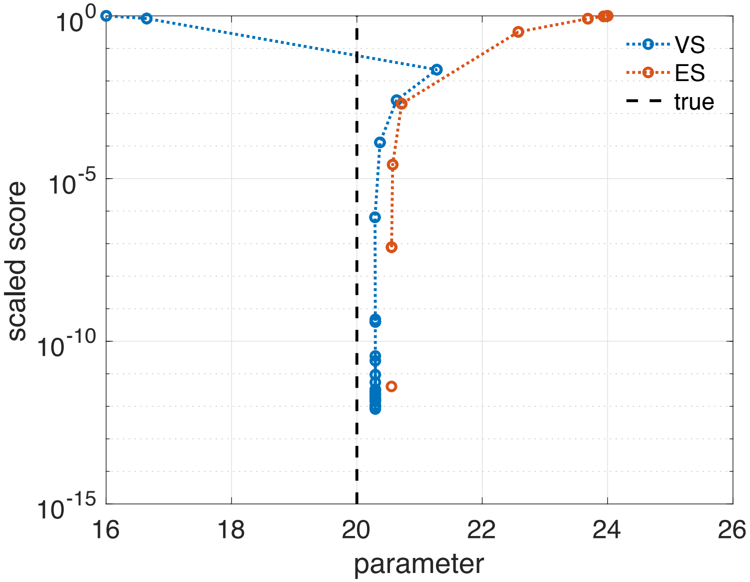


- Both scores seem to converge with 50+ observation batches; VS appears more robust

Power Grid Model – Optimization Result

- Use BFGS; finite difference gradient; 100 observation batches

$$m^* = \arg\min_{m} S(\boldsymbol{d}_T, \boldsymbol{d}; m)\,, \quad \text{s.t. } F(u, m; \xi) = 0$$



Summary on Probabilistic Predictions

- Increasingly, predictions are becoming stochastic (whether system is stochastic or not)
- Assessing the quality of probabilistic predictions is difficult
- Use ideas from statistics to generate rigorous metrics
- Can be used to compare models and calibrate physical models in order to improve predictions

with N. Petra, J. Bessac, C.G. Petra, Statistical Treatment of Inverse Problems Constrained by Differential Equations-based Models with Stochastic Terms, in prep, 2018

Summary

Making complex dynamic simulations predictive by using data

- Predictability implications for complex systems
- Part I: Numerical errors become important in complex systems
  – provide an efficient way to quantify and control
- Part II: Estimating model form errors – dominant errors in predictive simulations
  – can be estimated by using spatio-temporal stochastic processes
- Part III: Probabilistic predictions – are becoming pervasive
  – developed ways to calibrate models with probabilistic solutions