Towards Next Generation of Data Assimilation Systems

Răzvan Ştefănescu

\(^1\)Computational Science Laboratory, Department of Computer Science, Virginia Polytechnic Institute and State University
Outline

1. Current Research Interests
2. Data Assimilation
3. Reduced order data assimilation
4. On-going and future research
5. Education
6. Department improvement efforts
Current Research Interests

- Postdoc - FSU “Data Assimilation of lightning in WRF-4D VAR using nonlinear observation operators”, Prof. I.M. Navon.
- Researcher Associate - VT “Adjoint-based space-time adaptive solution algorithms for sensitivity analysis and inverse problems”, Prof. A. Sandu.

- Data Assimilation
- Reduced Order Modeling
- Uncertainty Quantification

- Applications: Numerical Weather Prediction, Fluid Dynamics, Population dynamics, Pharmacodynamics, Chemical Kinetics, etc.
Data Assimilation

- Data assimilation - process of combining information from models, measurements, and priors - all with associated uncertainties - to obtain the best estimate of the state of a physical system.

- Two families of methods: variational and ensemble based filters.

- Variational methods - rooted in control theory and require developments of tangent linear and adjoint models.

- Ensemble based sequential data assimilation schemes make use of statistical estimation theory.
4D-Var Data Assimilation

- The objective function $J$ to be optimized is defined based on model-data misfit penalty terms as:

$$J(x_0) = \frac{1}{2} (x^b - x_0)^T B_0^{-1} (x^b - x_0)$$

$$+ \frac{1}{2} \sum_{i=1}^{N} (y^i - H(x_i))^T R_i^{-1} (y^i - H(x_i)),$$

subject to

$$x_{i+1} = M_i x_i, \quad i = 0, \ldots, N - 1,$$

- The optimality conditions:

  Adjoint model: $\lambda_i = M_i^T \lambda_{i+1} + H^T R_i^{-1} (y^i - H(x_i)), \quad i = N - 1, 1$;

  $\lambda_N = H^T R_N^{-1} (y_N - H(x_N))$ and $\lambda_0 = M_0^T \lambda_1$.

Cost function gradient: $\nabla_{x_0} J = -B_0^{-1} (x^b - x_0) - \lambda_0 = 0$;
4D-Var Data Assimilation

- **Initial time** $t_0$
- **Analysis time** $t_a$
- **Model trajectory from corrected initial state**
- **Model trajectory from first guess** $x_b$

All observations $y_o$ between $t_a-9h$ and $t_a+3h$ are valid at their actual time.
The Weather Research and Forecasting (WRF) Model is a next-generation mesoscale numerical weather prediction system designed to serve both atmospheric research and operational forecasting needs (Skamarock et al. [2005, 2008]).

It features two dynamical cores, a data assimilation system, and a software architecture facilitating parallel computation and system extensibility.

The model serves: a wide range of meteorological applications across scales from tens of meters to thousands of kilometers; atmospheric simulations based on real data (observations, analyses) or idealized conditions.

WRF is currently in operational use at NCEP, NCAR.
Lightning Data Assimilation

- My research uses convective available potential energy (CAPE) as a proxy between lightning data and model variables (pressure, temperature, water vapor mixing ratio and geopotential height) Price and Rind [1992], Barthe et al. [2010].

- Three strategies: 3D-Var, 1D+3D-Var, 1D+4D-Var (Marécal and Mahfouf [2002, 2003]).

- 1D+nD-Var schemes avoid nonlinearity effects introduced by the lightning operator and generates temperature pseudo-observations assimilated as conventional observations.

- Development of tangent linear and adjoint lightning observation operators.

Incremental 4D-Var Data Assimilation Courtier et al. [1994]
Stage IV precipitation validation

Figure: 1 h precipitation (mm) ending at 2000 UTC 15 June from the control run, various assimilation procedures, and stage IV precipitation.
Reduced order data assimilation

\[ x_0 = x_b \]

Outer loop:

High-resolution non-linear trajectory

Reduced basis \( U \)

\[ \tilde{x}_0^i = U^T (x_0^i - \bar{x}) \]

Reduced-order nonlinear model

Reduced-order adjoint model

Iterative minimization algorithm

Inner loop:

\[ x_{i+1}^0 = \bar{x} + U\tilde{x}^a_i \]

High-resolution non-linear forecast

J

\( \nabla J \)
Why do we need reduced order data assimilation?

- Replace the current linearized cost function to be minimized in the inner loop
- Low-rank surrogate models that accurately represent sub-grid-scale processes
- Highly non-linear and non-smooth observation operators
- Increased space and time resolutions
- Reduced computational complexity
- Future DA system: real time analysis, accurate predictions with quantified uncertainties and efficient resource allocation.
- For three-dimensional studies this framework will dramatically reduce the computational time (by at least two orders of magnitude) without significantly affecting the quality of the optimal solution.
Reduced Order Modeling

Figure: The broad setup

- Approximation of Large-Scale Dynamical Systems, A.C. Antoulas
Reduced Order Modeling

**Figure:** Flowchart of approximation methods and their interconnections
Proper Orthogonal Decomposition

- The desired simulation is well approximated in the input collection Lumley [1967].
- Data analysis is conducted to extract basis functions, from experimental data or detailed simulations of high-dimensional systems.
- Galerkin projections that yield low dimensional dynamical models.
- Galerkin POD models (Aubry et al. [1988]): Its nonlinear reduced terms still have to be evaluated on the original state space making the simulation of the reduced-order system too expensive.
Reduced Order Modeling


- First time application of DEIM (Chaturantabut and Sorensen [2010]) to POD reduced order nonlinear shallow water equations (SWE) model.

- The basic step leading to application of POD/DEIM to 3D operational weather prediction models.

- Estimation of errors introduced by POD/DEIM and POD ADI SWE models (Gustafsson [1971]).

- Impacts of reduced order POD/DEIM SWE implicit and explicit schemes.
Reduced Order Modeling

The diagram illustrates the CPU time (in seconds) vs. the number of spatial discretization points for various methods. The methods include ADI SWE, POD ADI SWE, POD/DEIM ADI SWE, POD EE SWE, and POD/DEIM EE SWE. The data points are as follows:

- 2745, 4256, 10769, 16761, 66521

The graph shows a clear trend where the CPU time increases significantly as the number of spatial discretization points increases, with the methods differing in their performance patterns.
Reduced Order Modeling


- Application of tensorial calculus techniques to reduce the computational complexity of standard POD reduced polynomial nonlinear terms. The method is called tensorial POD.

- Development of ADI SWE model using sparse matrix environment and no equations decoupling.

- Computational complexity analysis of the reduced p-order polynomial nonlinearities for tensorial POD, standard POD and standard POD/DEIM.
Reduced Order Modeling

<table>
<thead>
<tr>
<th></th>
<th>Full ADI SWE</th>
<th>Standard POD</th>
<th>Tensorial POD</th>
<th>POD/DEIM $m=180$</th>
<th>POD/DEIM $m=70$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU time</td>
<td>950.0314s</td>
<td>161.907</td>
<td>2.125</td>
<td>0.642</td>
<td>0.359</td>
</tr>
<tr>
<td>$u$</td>
<td>-</td>
<td>5.358e-5</td>
<td>5.358e-5</td>
<td>5.646e-5</td>
<td>7.453e-5</td>
</tr>
<tr>
<td>$v$</td>
<td>-</td>
<td>2.728e-5</td>
<td>2.728e-5</td>
<td>3.418e-5</td>
<td>4.233e-5</td>
</tr>
<tr>
<td>$\phi$</td>
<td>-</td>
<td>8.505e-5</td>
<td>8.505e-5</td>
<td>8.762e-5</td>
<td>9.212e-5</td>
</tr>
</tbody>
</table>

Table: CPU time gains and the RMSE of model variables at $t_f = 3h$ for a 3h time window. Number of POD modes and DEIM points are $k = 50$ and $m = 180, 70$ and 103, 776 spatial points are used.

![Graph](image)

(a) On-line stage

(b) Off-line stage

Figure: CPU time vs. the number of spatial discretization points.
Reduced Order Modeling


- Development of the sparse matrix discrete empirical interpolation method (DEIM) to approximate parametric matrices such as time dependent Jacobians.

- The sparse algorithm relies on the discrete empirical interpolation method using only samples of nonzero entries of the matrix series.

- Application of sparse method to reduced order modeling.

- The new sparse strategy is validated against five existing methods for computing reduced Jacobians: a) matrix DEIM, b) DEIM, c) tensorial calculus, d) full Jacobian projection onto the reduced basis subspace, and e) directional derivatives of the model along the reduced basis.
Reduced Order Modeling

Figure: (a) Full Jacobians approximations - 2D-SWE - $23 \times 21$ space points; (b) Offline computational time performances of MDEIM and sparse MDEIM.
ROM 4D-Var DA systems


- Reduced order KKT conditions consistent with respect to the full optimality conditions. It means accurate low-rank approximations of the both adjoint and forward models.

- New bases selection strategy for Proper Orthogonal Decomposition (POD) reduced order data assimilation systems using both Galerkin and Petrov-Galerkin projections and non-linear models.

- Every type of reduced optimization involving adjoint models and projection based reduced order methods including reduced basis approach will benefit.
Figure: (a) The information from the adjoint equations has to be incorporated into POD basis. (b) Number of iterations and CPU time comparisons for the reduced Order SWE DA systems vs. full SWE DA system. \( n = 151 \times 111 \) space points, number of POD basis modes \( k = 50 \), MXFUN = 15 and \( \varepsilon_3 = 10^{-1} \).
On-going and future research

- Applications in nuclear proliferation area to develop new approaches for data assimilation, Bayesian inference and uncertainty quantification of radioactive substances and atmospheric flows.

- SRNL atmospheric transport and deposition model (Puff-Plume) – the dissolution of irradiated nuclear fuel, separation of the uranium and plutonium from fission products, emissions from the facility, including gases, particulates and waste heat.
On-going and future research

- Development of “Puff-Plume” adjoint model for adjoint sensitivity, data assimilation, uncertainty quantification and aposteriori error estimates and adaptivity.


- Model several main sources of uncertainty such as mean, standard deviation and turbulent wind components and measure the uncertainty in the gaseous and particulates dry and wet depositions of $^{238}\text{Pu}$, $^{239}\text{Pu}$, $^{252}\text{Cf}$, $^{244}\text{Cu}$, $^{60}\text{Co}$, FP, HTO, HT, $\text{Cl}_2$, $\text{H}_2\text{S}$, $\text{SO}_2$, $^{131}\text{I}$ and Noble Gases.

- Second Order Adjoint Sensitivity analysis procedure: allows the computation of all first and second order response sensitivities in less than $2(N_\alpha + 1)$ adjoint model runs - Cacuci [2014, 2015].
On-going and future research

- Exploit mesh hierarchies in computing probability densities on a fine mesh by summing cheaper density differences between successive coarser meshes (multilevel MCMC) (Heinrich [2001], Giles [2008], Ketelsen et al. [2013]).

- We will explore attaching model discrepancy terms directly to the model equations, and using newly proposed model error transport equations to propagate MFU into errors in the field solution variables and QoI, as it was done for discretization errors (Roy [2010], Phillips and Roy [2011, 2014]).

- Inverse algorithm: search for the amount and timing of fuel reprocessing that produces the best agreement with measured emission variables;

- Develop parallel 4D-Var framework using Augmented Lagrangian (Rao and Sandu [2015]) and use of HMCMC sampling for non-Gaussian distributions (Attia and Sandu [2013]), incorporating model dynamics via the gradient of the posterior distribution.
On-going and future research

- Mechanical draft cooling towers: remove waste heat from industrial processes, including suspected proliferators of weapons of mass destruction.

- The temperature of the air being exhausted from the MDCT is proportional to the amount of thermal energy being removed from the process cooling water. Savannah River National Laboratory, the Northrop-Grumman Corporation and the Aerospace Corporation.

- the inner model: computes the amount of cooling experienced by the water as it passes through the tower as a function of inlet cooling water temperature and ambient weather conditions (air temperature and humidity).

- the outer model: takes a remotely measured throat temperature and iterates on the inlet water temperature to match the target temperature of interest.
Outer model: build an adjoint model and implement a 3D-Var algorithm to retrieve the inlet water temperature and the corresponding mass flowrate from remotely measured throat temperature.

High order adjoint sensitivities to estimate the uncertainties in the temperature and mass flowrate of the effluent water by modeling the uncertainties in the input.
On-going and future research

- Coupling a nuclear reactor model, with an operational atmospheric model such as WRF-Chem or ATHAM model and city model using the capability of Fluidity (AMCG/ Imperial College group) Urban Terreno initial mesh generated to provide the topology of city canyons.
On-going and future research

- Obtain detailed air flow and radioactive material dispersion patterns.

- Develop of adjoint models for model calibration, data assimilation, adaptive sensor locations and uncertainty quantification.

- Develop new methods of detection of nuclear weapons and radioactive materials introduced by terrorists, or accidents in different environments including urban scales.

- Implementation of singular vectors (Palmer et al. [1998]), adjoint sensitivity algorithms such as optimization-constrained optimization technology (Dăescu and Navon [2004], Cioaca and Sandu [2014]) to acknowledge the mobile sensors deployment costs.

- Development and implementation of new sensor trajectory design methods which are robust to parameter uncertainties using the framework of stochastic optimization.
On-going and future research

- Multifidelity techniques: Local adaptive ROMs (Peherstorfer et al. [2014], Rapúñ and Vega [2010]) matching statistical QoIs and preserving the subsystems interconnections (via aposteriori error estimates); Reducing parameter space by discovering low-dimension structures via active subspace methods (Constantine et al. [2014]).

- We propose interpolatory methods for ROMs that match lower-order statistical moments of the original system with respect to QoIs, so surrogate ROMs function as unbiased estimators.

- Exploit the structure of the weak constraints variational approach (Trémolet [2006]), consistent reduced KKT conditions (Stefănescu et al. [2014]) and formulate a piecewise-in-space-time approximation strategy that uses different ROMs on different space locations and subintervals, and constructs them concurrently;
On-going and future research

- Develop a new adjoint free optimization system using Jacobian samples employed by the sparse matrix DEIM algorithm (Stefănescu and Sandu [2014]);

- Develop probabilistic error maps based on machine learning techniques such as neural network and Gaussian process kernel to quickly estimate the quality of reduced order solutions without the need of running expensive full version model or optimization framework (Drohmann and Carlberg [2014]);

- Develop a-posteriori error estimates and adaptive reduced order modeling data assimilation system using the super-Lagrangian technique (Alexe and Sandu [2014], Rao and Sandu [2014]);
Education

- Experience: “Al. I. Cuza” University of Iași, University of Medicine and Pharmacy “Gr. T. Popa” Iași, Florida State University, Virginia Tech.


- Teaching: motivating learning, structuring learning, encouraging activity and independence in learning, establishing interpersonal relations conducive to learning.
Education

- Teach undergraduate and graduate courses as needed within the Nuclear Engineering Program.

- Introduction to Nuclear Engineering.

- Develop graduate-level course in predictive modeling for nuclear engineering applications including validated modeling, adjoint sensitivity, uncertainty quantification and reduced order modeling using examples from nuclear energy related areas and my research.

- Developing a graduate-level course on “Algorithms for threat detections of radioactive materials and nuclear weapons proliferation development programs“ based on adjoint sensitivity, adaptive sensor locations and/or using satellite and/or drone platform imagery.
Department improvement efforts

- Place the foundation of a large and strong research group;
- Recruit, train and retain high quality students with expertise in sensitivity analysis, UQ and optimal predications based on data assimilation and large-scale computational model calibration.
- Advance the understanding in nuclear reactors, nuclear fuels and materials through multidisciplinary collaborations and making use of simulation codes, new developed methodologies and HPC facilities.
- Build multidisciplinary groups to attract financial support for novel nuclear engineering research projects which in turn will enrich substantially the existing curricula and establish a new paradigm for nuclear engineering educations.
Department improvement efforts

- Increase collaboration with industrial partners;

- Expand the current NNSA and SRNL collaborations on “Modeling of Nuclear Facilities for proliferation detection”.

- Develop a new research program in diagnostic medical imaging and interventional nuclear medicine - collaboration with NIH.

- Seek departmental collaborations for grant applications to DOE, Office of Naval Research and NSF programs (such as Computational and Data-Enabled Science and Engineering and Cyber-Innovation for Sustainability Science and Engineering).
Thank You


C. J. Roy. Review of discretization error estimators in scientific computing.


