NREL
NATIONAL RENEWABLE ENERGY LABORATORY

NREL is a national laboratory of the U. S. Department of Energy, Office of Energy Efficiency and Renewable Energy, operated by the Alliance for Sustainable Energy, LLC.

Fast Learning for Immersive Engagement in Energy Simulations

Bruce Bugbee,[1] Brian Bush,[2] Kenny Gruchalla,[1] Venkat Krishnan,[3] and Kristi Potter[1]

[1]Computational Science Center, [2]Strategic Energy Analysis Center, [3]Power Systems Engineering Center

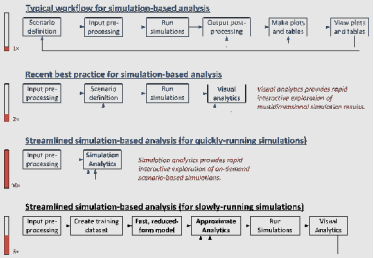
Overview

Motivation

- Visualization-driven design, exploration, and analysis of energy simulations is becoming central to leading-edge projects at NREL.
- Teams develop insights and test hypotheses more quickly when working in immersive, interactive workspaces.
- Energy simulation results typically contain low dimensional structures

Objectives

- Develop general methods for rapidly computing simplified versions of NREL's flagship, computation-intensive, energy simulations (ReEDS, BSM, etc.).
  - Statistically rigorous approximation to full simulation.
  - Low computational cost / quick response times.
  - Uncertainty quantification.
  - Interactive visualization.
- Insightfully characterize complex, high-dimensional input/output spaces associated with energy models.
  - Visualize and communicate feature discoveries.



Approach (4 phases of 6 months each)

1 Statistical methods for representing high dimensional data sources
- Artificial neural networks
- Linear, kernel, tensor, and non-linear principal components analysis
- Functional regression and principal components analysis
- Measurement of dimensionality
- Quantifying accuracy/fidelity of approximations

2 Applications to complex visualization
- Generic interactive visualization of multidimensional timeseries

3 Consolidating results into tools

4 Publication of results and release as open source

Aspects of Case Study Applications

Problem	• simplified visualization • embedding in another model • scenario planning • resource planning / design of experiments
Representation	• reduction of dimension • interpretability by humans • exploration of structure • simplicity of computation
Approximation	• availability of training data • dimensionality of input and output spaces • metrics for quality
Visualization	• discovery of structure • assessing quality of approximation • representation of dimensionality
Interaction	• navigating through multiple dimensions • selection, linking, and brushing

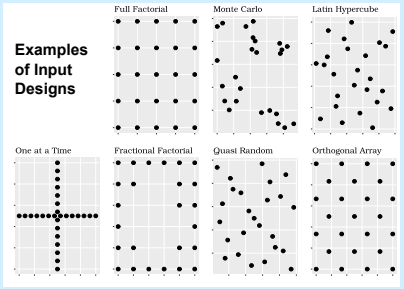
Impacts

- Rapid, interactive, visual design and exploration of scenarios with sponsors and stakeholders "in the loop".
- Reduced-form models for screening analyses.
- Resiliency assessment.
- Dramatically improved use of computational resources.

Input Data

Choice of Training and Testing Data

- Statistical models require a representative sample of "training" data.
- The training data should span the parameter ranges of interest.
  - Parameters not of interest should be set to their default values.
- The density of training data over the input parameter space should also match the level of interest in that part of the parameter space.
- The theory of the design of experiments supplies methods for sampling the input parameter space in a manner that uniformly covers the potentially high dimensional input space:
  - As a rule of thumb, the data set for testing the quality of the approximation should be about 25% of the size of the training dataset.

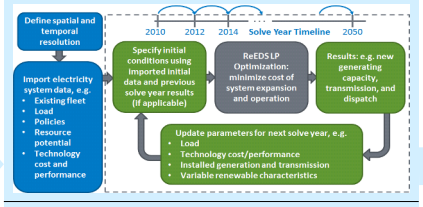


dimension reduction on input space

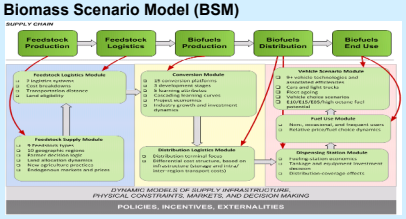
training, bypassing dimension reduction

Model / Simulation

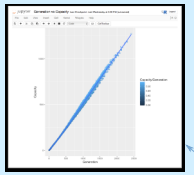
Regional Energy Deployment System (ReEDS)



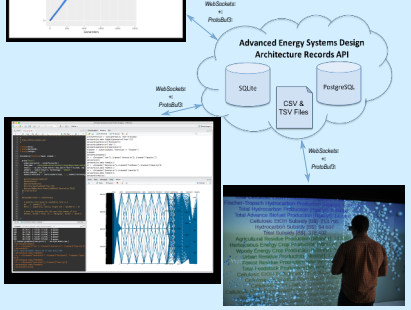
Biomass Scenario Model (BSM)



Output Data

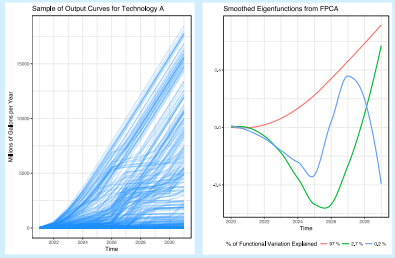


Output data from the simulations and the results of the fitting it with the reduced-form approximations are compared and visualized using both standard graphics tools and also the capabilities of NREL's Advanced Energy Systems Design Institute, such as immersive visualization in NREL's Energy Systems Integration Facility (ESIF).



dimension reduction on output space

Reduced-Dimension Input

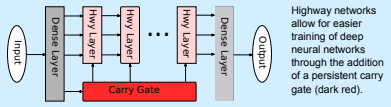


Often times both input and output data are more appropriately handled as functional objects rather than discrete points. This allows for the serial correlation structure to impact both modeling and inference, in order to better reflect the problem at hand.

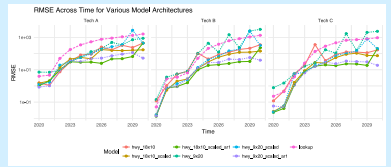
The common strategy for dealing with functional data involves some form of basis representation to turn high dimensional functions into low dimensional representations. Basis choice is driven by the particular application, with options including splines, wavelets, Fourier transforms, etc. A popular framework is to extend the idea of principal components to the functional domain (FPCA) and construct an empirical basis derived from covariance decompositions. The above graphic highlights an example FPCA representations.

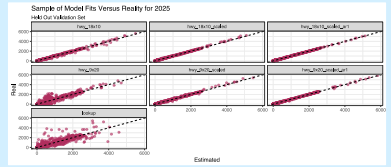
training

Reduced-Form Model



Highway networks allow for easier training of deep neural networks through the addition of a persistent carry gate (dark red).





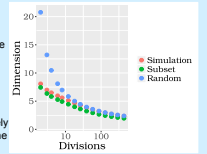
prediction, bypassing dimension reduction

prediction

Reduced-Dimension Output

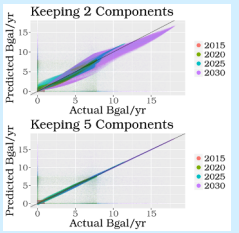
Measurement of Dimensionality

- The empirical dimensionality of a multivariate dataset is estimated using a "box counting" technique.
- This estimate provides guidance as to the dimensionality of the space in which the output could be embedded.
- The plot at the right suggests that the 21 dimensions of output of the Biomass Scenario Learning Model could be approximated by a manifold approximately 8 dimensions, considerably simplifying the output's representation.



Reduction of Dimensionality

- Tensor Principal Components Analysis (TPCA) is a multivariate generalization of the well know method PCA.
- TPCA separates a tensor (e.g., a muliply indexed array or data) into an optimally determined product of smaller tensors.
- The plots like those at the right demonstrate the successful application of PCA and TPCA to reduce the dimensionality of output from the Biomass Scenario Learning Model from 23 variables at 21 times to 12 pseudovariables at 5 pseudotimes.



This work was supported by the U.S. Department of Energy under Contract No. DE-AC36-08GO28308 with Alliance for Sustainable Energy, LLC, the Manager and Operator of the National Renewable Energy Laboratory. Funding provided by the Laboratory Directed Research & Development Program. The research was performed using computational resources sponsored by the Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory.

NREL LDRD Poster Session
1 June 2017
NREL/PO-6A20-70089