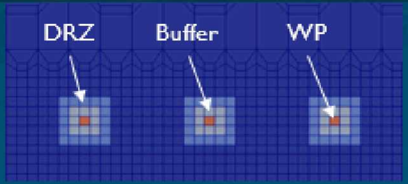
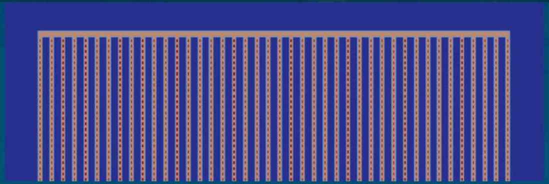
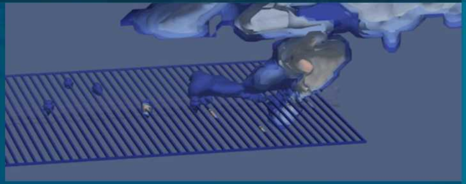
SAND2019-13375C

Lessons Learned in the Development of Source Term Surrogate Models for Repository Performance Assessment







Paul E. Mariner[1], Bert J. Debusschere[2], James Jerden[3], D. Thomas Seidl[1], Laura P. Swiler[1], Jonathan Vo[2]

[1]Sandia National Laboratories, New Mexico, U.S.A.; [2]Sandia National Laboratories, California, U.S.A.; [3]Argonne National Laboratory, Illinois, U.S.A

PRESENTED BY

Paul Mariner

November 5, 2019

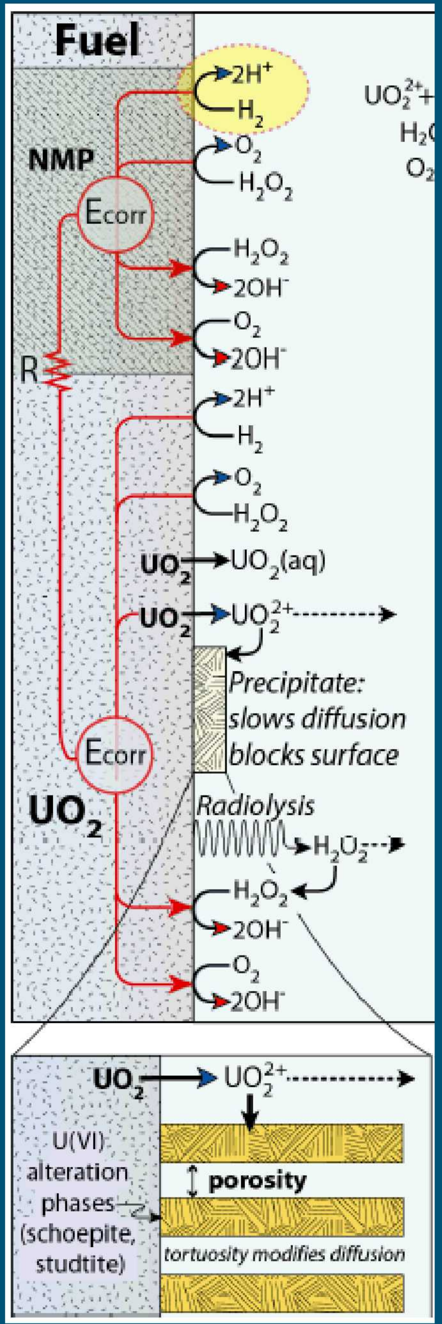
DECOVALEX Symposium 2019
Brugg, Switzerland

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Objectives

- Develop surrogate models
  - For the Fuel Matrix Degradation (FMD) process model
  - For use by PFLOTRAN in GDSA Framework[1]
- Assess performance
  - Accuracy
  - Speed
- Lessons learned

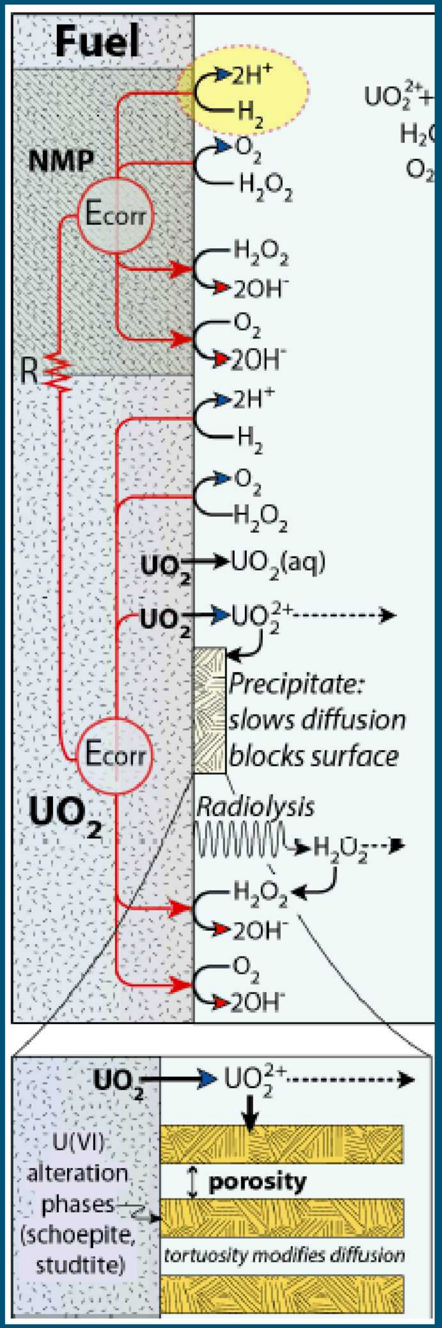
[1]GDSA Framework is U.S. DOE's Geologic Disposal Safety Assessment (GDSA) framework for probabilistically assessing the performance of geologic nuclear waste repository concepts (pa.sandia.gov)



(Figure adapted from Jerden et al. 2017)

Fuel Matrix Degradation (FMD) Process Model

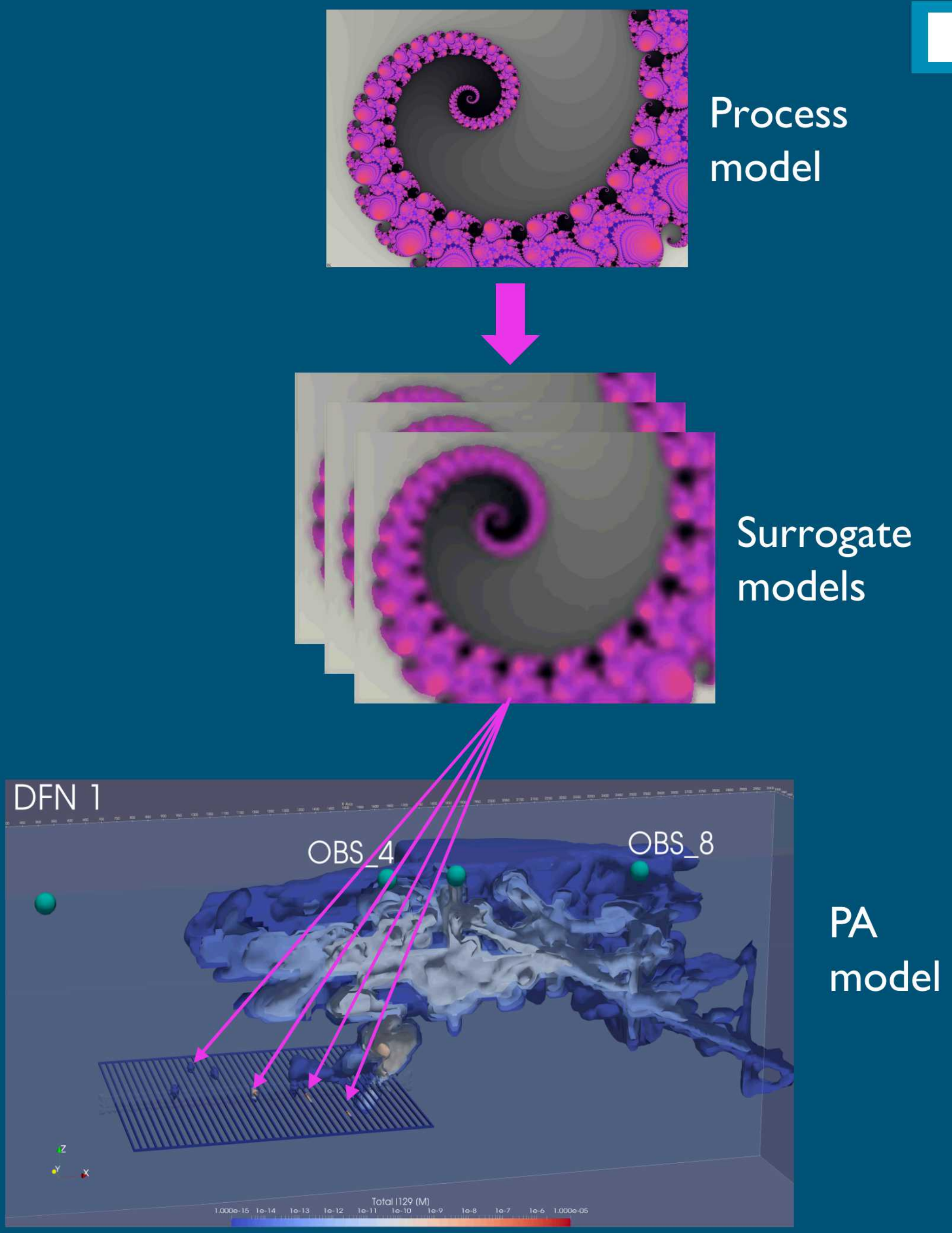
- Complex set of processes
  - Radiolysis
  - Oxidation of H_2 via noble metal particle (NMP) catalyst
  - 1-D reactive transport through alteration layer
  - Growth of the alteration layer
  - Diffusion of reactants and products through the alteration layer
- Expensive in a repository PA calculation



(Figure adapted from Jerden et al. 2017)

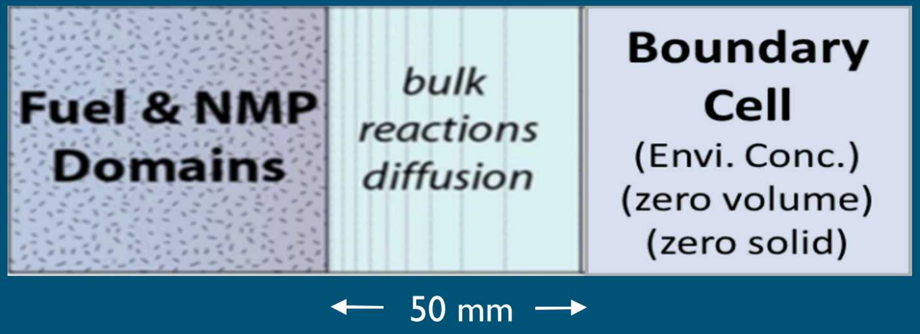
Surrogate models

- Developed 3 surrogate models
  - 2 parametric
    - Polynomial regression
    - Neural network regression
  - 1 non-parametric
    - k-Nearest Neighbors regression (kNNr)
- Used training data from FMD process model
    - MATLAB



Matlab FMD process model domain and inputs/outputs

- Domain
  - 1D, fuel surface to bulk water, 50 mm



- Inputs/outputs each time step

Inputs	Outputs
• Initial concentration profiles across 1D corrosion/water layer (UO_2(s), UO_3(s), UO_4(s), H_2O_2, UO_2^{2+}, UCO_3^{2-}, UO_2, CO_3^{2-}, O_2, Fe^{2+}, and H_2) • Initial corrosion layer thickness • Dose rate at fuel surface (= f (time, burnup)) • Temperature • Time, time step length • Environmental concentrations (CO_3^{2-}, O_2, Fe^{2+}, and H_2)	• Final concentration profiles across 1D corrosion/water layer • Final corrosion layer thickness • Fuel dissolution rate

Surrogate training / testing data

- Six-dimensional space
- Matlab FMD process model simulations
  - Latin hypercube sampling (LHS)
  - Each sim with 101 points in time, logarithmically spaced from 0 to 10^5 yr
  - 1,908 training simulations for parametric surrogates
  - 15,169 training simulations for kNNr
- Log transformations for regressions
  - Inputs (not temperature)
  - Outputs

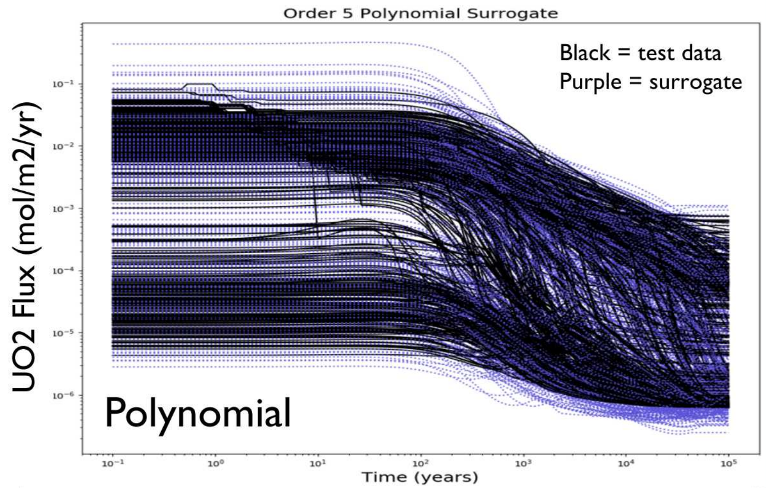
Parameter	Dist.	Min.	Max.
Init. Temp. (C)	Uniform	298	373
Burnup (Gwd/MTU)	Uniform	20	90
Env. CO_3^{2-} (mol/m³)	Log-uniform	10^{-6}	10^{0}
Env. O2 (mol/m3)	Log-uniform	10^{-6}	10^{-1}
Env. Fe^{2+} (mol/m³)	Log-uniform	10^{-6}	10^{-5}
Env. H_2 (mol/m³)	Log-uniform	10^{-6}	10^{-1}

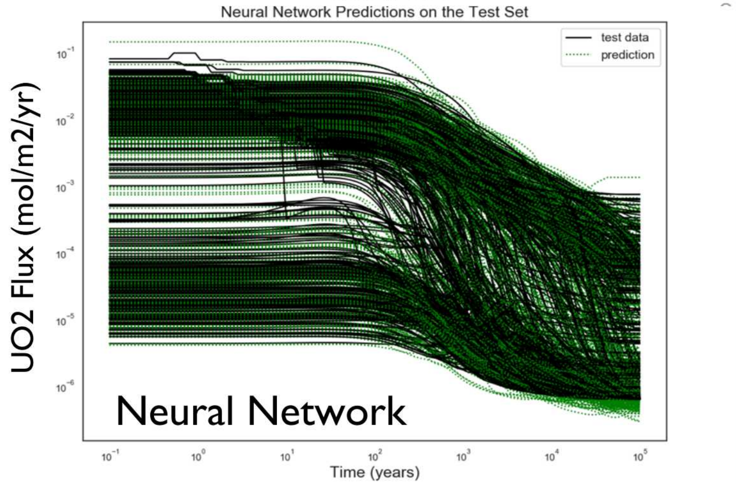
Error analysis

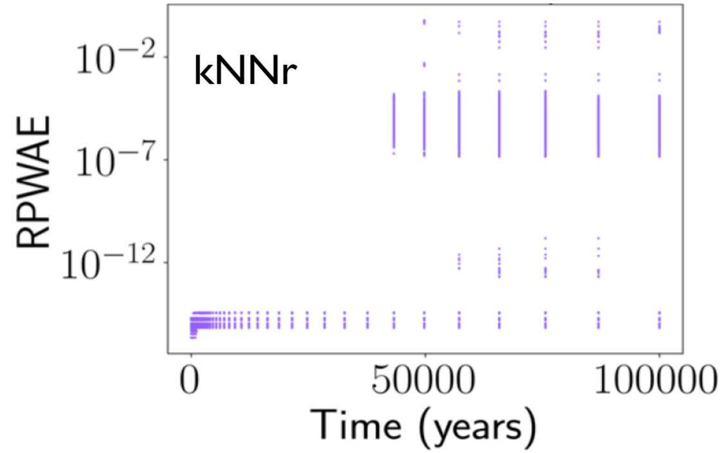
- Relative pointwise absolute error (RPWAE) at each point

$$RPWAE = \frac{|y_{pred} - y_{true}|}{y_{true}}$$

Surrogate	Terms / Coeffi-cients	Train R^2	Train Mean RPWAE	Test R^2	Test Mean RPWAE
Polynomial (Order 5)	462	0.952	0.858	0.942	0.898
Neural Network	801	0.978	0.40	0.972	0.635
kNNr (k=7)	NA	NA	NA	NA	1×10^{-5}

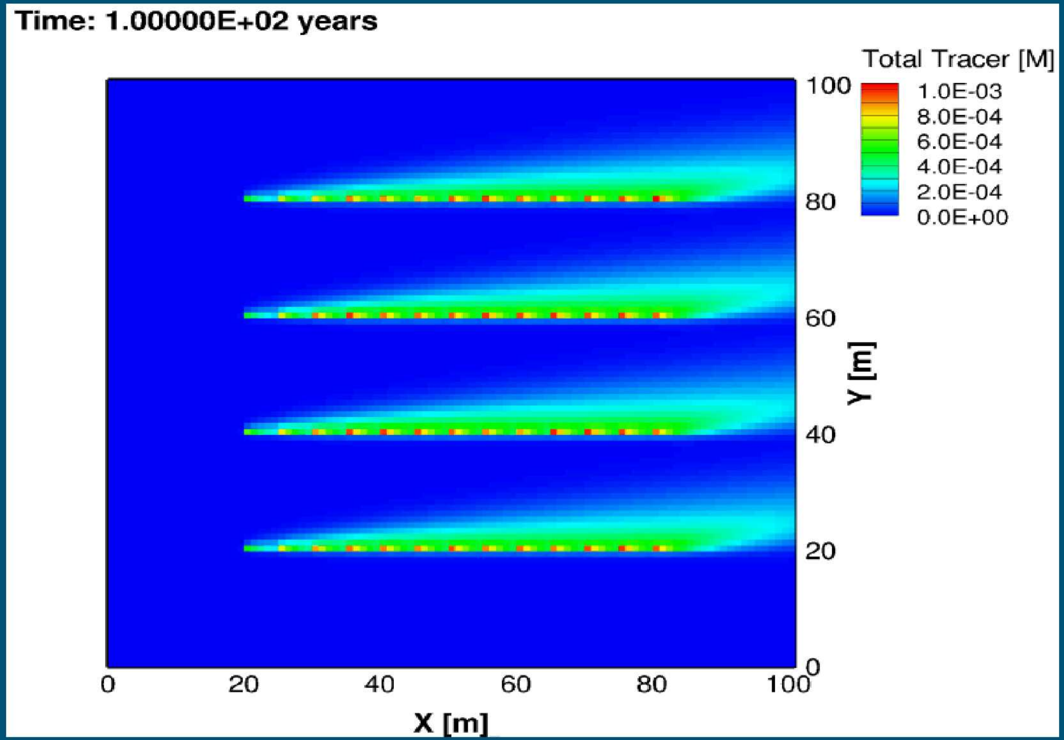






Speed

- Polynomial surrogate model coupled to PFLOTRAN
  - Adapted to earlier version of process model coupled to PFLOTRAN
- Tested on 2D example
  - 52 waste packages
- Coupled polynomial surrogate ~200× faster than coupled process model
- Neural network and kNNr surrogates not yet coupled and tested for run time effects

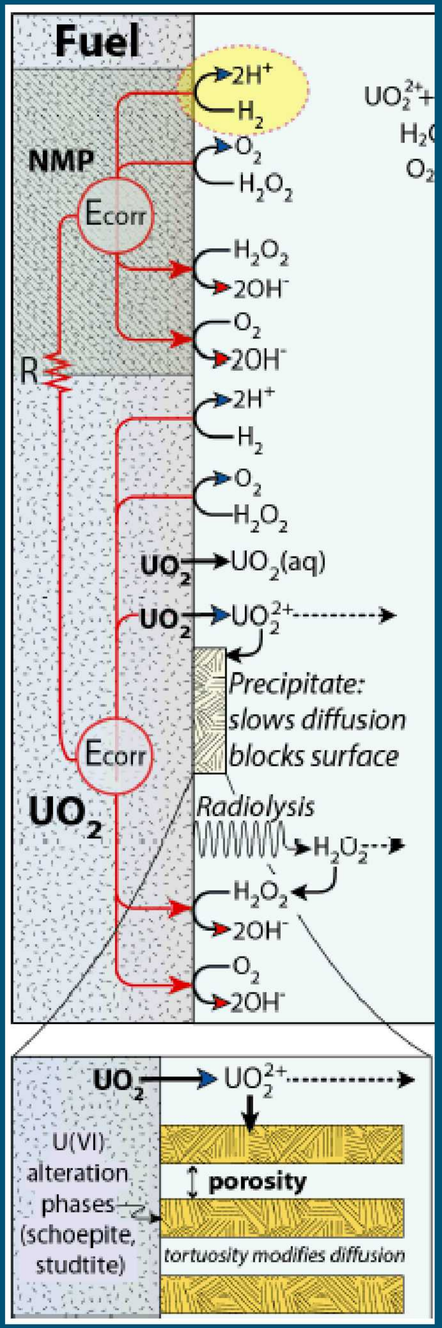


Run Time (seconds)

Module	Coupled FMD Process Model	Coupled Polynomial Surrogate Model
Flow	168	194
Transport	244	278
Waste Form	1522	8

Lessons learned / reaffirmed (1/3)

- Understand the process model
  - Identify all potential predictors – full set may not be obvious at first
  - The optimal predictors used by the surrogate may be quite different from the inputs used by the process model
    - Significant vs. insignificant inputs vs. lumped parameters
    - State variables from previous time step may not be needed for surrogates
  - The surrogate model may need to calculate values for predictors and store them for use in the following time step

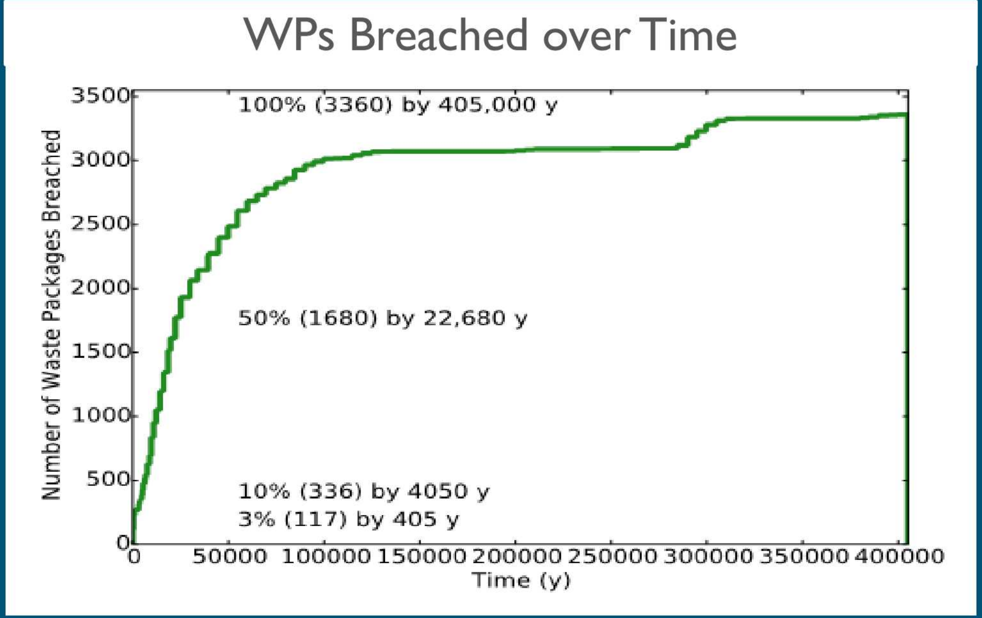


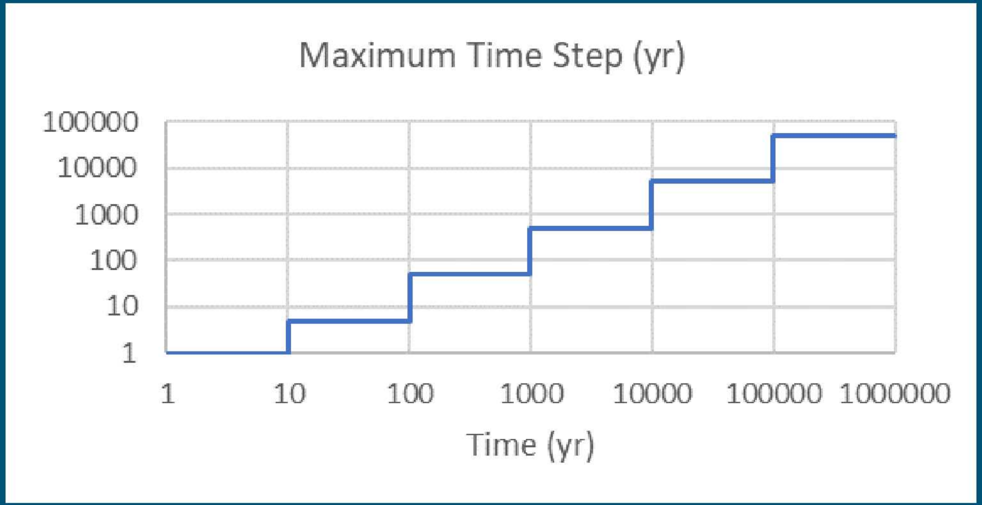
(Figure adapted from Jerden et al. 2017)

Lessons learned / reaffirmed (2/3)

- Consider how the surrogate model will be applied to the performance assessment model
  - e.g., PA model time frame and time step size
- Include explicit process model calculations in the surrogate model
  - e.g.,

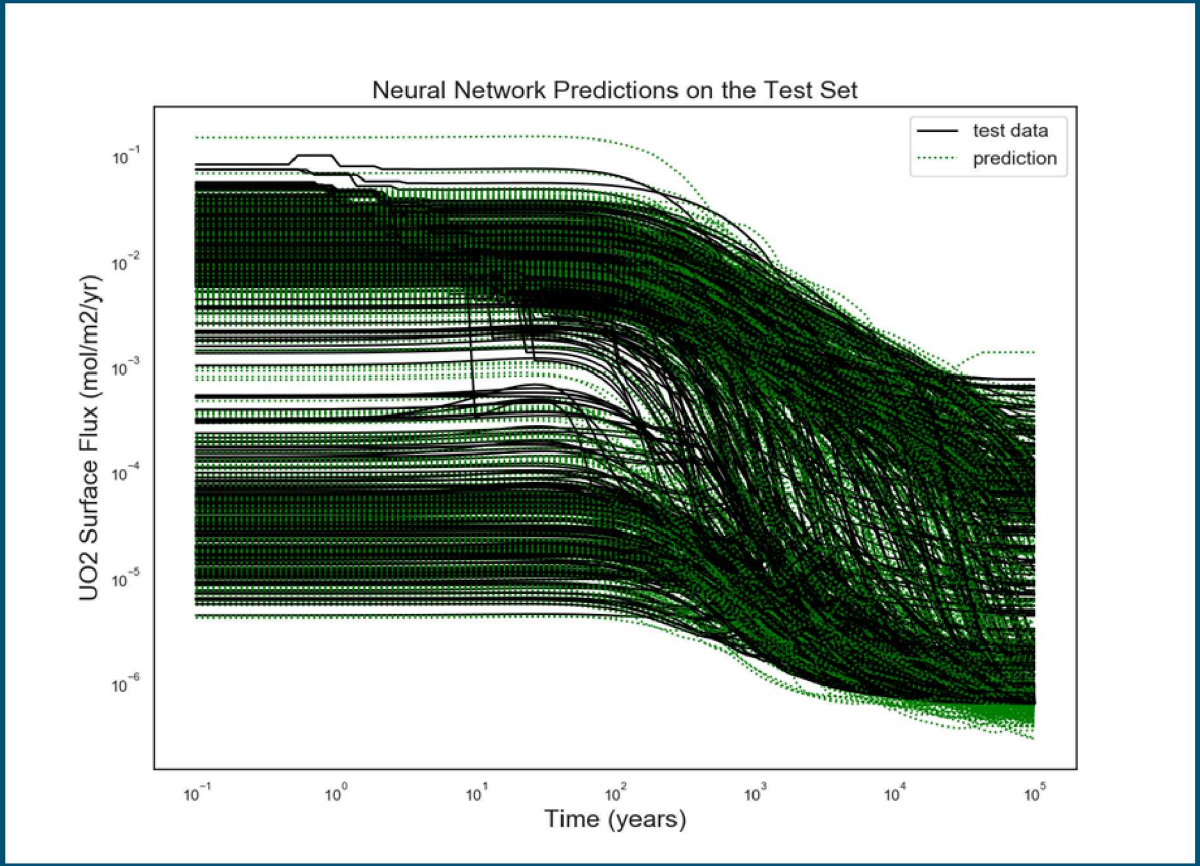
    $$Dose\ Rate = f(time, burnup)$$





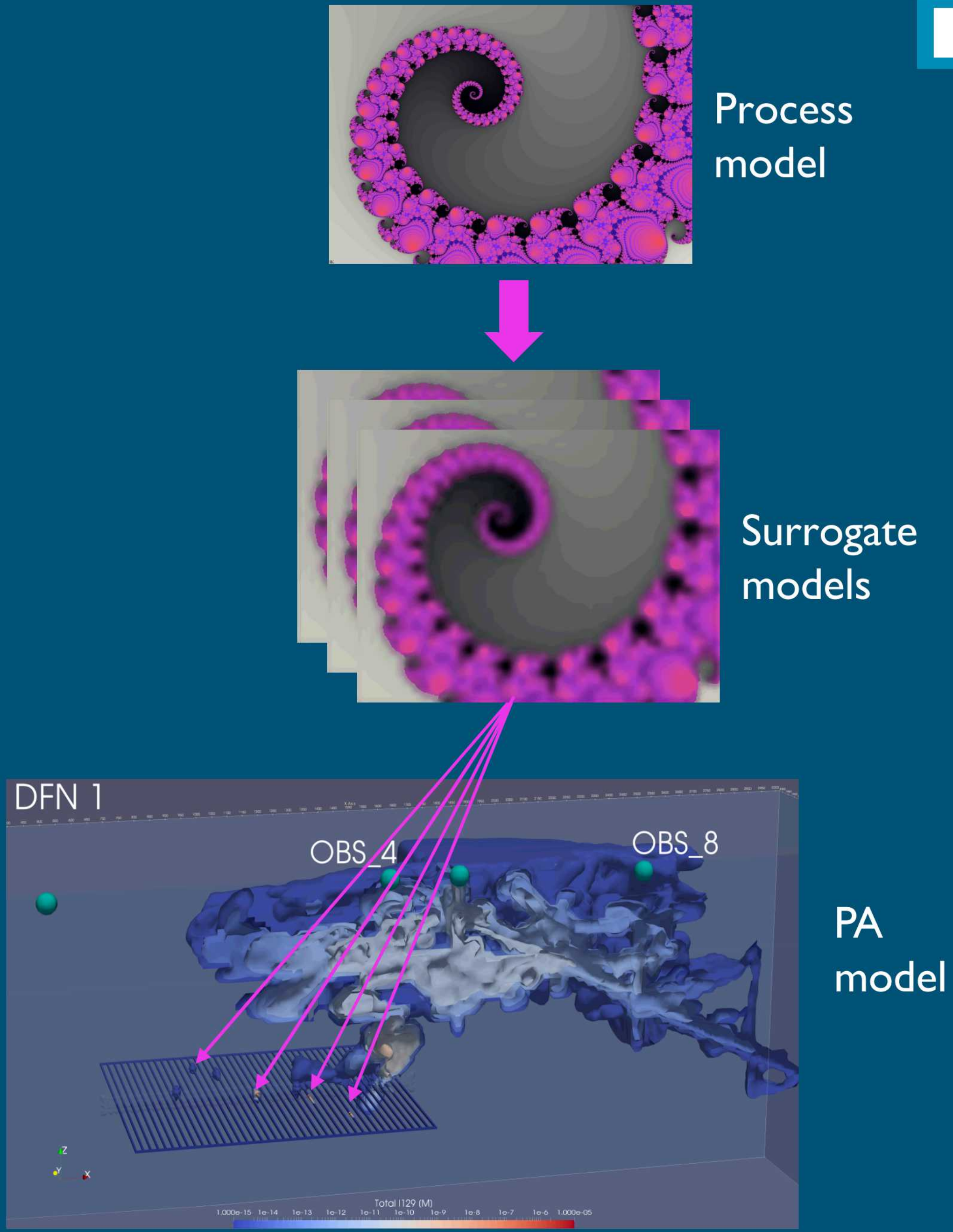
Lessons learned / reaffirmed (3/3)

- Prior to generating training data
  - Perform a spatial and temporal convergence study on the process model
  - Identify outliers that may indicate a potential problem with the process model
- Generate training and testing data for the realm of interest
  - Do not include outputs much beyond the domain of the application
  - Process model time steps may be very different than needed for PA model



Surrogate model comparison

- All three surrogate models
  - Highly accurate and fast
- Of the parametric FMD surrogate models
  - Neural network surrogate provided more accuracy than the polynomial surrogate
- Of all three surrogates
  - Non-parametric kNNr surrogate provided most accuracy
- Future work
  - Couple all surrogates to PFLOTRAN for fair accuracy/speed comparison
  - Evaluate relative cost of development
  - Identify additional potential improvements



References

- Ben-David, S. and S. Shalev-Shwartz (2014). Understanding Machine Learning: From Theory to Algorithms. Cambridge, United Kingdom, Cambridge University Press.
- Jerden, J., J. M. Copple, K. E. Frey and W. Ebert (2015a). Mixed Potential Model for Used Fuel Dissolution - Fortran Code. O. o. U. N. F. Disposition. FCRD-UFD-2015-000159. U.S. Department of Energy, Washington, DC.
- Jerden, J., G. Hammond, J. M. Copple, T. Cruse and W. Ebert (2015b). Fuel Matrix Degradation Model: Integration with Performance Assessment and Canister Corrosion Model Development. FCRD-UFD-2015- 000550. U.S. Department of Energy, Washington, DC.
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay (2011). "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, 12:2825-2830.
- Rasmussen, C. E. and C. K. I. Williams (2006). Gaussian Processes for Machine Learning, MIT Press.
- Santner, T., B. Williams and W. Notz (2003). The Design and Analysis of Computer Experiments. New York, New York, Springer.
- Seber, G. A. F. and C. J. Wild (2003). Nonlinear Regression. New York, New York, Wiley & Sons.
- Simpson, T. W., V. Toropov, V. Balabanov and V. F.A.C. (2008). "Design and analysis of computer experiments in multidisciplinary design optimization: A review of how far we have come or not," Proceedings of the 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Victoria, British Columbia, Canada. AIAA Paper 2008-5802.
- SNL. (2017). "GDSA Framework: A Geologic Disposal Safety Assessment Modeling Capability," pa.sandia.gov.
- Storlie, C. B., L. P. Swiler, J. C. Helton and C. J. Sallaberry (2009). "Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models," Reliability Engineering & System Safety, 94(11):1735-1763.
- Swiler, L. P., J. C. Helton, E. Basurto, D. M. Brooks, P. E. Mariner, L. M. Moore, S. Mohanty, S. D. Sevougian and E. R. Stein (2019). Status Report on Uncertainty Quantification and Sensitivity Analysis Tools in the Geologic Disposal Safety Assessment (GDSA) Framework. M2SF-19SN010304031. Sandia National Laboratories, Albuquerque, New Mexico.

Questions? Comments?