To establish the best practise for sim-exp benchmarking and comparisons

Juncheng E

Keywords

- Protocol for comparison of raw simulated to raw experimental data
- benchmarking experiment and simulation
- compare multidimensional data

Implementation in scientific research

Correlation plot

Partitioning of Amino Acid Side Chains into Lipid Bilayers: Results from Computer Simulations and Comparison to Experiment

Justin L. MacCallum, W.F. Drew Bennett, and D. Peter Tieleman

Department of Biological Sciences, University of Calgary, Calgary AB T3A 2H4, Canada

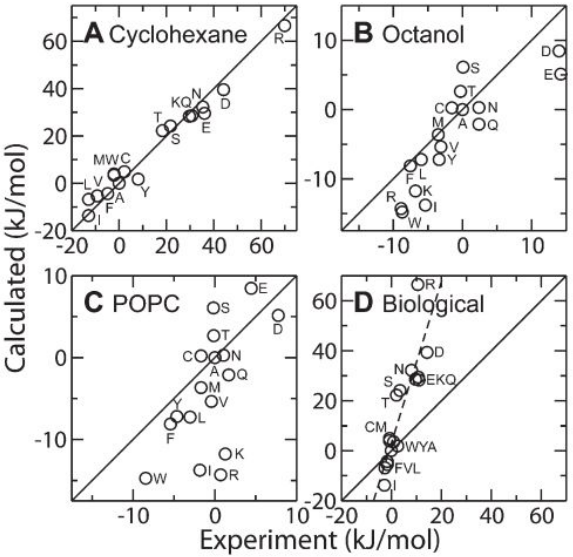


Figure 3. Comparison of calculated results (vertical axis) to experimental results (horizontal axis). Residues are indicated by a circle and labeled with the corresponding single letter amino acid code. Solid lines indicate perfect agreement between calculation and experiment. In all cases, the experimental and calculated

J. MacCallum et al., Journal of General Physiology (2007)

Pearson product-moment correlation coefficient

$$r^2 = \frac{\sigma_{xy}^2}{\sigma_x \sigma_y}$$

where,

$\sigma_x = \sum_{i=1}^{n}(x_i - \bar{x})^2$, $\sigma_y = \sum_{i=1}^{n}(y_i - \bar{y})^2$, and $\sigma_{xy} = \sum_{i=1}^{n}(x_i y_i - \bar{x}\bar{y})$, x and y are two sets of data to be correlated. As see in Fig. 2 the contour plot is a representation of a 2-D matrix, and to use the Pearson product-moment correlation, the matrix should be rearranged to an array of data.

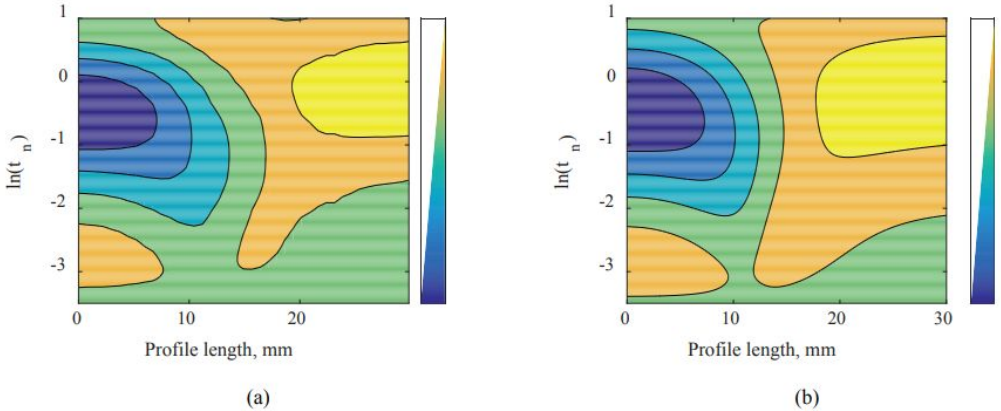


FIGURE 5. A comparison of, (a) experimental and (b) numerical, 2d contour plots of 1" diameter and 0.085" depth axisymmetric defect

TABLE 1. 2d correlation coefficients of defects considered

			Experimental								
	Diameter /inch		0.375			0.750			1.000		
		Depth/inch	0.085	0.160	0.265	0.085	0.160	0.265	0.085	0.160	0.265
Numerical	0.375	0.085	**0.94**	0.80	0.79	0.57	0.32	0.44	0.12	-0.02	0.26
		0.160	0.83	**0.92**	0.90	0.58	0.55	0.64	0.19	0.21	0.43
		0.265	0.70	0.75	**0.94**	0.39	0.35	0.75	0.05	0.06	0.62
	0.750	0.085	0.64	0.68	0.52	**0.92**	0.72	0.45	0.70	0.54	0.25
		0.160	0.29	0.61	0.52	0.67	**0.91**	0.72	0.68	0.85	0.57
		0.265	0.36	0.53	0.78	0.30	0.44	**0.90**	0.19	0.31	0.88
	1.000	0.085	0.19	0.35	0.23	0.74	0.75	0.44	**0.90**	0.81	0.36
		0.160	-0.05	0.26	0.18	0.49	0.80	0.54	0.73	**0.94**	0.50
		0.265	0.15	0.29	0.57	0.13	0.30	0.84	0.16	0.29	**0.93**

L. Sripragash and M. Sundaresan, AIP Conference Proceedings 1806, 100002 (2017)

Annals of Nuclear Energy 142 (2020) 107428

Contents lists available at ScienceDirect

Annals of Nuclear Energy

journal homepage: www.elsevier.com/locate/anucene

Results from a multi-physics numerical benchmark for codes dedicated to molten salt fast reactors

Marco Tiberga [a,*], Rodrigo Gonzalez Gonzaga de Oliveira [b], Eric Cervi [c], Juan Antonio Blanco [d], Stefano Lorenzi [c], Manuele Aufiero [d], Danny Lathouwers [a], Pablo Rubiolo [d]

[a] Delft University of Technology, Department of Radiation Science and Technology, Mekelweg 15, 2629 JB Delft, The Netherlands
[b] Paul Scherrer Institute, Nuclear Energy and Safety Research Division, Laboratory for Scientific Computing and Modelling – PSI, 5232 Villigen, Switzerland
[c] Politecnico di Milano, Department of Energy, Nuclear Engineering Division, Via La Masa 34, 20156 Milan, Italy
[d] LPSC, Université Grenoble-Alpes, CNRS/IN2P3, 53 rue des Martyrs, F-38026 Grenoble Cedex, France

ARTICLE INFO

Article history:
Received 20 December 2019
Received in revised form 10 February 2020
Accepted 24 February 2020
Available online 11 March 2020

Keywords:
Benchmark
Multi-physics
Molten salt reactor
Fast-spectrum
Thermal-hydraulics
Neutronics
Code-to-code comparison

ABSTRACT

Verification and validation of multi-physics codes dedicated to fast-spectrum molten salt reactors (MSR) is a very challenging task. Existing benchmarks are meant for single-physics codes, while experimental data for validation are absent. This is concerning, given the importance numerical simulations have in the development of fast MSR designs. Here, we propose the use of a coupled numerical benchmark specifically designed to assess the physics-coupling capabilities of the aforementioned codes. The benchmark focuses on the specific characteristics of fast MSRs and features a step-by-step approach, where physical phenomena are gradually coupled to easily identify sources of error. We collect and compare the results obtained during the benchmarking campaign of four multi-physics tools developed within the SAMOFAR project. Results show excellent agreement for all the steps of the benchmark. The benchmark generality and the broad spectrum of results provided constitute a useful tool for the testing and development of similar multi-physics codes.

Discrepancy

$$\epsilon_c = \sqrt{\frac{\sum_{i=1}^{N_p} \left(Q_c(\mathbf{r}_i) - Q_{avg}(\mathbf{r}_i)\right)^2}{\sum_{i=1}^{N_p} Q_{avg}^2(\mathbf{r}_i)}},$$

$$Q_{avg}(\mathbf{r}_i) = \frac{1}{N_c} \sum_{c=1}^{N_c} Q_c(\mathbf{r}_i)$$

Tiberga et al., Annals of Nuclear Energy, 2020

R-factor

$$R(d) = \int_{Q<Q_{\max}(d)} \left| \frac{\sqrt{N_{\text{real}}(\Omega)}}{\int_{Q'<Q_{\max}(d)} \sqrt{N_{\text{real}}(\Omega')}d\Omega'} - \frac{\sqrt{N_{\text{ideal}}(\Omega)}}{\int_{Q'<Q_{\max}(d)} \sqrt{N_{\text{ideal}}(\Omega')}d\Omega'} \right| d\Omega, \quad (21)$$

where

$$N_{\text{real}}(\Omega) = \frac{dN_S}{d\Omega} = \int_{-\infty}^{\infty} J(t)\frac{d\sigma_S}{d\Omega}(t)dt \quad (22)$$

is the number of photons (per unit solid angle) scattered from the sample undergoing x-ray-induced electronic damage, and

$$N_{\text{ideal}}(\Omega) = \left(\int_{-\infty}^{\infty} J(t)dt\right)\frac{d\sigma_S}{d\Omega}\Big|_{\text{neutral}} = F\frac{d\sigma_S}{d\Omega}\Big|_{\text{neutral}} \quad (23)$$

is the number of photons (per unit solid angle) scattered from the undamaged sample, which is given by the fluence (F) times the differential cross section of the undamaged sample. In our case, the undamaged sample is the neutral carbon atom in its ground configuration.

Son, S.-K., Young, L. & Santra, R., Physical Review A 83, (2011).

Mathematical discussion

Cosine similarity

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

Euclidean Distance

Introduction to Data Mining

1-norm distance $= \sum_{i=1}^{n} |x_i - y_i|$

2-norm distance $= \left(\sum_{i=1}^{n} |x_i - y_i|^2 \right)^{1/2}$ ⟶ Euclidean Distance

p-norm distance $= \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}$

infinity norm distance $= \lim_{p \to \infty} \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}$

$= \max \left(|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n| \right).$

Higher dimension?

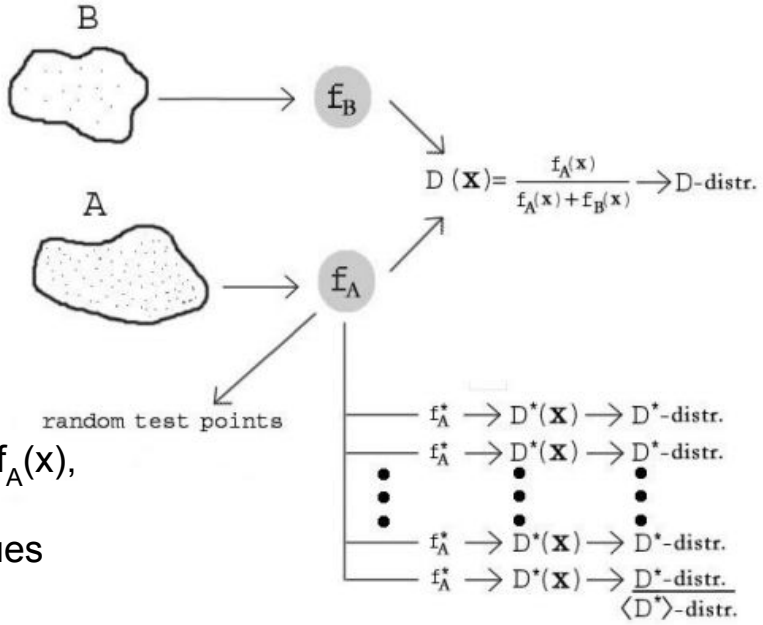
Comparing N-dimensional Distributions

PHYSTAT2003, SLAC, Stanford, California, September 8-11, 2003

A Multivariate Method for Comparing N-dimensional Distributions

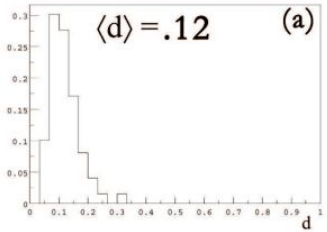
James D. Loudin and Hannu E. Miettinen
Rice University, Houston, TX 77005, U.S.A.

We propose a new multivariate method for comparing two N-dimensional distributions. We first use kernel estimation to construct probability densities for the two data sets, and then define two discriminant functions, one appropriate for the null hypothesis and another appropriate for the actual data. Distributions of the two discriminant functions at random test points are then compared using the one-dimensional K-S test. The performance of the method is illustrated with Monte Carlo data.

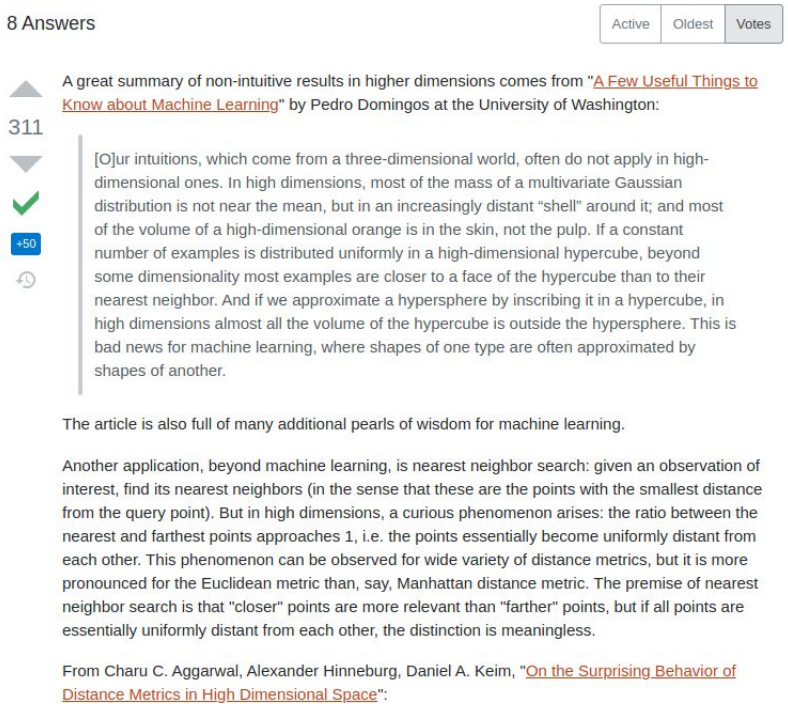


1. Construct probability densities using kernel estimation $f_A(x)$, $f_B(x)$
2. Calculate the discriminant function D and its mean values

In the 10% case the mean value is 0.12, indicating that there is a ~30% probability that the average d would exceed this value for two identical densities



Why is Euclidean distance not a good metric in high dimensions?



8 Answers

Active | Oldest | Votes

311

+50

A great summary of non-intuitive results in higher dimensions comes from "A Few Useful Things to Know about Machine Learning" by Pedro Domingos at the University of Washington:

> [O]ur intuitions, which come from a three-dimensional world, often do not apply in high-dimensional ones. In high dimensions, most of the mass of a multivariate Gaussian distribution is not near the mean, but in an increasingly distant "shell" around it; and most of the volume of a high-dimensional orange is in the skin, not the pulp. If a constant number of examples is distributed uniformly in a high-dimensional hypercube, beyond some dimensionality most examples are closer to a face of the hypercube than to their nearest neighbor. And if we approximate a hypersphere by inscribing it in a hypercube, in high dimensions almost all the volume of the hypercube is outside the hypersphere. This is bad news for machine learning, where shapes of one type are often approximated by shapes of another.

The article is also full of many additional pearls of wisdom for machine learning.

Another application, beyond machine learning, is nearest neighbor search: given an observation of interest, find its nearest neighbors (in the sense that these are the points with the smallest distance from the query point). But in high dimensions, a curious phenomenon arises: the ratio between the nearest and farthest points approaches 1, i.e. the points essentially become uniformly distant from each other. This phenomenon can be observed for wide variety of distance metrics, but it is more pronounced for the Euclidean metric than, say, Manhattan distance metric. The premise of nearest neighbor search is that "closer" points are more relevant than "farther" points, but if all points are essentially uniformly distant from each other, the distinction is meaningless.

From Charu C. Aggarwal, Alexander Hinneburg, Daniel A. Keim, "On the Surprising Behavior of Distance Metrics in High Dimensional Space":

https://stats.stackexchange.com/questions/99171/why-is-euclidean-distance-not-a-good-metric-in-high-dimensions

Discussion

1. Should we narrow the focus to cases with same binning (simulation can always be set to real detector binning), or consider how to compare data with different binnings?
2. Is there a way to compare event data that was never in a histogram to begin with?
3. Assuming we find a systematic way to do this comparison, how do we implement it? As a little library we all use in ViNYL?

Why do we want to compare simulations with experiments

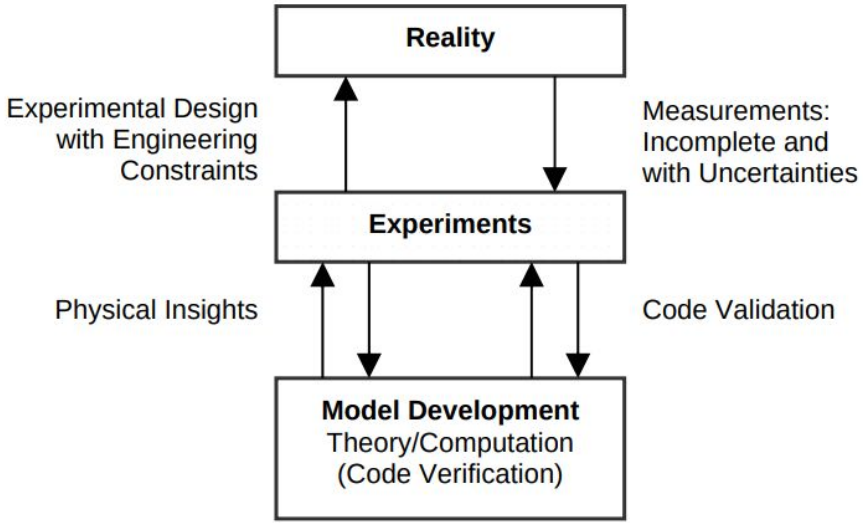


Figure 1. The interrelationship between simulations, experiments and physical reality are illustrated along with the processes that connect them.

Verification assesses the degree to which a code correctly implements the chosen physical model

Validation assesses the degree to which a code describes the real world.

Greenwald, M., Computer Physics Communications 164, 1–8 (2004)

Why do we want to compare simulations with experiments

Differences between measurements and code results can arise through several sources.

- Model formulation errors – missing or incorrect physics
- Numerical solution errors due to discretization, boundary conditions or implementation
- Measurement errors and scarcity

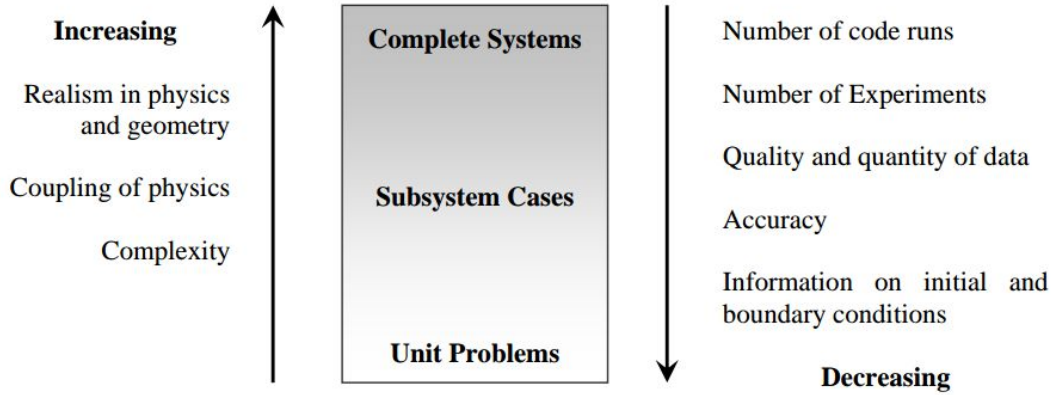


Figure 2. The hierarchy of validation begins with experiments designed to isolate particular physical phenomena and proceeds to more complex and realistic cases. As the hierarchy is traversed, the quantity and quality of data decreases, making comparisons less definitive.

Greenwald, M., Computer Physics Communications 164, 1–8 (2004)

Why do we want to compare simulations with experiments

1. Verify codes first.
2. Plan a hierarchy of experiments beginning with the simplest physics and geometry.
3. Design experiments jointly by experimentalists and computationalists.
4. Experiments should test crucial features of the model, especially its assumptions or important simplifications. Perturbing effects should be minimized. Geometry, boundary and initial conditions must be well characterized and documented. Critical measurements should be defined and limitations, uncertainties, and sources of error discussed with openness and candor.
5. Document code predictions in advance.
6. While jointly designed, carry out experiments and code runs independently.
7. Make as complete measurements as possible when carrying out experiments. Multiple diagnostics to measure the same quantities are desirable. Statistically sufficient data sets should be collected, repeating runs as required. It can be valuable to conduct experiments at more than one facility if this is practical.
8. Pay special attention to analysis of errors and uncertainties. Use modern statistical techniques to design experiments and to identify random and bias errors.
9. When analyzing results, don’t paper over differences. The goal is not to prove that a code is correct, but to assess its reliability and point the way towards improvement.
10. Document process and results including data reduction techniques and error analysis.

Why do we want to compare simulations with experiments

In many cases, it is preferable to make comparisons through “synthetic” diagnostics – that is by post-processing simulation data in a manner which is as analogous as possible to the physical diagnostic.

Databases dedicated for comparisons between simulations and experiments:
- Dynamic and interactive – able to be updated, annotated, appended
- Include metadata (data about the data) for every data item. This would document, for example, where the data came from, when it was written, who was responsible for it as well as basic information on the data type, size, structure and so forth creating a complete coherent self-descriptive structure
- Include both experimental and modeling data
- Contain all auxiliary data, assumptions, geometry, boundary and initial conditions
- Contain estimations of error
- Regimes well defined
- Able to be queried - searchable by content or by address
- Able to be browsed
- Linked to publications

Greenwald, M., Computer Physics Communications 164, 1–8 (2004)

research papers

Journal of
Synchrotron
Radiation

ISSN 0909-0495

Received 26 March 2010
Accepted 9 December 2010

Coherent diffraction microscopy at SPring-8: instrumentation, data acquisition and data analysis

Rui Xu,[a,b]* Sara Salha,[a] Kevin S. Raines,[a] Huaidong Jiang,[c] Chien-Chun Chen,[a] Yukio Takahashi,[d] Yoshiki Kohmura,[b] Yoshinori Nishino,[e] Changyong Song,[b] Tetsuya Ishikawa[b] and Jianwei Miao[a]

[a]Department of Physics and Astronomy and California NanoSystems Institute, University of California, Los Angeles, CA 90095, USA, [b]RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan, [c]State Key Laboratory of Crystal Materials, Shandong University, Jinan 250100, People's Republic of China, [d]Frontier Research Base for Global Young Researchers, Frontier Research Center, Graduate School of Engineering, Osaka University, 2-1 Yamada-oka, Suita, Osaka 565-0871, Japan, and [e]Research Institute for Electronic Science, Hokkaido University, Kita 21 Nishi 10, Kita-ku, Sapporo 001-0021, Japan. E-mail: xur@physics.ucla.edu

Since the first demonstration of coherent diffraction microscopy in 1999, this lensless imaging technique has been experimentally refined by continued developments. Here, instrumentation and experimental procedures for measuring oversampled diffraction patterns from non-crystalline specimens using an undulator beamline (BL29XUL) at SPring-8 are presented. In addition, detailed post-experimental data analysis is provided that yields high-quality image reconstructions. As the acquisition of high-quality diffraction patterns is at least as important as the phase-retrieval procedure to guarantee successful image reconstructions, this work will be of interest for those who want to apply this imaging technique to materials science and biological samples.

$$R_{\mathrm{f}} = \sum ||F_{\mathrm{cal}}| - |F_{\mathrm{exp}}|| / |F_{\mathrm{exp}}|,$$

Oversampling requirements: A function of the X-ray wavelength, the sample size and the detector pixel size (Miao et al. 2003b)

Quantitative density map (Song et al., 2008)