NCSA Industry Overview with Computational Breakthroughs and Synergies with Artificial Intelligence

Brendan McGinty
Program Director

Seid Korić
Technical Director

Shirui Luo
Research Scientist

National Center for Supercomputing Applications
UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

With NCSA: Six Months Ahead of Competition

Industry Dedicated
- Technical Teams
- HPC Resources
- Business Leadership and Project Management

Tradition
- Industry part of NCSA's mission for more than 30 years
- Most decorated: many HPC awards and world records

Culture
- Work under industrial pace and NDAs
- Deliver on time and under budget
- Collaborative environment

Largest and Oldest Industrial HPC Program in the World

Industry Partners – 1 of 3

CAT
syngenta
abbvie
MAYO CLINIC
DANIELI
Rolls-Royce
SIEMENS
leidos
ExxonMobil
synchrony
Capital One
JOHN DEERE
PHILLIPS 66
BRUNSWICK
P&G
HITACHI
SEAL OF THE STATE OF ILLINOIS AUG. 26TH 1818
CU Aerospace
CASI
HATCH AG GROUP
AXIS
ILLINOIS ROCSTAR
SIMBIOSYS

Industry Partners – 2 of 3

Industry Partners – 3 of 3

IBM Watson

LSTC
Livermore Software Technology Corp.

BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

M D
The Digital Manufacturing Institute

State Farm®

Sentieon

affoa

GENALICE
TECHNOLOGY FOR PEOPLE & SCIENCE

The Cyprus Institute
RESEARCH•TECHNOLOGY•INNOVATION

ANSYS®

CD-adapco
www.cd-adapco.com

Continuous Casting Consortium

Legacy Partners

Allstate
Microsoft
American Airlines
Shell
BOEING
Schlumberger
ADM
MCDONNELL DOUGLAS
Kellogg's
ACNielsen
MOTOROLA
Ford
bp
J.P.Morgan
Lilly
AT&T
FMC
KODAK
State Farm
DOW
nag
sears

History

1986 – Program founded with first industry partner, Eastman Kodak

1992 – First Grand Challenge Award: Eli Lilly

1993 – Caterpillar joins, wins Grand Challenge Award

2004 – Boeing recognized with Grand Challenge Award

2011 – iForge industrial cluster becomes available

2014 and 2017 – Winner of HPCwireTop Supercomputing Achievement

2017 – ExxonMobil sets sector world record

- Oil reservoir model: 3 months to 10 minutes, 719000 cores, $1B+ ROI

2020 – Majority of Industrial engagement becomes AI-oriented

Engagement Model: Current Partners

Discover

Initial meetings
Identify needs
Define scope
Set timelines
Define budget
Create work plan

Build

Design solutions
Develop
Test
Loop as necessary

Deliver

Implement
Interview stakeholders
Evaluate effectiveness
Calculate ROI

Engagement Model: Prospective Partners

- Identify **challenges for companies** that match team skills

abbvie → gsk GlaxoSmithKline

- Be **consultative**: listen to needs and challenges
- **Match needs with specific skills** within team or with strategic partners
- Define **value** proposition: what company gets from engagement

NCSA Industry Technical Team Expertise

- Modeling and Simulation
- Bioinformatics and Genomics
- “Big” Data Analytics, GIS, and AI
- Code Profiling and Optimization
- Rapid User Support and Domain/HPC Training
- Cyberinfrastructure and Security
- Visualization
- Much more at NCSA and the University of Illinois

National Petascale Computing Facility

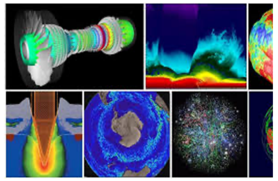
World-Class Data Center

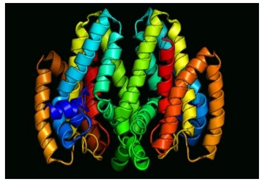
- Dept. of Energy-like security
- 88000 sqft
- 25 MW of power; LEED Gold
- 400+ Gb/sec bandwidth

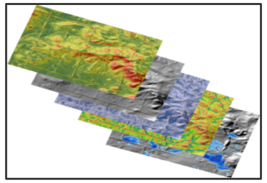
Hosting Benefits to Industry

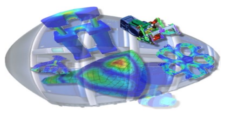
- Low-cost power & cooling
- 24/7/365 Help Desk
- Adjacent to and aligned with UIUC Research Park

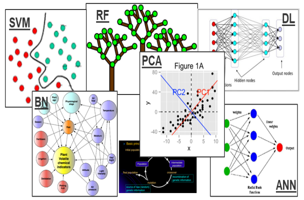
*Forge – The HPC Environment for Industry











- **Latest and best**
  - Computing (Intel/Skylake 192-256 GB)
  - GPU driven AI technologies (V100)
- **99% uptime and live upgrades**
- **Development and production workhorse**
- **Rapid user support and advanced consulting**
- **Built exclusively for Industry’s applications and workflows**

Engineering Application Breakthroughs on Blue Waters 2013-2020

64,000+ cores
LS-DYNA (Cray, RRC, P&G, NCSA)

HTC, 600TB
H3Africa (IGB, HPCBio, U of C. Town, NCSA)

100,000+ cores
Alya Multiphysics
~**90%** PE @ 100K
!(BSC & NCSA)

716,800+ cores
Oil & Gas Reservoir Modeling (Exxon & NCSA)

114,000+ cores
Ansys-Fluent
(Cray, Dell, NCSA)

512 XK7 nodes
ACCEL_WSMP
(NVIDIA, IBM-Watson, NCSA)

65,000+ cores
WSMP (IBM-Watson, NCSA, BSC, RRC, Repsol)

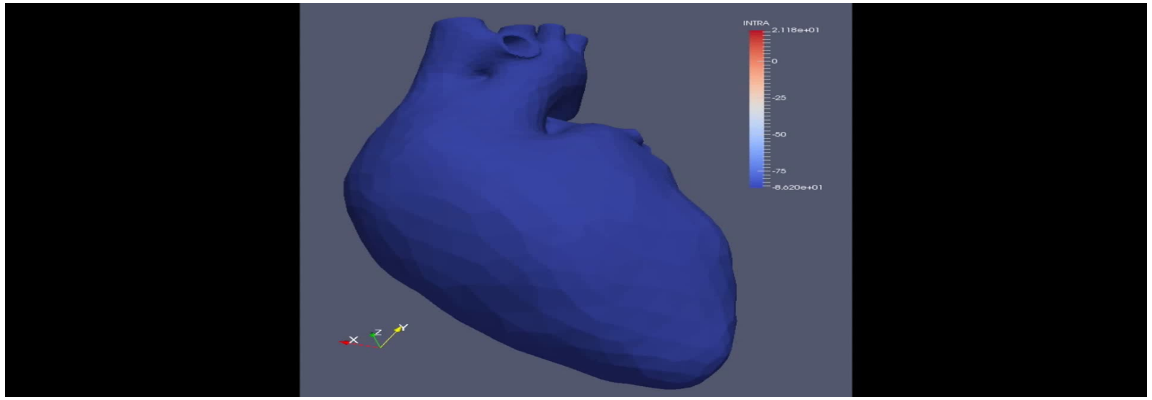
Two Real-World Cases Solved with Alya Multiphysics Code from BSC on NCSA's Blue Waters

Human Heart

Non-linear solid mechanics

Coupled with electrical propagation

3.4 billion elements, scaled to 100,000 cores



Kiln Furnace

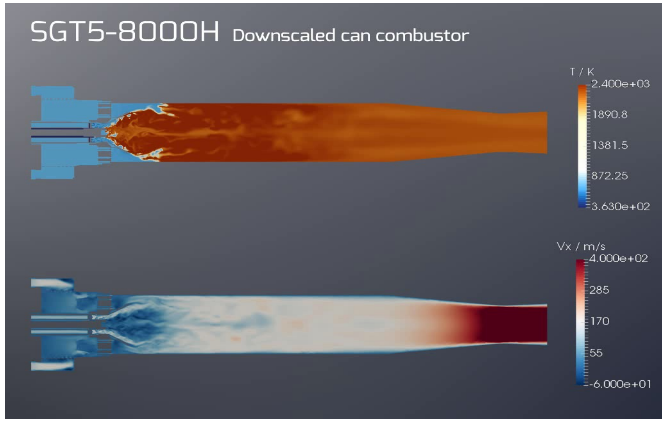
Transient incompressible turbulent flow

Coupled with energy and combustion

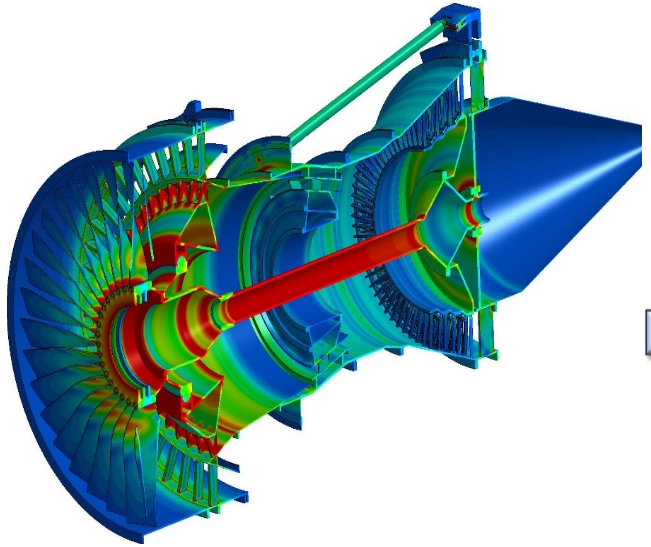
4.22 billion elements

Scaled to 100,000 cores @90% parallel efficiency

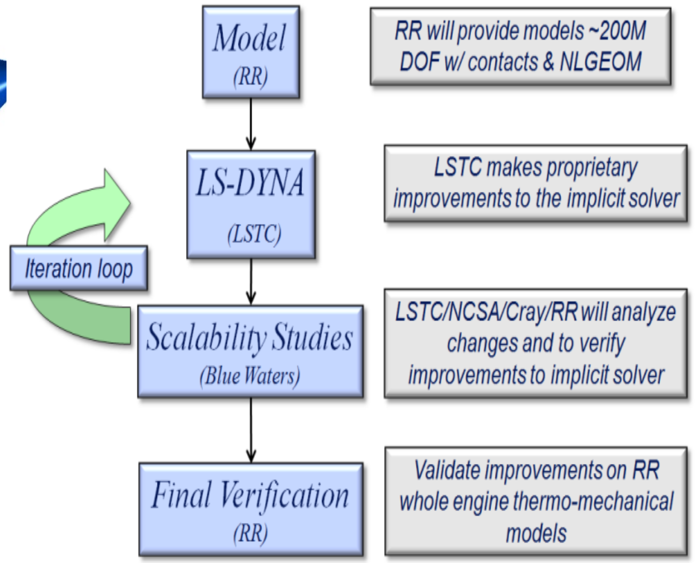
17.4 years on a serial PC down to **1.8 hours** on BW

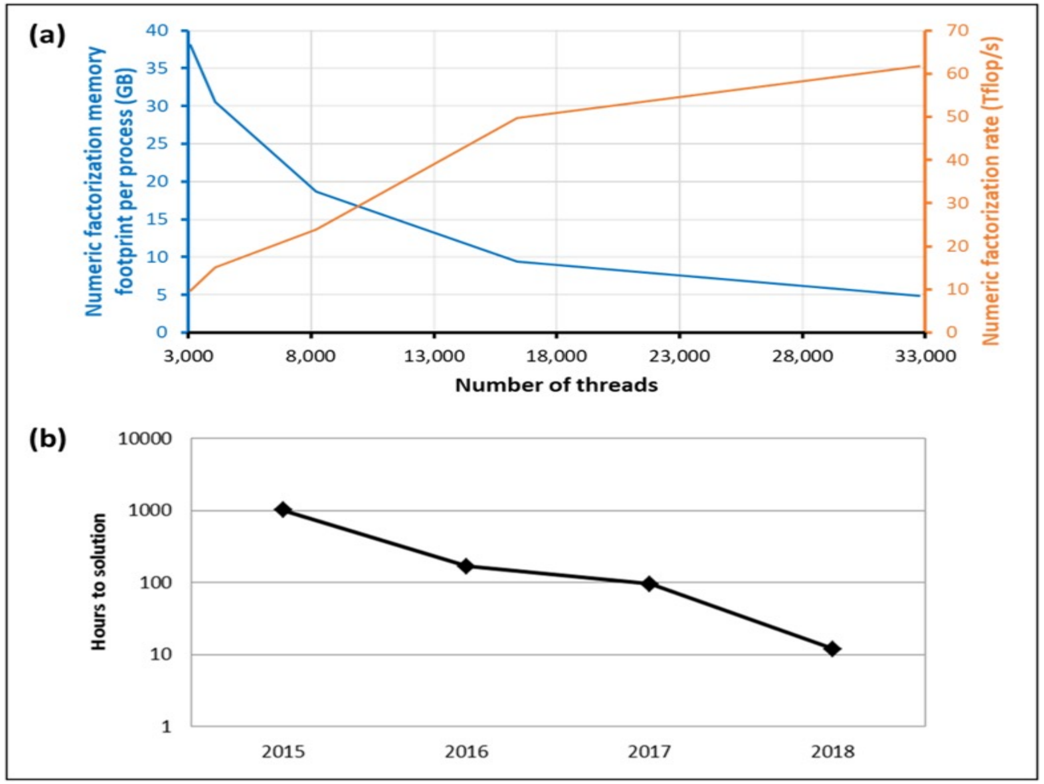


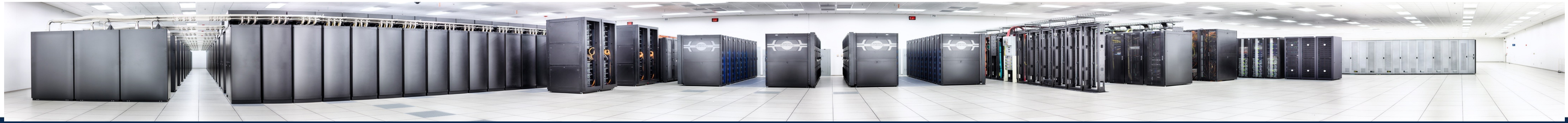
Reducing the Time-to-Solution for High Fidelity Finite Element Analysis of Gas Turbine Engines - from Months to Hours, 2015-2018



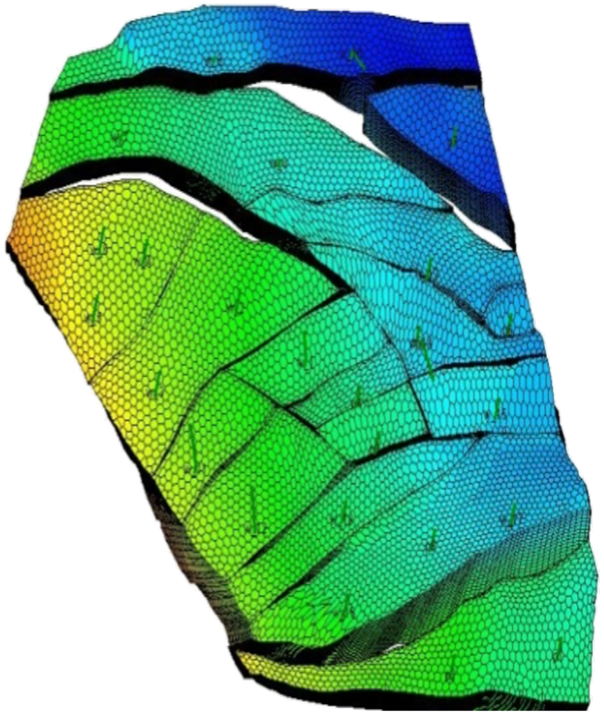
Rolls-Royce engine model for thermo-mechanical analysis, >200M DOFs







Massively Parallel Modeling in Oil & Gas & ROI

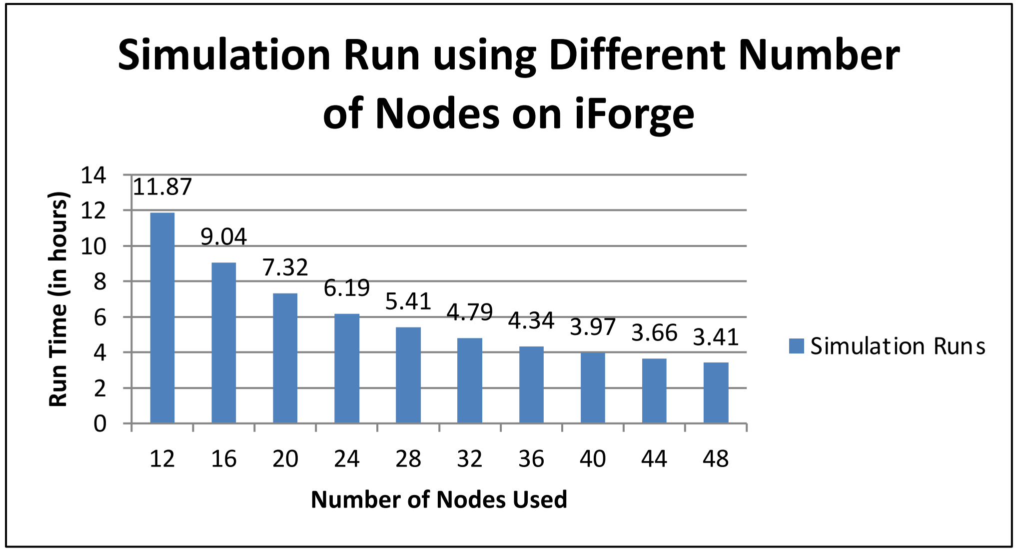


ExxonMobil

- Reservoir simulation models the complex subsurface flows of fluids in oil and natural gas reservoirs
- Previous runtime: 3.5 months on prem
- Optimized: 10 minutes on Blue Waters
- 716800 MPI processes, was the entire engineering sector world record for degree of parallelism

- **Minimized costs and environmental impact**
- **ROI: USD$1+B**

Large Scale Statistical HPC Analysis in Agriculture

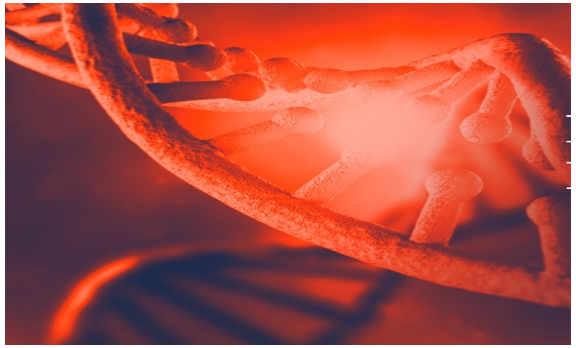


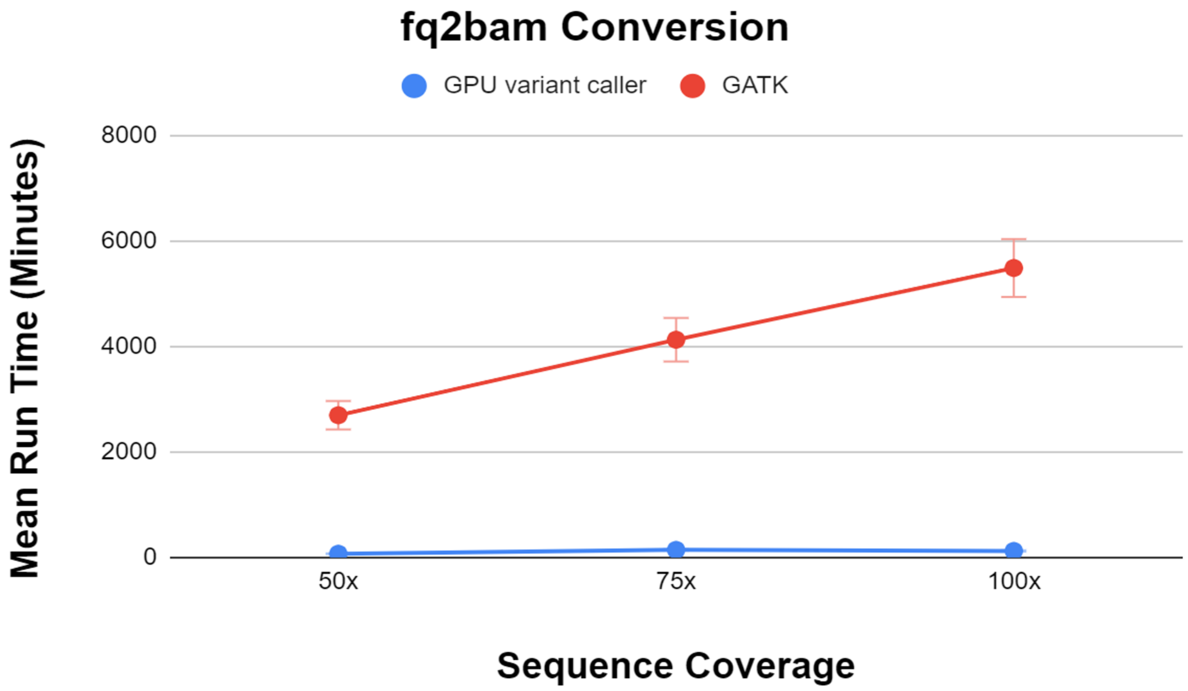
- Power statistical analysis uses massive data collected from farm field trials to allow an agriculture partner of NCSA to assess quality of their experimental designs
- NCSA has developed an efficient and scalable implementation in **R** to perform massive simulation using multi-node parallelization and variable instantiation techniques
- Our new implementation decreases the size of the program from over 50,000 lines to less than 100 lines, reduces the processing time for a simulation with over 70,000 cases from **175 days** (@partner) **to less than 3.5 hours)** (@HPC/iForge)

Courtesy of Dr. Dora Cai and an Industrial Partner of NCSA

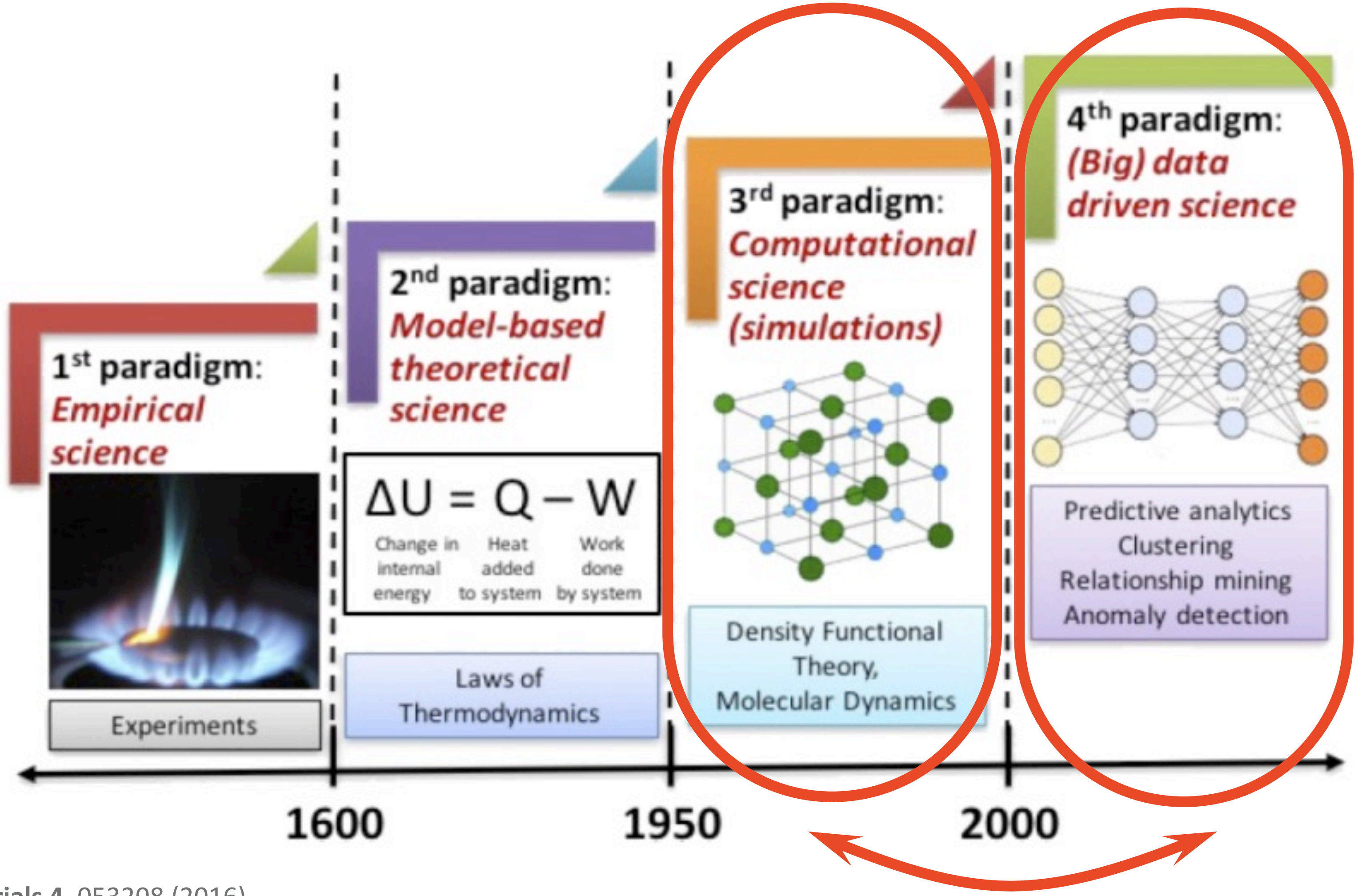
Benchmarking of new variant calling tools on GPUs

- Benchmarked a new genomic variant calling software which **runs on GPU only**
- Tested **multiple tools within the suite,** determined the speed up of this software with respect to the industry standard GATK
- Evaluated the **biological accuracy** by comparing results to GATK, the gold standard of variant calling.
- Tested the **scalability** of this software with different sizes of genomic data to determine its robustness.
- Worked with our **industry partners** to test against their variant calling tools.





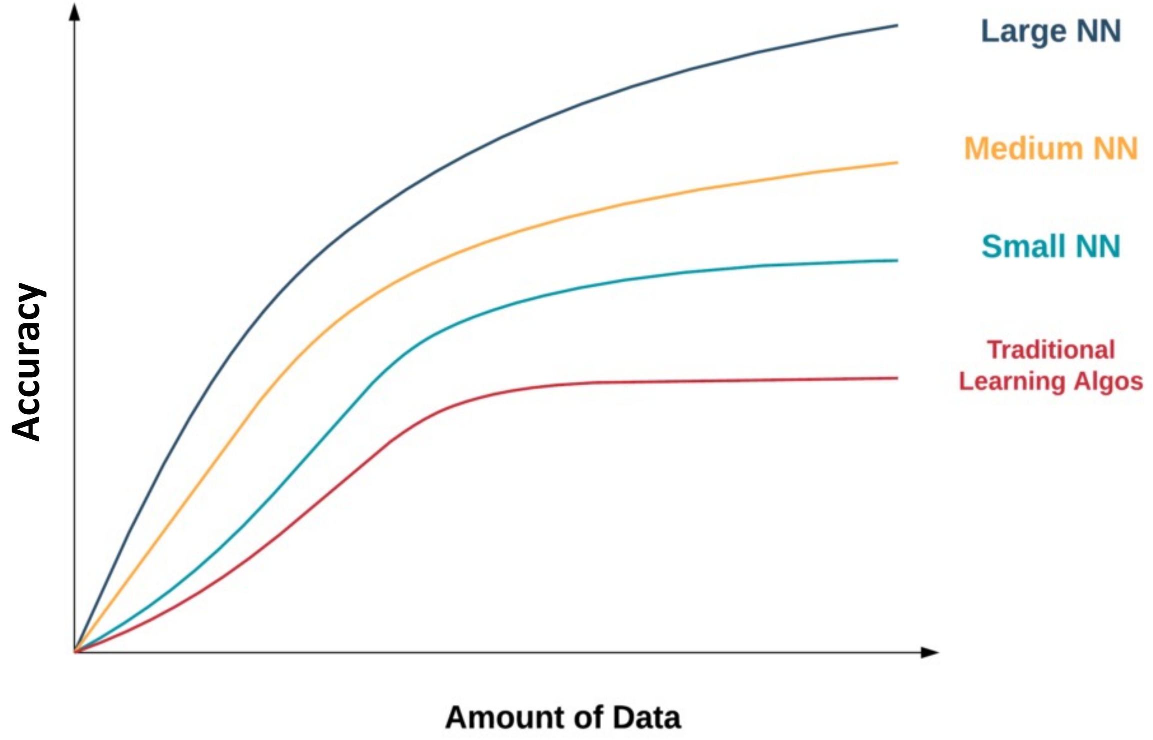
Four Paradigms in Science and Engineering

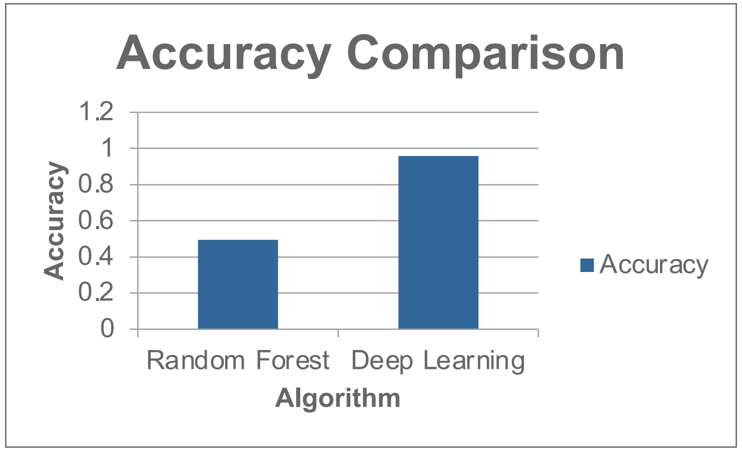


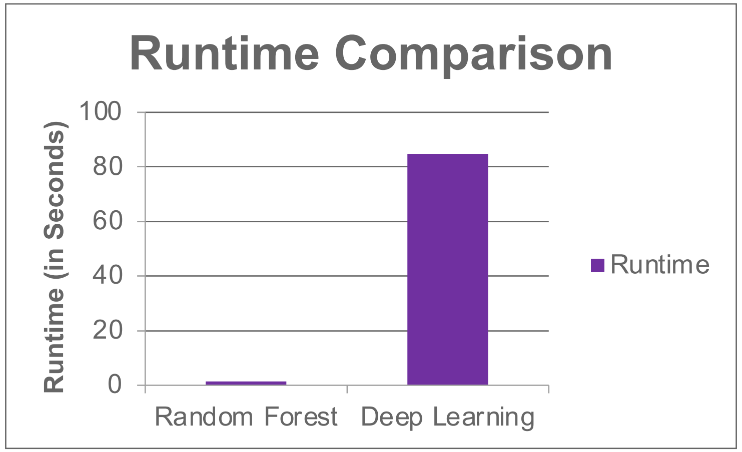
"AI is the new electricity"
Prof. Andrew Ng, Stanford, Coursera founder

APL Materials 4, 053208 (2016)

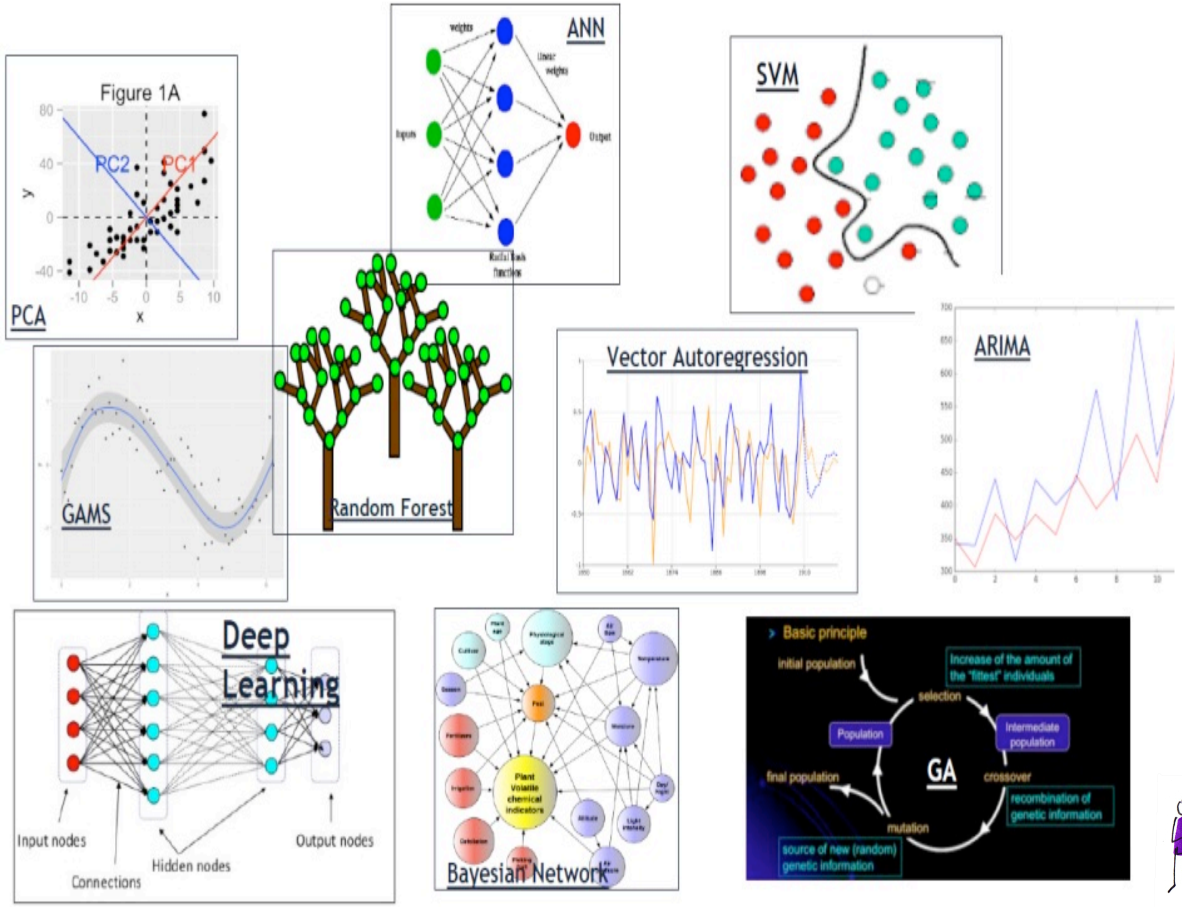
Big Data and HPC Driven Deep Learning





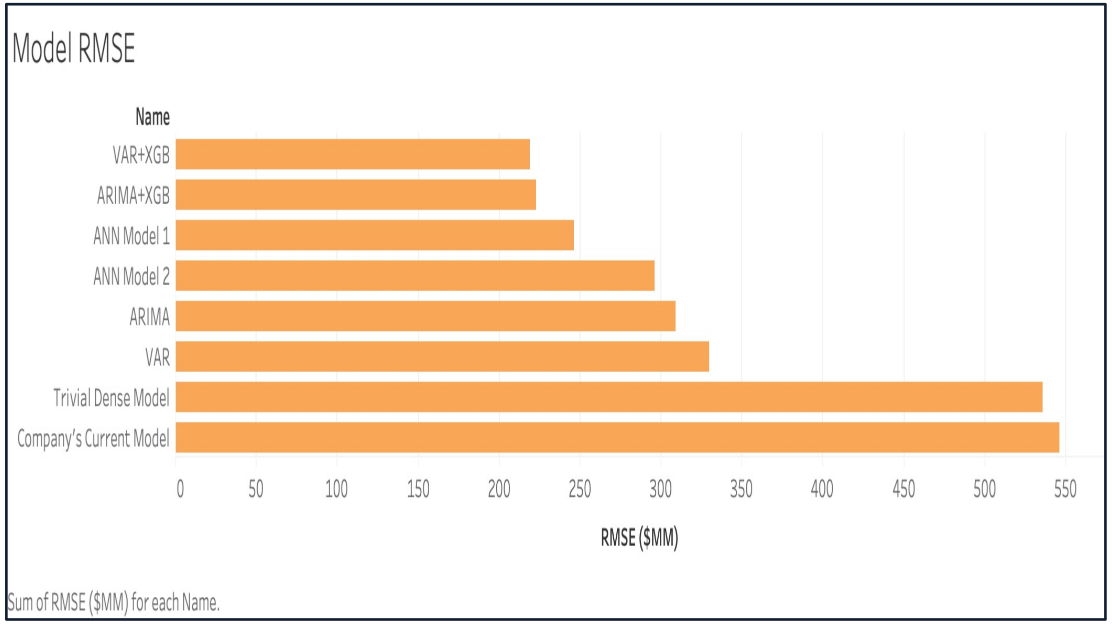


Choosing and Applying Best Machine Learning Algorithm

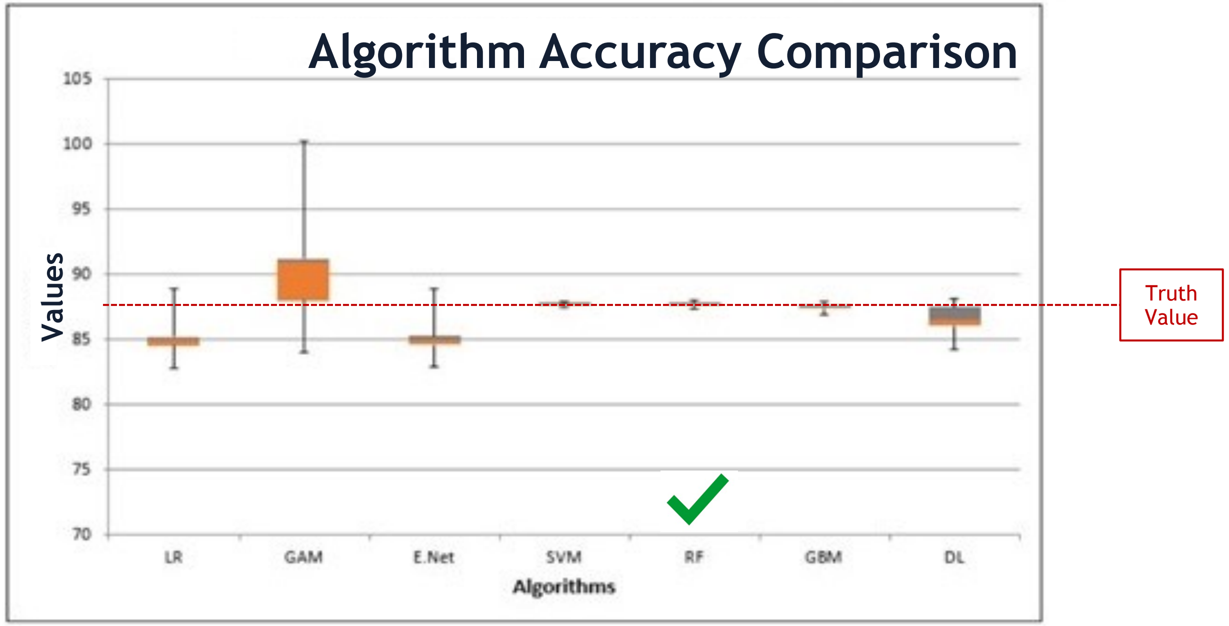


Choosing Best Machine Learning Algorithm

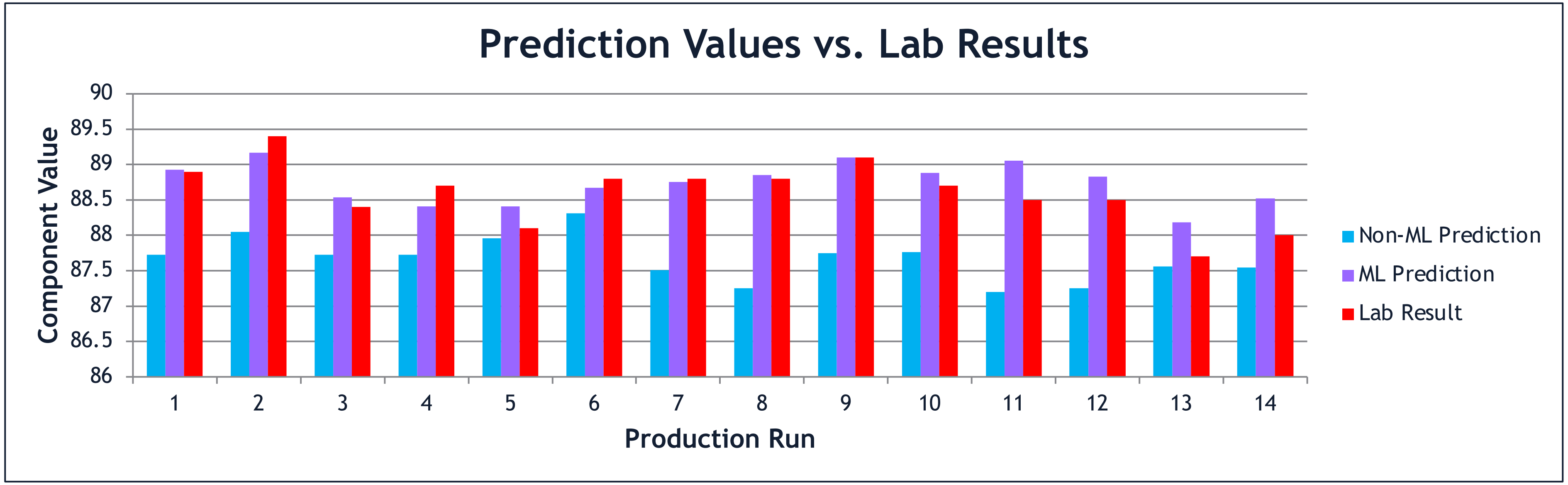
- Based on Root Mean Square Errors (RMSE)



- Based on Median Values and Standard Deviation



Reduce Production Cost using Machine Learning



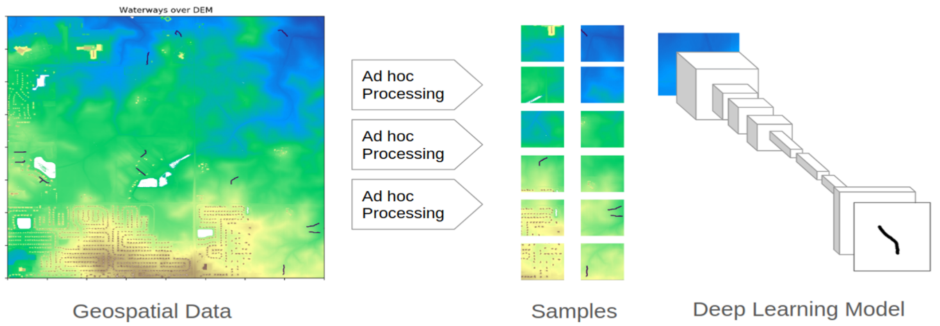
- Optimize ingredient recipes using Machine Learning predictive models
- Make the predicted values closer to the real lab test results (ground truth)
- Reduce *Mean Absolute Errors* (**MAE**) from 0.73 to 0.43
- ROI: USD**$18 million annually** by reducing the production cost

Connecting Industrial Geospatial and AI Communities

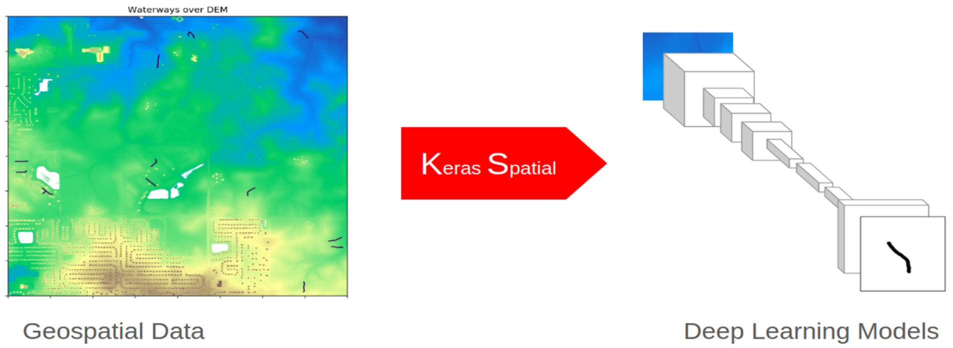
Novel Spatial Data Generators to connect TensorFlow models with geospatial data :

- Handles geospatial data in consumable formats by AI models without worrying about their specs such as projection, resolution, etc.
- Harmonizes multiple data sources and feeds them directly to the same AI model during the training phase.
- Scales processing of archives of geospatial data during the prediction phase.

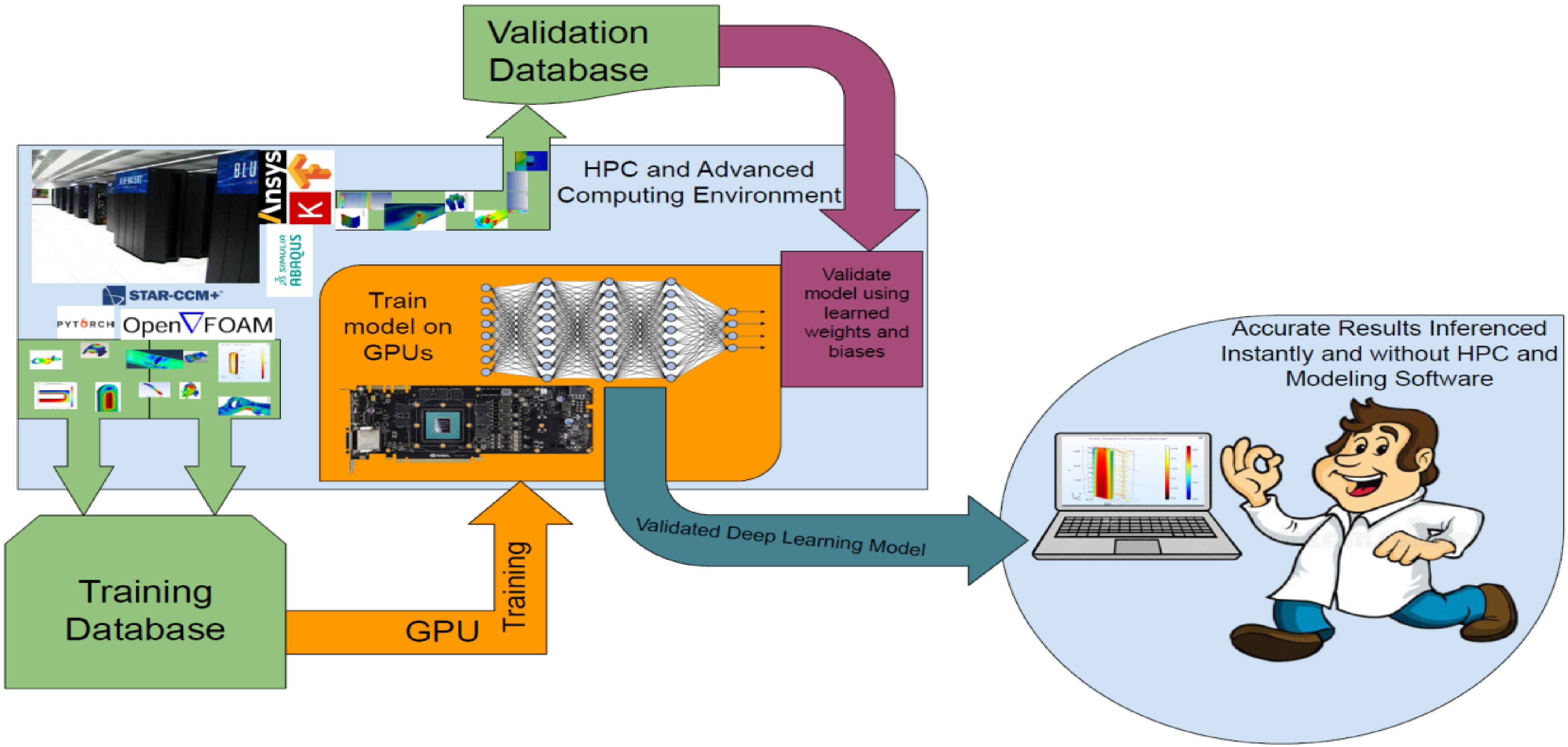
Separate Worlds



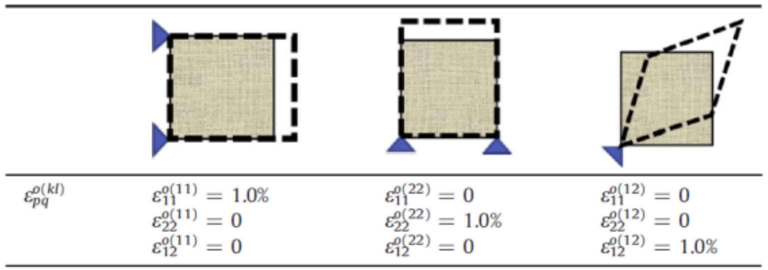
Connecting Two Worlds



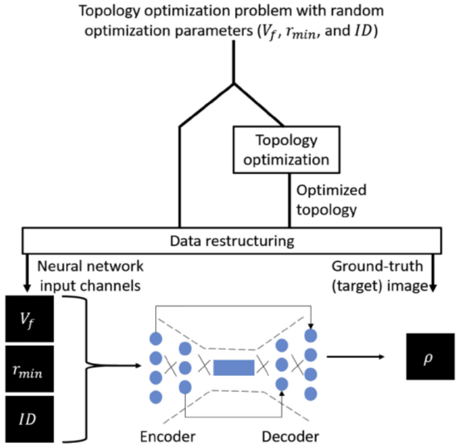
Surrogate Data-Driven Deep Learning Model

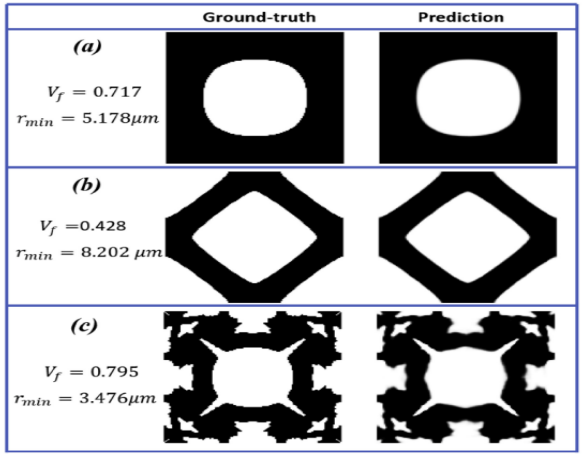


Deep Learning for Topological Optimization of Metamaterials

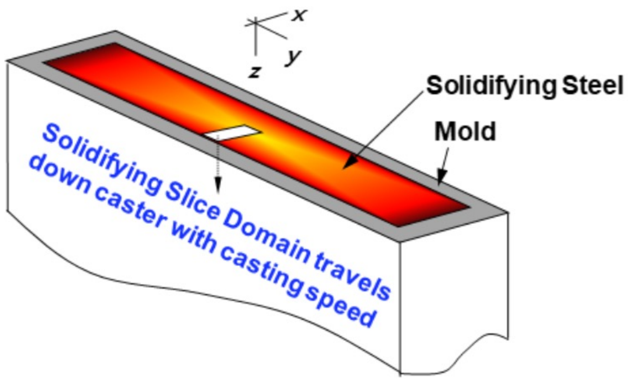


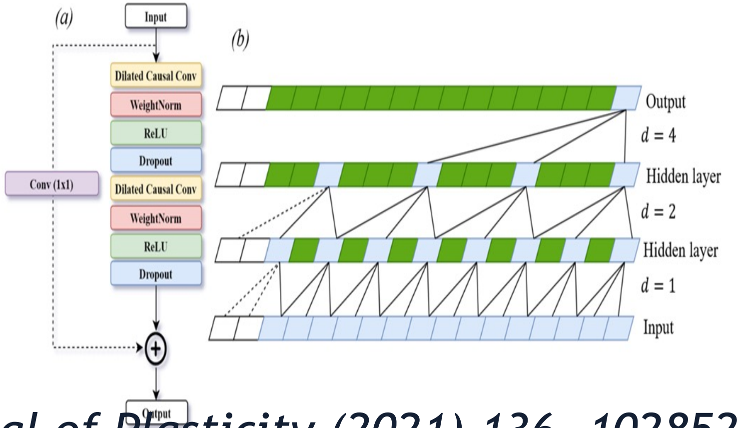
Kollmann et al., Materials & Design (2020), 109098

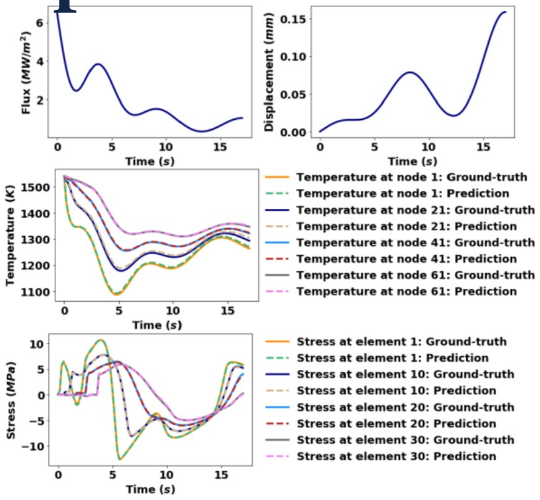




Deep Learning for Multiphysics Modeling of Visco-plastic Materials



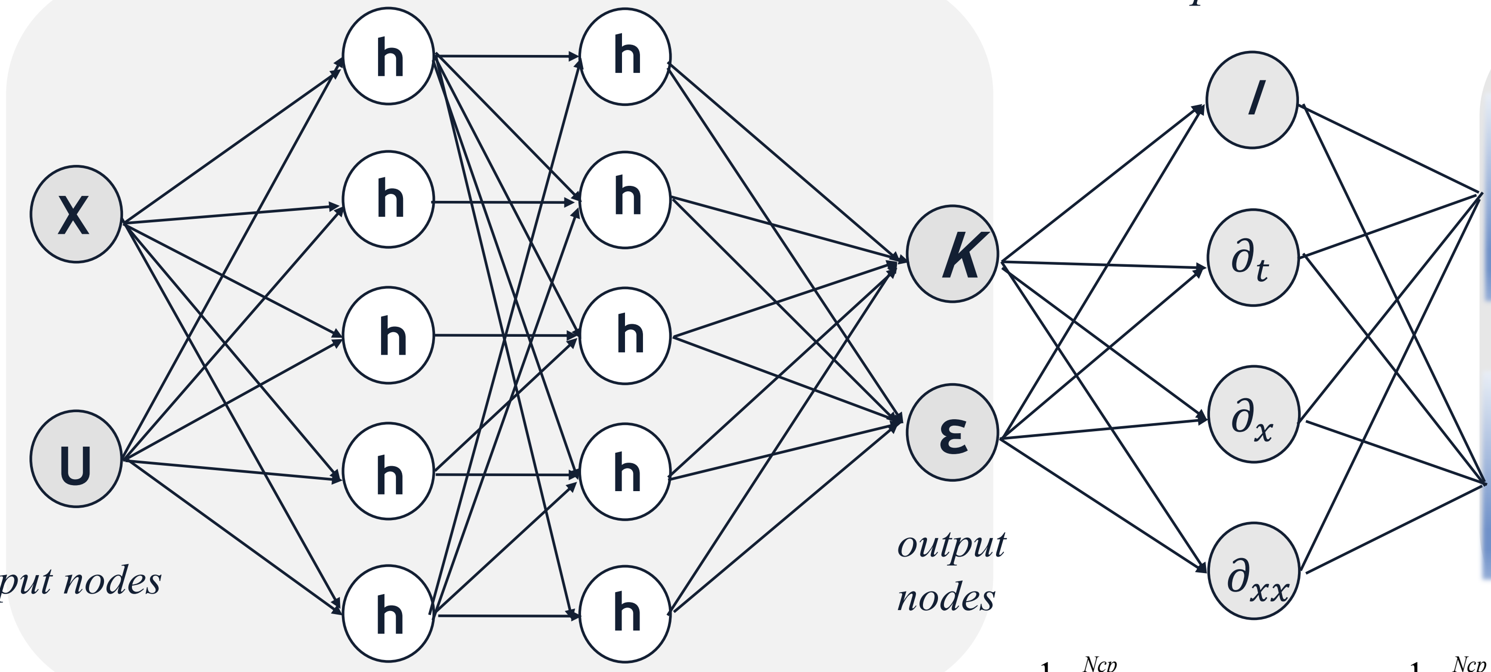




Abueidda et al., International Journal of Plasticity (2021) 136, 102852

Physics Informed Neural Network (PINN) Tuning K-ε Turbulence Model

Feedforward neural network | *Fluid physics constraints*



$$f:\ \frac{\partial K}{\partial t}+\bar{u}_i\frac{\partial K}{\partial x_i}+\tau_{ij}\frac{\partial \bar{u}_i}{\partial x_j}+\varepsilon-\frac{\partial}{\partial x_i}\left(\frac{\nu_T}{\sigma_K}\frac{\partial K}{\partial x_i}\right)-\nu_T\frac{\partial^2 K}{\partial x_i\partial x_i}$$

$$\nu_T=C_\mu\frac{K^2}{\varepsilon}$$

$$g:\ \frac{\partial \varepsilon}{\partial t}+\bar{u}_i\frac{\partial \varepsilon}{\partial x_i}+C_{\varepsilon 1}\frac{\varepsilon}{K}\tau_{ij}\frac{\partial \bar{u}_i}{\partial x_j}-\frac{\partial}{\partial x_i}\left(\frac{\nu_T}{\sigma_\varepsilon}\frac{\partial \varepsilon}{\partial x_i}\right)+C_{\varepsilon 2}\frac{\varepsilon^2}{K}-\nu_T\frac{\partial^2 \varepsilon}{\partial x_i\partial x_i}$$

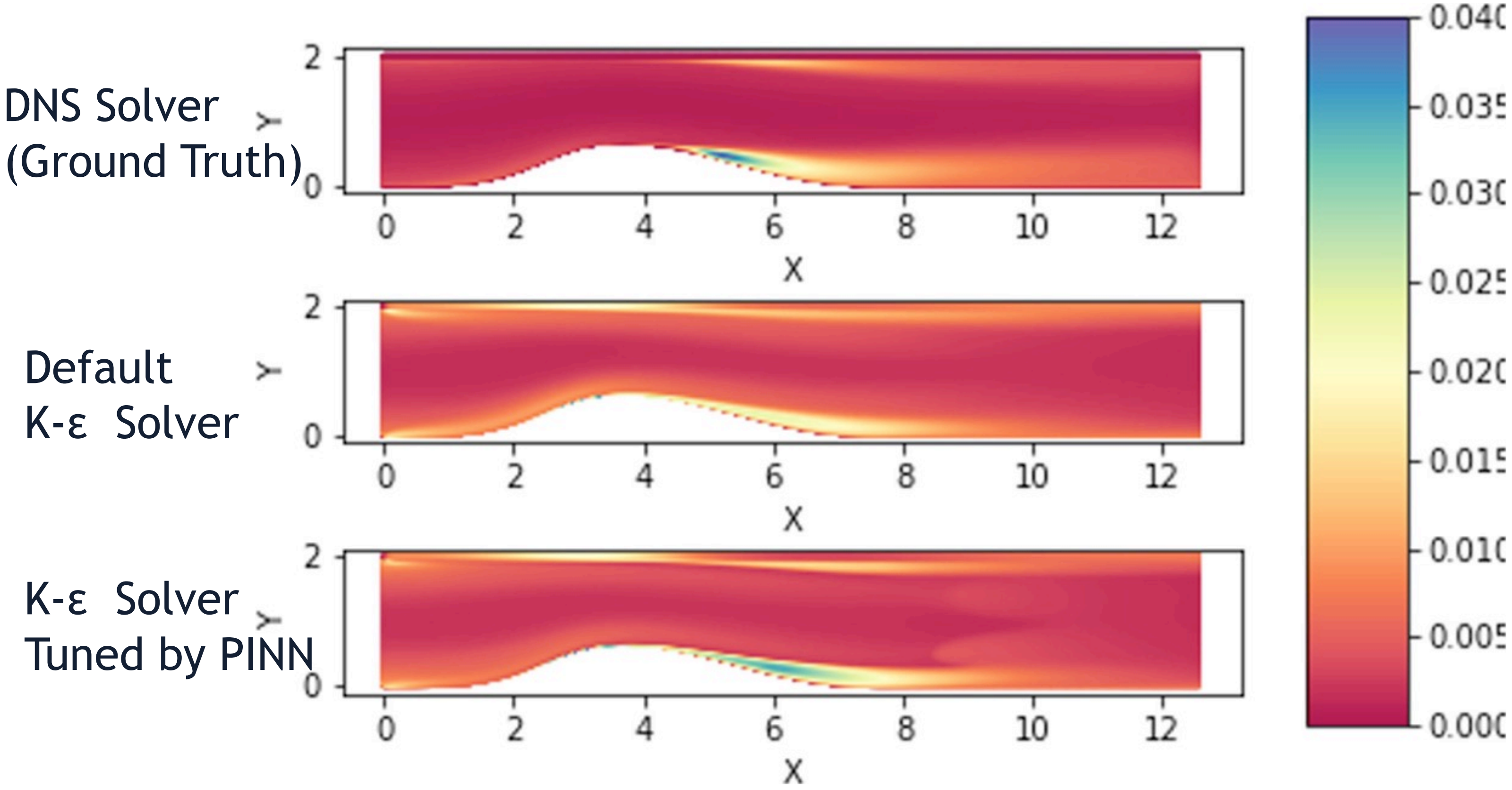
$$Loss=\frac{1}{N_{cp}}\sum_{i=1}^{Ncp}(K_i^{DNS}-K_i^{pred})^2+\frac{1}{N_{cp}}\sum_{i=1}^{Ncp}(\varepsilon_i^{DNS}-\varepsilon_i^{pred})^2+\omega_f*\frac{1}{N_{cp}}\sum_{i=1}^{Ncp}(f_i^{pred})^2+\omega_g*\frac{1}{N_{cp}}\sum_{i=1}^{Ncp}(g_i^{pred})^2$$

Luo et al., International Supercomputing Conference (ISC) 2020

Five Parameters $C_{\varepsilon 1}$, $C_{\varepsilon 2}$, C_μ, σ_K, σ_ε tuned by TF as 5 extra Hyperparameters to additionally minimize Loss

Comparison of the time-averaged Turbulent Kinetic Energy

Five constant	Empirical (Default)	NN-pred Fix C_μ
$C_{\varepsilon 1}$	1.44	1.302
$C_{\varepsilon 2}$	1.92	1.862
C_μ	0.09	**0.09**
σ_κ	1.0	0.751
σ_ε	1.3	0.273

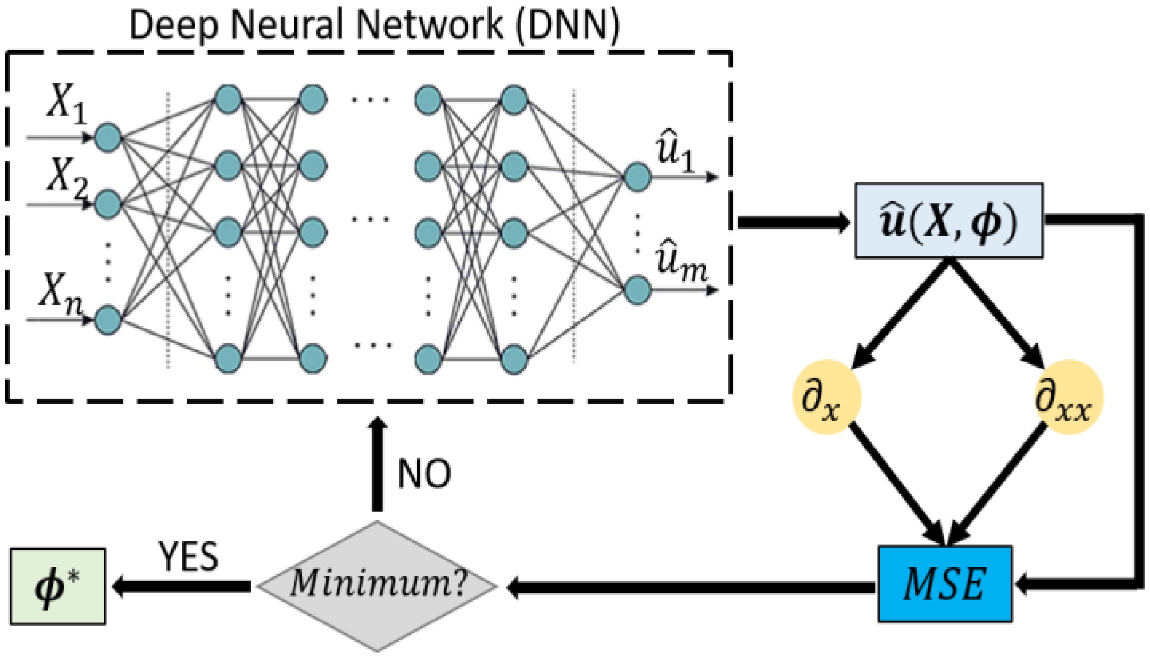


Luo et al., International Supercomputing Conference (ISC) 2020

DNS Simulation ~ Weeks and Months
K-ε Simulation ~ Minutes and Hours

ILLINOIS NCSA

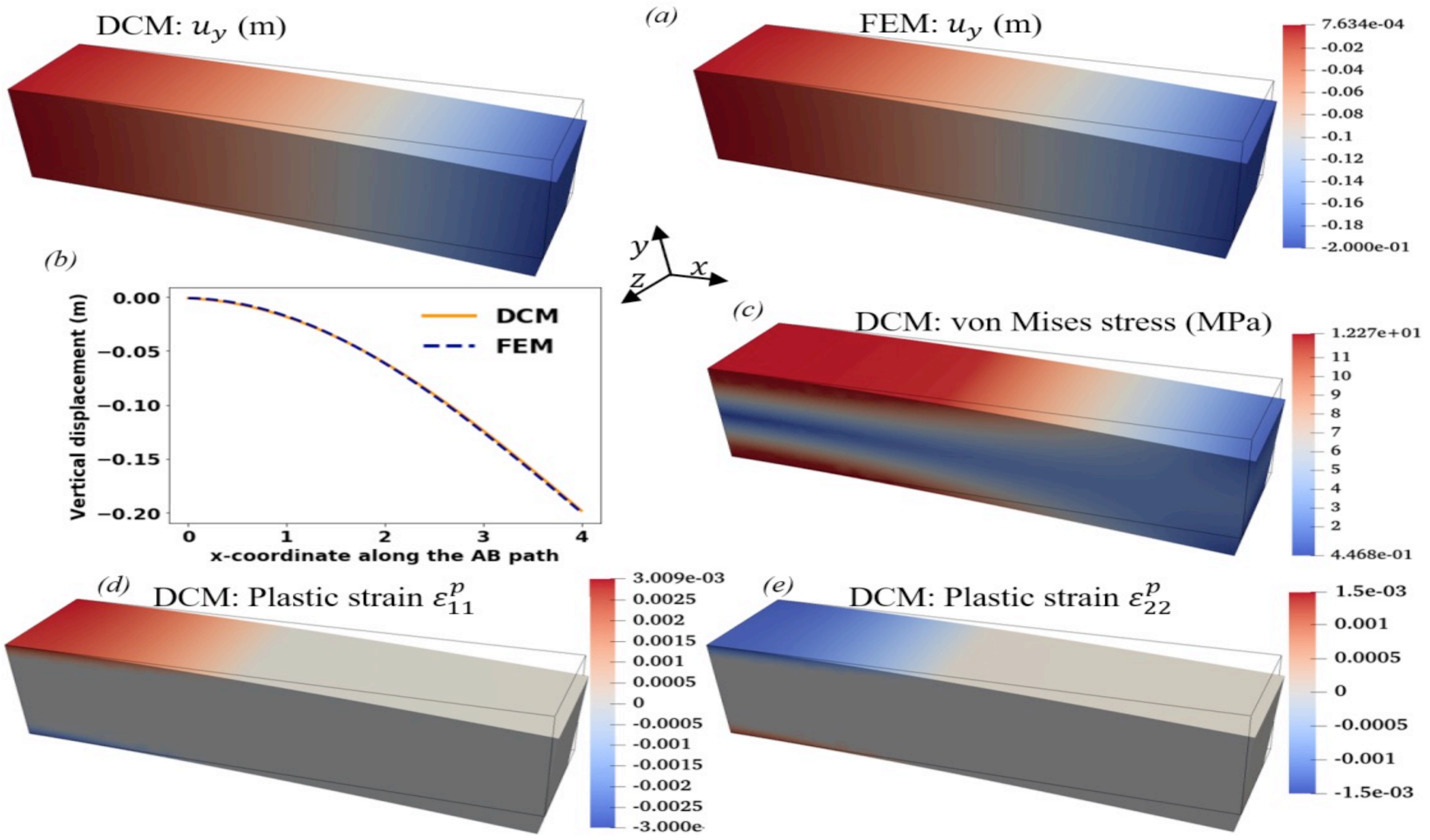
Meshless Physics Informed Deep Learning Method For 3-D Solid Mechanics



Abueidda et al., International Journal for Numerical Methods in Engineering, In Revision, 2021

- **Strong Form (Governing PDE) is NN's Loss,** enforced on collocation points in the domain and on essential and natural Boundary Conditions
- **No data generation** is needed, such as in surrogate deep learning models
- Relies on simple linear algebra kernels, which are **better optimized on GPU**-s then sparse direct solvers in implicit Finite Element Method (FEM) and use significantly less memory
- **Easier Discretization** of complex geometries without issues with the aspect ratio of elements, connectivity, and assembling into a global stiffness matrix
- **Globally smoother solution** without discontinuity in displacement solution between lower order elements in FEM

Meshless Physics Informed Deep Learning Method For 3-D Solid Mechanics



Abueidda et al., International Journal for Numerical Methods in Engineering, In Revision, 2021

The Ultimate Singularity in AI?

AI Reality Checks:

- No, machines can't read better than humans (2018)
  - *https://www.theverge.com/2018/1/17/16900292/ai-reading-comprehension-machines-humans*

- How IBM Watson Overpromised and Under-delivered on AI Health Care, IEEE Spectrum By Eliza Strickland, April 2019

- DeepMind's Latest A.I. Health Breakthrough Has Some Problems, by Julia Powles, August 2019

AI machines can "learn" but not yet "think" (at least not like humans), and it remains to be seen if, how, and when the major AI singularity point of true intelligence will be reached?

But be careful what you wish for!

Thanks to machine-learning algorithms,
the robot apocalypse was short-lived.

Acknowledgements:

- **IA Confluence Team Members**: Diab Abueidda, Shirui Luo, Nahil Sobh, Hunter Kollmann and Erman Guleyruz
- Dora Cai, Aiman Soliman, Jeff Terstriep & NCSA Data Analytics Team
- Christina Fliege, Liudmila Mainzer & NCSA Genomics Team
- NCSA Industry Program
- AI Center for Excellence at UIUC

Thank you!

Brendan McGinty – bmcginty@Illinois.edu

Dr. Seid Korić – koric@Illinois.edu

NCSA.Illinois.edu/Industry