MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

Climate simulation analysis: approach by Data Science

Data Science and Machine Learning

Data Science and Machine Learning

- Team profiles
  - Computational fluid dynamics modelers
  - Machine learning professionals
  - Computer science
  - Meteorologist
  - Statistician
  - Geodesist
  - Mathematician

Data Science and Machine Learning

- Climate change and risks: research opportunities
  - Data assimilation and model reduction
  - Data dimension reduction
  - Climate change and impact on urban atmospheric pollution
  - Extreme weather events

Data Science and Machine Learning

- Data assimilation and model reduction
  - High performance computing: hardware
  - Improving the computer power
  - Processor (transistor(s))
  - Vector processor (shared memory)
  - Multi-processor (shared memory)
  - Multi-processor (distributed memory)
  - Cloud and grid-computing
  - Multi-processor (multi-core processor)
  - Multi-processor (multi-core processor) + co-processor

Data Science and Machine Learning

- Data assimilation and model reduction
  - High performance computing: software
  - One example: FFT (changed the world)
  - Algorithm complexity
  - Lesson to be learned: do the same thing
  - Lesson to be learned: do in another way (efficiency)

Data Science and Machine Learning

- Data assimilation and **model reduction**
  - Do same thing
  - … but more efficient
  - Lesson to be learned!

Received: 26 February 2017 | Revised: 14 July 2017 | Accepted: 14 July 2017
DOI: 10.1002/fld.4416

RESEARCH ARTICLE

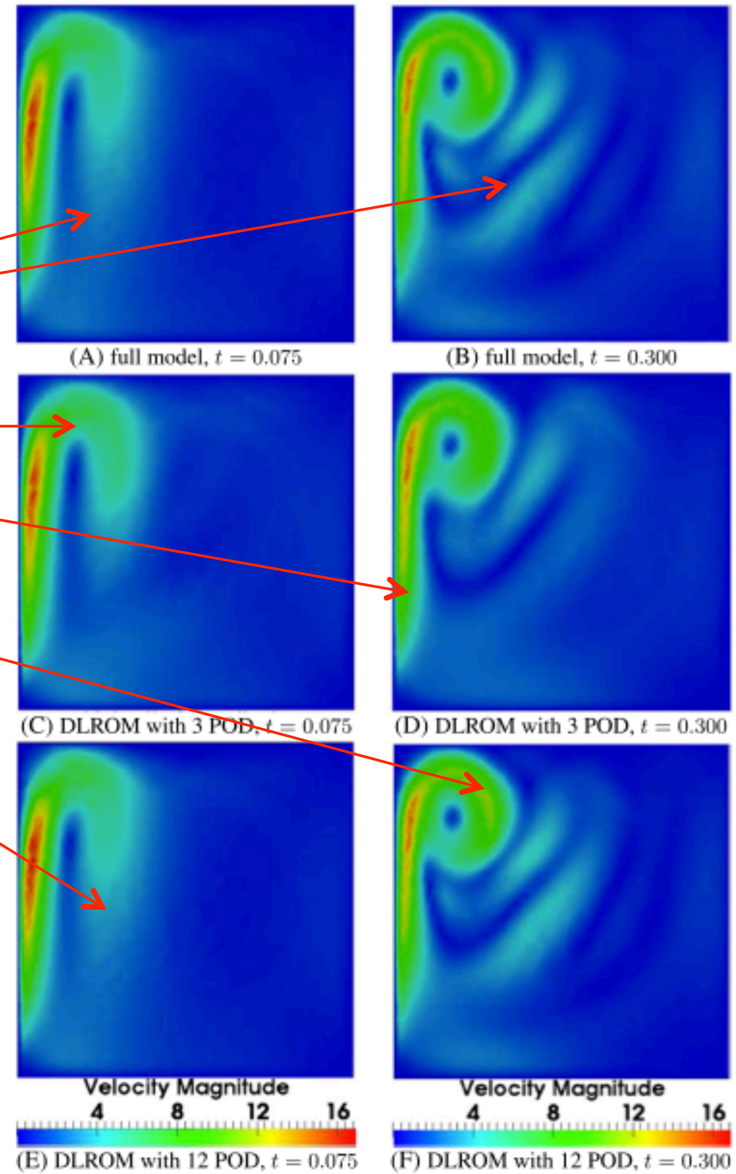
WILEY

Model identification of reduced order fluid dynamics systems using deep learning

Z. Wang[1] | D. Xiao[2,4,5] | F. Fang[2,5] | R. Govindan[3] | C. C. Pain[2,5] | Y. Guo[5]

Data Science and Machine Learning

- DA and **model reduction**
  - Deep Learning for ROM
  - Full (“Exact”)
  - 3 POD
  - 12 POD
  - (POD: Proper Orthogonal Decomposition
  - (ROM: Reduced Order Model)



Data Science and Machine Learning

- Data assimilation and model reduction
  - Do same thing … but more efficient
  - NN doing same thing than LETKF (COAPS-FSU)

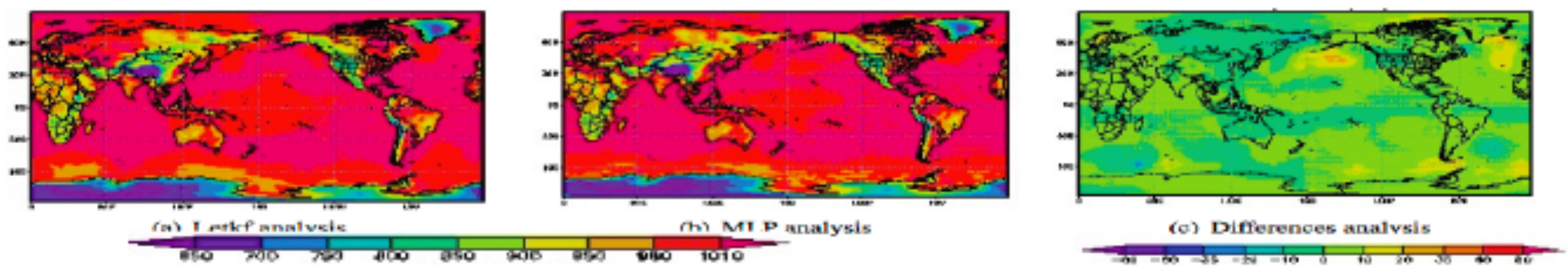


Fig. 7 -Surface Pressure (PS) [hPa] Fields 03/01/2004 at 06 UTC. (a) LETKF analysis (b) MLP-DA analysis (c) differences between LETKF and MLP-DA analyses.

Total running time of 124 cycles or January/2004.

Execution of 124 cycles	MLP-DA (hour:min:sec)	LETKF (hour:min:sec)
Analysis time	00:02:25	11:01:20
Ensemble time	00:00:00	15:50:40
Parallel model time	00:27:20	00:00:00
Total Time	00:29:45	26:52:00

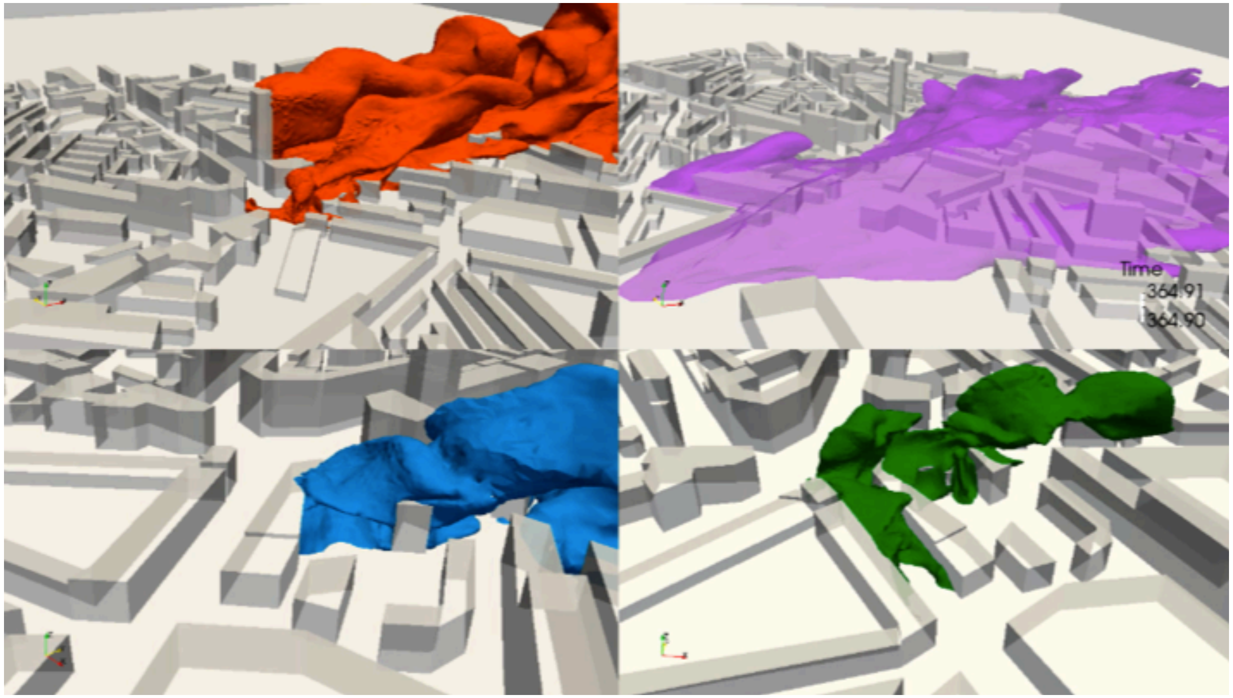
AGU FALL MEETING
San Francisco | 14–18 December 2015

Data Science and Machine Learning

- Data dimension reduction
  - As already mentioned: SVD_POD
  - Methodologies based on computational intelligence
  - Kaizen Programming
  - Decision Tree
    - Random Forest
  - ANN:
    - MLP (S-NR), RBF (S-NR), Elman/Jordan (S-R),
    - Hopfield (unS-R), SOM (unS-R)
    - Deep-Learning
    - Kaizen Programming
  - Fuzzy, Neuro-Fuzzy

Data Science and Machine Learning

- CC and impact on Urban Atmospheric Pollution
  - Impact on health of communities
  - Sustainable cities: VAPOR (**V**irtual City **A**ir **PO**llution Fast **R**esponse Model)



Data Science and Machine Learning

- CC and impact on Urban Atmospheric Pollution
  - Sustainable cities: VAPOR
  - IPCC scenarios: downscaling techniques
  - Analysis of different scenarios:
    - under detailed simulations: reduction of problem dimensionality

Data Science and Machine Learning

- Extreme weather events
  - IPCC report: more intense and more frequent extreme events

Data Science and Machine Learning

- Extreme weather events
  - Intense rain fall (storms, hurricane/typhoon)
  - Floods
  - Deep drought
  - Heatwaves
  - Scales:
    - Short and medium term (1 day – 10 days)
    - Sub-seasonal (2 up to 4 weeks)
    - Climate (more than 4 weeks)

Data Science and Machine Learning

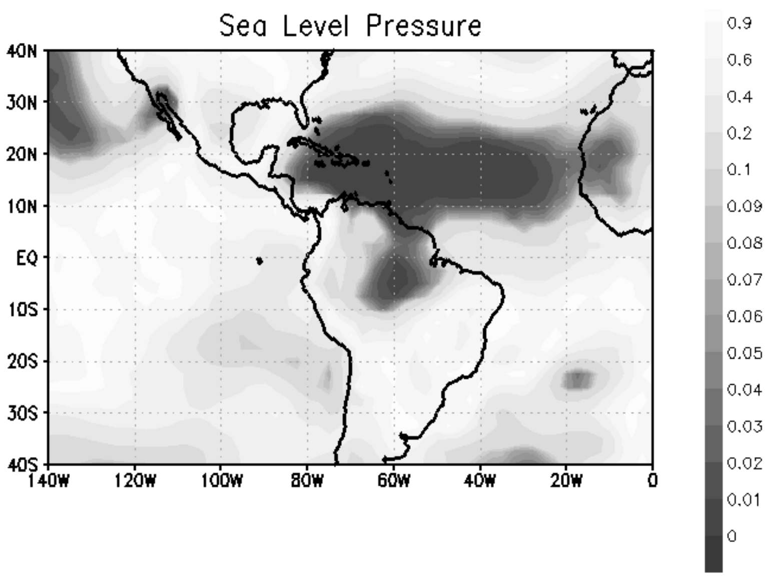
- Extreme weather events (methods for detections)
  - Data reduction:
    - Raizen approach
    - P-value evaluation
    - Rough set theory
    - Map reduce
  - Detection:
    - Deep Learning (LSTM: Long-Short Term Memory)
    - Decision Tree
    - Random forest
  - Example:
    - P-value (reduction) + Decision Tree (detection)

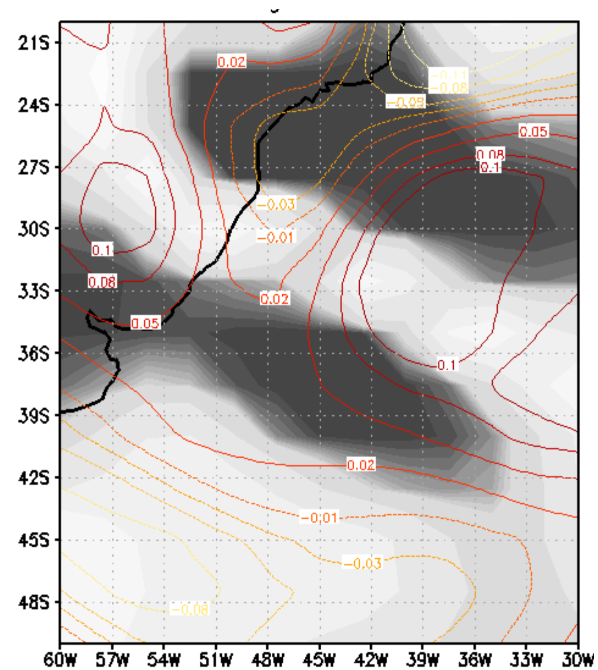
Data Science and Machine Learning

- Extreme weather events (methods for detections)
  - Example: (Nov/2008: Santa Catarina state, Brazil)
    - P-value (reduction) + Decision Tree (detection)

Data Science and Machine Learning

- Extreme weather events (methods for detections)
  - Example: (Nov/2008: Santa Catarina state, Brazil)
    - **P-value (reduction)** + Decision Tree (detection)





Data Science and Machine Learning

- Extreme weather events (methods for detections)
  - Example: (Nov/2008: Santa Catarina state, Brazil)
    - P-value (reduction) + **Decision Tree (detection)**