Imperial College London

ESA-ECMWF Workshop 2022

Data Learning for more reliable digital twins

Dr Rossella Arcucci

Department of Earth Science & Engineering,
Data Learning working group,
Data Science Institute,
AI network speaker at ICL (~250 academics),
World Meteorological Organization wg-member,

r.arcucci@imperial.ac.uk
https://www.imperial.ac.uk/people/r.arcucci

CONTENT

- AI & Digital Twins (Intro)
- Data Learning (models)
- Examples (**air pollution**, **energy** convertors, energy control systems, **wildfires**, fluids flow in pipes, **ocean**)

EFFICIENCY

ACCURACY

mocean energy

dyson

LEVERHULME Centre for Wildfires, Environment and Society

BREATHE LONDON

BR PETROBRAS

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici

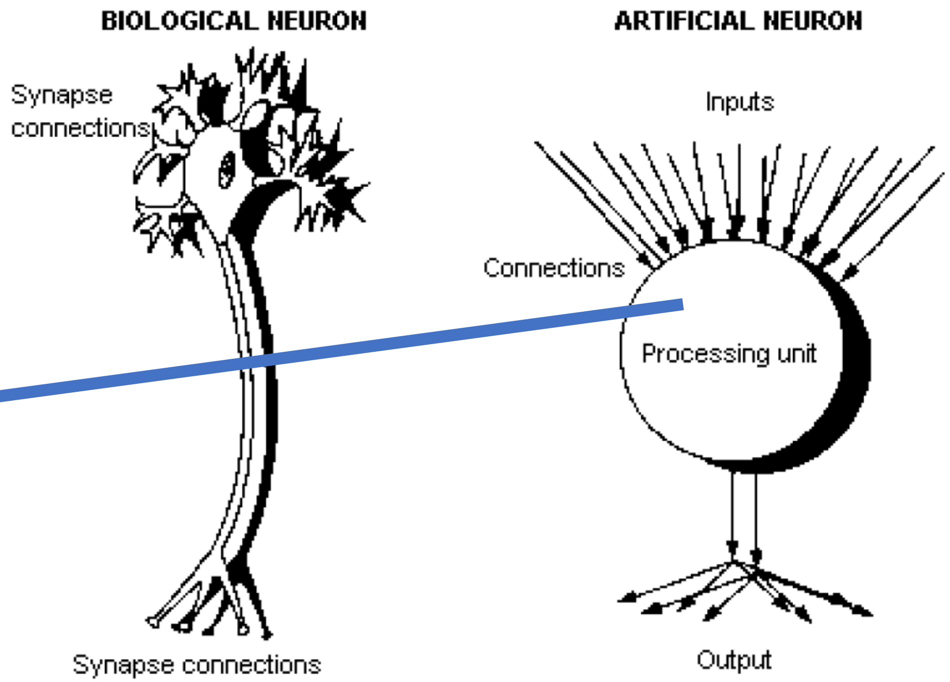
What is AI?

Artificial Intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.

What is AI?

Artificial Intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.

ARTIFICIAL NEURAL NETWORK



What is AI?

Artificial Intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.

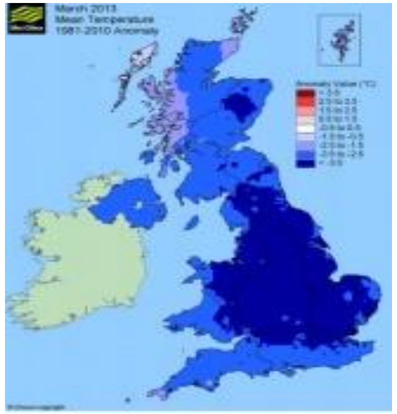
DATA

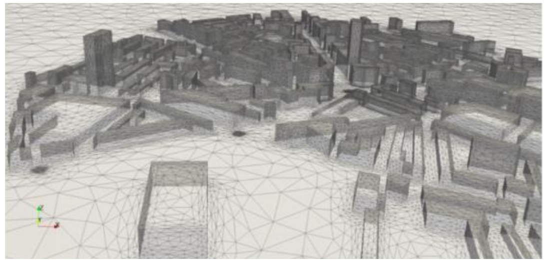
SYNONYMS

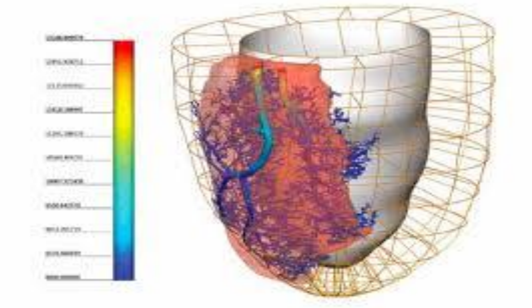
facts, figures, statistics, details, particulars, specifics, features
information, evidence, intelligence, material, background, input
proof, fuel, ammunition
statement, report, return, dossier, file, documentation, archive, archives

… the era of the data!

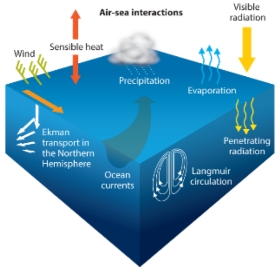
High resolution Models…

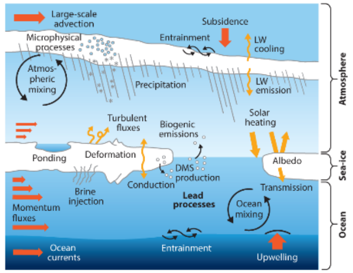


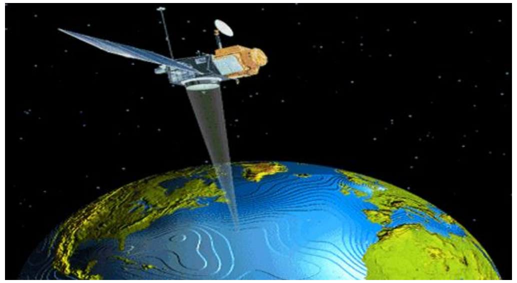


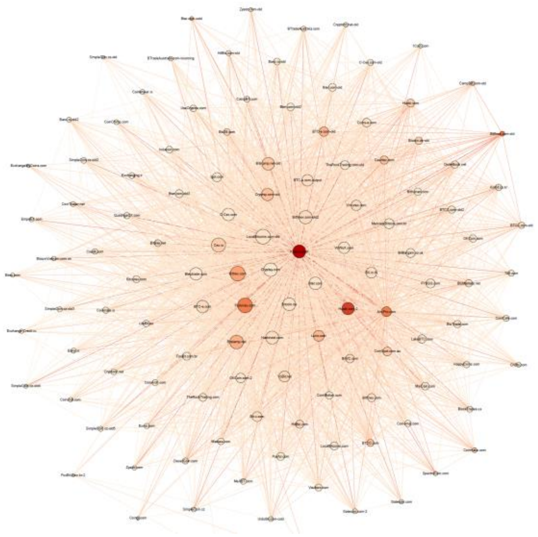


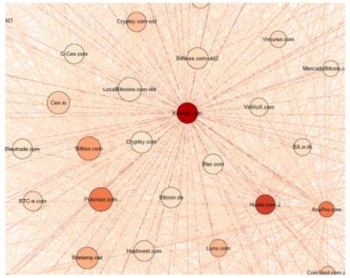
Real observations…

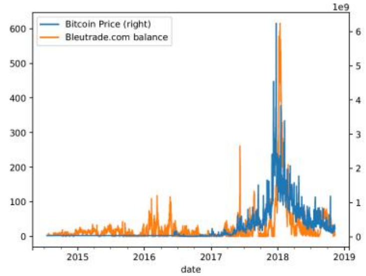








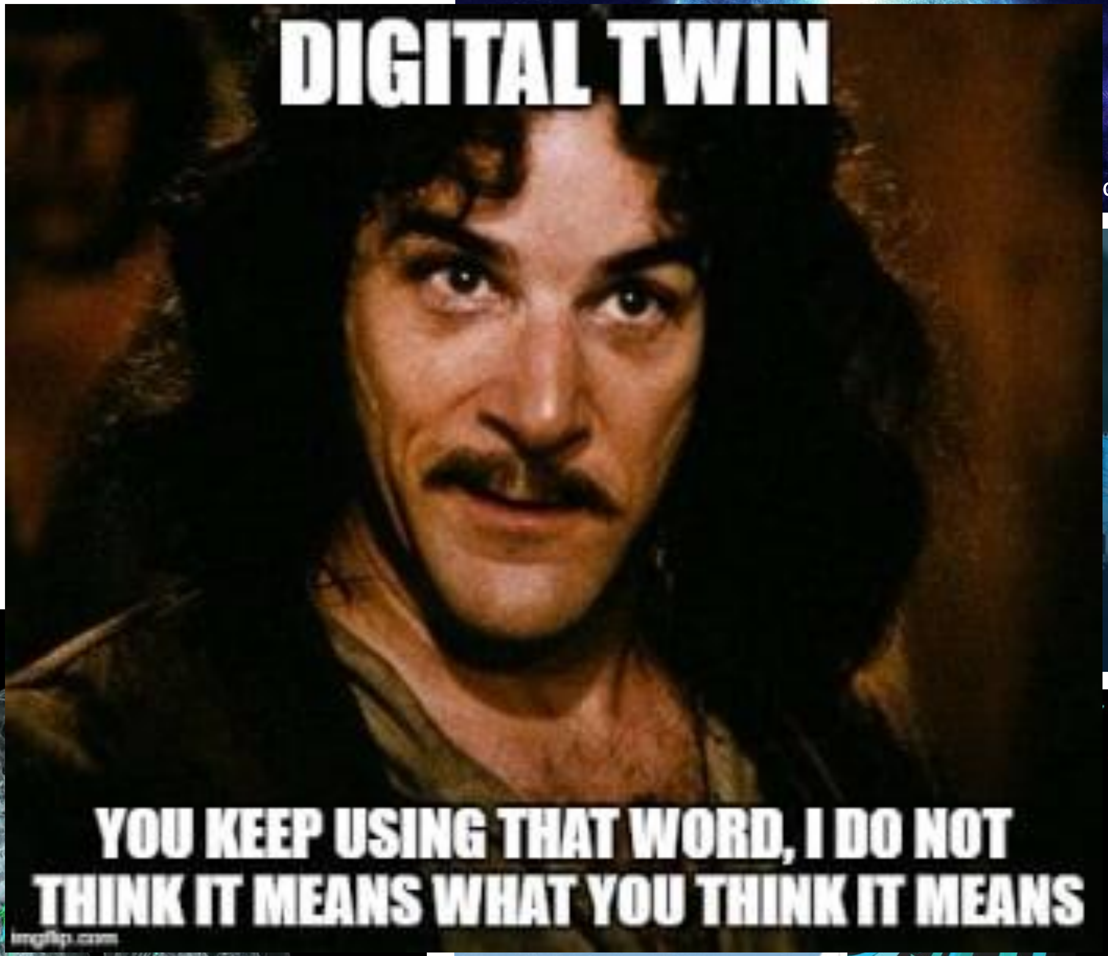


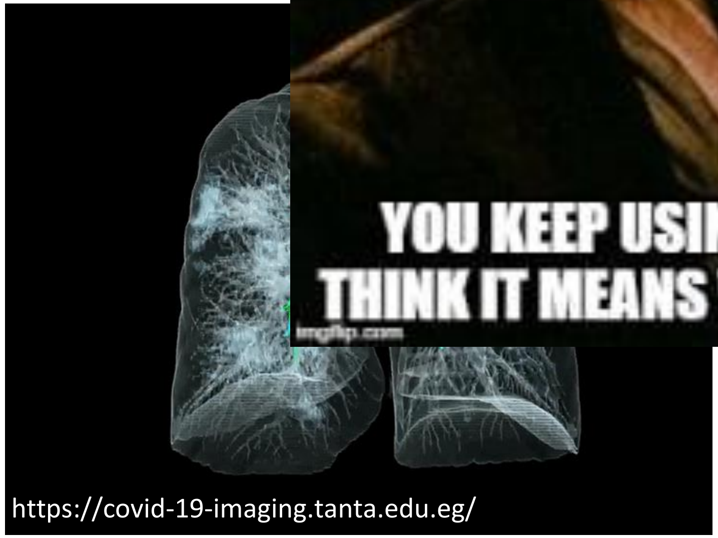


main motivations/contributions of Data Learning

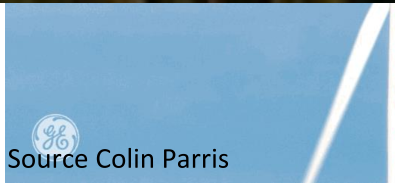
Source Microsoft Blog Europe's open data revolution: the road to collaboration

ource ESA





https://covid-19-imaging.tanta.edu.eg/



Source Colin Parris

Digital Twins

THE ERA OF ... DIGITAL TWINS for managing risks

Neurocomputing
Volume 470, 22 January 2022, Pages 11-28

Digital twins based on bidirectional LSTM and GAN for modelling the COVID-19 pandemic

César Quilodrán-Casas [a, b], Vinicius L.S. Silva [b], Rossella Arcucci [a, b], Claire E. Heaney [b], YiKe Guo [a], Christopher C. Pain [a, b]

applied sciences
MDPI

Review

Is Digital Twin Technology Supporting Safety Management? A Bibliometric and Systematic Review

Giulio Paolo Agnusdei [1,2,*], Valerio Elia [1] and Maria Grazia Gnoni [1]

Hindawi
Advances in Civil Engineering
Volume 2020, Article ID 8888876, 10 pages
https://doi.org/10.1155/2020/8888876

Research Article

Digital Twin-based Safety Evaluation of Prestressed Steel Structure

Zhansheng Liu, Wenyan Bai, Xiuli Du, Anshan Zhang, Zezhong Xing, and Antong Jiang

BIC MAGAZINE
BUSINESS & INDUSTRY CONNECTION

Digital twin technology promotes safety, reduces costs

BY NANCY FORD JANUARY 27, 2021 2:06 PM

According to Greg Withers, projects modernization and transformation director for BP, the advantages for companies introducing digital twin

main motivations/contributions of Data Learning

Digital Twins

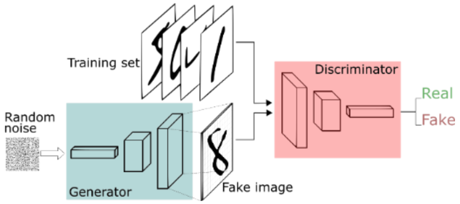
- **Data driven models - Surrogate Models - Physics Informed Machine Learning**
- **Uncertainty Quantification and minimization**
- **Explainable AI**

Main Challenges/Questions when working with data

Assuming your data is meaningful

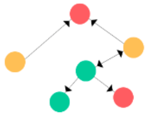
1. enough/not enough
2. structured/unstructured
3. too big (Big Data)
4. updated

1. Create synthetic-realistic data



Generative ML models

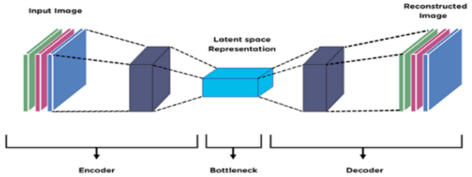
2. Manage unstructured data



Graph NN

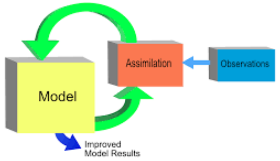
3. Compress the data (to develop reduce order models)

Encoder-Decoder



4. Fine tune your systems as soon as new info come available

Data Assimilation

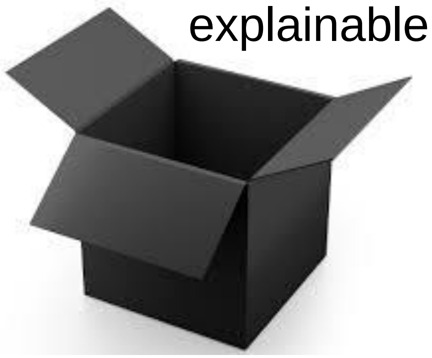


Managing Data for AI

Building data-driven models becomes difficult in many real-world scenarios due to:

- **Dimensionality constraints**: matrices become so large that they are difficult to work with.
- **Noisy data**: uncertainty and noise in the data creates serious error propagation
- **Low-quality data**: the data do not provide meaningful information over the whole field

DA for more explainable AI



Data Assimilation is the missing piece!!!

Uncertainty quantification and minimization:

Data Assimilation + Machine Learning = Data Learning models

Data Assimilation + Machine Learning = Data Learning

data learning

All the models and the technologies which have been developed at DataLearning working group are completely general and applied to a lot of different real world applications.

Grants:

EPSRC Engineering and Physical Sciences Research Council

INHALE

MAGIC
Envisaging a world with greener cities

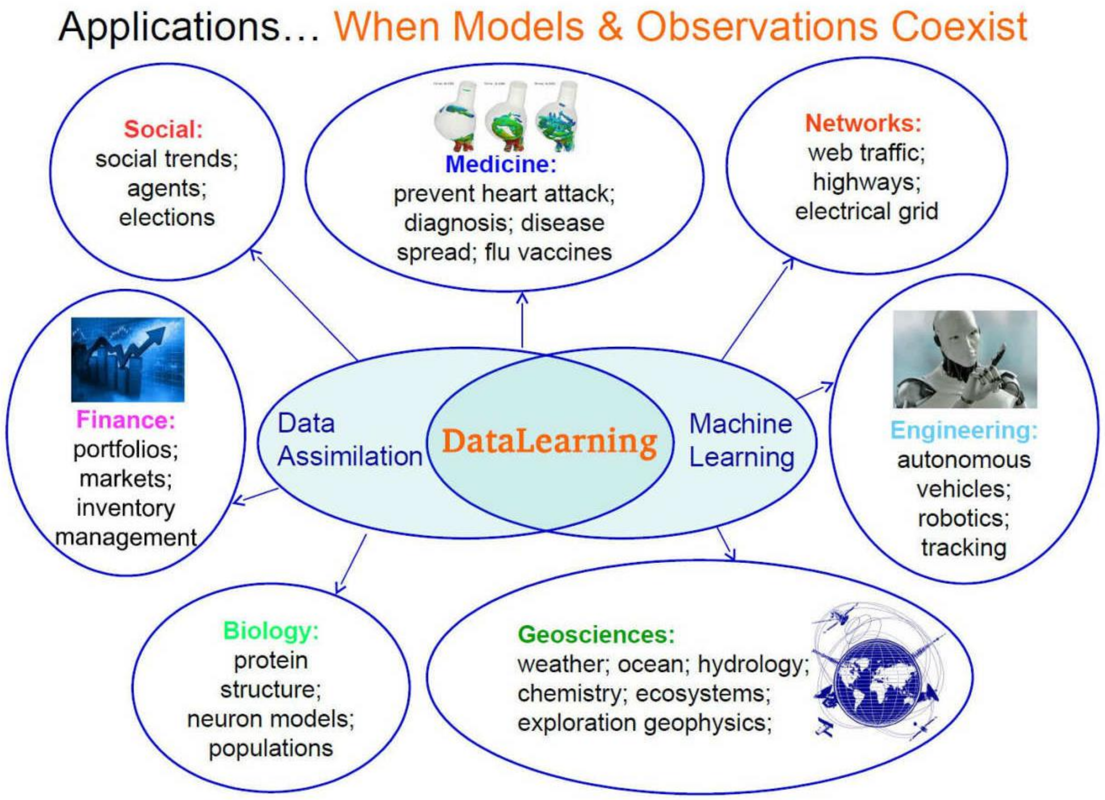
PREMIERE

RELIANT

LEVERHULME
Centre for Wildfires, Environment and Society

Leonardo Centre for Sustainable Business

WAVE-SUITE



Our Academic Collaborations:

UNIVERSITY OF CAMBRIDGE

CITY UNIVERSITY OF LONDON EST 1894

UNIVERSITY OF SURREY

University of Glasgow

KING'S College LONDON

Imperial College Business School

UCL

UNIVERSITY OF BIRMINGHAM

THE UNIVERSITY OF EDINBURGH

https://www.imperial.ac.uk/data-science/research/research-themes/datalearning/

Connecting the dots

Imperial College Business School

… Enthusiasm is the key for a successful FUTURE!!!

Journal of Computational Science 58 (2022) 101525

Contents lists available at ScienceDirect

Journal of Computational Science

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

Data Learning: Integrating Data Assimilation and Machine Learning☆

Caterina Buizza[b], César Quilodrán Casas[a], Philip Nadler[a], Julian Mack[a], Stefano Marrone[a,f], Zainab Titus[c], Clémence Le Cornec[d], Evelyn Heylen[e], Tolga Dur[a], Luis Baca Ruiz[a,g], Claire Heaney[c], Julio Amador Díaz Lopez[a,h], K.S. Sesh Kumar[a], Rossella Arcucci[a,c,*]

[a] *Data Science Institute, Imperial College London, UK*
[b] *Personal Robotics Lab, Department of EEE, Imperial College London, UK*
[c] *Department of Earth Science and Engineering, Imperial College London, UK*
[d] *Department of Civil and Environmental Engineering, Imperial College London, UK*
[e] *Control and Power Group, Department of EEE, Imperial College London, UK*
[f] *DIETI, University of Naples Federico II, Italy*
[g] *Department of Computer Science and Artificial Intelligence, University of Granada, Spain*
[h] *Data Science Intitute, London School of Economics and Political Science, UK*

Data Learning: a modular approach

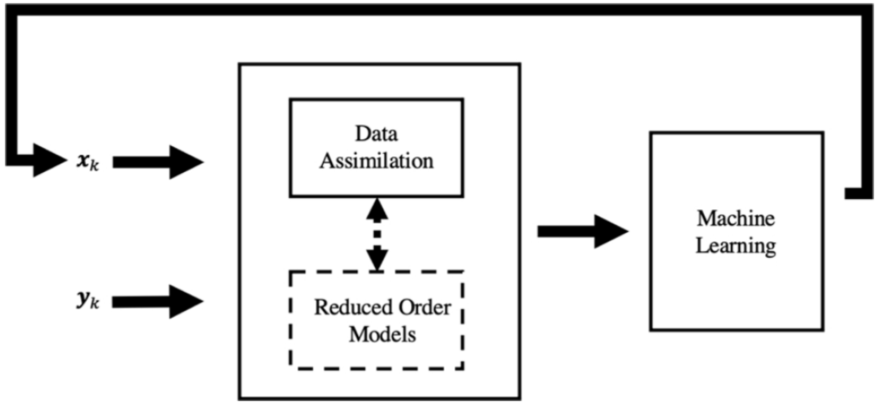
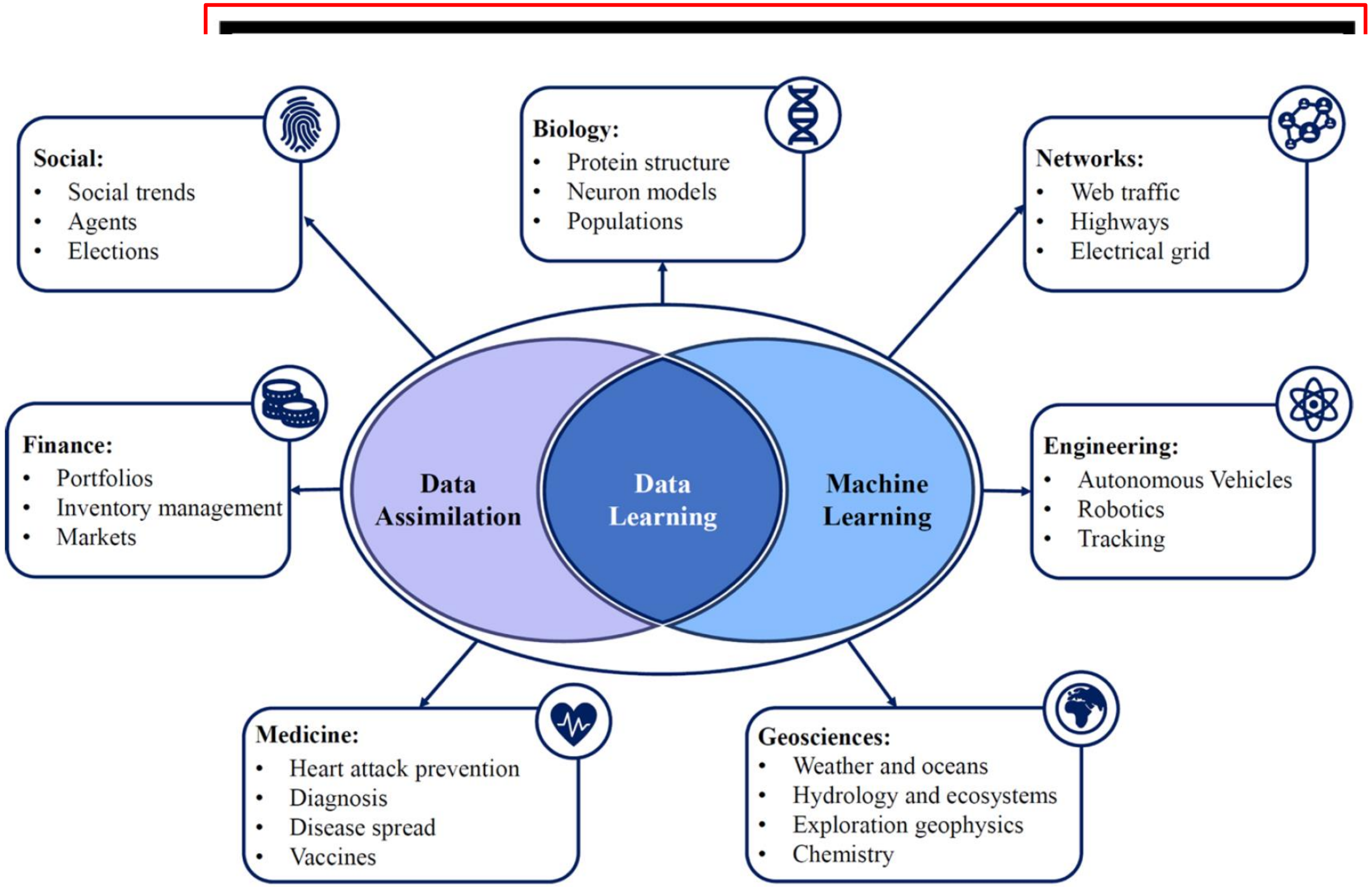
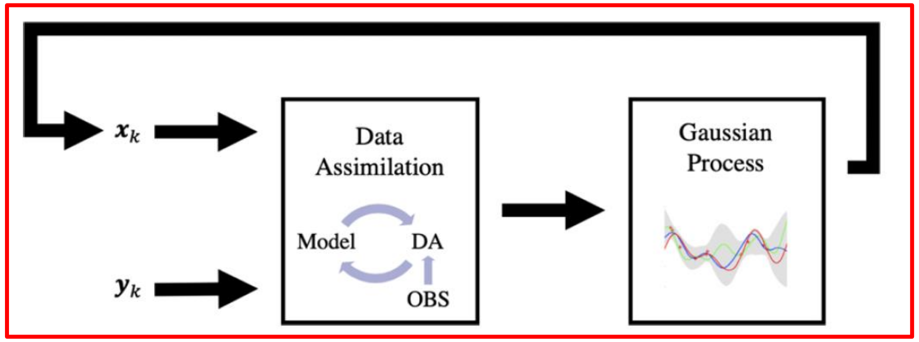
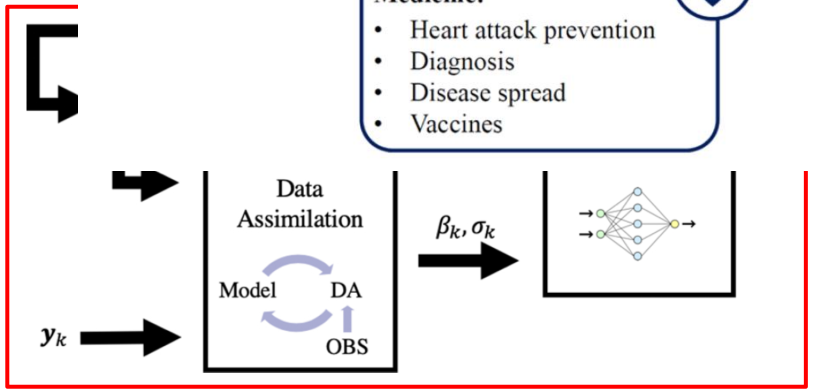
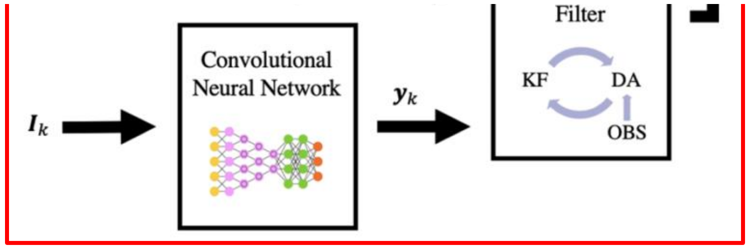


Fig. 1. General Data Learning framework.









Our main models/approaches

data learning

ACCURACY **(ERROR)**

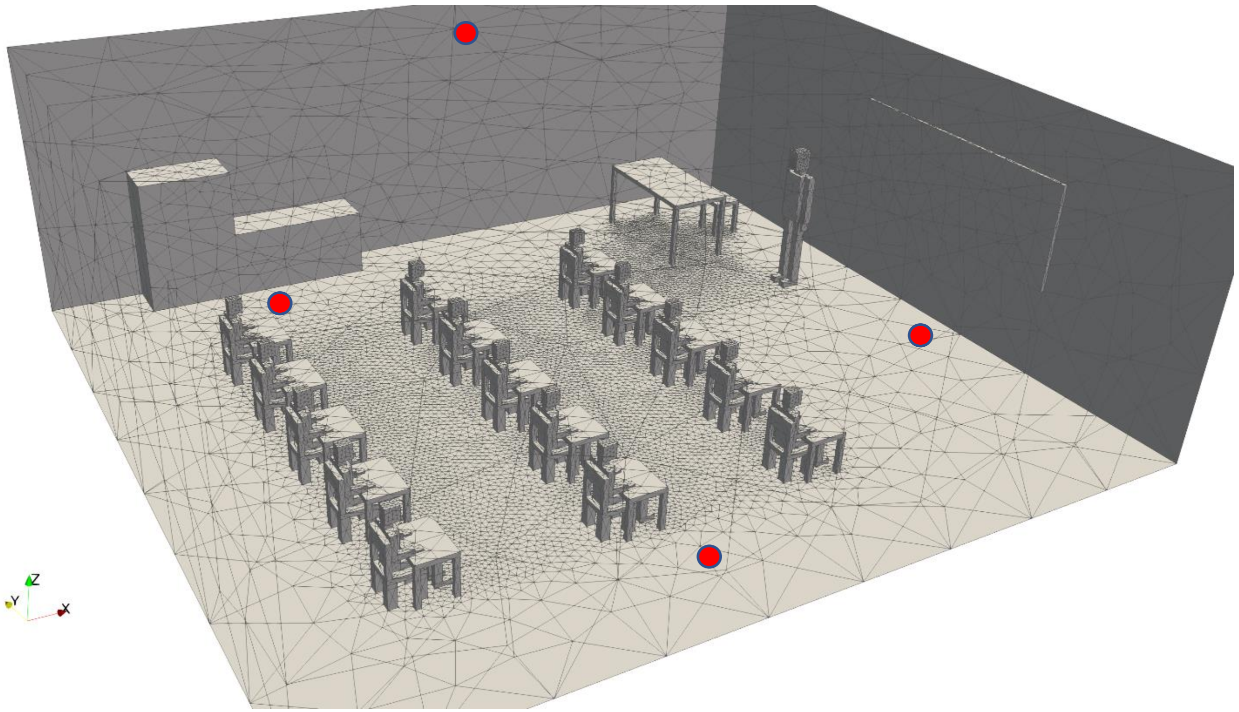
EFFICIENCY **(TIME)**

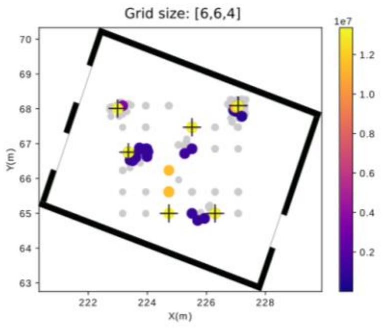
	OFFLINE: R&D (CLEANING, TRAINING)	**ONLINE: PRODUCTION** (ADJUSTING, RUNNING)
ACCURACY (ERROR)	Optimal Data Selection Parameters Estimation Data Augmentation	Data Assimilation
EFFICIENCY (TIME)	Surrogate models (training) Data Driven models	Data Learning Surrogate models (forecasting)

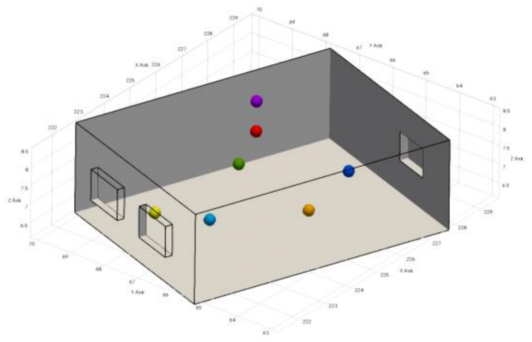
PRE-PROCESS: Error Analysis, Error Distribution, Error Covariance

Decision-making

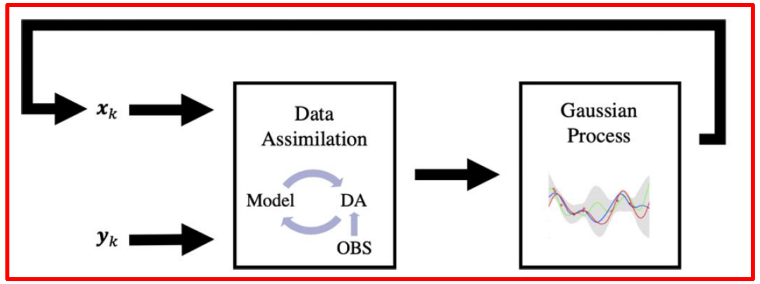
Optimal data selection, Optimal sensor placement, Big Data

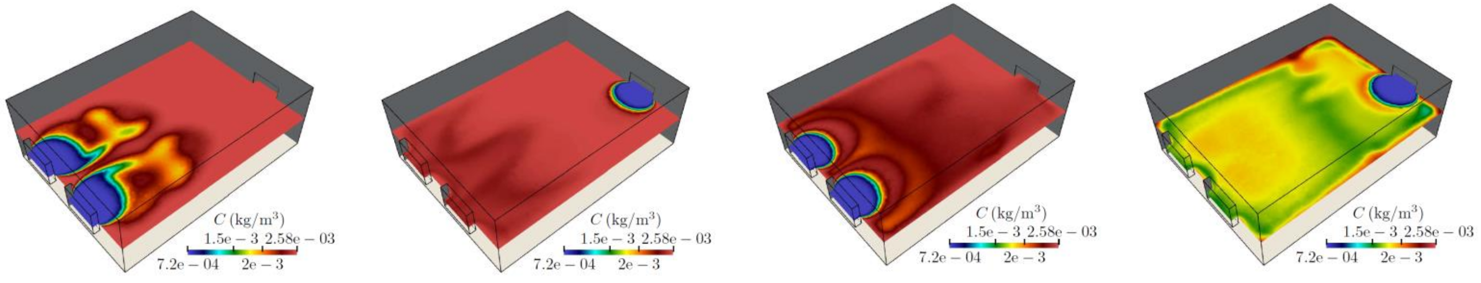






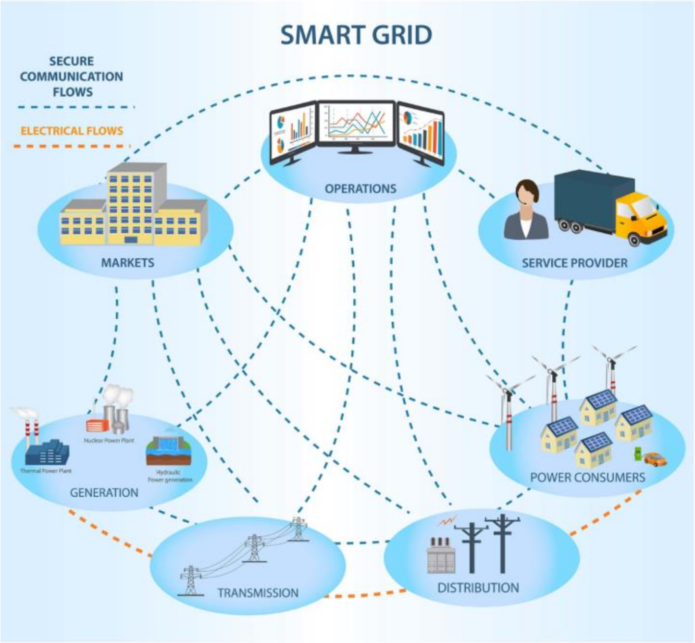
Our model is based on **Gaussian processes, Mutual Information and Data Assimilation**



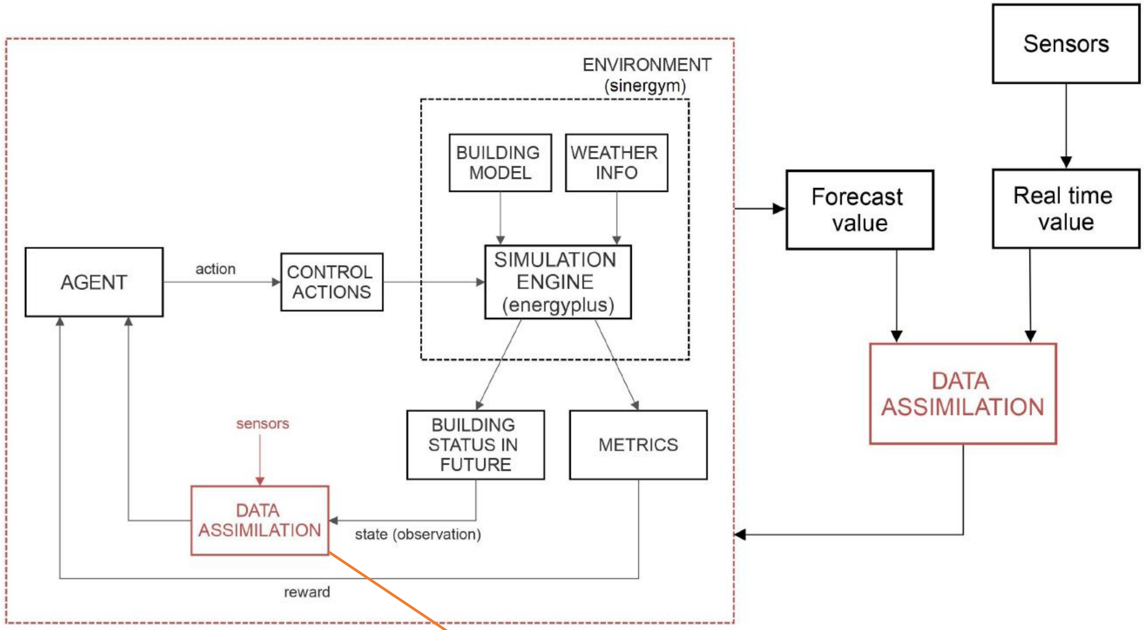


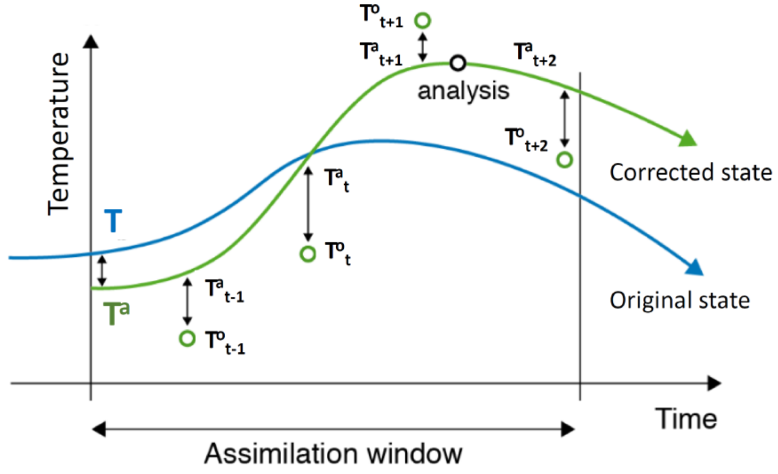
(a) $t = 15\,\mathrm{sec}$ (b) $t = 1\,\mathrm{min}$ (c) $t = 2\,\mathrm{min}\ 30\,\mathrm{sec}$ (d) $t = 5\,\mathrm{min}$

Energy Control Systems with Data Assimilation



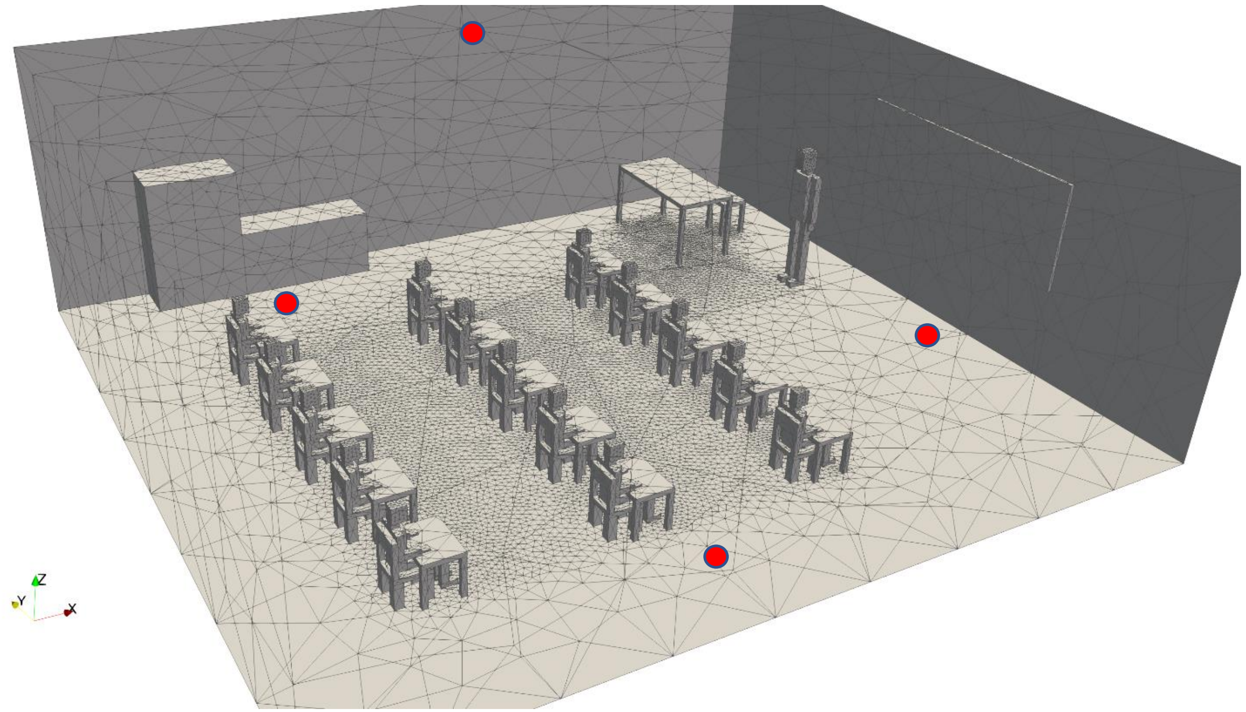
***with Alex Dmitrewski (2021)**





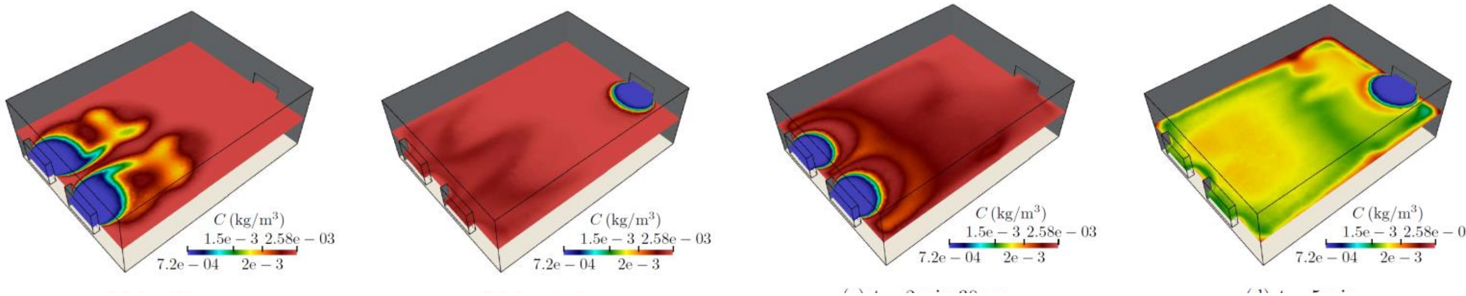
[*] Dmitrewski, A., Molina-Solana, M., & Arcucci, R. (2022). CntrlDA: A building energy management control system with real-time adjustments. Application to indoor temperature. *Building and Environment*, 108938.

Optimal data selection, Optimal sensor placement, Big Data



Our model is based on **Gaussian processes, Mutual Information and Data Assimilation**

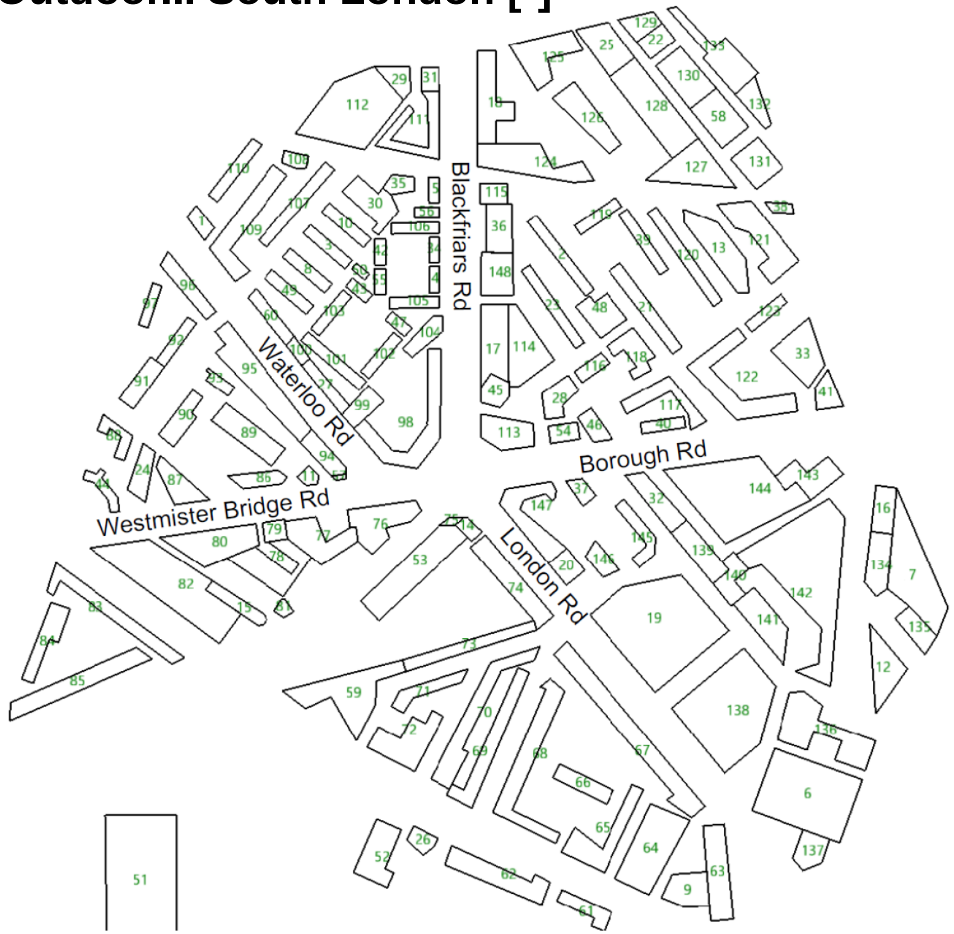
Assimilating the optimal positions, the **error of the predictive model**, i.e. **Fluidity**, is reduced by up to three order of magnitude: MSE(C^n) = 0,17 and MSE(C^{DA}) = 0,0005

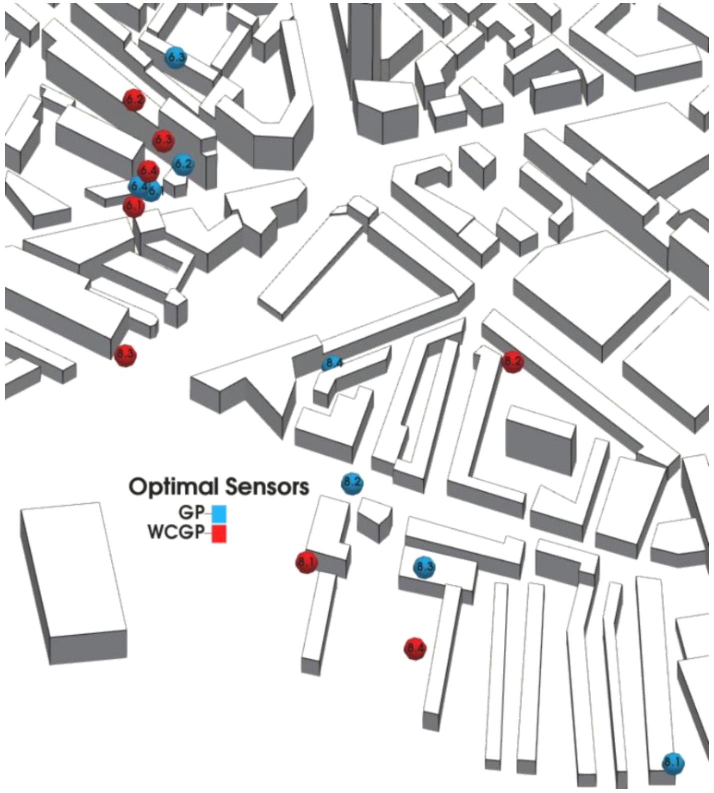


[*] T. Dur, R. Arcucci, L. Mottet, M. Molina Solana, C. Pain, Y. Guo - **Weak Constraint Gaussian Process for optimal sensor placement**-Journal of Computational Science
[**] G. Tajnafoi, R. Arcucci, L. Mottet, Molina Solana, C. Pain, Y. Guo - **Variational Gaussian Processes for optimal sensor placement**-Journal of Applied Mathematics

Machine Learning models for optimal sensor placement

Outdoor... South London [*]





	Real Mean	Estimated Mean	$MSE(\mathbf{x}^{\mathcal{M}})$	$MSE(\mathbf{x}^{DA})$
Original Algorithm	2.4662e-01	1.9598e-01	2.24e-01	5.25e-02
Data Learning (GP+DA)	2.4662e-01	2.2771e-01	1.77e-01	3.35e-02
Random	2.4662e-01	2.3900e02	6.54e00	8.90e-01

EPSRC

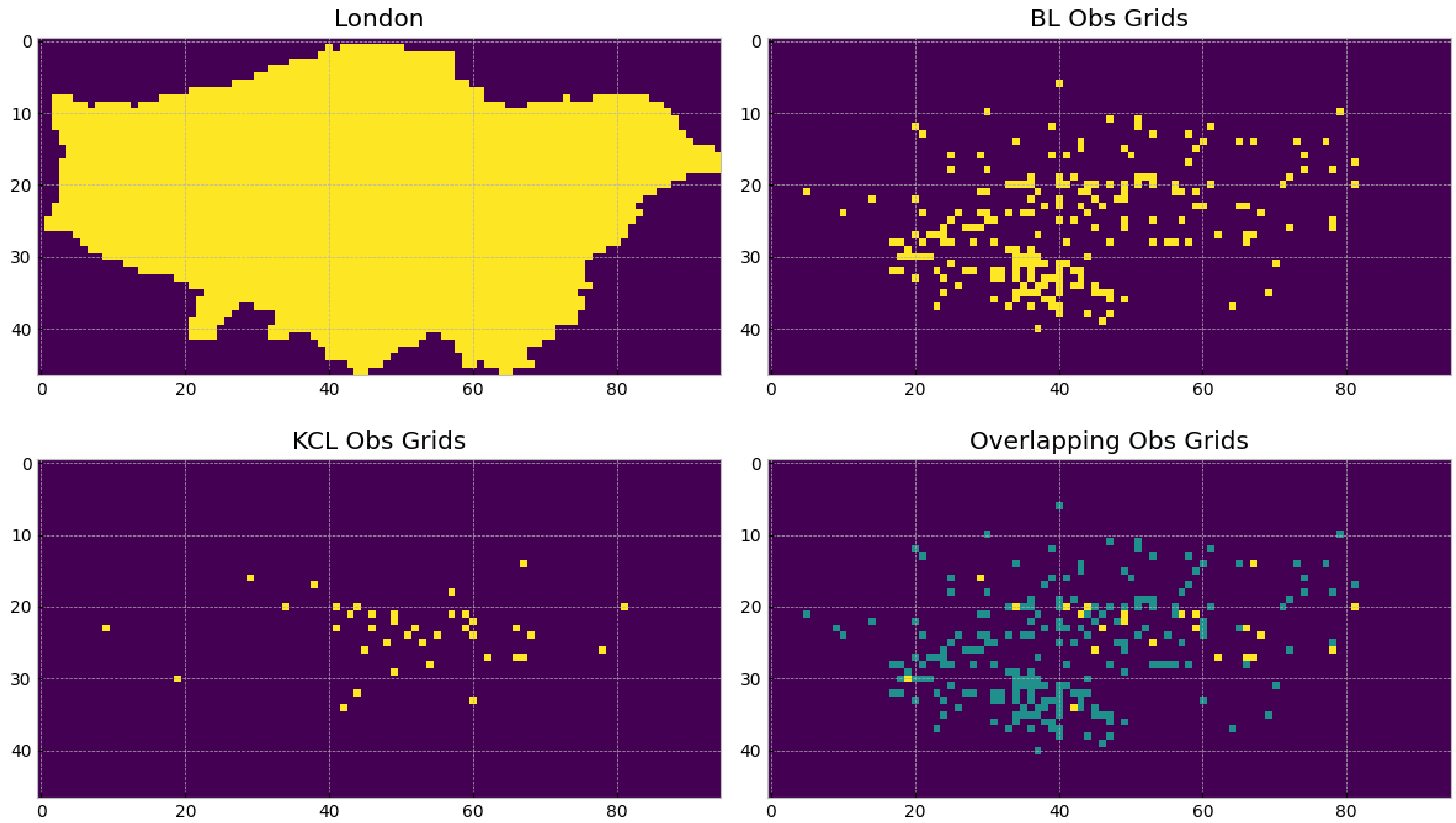
INHALE

MAGIC
Envisaging a world with greener cities

[*] T. Dur, R. Arcucci, L. Mottet, M. Molina Solana, C. Pain, Y. Guo - **Weak Constraint Gaussian Process for optimal sensor placement**-Journal of Computational Science
[**] G. Tajnafoi, R. Arcucci, L. Mottet, Molina Solana, C. Pain, Y. Guo - **Variational Gaussian Processes for optimal sensor placement**-Journal of Applied Mathematics

Breath London and London Regulatory Monitor Network

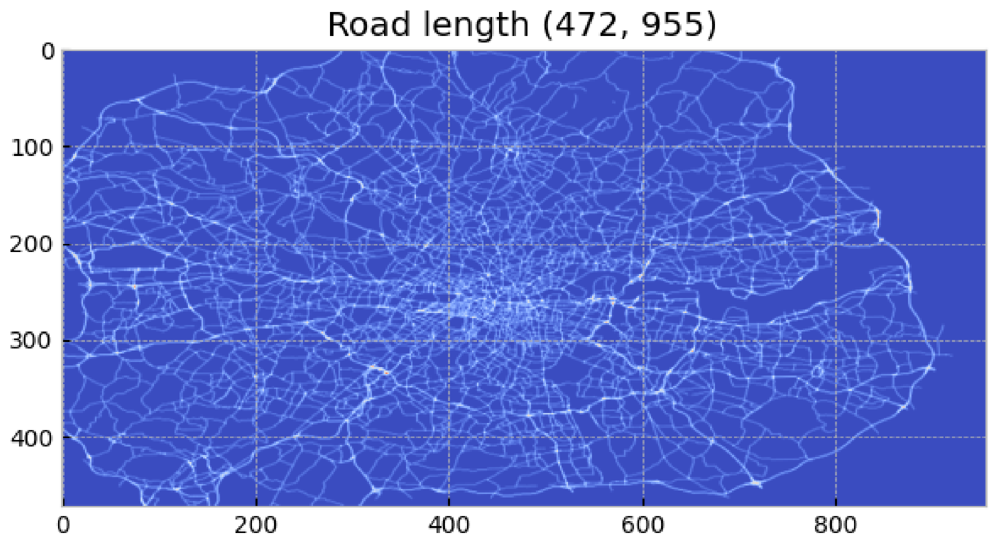
Distribution of PM sites (1km) resolution)

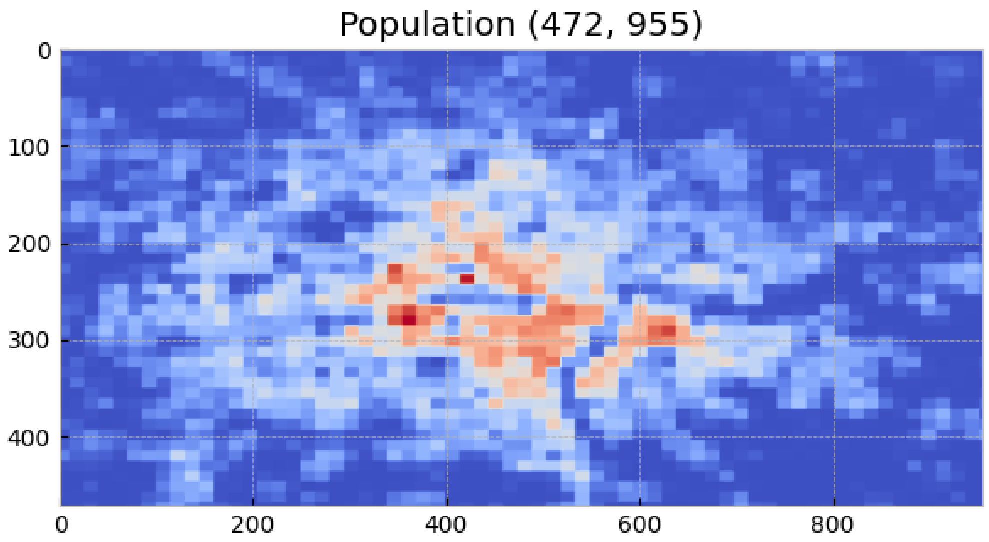


Hongwei Fan

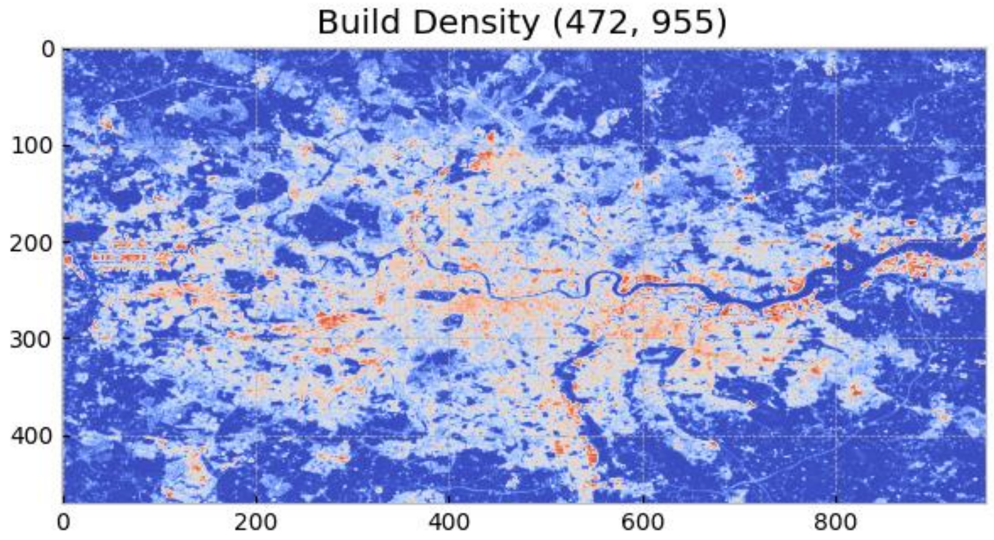
KCL is Regualtory Monitoring Network

Land use characteristics

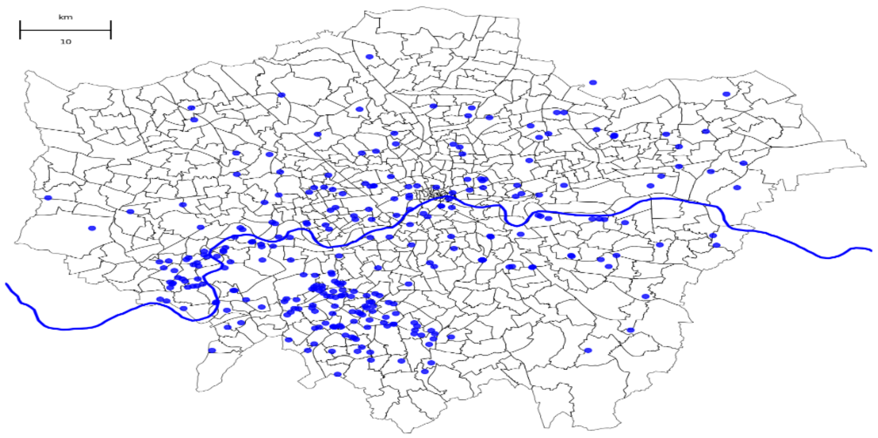




Hongwei Fan



Weather data

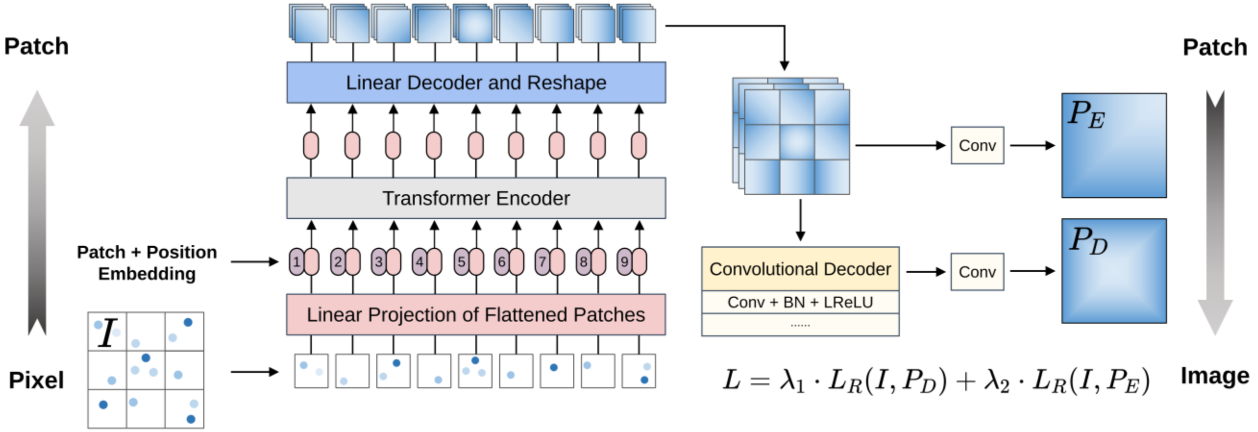


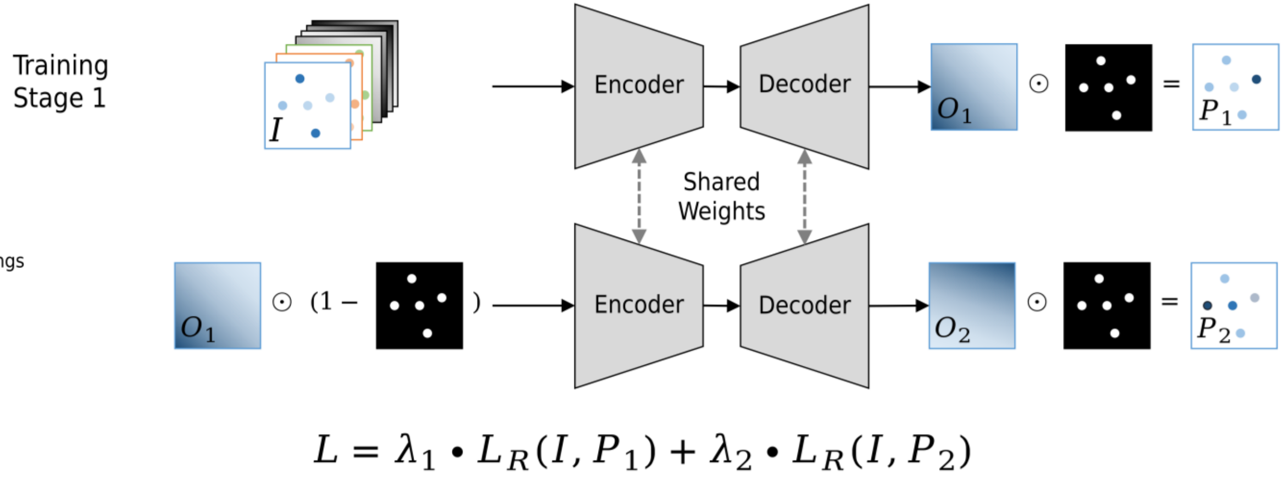
BL nodes also measures temperature and humidity

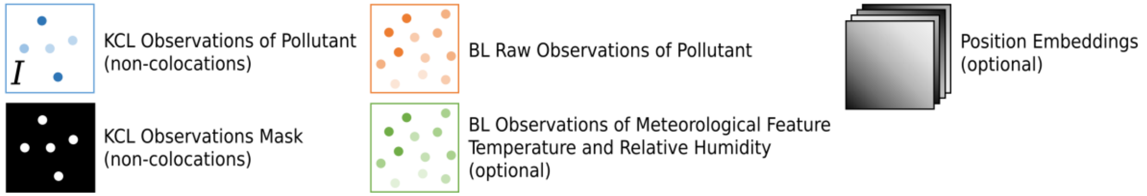
A deep learning model for fine-scale air quality estimation (0.1km)

we propose a novel vision transformer-based autoencoder (ViTAE) deep learning model for large-scale and complex field reconstruction.

Hongwei Fan



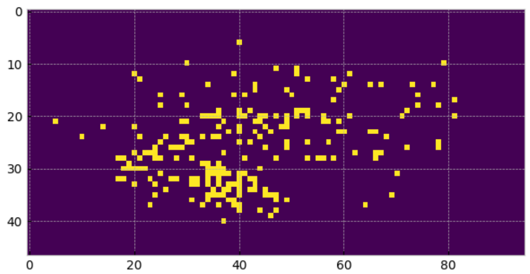




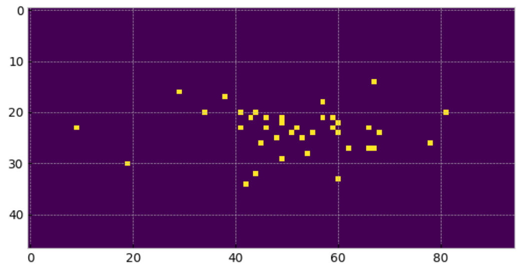
A deep learning model for fine-scale air quality estimation (0.1km)

Here we present the results of PM2.5 estimation with different data.

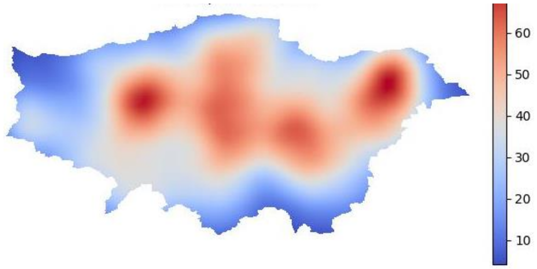
Low-cost sensors+ Regulatory sensors



+



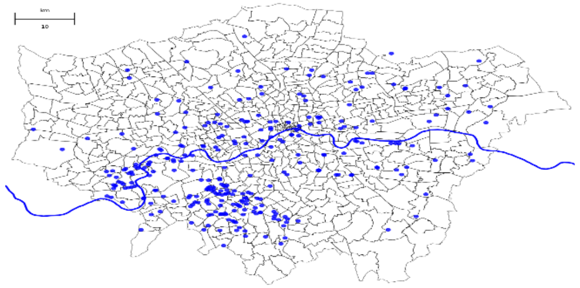
=



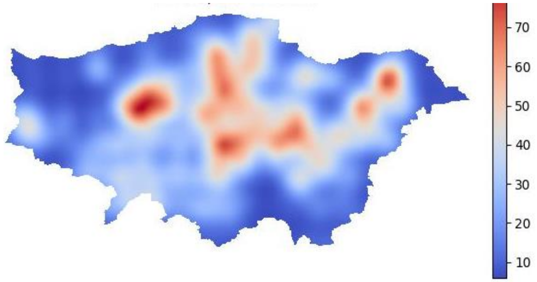
Hongwei Fan

+ Weather data

+



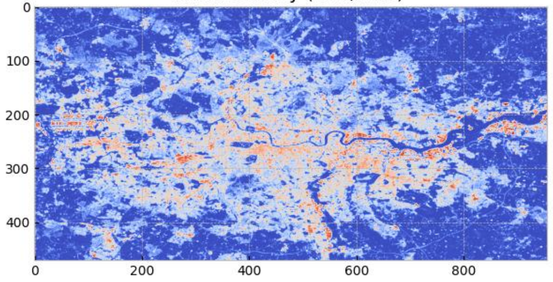
=



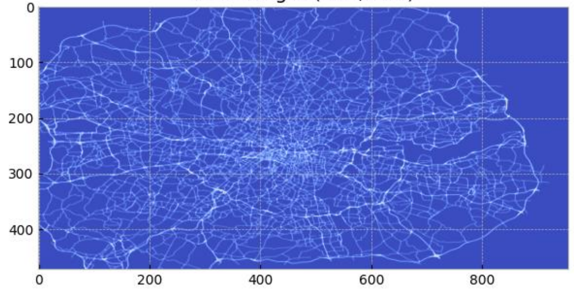
+ Land use characteristics

+

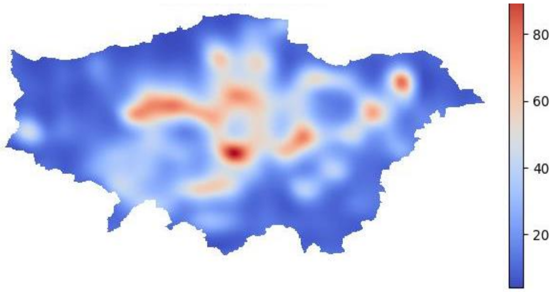
Build Density (472, 955)



Road length (472, 955)



=



CNN	0.3978	0.3527
ViTAE	**0.3570**	**0.3434**
Kriging	0.8599	0.8195
Monitor Percent	0.13%	0.52%

CNN	0.3261	0.3220
ViTAE	**0.3129**	**0.3041**
Kriging	0.5520	0.4918
Monitor Percent	2.08%	4.68%

Our main models/approaches

data learning

OFFLINE: R&D

(CLEANING, TRAINING)

ONLINE: PRODUCTION

(ADJUSTING, RUNNING)

ACCURACY **(ERROR)**

Optimal Data Selection

Parameters Estimation

Data Augmentation

Data Assimilation

EFFICIENCY **(TIME)**

Surrogate models (training)

Data Driven models

Data Learning

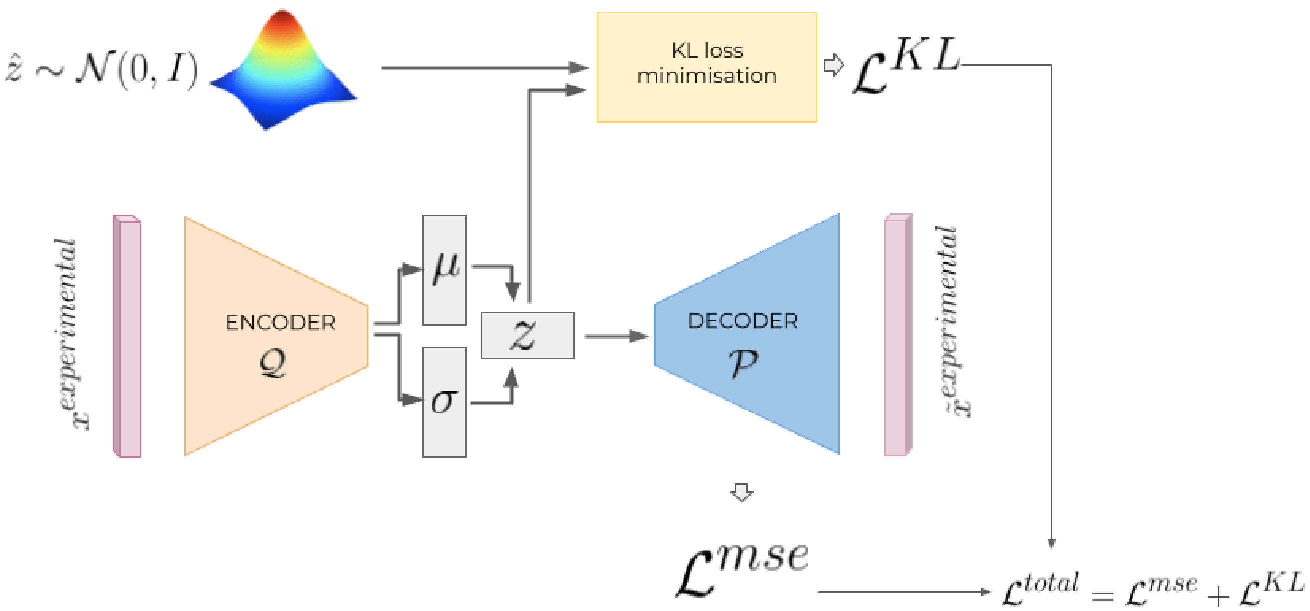
Surrogate models (forecasting)

PRE-PROCESS: Error Analysis, Error Distribution, Error Covariance

Decision-making

Synthetic experimental data generation using Variational Auto Encoders (VAEs)

Dr CQC



Synthetic experimental data generation using Variational Auto Encoders (VAEs)

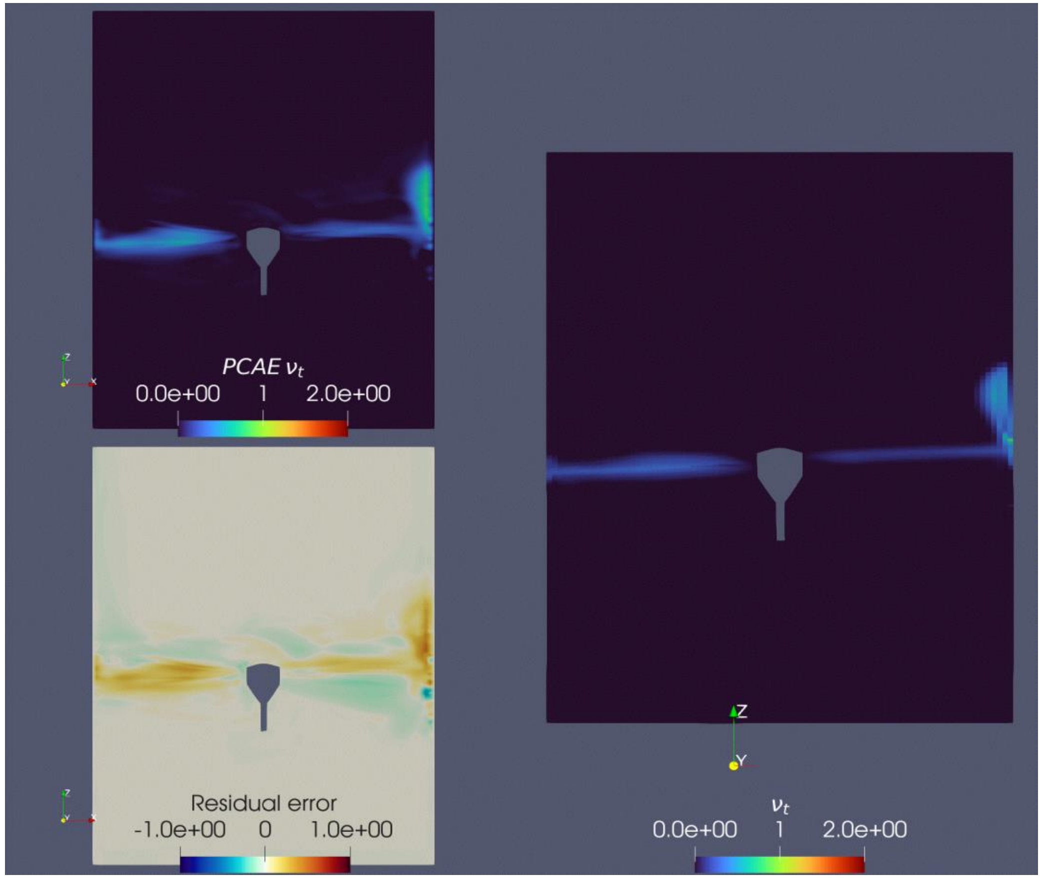

Dr CQC

The idea …
https://thispersondoesnotexist.com/

What if you want to generate 3D data with a physical meaning???

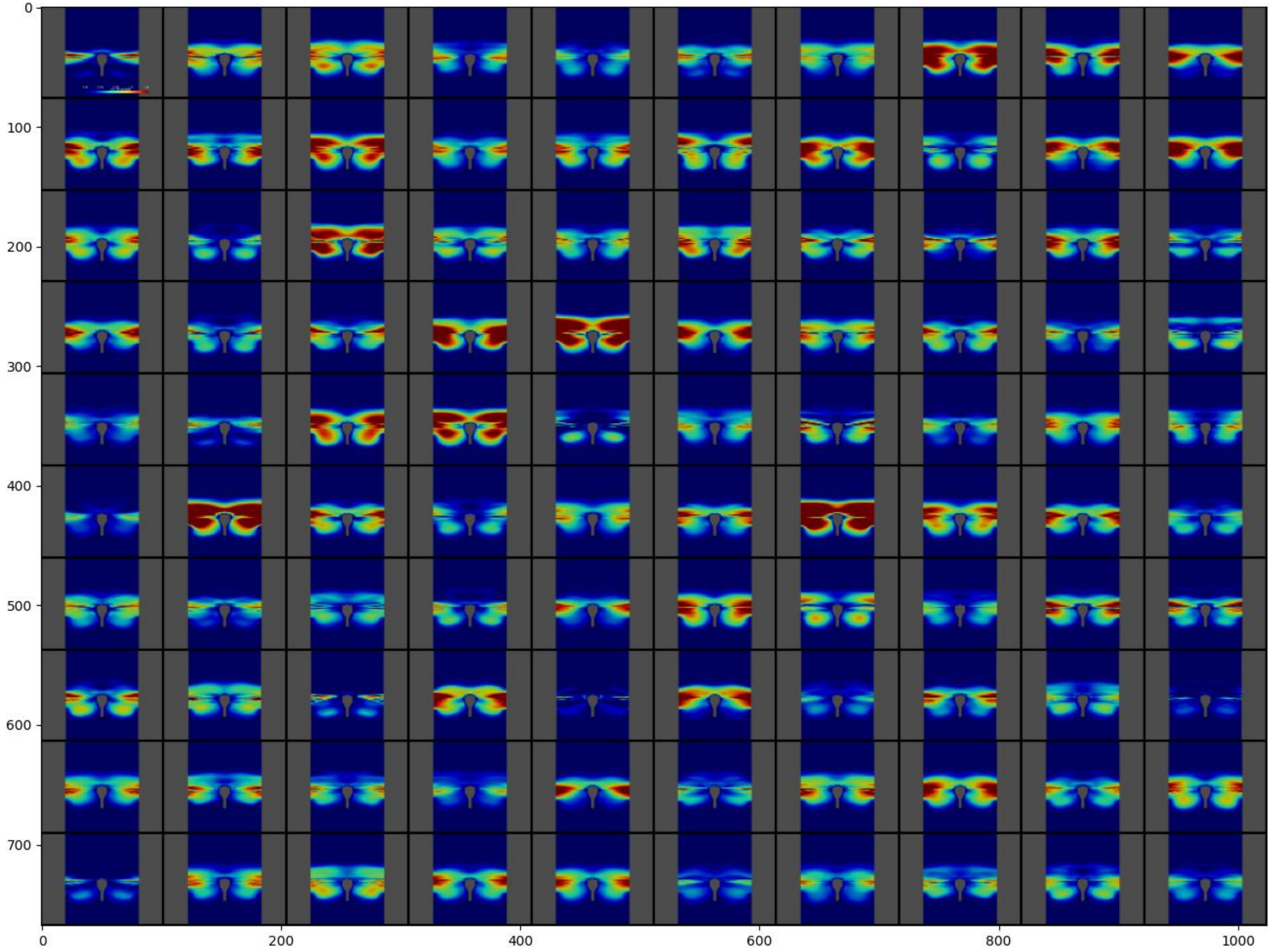
Test case: Synthetic data applied to Wave Energy converters

- Compression from 800k to 16 dimensions
- Right: GT
- Left: Prediction and residual error
- ~2 weeks to simulate



Synthetic experimental data generation using Variational Auto Encoders (VAEs)

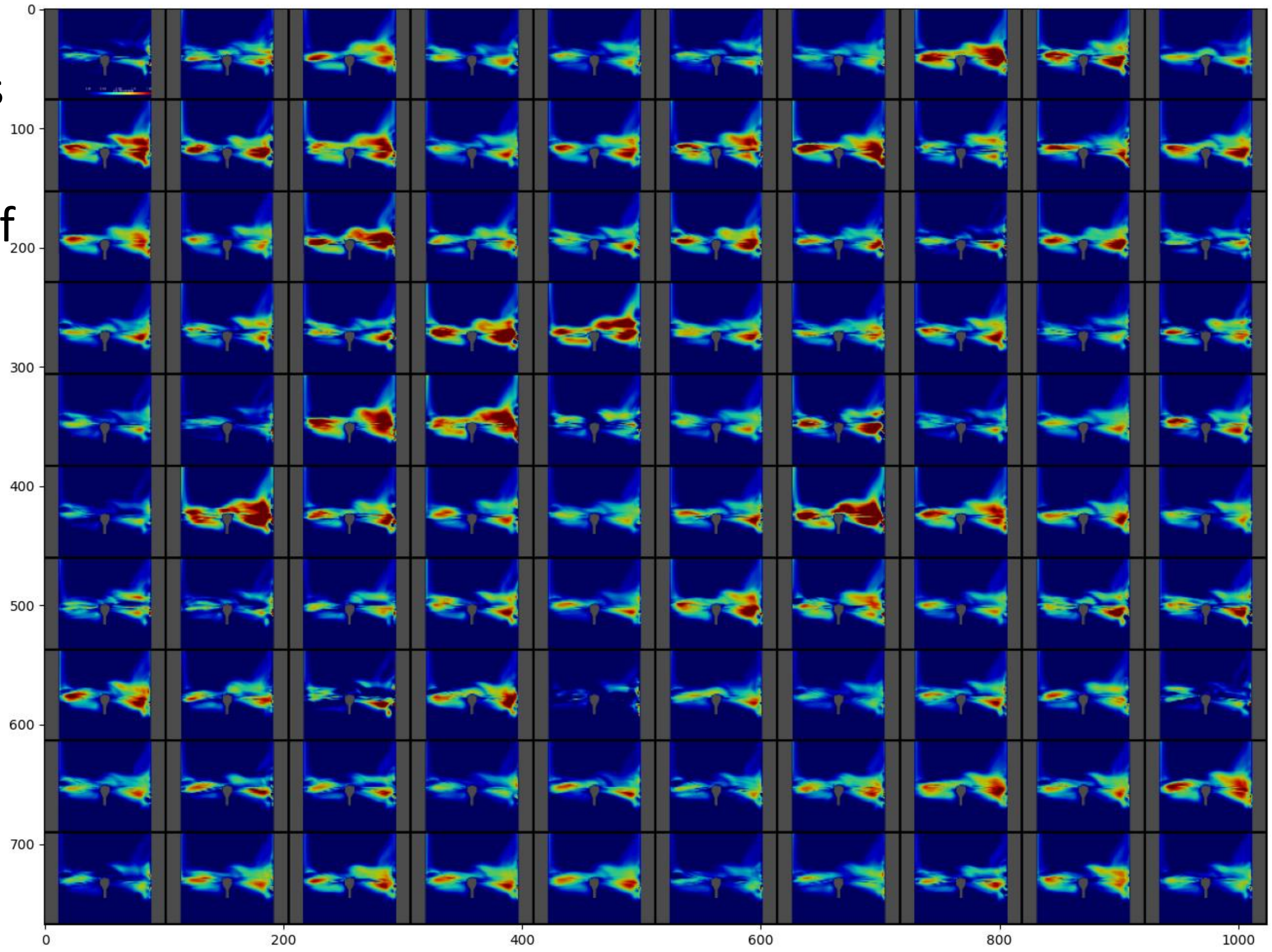
- Using an adversarial auto encoder
- 100 new samples of dynamic viscosity
- 0.05 [s] for 100 samples in PC-space
- 3.47 [s] for 100 samples in Physical Space
- YZ projection



Synthetic experimental data generation using Variational Auto Encoders (VAEs)

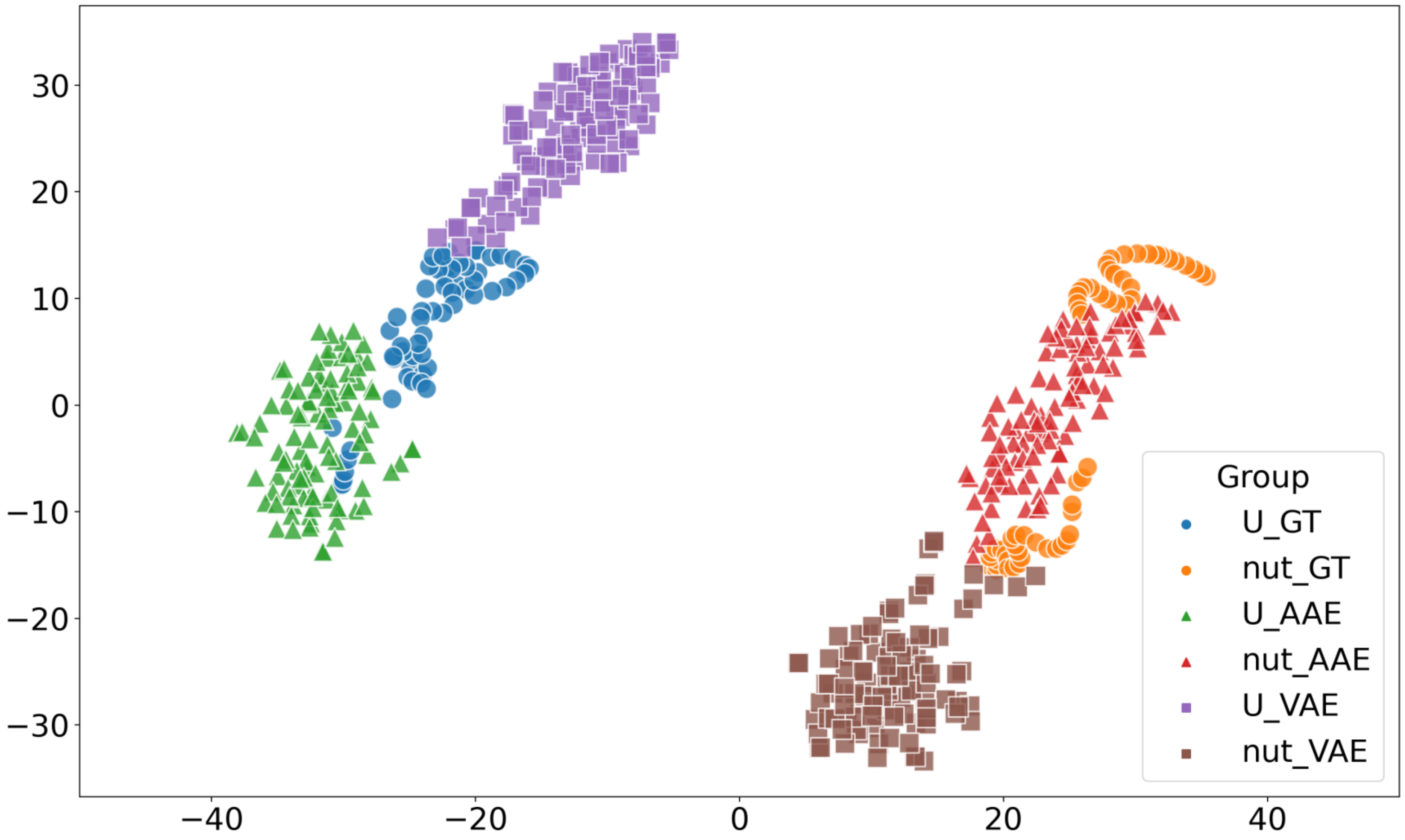
Dr COC

- Using VAE (same as case study 3)
- 100 new samples of dynamic viscosity
- XZ projection



Synthetic experimental data generation using Variational Auto Encoders (VAEs)

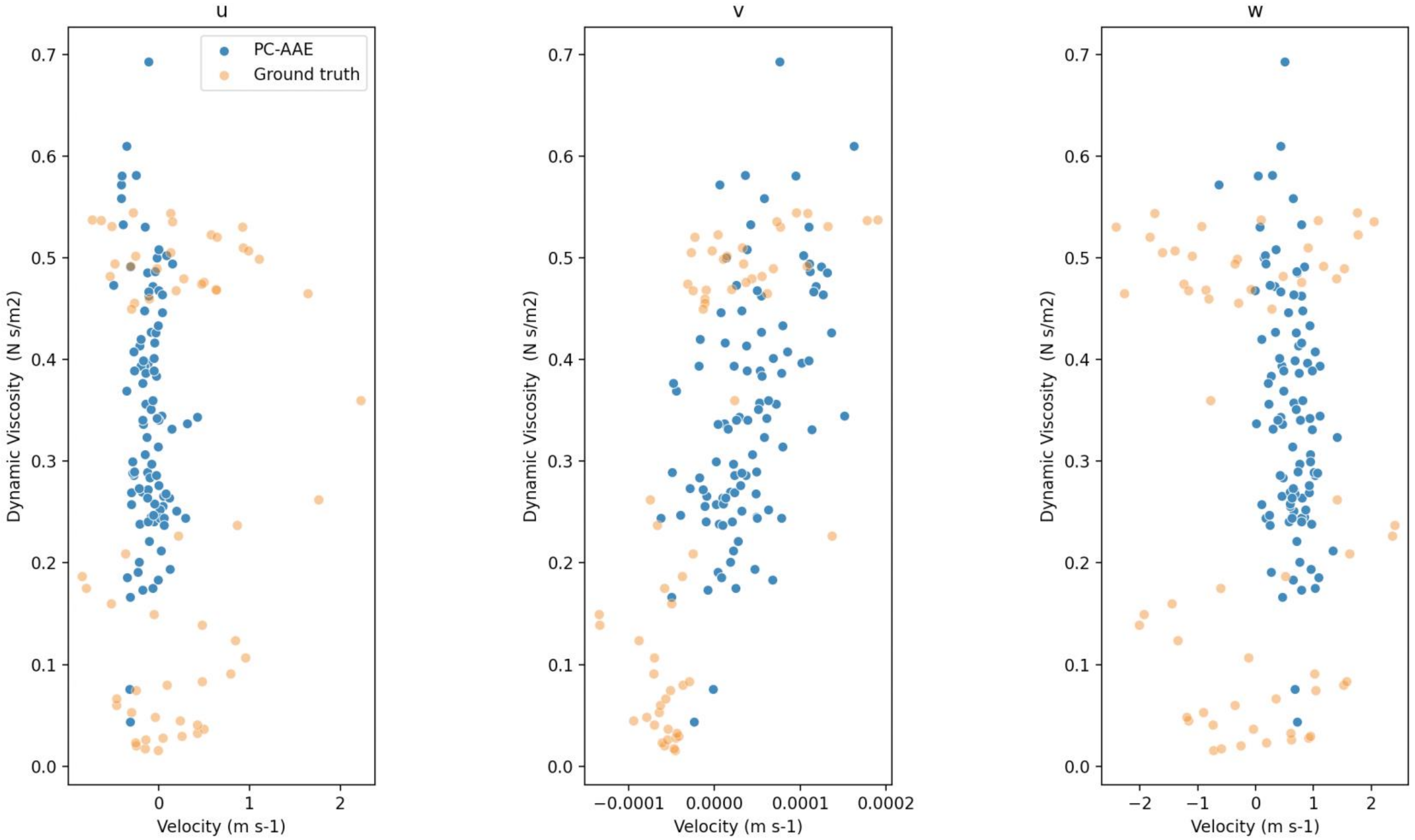
Dr CQC



Physical relation between U and dynamic viscosity

Dr CQC

Averaged over the number of nodes (850k)



Our main models/approaches

OFFLINE: R&D

(CLEANING, TRAINING)

ONLINE: PRODUCTION

(ADJUSTING, RUNNING)

ACCURACY **(ERROR)**

Optimal Data Selection

Parameters Estimation

Data Augmentation

Data Assimilation

EFFICIENCY **(TIME)**

Surrogate models (training)

Data Driven models

Data Learning

Surrogate models (forecasting)

PRE-PROCESS: Error Analysis, Error Distribution, Error Covariance

Decision-making

Data Assimilation or Latent Assimilation?

$$J(\mathbf{u}) = \alpha \|\mathbf{u} - \mathbf{u}_0\|^2_{\mathbf{B}^{-1}} + \|\mathbf{G}\mathbf{u} - \mathbf{v}\|^2_{\mathbf{R}^{-1}}$$

DA function

3DVar in the control space

$$\mathbf{w} = \mathbf{V}^{+}\delta\mathbf{u}$$
$$\mathbf{B} = \mathbf{V}\mathbf{V}^T$$

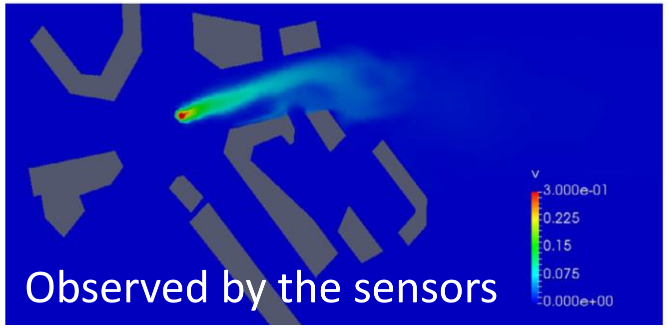
Reduced space, TSVD

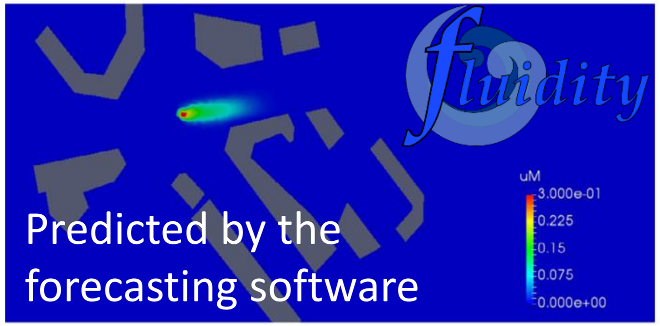
$$\mathbf{w}^{DA} = argmin_{\mathbf{w}\in\Re^{NP\times N}} J(\mathbf{w})$$

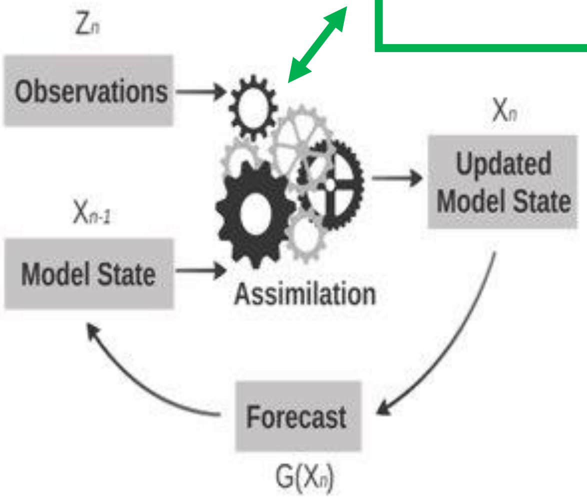
with

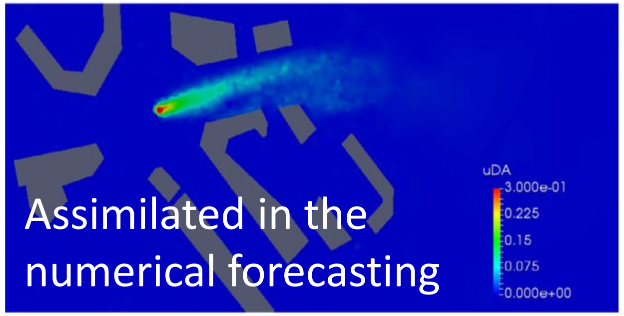
σ=μcδ

$$J(\mathbf{w}) = \frac{1}{2}\alpha\mathbf{w}^T\mathbf{w} + \frac{1}{2}(\mathbf{G}\mathbf{V}\mathbf{w} - \mathbf{d})^T\mathbf{R}^{-1}(\mathbf{G}\mathbf{V}\mathbf{w} - \mathbf{d})$$









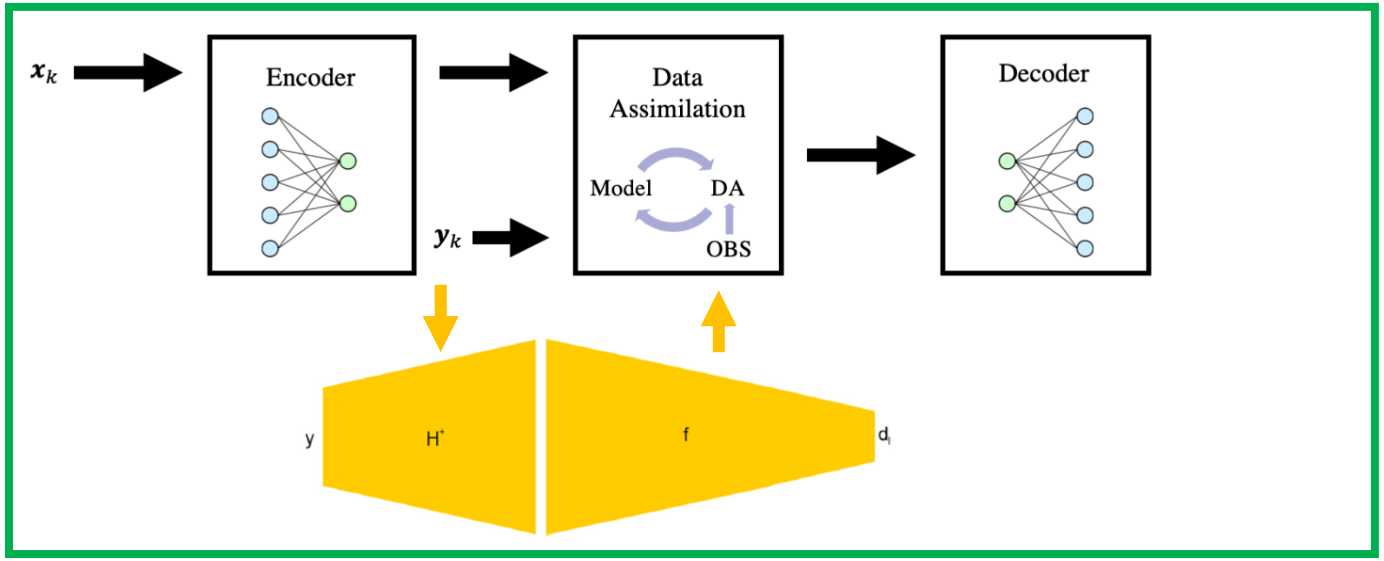
EPSRC

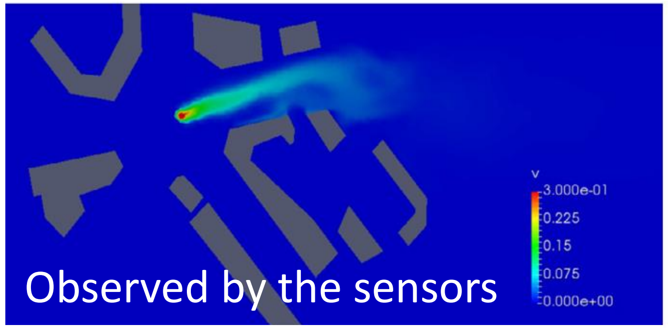
INHALE

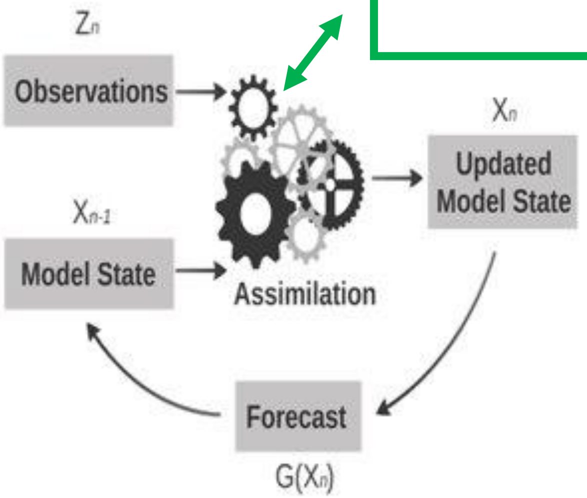
MAGIC
Envisaging a world with greener cities

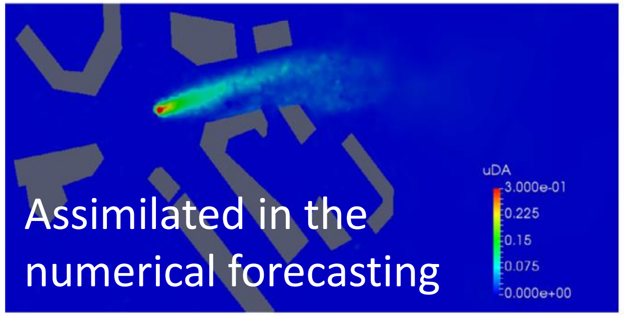
[*] R. Arcucci, L. Mottet, C. Pain and Y. Guo - **Optimal reduced space for Variational Data Assimilation** -Journal of Computational Physics
[**] R. Arcucci, C. Pain, Y. Guo, **Effective variational data assimilation in air-pollution prediction**, Big Data Mining and Analytics
[***] Mack, J., Arcucci, R., Molina-Solana, M., & Guo, Y. K. (2020). **Attention-based Convolutional Autoencoders for 3D-Variational Data Assimilation**. *Computer Methods in Applied Mechanics and Engineering*

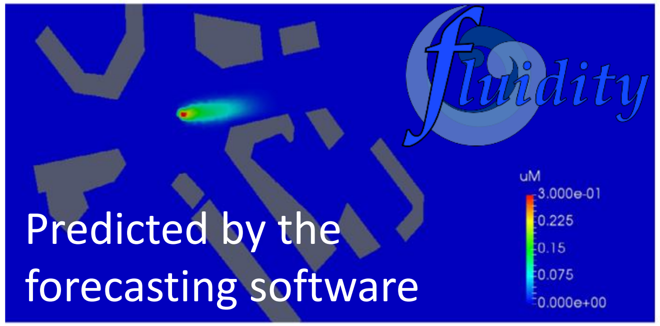
Data Assimilation or Latent Assimilation?











EPSRC

INHALE

MAGIC
Envisaging a world with greener cities

[*] R. Arcucci, L. Mottet, C. Pain and Y. Guo - **Optimal reduced space for Variational Data Assimilation** -Journal of Computational Physics

[**] R. Arcucci, C. Pain, Y. Guo, **Effective variational data assimilation in air-pollution prediction**, Big Data Mining and Analytics

[***] Mack, J., Arcucci, R., Molina-Solana, M., & Guo, Y. K. (2020). **Attention-based Convolutional Autoencoders for 3D-Variational Data Assimilation**. *Computer Methods in Applied Mechanics and Engineering*

Data Assimilation in a latent space

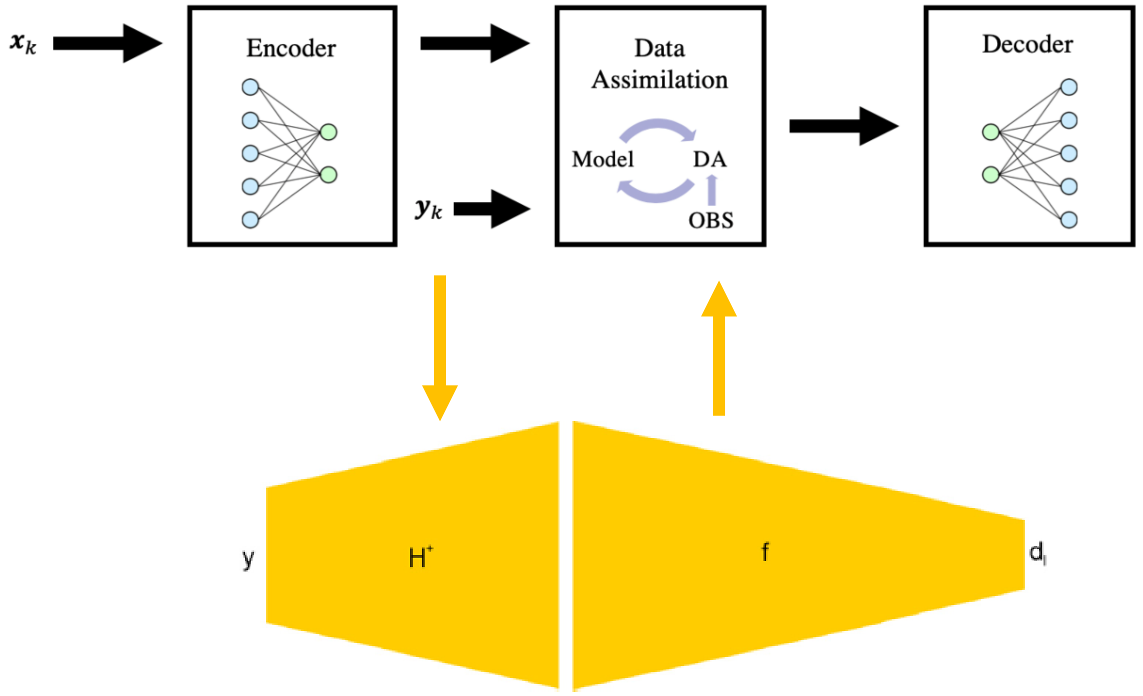
***with Julian Mack - 2019**

3DVar in the latent space

$$\mathbf{w}_l^{DA} = \arg\min_{\mathbf{w}_l} J(\mathbf{w}_l)$$

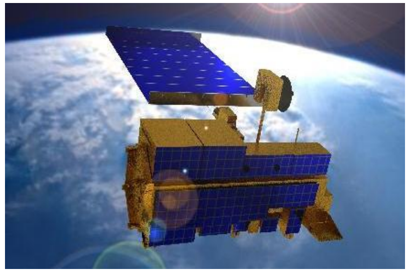
$$J(\mathbf{w}_l) = \frac{1}{2}\mathbf{w}_l^T \mathbf{w}_l + \frac{1}{2}\|\boldsymbol{d}_l - \boldsymbol{V}_l \mathbf{w}_l\|^2_{\boldsymbol{R}_l^{-1}}$$

Model	MSE	Execution Time (s)
Ref MSE	1.0001	-
PCA, $\nu = 32$, $m = n$	0.1270	1.8597
PCA, $\nu = 32$, $m = 0.1n$	0.1270	0.2627
PCA, $\nu = 32$, $m = 0.01n$	0.1334	0.0443
PCA, $\nu = 32$, $m = 0.001n$	0.1680	0.0390
Data Learning with Tucodec-NeXt	0.0787	0.0537

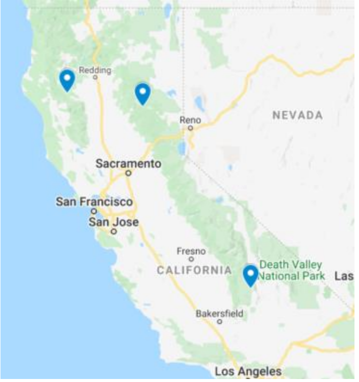


[*] Mack, J., Arcucci, R., Molina-Solana, M., & Guo, Y. K. (2020). **Attention-based Convolutional Autoencoders for 3D-Variational Data Assimilation.** *Computer Methods in Applied Mechanics and Engineering*, *372*, 113291.

Wildfire forecasting

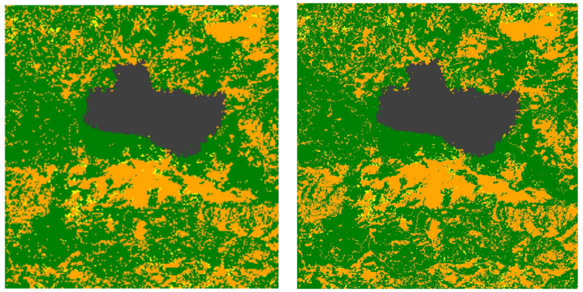


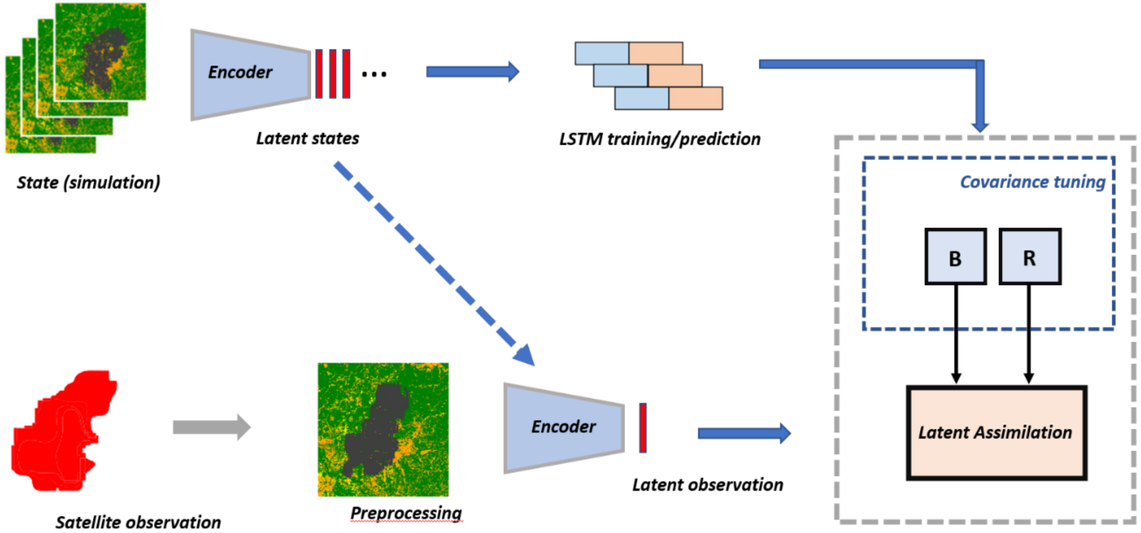
MODIS: every 1-2 days at 1km resolution

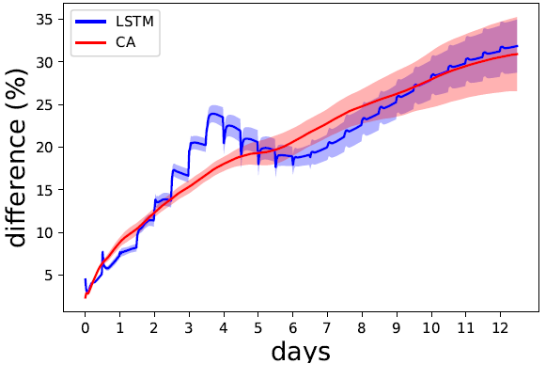


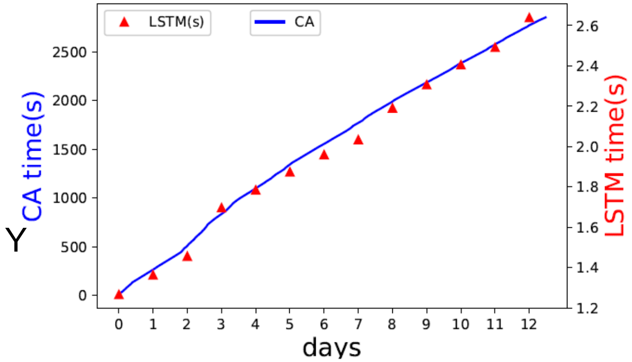
- Learning from simulation data
- Using satellite observations to validate/assimilate

observation prediction





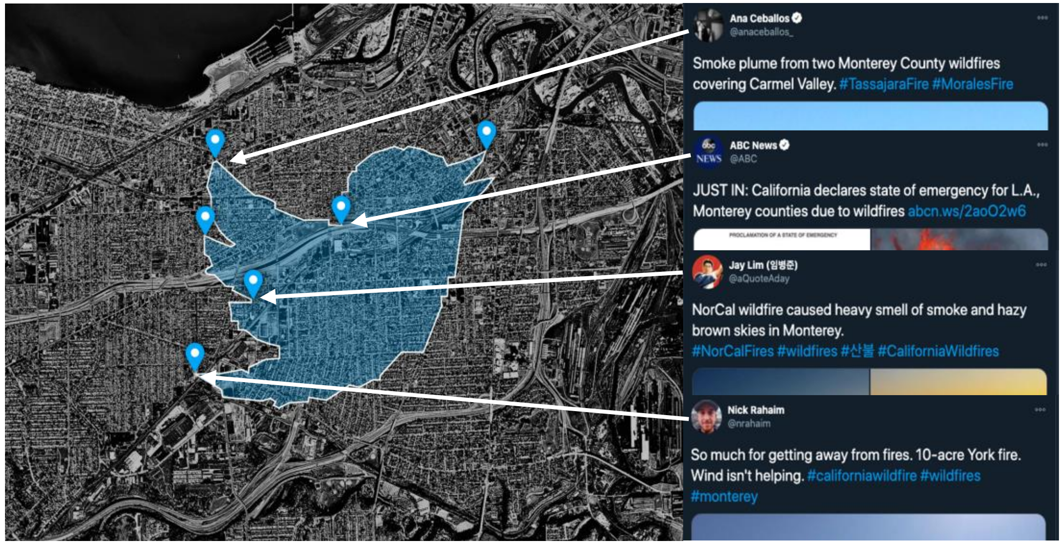




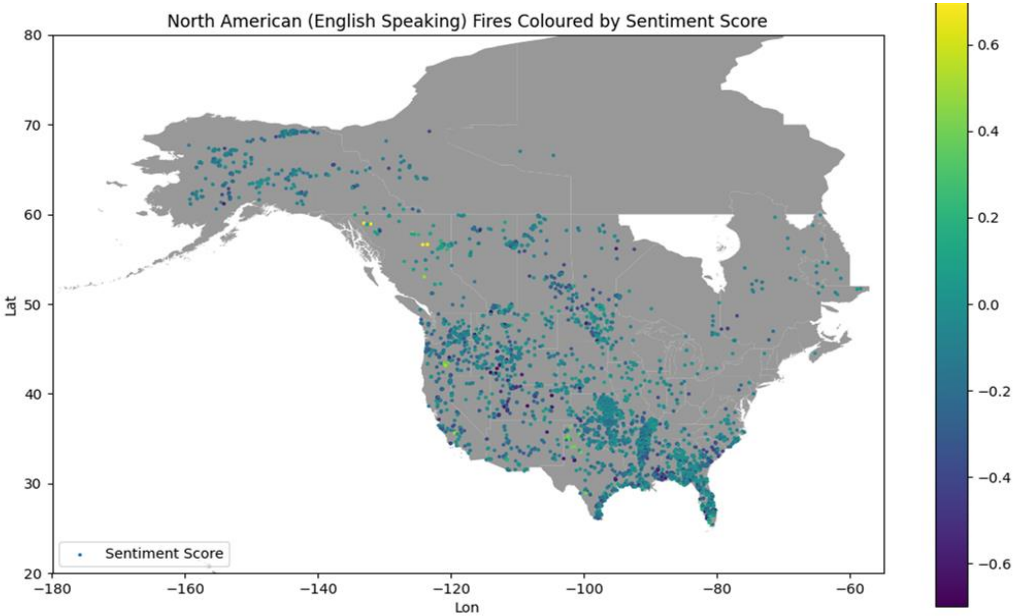
[*] Data-driven surrogate model with latent data assimilation: Application to wildfire forecastingS Cheng, IC Prentice, Y Huang, Y Jin, YK Guo, R Arcucci - Journal of Computational Physics, 111302

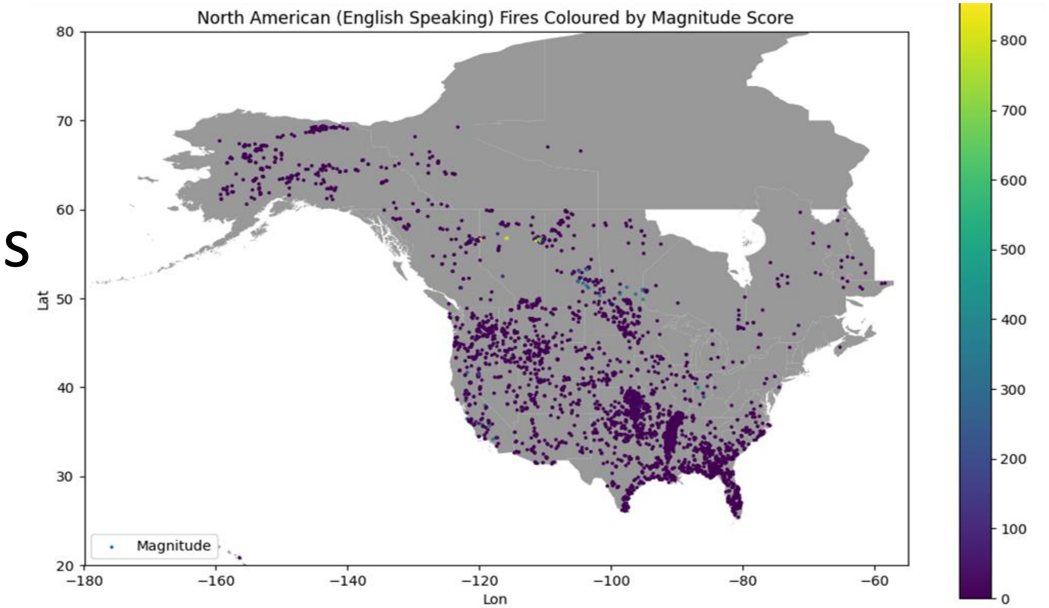
[**] Parameter Flexible Wildfire Prediction Using Machine Learning Techniques: Forward and Inverse ModellingS Cheng, Y Jin, SP Harrison, C Quilodrán-Casas, IC Prentice, YK Guo, ..., R.Arcucci - Remote Sensing 14 (13), 3228

Social data and satellites: Wildfires System



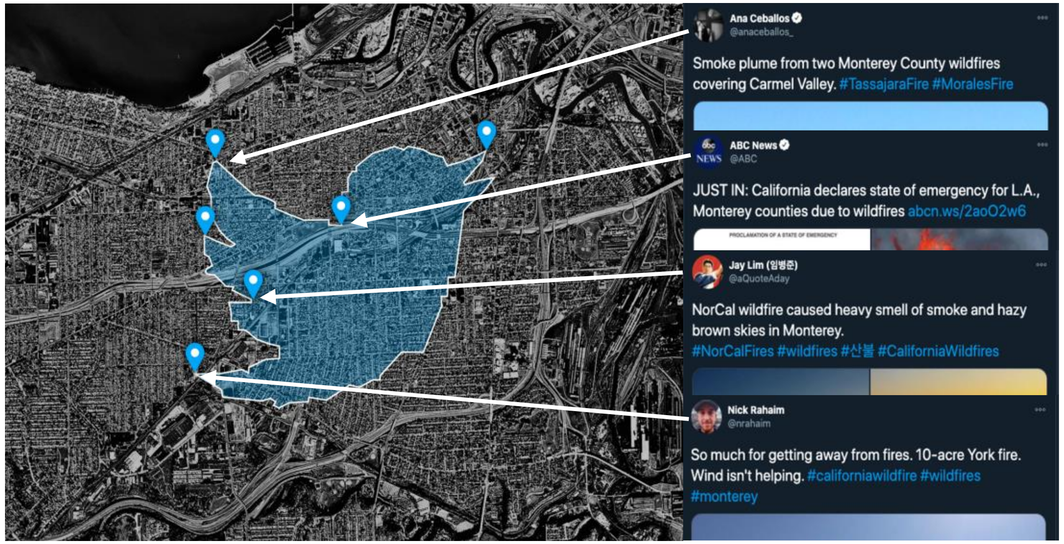
Jake Lever



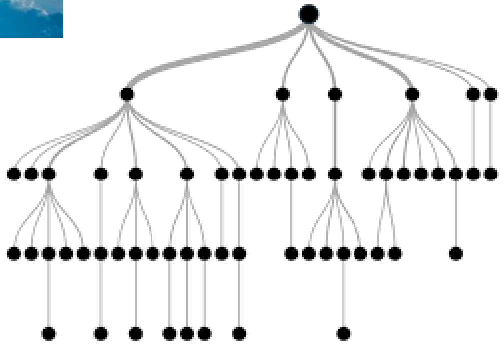


- - How do people perceive wildfires? Can this be measured or modelled?
- - Idea: can we collect many subjective opinions on certain natural events, and are these opinions reflective of the size and severity of the event?
- - Social media and Twitter - human sensors; Sentiment analysis - Converting emotional leaning in a passage of text into a numerical value, evaluating the positivity (or negativity) of

Predictive data driven model based on HUMAN SENSORS



Jake Lever

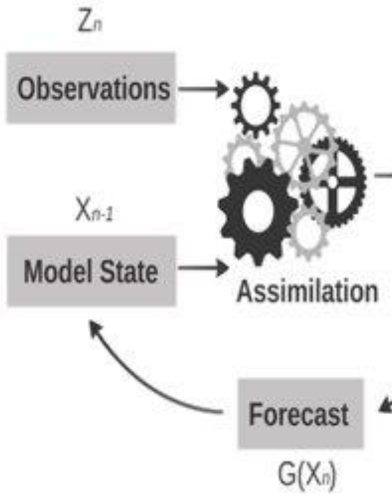


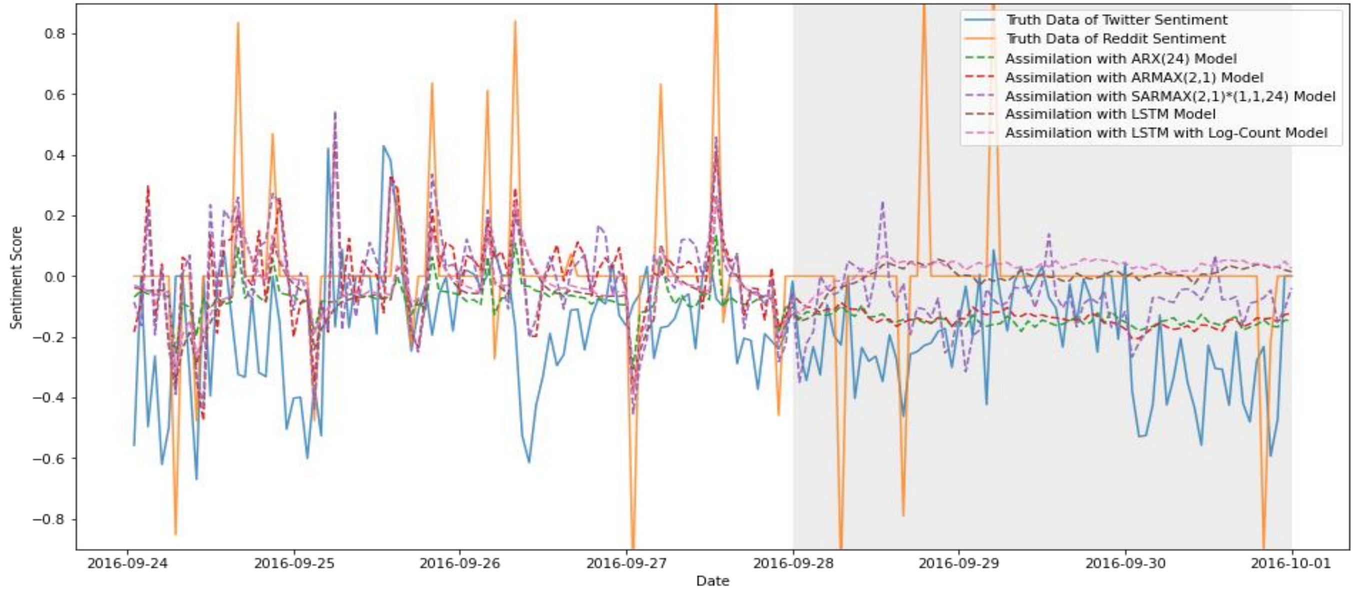
- Historical Wildfire Data - global fire atlas 2016 ignitions.
- Used full archive search of Twitter to find tweets relevant to individual wildfire events
- Results show predictive power for predicting some physical wildfire variables from social sentiment

Variable	MAE
Latitude	3.958
Longitude	6.661
Size	6.29
Perimeter	5.19
Duration	0.51
Speed	0.38
Expansion	0.52
Population Density	92.56

[*] Sentimental wildfire: a social-physics machine learning model for wildfire nowcasting , J Lever, R Arcucci, Journal of Computational Social Science, 1-39

Social data and satellites: Wildfires System





[*] Social Data Assimilation of Human Sensor Networks for Wildfires
J Lever, R Arcucci, J Cai - Proceedings of the 15th International Conference PETRA
[**] Sentimental wildfire: a social-physics machine learning model for wildfire nowcasting
J Lever, R Arcucci - Journal of Computational Social Science, 1-39

Our main models/approaches

data learning

ACCURACY **(ERROR)**

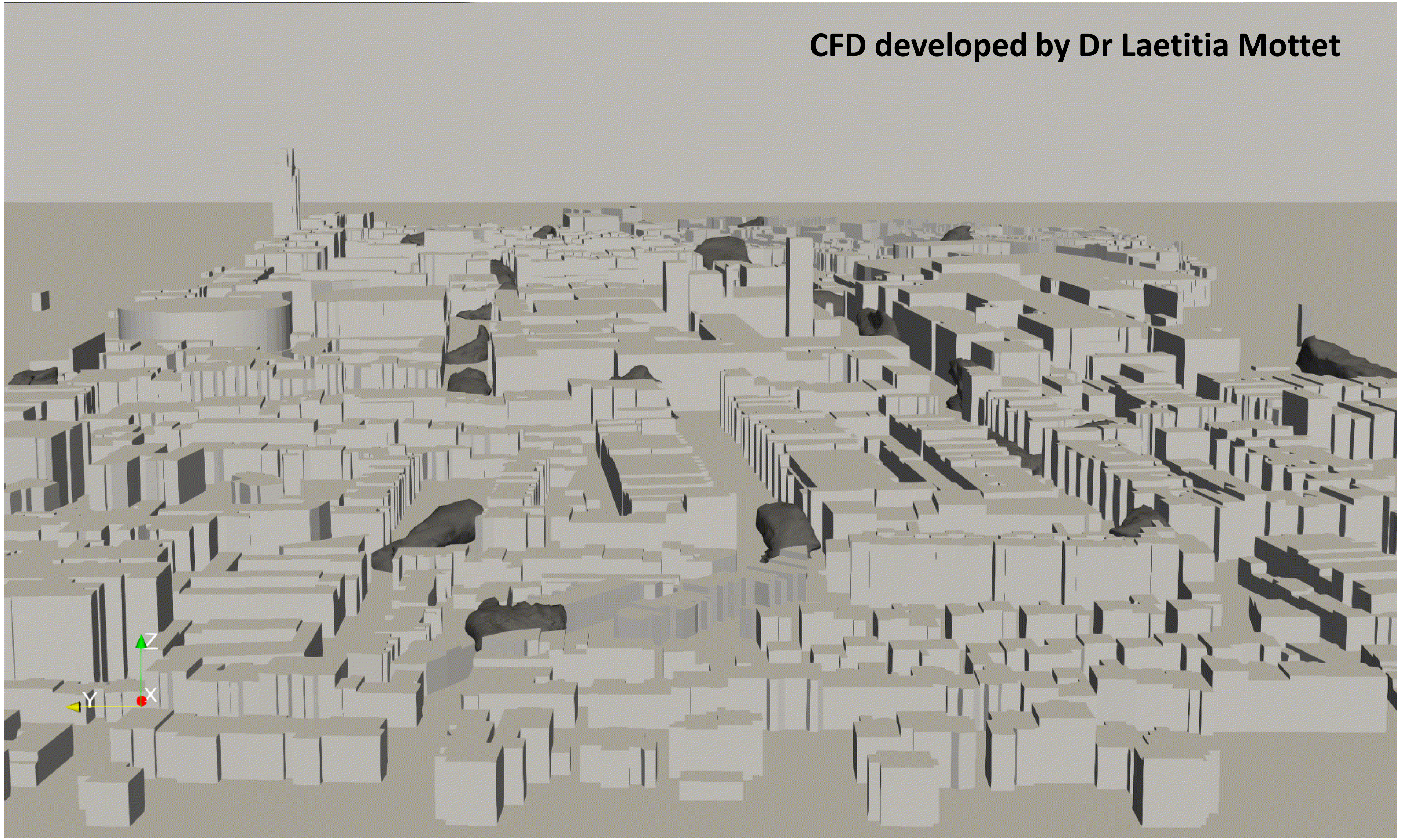
EFFICIENCY **(TIME)**

	OFFLINE: R&D **(CLEANING, TRAINING)**	**ONLINE: PRODUCTION** **(ADJUSTING, RUNNING)**
ACCURACY (ERROR)	Optimal Data Selection Parameters Estimation Data Augmentation	Data Assimilation
EFFICIENCY (TIME)	Surrogate models (training) Data Driven models	Data Learning Surrogate models (forecasting)

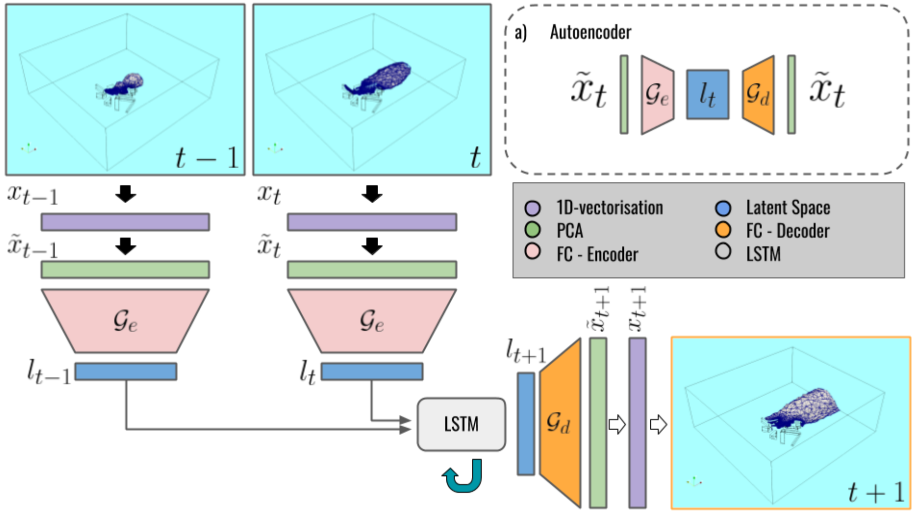
PRE-PROCESS: Error Analysis, Error Distribution, Error Covariance

Decision-making

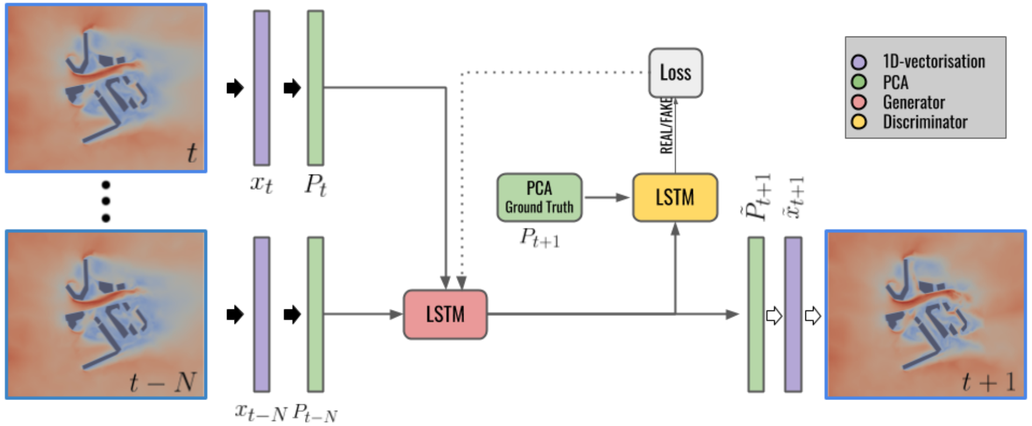
Surrogate Models: fast ML models to emulate CFD simulations



Surrogate Models: fast ML models to emulate CFD simulations



[*]

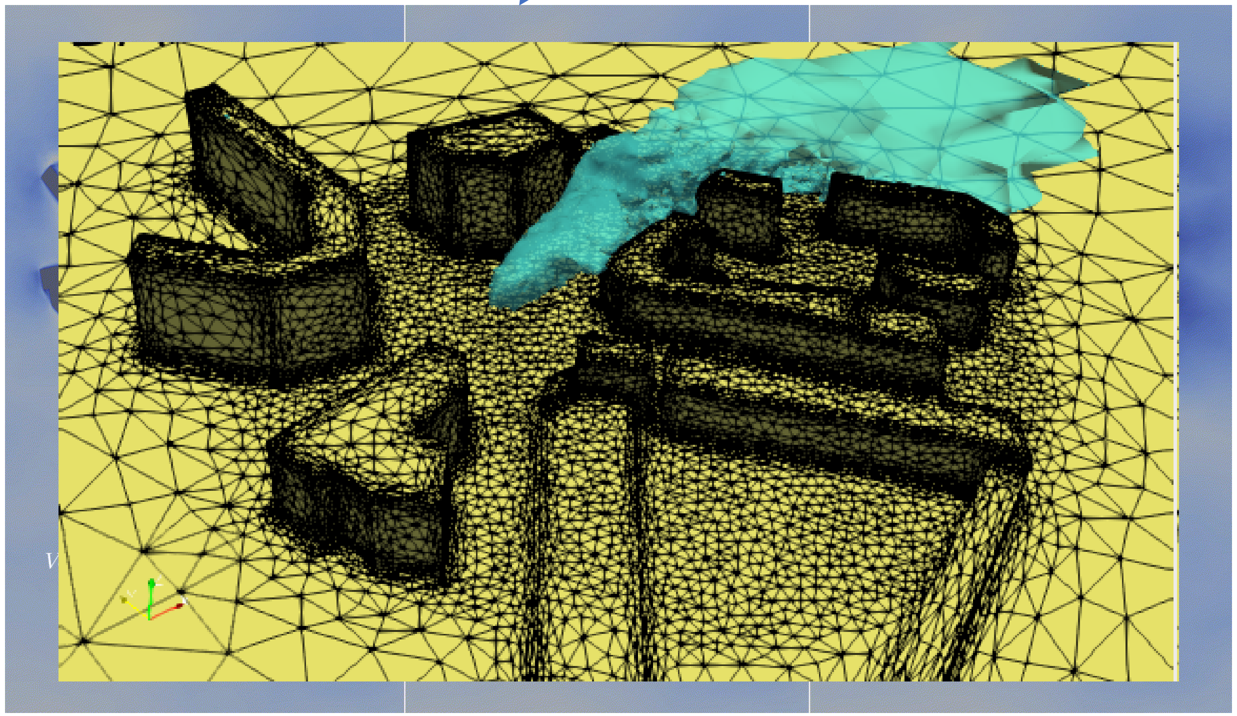


[**]

CFD simulation

previous surrogate models after few free time steps

OUR surrogate models after few free time steps



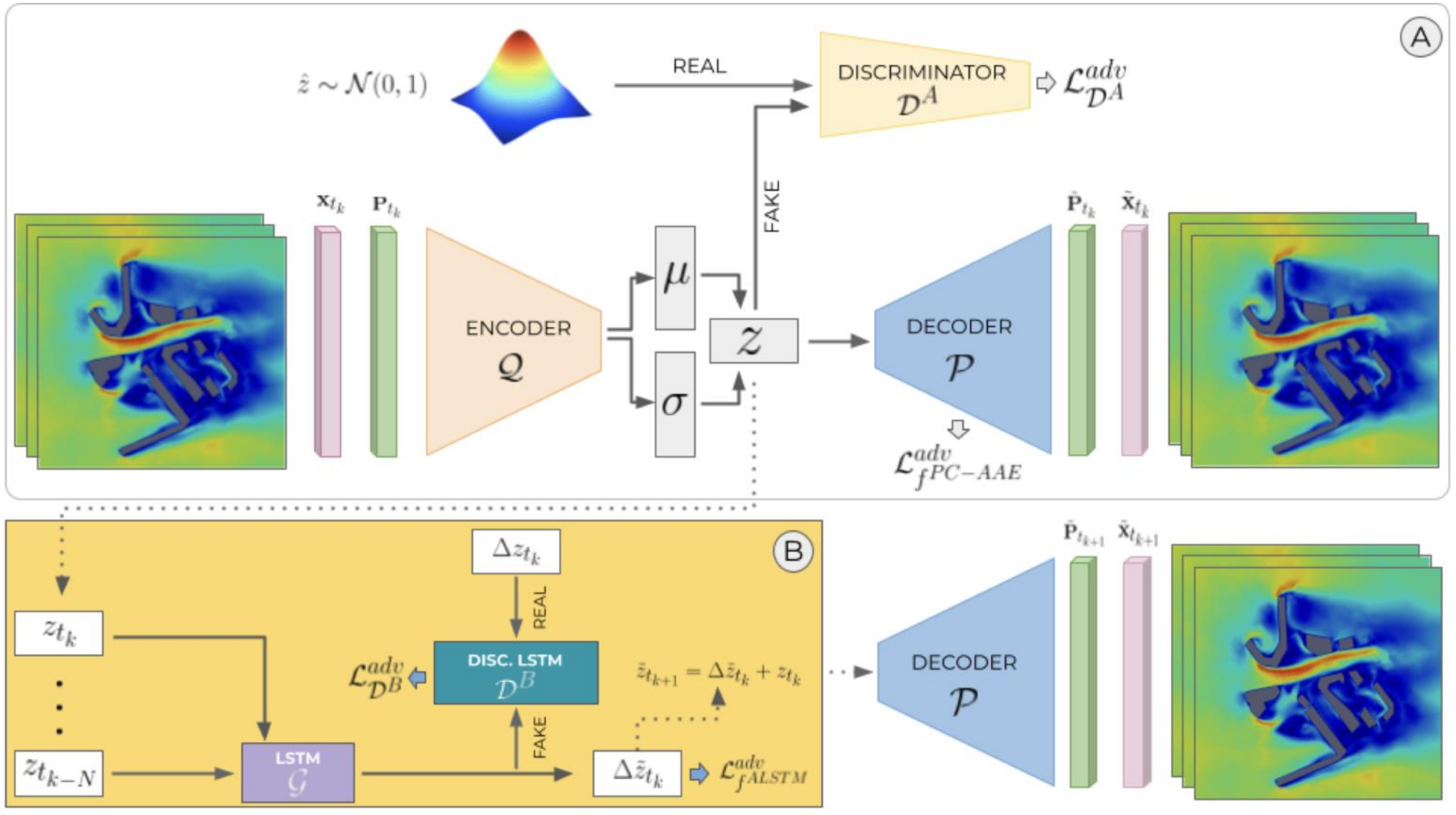
Unstructured mesh

The forecasts are $O(10^4)$ faster than the CFD simulation

[*] C. Quilodran Casas, R. Arcucci, Y. Guo - **Urban Air Pollution Forecasts Generated from Latent Space Representations**

[**] C. Quilodran Casas, R. Arcucci, C. Pain, Y. Guo - **Adversarially trained LSTMs on reduced ordermodels of urban air pollution simulations.**

Data Driven Surrogate Models

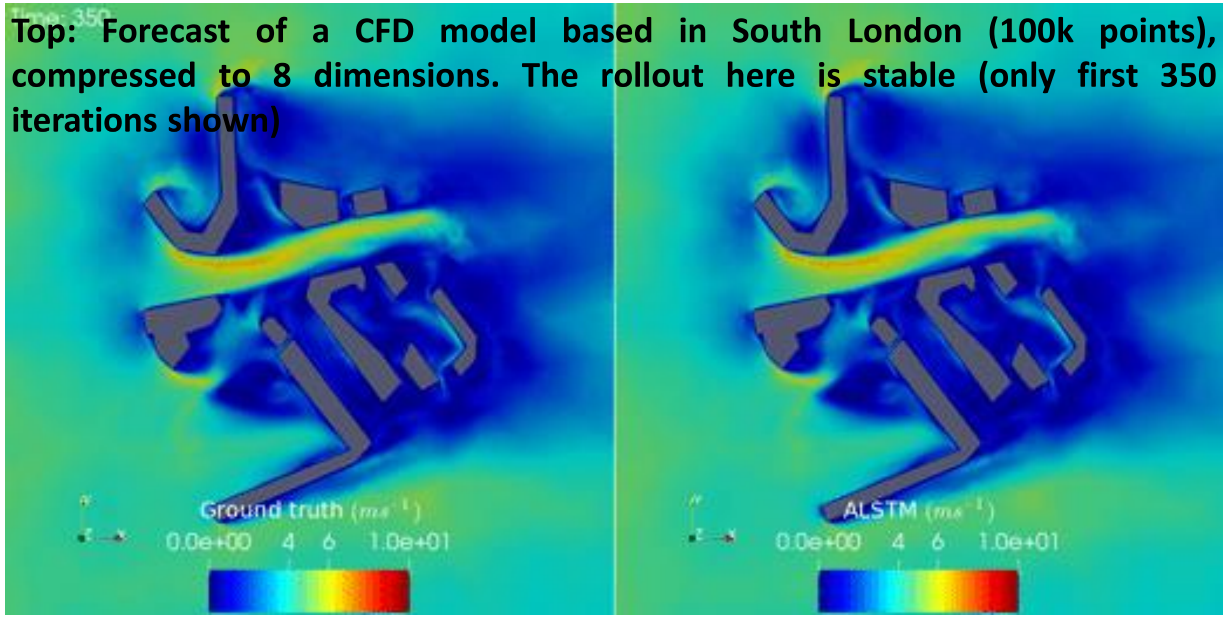


Network A: Does a compression of the model by reducing the number of dimensions of the Principal Components. This is an adversarial AE which maps the latent space into a Gaussian distribution

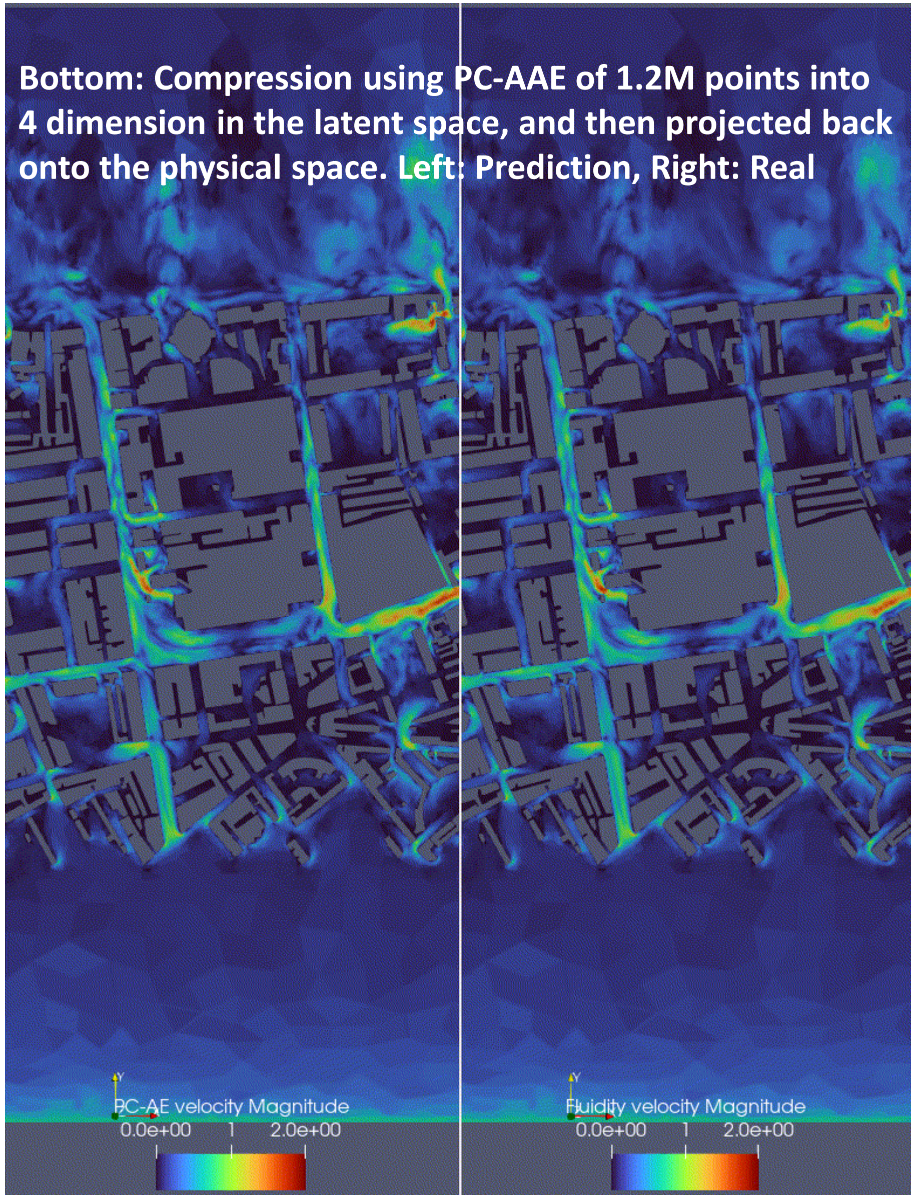
Network B: Uses an adversarial LSTM to forecast the Gaussian latent space, this makes the forecasts more robust as they stay within the data distribution and improves the rollout.

[*] Quilodrán-Casas, C., Arcucci, R., Mottet, L., Guo, Y., & Pain, C. (2021). Adversarial autoencoders and adversarial LSTM for improved forecasts of urban air pollution simulations. *arXiv preprint arXiv:2104.06297*.Work presented at ICLR SimDL 2021

Some applications

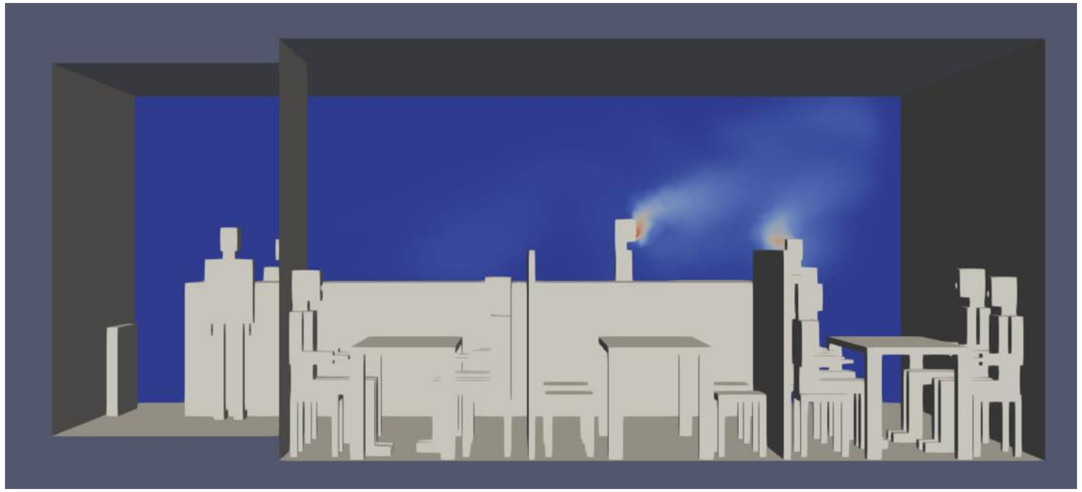


Top: Forecast of a CFD model based in South London (100k points), compressed to 8 dimensions. The rollout here is stable (only first 350 iterations shown)



Bottom: Compression using PC-AAE of 1.2M points into 4 dimension in the latent space, and then projected back onto the physical space. Left: Prediction, Right: Real

Covid risk assessment, air flow, PUB simulation



Surrogate modelling for global fire risk modelling for unseen scenarios

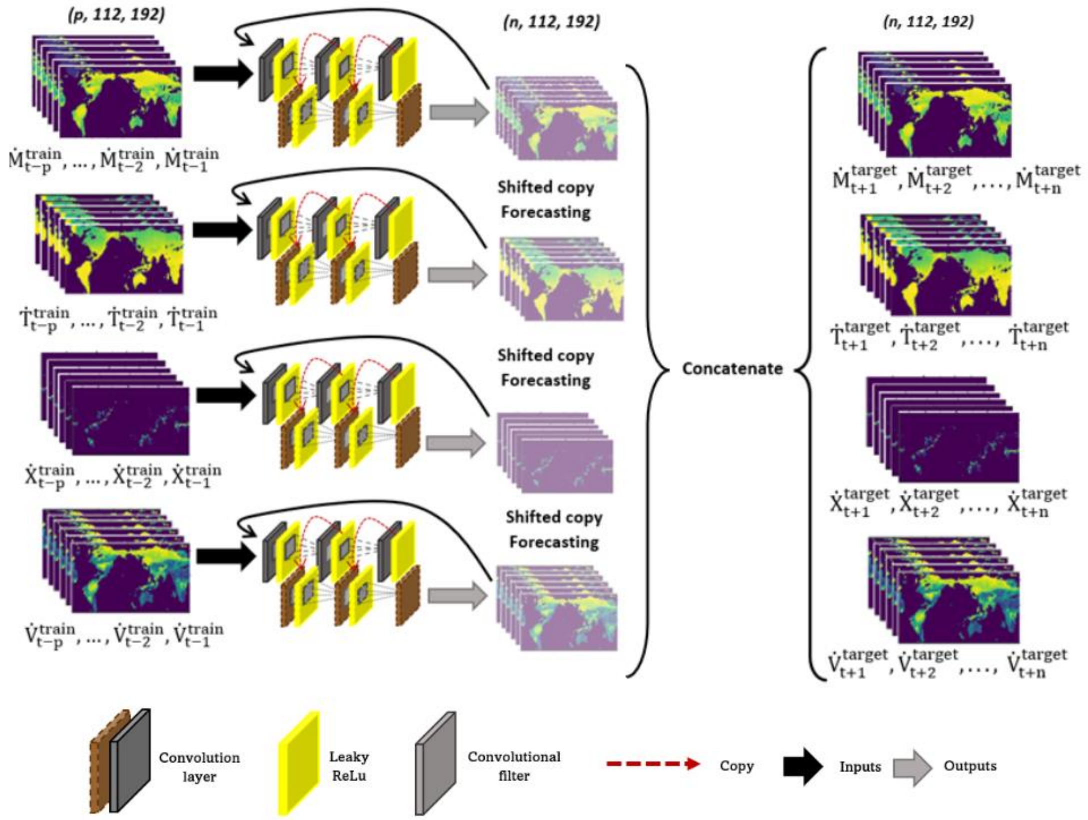
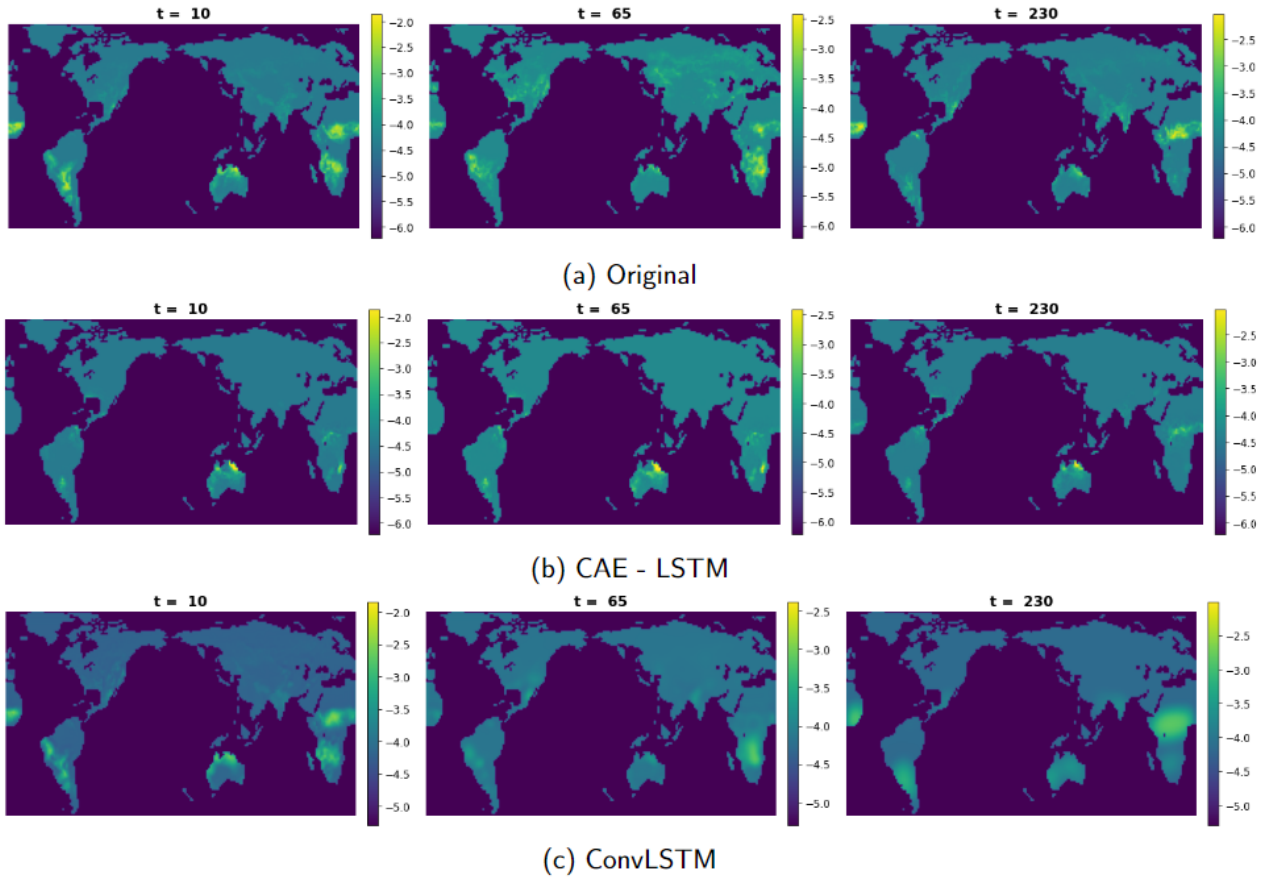


Figure 2: *Joint* ConvLSTM's architecture



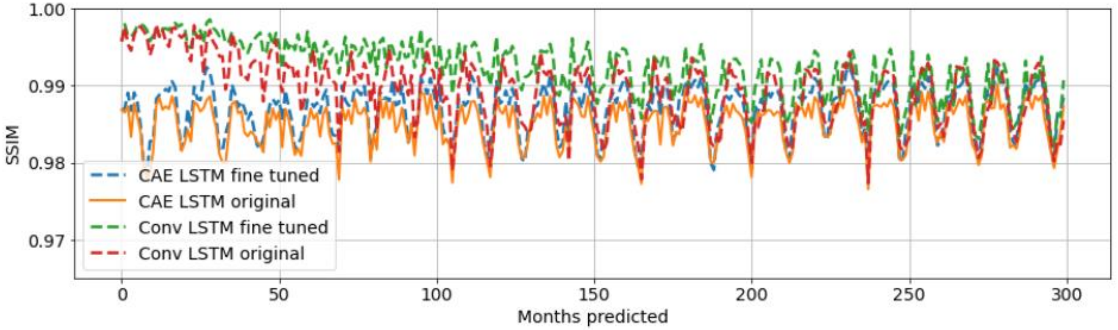


Figure 6: SSIM of each forecast for the best models before and after fine tuning

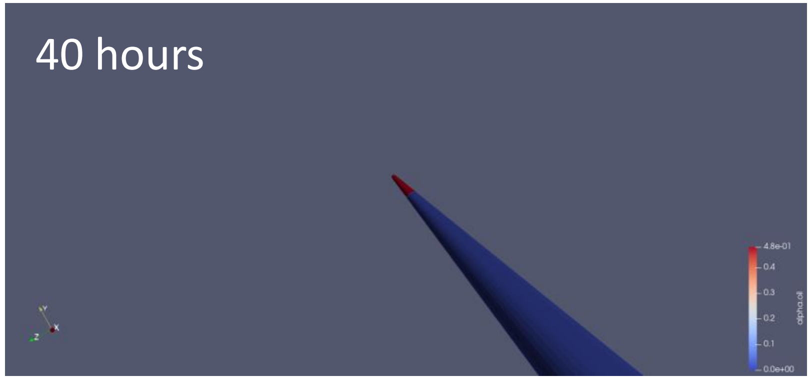
Surrogate Models: Flow pattern transition of two-phase flow in pipes

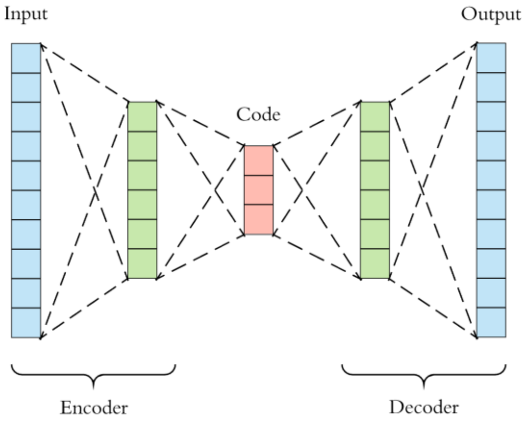
In collaboration with:

UCL

- 180,000 nodes
- Around 40 hours for 1 CFD simulation

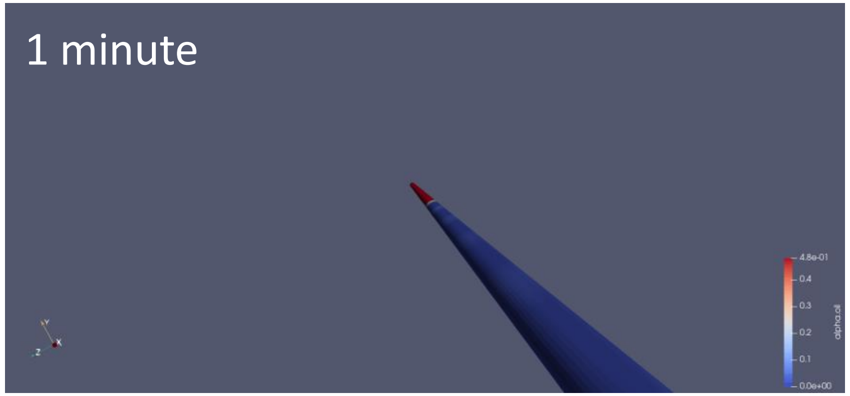
CFD





180,000 to 30

Latent prediction



Compare different approaches of auto-encoder + LSTM

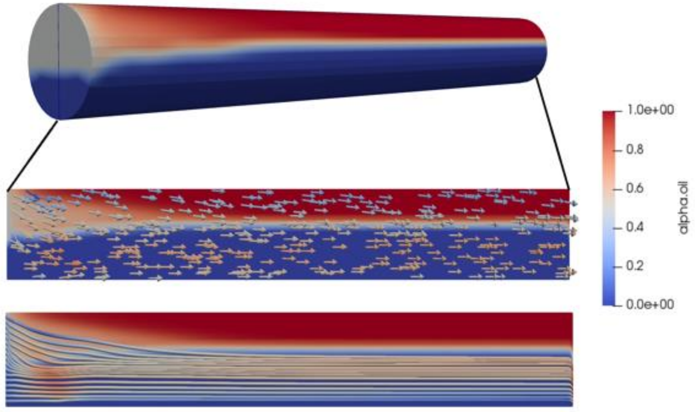
POD

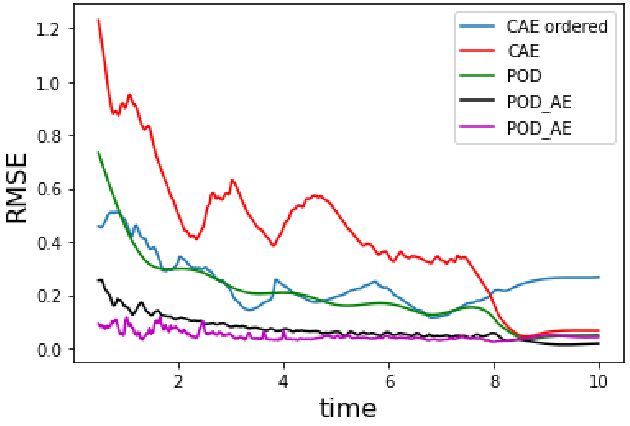
CAE

Ordered CAE

POD AE

GCN





Our main models/approaches

data learning

ACCURACY **(ERROR)**

EFFICIENCY **(TIME)**

	OFFLINE: R&D **(CLEANING, TRAINING)**	**ONLINE: PRODUCTION** **(ADJUSTING, RUNNING)**
ACCURACY	Optimal Data Selection Parameters Estimation Data Augmentation	Data Assimilation
EFFICIENCY	Surrogate models (training) Data Driven models	Data Learning Surrogate models (forecasting)

PRE-PROCESS: Error Analysis, Error Distribution, Error Covariance

Decision-making

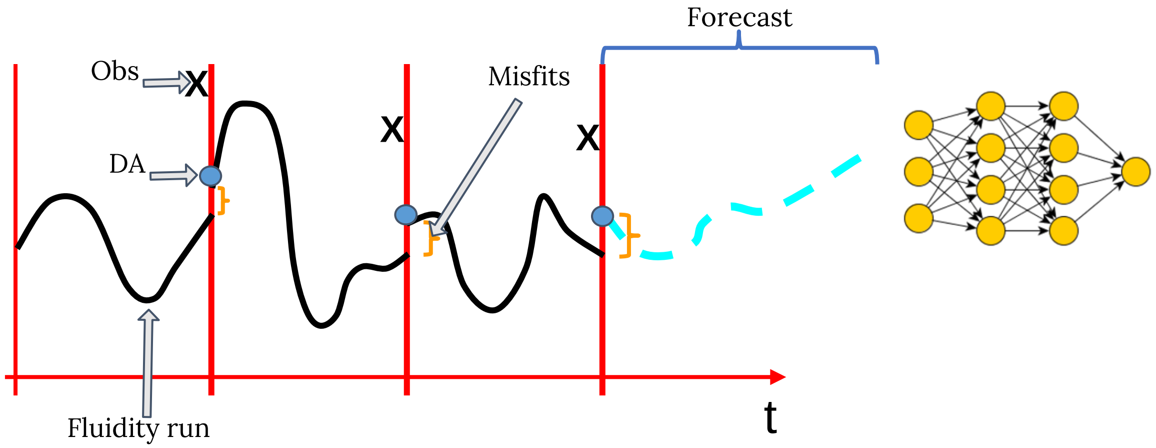
Data Learning to reduce the errors in the solution of existing systems

having benefit from AI without changing your existing system

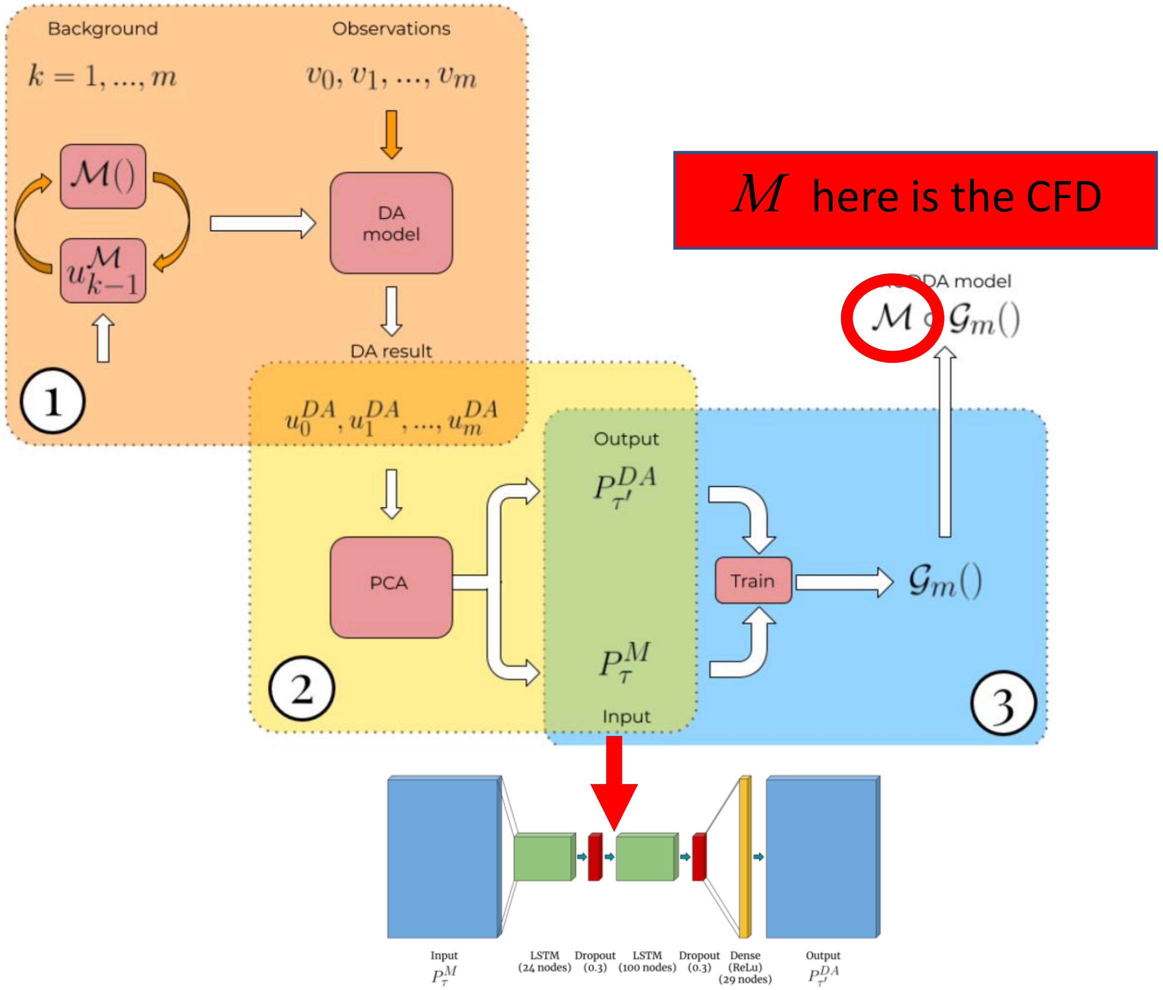
What if the observations are not available?

DDA … learning the Data Assimilation process

The idea:
Data Assimilation at each time step give us a misfit (DA - fluidity background),
the saved misfits are trained using a Long short-term memory (LSTM) network
and used for future forecasts.

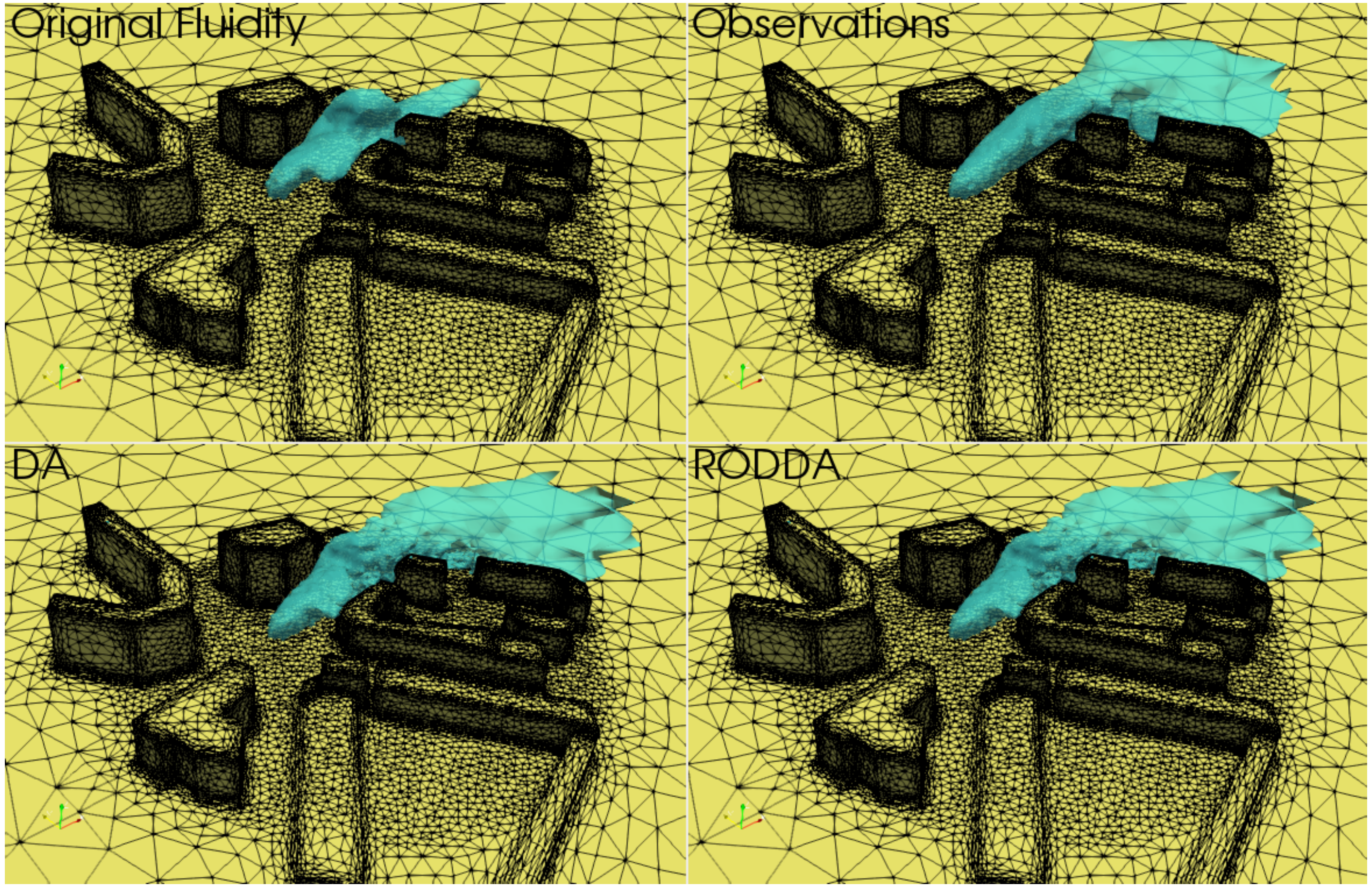


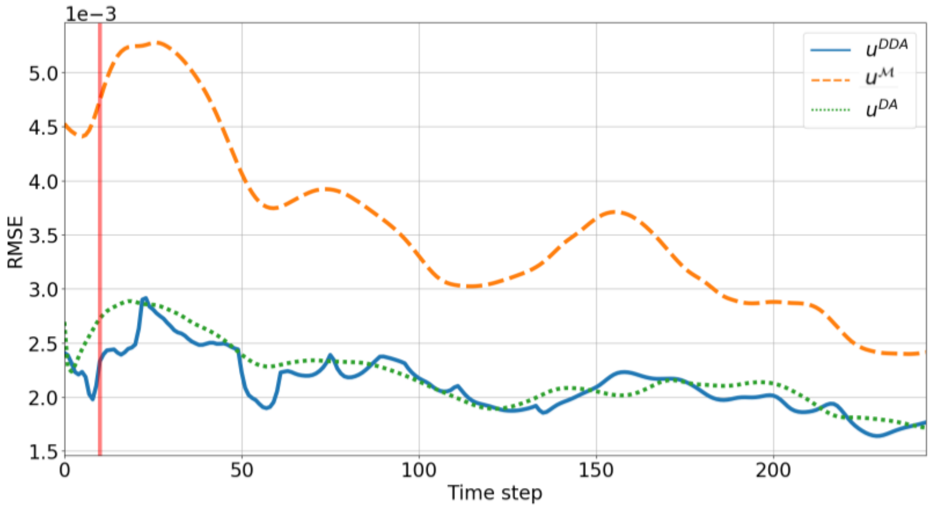
Reduced Order Deep Data Assimilation (RODDA)



Data Learning to reduce the errors in the solution of existing systems

having benefit from AI without changing your existing system





Same accuracy but RODDA is 1000 times faster than DA

[*] R. Arcucci, J. Zhu, S. Hu, YK Guo, **Deep data assimilation: Integrating deep learning with data assimilation** - Applied Sciences

[**] C. Quilodran Casas, R. Arcucci**,** P. Wu, C. Pain, Y. Guo - **A Reduced Order Deep Data Assimilation model** – Physica D: nonlinear phenomena

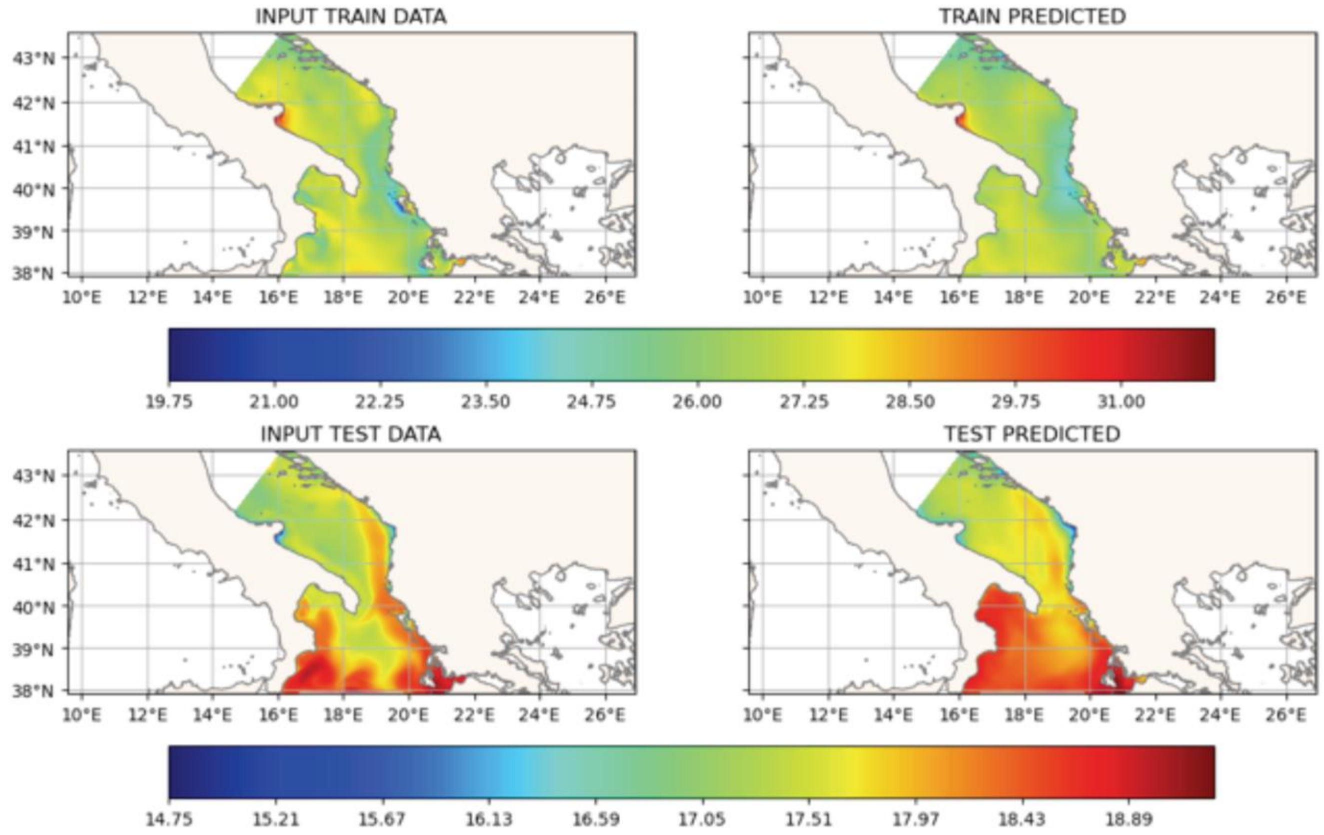
data learning

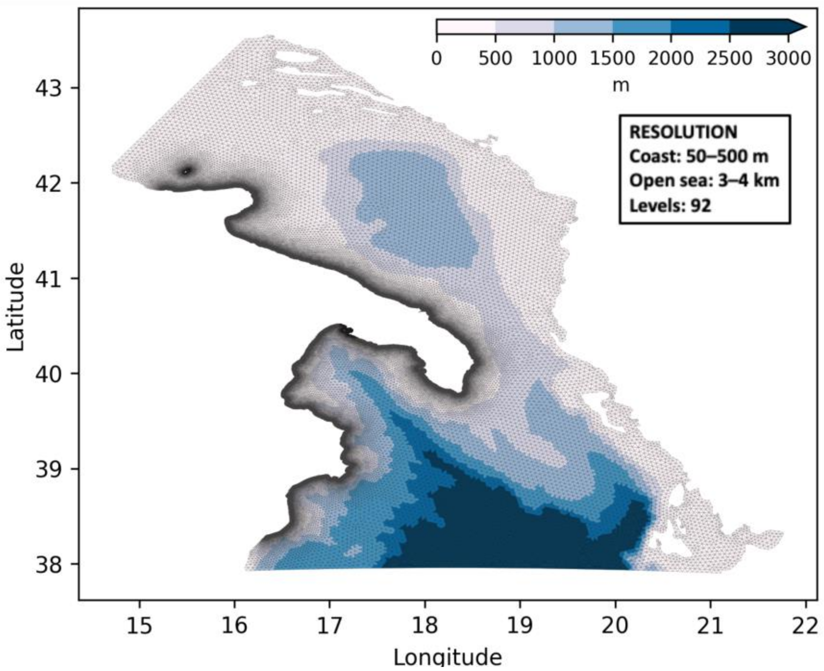
Data Learning to reduce the errors in the solution of existing systems

having benefit from AI without changing your existing system

cmcc
Centro Euro-Mediterraneo sui Cambiamenti Climatici

Marco Stefanelli





DDA for Sea Surface Temperature

… we are happy to share

Weekly meetings with invited speakers from other universities or companies:

We meet every Tuesday at 4pm (UK time) on Zoom
Our mailing list:
https://mailman.ic.ac.uk/mailman/listinfo/datalearning

All the talks are recorded and uploaded on our YouTube Channel – Data Learning

International Conference:

Every year, the DataLearning group organises a workshop on **Machine Learning and Data Assimilation for Dynamical Systems (MLDADS)**, as part of the International Conference on Computational Science (ICCS).

London - ICCS 2022   Poland - ICCS 2021   Amsterdam - ICCS 2020   Faro, Portugal - ICCS 2019

Sharing contents with our community worldwide:
To get access to our **codes**: **Our GitHub** https://github.com/DL-WG

3 Open special Issues
ELSEVIER

There is nothing measured that doesn't exist.

Thank you!

Some other papers and applications:
https://sites.google.com/view/rossella-arcucci/datalearning