Imperial College London

Artificial Neural Network at the service of Data Assimilation (and vice versa)

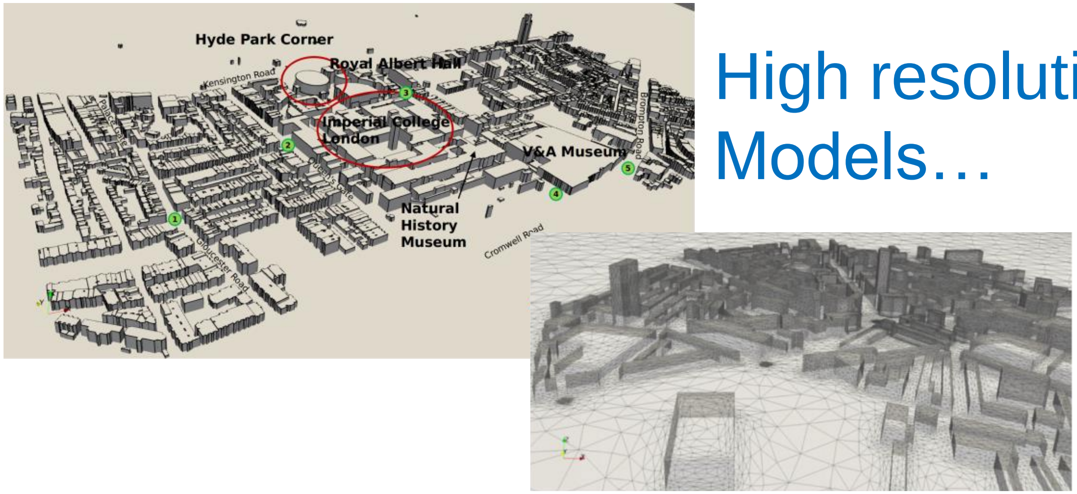
Dr. Rossella Arcucci

Data Science Institute, Imperial College London

06 October 2020

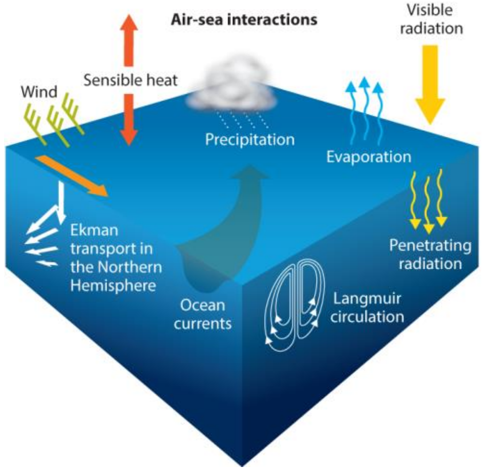
ECMWF - ESA Workshop on Machine Learning for Earth System Observation and Prediction

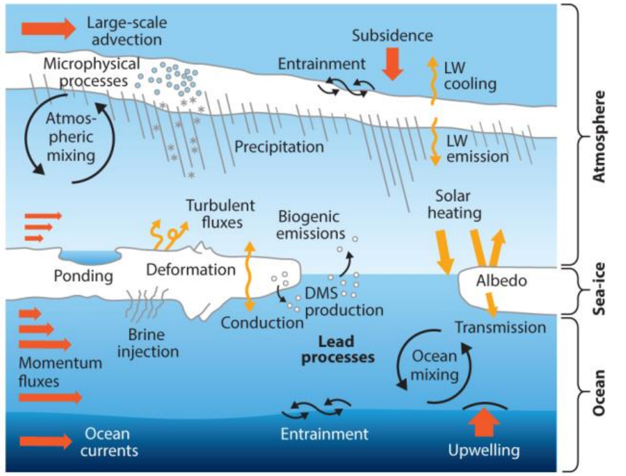
… the era of the data!



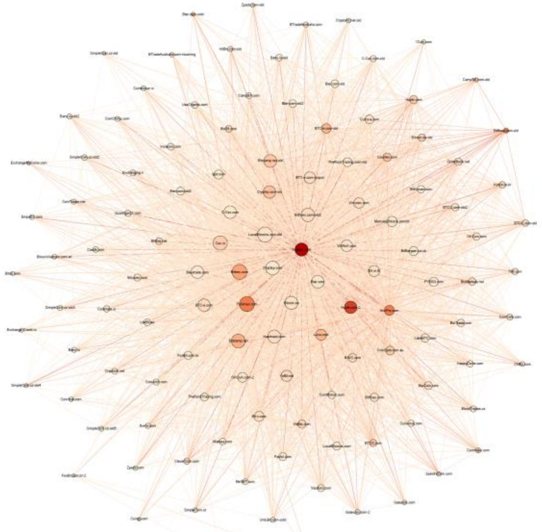
High resolution Models…

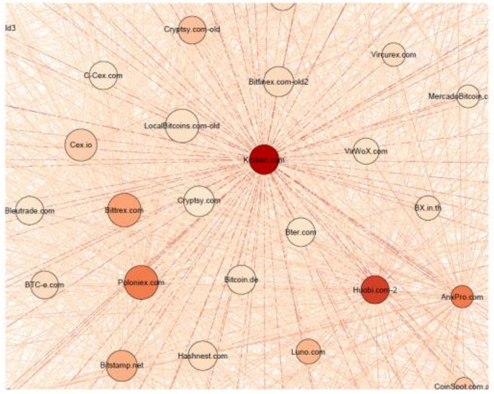
Coupled models…

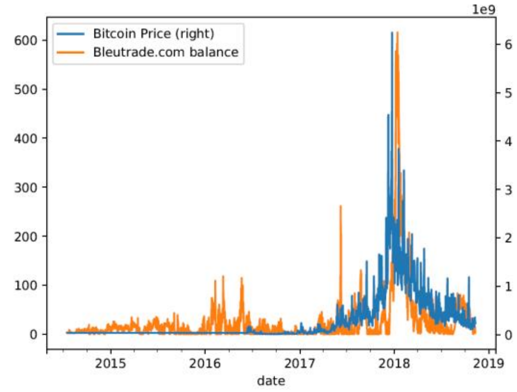




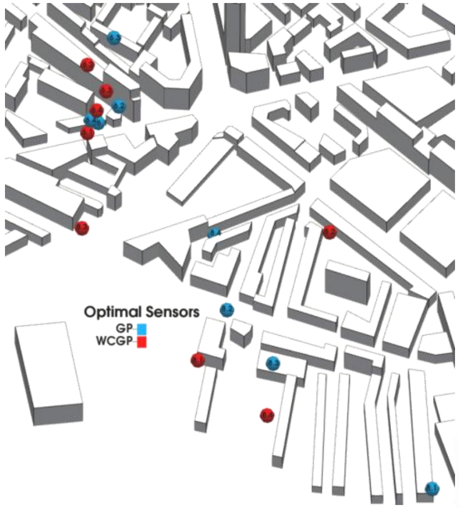
Observations…



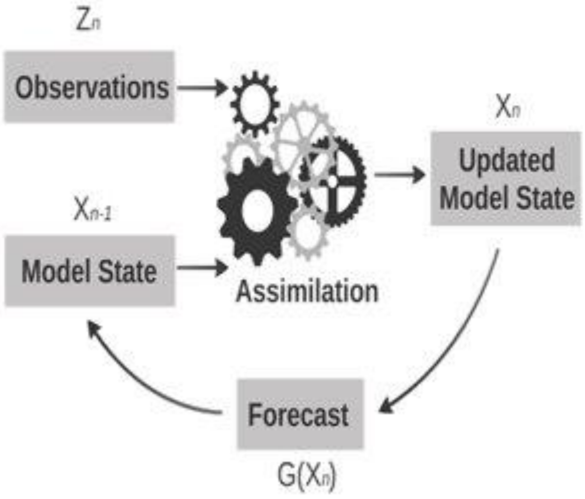


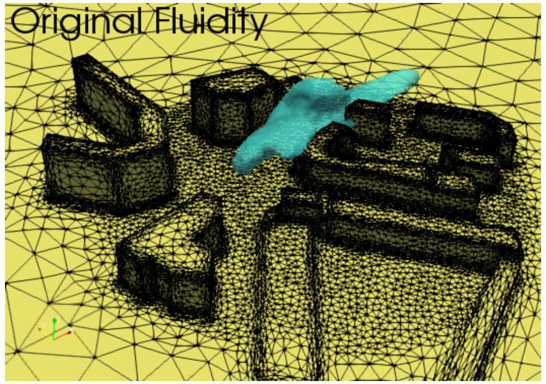


Content

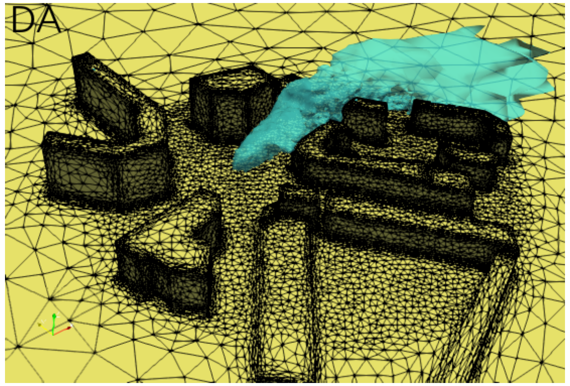


sensors





simulation



data assimilation

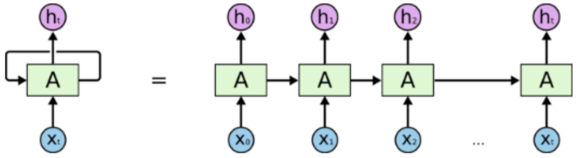
To have the expected benefit from this technology, we need:

- **Good efficiency** (accelerate the execution time)
- **Good accuracy** (reduce the errors propagation in the models)
- **Real world scenarios** (develop models that can be used in real world scenarios – face **Big Data** problems)
- **Good quality data** (optimal placements/locations)

ANN at the service of DA???

Conventional methods for DA have increased in sophistication to better fit their application requirements and circumvent their implementation issues. Nevertheless, these approaches are incapable of fully overcoming some unrealistic assumptions.

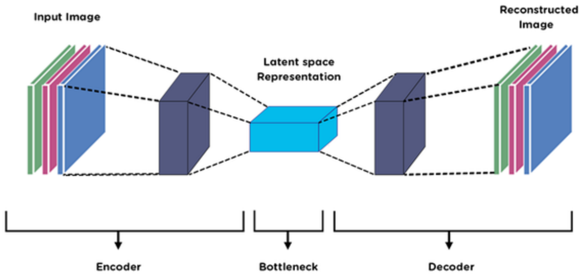
- **Nonlinearities:** Can we **replace** part of the **DA** process (or the entire DA process) **by a ML** technology?
- **Heavy background error covariance matrices**: can we compute them in the **latent space** of an Auto-Encoder?
- **Computationally expensive CFD software and difficulties to compute the adjoints of the models**: can we use **surrogate models** in the DA process?



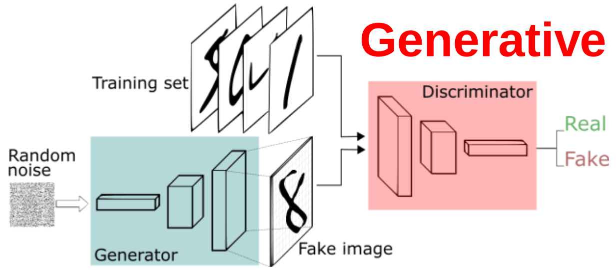
An unrolled recurrent neural network.

Recurrent NN

Encoder-Decoder



Generative Adversarial Networks



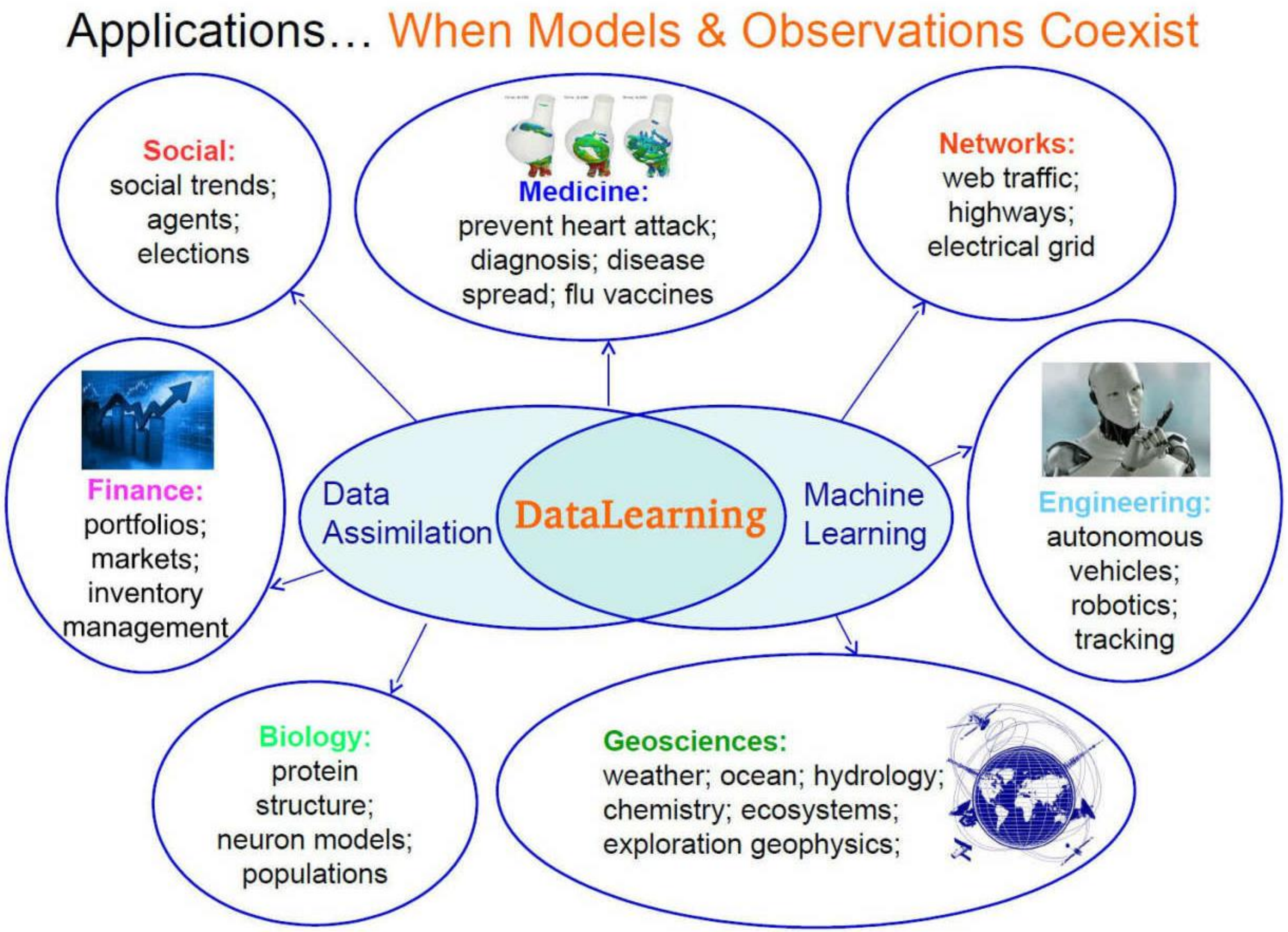
And vice versa???

Building Machine Learning 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

Data Assimilation is the missing piece!!!

Data Learning



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

How to improve estimations or predictions? … Data Assimilation

How to have real time predictions? … Reduced Order Models

How to achieve both accuracy and efficiency? … Machine Learning with Data Assimilation

How to face Big Data problems? … introducing Domain Decomposition

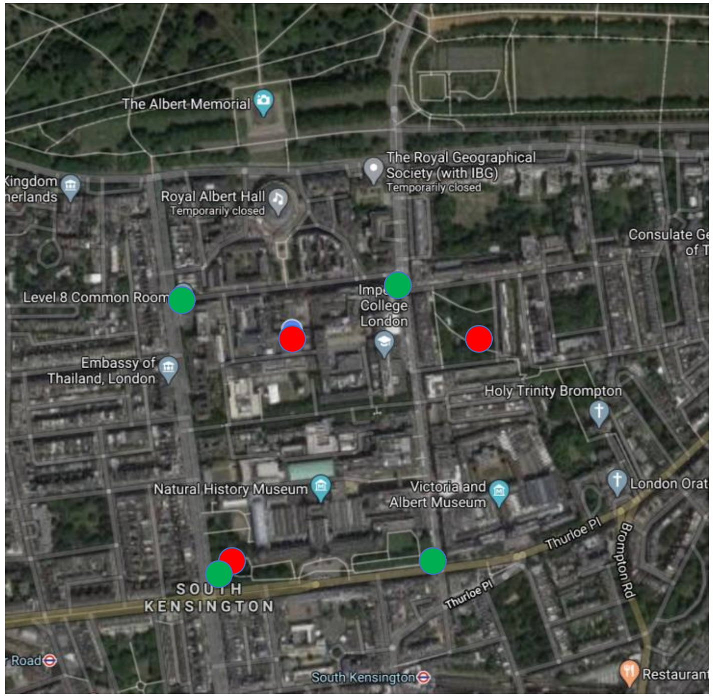
In the following … Air pollution

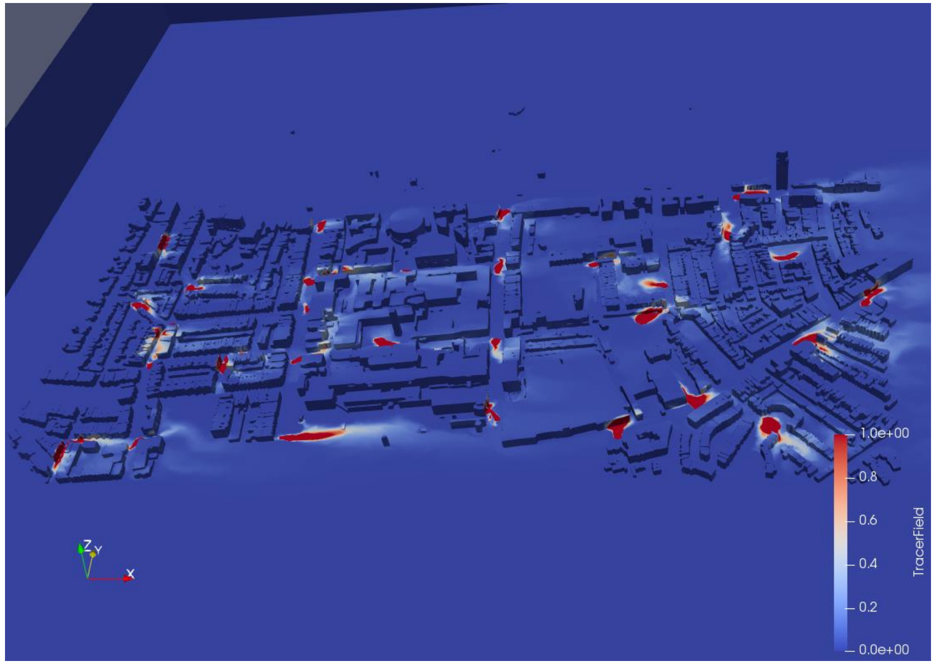
The CFD software is named FLUIDITY

3D non-hydrostatic Navier-Stokes equations

$$
\begin{aligned}
\nabla \cdot \mathbf{u} &= 0, \\
\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} &= -\nabla p + \nabla \cdot \boldsymbol{\tau},
\end{aligned}
$$

fluidity





Left: Satellite image of the South Kensington area around Imperial college London (Google maps),

Right: Computational Fluid dynamics simulation of the same area (Fluidity).

INHALE

MAGIC
Envisaging a world with greener cities

Imperial College London

UNIVERSITY OF CAMBRIDGE

UNIVERSITY OF SURREY

THE UNIVERSITY OF EDINBURGH

EPSRC

UKRI Engineering and Physical Sciences Research Council

Data Assimilation

Algorithm 1: DA

Input: for $k = 0, \dots, m$ temporal steps: observations y_k, matrices R_k, model M, function H_k, background u_0, historical data $S = \{u_j\}_j = 1, \dots, n$,

Compute B_0 from u_0 and S

Initialize iteration $k = 1$

while $k < n$ **do**

    Compute $u_k = M\ u_{k-1}$

    Compute B_k from u_k and S

    Compute

$$u_k^{DA} = argmin_u \left\{ \|u - u_k\|_{B_k^{-1}} + \|v_k - H_k u\|_{R_k^{-1}} \right\}$$

    Count up k for the next iteration

end

Output: u^{DA}

Data Assimilation with Machine Learning

Algorithm 1: DA

Input: for $k = 0, \ldots, m$ temporal steps: observations y_k, matrices R_k, model M, function H_k, background u_0, historical data $S = \{u_j\}_j = 1, \ldots, n$,

Compute B_0 from u_0 and S

Initialize iteration $k = 1$

while $k < n$ **do**

  Compute $u_k = \boxed{M}\, u_{k-1}$

  Compute $\boxed{B_k}$ from u_k and S

  Compute

$$u_k^{DA} = argmin_u \left\{ \|u - u_k\|_{B_k^{-1}} + \|v_k - H_k u\|_{R_k^{-1}} \right\}$$

  Count up k for the next iteration

end

Output: u^{DA}

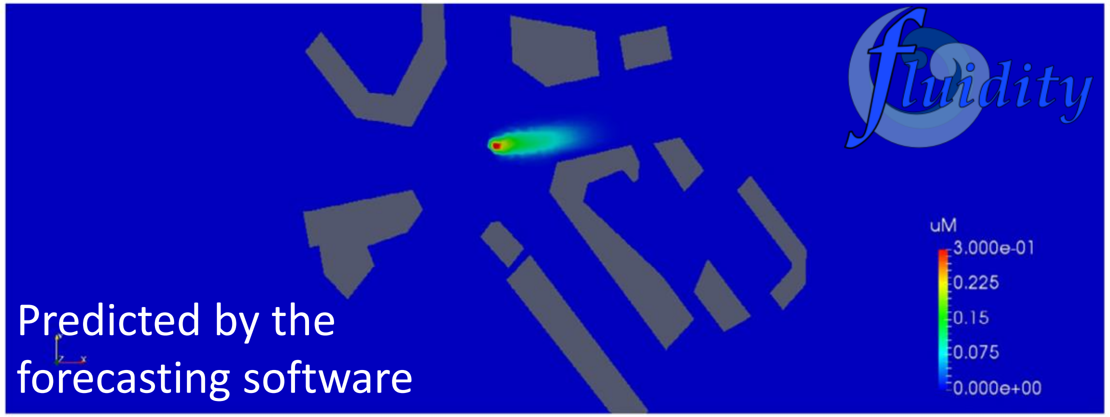
efficiency

machine learning to accurately **model the dynamic systems** reducing the CPU time

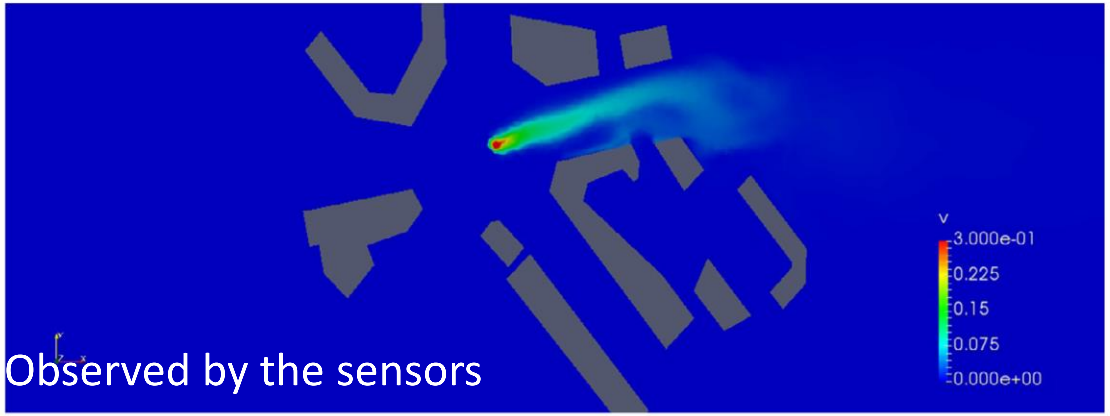
accuracy

machine learning models to **reduce the errors** in the assimilated data

Data Assimilation



(a) Values of $\mathbf{u}_0$: predicted pollutant concentration field



(b) Values of $\mathbf{v}$: observed pollutant concentration field

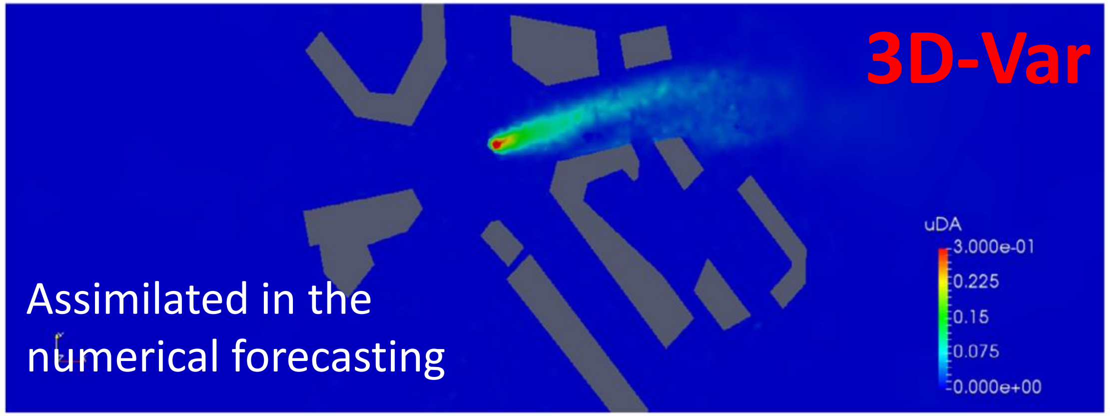
$$J(\mathbf{u}) = \alpha \|\mathbf{u} - \mathbf{u}_0\|^2_{\mathbf{B}^{-1}} + \|\mathbf{G}\mathbf{u} - \mathbf{v}\|^2_{\mathbf{R}^{-1}}$$ ← DA function

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

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



(c) Values of $\mathbf{u}^{DA}$: assimilated pollutant concentration field

[*] R. Arcucci, L. Mottet, C. Pain and Y. Guo - **Optimal reduced space for Variational Data Assimilation** -Journal of Computational Physics, Vol 379, pag: 51-69
[**] R. Arcucci, C. Pain, Y. Guo, **Effective variational data assimilation in air-pollution prediction**, Big Data Mining and Analytics, Vol 1, Issue 4 pag: 297 - 307, 2018

Data Assimilation in a latent space

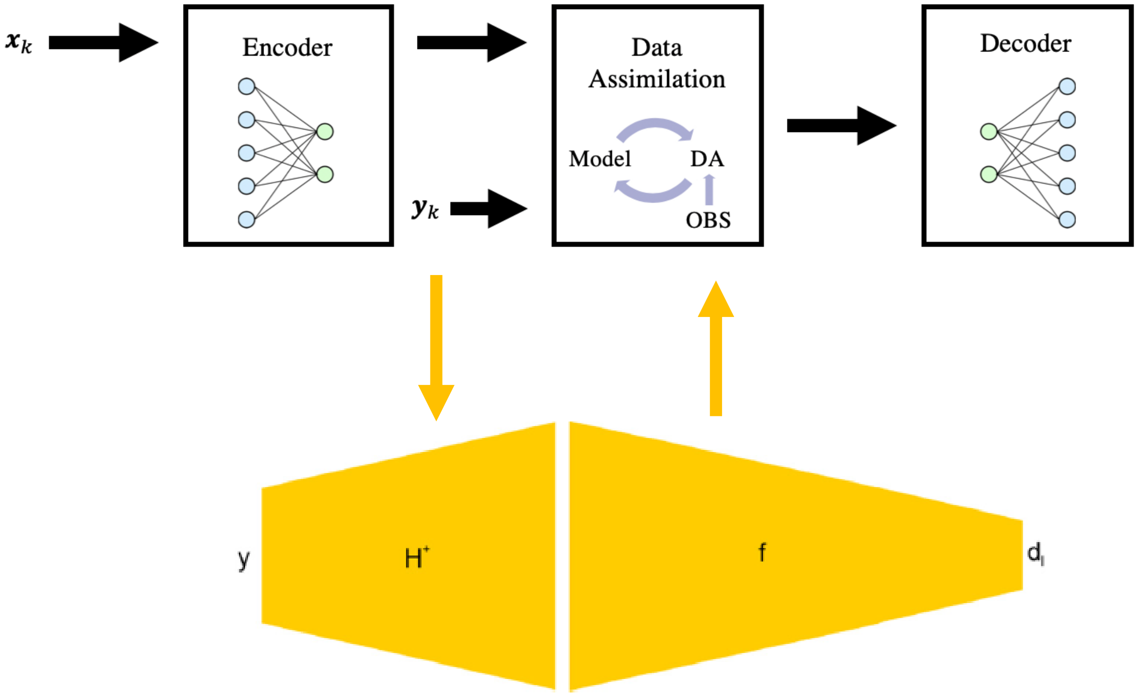
***with Julian Mack - 2019**

3DVar in the latent space

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

$$J(\mathbf{w}_l) = \frac{1}{2}\mathbf{w}_l^T\mathbf{w}_l + \frac{1}{2}\|d_l - V_l\mathbf{w}_l\|^2_{R_l^{-1}}$$

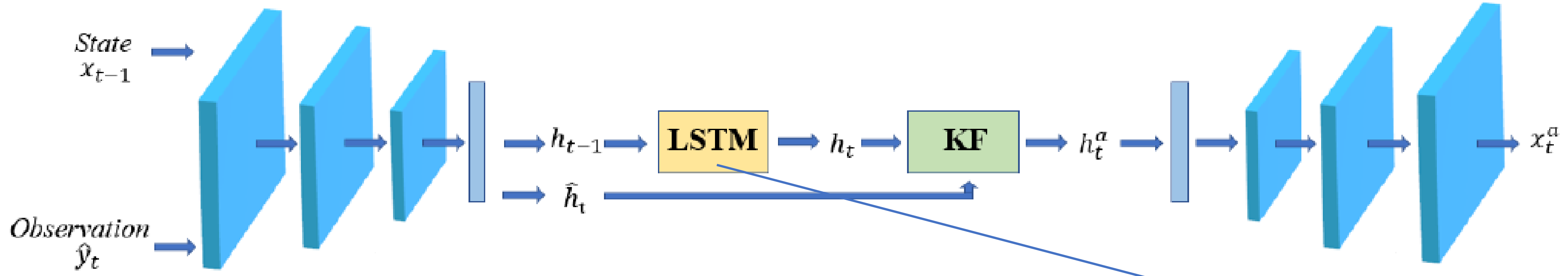
Model	MSE	Execution Time (s)
Ref MSE	1.0001	-
PCA, $\nu = 32$, $m = n$	0.1270	1.8597
PCA, $\nu = 32$, $m = 0.1n$	0.1270	0.2627
PCA, $\nu = 32$, $m = 0.01n$	0.1334	0.0443
PCA, $\nu = 32$, $m = 0.001n$	0.1680	0.0390
Data Learning with Tucodec-NeXt	0.0787	0.0537

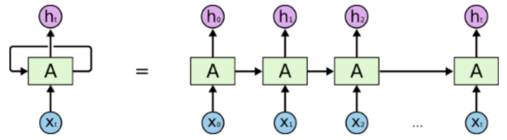


[*] Mack, J., Arcucci, R., Molina-Solana, M., & Guo, Y. K. (2020). Attention-based Convolutional Autoencoders for 3D-Variational Data Assimilation. *Computer Methods in Applied Mechanics and Engineering*, *372*, 113291.

Latent Assimilation

***with Maddalena Amendola - 2020**





An unrolled recurrent neural network.

Time Update ("Predict")

(1) Project the state ahead

$$h_{t+1} = l(h_{t,q})$$

Measurement Update ("Correct")

(1) Compute the Kalman Gain

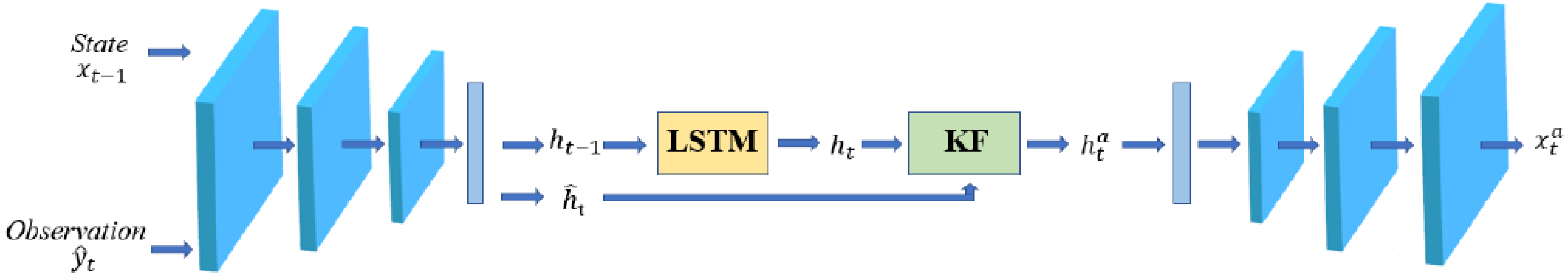
$$\widehat{K}_{t+1} = \widehat{Q}\widehat{H}^T(\widehat{H}\widehat{Q}\widehat{H}^T + \widehat{R}_{t+1})^{-1}$$

(2) Update estimate with measurement

$$h_{t+1}^a = h_{t+1} + \widehat{K}_{t+1}(\hat{h}_{t+1} - \widehat{H}\hat{h}_{t+1})$$

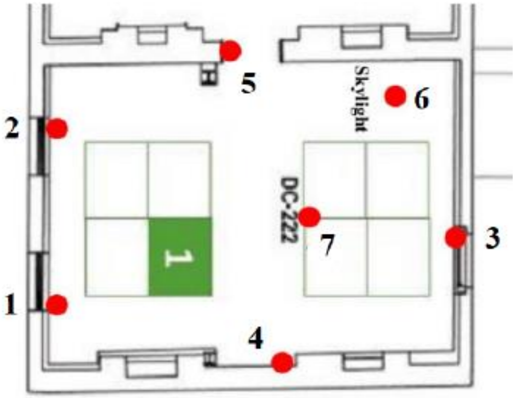
Latent Assimilation

***with Maddalena Amendola - 2020**

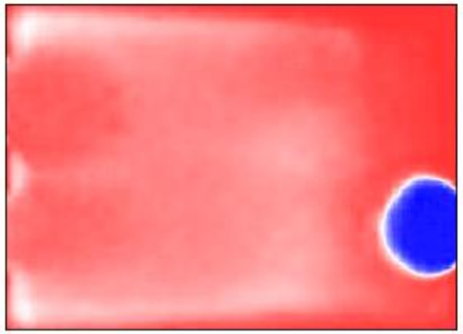


Measurements of carbon dioxide (CO_2) in a room

MAGIC
Envisaging a world with greener cities



sensors



CFD

$\hat{R}$	cov-matrix from data	0.01 I	0.001 I	0.0001 I
MSE	3.356e-02	6.933e-04	1.211e-04	2.691e-06
Time (sec)	3.191e+00	2.899e+00	2.896e+00	2.896e+00

Table 6.11: Values of MSE of of x_t^a in the Physical space for different values of the observations errors covariance matrix $\hat{R}$ with the Structured dataset.

R	cov-matrix from data	0.01 I	0.001 I	0.0001 I
MSE	5.179e-02	6.928e-03	6.928e-03	6.997e-03
Time (sec)	2.231e+03	2.148e+03	2.186e+03	2.159e+03

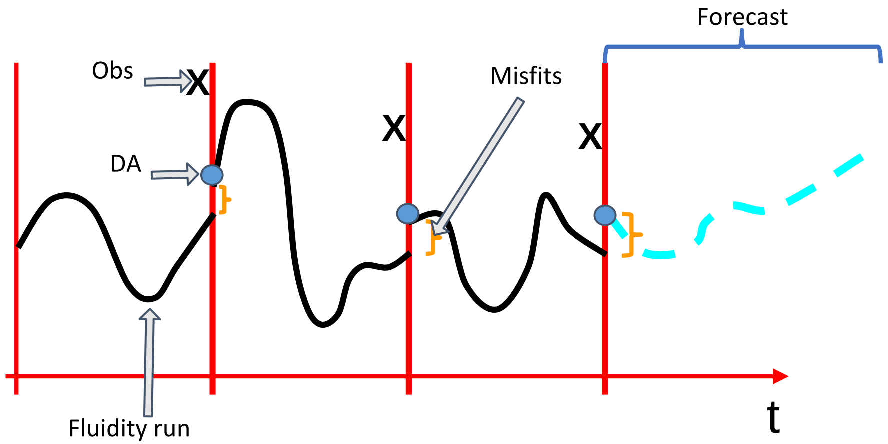
Table 6.12: Standard Assimilation in the physical space performed by a KF

Deep Data Assimilation (DDA)

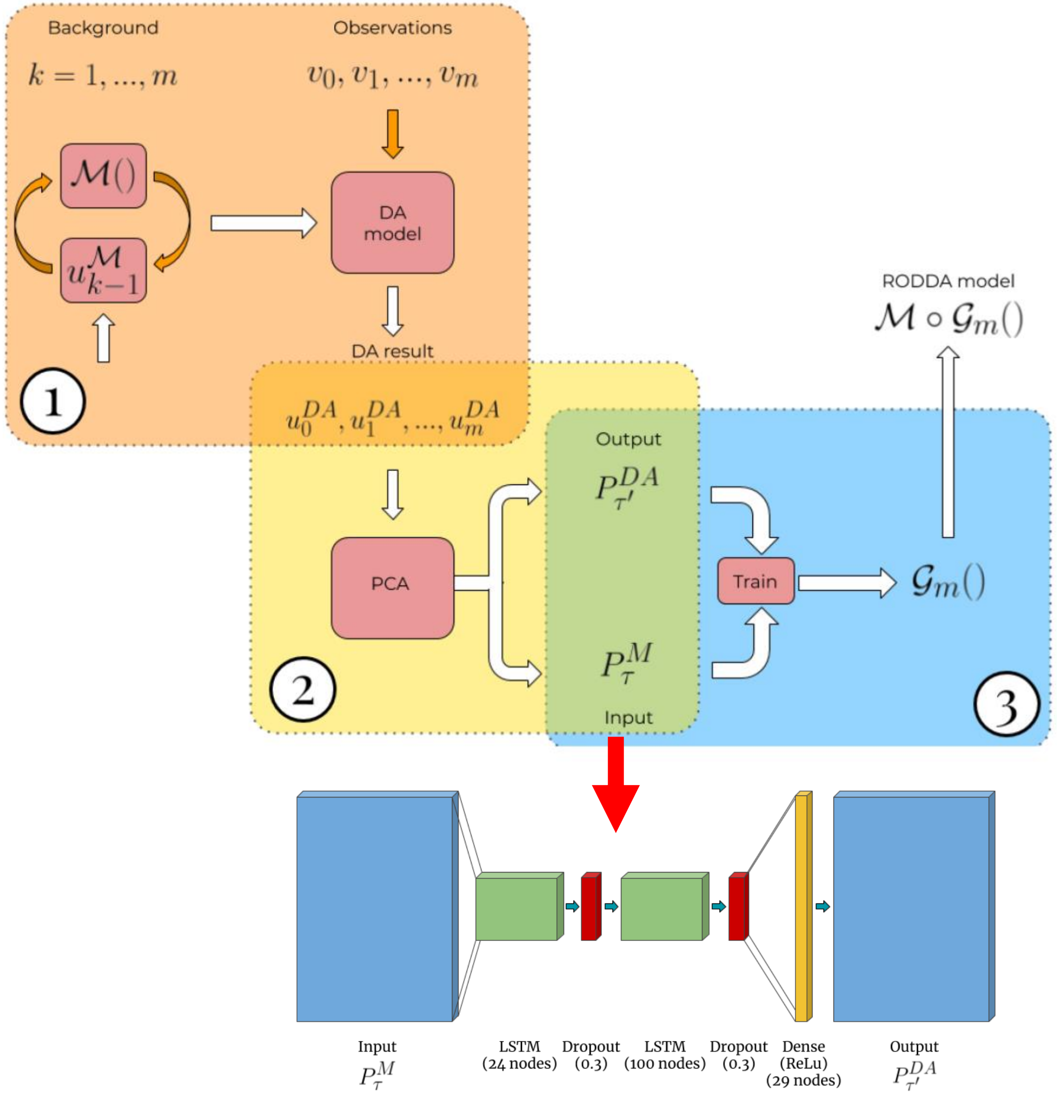
What if the observation are not available?

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)

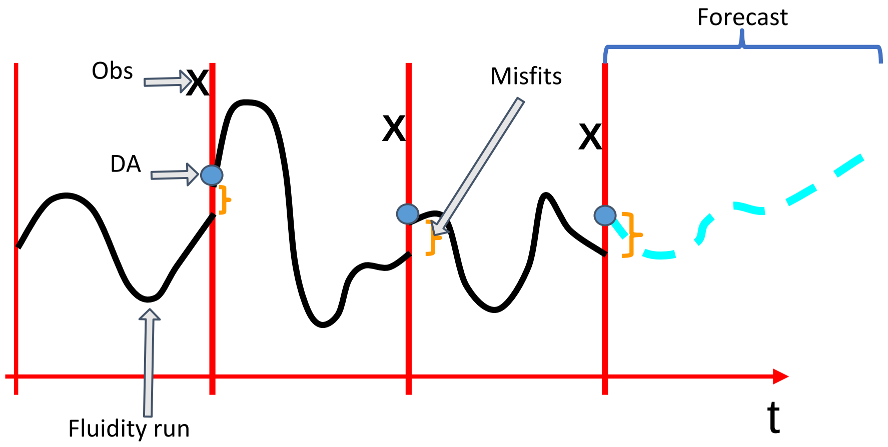


Deep Data Assimilation (DDA)

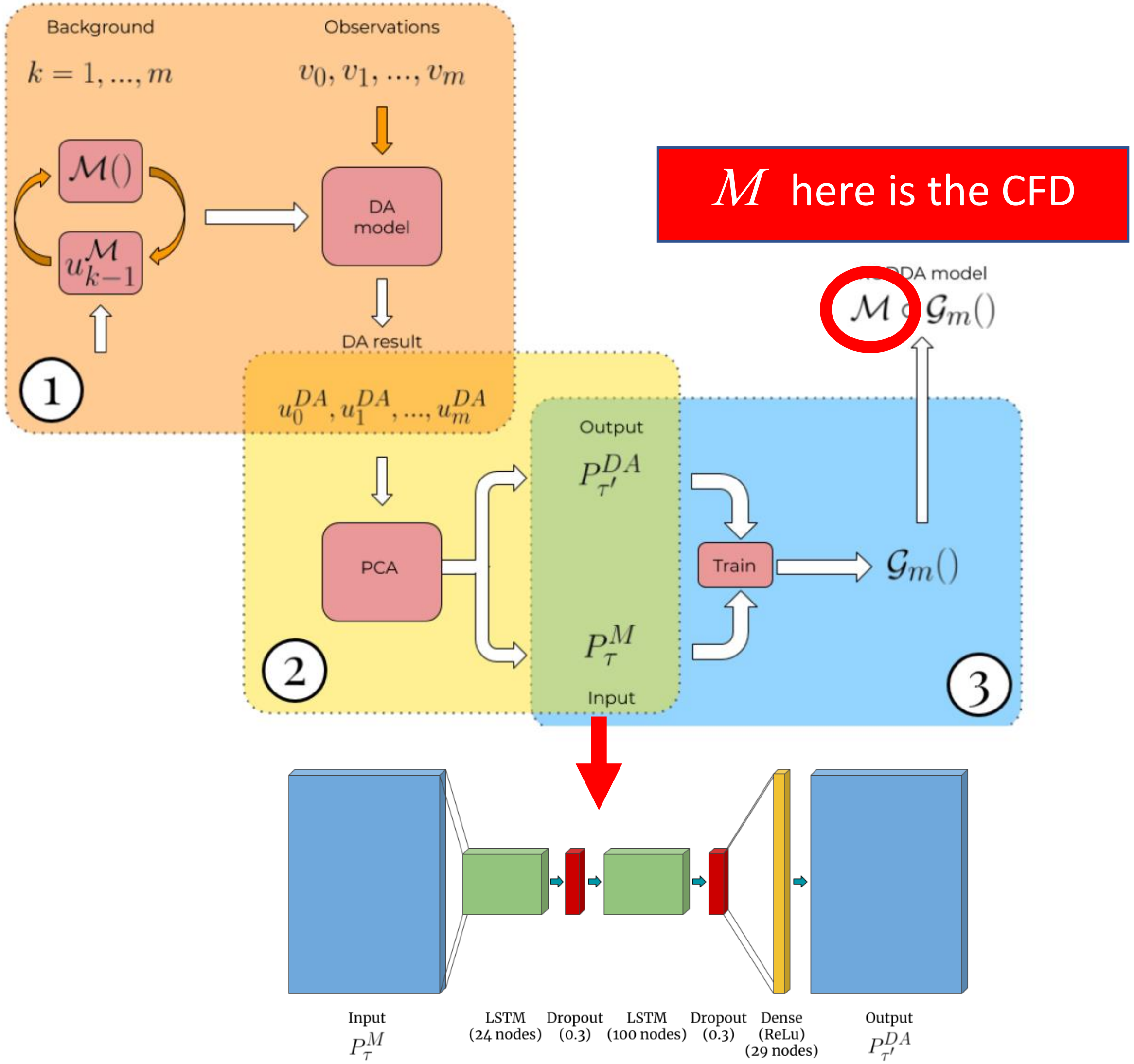
What if the observation are not available?

DDA … learning the Data Assimilation process

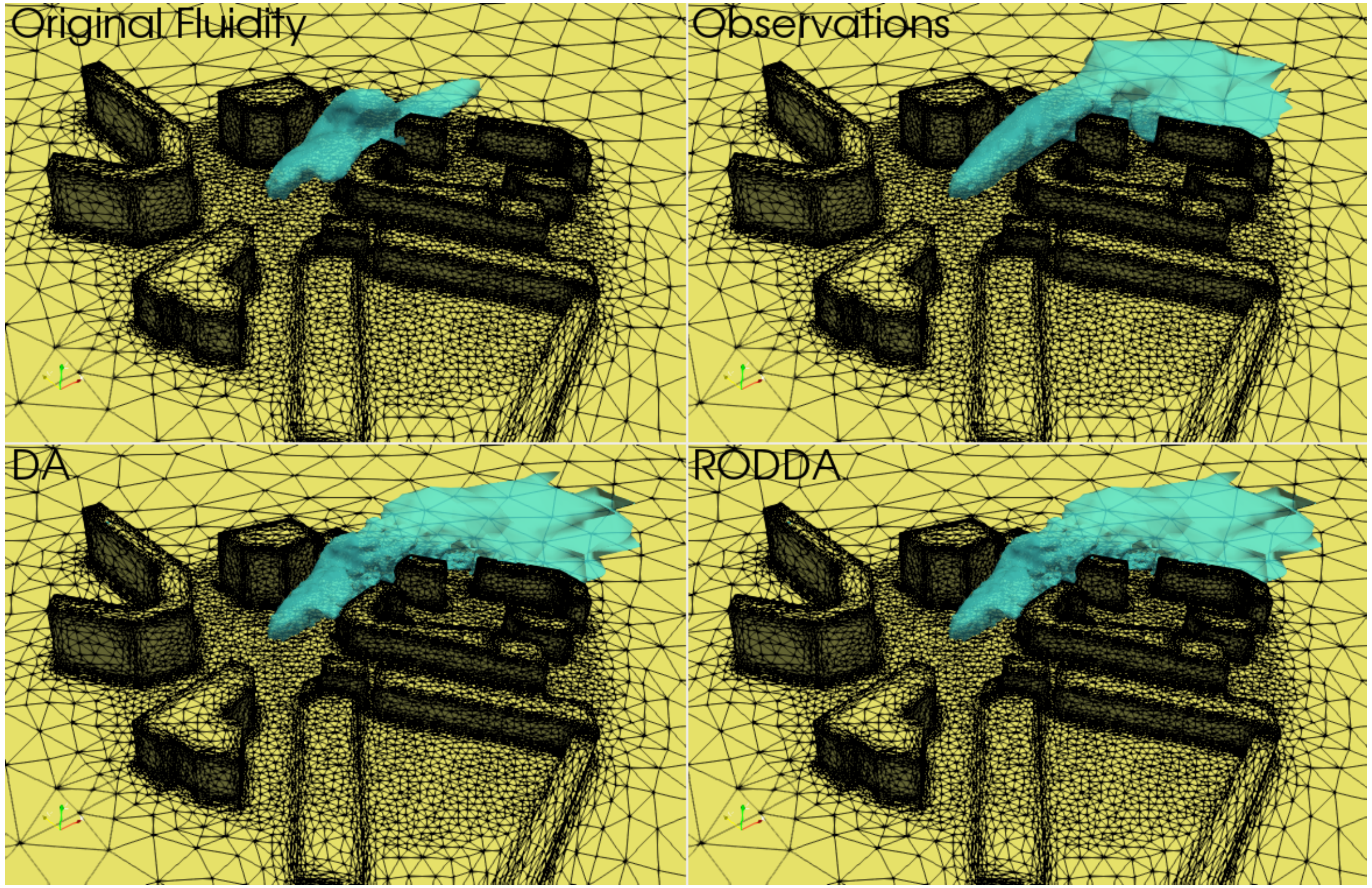
The idea:
Data Assimilation at each time step give us a misfit (DA - fluidity background),
the saved misfits are trained using a Long short-term memory (LSTM) network
and used for future forecasts.

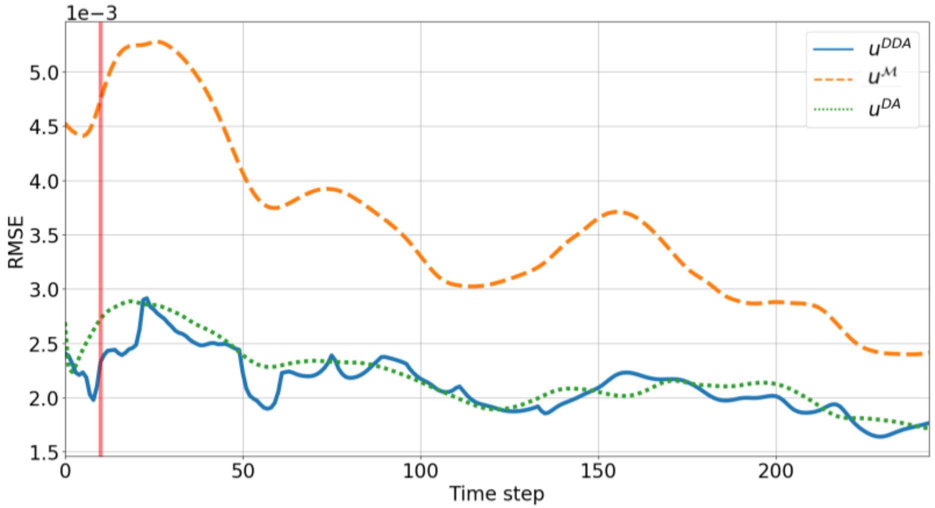


Reduced Order Deep Data Assimilation (RODDA)



Reduced Order Deep Data Assimilation (RODDA)



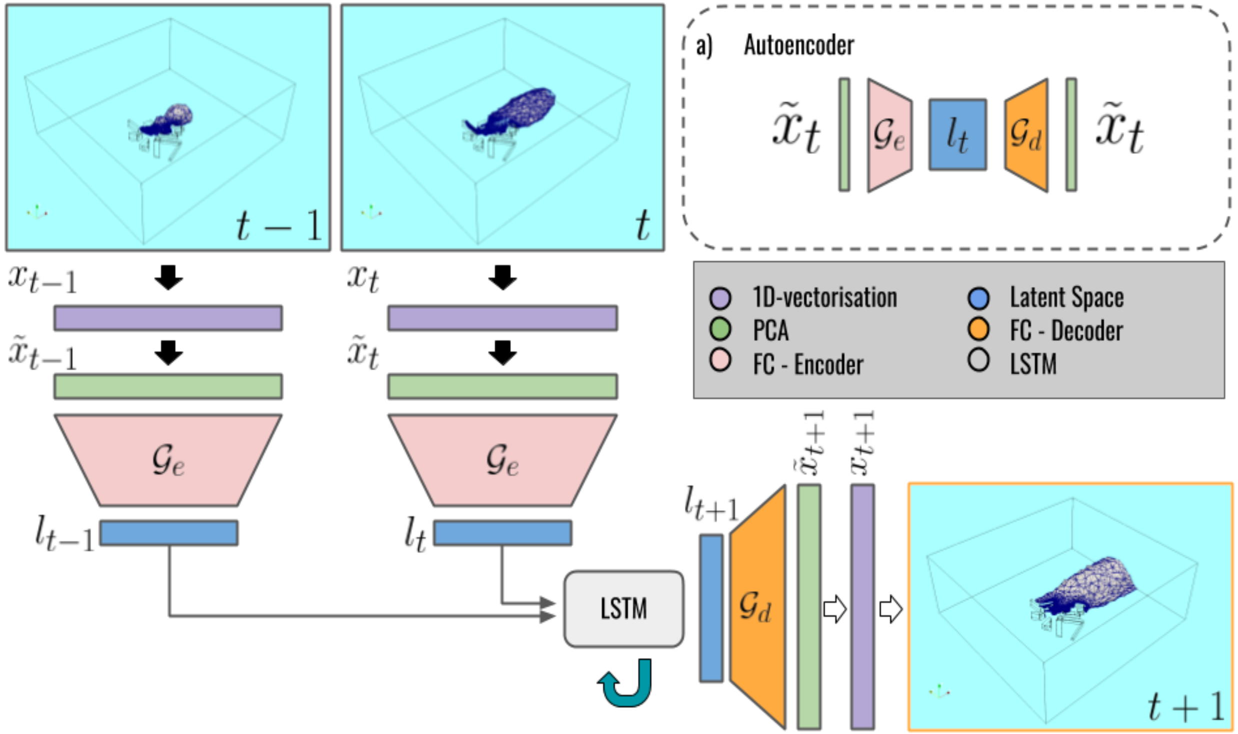


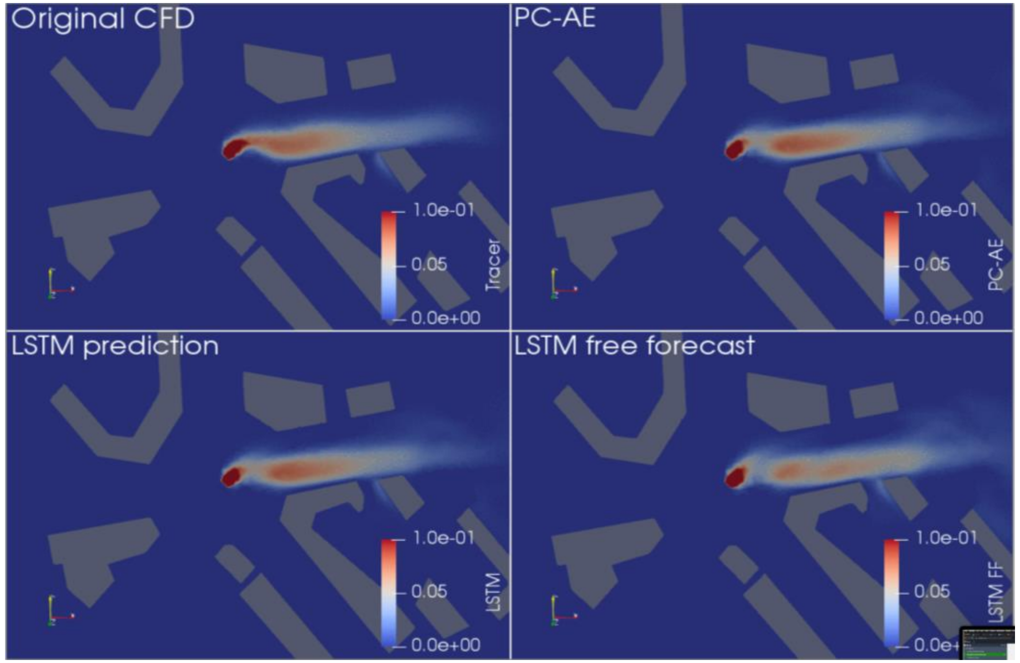
Same accuracy but RODDA is 1000 times faster than DA

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

Surrogate models: Machine Leaning

***with Dr C. Quilodran Casas**





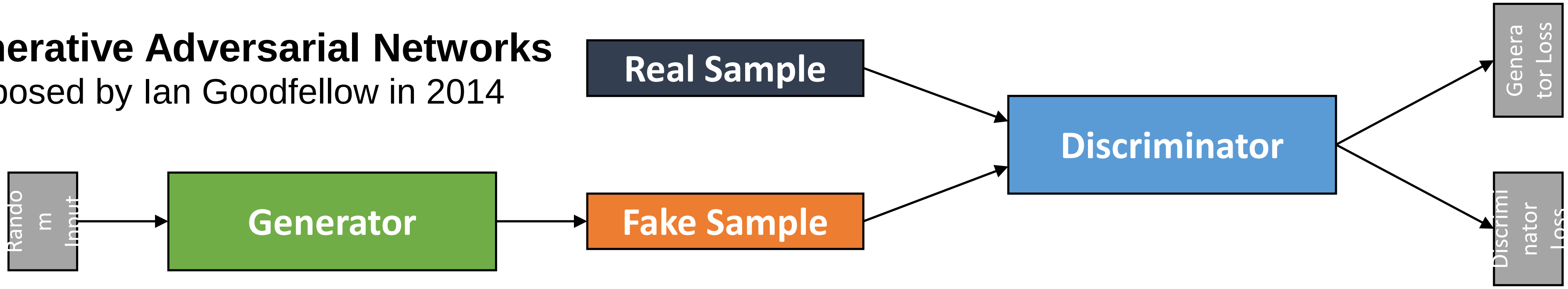
The LSTM running freely (bottom right) after 8 time-steps with no external input.

[*] C. Quilodran Casas, R. Arcucci, Y. Guo - **Urban Air Pollution Forecasts Generated from Latent Space Representations** - International Conference on Learning Representations (ICLR)

The use of GAN for fast predictions

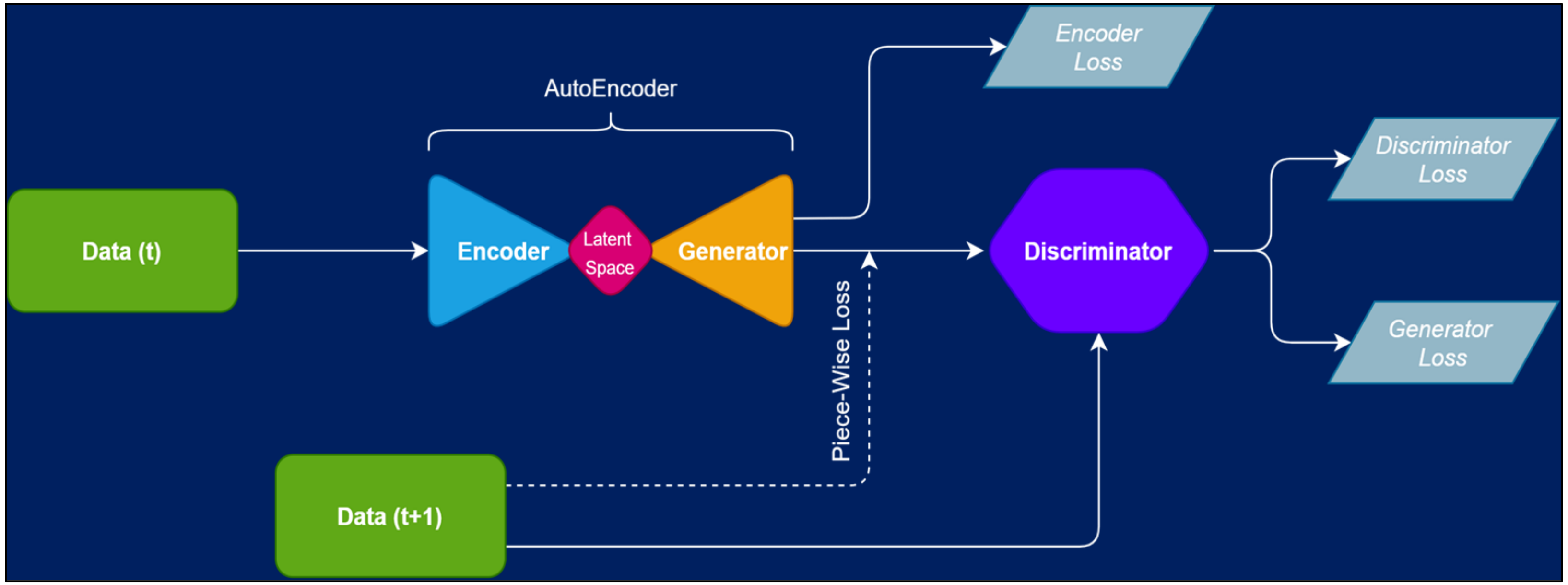
***with Jamal Afzali**

Generative Adversarial Networks
Proposed by Ian Goodfellow in 2014



Generator: aims to produce data that is as close as possible to the original dataset.

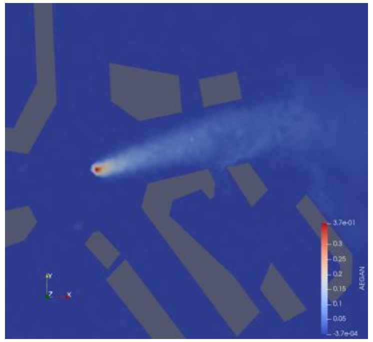
Discriminator: aims to distinguish between real data and generated data



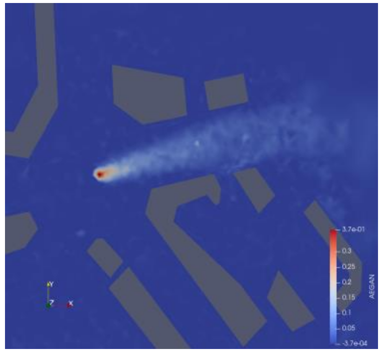
Latent GAN with DA

***with Jamal Afzali**

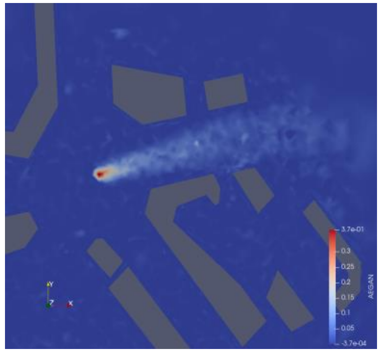
t = 989



t = 990

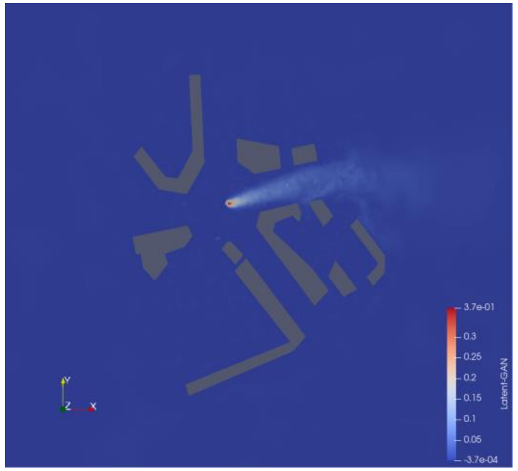


t = 991

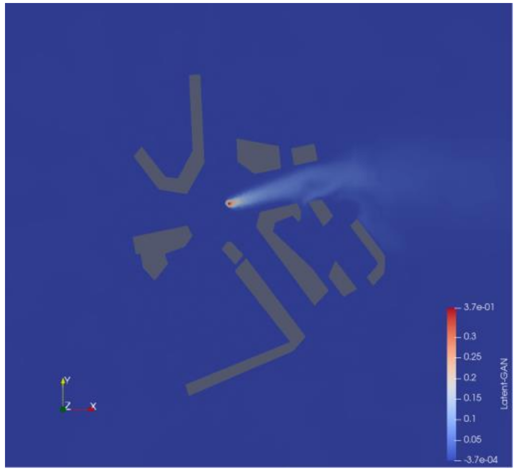


What if we assimilate during the training???

t = 989

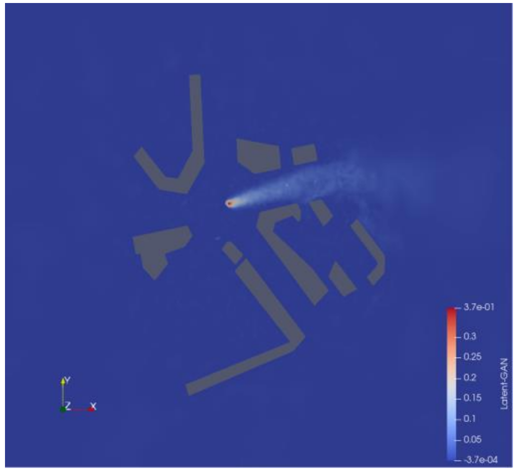


t = 990



assim

t = 991



- **Simulation** for a single timestep took **172 seconds** on a single core of **bi-Xeon E5-2650 v3 CPU** with 250GB RAM
- **Latent GAN prediction** for a single timestep took **1.1 seconds for Tracers** and **1.3 seconds for Velocity Fields** on a single core of **i7-4790k** with 16GB RAM
- Running on **GeForce GTX 970 with 4GB vRAM** took on average **0.25 seconds for Tracers** and **0.3 seconds for Velocity Fields**

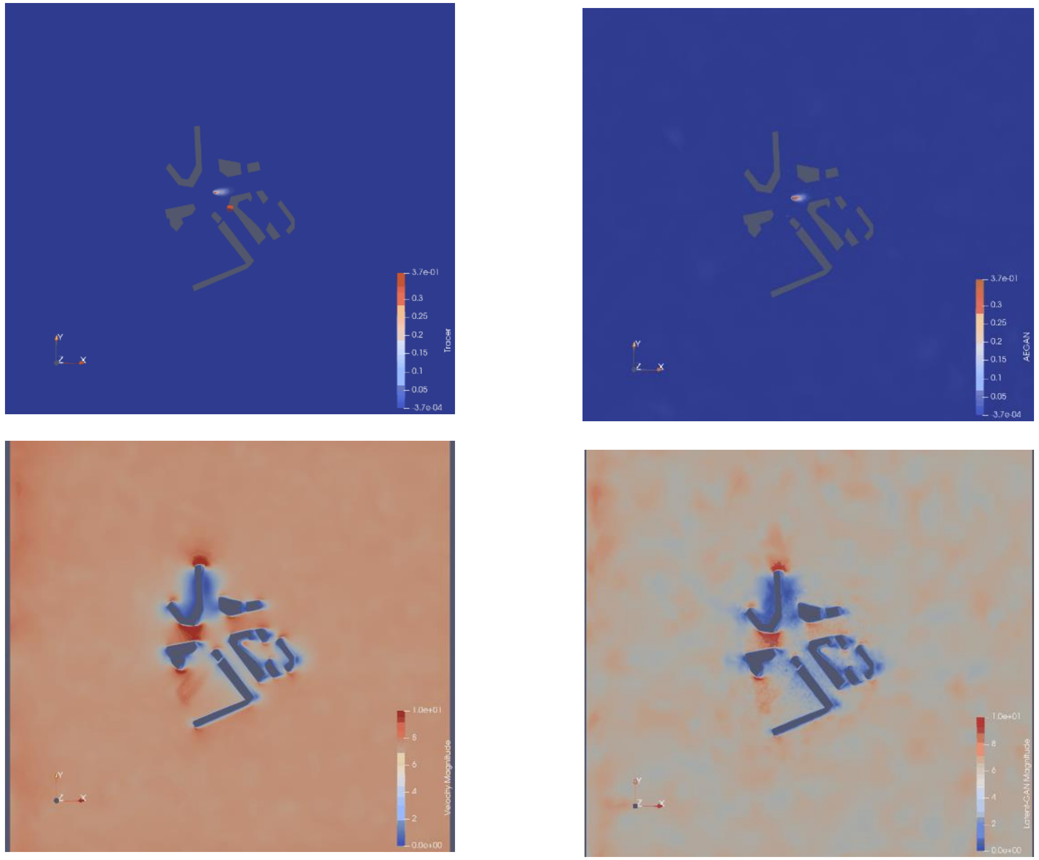
Latent GAN with DA

***with Jamal Afzali**

Ground Truth | **Prediction**

Tracers

Velocity Fields



- **Simulation** for a single timestep took **172 seconds** on a single core of **bi-Xeon E5-2650 v3 CPU** with 250GB RAM
- **Latent GAN prediction** for a single timestep took **1.1 seconds for Tracers** and **1.3 seconds for Velocity Fields** on a single core of **i7-4790k** with 16GB RAM
- Running on **GeForce GTX 970 with 4GB vRAM** took on average **0.25 seconds for Tracers** and **0.3 seconds for Velocity Fields**

Bus stops … ANN + DA

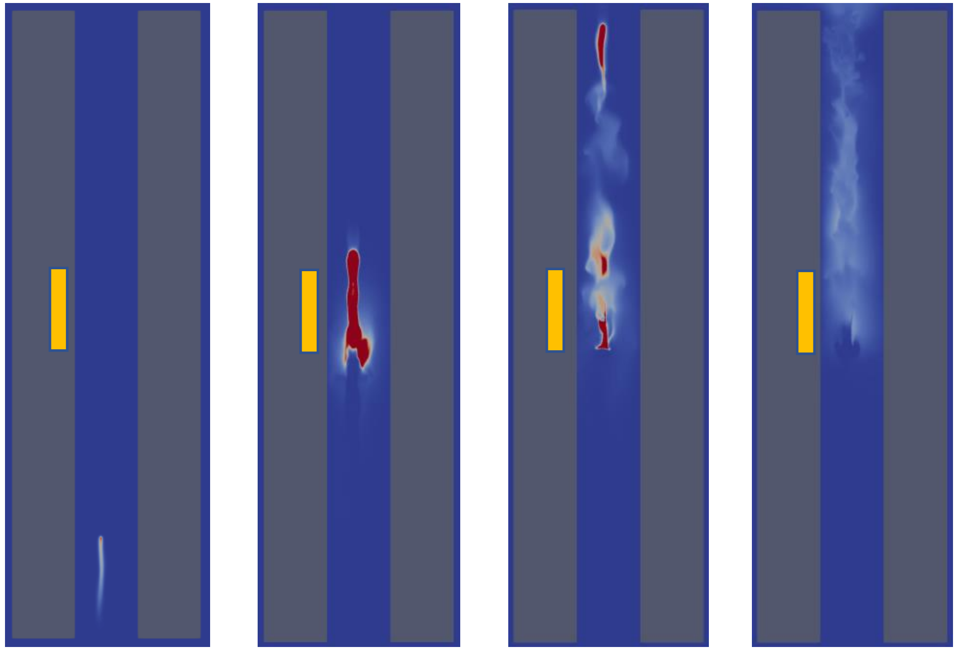
3D Canion

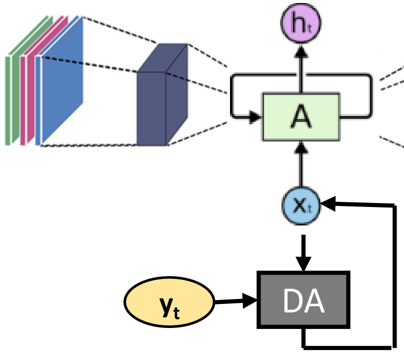
***with Thibault De Rycke and Dr Huw Woodward**

How can we estimate something which has been not CFD simulated before ?

CFD

1 Bus

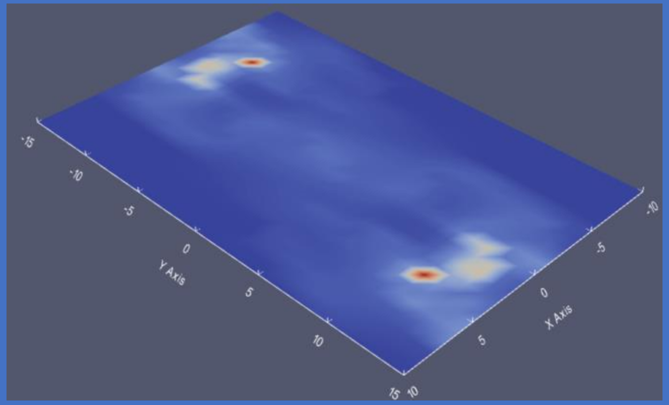


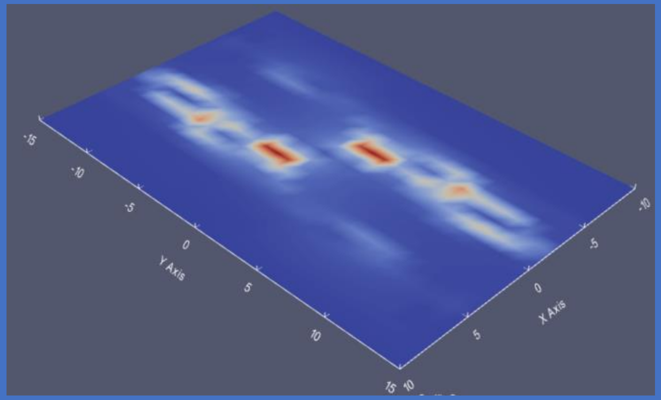


AE + LSTM + DA

AI

2 Busses crossing





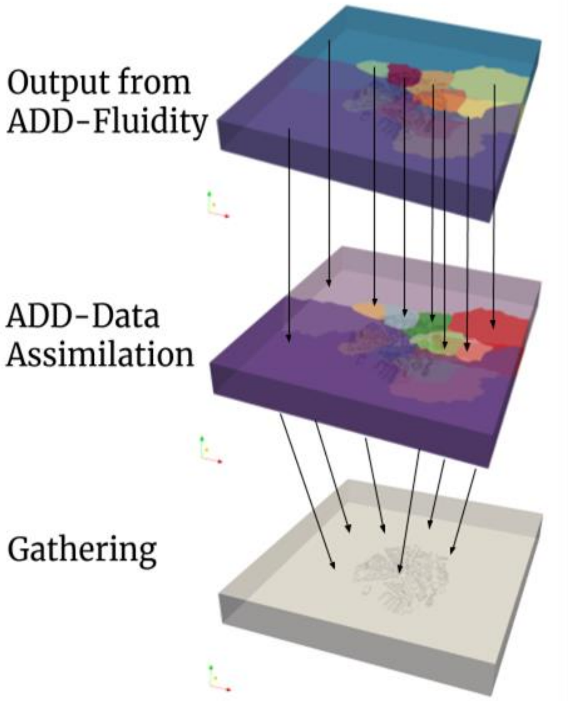
Model	Computation time for 1 epoch
LSTM network	1.2 seconds
LSTM network with OI	2.1 seconds
LSTMs Kalman filter	24.2 seconds

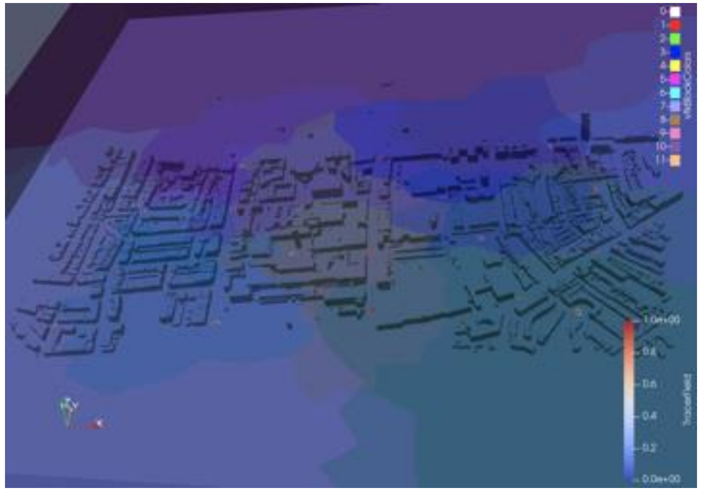
MAGIC

Envisaging a world with greener cities

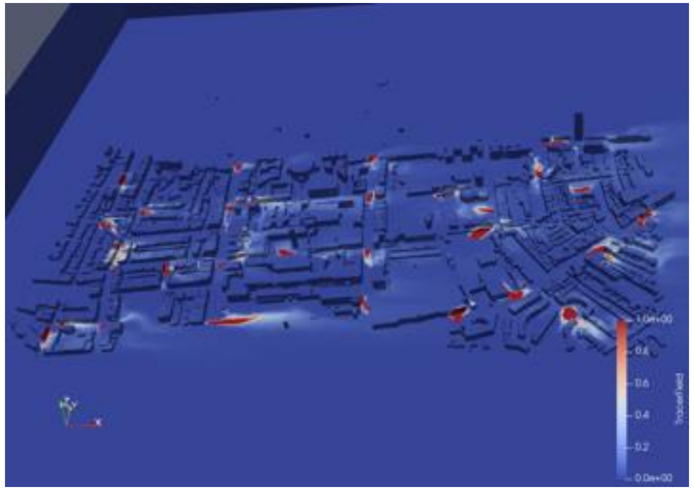
Domain Decomposition for DA and ML

How to face a computationally expensive problem (like Data Assimilation and Machine Learning) on a domain such as a BIG city

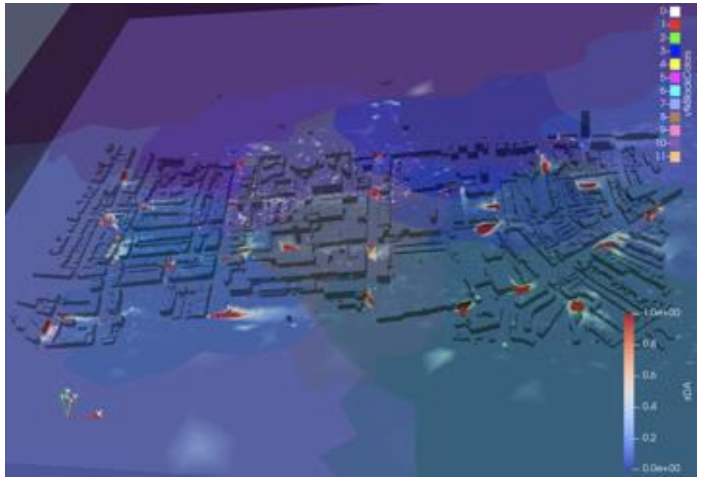




Predicted by the forecasting software



Observed by the sensor



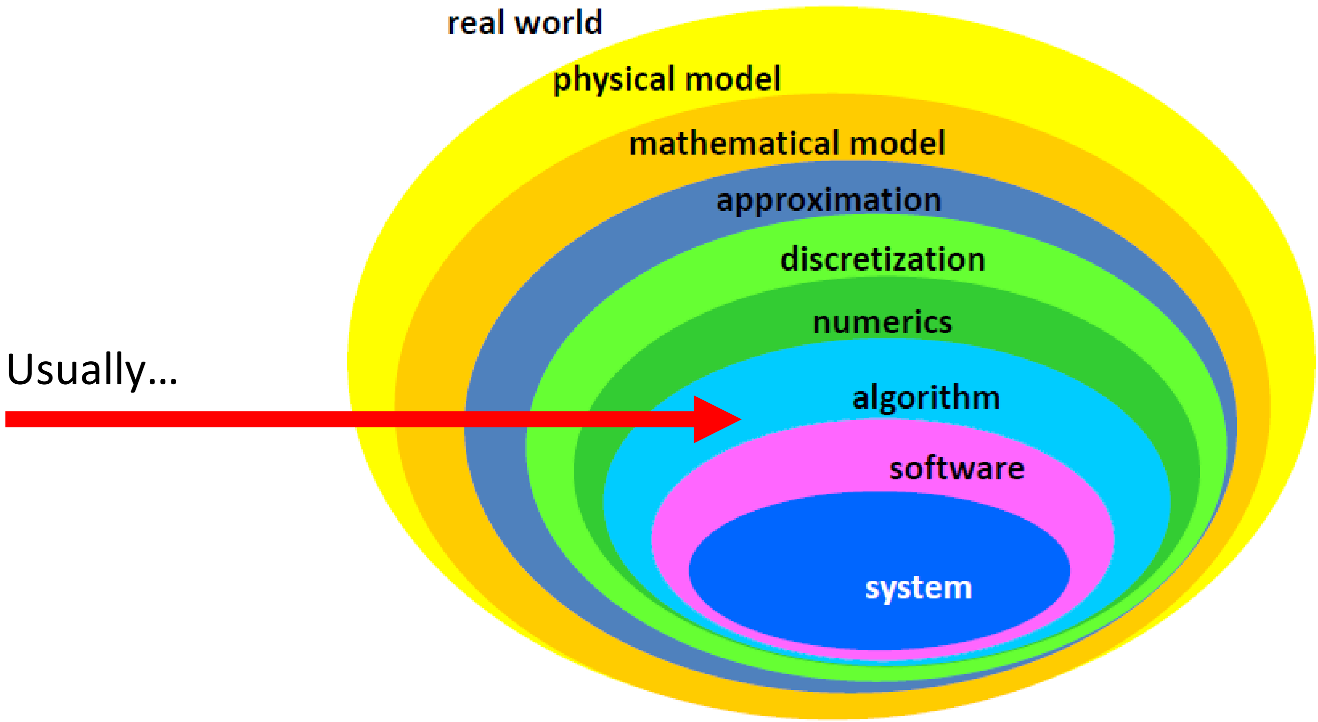
Assimilated in the numerical forecasting

Iso-surface of the pollutant concentration computed in parallel with 12 processors

Data Assimilation with Machine Learning…

New formulations can be introduced at different levels of the

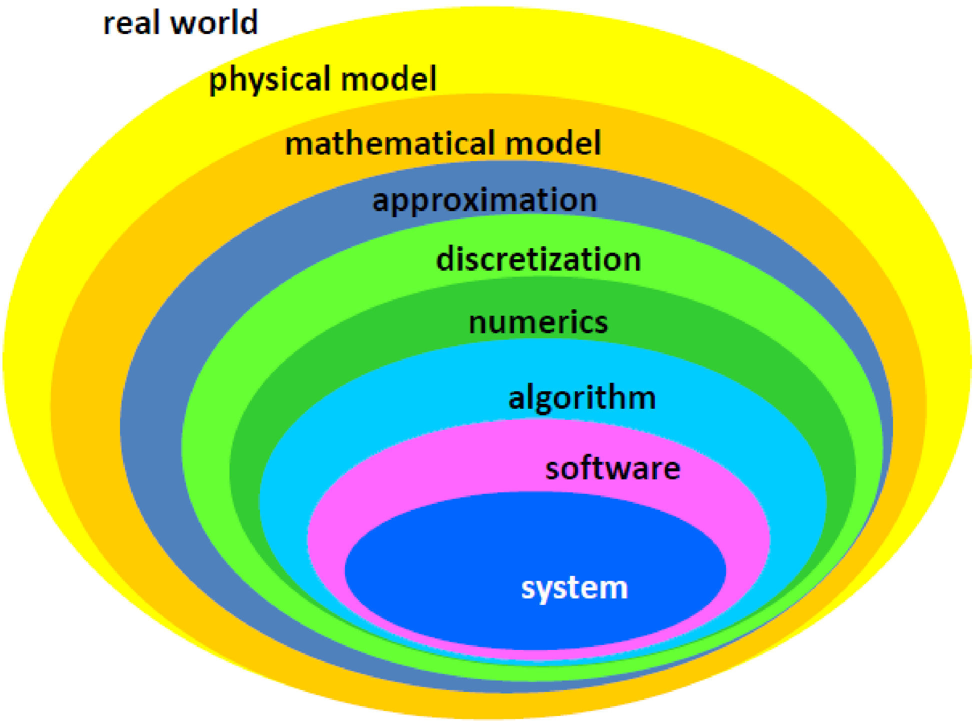
Mathematical Stack



Data Assimilation with Machine Learning…

«Adapting old programs to fit new machines usually means adapting new machines to behave like old ones.»
Alan Perlis

For programmers



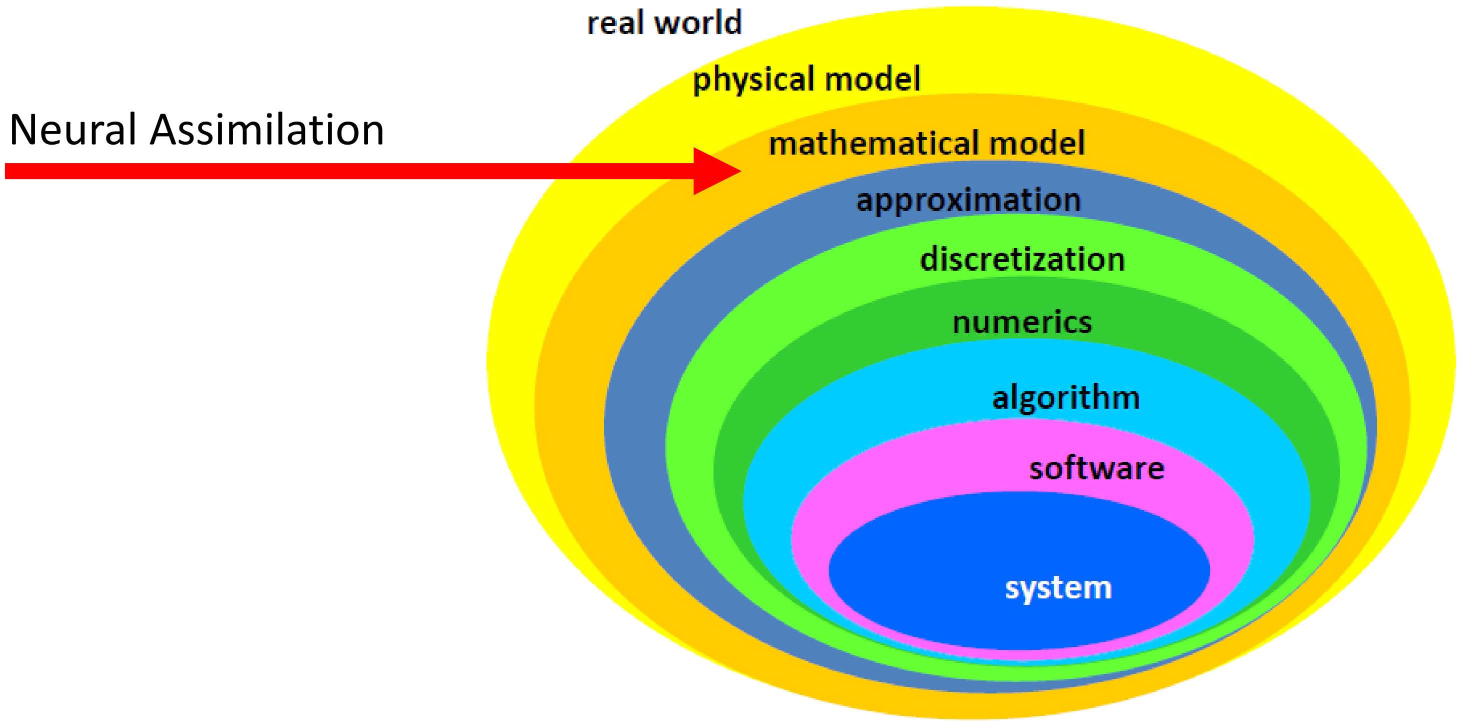
Data Assimilation with Machine Learning…

«Adapting old programs to fit new machines usually means adapting new machines to behave like old ones.»
Alan Perlis

For programmers

«Adapting old algorithms to fit new data sets usually means adapting new data sets to behave like old ones.»
Rossella Arcucci

For data scientists



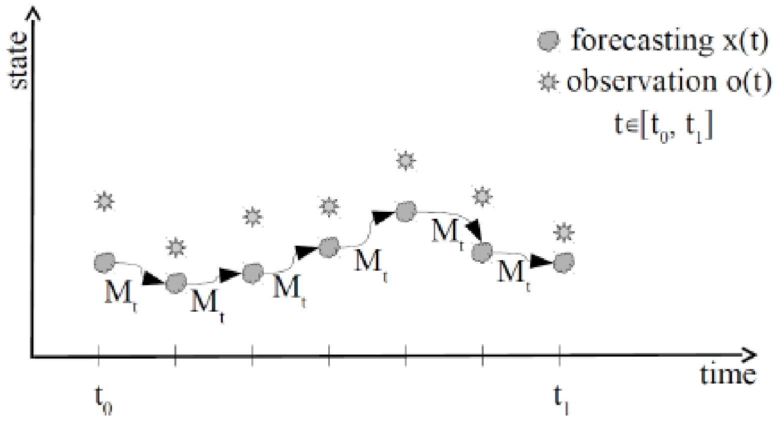
Neural Assimilation

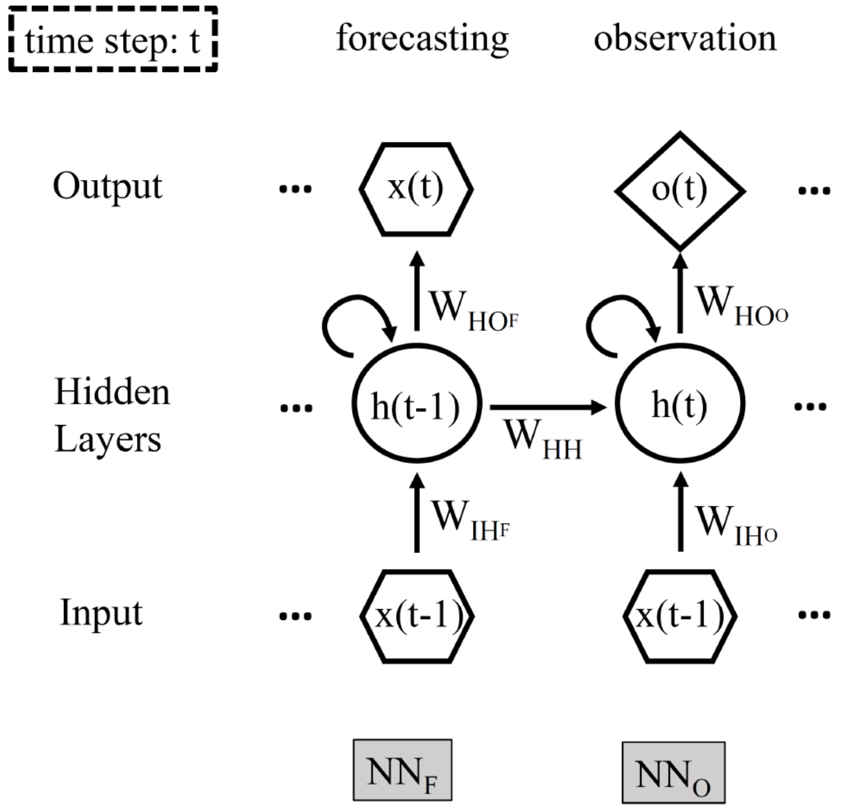
… a coupled neural network made of two Recurrent Neural Networks trained on forecasting data and observed data respectively.

$$\bar{o}(t) = f_{O_O}\left(W_{HO_O} h(t-1)\right)$$

$$h(t) = f_H\left(W_{IH}\bar{x}(t-1) + W_{HH}h(t-1)\right)$$

$$\bar{x}(t) = f_{O_F}\left(W_{HO_F} h(t)\right)$$



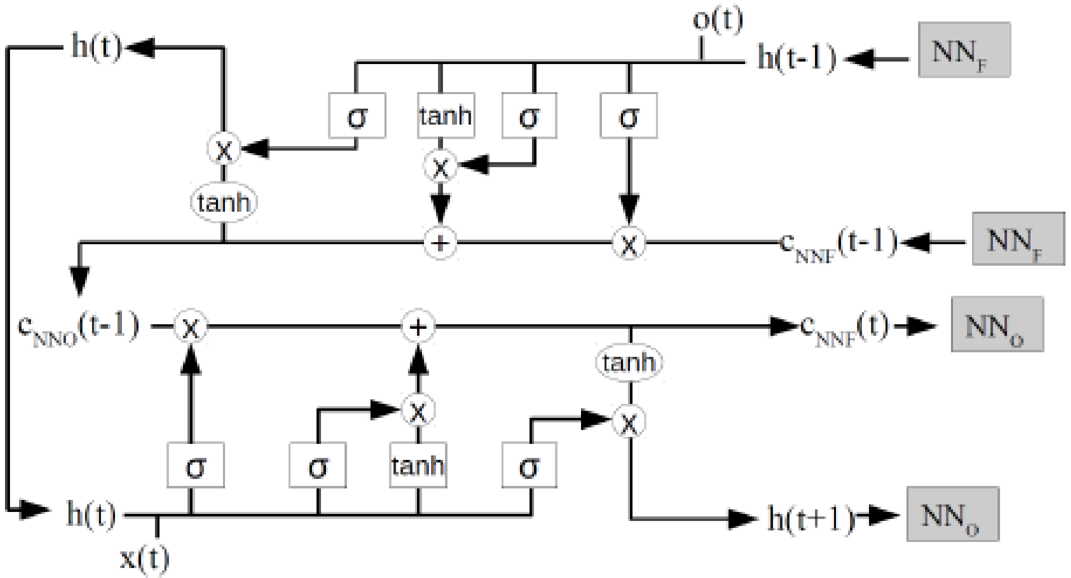


[*] R. Arcucci, L. Moutiq, Y. Guo - Neural Assimilation - Lecture Notes in Computer Science book series Volume 12142

… Neural Assimilation

… adopting Long short-term memory (LSTM) architecture for the two RNNs.

Why LSTM???



- LSTMs are suitable to contain information outside the normal flow of the recurrent network so it is easier to plug two networks together
- LSTMs allow to preserve the error that can be backpropagated through time and layers which is a very important point for discrete forecasting models.

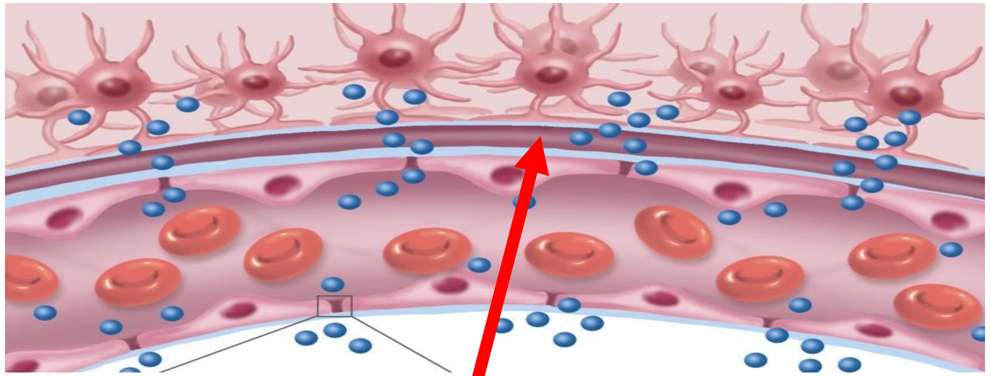
… Neural Assimilation

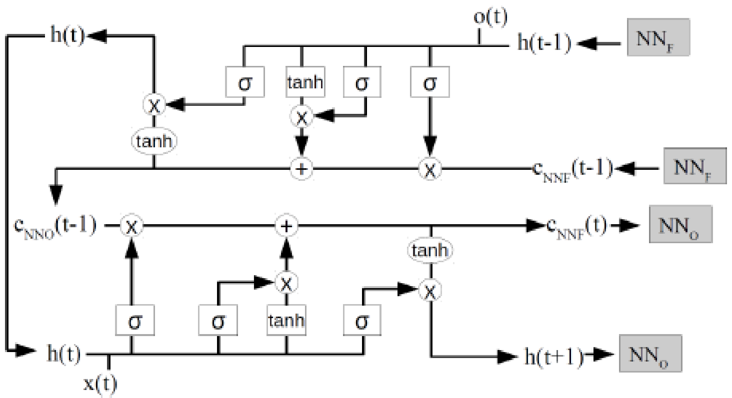
predict the oxygen diffusion across the Blood-Brain Barrier

$$\begin{cases} \dfrac{\partial x}{\partial t} = D\dfrac{\partial^2 x}{\partial y^2} \\ x(0, y) = x_{0,y} \\ x(t, 0) = x_{t,0} \\ x(t, L) = x_{t,L} \end{cases}$$

$$L = 400nm$$

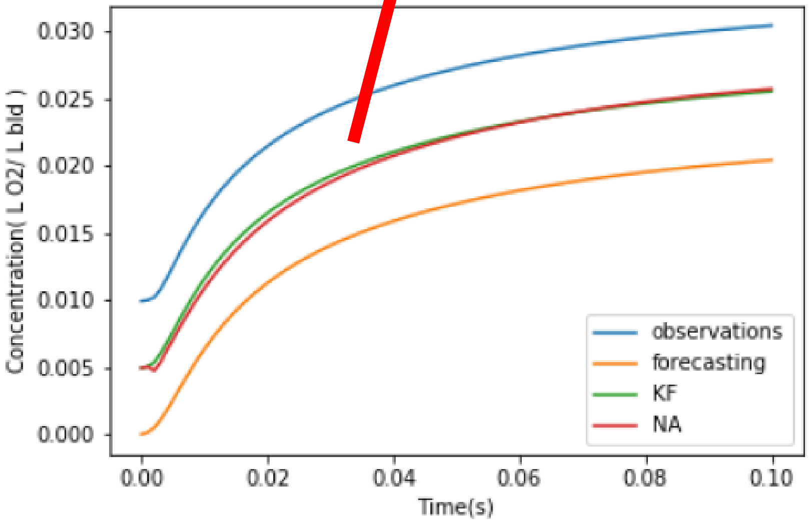
$$t \in [0, 10ms]\ (ms \text{ denotes microsecond})$$





NA has been trained using the 85% of the data and tested on the remaining 15%

Temporal evolution of the concentration at y = 12nm



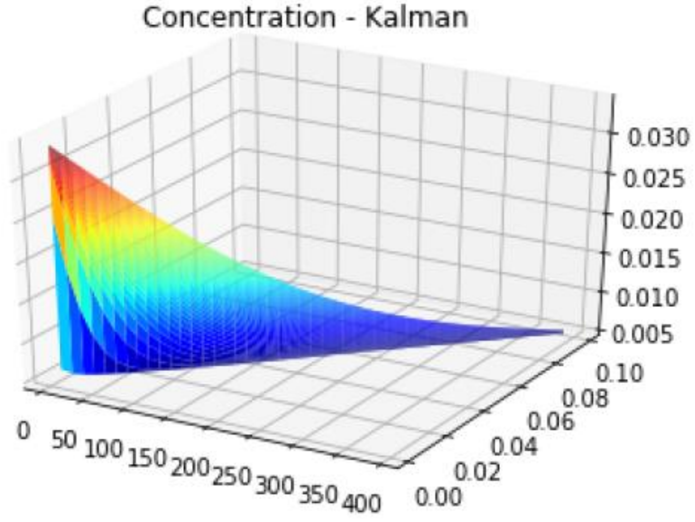
[*] R. Arcucci, L. Moutiq, Y. Guo - Neural Assimilation - Lecture Notes in Computer Science bookseries Volume 12142

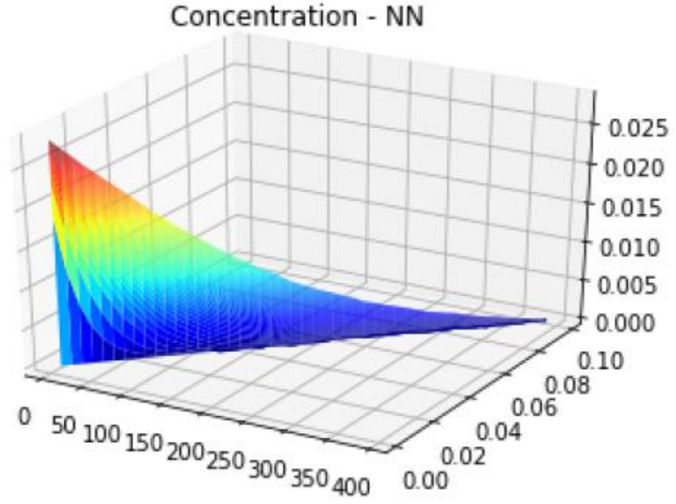
Neural Assimilation

predict the oxygen diffusion across the Blood-Brain Barrier

$$e_{NA}(t,y) = |z(t,y) - h(t,y)|$$

Time step t	$e_{NA}(t,y)$, $y = 12nm$	$e_{NA}(t,y)$, $y = 35nm$
0	0	0
10	7.05e-04	6.11e-04
20	4.17e-04	4.88e-04
30	4.29e-04	1.91e-04
40	1.52e-04	6.05e-07
50	2.51e-04	9.11e-05
60	3.40e-05	1.11e-04
70	4.13e-05	1.05e-04
80	4.72e-05	7.35e-05
90	1.11e-04	1.89e-05
100	1.60e-04	3.18e-05





Neural Assimilation

predict the oxygen diffusion across the Blood-Brain Barrier

$$e_{NA}(t, y) = |z(t, y) - h(t, y)|$$

Time step t	$e_{NA}(t, y)$, $y = 12nm$	$e_{NA}(t, y)$, $y = 35nm$
0	0	0
10	7.05e-04	6.11e-04
20	4.17e-04	4.88e-04
30	4.29e-04	1.91e-04
40	1.52e-04	6.05e-07
50	2.51e-04	9.11e-05
60	3.40e-05	1.11e-04
70	4.13e-05	1.05e-04
80	4.72e-05	7.35e-05
90	1.11e-04	1.89e-05
100	1.60e-04	3.18e-05

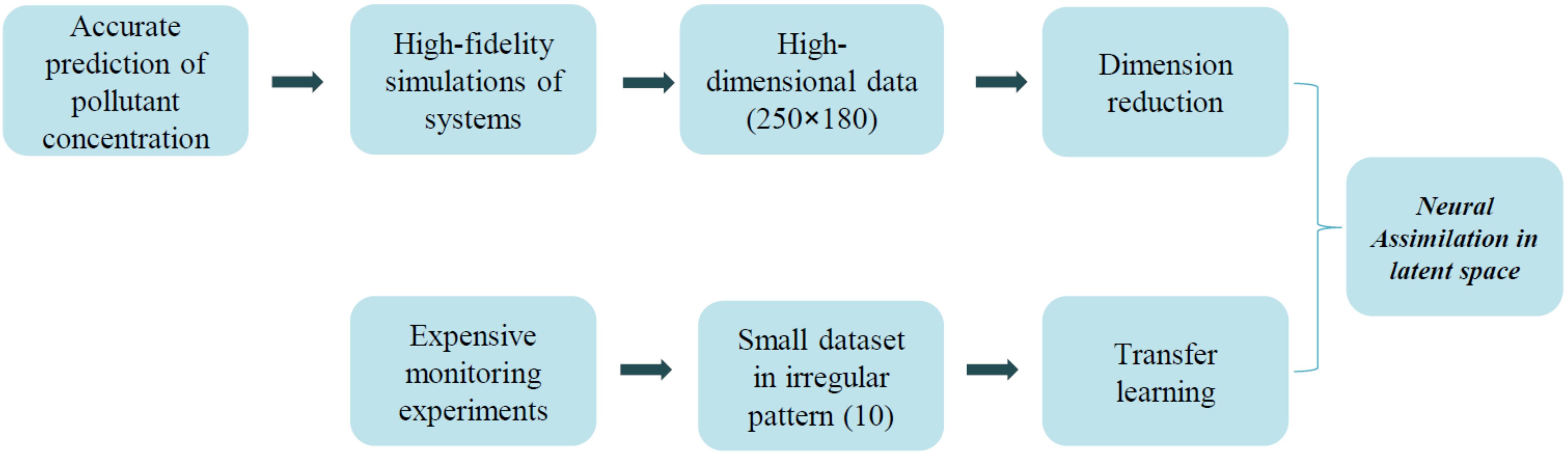
	Executing Time (s)
Neural Assimilation (training)	121.47
Neural Assimilation (prediction)	0.117
Kalman filter (prediction)	138

What if I want to add new time steps? … fine tuning!

Neural Assimilation

Pollutant concentration London (UK)

***with Yiwen Xu - 2020**



INHALE

MAGIC
Envisaging a world with greener cities

EPSRC
Engineering and Physical Sciences Research Council

The CFD software is FLUIDITY

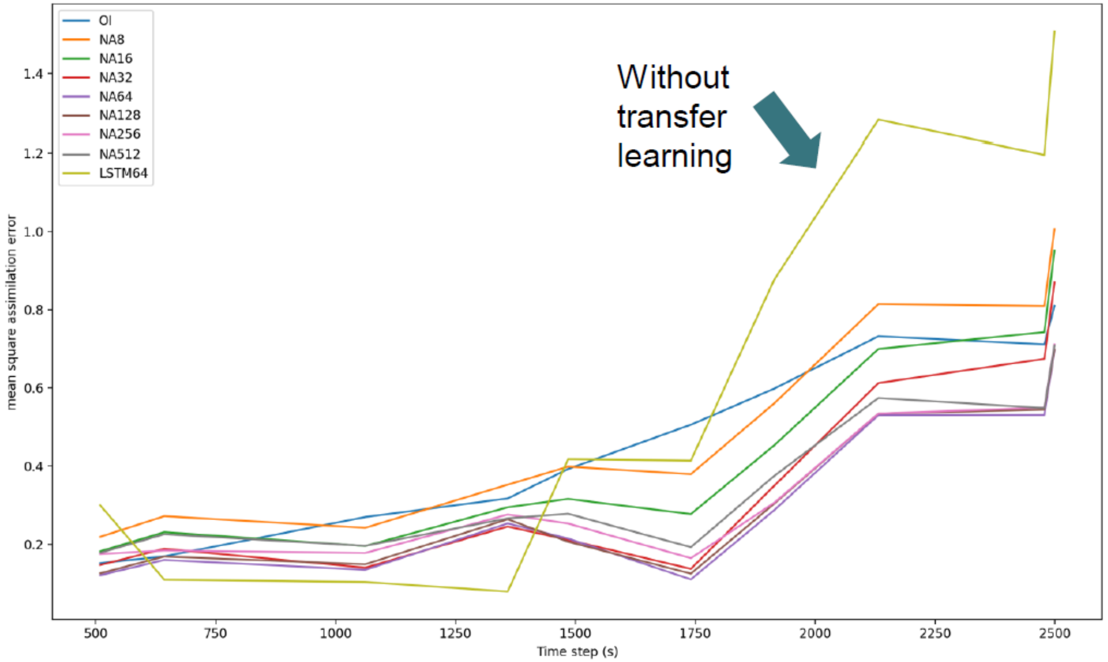
Neural Assimilation

Pollutant concentration London (UK)

***with Yiwen Xu - 2020**

reduced method	NN_F training(s)	NN_O training(s)	NA prediction(s)	error
8 PCA	73.04	10.64	0.032	0.39
16 PCA	69.87	9.53	0.032	0.33
32 PCA	69.43	9.54	0.032	0.26
64 PCA	70.44	9.51	0.033	0.24
128 PCA	71.39	9.12	0.033	0.26
256 PCA	73.46	11.68	0.033	0.27
512 PCA	79.85	16.4	0.034	0.28
8 PCAE	63.42	12.68	0.023	0.51
64 Autoencoder	112	9.09	0.014	0.15
128 Autoencoder	120	9.72	0.015	0.16
256 Autoencoder	108	7.86	0.014	0.53

Reference OI: 0.34
Reference OI: 267.43 (s) per time step

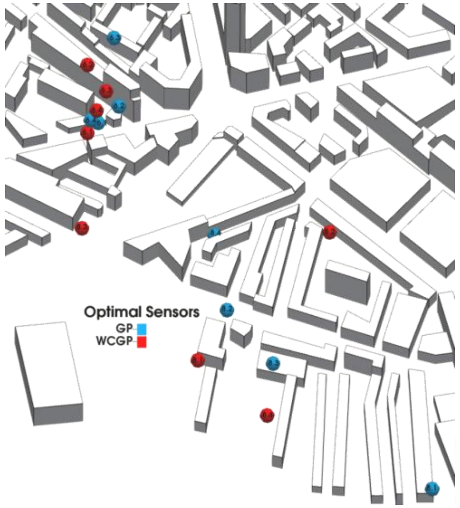


INHALE

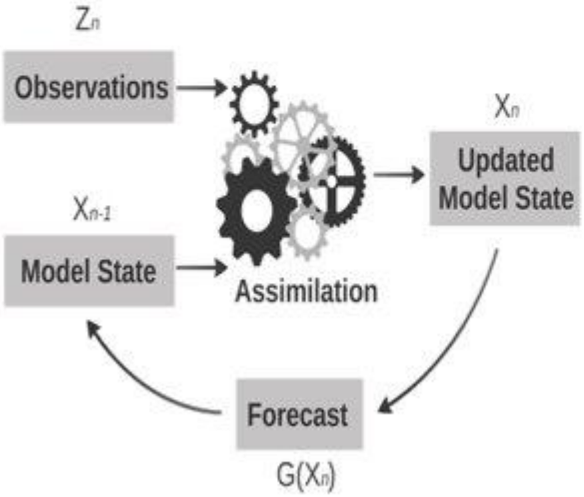
MAGIC
Envisaging a world with greener cities

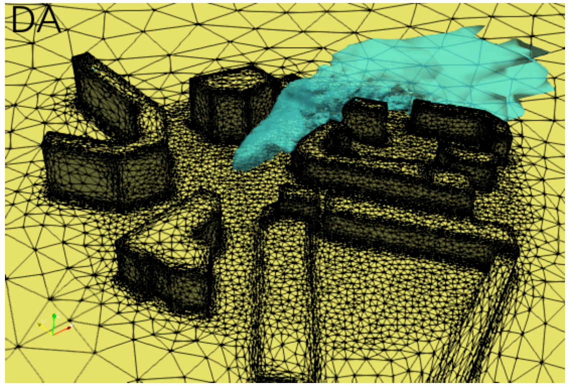
EPSRC
Engineering and Physical Sciences Research Council

Data Assimilation

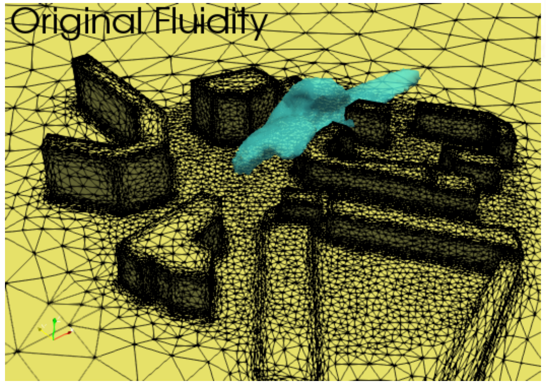


sensors





data assimilation

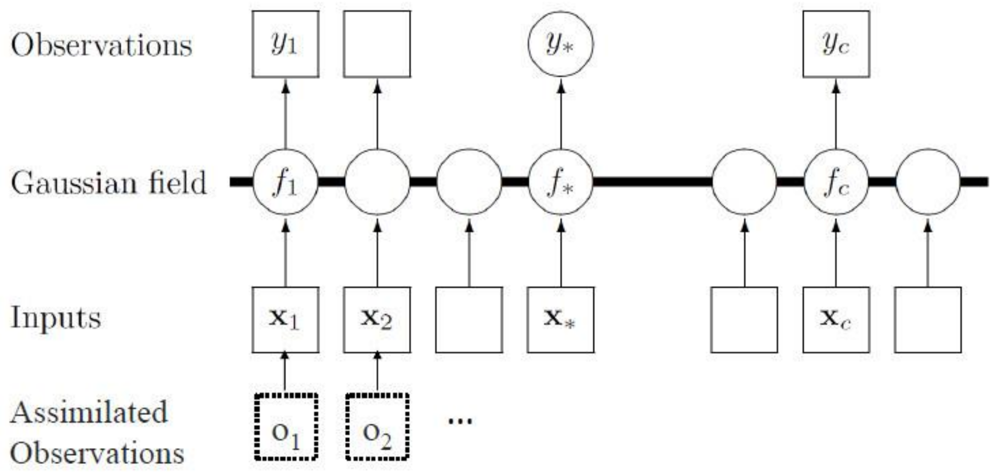


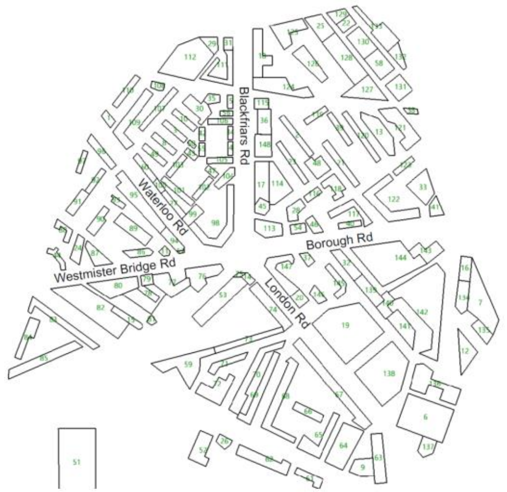
simulation

To have the expected benefit from this technology, we need:

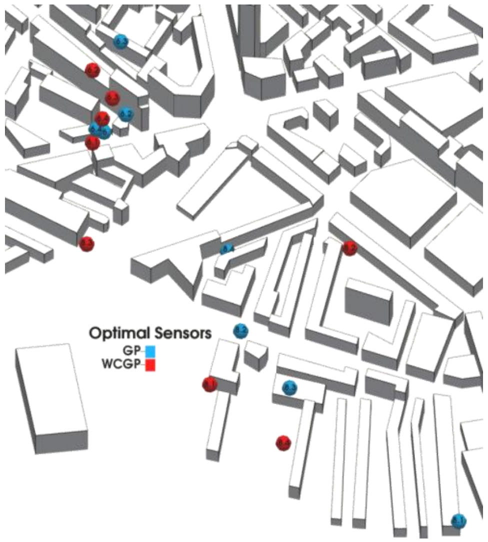
- **Good efficiency** (accelerate the execution time)
- **Good accuracy** (reduce the errors propagation in the models)
- **Real world scenarios** (develop models that can be used in real world scenarios – face **Big Data** problems)
- **Good quality data** (optimal locations)

DA+ML … for optimal sensor placement or optimal data collection





Outdoor... South London [*]



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 [**].

	Real Mean	Estimated Mean	MSE($\mathbf{x}^{\mathcal{M}}$)	MSE($\mathbf{x}^{\mathrm{DA}}$)
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

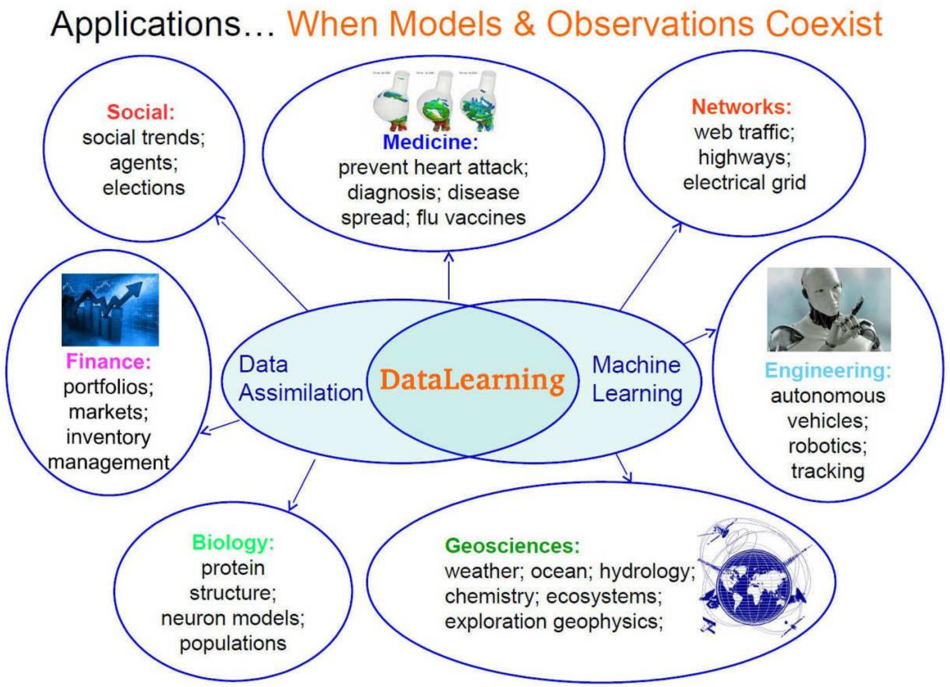
[*] T. Dur, R. Arcucci, L. Mottet, M. Molina Solana, C. Pain, Y. Guo - **Weak Constraint Gaussian Process for optimal sensor placement** - Journal of Computational Science vol 42,pag.101-110 DOI:10.1016/j.jocs.2020.101110

[**] G. Tajnafoi, R. Arcucci, L. Mottet, C. Vouriot, Molina Solana, C. Pain, Y. Guo - **Variational Gaussian Processes for optimal sensor placement** - Journal of Applied Mathematics

… we are happy to share

To get access to our **codes**:

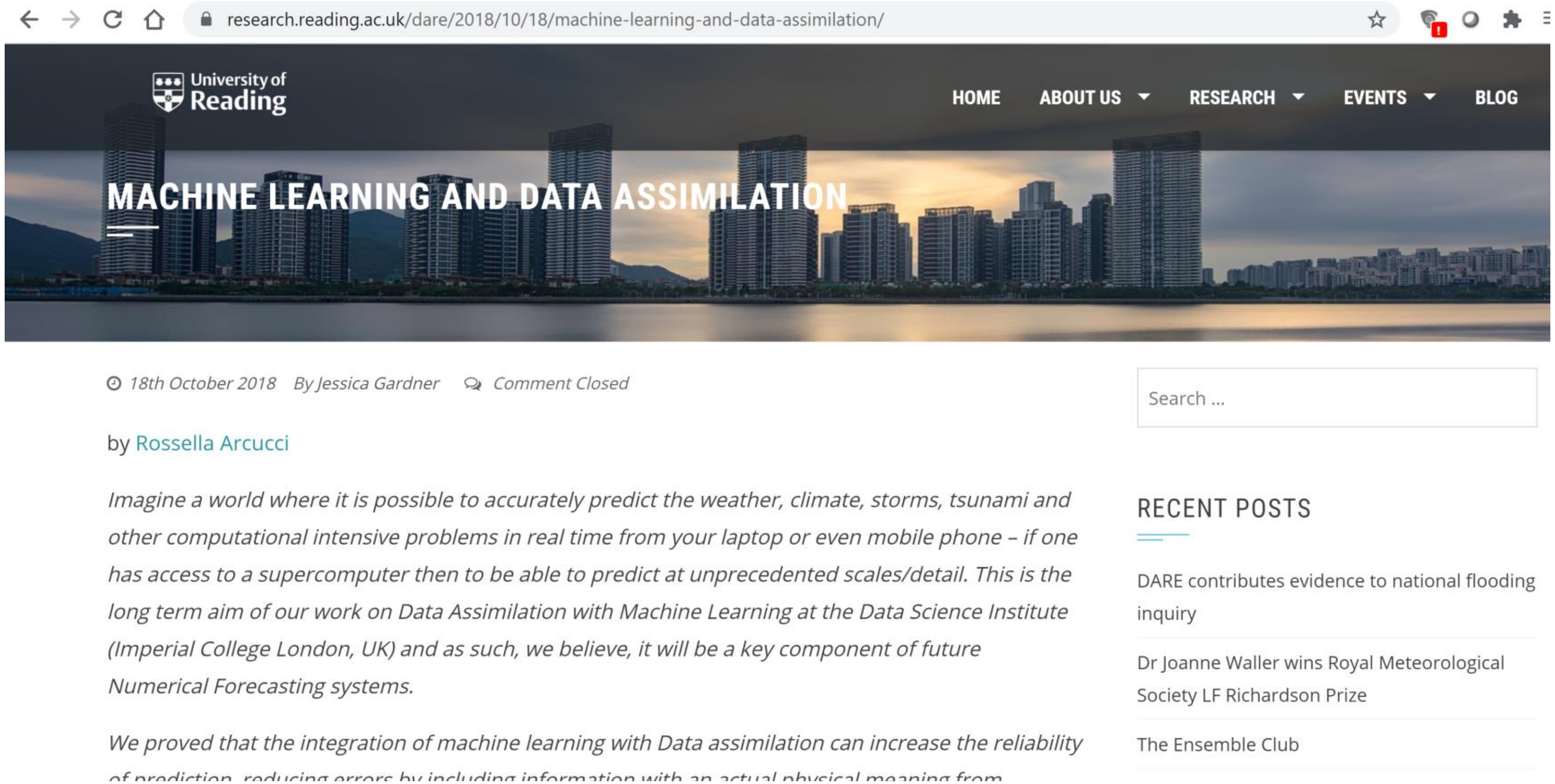
- Send me an email r.arcucci@imperial.ac.uk
- Google "DataLearning Data Science Institute Imperial College London"… you will find our DataLearning group https://www.imperial.ac.uk/data-science/research/research-themes/datalearning/



Imagine it, then do it!

The true sign of Intelligence is not Knowledge but Imagination. Albert Einstein

Intelligence / Artificial intelligence ? … Imagination is the key!


research.reading.ac.uk/dare/2018/10/18/machine-learning-and-data-assimilation/

University of Reading

HOME ABOUT US RESEARCH EVENTS BLOG

MACHINE LEARNING AND DATA ASSIMILATION

18th October 2018 By Jessica Gardner Comment Closed

by Rossella Arcucci

Imagine a world where it is possible to accurately predict the weather, climate, storms, tsunami and other computational intensive problems in real time from your laptop or even mobile phone – if one has access to a supercomputer then to be able to predict at unprecedented scales/detail. This is the long term aim of our work on Data Assimilation with Machine Learning at the Data Science Institute (Imperial College London, UK) and as such, we believe, it will be a key component of future Numerical Forecasting systems.

We proved that the integration of machine learning with Data assimilation can increase the reliability of prediction, reducing errors by including information with an actual physical meaning from

Search ...

RECENT POSTS

DARE contributes evidence to national flooding inquiry

Dr Joanne Waller wins Royal Meteorological Society LF Richardson Prize

The Ensemble Club

https://research.reading.ac.uk/dare/2018/10/18/machine-learning-and-data-assimilation/

Thank You

r.arcucci@imperial.ac.uk

Data Assimilation with Machine Learning…

www.imperial.ac.uk/people/r.arcucci/publications

- J. Zhu, S. Hu, R. Arcucci, C. Xu, J. Zhu and Y. Guo - Model Error Correction in Data Assimilation by Integrating Neural Networks - Journal of Big Data Mining and Analytics Vol 2 (Issue 2) pag. 83 – 91
- R. Arcucci, L. Mottet, C. Pain and Y. Guo - Optimal reduced space for Variational Data Assimilation -Journal of Computational Physics, Vol 379, pag: 51-69
- 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 vol 42, pag.101-110
- P. Wu, X. Chang, W. Zhang, R. Arcucci, Y. Guo, C. Pain, Data-driven reduced order model with temporal convolutional neural network - Computer Methods in Applied Mechanics and Engineering, Volume 360,
- R. Arcucci, D. McIlwraith and Y. Guo - Scalable Weak Constraint Gaussian Processes - Lecture Notes in Computer Science book series (ICCS 2019) vol. 11539, pag. 111-125
- E. Aristodemou, R. Arcucci, L. Mottet, A. Robins, C. Pain, Y. Guo, Enhancing CFD-LES air pollution predictions using data assimilation - Journal of Building and Environment, Volume 165
- R. Arcucci, C. Pain, Y. Guo, Effective variational data assimilation in air-pollution prediction, Big Data Mining and Analytics, Vol 1, Issue 4 pag: 297 - 307, 2018
- E. Lim, R. Arcucci, M. Molina Solana, C. Pain, Y. Guo - Hybrid Data Assimilation: an Ensemble-Variational Approach - 15th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS), IEEE
- R. Arcucci- Effective Data Assimilation with Machine Learning - Data Science Book 2020
- J. Mack, R. Arcucci, M. Molina, Y. Guo - Attention-based Convolutional Autoencoders for 3D-VariationalData Assimilation - Journal Computer Methods in Applied Mechanics and
- C. Quilodran Casas, R. Arcucci, P. Wu, C. Pain, Y. Guo - A Reduced Order Deep Data Assimilation model - Journal Physica D: nonlinear phenomena
- G. Tajnafoi, R. Arcucci, L. Mottet, C. Vouriot, Molina Solana, C. Pain, Y. Guo - Variational Gaussian Processes for optimal sensor placement - Journal of Applied Mathematics
- …