Imperial College London

Wildfire project

Applying Data Science

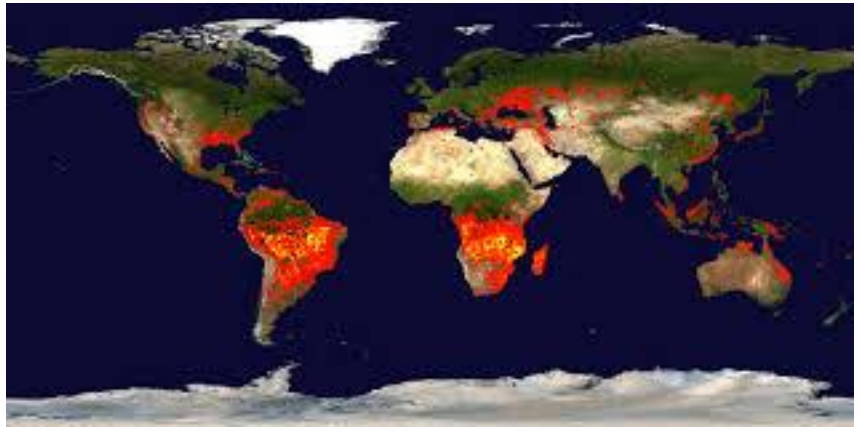
Deadlines:

Code: Friday, 24th May 2024 at **12:00** BST

Presentation: Friday, 24th May 2024 at **16:00** BST

WILDFIRE PROJECT OVERVIEW

Predicting wildfires is essential for protecting human life, property, and the environment. It plays a critical role in disaster preparedness and response, enabling proactive measures to mitigate the devastating impacts of wildfires. As the number of wildfires increases with climate change, the development and implementation of advanced predictive models become even more crucial for ensuring safety, sustainability, and resilience in vulnerable regions.



The goal of this project is to develop a comprehensive system that predicts a wildfire behavior using Recurrent Neural Networks (RNN), Generative AI, and data assimilation techniques. This system will use historical wildfire data, and satellite imagery feeds, to enhance prediction accuracy and provide actionable insights.

DATA: FERGUSON WILDFIRE

- Model data (already pre-processed):
  - Ferguson_fire_train: training data obtained from wildfires simulations

    Source Ferguson_fire_train.zip
  - Ferguson_fire_test: similar to Ferguson_fire_train but obtained from different simulations

    Source Ferguson_fire_test.zip
  - Ferguson_fire_background: model data to be used for the data assimilation:

    Source Ferguson_fire_background.zip

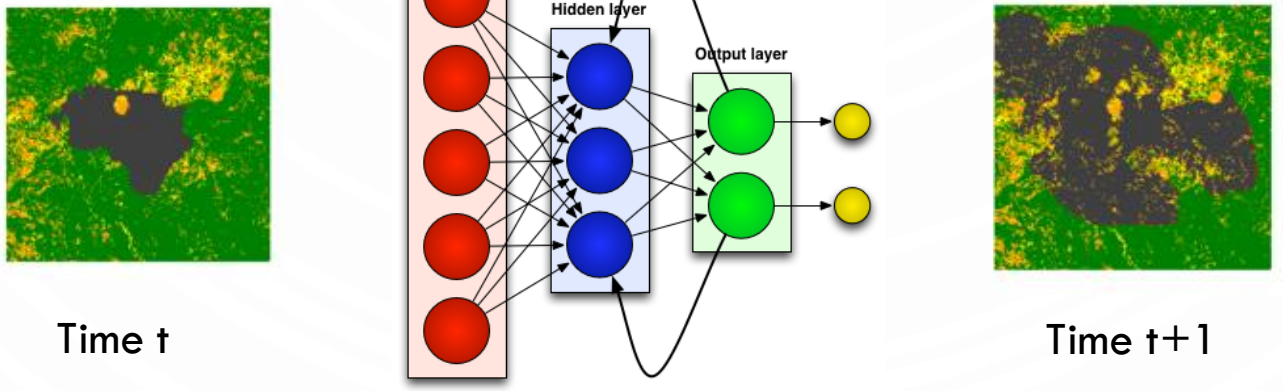
- Satellite data (already pre-processed):
  - Ferguson_fire_obs: Observation data at different days after ignition (only one trajectory)

    Source Ferguson_fire_obs.npy

Note: in the background file you will find model data already selected in time steps corresponding to satellites observed data.

OBJECTIVE 1

- **Build a SURROGATE MODEL using RNN** - Use a recurrent neural network to train a surrogate model of a wildfires predictive model.
    - **Step 1**: Use the model data for training (Ferguson_fire_train) and test the model using the model data for testing (Ferguson_fire_test).

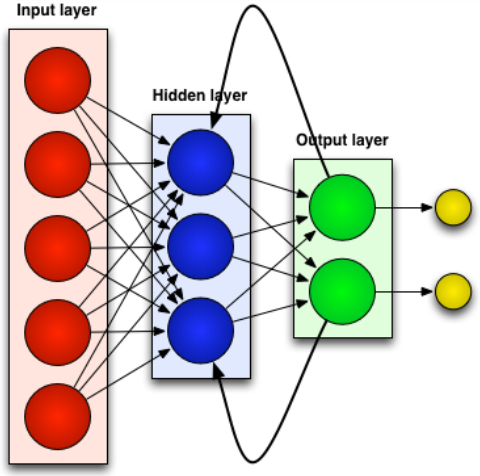


<u>Note: you can use any recurrent neural network. LSTM could be a good option.</u>

OBJECTIVE 1

- **Build a SURROGATE MODEL using RNN** - Use a recurrent neural network to train a surrogate model of a wildfires predictive model.
    - **Step 2**: Use the RNN model with background data (Ferguson_fire_background) to make a forecast.

Time t

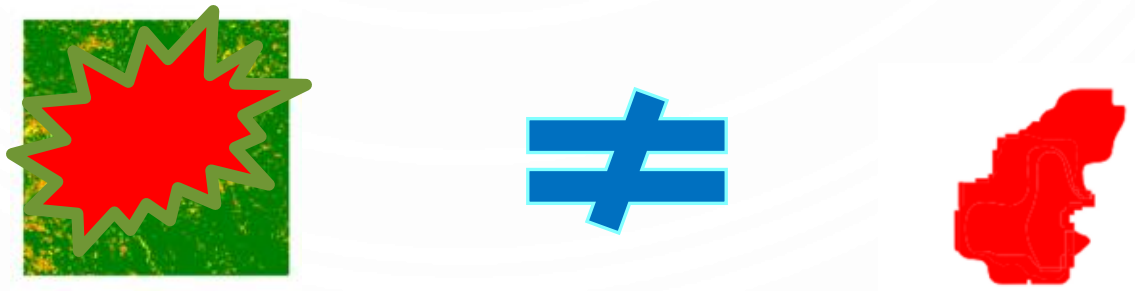


Time t+1

What does it look like?

OBJECTIVE 1

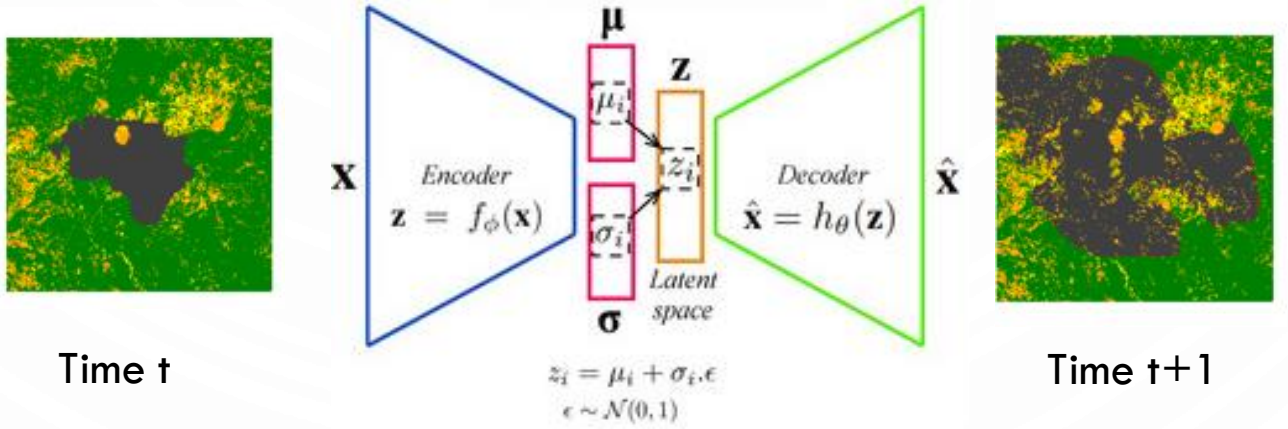
- **Build a SURROGATE MODEL using RNN** - Use a recurrent neural network to train a surrogate model of a wildfires predictive model.
  - **Step 3**: compare your forecasted results with satellite data (Ferguson_fire_obs) and compute the MSE between your forecast and the satellite data. Submit both code (all three steps in the same notebook) and MSE.



Time t+1

OBJECTIVE 2

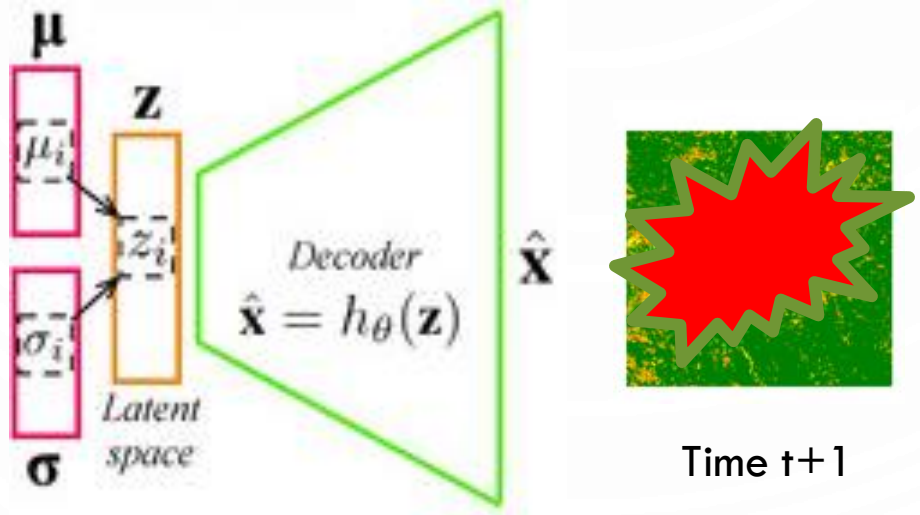
- **Build a SURROGATE MODEL using GENERATIVE AI** - Use a generative AI method to train a wildfires generative model.
    - **Step 1**: Use the model data for training (Ferguson_fire_train) and test the model using the model data for testing (Ferguson_fire_test).



Note: you can use any generative model. Variational AE could be a good option.

OBJECTIVE 2

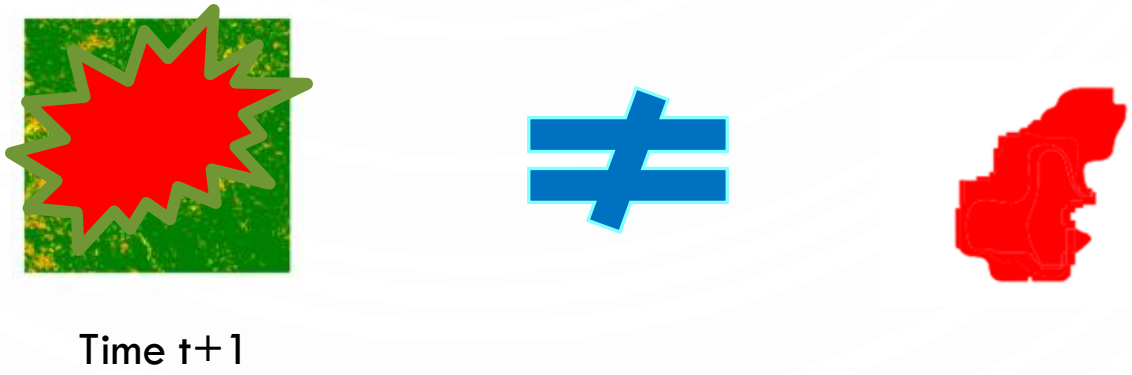
- **Build a SURROGATE MODEL using GENERATIVE AI -** Use a generative AI method to train a wildfires generative model.
  - **Step 2**: Use the wildfire generative model to make a forecast.



What does it look like?

OBJECTIVE 2

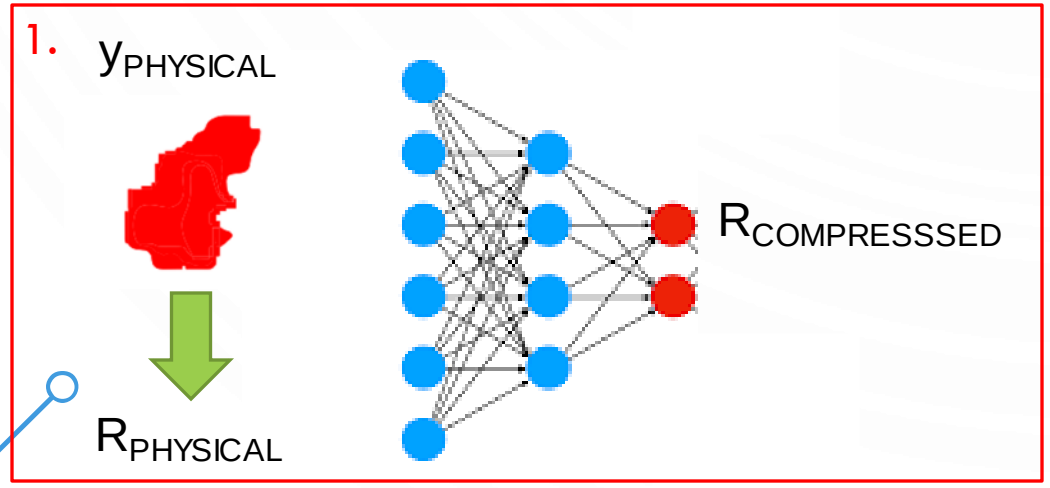
- **Build a SURROGATE MODEL using GENERATIVE AI** - Use a generative AI method to train a wildfires generative model
  - **Step 3**: compare your forecasted results with satellite data (Ferguson_fire_obs) and compute the MSE between your forecast and the satellite data. Submit both code (all 3 steps in the same notebook) and MSE.

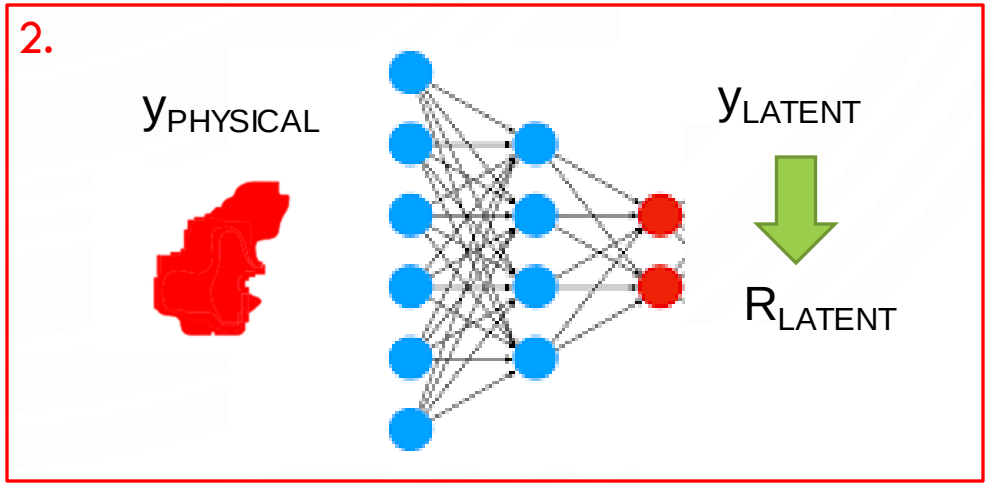


Time t+1

OBJECTIVE 3

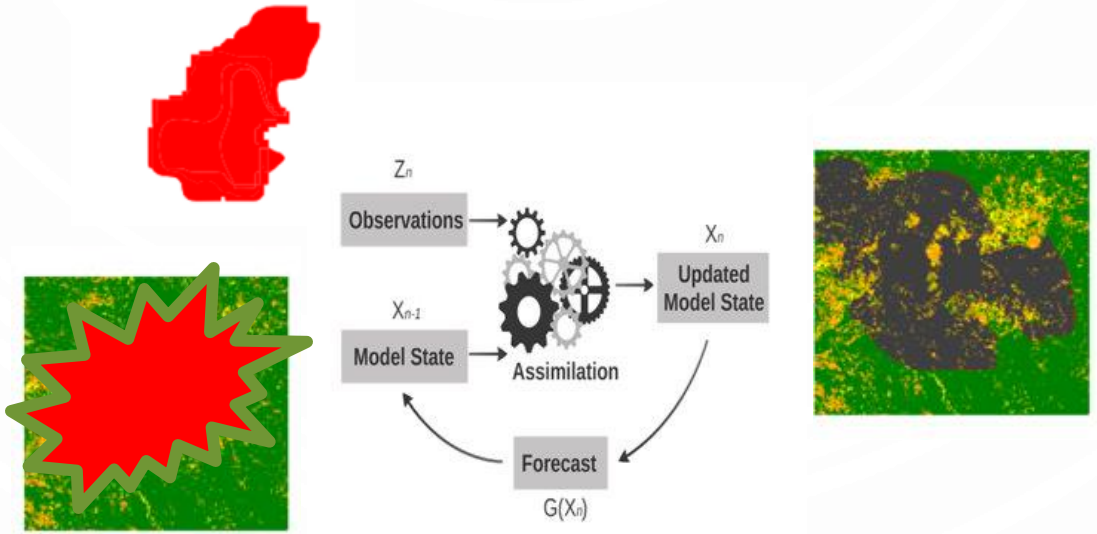
- **CORRECTION using Data Assimilation** - Perform Data Assimilation with the results of your RNN model and your wildfires generative model. Data Assimilation **must be performed in a reduced space**.
  - **Step 1**: Compute the error covariance matrices for the background data (matrix B) and satellite data (matrix R) in the data assimilation model. The observation error covariance matrix R is usually diagonal. This time it must be computed using satellite data and then used to perform DA in a reduced space. There are two main strategies:
    1. **compute R in the physical space and then compress it**
    2. **compress the data and compute R in the compressed space**





OBJECTIVE 3

- **CORRECTION using Data Assimilation** - Perform Data Assimilation with the results of your RNN model and your Generative model. Data Assimilation must be performed in a reduced space.
    - **Step 2**: perform **data assimilation** in a reduced space using satellite data (Ferguson_fire_obs) and background data (Ferguson_fire_background) and submit both code and MSE before and after performing DA.



Note: in the folder with the background files you will have the model data already selected in time steps corresponding to the observations.

AI FOR NATURAL HAZARDS

EVERY SCIENTIFIC CONTRIBUTION CAN HELP SAVE MILLIONS OF LIVES DAILY.

THANK YOU FOR YOUR EFFORTS TODAY.