Using optimisation method for input data selection in AI-based time-series flood forecasting models

UNIVERSITY OF WEST LONDON
The Career University

Using Optimisation Method for Input Data Selection in AI-Based Time-Series Flood Forecasting Models

IIOT research group

Principal supervisor
Dr Kourosh Behzadian

Secondary supervisor
Prof. Amir Alani

Farzad Piadeh
PhD Student of civil engineering

July 2022

CONTENTS

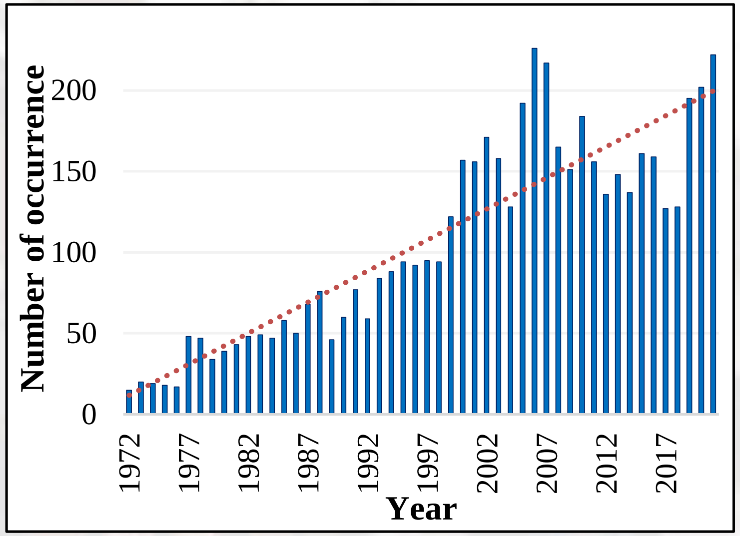
Introduction

In recent 50 years, floods:

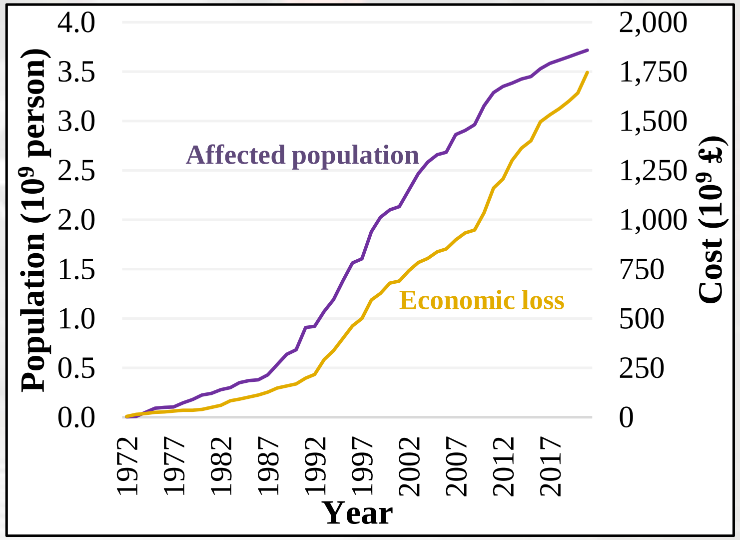
❖ **1,750 £ billion** economy damages

❖ **3.7 billion people** are affected

❖ **235,000 people** are killed



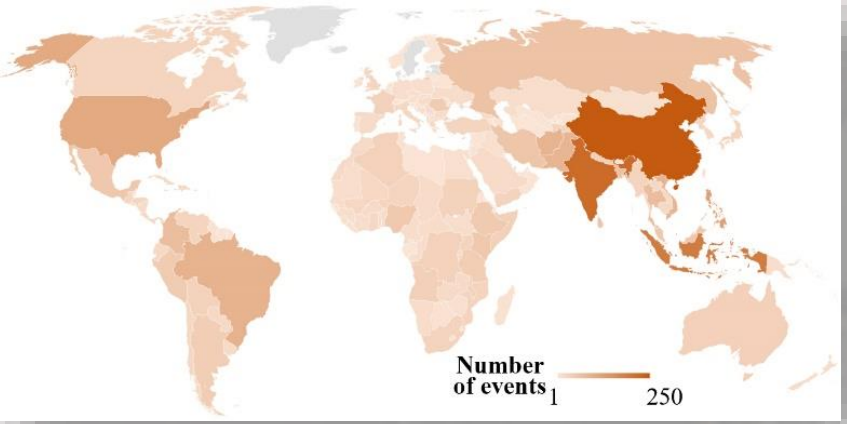
Number of flood occurrences



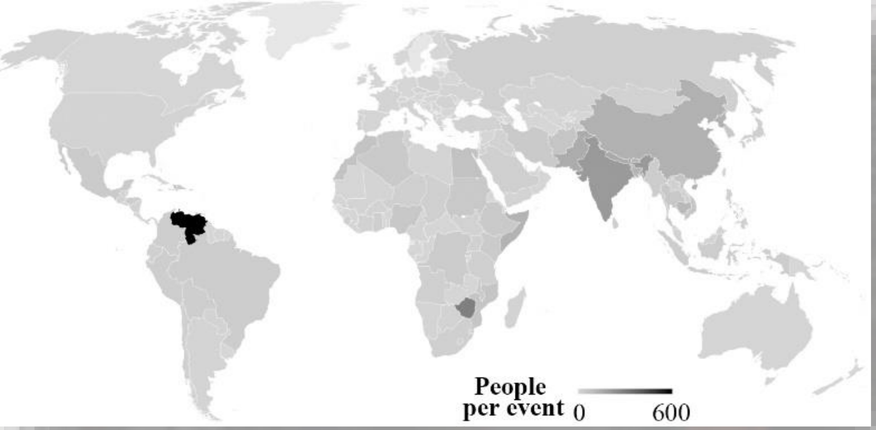
Cumulative social and economic loss

"Multi-Step Flood Prediction in Drainage Systems Using Time-series Data Mining Techniques", **Piadeh F.**, Behzadian K. Alani A.M., Water Efficiency Conference, West Indies, Trinidad and Tobago, 2022. [Under review]

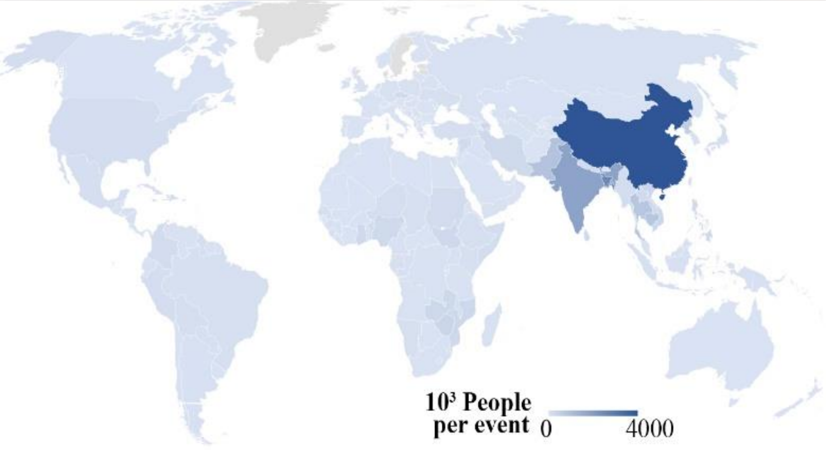
Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk



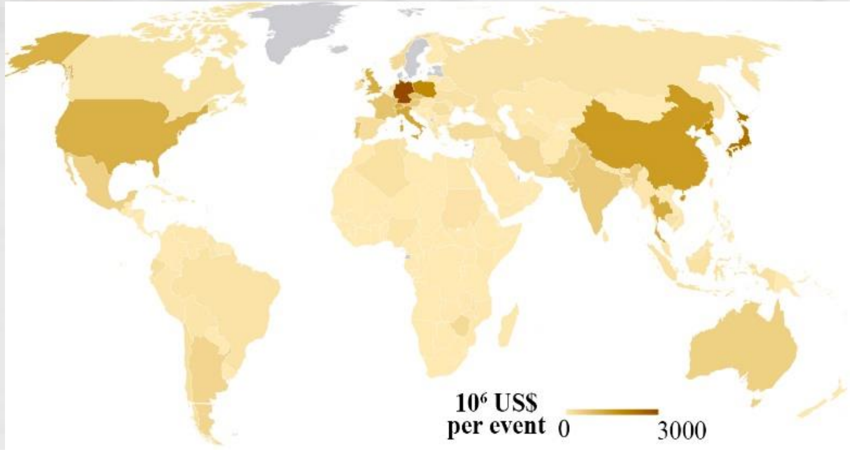
Number of flood events



Average human loss



Average affected people



Average economic loss

“A Critical Review of Real-Time Modelling of Flood Forecasting in Urban Drainage Systems”, **Piadeh F.**, Behzadian K. Alani A.M., *Journal of Hydrology*, 2022; 607: 127476

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

Characteristics	Drainage systems	
	Urban areas	Non-urban areas
Flood description	- Overflow of urban drainage infrastructures due to lack of proper drainage in an urban area	- Overflow or rise of water bodies such as rivers, streams, sea level and reservoirs
Flood causalities	- Mainly fast surface runoff generated by rainfall	- Mainly high intensity of rainfall or accumulation of surface runoff
Flood duration	- Between a few minutes to a couple of days	- Part of days to a week
Spatial flood impacts	- Small areas i.e. streets to neighbourhoods, can be extended to all urban areas, but highly distributed	- Large scale such as vulnerable zones, and river riparian zones
Spatial restrictions for flood management	- No flexibility in land surfaces or underground modification as previously occupied. - Fast variation in land use	- High flexibility in non-urban areas
Main types of impacts	- Economic loss and business interruption - Human loss, - Mental and social problems - Urban structure and infrastructure damages	- Soil erosion - Wasting crops and livestock - Natural habitat loss - Water pollution - Reservoir or water infrastructure damages

“A Critical Review of Real-Time Modelling of Flood Forecasting in Urban Drainage Systems”, **Piadeh F.**, Behzadian K. Alani A.M., *Journal of Hydrology*, 2022; 607: 127476

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

Flood Event

Mitigation
addressing to "**plan and design**" of proper strategies or actions, including predictions, and determining suitable constructions.

Response
focusing on maintaining designed plans by "implementation" of proper actions, particularly public awareness or running green-blue infrastructures.

Recovery
"evaluating" designed plans and related responses, to correct future planning.

Structural approach
Designing suitable construction facilities for water balance and flood source control.

+

Non-structural approach
Excluding policy and strategy creation, this part **predicts** event's spreads before occurring.

Rainfall forecasting
- What is the nature of rainfall?
- When is it coming?
- Where is it coming?

Flood forecasting
- What does happen in affected areas?

Detection instruments
Emphasising on upgrading performance of instruments and devices

Better prediction methods
Tackling more accurate spatial and temporal prediction

- **Empirical models**
- **Conceptual models**
- **Physical models**

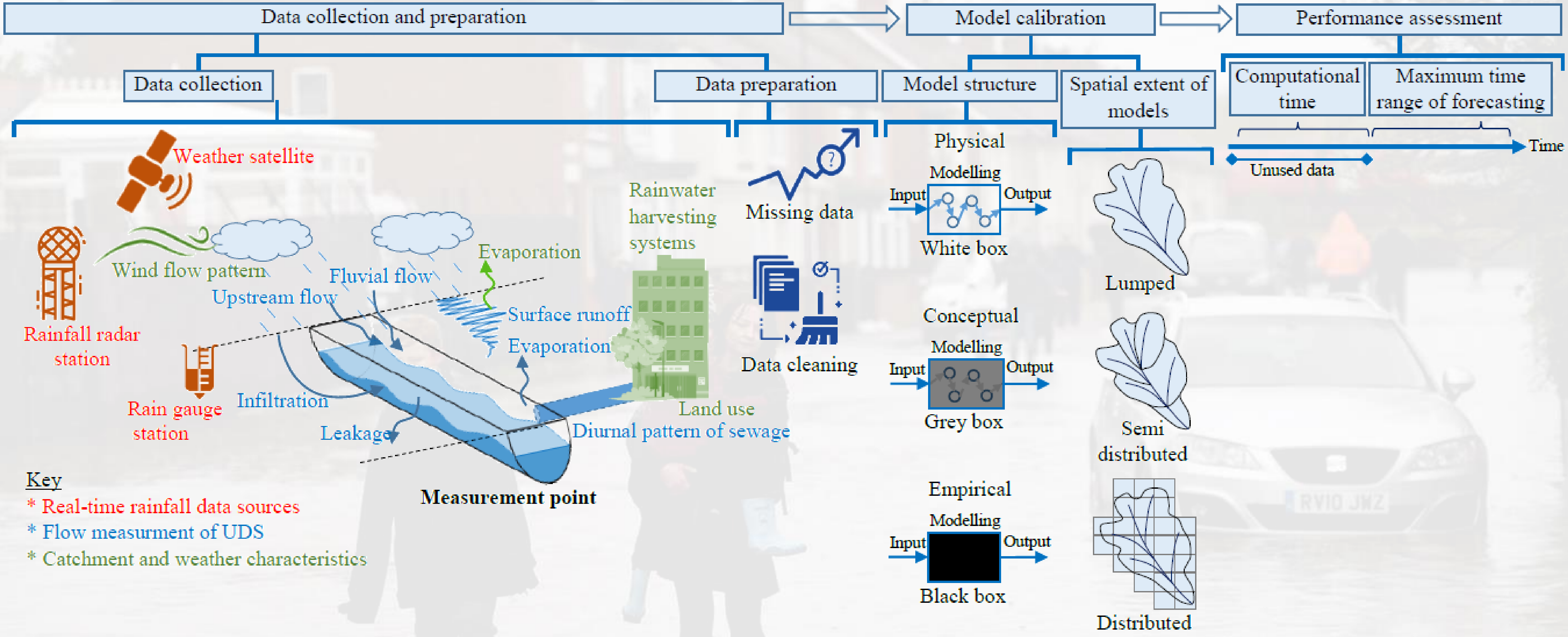
Artificial intelligence-based model

Artificial Neural Network (NN) models

Time-series real-time NN

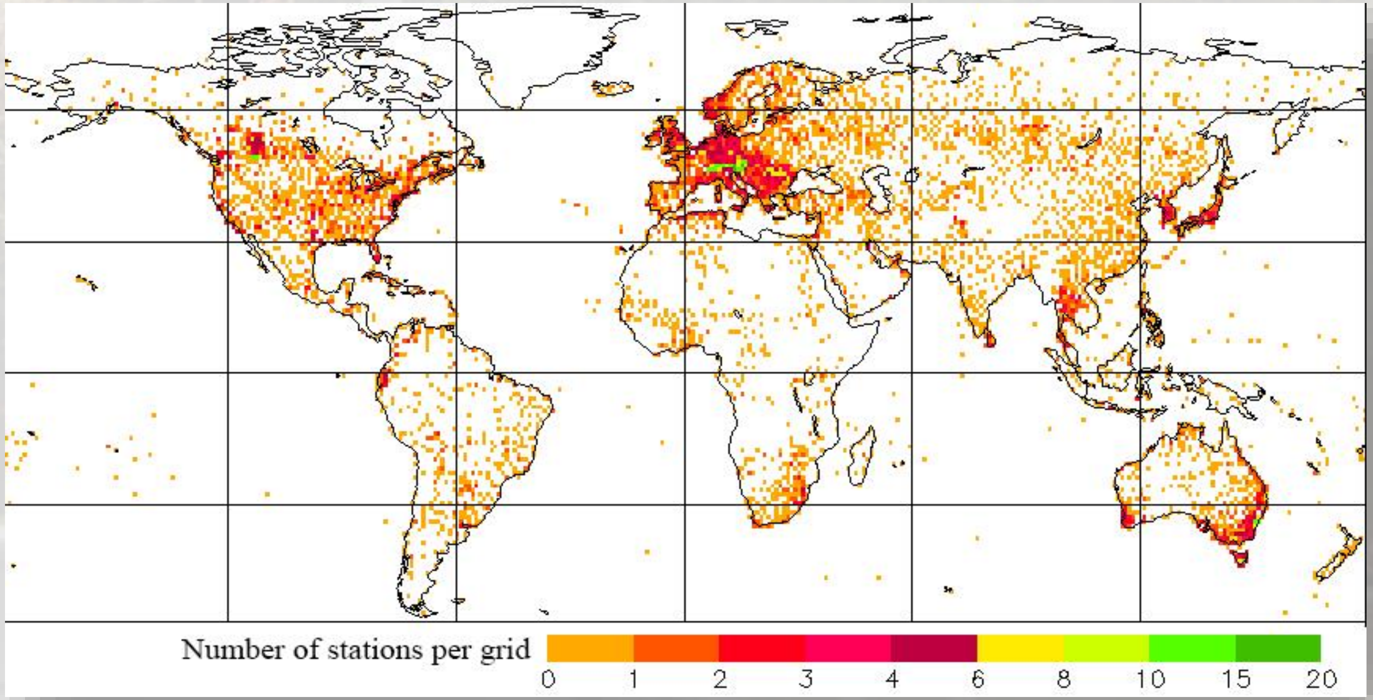
"The Role of Event Identification in Translating Performance Assessment of Time-Series Real-Time Urban Flood Forecasting", **Piadeh F.**, Behzadian K. Alani A.M., 15th UWL Doctorial Conference, London, UK, 2021. [Poster presentation]

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk



"A Critical Review of Real-Time Modelling of Flood Forecasting in Urban Drainage Systems", **Piadeh F.**, Behzadian K. Alani A.M., *Journal of Hydrology*, 2022; 607: 127476

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

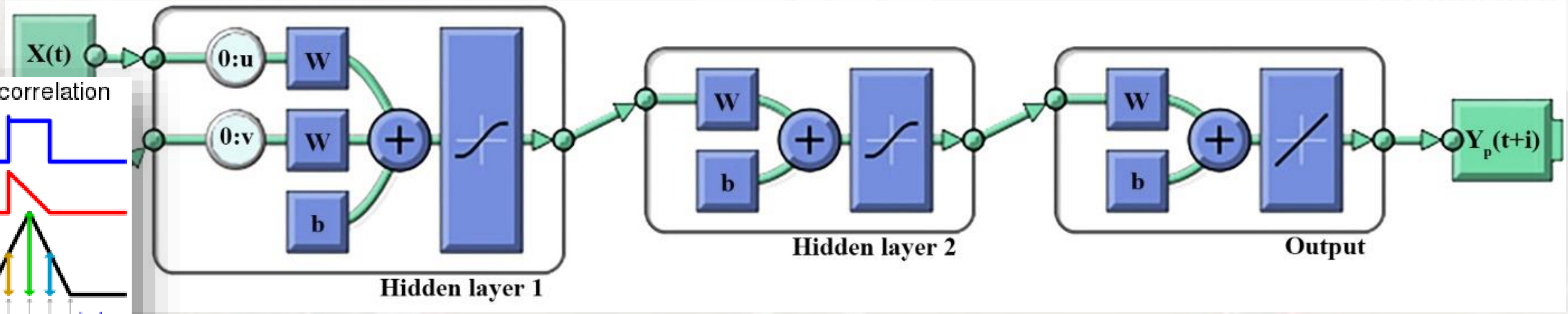


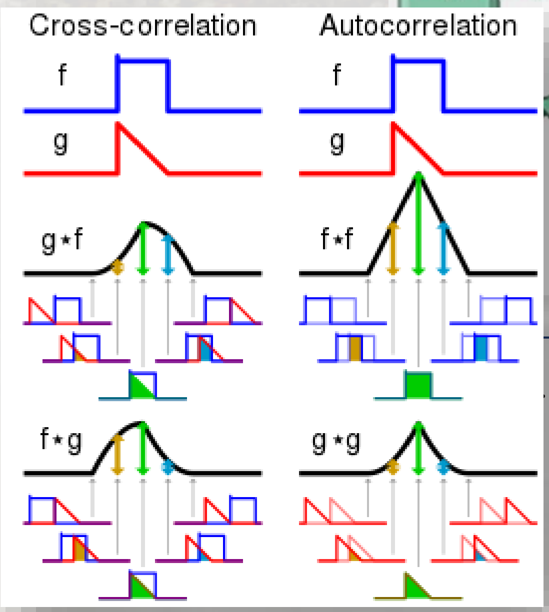
Global distribution of installed rain gauge stations

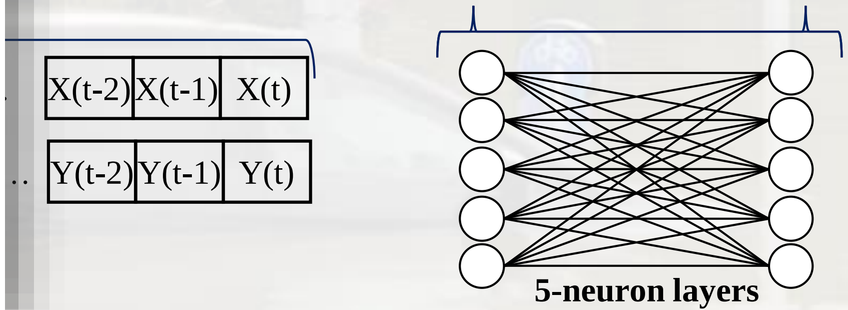
National Centre for Atmospheric Research (NCAR). (2012). *Number of stations used by GPC for May 2012.* [Online]. Available at https://climatedataguide.ucar.edu, [Accessed 8 Jan. 2022].

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

Recommendation for range of input data: Cross correlation and Cross covariance







Key

X: Observant rainfall intensity

Y: Observant water depth

Y_p: Predicted water depth

u: Delay factor (range of pervious X data)

v: Feedback factor (range of pervious Y)

w: Weight value

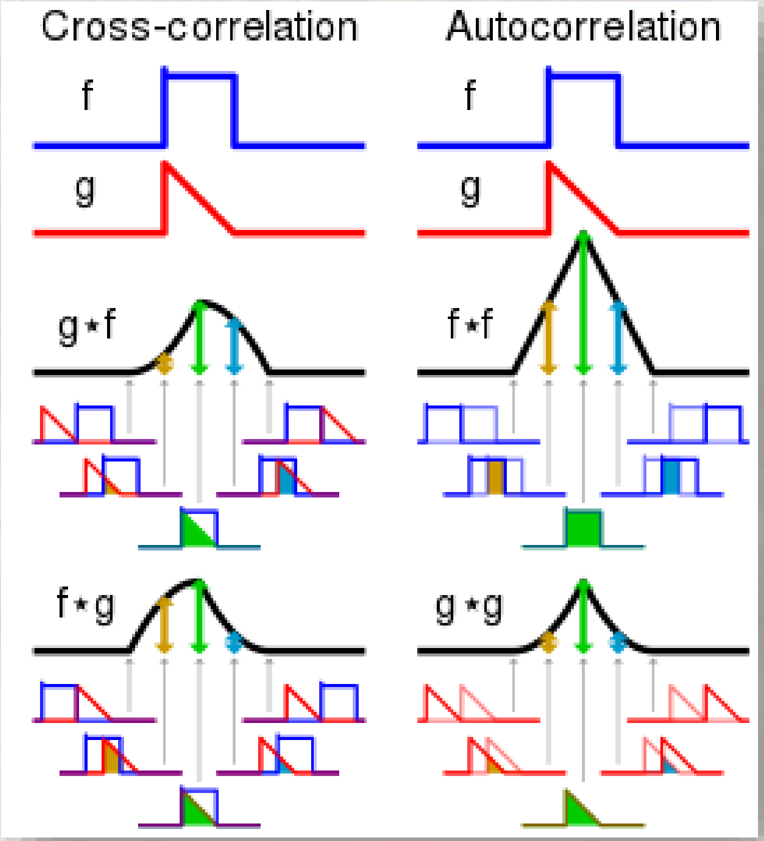
b: Bias value

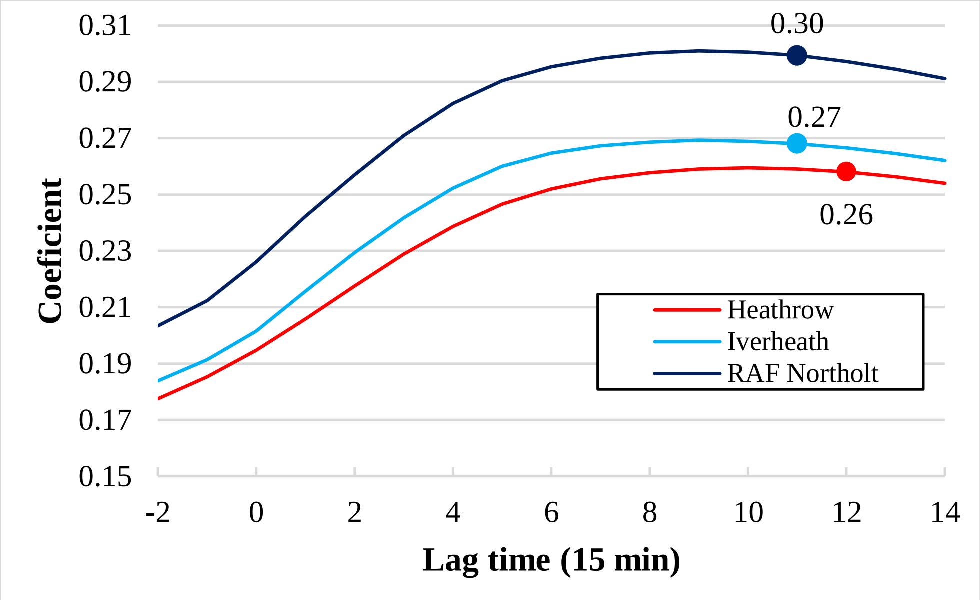
i: Time-step ahead

Linkedin.com/in/farzad-piadeh-b9313450

Email: 21452390@student.uwl.ac.uk

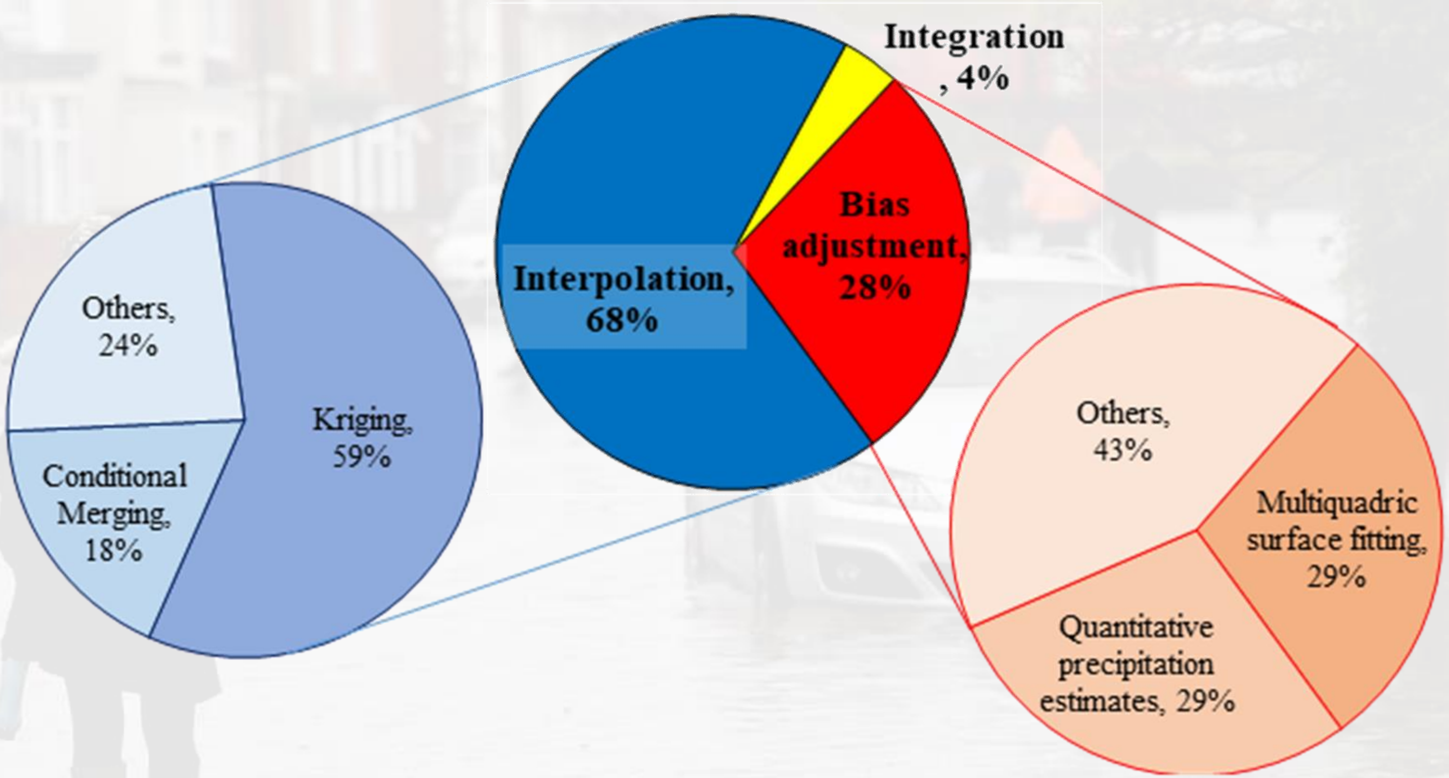
Recommendation for range of input data: Cross correlation and Cross covariance





Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

- Integration techniques have more capability to increase the model accuracy than other techniques
- Integration has not much interest due mainly to requirements for :
    - More model complexity
    - More data records
    - Higher computational efforts



"A Critical Review of Real-Time Modelling of Flood Forecasting in Urban Drainage Systems", **Piadeh F.**, Behzadian K. Alani A.M., *Journal of Hydrology*, 2022; 607: 127476

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

Research question:

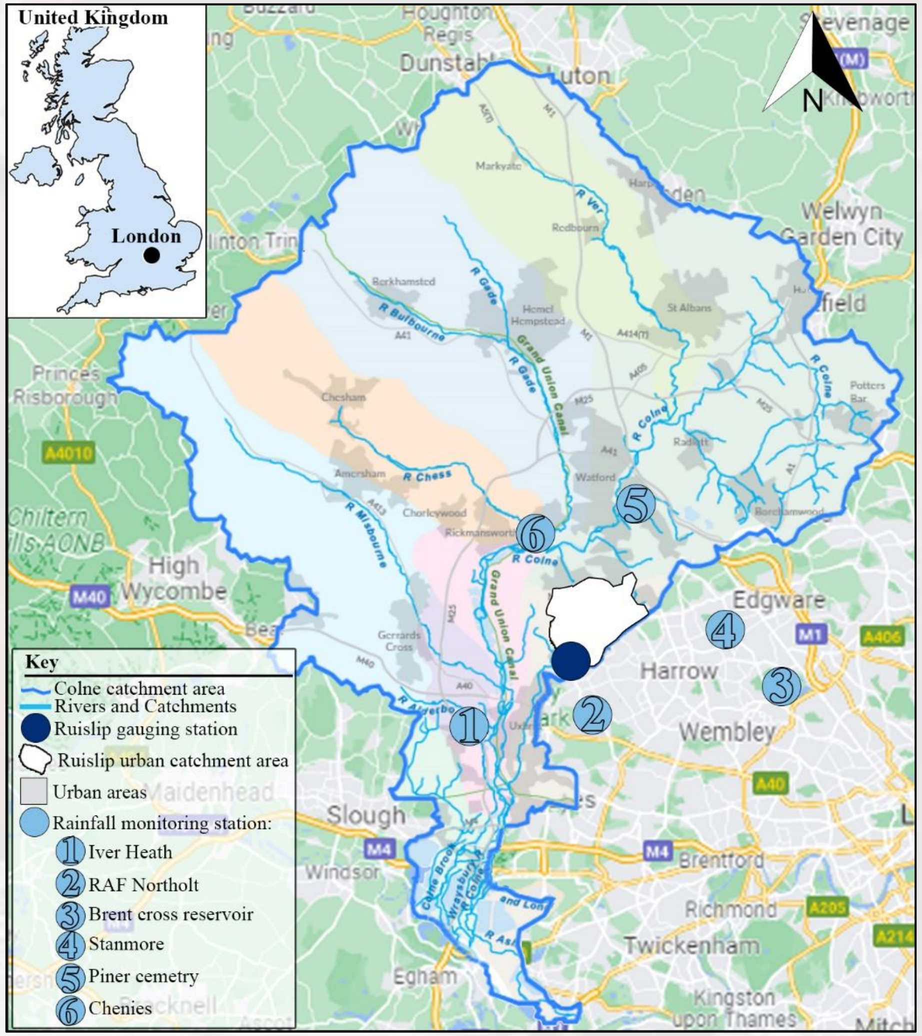
- Which type of data merging is more effective?
- How many of data should be add to the time-series model for each iteration?
- Recommendations are compatible with optimisation of used range of data?

“A Critical Review of Real-Time Modelling of Flood Forecasting in Urban Drainage Systems”, **Piadeh F.**, Behzadian K. Alani A.M., *Journal of Hydrology*, 2022; 607: 127476

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

Method and material

❖ *Case study description*



Location of the case study

Inputs:

- **Rainfall data**
- **Water level**

Output: Water level for lead time

Selected model: NARX (The nonlinear autoregressive network with exogenous inputs)

Pre-phase: Testing the model for 4-step ahead:

- **Best correlated rainfall data and water level**
- **Interpolating all rainfall data (Kriging with external draft)**
- **Bias adjustment of best rainfall data with others (Multiquadric fitting)**
- **Integration of all rainfall data**

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

❖ *Method and concept*

Data collection and preparation

- Rainfall data of 6 triple bucket rain gauge stations (X)
- Water level data of ultrasonic gauge station (X)
- Linear regression for infilling missing data

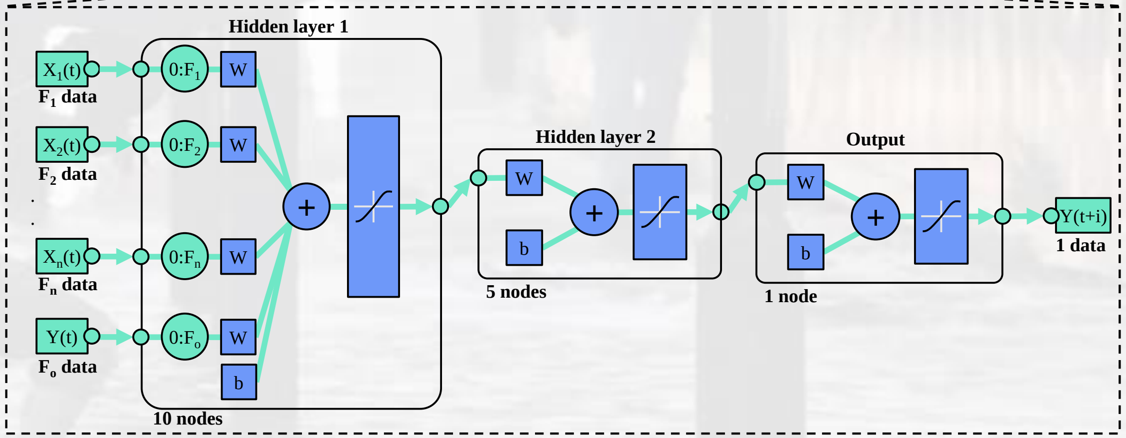
Model development

- Cross-covariance between “X” and “Y” data
- Auto-correlation between “Y” and “Y” data
- Cross-correlation between “X” and “Y” data
- Using shuffled frog leaping optimisation algorithm

- Conventional model with one rainfall station
- Conventional model with 3 rainfall stations
- Conventional model with 6 rainfall stations
- Optimised model

Performance assessment

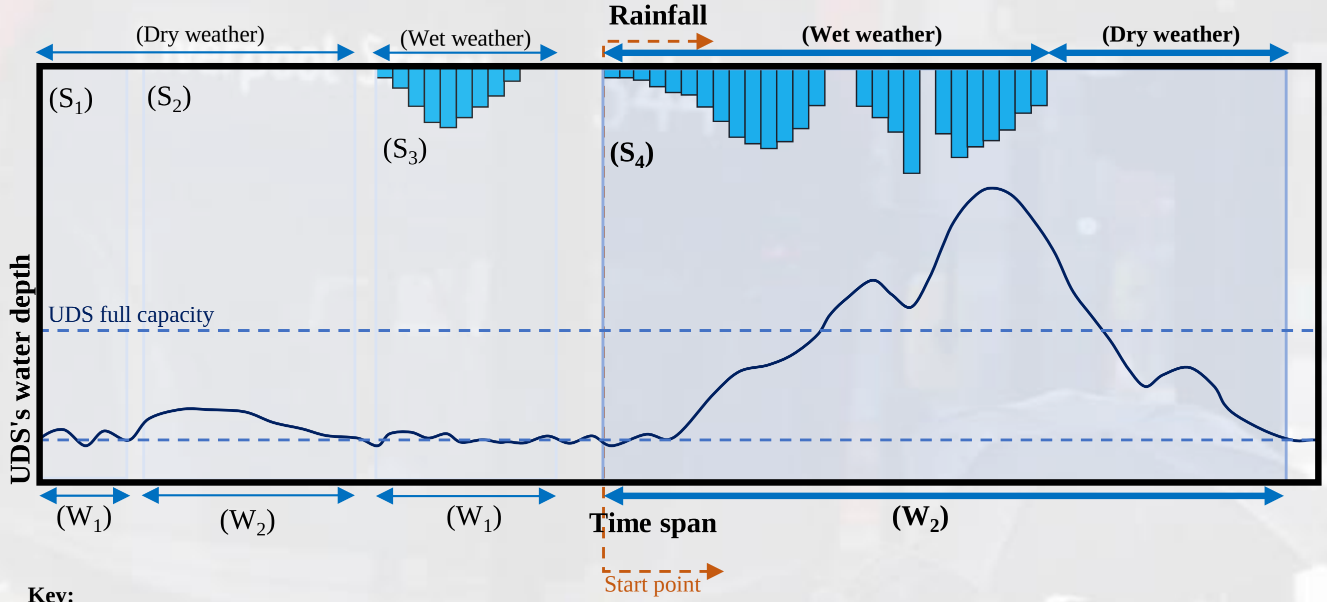
- Event identification
- RMSE
- NNSE
- Accuracy of overflow detection



All the models were developed on a laptop with Intel i7-6700 HQ CPU @ 2.60GHz and 16 GB RAM Memory

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

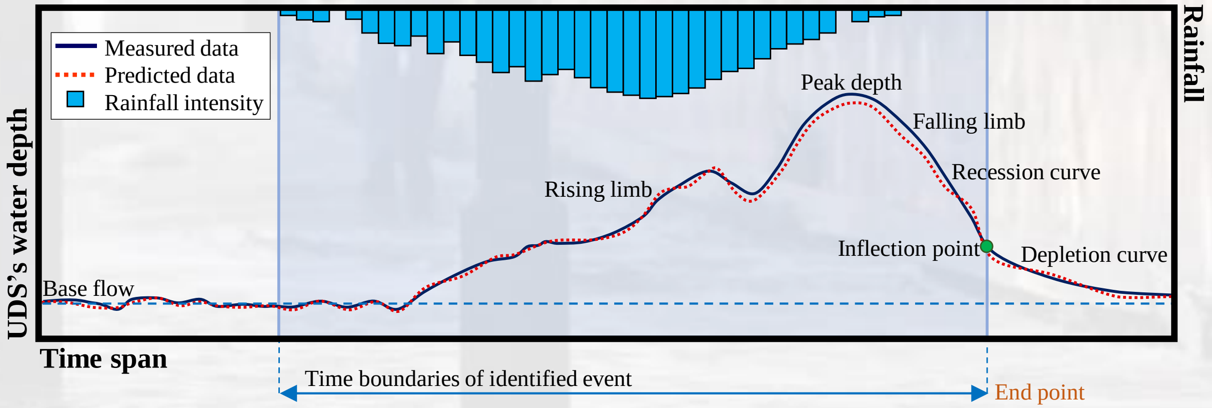
❖ *Typical flood event with temporal boundaries*



Key:

State	Captured data: Rainfall intensity	Captured data: Water depth
(S1): Dry weather, non-flood event	(R_1): -	(W_1): -
(S2): Sudden rising flow, non-flood event	(R_1): -	(W_2): +
(S3): Ineffective precipitation, non-flood event	(R_2): +	(W_1): -
(S4): Flood event	(R_2): +	(W_2): +

-: No rainfall, no change for water depth
+: Rainfall, net change (increase or decrease) for water depth

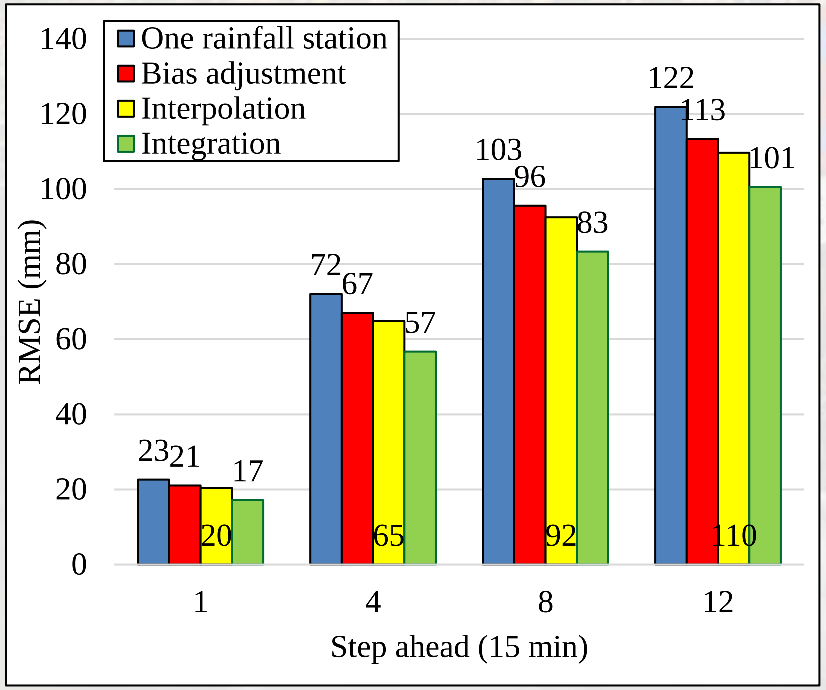


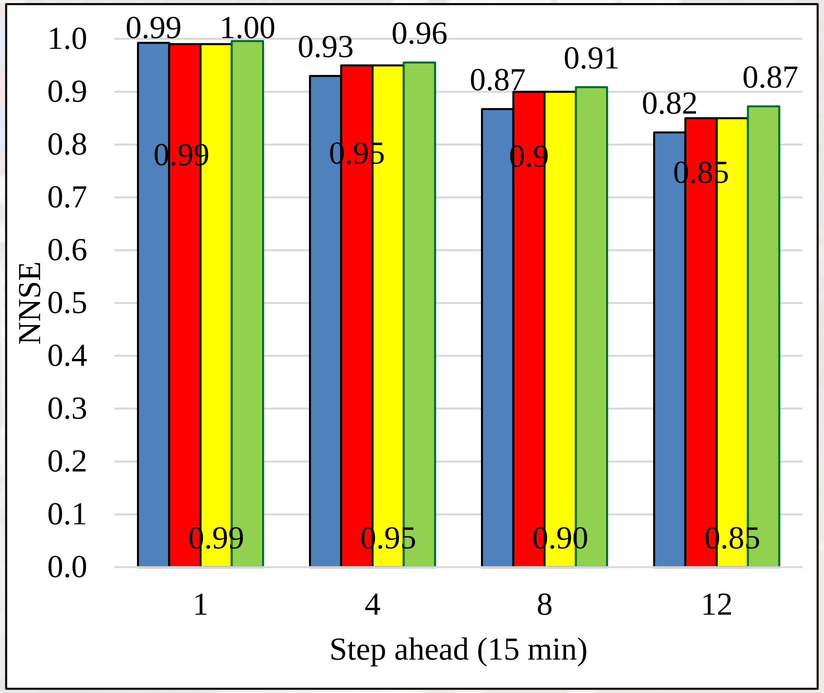
"A New Event-based Framework for Artificial Intelligence-based Modelling of Real-Time Flood Forecasting", **Piadeh F.**, Behzadian K. Alani A.M., Chen A.S., Campos L.C., *Water research,* Jun. 2022 [Under review]

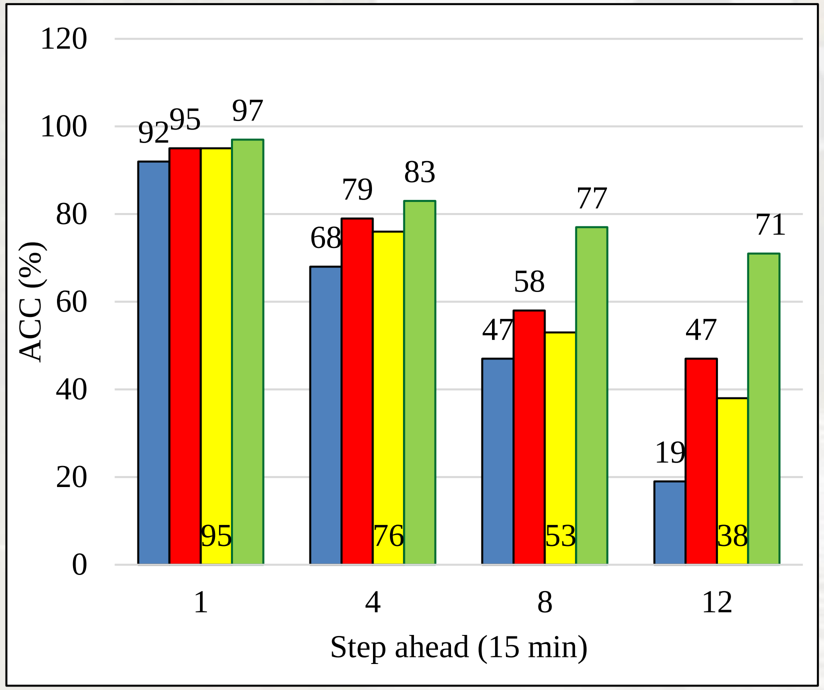
Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

Results and discussion

❖ *Pre-phase analysis*

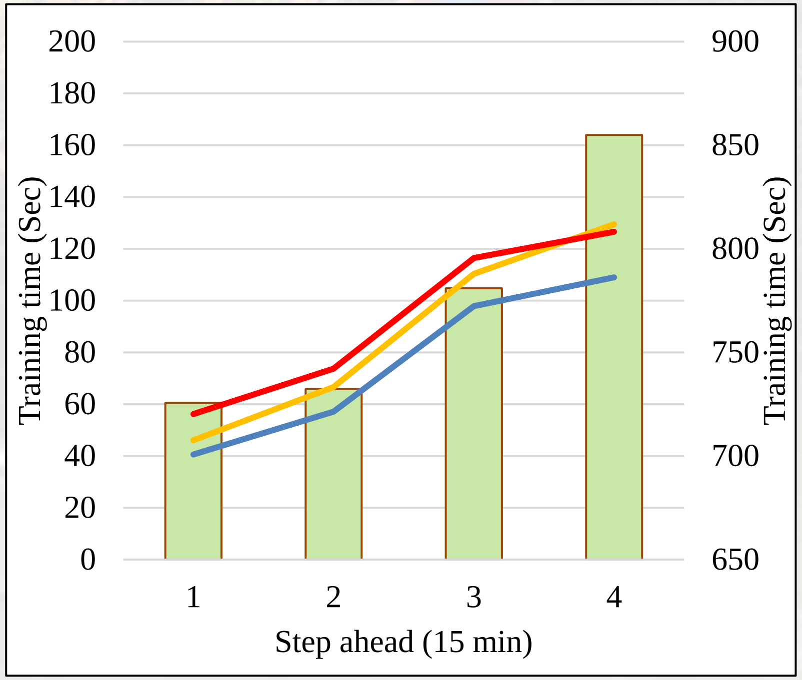


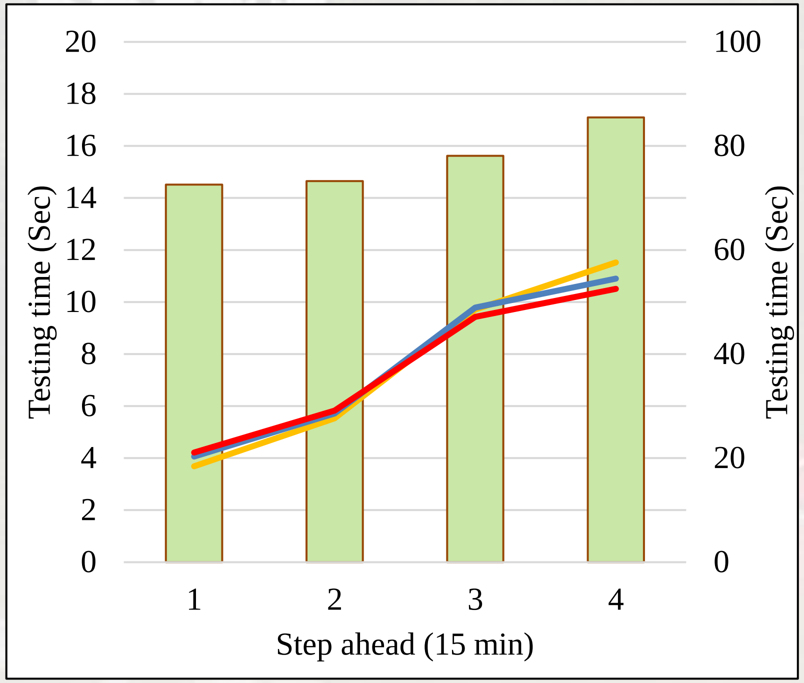




Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

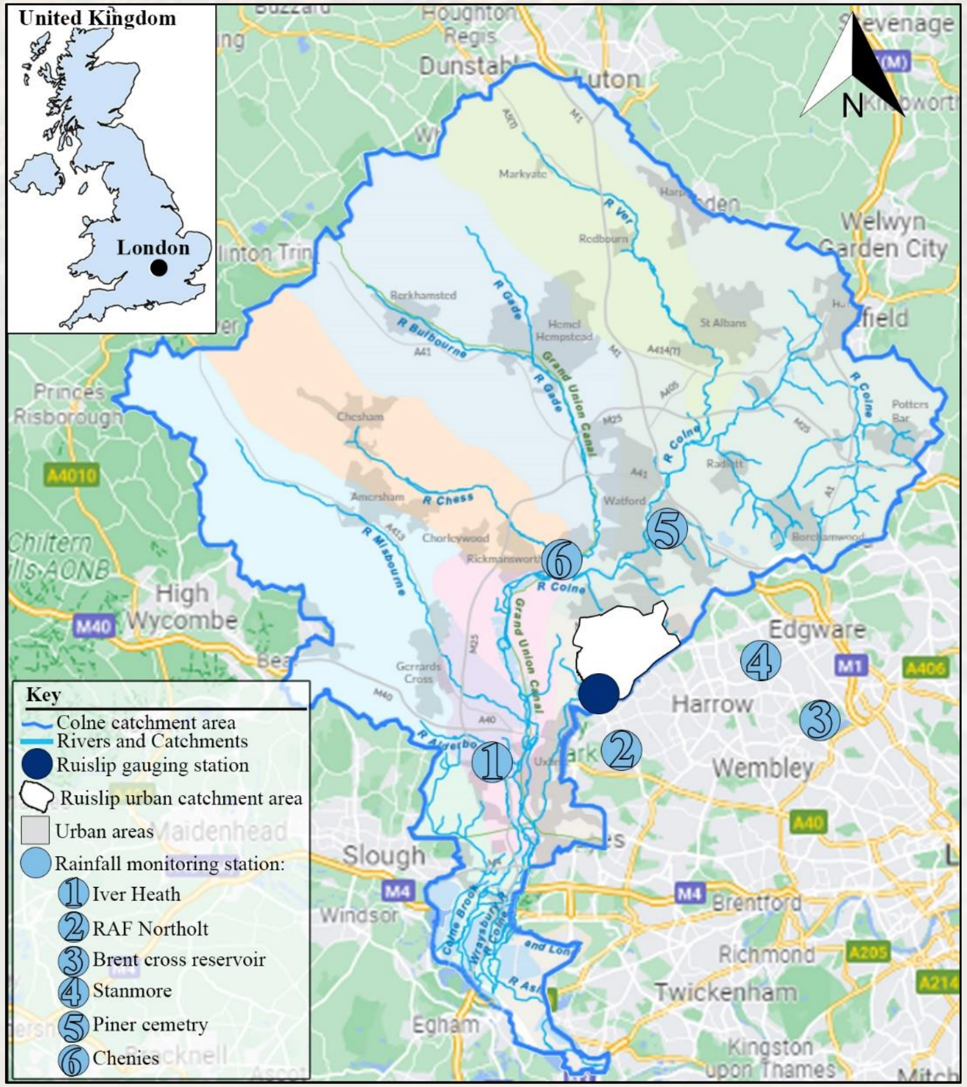
❖ Pre-phase analysis

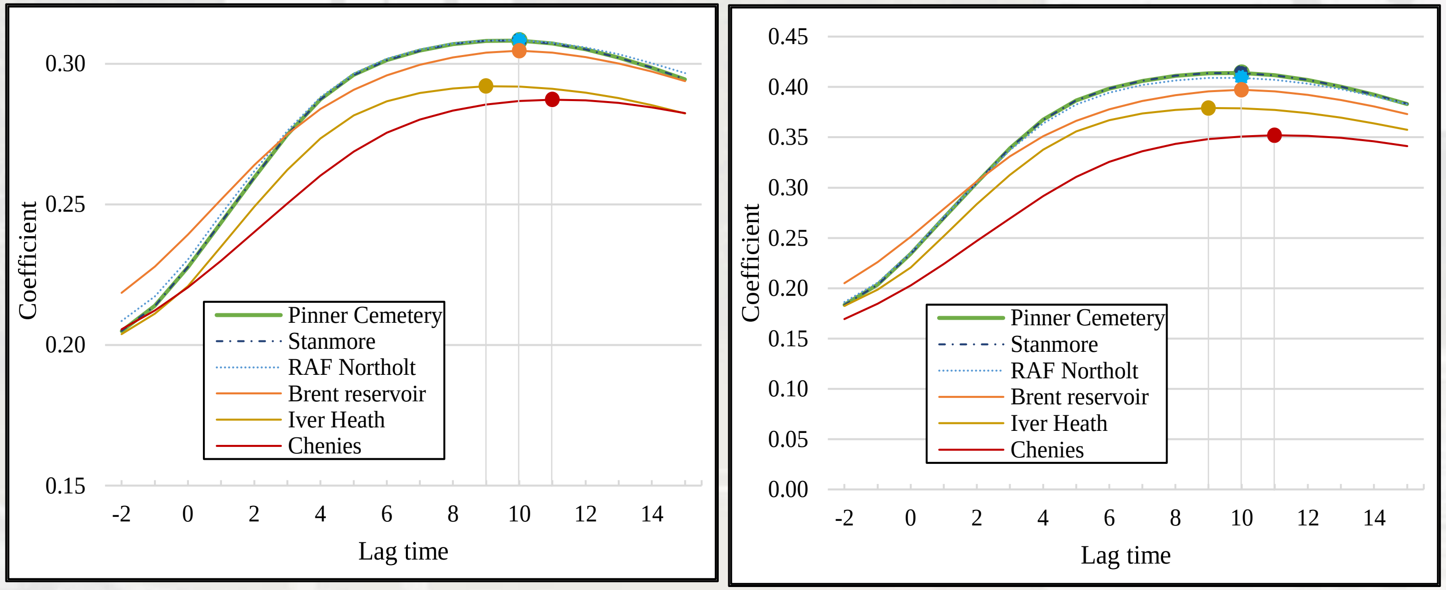




Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

❖ *Integration method analysis*

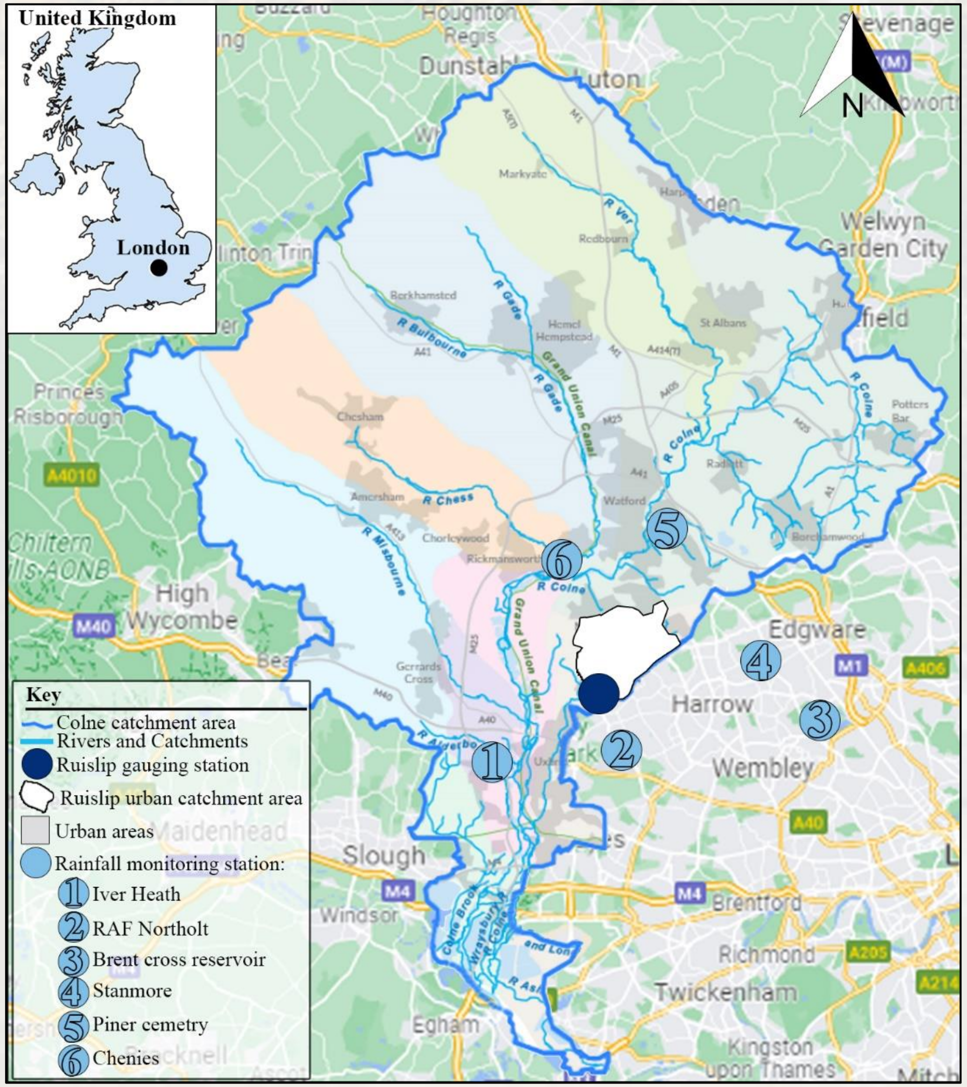


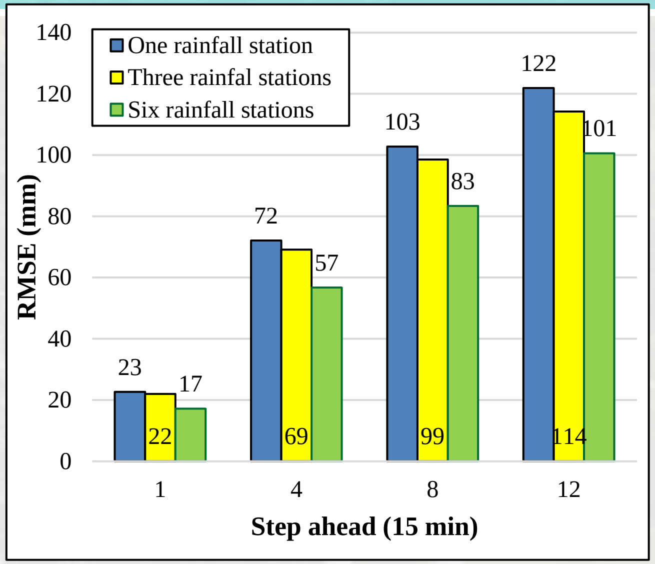


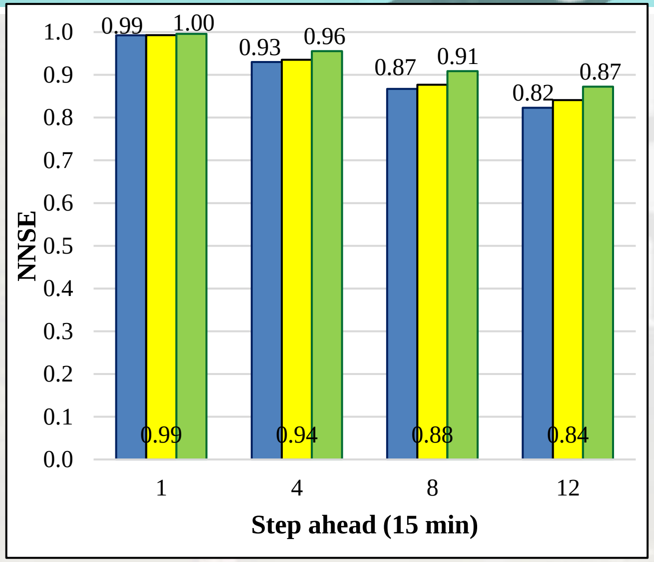
Data relationship between different selected rainfall stations and water level
(Left): Cross-correlation
(Right): Cross-covariance

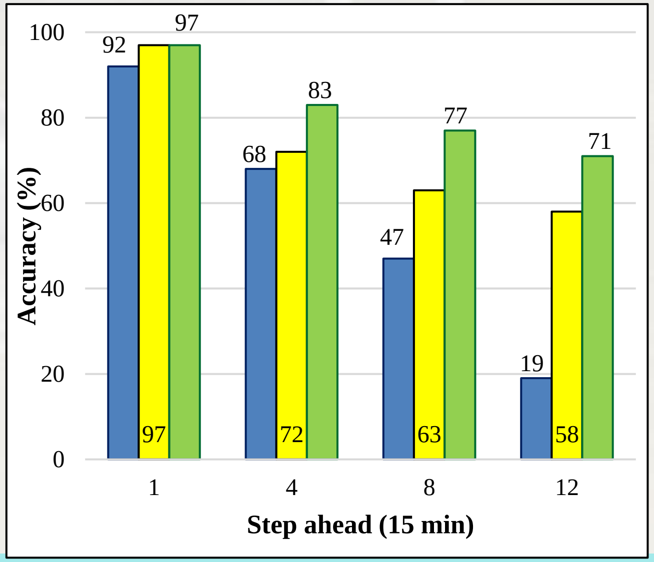
Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

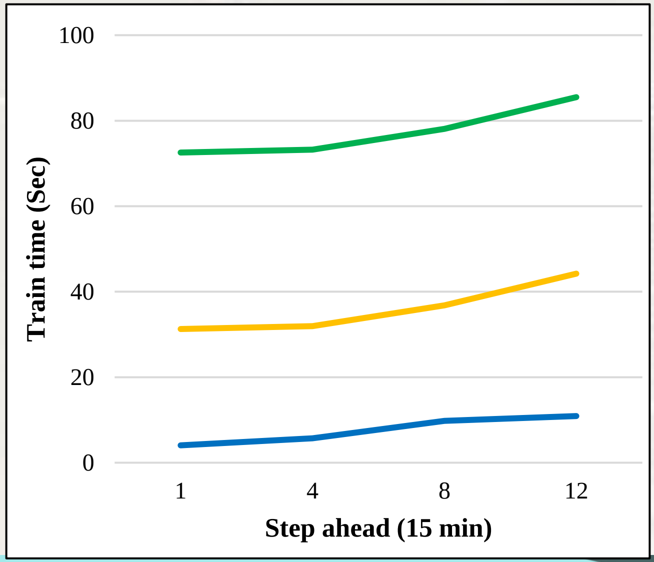
❖ *Integration method analysis*











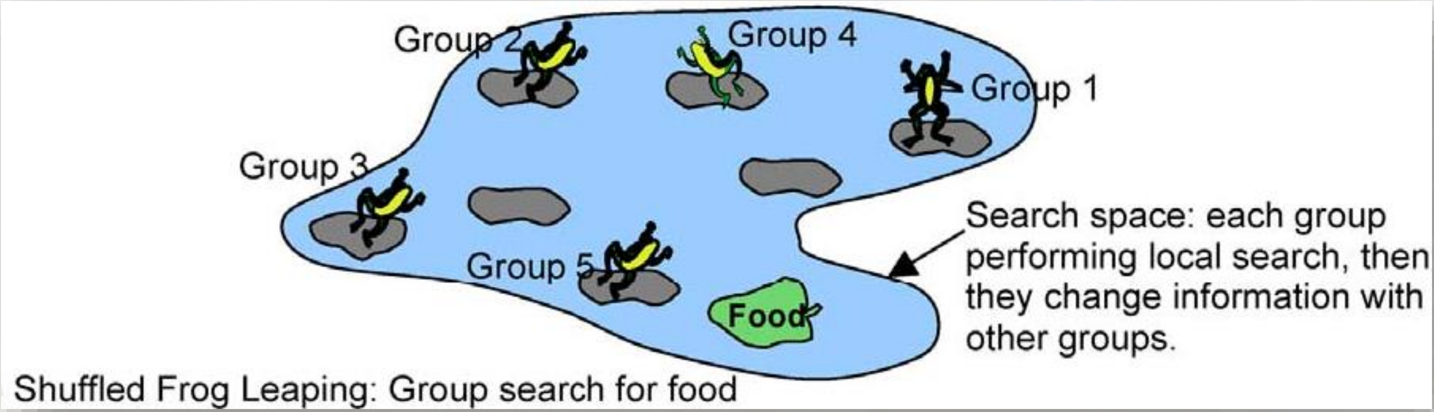
Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

Selected optimisation algorithm: Shuffled frog-leaping algorithm (SFLA)

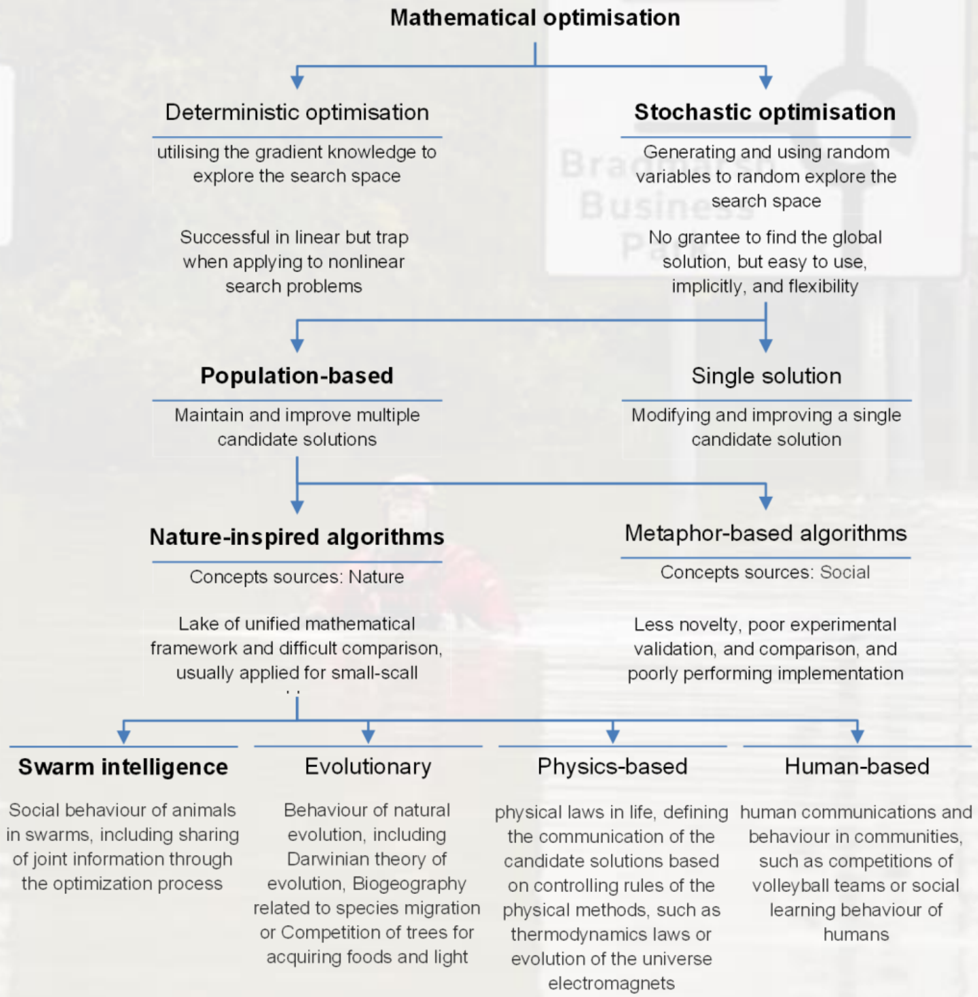
Decision variables: Range of time-series input data

Each trial: 4 shuffle sample for exploration step
4 shuffle sample for exploitation step

Objectives: RMSE, NSE, and ACC enhancement
Stopping criteria: Improvement less than 0.01%



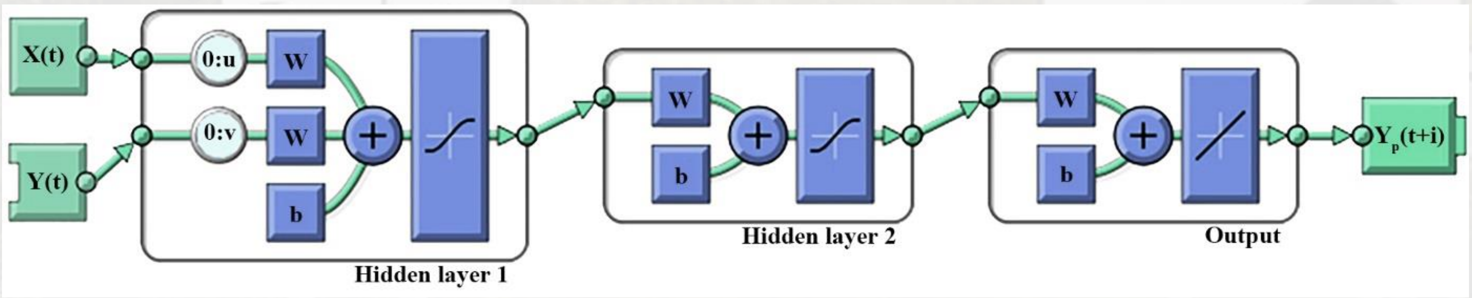
Shuffled Frog Leaping: Group search for food

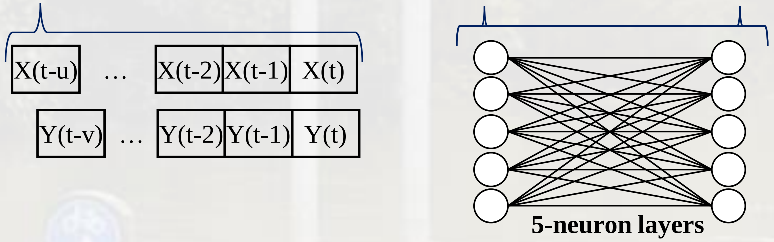


Bui, Q., Nguyen, Q., Nguyen, X., Pham, V., Nguyen, H., Pham, V. (2020). Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. *Journal of Hydrology*, 581: 124379.

❖ *Optimisation analysis*

Trial	Lag times of dataset						
	R1	R2	R3	R4	R5	R6	Water level
0 (Conventional approach)	9	10	10	10	10	11	1
1	7	7	3	4	8	8	6
2	4	5	9	10	2	11	3
3	4	4	4	9	5	7	2
4	3	3	9	8	5	4	8
5	7	9	6	7	2	5	9
6	2	3	5	3	5	7	2
7	5	5	2	6	8	7	5
8	4	8	8	5	7	7	2
9	3	5	9	6	7	7	5
10	4	5	9	6	7	7	5





Key
X: Observant rainfall intensity
Y: Observant water depth
Y_p: Predicted water depth
u: Delay factor (range of pervious X data)
v: Feedback factor (range of pervious Y)
w: Weight value
b: Bias value
i: Time-step ahead

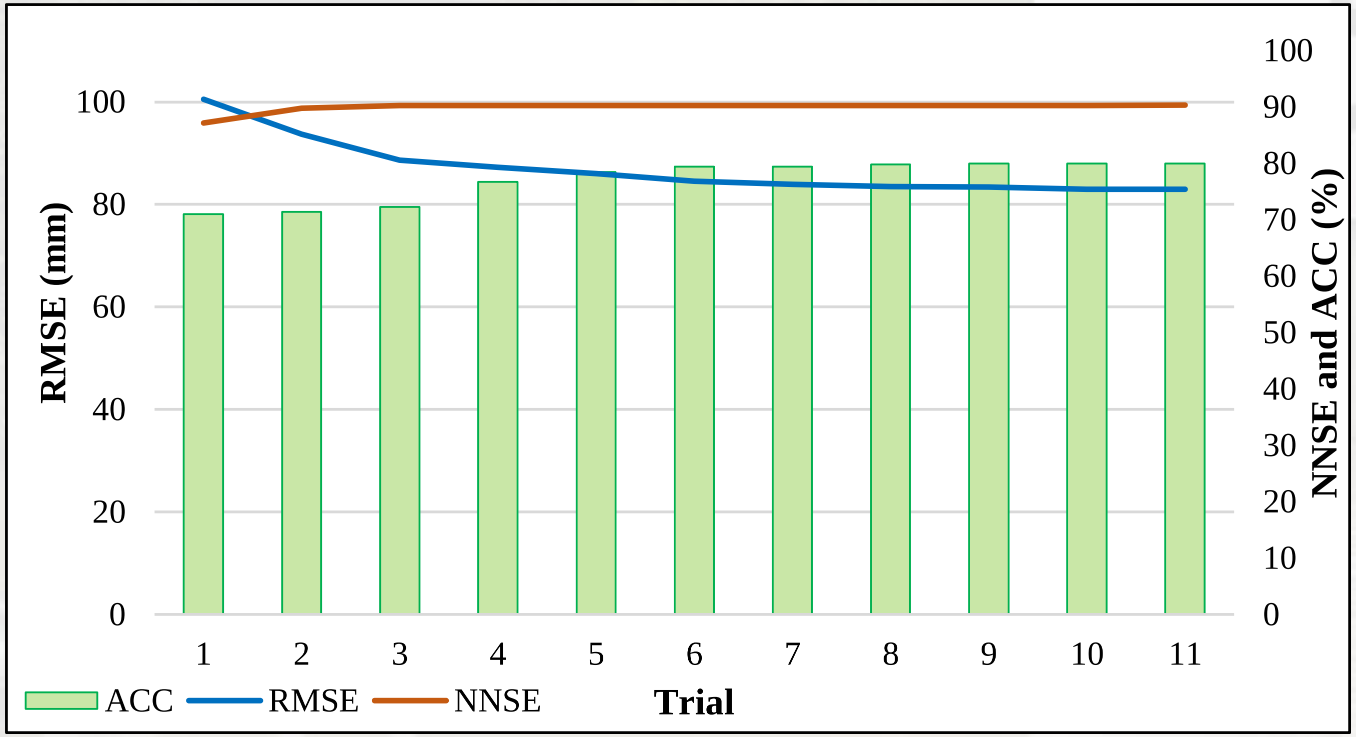
$10 \times 10 \times 10 \times 10 \times 10 \times 10 \times 11 \times 10 \times 15$ (min) = 314 Years !!!!!!

$10 \times 8 \times 15$ (min) = 20 Hrs

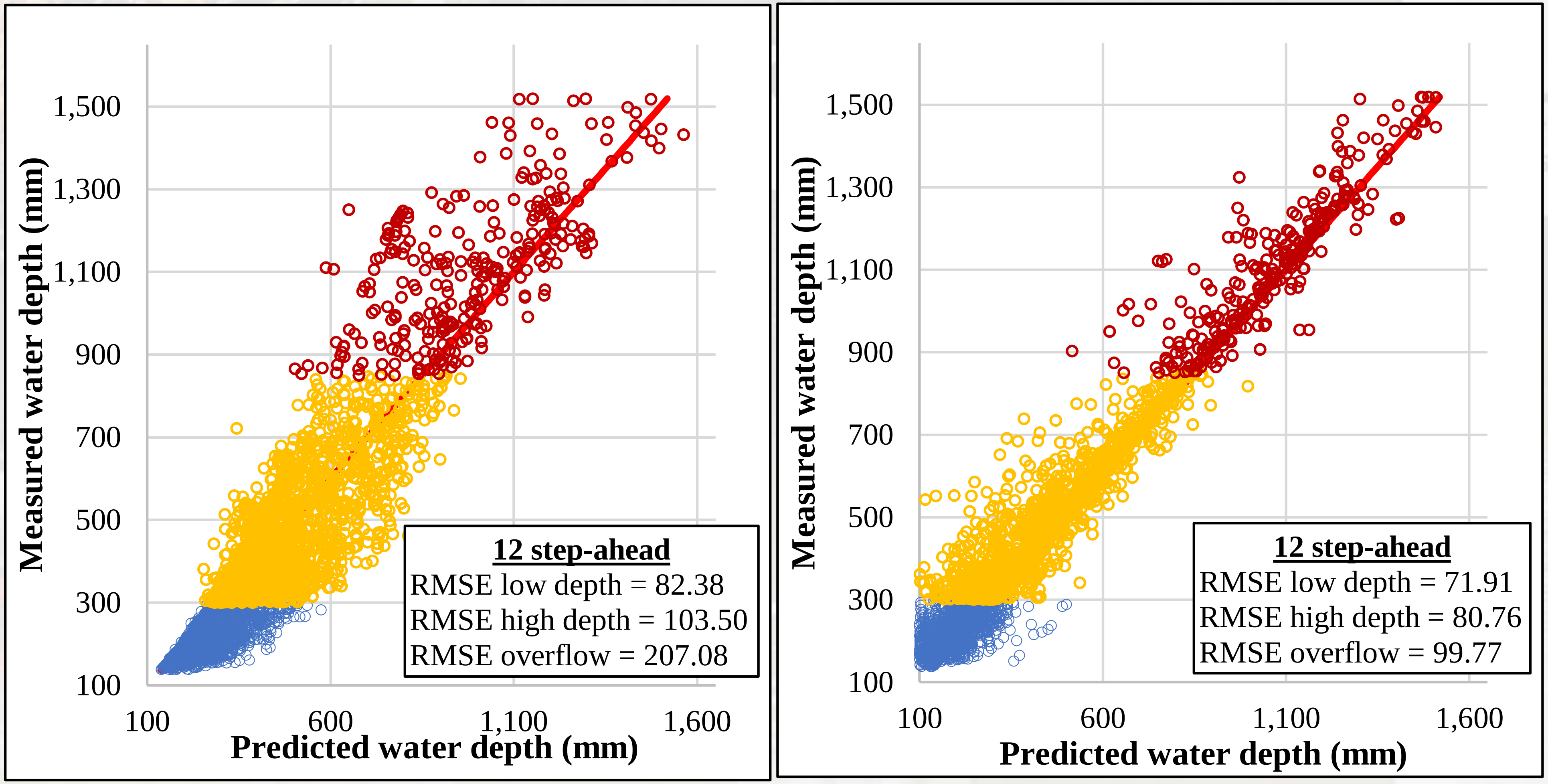
Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

❖ *Optimisation analysis*

Trial	Lag times of dataset						
	R1	R2	R3	R4	R5	R6	Water level
0 (Conventional approach)	9	10	10	10	10	11	1
1	7	7	3	4	8	8	6
2	4	5	9	10	2	11	3
3	4	4	4	9	5	7	2
4	3	3	9	8	5	4	8
5	7	9	6	7	2	5	9
6	2	3	5	3	5	7	2
7	5	5	2	6	8	7	5
8	4	8	8	5	7	7	2
9	3	5	9	6	7	7	5
10	4	5	9	6	7	7	5



Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk



Model performance of (Left) Best conventional model, (Right) Optimised model

Linkedin.com/in/farzad-piadeh-b9313450
Email: 21452390@student.uwl.ac.uk

Conclusion

01 Merging method

Integration is more effective but time consuming

02 Integrated method

Regardless of rainfall correlation, increasing input data causes outperformance

03 Input selection

Optimisation result is not compatible with conventional approach


Thank You For Your Attention