Multi-objective evolutionary feature selection with deep learning applied to air quality spatio-temporal forecasting

Artificial Intelligence in Research and Applications Seminar (AIRA)

Raquel Espinosa Fernández

Universidad de Murcia

October 19, 2023

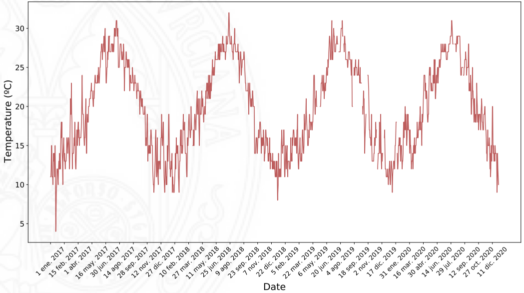
Overview

Overview

Motivation

Time series data have been used in a variety of:

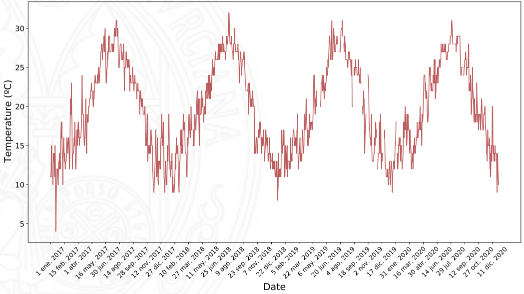
- Domains
  - ➢ Financial markets.
  - ➢ Internet of things.
  - ➢ Air quality.
- Applications
  - ➢ Pattern recognition.
  - ➢ Forecasting.
  - ➢ Signal processing.



Motivation

Time series data have been used in a variety of:

- Domains
  - ➢ Financial markets.
  - ➢ Internet of things.
  - ➢ Air quality.
- Applications
  - ➢ Pattern recognition.
  - ➢ Forecasting.
  - ➢ Signal processing.

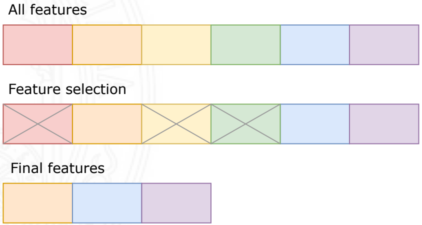


Increasing amount of data $\rightarrow$ Curse of dimensionality

Feature selection

Feature selection (FS) is a process in which relevant features are selected from a data set. Three main methods:

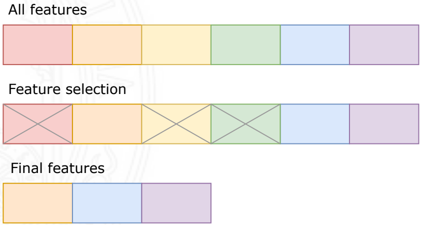
- **Filter**. Separates attribute selection from the learning algorithm.
- **Wrapper**. Uses predictive accuracy of a learning algorithm to select the attributes.
- **Embedded**. Feature selection integrated into the learning algorithm.



Feature selection

Feature selection (FS) is a process in which relevant features are selected from a data set. Three main methods:

- **Filter**. Separates attribute selection from the learning algorithm.
- **Wrapper**. Uses predictive accuracy of a learning algorithm to select the attributes.
- **Embedded**. Feature selection integrated into the learning algorithm.



Wrappers find more precise combinations of attributes → Computationally expensive

Area of application

The main area of application of the proposed techniques is **air quality prediction**.

- The emission of certain gases, such as CO_2, NO_2, NO_x, or PM, is deteriorating air quality.
- According to the WHO, in 2022, 99% of the population has been exposed to areas where air quality limits are exceeded.
- Prolonged exposure to noxious gases can cause various diseases and even premature death.
- The environment and ecosystems are negatively affected by the deterioration of air quality.

Source: *WHO Website*

Overview

General objective

Develop efficient and effective feature selection techniques for deep learning through multi-objective evolutionary algorithms and application of the created methods for time series forecasting in different areas of interest.

Specific objectives

- **SO1: Develop a comprehensive methodology and implement a multi-criteria decision-making process for the comparison and evaluation of predictive models for time series forecasting.**

Specific objectives

- **SO1:** Develop a comprehensive methodology and implement a multi-criteria decision-making process for the comparison and evaluation of predictive models for time series forecasting.
- **SO2: Study, design and develop a multi-objective evolutionary approach based on spatio-temporal characteristics within the Autonomous Region of Murcia.**

Specific objectives

- **SO1:** Develop a comprehensive methodology and implement a multi-criteria decision-making process for the comparison and evaluation of predictive models for time series forecasting.
- **SO2:** Study, design and develop a multi-objective evolutionary approach based on spatio-temporal characteristics within the Autonomous Region of Murcia.
- **SO3: Define multi-objective optimization problems for FS, with objectives of different nature.**

Specific objectives

- **SO1:** Develop a comprehensive methodology and implement a multi-criteria decision-making process for the comparison and evaluation of predictive models for time series forecasting.
- **SO2:** Study, design and develop a multi-objective evolutionary approach based on spatio-temporal characteristics within the Autonomous Region of Murcia.
- **SO3:** Define multi-objective optimization problems for FS, with objectives of different nature.
- **SO4: Solve the proposed optimization problems by identifying the best MOEAs and developing surrogate-assisted approaches to reduce the computational cost of the algorithms.**

Specific objectives

- **SO1:** Develop a comprehensive methodology and implement a multi-criteria decision-making process for the comparison and evaluation of predictive models for time series forecasting.
- **SO2:** Study, design and develop a multi-objective evolutionary approach based on spatio-temporal characteristics within the Autonomous Region of Murcia.
- **SO3:** Define multi-objective optimization problems for FS, with objectives of different nature.
- **SO4:** Solve the proposed optimization problems by identifying the best MOEAs and developing surrogate-assisted approaches to reduce the computational cost of the algorithms.
- **SO5: Identify metrics to quantify the variability between surrogate-assisted approaches and facilitate the establishment of qualitative analysis.**

Specific objectives

- **SO1:** Develop a comprehensive methodology and implement a multi-criteria decision-making process for the comparison and evaluation of predictive models for time series forecasting.
- **SO2:** Study, design and develop a multi-objective evolutionary approach based on spatio-temporal characteristics within the Autonomous Region of Murcia.
- **SO3:** Define multi-objective optimization problems for FS, with objectives of different nature.
- **SO4:** Solve the proposed optimization problems by identifying the best MOEAs and developing surrogate-assisted approaches to reduce the computational cost of the algorithms.
- **SO5:** Identify metrics to quantify the variability between surrogate-assisted approaches and facilitate the establishment of qualitative analysis.
- **SO6: Evaluate, validate and compare the developed FS methods with time series data for air quality forecasting in the context of the Autonomous Region of Murcia, as well as in other geographic locations and in other time series forecasting problems.**

Specific objectives

- **SO1:** Develop a comprehensive methodology and implement a multi-criteria decision-making process for the comparison and evaluation of predictive models for time series forecasting.
- **SO2:** Study, design and develop a multi-objective evolutionary approach based on spatio-temporal characteristics within the Autonomous Region of Murcia.
- **SO3:** Define multi-objective optimization problems for FS, with objectives of different nature.
- **SO4:** Solve the proposed optimization problems by identifying the best MOEAs and developing surrogate-assisted approaches to reduce the computational cost of the algorithms.
- **SO5:** Identify metrics to quantify the variability between surrogate-assisted approaches and facilitate the establishment of qualitative analysis.
- **SO6:** Evaluate, validate and compare the developed FS methods with time series data for air quality forecasting in the context of the Autonomous Region of Murcia, as well as in other geographic locations and in other time series forecasting problems.

Overview

Multi-objective evolutionary algorithms

Multi-objective evolutionary algorithms (MOEAs) are multi-objective global search and optimization techniques. Key characteristics:

- Conflicting objective functions.
- Non-dominated solutions.
- Pareto front.
- Solve complex problems.

Wrapper methods for FS can be defined as a **multi-objective optimization problem**.

Multi-objective evolutionary algorithms

Multi-objective evolutionary algorithms (MOEAs) are multi-objective global search and optimization techniques. Key characteristics:

- Conflicting objective functions.
- Non-dominated solutions.
- Pareto front.
- Solve complex problems.

Wrapper methods for FS can be defined as a **multi-objective optimization problem**.

Problem

High computational cost to converge in a set of diverse non-dominated solutions

Surrogate models

Solution

Use surrogate models to approximate the objective function of a MOEA

Surrogate models

Solution

Use surrogate models to approximate the objective function of a MOEA

Surrogate models simulate the behaviour of a model and try to approximate its results, which makes possible a reduction in computational costs. Key characteristics:

- Evaluate candidate solutions instead of using the real objective function.
- Reduce computational cost of the optimization.

Overview

Datasets

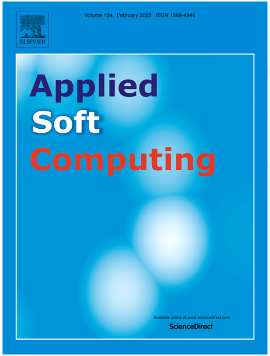
Type	Location	Task	Duration	Frequency	Instances	Attributes
Air quality	Wrocław (Poland)	Regression	2015 - 2017	Hourly	26304	9
Air quality	La Aljorra (Spain)	Regression/Classification	2017 - 2020	Daily	1461	17
Domotic house	Valencia (Spain)	Regression/Classification	March - May 2012	15 minutes	4137	24
Smart building	Vienna (Austria)	Classification/Clustering	2013 - 2022	Hourly	32323407	14

Table 1: Summary of datasets.

La Aljorra monitoring station.
Source: *Google Images*

Smart buildings from Vienna.
Source: *Siemens AG*

A time series forecasting based multi-criteria methodology for air quality prediction



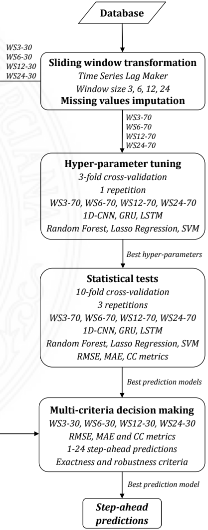
Title	A time series forecasting based multi-criteria methodology for air quality prediction
Authors	Raquel Espinosa, José Palma, Fernando Jiménez, Joanna Kamińska, Guido Sciavicco, Estrella Lucena-Sánchez
Journal	Applied Soft Computing
Impact factor (2021)	8.263
JCR Rank (2021)	Computer science, interdisciplinary applications: 11/112 (D1) Computer science, artificial intelligence: 23/145 (Q1)
Cited by	41
Publisher	Elsevier
Date	7 September 2021
ISSN	1568-4946
State	Published
Contribution	Term, conceptualization, methodology, software, validation, investigation, resources, data curation, writing – original draft, writing – review and editing, visualization, project administration

Methodology for the identification of DL architectures and comparison of predictive models

Dataset: air quality from Wrocław.

Phases of the proposed methodology:

1. Sliding window transformation.
   - ➢ Window sizes: 3, 6, 12, 24.
2. Hyper-parameter tuning.
   - ➢ Machine learning: RF, Lasso, SVM.
   - ➢ Deep learning: 1D-CNN, GRU, LSTM.
3. Statistical tests.
4. Multi-criteria decision making.
5. Step ahead predictions.



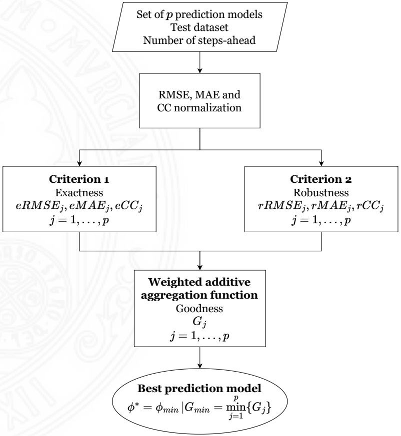
Multi-criteria decision making

To measure the **goodness** of the prediction models. Two criteria:

- Exactness.
- Robustness.

For validation:

- Error metrics: RMSE, MAE and CC.
- 6 best prediction models.
- 24-steps ahead.



Results

Model	Wins	Losses	Wins−Losses
RF-WS24-70	21	0	21
GRU-WS24-70	20	0	20
LSTM-WS24-70	**19**	**0**	**19**
Lasso-WS24-70	20	1	19
LSTM-WS3-70	12	3	9
GRU-WS6-70	9	4	5
RF-WS3-70	8	4	4
RF-WS12-70	7	4	3
1DCNN-WS3-70	7	4	3
LSTM-WS6-70	7	4	3
RF-WS6-70	7	5	2
GRU-WS12-70	6	4	2
GRU-WS3-70	6	5	1
1DCNN-WS6-70	4	4	0
LSTM-WS12-70	4	5	-1
1DCNN-WS24-70	4	6	-2
1DCNN-WS12-70	4	7	-3
Lasso-WS12-70	6	11	-5
Lasso-WS6-70	5	14	-9
Lasso-WS3-70	4	15	-11
SVMRadial-WS3-70	3	20	-17
SVMRadial-WS6-70	2	21	-19
SVMRadial-WS12-70	1	22	-21
SVMRadial-WS24-70	0	23	-23

(a) MAE

Model	Wins	Losses	Wins−Losses
Lasso-WS24-70	20	0	20
RF-WS24-70	20	0	20
GRU-WS24-70	20	0	20
LSTM-WS24-70	**14**	**0**	**14**
LSTM-WS3-70	9	3	6
RF-WS3-70	7	3	4
RF-WS6-70	7	3	4
GRU-WS6-70	7	3	4
RF-WS12-70	7	4	3
1DCNN-WS3-70	6	3	3
LSTM-WS6-70	6	4	2
1DCNN-WS6-70	5	3	2
GRU-WS3-70	5	4	1
1DCNN-WS24-70	5	4	1
1DCNN-WS12-70	4	4	0
GRU-WS12-70	4	5	-1
LSTM-WS12-70	4	6	-2
Lasso-WS12-70	6	8	-2
Lasso-WS6-70	5	12	-7
Lasso-WS3-70	4	16	-12
SVMRadial-WS3-70	3	20	-17
SVMRadial-WS6-70	2	21	-19
SVMRadial-WS12-70	1	22	-21
SVMRadial-WS24-70	0	23	-23

(b) RMSE

Model	Wins	Losses	Wins−Losses
RF-WS24-70	10	0	10
LSTM-WS3-70	9	0	9
Lasso-WS24-70	8	0	8
RF-WS3-70	7	0	7
RF-WS6-70	7	0	7
RF-WS12-70	7	0	7
GRU-WS24-70	7	0	7
GRU-WS6-70	7	0	7
1DCNN-WS3-70	6	0	6
LSTM-WS6-70	6	0	6
GRU-WS3-70	5	0	5
LSTM-WS24-70	**4**	**0**	**4**
1DCNN-WS12-70	4	0	4
1DCNN-WS6-70	4	0	4
GRU-WS12-70	4	2	2
1DCNN-WS24-70	4	3	1
LSTM-WS12-70	4	3	1
Lasso-WS12-70	6	6	0
Lasso-WS6-70	5	11	-6
Lasso-WS3-70	4	13	-9
SVMRadial-WS3-70	3	20	-17
SVMRadial-WS6-70	2	21	-19
SVMRadial-WS12-70	1	22	-21
SVMRadial-WS24-70	0	23	-23

(c) CC

Table 2: Ranking of the NO_2 models.

Results

Model	Goodness
LSTM-WS24-70	1.277197
Lasso-WS24-70	1.368710
RF-WS24-70	1.554364
LSTM-WS3-70	1.729293
GRU-WS6-70	1.804245
GRU-WS24-70	1.845464
RF-WS3-70	1.961961
RF-WS12-70	2.037450
RF-WS6-70	2.183755

Table 3: Set of the competing models in the multi-criteria decision-making process and goodness obtained for NO_2 prediction.

Results

Model	Goodness
LSTM-WS24-70	**1.277197**
Lasso-WS24-70	1.368710
RF-WS24-70	1.554364
LSTM-WS3-70	1.729293
GRU-WS6-70	1.804245
GRU-WS24-70	1.845464
RF-WS3-70	1.961961
RF-WS12-70	2.037450
RF-WS6-70	2.183755

Table 3: Set of the competing models in the multi-criteria decision-making process and goodness obtained for NO_2 prediction.

Results

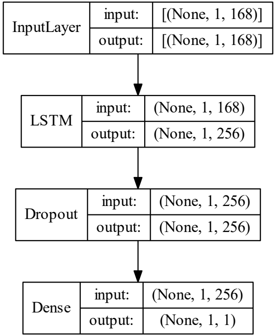


Figure 1: Architecture of the LSTM-WS24 deep learning model for NO_2 prediction.

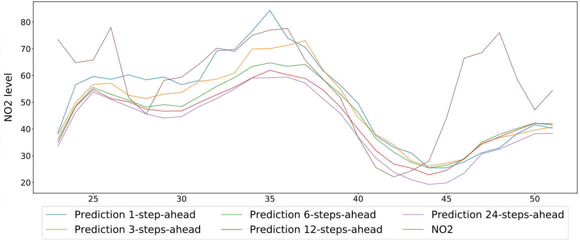


Figure 2: NO_2 level predicted time series in test with LSTM-WS24.

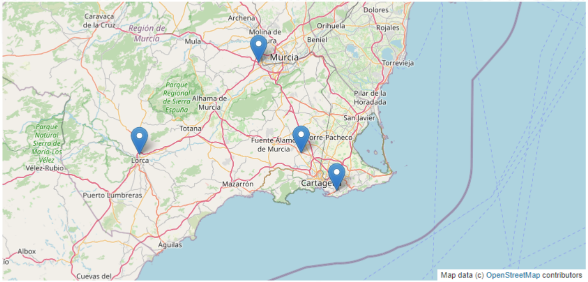
Multi-objective evolutionary spatio-temporal forecasting of air pollution

Title	Multi-objective evolutionary spatio-temporal forecasting of air pollution
Authors	Raquel Espinosa, Fernando Jiménez, José Palma
Journal	Future Generation Computer Systems
Impact factor (2021)[1]	7.307
JCR Rank (2021)[1]	Computer science, theory & methods: 10/110 (D1)
Cited by	9
Publisher	Elsevier
Date	31 May 2022
ISSN	0167-739X
State	Published
Contribution	Conception and design of the study, acquisition of data, analysis and interpretation of data, point by point revision of reviewer's comments

[1] At the time of publication of this thesis, the impact factor and JCR rank data for the year 2022 were not available.

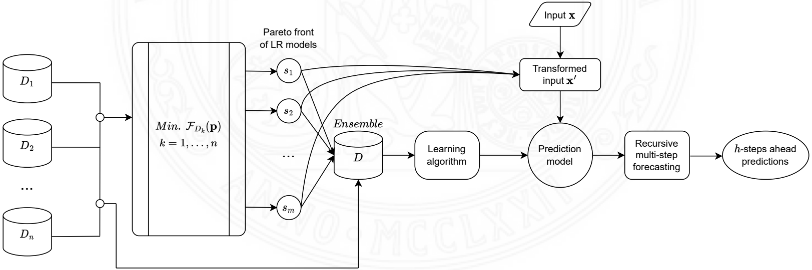
Multi-objective optimization based spatio-temporal approach

- 3 objectives: RMSE of a LR predictive model of the monitoring stations.
- MOEAs: NSGA-II, MOEA/D, SPEA2.
- Ensemble learning approach based on **stacking**.
- Learning algorithms: RF, LR, SVM, QRNN, MLP, kNN, ZeroR.
- Comparison: interpolation based on an IDW.



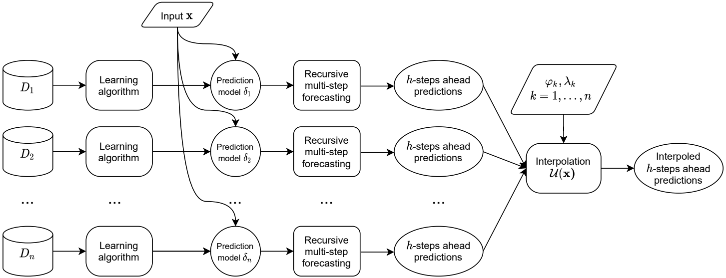
Multi-objective optimization based spatio-temporal approach

- 3 objectives: RMSE of a LR predictive model of the monitoring stations.
- MOEAs: NSGA-II, MOEA/D, SPEA2.
- Ensemble learning approach based on **stacking**.
- Learning algorithms: RF, LR, SVM, QRNN, MLP, kNN, ZeroR.
- Comparison: interpolation based on an IDW.



Multi-objective optimization based spatio-temporal approach

- 3 objectives: RMSE of a LR predictive model of the monitoring stations.
- MOEAs: NSGA-II, MOEA/D, SPEA2.
- Ensemble learning approach based on **stacking**.
- Learning algorithms: RF, LR, SVM, QRNN, MLP, kNN, ZeroR.
- Comparison: **interpolation based on an IDW**.



Results

Algorithm	Best	Worse	Average	SD
NSGA-II	0.35393	0.22136	0.30934	0.03169
MOEA/D	0.33819	0.17268	0.26805	0.04673
SPEA2	0.25006	0.14090	0.19320	0.02874

Table 4: Summary of the results of the MOEAs, 10,000 evaluations, 30 runs.

MOEA	Win	Loss	Win – Loss
NSGA-II	2	0	2
MOEA/D	1	1	0
SPEA2	0	2	-2

Table 5: Ranking of MOEAs sorted from best to worse.

Results

Algorithm	Best	Worse	Average	SD
NSGA-II	**0.35393**	**0.22136**	**0.30934**	**0.03169**
MOEA/D	0.33819	0.17268	0.26805	0.04673
SPEA2	0.25006	0.14090	0.19320	0.02874

Table 4: Summary of the results of the MOEAs, 10,000 evaluations, 30 runs.

MOEA	Win	Loss	Win – Loss
NSGA-II	2	0	2
MOEA/D	1	1	0
SPEA2	0	2	-2

Table 5: Ranking of MOEAs sorted from best to worse.

Results

Algorithm	Best	Worse	Average	SD
NSGA-II	**0.35393**	**0.22136**	**0.30934**	**0.03169**
MOEA/D	0.33819	0.17268	0.26805	0.04673
SPEA2	0.25006	0.14090	0.19320	0.02874

Table 4: Summary of the results of the MOEAs, 10,000 evaluations, 30 runs.

MOEA	Win	Loss	Win – Loss
NSGA-II	**2**	**0**	**2**
MOEA/D	1	1	0
SPEA2	0	2	-2

Table 5: Ranking of MOEAs sorted from best to worse.

Results

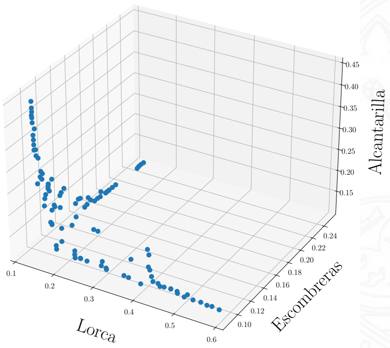


Figure 3: 3D pareto front with NSGA-II.

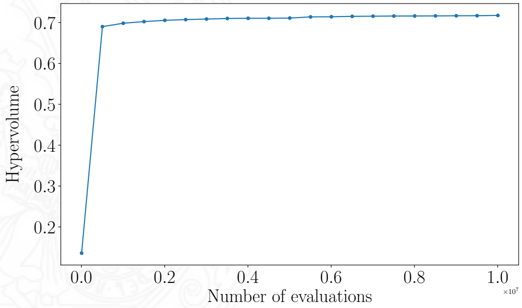


Figure 4: Hypervolume evolution with NSGA-II.

Results

Algorithm	Ensemble
LR	0.108788
RF	0.107863
SVM	0.108686
MLP	0.109847
kNN	0.142134
QRNN	0.107168
ZeroR	0.153978

Table 6: RMSE evaluation, 10-fold cross-validation, 10 repetitions of ensemble learning algorithms

Algorithm	Win	Loss	Win-Loss
RF	6	0	6
LR	5	0	5
SVM	5	0	5
QRNN	4	0	4
MLP	4	4	0
kNN	1	10	-9
ZeroR	0	11	-11

Table 7: Win-loss statistical test of algorithms.

Results

Algorithm	Ensemble
LR	0.108788
RF	0.107863
SVM	0.108686
MLP	0.109847
kNN	0.142134
QRNN	**0.107168**
ZeroR	0.153978

Table 6: RMSE evaluation, 10-fold cross-validation, 10 repetitions of ensemble learning algorithms

Algorithm	Win	Loss	Win-Loss
RF	6	0	6
LR	5	0	5
SVM	5	0	5
QRNN	4	0	4
MLP	4	4	0
kNN	1	10	-9
ZeroR	0	11	-11

Table 7: Win-loss statistical test of algorithms.

Results

Algorithm	Ensemble
LR	0.108788
RF	0.107863
SVM	0.108686
MLP	0.109847
kNN	0.142134
QRNN	**0.107168**
ZeroR	0.153978

Table 6: RMSE evaluation, 10-fold cross-validation, 10 repetitions of ensemble learning algorithms

Algorithm	Win	Loss	Win-Loss
RF	**6**	**0**	**6**
LR	**5**	**0**	**5**
SVM	**5**	**0**	**5**
QRNN	**4**	**0**	**4**
MLP	4	4	0
kNN	1	10	-9
ZeroR	0	11	-11

Table 7: Win-loss statistical test of algorithms.

Results

Models	RF	LR	SVM	QRNN
Training MOEA + Ensemble	0.1306	0.2181	0.2063	0.1732
Training Interpolation	0.0819	0.2182	0.1370	0.1518
La Aljorra Test MOEA + Ensemble	0.2694	0.2504	0.2825	0.2677
La Aljorra Test Interpolation	0.3397	0.3268	0.3296	0.3021

Table 8: Goodness of the prediction models with RF, LR, SVM and QRNN.

Method	Metric	1-step ahead	2-steps ahead	3-steps ahead	4-steps ahead	5-steps ahead	6-steps ahead	7-steps ahead
MOEA + Ensemble	RMSE	0.1118	0.1207	0.1229	0.1233	0.1207	0.1261	0.1408
	MAE	0.0890	0.0960	0.0980	0.0984	0.0961	0.1014	0.1154
	CC	0.5410	0.4863	0.4745	0.4724	0.4860	0.4581	0.3830
Interpolation	RMSE	0.1175	0.1207	0.1212	0.1296	0.1322	0.1346	0.1352
	MAE	0.0919	0.0936	0.0939	0.1014	0.1027	0.1035	0.1035
	CC	0.3199	0.2673	0.2393	0.2420	0.2574	0.2057	0.1874

Table 9: Results of the evaluation of models on La Aljorra test set with LR.

Results

Models	RF	LR	SVM	QRNN
Training MOEA + Ensemble	0.1306	0.2181	0.2063	0.1732
Training Interpolation	0.0819	0.2182	0.1370	0.1518
La Aljorra Test MOEA + Ensemble	0.2694	**0.2504**	0.2825	0.2677
La Aljorra Test Interpolation	0.3397	0.3268	0.3296	0.3021

Table 8: Goodness of the prediction models with RF, LR, SVM and QRNN.

Method	Metric	1-step ahead	2-steps ahead	3-steps ahead	4-steps ahead	5-steps ahead	6-steps ahead	7-steps ahead
MOEA + Ensemble	RMSE	0.1118	0.1207	0.1229	0.1233	0.1207	0.1261	0.1408
	MAE	0.0890	0.0960	0.0980	0.0984	0.0961	0.1014	0.1154
	CC	0.5410	0.4863	0.4745	0.4724	0.4860	0.4581	0.3830
Interpolation	RMSE	0.1175	0.1207	0.1212	0.1296	0.1322	0.1346	0.1352
	MAE	0.0919	0.0936	0.0939	0.1014	0.1027	0.1035	0.1035
	CC	0.3199	0.2673	0.2393	0.2420	0.2574	0.2057	0.1874

Table 9: Results of the evaluation of models on La Aljorra test set with LR.

Results

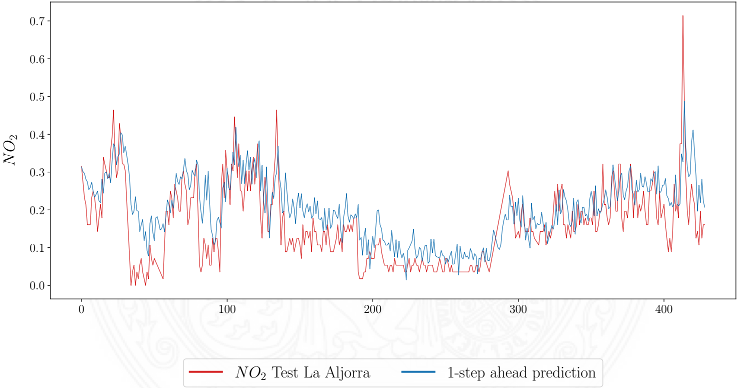


Figure 5: Original and predicted NO_2 time series for 1-step ahead built with the multi-objective optimization based spatio-temporal approach for LR.

Surrogate-assisted and filter-based multi-objective evolutionary FS for DL

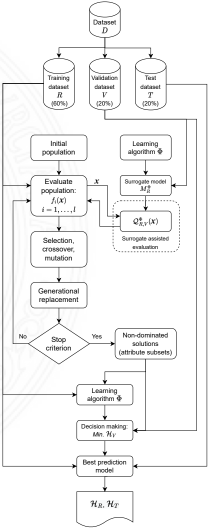
IEEE TRANSACTIONS ON
NEURAL NETWORKS AND LEARNING SYSTEMS
A PUBLICATION OF THE IEEE COMPUTATIONAL INTELLIGENCE SOCIETY
www.ieee-cis.org/pubs/tnnls

Title	Surrogate-assisted and filter-based multi-objective evolutionary feature selection for deep learning
Authors	Raquel Espinosa, Fernando Jiménez, José Palma
Journal	IEEE Transactions on Neural Networks and Learning Systems
Impact factor (2021)[2]	14.255
JCR Rank (2021)[2]	Computer science, artificial intelligence: 6/145 (D1) Computer science, theory & methods: 4/110 (D1) Computer science, hardware & architecture: 1/54 (D1)
Cited by	5
Publisher	IEEE
Date	12 January 2023
ISSN	2162-237X
State	Published
Contribution	Conceptualization, methodology, visualization, investigation, software, writing, point by point revision of reviewer's comments

[2] At the time of publication of this thesis, the impact factor and JCR rank data for the year 2023 were not available.

Surrogate-assisted and filter-based multi-objective evolutionary FS for DL

- Multi-objective evolutionary problems which combine 4 objectives:
  - *O1* RMSE of a surrogate model.
  - *O2* Number of attributes.
  - *O3* Correlation.
  - *O4* ReliefF.
- Regression problems: air quality and indoor temperature.
- MOEAs: NSGA-II, NSGA-III, MOEA/D, SPEA2, IBEA, ϵ-MOEA, ϵ-NSGA-II.
- New multi-criteria performance metric: $\mathcal{H}$.
- Comparison with other FS methods.



Surrogate-assisted and filter-based multi-objective evolutionary FS for DL

- Multi-objective evolutionary problems which combine 4 objectives:
  - *O1* **RMSE of a surrogate model**.
  - *O2* Number of attributes.
  - *O3* Correlation.
  - *O4* ReliefF.
- Regression problems: air quality and indoor temperature.
- MOEAs: NSGA-II, NSGA-III, MOEA/D, SPEA2, IBEA, ϵ-MOEA, ϵ-NSGA-II.
- New multi-criteria performance metric: $\mathcal{H}$.
- Comparison with other FS methods.

Algorithm 1 Fitness function for objective function *O1* (RMSE of a surrogate model)

Require: $I = \{b_1^I, \ldots, b_w^I\}$ {Individual}
Require: M_R^Φ {Surrogate prediction model built with learning algorithm Φ and trained with dataset R with all attributes}
Require: $V \subset D$ {Validation dataset}
Require: α {Imputation constant}

```
1: V' ← V
2: for i = 1 to w do
3:     if b_i^I = 0 then
4:         for d_t ∈ V' do
5:             d_t^i = α
6:         end for
7:     end if
8: end for
9: return RMSE(M_R^Φ, V') {RMSE of surrogate prediction model M_R^Φ evaluated in dataset V'}
```

Surrogate-assisted and filter-based multi-objective evolutionary FS for DL

- Multi-objective evolutionary problems which combine 4 objectives:
  - *O*1 RMSE of a surrogate model.
  - *O*2 **Number of attributes**.
  - *O*3 Correlation.
  - *O*4 ReliefF.
- Regression problems: air quality and indoor temperature.
- MOEAs: NSGA-II, NSGA-III, MOEA/D, SPEA2, IBEA, ϵ-MOEA, ϵ-NSGA-II.
- New multi-criteria performance metric: $\boldsymbol{\mathcal{H}}$.
- Comparison with other FS methods.

$$\mathcal{C}(\boldsymbol{x}) = \sum_{i=1}^{w} \mathcal{N}(x_i)$$

where $\mathcal{N}$ is a function that transforms a boolean value into numeric ($true = 1$ and $false = 0$) and w is the number of input attributes.

Surrogate-assisted and filter-based multi-objective evolutionary FS for DL

- Multi-objective evolutionary problems which combine 4 objectives:
  - *O1* RMSE of a surrogate model.
  - *O2* Number of attributes.
  - *O3* **Correlation**.
  - *O4* ReliefF.
- Regression problems: air quality and indoor temperature.
- MOEAs: NSGA-II, NSGA-III, MOEA/D, SPEA2, IBEA, ϵ-MOEA, ϵ-NSGA-II.
- New multi-criteria performance metric: $\mathcal{H}$.
- Comparison with other FS methods.

$$\mathcal{P}_R(\boldsymbol{x}) = \sum_{\substack{i=1 \\ x_i=true}}^{w} \rho_i^R$$

where ρ_i^R is the normalized Pearson's correlation coefficient between the selected attribute i and the output in dataset R.

Surrogate-assisted and filter-based multi-objective evolutionary FS for DL

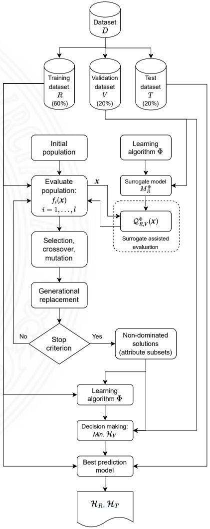
- Multi-objective evolutionary problems which combine 4 objectives:
  - *O*1 RMSE of a surrogate model.
  - *O*2 Number of attributes.
  - *O*3 Correlation.
  - *O*4 **ReliefF**.
- Regression problems: air quality and indoor temperature.
- MOEAs: NSGA-II, NSGA-III, MOEA/D, SPEA2, IBEA, ϵ-MOEA, ϵ-NSGA-II.
- New multi-criteria performance metric: $\boldsymbol{\mathcal{H}}$.
- Comparison with other FS methods.

$$\mathcal{S}_R(\boldsymbol{x}) = \sum_{\substack{i=1 \\ x_i=true}}^{w} \sigma_i^R$$

where σ_i^R is normalized *reliefF* score of attribute i in dataset R.

Surrogate-assisted and filter-based multi-objective evolutionary FS for DL

- Multi-objective evolutionary problems which combine 4 objectives:
  - *O1* RMSE of a surrogate model.
  - *O2* Number of attributes.
  - *O3* Correlation.
  - *O4* ReliefF.
- Regression problems: air quality and indoor temperature.
- MOEAs: NSGA-II, NSGA-III, MOEA/D, SPEA2, IBEA, ϵ-MOEA, ϵ-NSGA-II.
- New multi-criteria performance metric: $\mathcal{H}$.
- Comparison with other FS methods.



Results

Optimization model	$\mathcal{Q}^{\Phi}_{R,V}(\boldsymbol{x})$	Number of selected attributes	$\mathcal{H}_R$	$\mathcal{H}_V$	$\mathcal{H}_T$	Run time (minutes)
O1O2	0.2403	2	0.1234	0.2328	0.1442	9.80
O1O2O3	0.1939	16	0.0807	0.2182	0.1298	15.76
O1O2O4	0.2462	6	0.0928	0.2447	0.1647	18.33
O2O3O4	–	19	0.1006	0.2746	0.1965	7.65
O1O2O3O4	0.2006	4	0.1210	0.2347	0.1447	28.41

Table 10: Prediction model evaluation with NSGA-II for the air quality problem, 100,000 evaluations and 10 runs.

Optimization model	Win	Loss	Win-Loss
O1O2O3	4	0	4
O1O2O4	3	1	2
O1O2	1	2	-1
O2O3O4	1	2	-1
O1O2O3O4	0	4	-4

Table 11: Ranking of multi-objective optimization models.

MOEA	Win	Loss	Win-Loss
NSGA-II	6	0	6
IBEA	4	1	3
ϵ-MOEA	4	1	3
ϵ-NSGA-II	3	3	0
NSGA-III	2	4	-2
SPEA2	1	5	-4
MOEA/D	0	6	-6

Table 12: Ranking of the optimization algorithms with hypervolume metric.

Results

Optimization model	$\mathcal{Q}^{\Phi}_{R,V}(\boldsymbol{x})$	Number of selected attributes	$\mathcal{H}_R$	$\mathcal{H}_V$	$\mathcal{H}_T$	Run time (minutes)
O1O2	0.2403	2	0.1234	0.2328	0.1442	9.80
O1O2O3	**0.1939**	**16**	**0.0807**	**0.2182**	**0.1298**	**15.76**
O1O2O4	0.2462	6	0.0928	0.2447	0.1647	18.33
O2O3O4	–	19	0.1006	0.2746	0.1965	7.65
O1O2O3O4	0.2006	4	0.1210	0.2347	0.1447	28.41

Table 10: Prediction model evaluation with NSGA-II for the air quality problem, 100,000 evaluations and 10 runs.

Optimization model	Win	Loss	Win-Loss
O1O2O3	4	0	4
O1O2O4	3	1	2
O1O2	1	2	-1
O2O3O4	1	2	-1
O1O2O3O4	0	4	-4

Table 11: Ranking of multi-objective optimization models.

MOEA	Win	Loss	Win-Loss
NSGA-II	6	0	6
IBEA	4	1	3
ϵ-MOEA	4	1	3
ϵ-NSGA-II	3	3	0
NSGA-III	2	4	-2
SPEA2	1	5	-4
MOEA/D	0	6	-6

Table 12: Ranking of the optimization algorithms with hypervolume metric.

Results

Optimization model	$\mathcal{Q}^{\Phi}_{R,V}(\boldsymbol{x})$	Number of selected attributes	$\mathcal{H}_R$	$\mathcal{H}_V$	$\mathcal{H}_T$	Run time (minutes)
O1O2	0.2403	2	0.1234	0.2328	0.1442	9.80
O1O2O3	**0.1939**	**16**	**0.0807**	**0.2182**	**0.1298**	**15.76**
O1O2O4	0.2462	6	0.0928	0.2447	0.1647	18.33
O2O3O4	–	19	0.1006	0.2746	0.1965	7.65
O1O2O3O4	0.2006	4	0.1210	0.2347	0.1447	28.41

Table 10: Prediction model evaluation with NSGA-II for the air quality problem, 100,000 evaluations and 10 runs.

Optimization model	Win	Loss	Win-Loss
O1O2O3	**4**	**0**	**4**
O1O2O4	3	1	2
O1O2	1	2	-1
O2O3O4	1	2	-1
O1O2O3O4	0	4	-4

Table 11: Ranking of multi-objective optimization models.

MOEA	Win	Loss	Win-Loss
NSGA-II	6	0	6
IBEA	4	1	3
ϵ-MOEA	4	1	3
ϵ-NSGA-II	3	3	0
NSGA-III	2	4	-2
SPEA2	1	5	-4
MOEA/D	0	6	-6

Table 12: Ranking of the optimization algorithms with hypervolume metric.

Results

Optimization model	$\mathcal{Q}^{\Phi}_{R,V}(\boldsymbol{x})$	Number of selected attributes	$\mathcal{H}_R$	$\mathcal{H}_V$	$\mathcal{H}_T$	Run time (minutes)
O1O2	0.2403	2	0.1234	0.2328	0.1442	9.80
O1O2O3	**0.1939**	**16**	**0.0807**	**0.2182**	**0.1298**	**15.76**
O1O2O4	0.2462	6	0.0928	0.2447	0.1647	18.33
O2O3O4	–	19	0.1006	0.2746	0.1965	7.65
O1O2O3O4	0.2006	4	0.1210	0.2347	0.1447	28.41

Table 10: Prediction model evaluation with NSGA-II for the air quality problem, 100,000 evaluations and 10 runs.

Optimization model	Win	Loss	Win-Loss
O1O2O3	**4**	**0**	**4**
O1O2O4	3	1	2
O1O2	1	2	-1
O2O3O4	1	2	-1
O1O2O3O4	0	4	-4

Table 11: Ranking of multi-objective optimization models.

MOEA	Win	Loss	Win-Loss
NSGA-II	**6**	**0**	**6**
IBEA	4	1	3
ϵ-MOEA	4	1	3
ϵ-NSGA-II	3	3	0
NSGA-III	2	4	-2
SPEA2	1	5	-4
MOEA/D	0	6	-6

Table 12: Ranking of the optimization algorithms with hypervolume metric.

Results

Set	**Metric**	**1-step ahead**	**2-steps ahead**	**3-steps ahead**	**4-steps ahead**	**5-steps ahead**	**6-steps ahead**	**7-steps ahead**
R	RMSE	0.0761	0.0744	0.0749	0.0751	0.0755	0.0758	0.0773
	MAE	0.0533	0.0529	0.0534	0.0535	0.0537	0.0538	0.0548
	CC	0.8892	0.8916	0.8900	0.8885	0.8861	0.8846	0.8805
T	RMSE	0.0814	0.0819	0.0822	0.0829	0.0832	0.0848	0.0873
	MAE	0.0494	0.0498	0.0503	0.0505	0.0507	0.0517	0.0537
	CC	0.7535	0.7508	0.7507	0.7478	0.7467	0.7380	0.7257

Table 13: Results of the best prediction model (*O1O2O3-NSGA-II*) for air quality evaluated on the training set *R* and test set *T*.

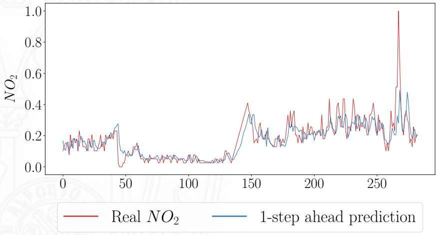


Figure 6: 1-step ahead time series predictions for NO_2 evaluated with *O1O2O3-NSGA-II* with test *T*.

Results

Method	$\mathcal{H}_R$	$\mathcal{H}_T$	Number of selected attributes	Run time (minutes)
O1O2O3-NSGA-II	0.0807	0.1298	16	15.76
M1	0.1246	0.1437	2	3.93
M2	0.1246	0.1437	2	4.09
M4	0.1235	0.1876	6	22.60
M6	0.1701	0.2069	1	2.16
M3	0.1227	0.2243	14	5.45
M5	0.0602	0.2452	84	0.07
All attributes	0.0560	0.2763	84	0.01

Table 14: Comparison of FS methods for the air quality problem.

- ***M1***. Hybrid filter-wrapper based on correlation and LSTM with deterministic search.
- ***M2***. Hybrid filter-wrapper based on ReliefF and LSTM with deterministic search.
- ***M3***. Wrapper multi-objective evolutionary FS method based on LR.
- ***M4***. Wrapper multi-objective evolutionary FS method based on RF.
- ***M5***. CancelOut.
- ***M6***. RF.

Results

Method	$\mathcal{H}_R$	$\mathcal{H}_T$	**Number of selected attributes**	**Run time (minutes)**
O1O2O3-NSGA-II	0.0807	0.1298	16	15.76
M1	0.1246	0.1437	2	3.93
M2	0.1246	0.1437	2	4.09
M4	0.1235	0.1876	6	22.60
M6	0.1701	0.2069	1	2.16
M3	0.1227	0.2243	14	5.45
M5	0.0602	0.2452	84	0.07
All attributes	0.0560	0.2763	84	0.01

Table 14: Comparison of FS methods for the air quality problem.

- ***M1***. Hybrid filter-wrapper based on correlation and LSTM with deterministic search.
- ***M2***. Hybrid filter-wrapper based on ReliefF and LSTM with deterministic search.
- ***M3***. Wrapper multi-objective evolutionary FS method based on LR.
- ***M4***. Wrapper multi-objective evolutionary FS method based on RF.
- ***M5***. CancelOut.
- ***M6***. RF.

Results

Method	$\mathcal{H}_R$	$\mathcal{H}_T$	**Number of selected attributes**	**Run time (minutes)**
O1O2O3-NSGA-II	0.0807	0.1298	16	15.76
M1	0.1246	0.1437	2	3.93
M2	0.1246	0.1437	2	4.09
M4	0.1235	0.1876	6	22.60
M6	0.1701	0.2069	1	2.16
M3	0.1227	0.2243	14	5.45
M5	0.0602	0.2452	84	0.07
All attributes	0.0560	0.2763	84	0.01

Table 14: Comparison of FS methods for the air quality problem.

- ***M1***. Hybrid filter-wrapper based on correlation and LSTM with deterministic search.
- ***M2***. Hybrid filter-wrapper based on ReliefF and LSTM with deterministic search.
- ***M3***. Wrapper multi-objective evolutionary FS method based on LR.
- ***M4***. Wrapper multi-objective evolutionary FS method based on RF.
- ***M5***. CancelOut.
- ***M6***. RF.

Results

Method	$\mathcal{H}_R$	$\mathcal{H}_T$	Number of selected attributes	Run time (minutes)
O1O2O3-NSGA-II	0.0807	0.1298	16	15.76
M1	0.1246	0.1437	2	3.93
M2	0.1246	0.1437	2	4.09
M4	0.1235	0.1876	6	22.60
M6	0.1701	0.2069	1	2.16
M3	0.1227	0.2243	14	5.45
M5	0.0602	0.2452	84	0.07
All attributes	0.0560	0.2763	84	0.01

Table 14: Comparison of FS methods for the air quality problem.

- ***M1***. Hybrid filter-wrapper based on correlation and LSTM with deterministic search.
- ***M2***. Hybrid filter-wrapper based on ReliefF and LSTM with deterministic search.
- ***M3***. Wrapper multi-objective evolutionary FS method based on LR.
- ***M4***. Wrapper multi-objective evolutionary FS method based on RF.
- ***M5***. CancelOut.
- ***M6***. RF.

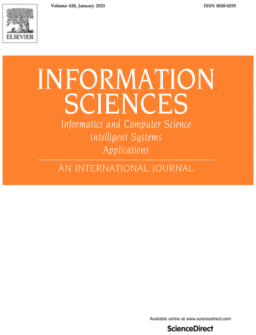
Results

Method	$\mathcal{H}_R$	$\mathcal{H}_T$	Number of selected attributes	Run time (minutes)
O1O2O3-NSGA-II	**0.0807**	**0.1298**	**16**	**15.76**
M1	0.1246	0.1437	2	3.93
M2	0.1246	0.1437	2	4.09
M4	0.1235	0.1876	6	22.60
M6	0.1701	0.2069	1	2.16
M3	0.1227	0.2243	14	5.45
M5	0.0602	0.2452	84	0.07
All attributes	0.0560	0.2763	84	0.01

Table 14: Comparison of FS methods for the air quality problem.

- ***M1***. Hybrid filter-wrapper based on correlation and LSTM with deterministic search.
- ***M2***. Hybrid filter-wrapper based on ReliefF and LSTM with deterministic search.
- ***M3***. Wrapper multi-objective evolutionary FS method based on LR.
- ***M4***. Wrapper multi-objective evolutionary FS method based on RF.
- ***M5***. CancelOut.
- ***M6***. RF.

Multi-surrogate assisted multi-objective evolutionary algorithms for feature selection in regression and classification problems with time series data

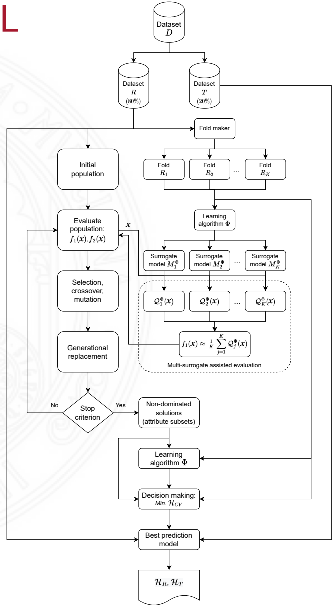


Title	Multi-surrogate assisted multi-objective evolutionary algorithms for feature selection in regression and classification problems with time series data
Authors	Raquel Espinosa, Fernando Jiménez, José Palma
Journal	Information Sciences
Impact factor (2021)[3]	8.233
JCR Rank (2021)[3]	Computer science, information systems: 16/164 (D1)
Cited by	13
Publisher	Elsevier
Date	10 December 2022
ISSN	0020-0255
State	Published
Contribution	Project administration, conceptualization, methodology, data curation, visualization, investigation, software, writing

[3] At the time of publication of this thesis, the impact factor and JCR rank data for the year 2022 were not available.

Multi-surrogate assisted MOEA for FS with DL

- Use of multiple surrogate assisted models.
  - ➢ K-fold cross-validation.
- Learning algorithms:
  - ➢ Regression: RF and LSTM.
  - ➢ Classification: RF and SVM.
- New variability metric to qualitatively analyze FS results: V_X
  - ➢ Differences between multi-surrogate and conventional wrapper approaches.
- Comparison with conventional FS wrapper methods.



Multi-surrogate assisted MOEA for FS with DL

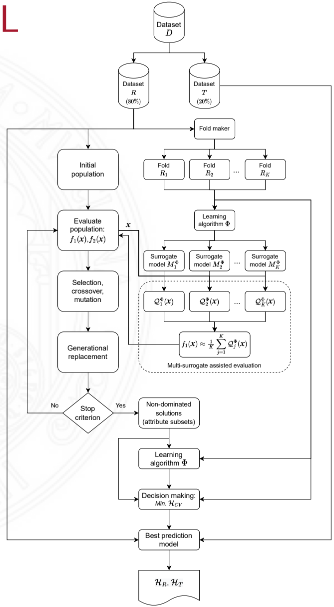
- Use of multiple surrogate assisted models.
  - ➢ K-fold cross-validation.
- Learning algorithms:
  - ➢ Regression: RF and LSTM.
  - ➢ Classification: RF and SVM.
- **New variability metric to qualitatively analyze FS results: V_X**
  - ➢ Differences between multi-surrogate and conventional wrapper approaches.
- Comparison with conventional FS wrapper methods.

$$V_X(A, B) = \frac{1}{C} \sum_{s=1}^{C} \left|\delta_A(\boldsymbol{x}_s) - \delta_B(\boldsymbol{x}_s)\right|$$

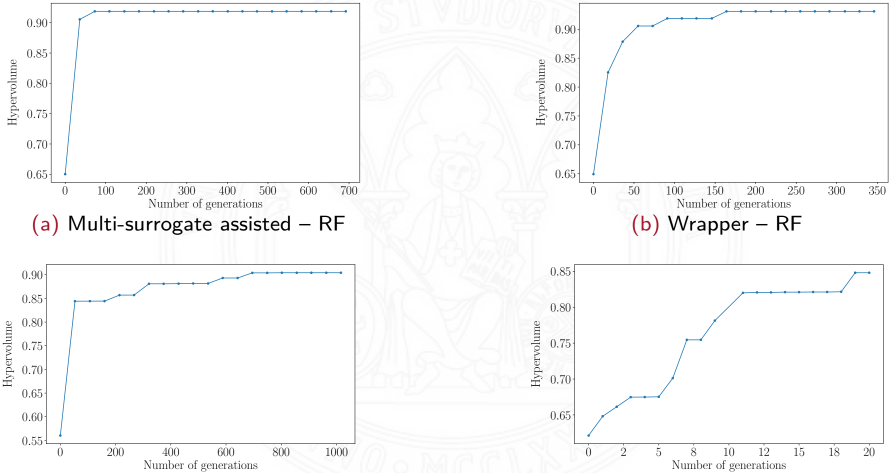
where $\delta_A(\boldsymbol{x}_s)$ is the ranking of the attribute subset $\boldsymbol{x}_s$ obtained with the FS method A and $\delta_B(\boldsymbol{x}_s)$ is the ranking of subset $\boldsymbol{x}_s$ obtained with the FS method B. C is the cardinality of the set X calculated as $C = 2^w - 1$.

Multi-surrogate assisted MOEA for FS with DL

- Use of multiple surrogate assisted models.
  - ➢ K-fold cross-validation.
- Learning algorithms:
  - ➢ Regression: RF and LSTM.
  - ➢ Classification: RF and SVM.
- New variability metric to qualitatively analyze FS results: V_X
  - ➢ Differences between multi-surrogate and conventional wrapper approaches.
- Comparison with conventional FS wrapper methods.



Hypervolume evolution for regression problem



(a) Multi-surrogate assisted – RF

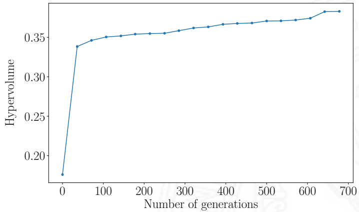
(b) Wrapper – RF

(c) Multi-surrogate assisted – LSTM

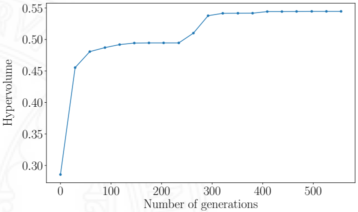
(d) Wrapper – LSTM

Figure 7: Hypervolume evolution of NSGA-II for the air quality regression problem.

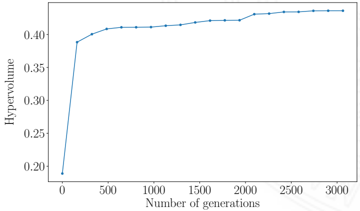
Hypervolume evolution for classification problem



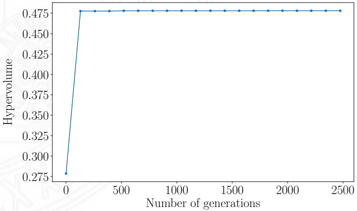
(a) Multi-surrogate assisted – RF



(b) Wrapper – RF



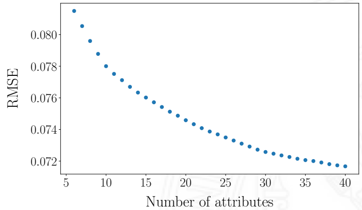
(c) Multi-surrogate assisted – SVM



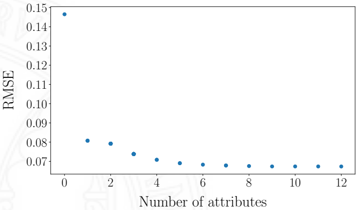
(d) Wrapper – SVM

Figure 8: Hypervolume evolution of NSGA-II for the air quality classification problem.

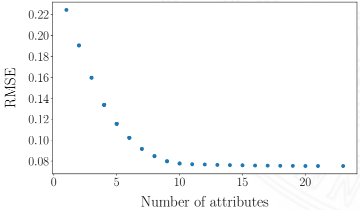
Pareto fronts for regression problem



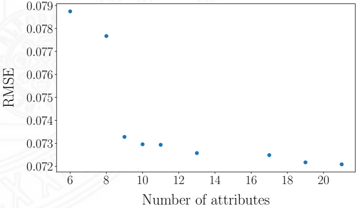
(a) Multi-surrogate assisted – RF



(b) Wrapper – RF



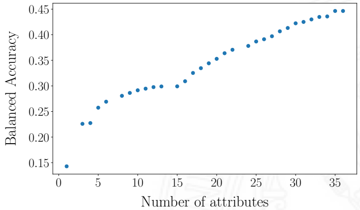
(c) Multi-surrogate assisted – LSTM



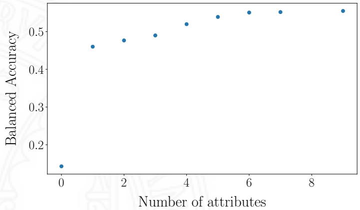
(d) Wrapper – LSTM

Figure 9: Pareto fronts obtained with NSGA-II for the air quality regression problem.

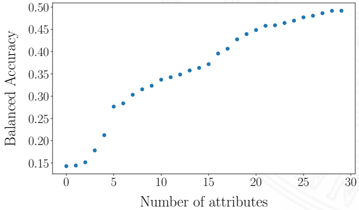
Pareto fronts for classification problem



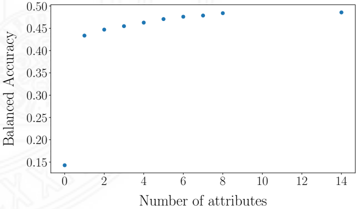
(a) Multi-surrogate assisted – RF



(b) Wrapper – RF



(c) Multi-surrogate assisted – SVM



(d) Wrapper – SVM

Figure 10: Pareto fronts obtained with NSGA-II for the air quality classification problem.

Results

Method	Number of selected atributes	Number of MOEA generations	RMSE	MAE	CC	$\mathcal{H}_R$
Multi-surrogate assisted – RF	39	692	0.0268	0.0178	0.9848	0.0497
Wrapper – RF	10	346	0.0262	0.0178	0.9851	0.0636
Multi-surrogate assisted – LSTM	19	1016	0.0723	0.0501	0.8697	0.1014
Wrapper – LSTM	19	20	0.0702	0.0488	0.8776	0.1023

Table 15: Evaluation of prediction models on the training dataset R for the air quality regression problem.

Method	RMSE	MAE	CC	$\mathcal{H}_T$
Multi-surrogate assisted – RF	0.0533	0.0385	0.8068	0.1816
Wrapper – RF	0.0573	0.0436	0.7728	0.2078
Multi-surrogate assisted – LSTM	0.0489	0.0316	0.8205	0.1217
Wrapper – LSTM	0.0493	0.0322	0.8177	0.1296

Table 16: Evaluation of prediction models on the test dataset T for the air quality regression problem.

Results

Method	**Number of selected atributes**	**Number of MOEA generations**	**RMSE**	**MAE**	**CC**	$\mathcal{H}_R$
Multi-surrogate assisted – RF	39	692	0.0268	0.0178	0.9848	0.0497
Wrapper – RF	10	346	0.0262	0.0178	0.9851	0.0636
Multi-surrogate assisted – LSTM	19	1016	0.0723	0.0501	0.8697	0.1014
Wrapper – LSTM	19	20	0.0702	0.0488	0.8776	0.1023

Table 15: Evaluation of prediction models on the training dataset R for the air quality regression problem.

Method	**RMSE**	**MAE**	**CC**	$\mathcal{H}_T$
Multi-surrogate assisted – RF	0.0533	0.0385	0.8068	0.1816
Wrapper – RF	0.0573	0.0436	0.7728	0.2078
Multi-surrogate assisted – LSTM	**0.0489**	**0.0316**	**0.8205**	**0.1217**
Wrapper – LSTM	0.0493	0.0322	0.8177	0.1296

Table 16: Evaluation of prediction models on the test dataset T for the air quality regression problem.

Results

Method	Number of selected atributes	Number of MOEA generations	BA	AUC	$\mathcal{H}_R$
Multi-surrogate assisted – RF	32	678	1.0000	1.0000	0.0000
Wrapper – RF	9	555	1.0000	1.0000	0.3476
Multi-surrogate assisted – SVM	21	5567	0.8647	0.8314	0.2340
Wrapper – SVM	16	156	0.8230	0.7938	0.4688

Table 17: Evaluation of prediction models on the training dataset R for the air quality classification problem.

Method	BA	AUC	$\mathcal{H}_T$
Multi-surrogate assisted – RF	0.3530	0.8203	0.4471
Wrapper – RF	0.3341	0.7453	0.5122
Multi-surrogate assisted – SVM	0.5437	0.8079	0.4800
Wrapper – SVM	0.3441	0.8039	0.4920

Table 18: Evaluation of prediction models on the test dataset T for the air quality classification problem.

Results

Method	**Number of selected atributes**	**Number of MOEA generations**	**BA**	**AUC**	$\mathcal{H}_R$
Multi-surrogate assisted – RF	32	678	1.0000	1.0000	0.0000
Wrapper – RF	9	555	1.0000	1.0000	0.3476
Multi-surrogate assisted – SVM	21	5567	0.8647	0.8314	0.2340
Wrapper – SVM	16	156	0.8230	0.7938	0.4688

Table 17: Evaluation of prediction models on the training dataset R for the air quality classification problem.

Method	**BA**	**AUC**	$\mathcal{H}_T$
Multi-surrogate assisted – RF	**0.3530**	**0.8203**	**0.4471**
Wrapper – RF	0.3341	0.7453	0.5122
Multi-surrogate assisted – SVM	0.5437	0.8079	0.4800
Wrapper – SVM	0.3441	0.8039	0.4920

Table 18: Evaluation of prediction models on the test dataset T for the air quality classification problem.

Results

Evaluation dataset	Performance metric	1-step ahead	2-steps ahead	3-steps ahead	4-steps ahead	5-steps ahead	6-steps ahead	7-steps ahead
	RMSE	0.0780	0.0790	0.0801	0.0805	0.0807	0.0810	0.0820
R	MAE	0.0549	0.0558	0.0569	0.0573	0.0576	0.0580	0.0586
	CC	0.8469	0.8395	0.8345	0.8313	0.8293	0.8269	0.8235
	RMSE	0.0633	0.0715	0.0746	0.0758	0.0764	0.077	0.0795
T	MAE	0.0435	0.0493	0.0528	0.0545	0.0552	0.0554	0.0578
	CC	0.6967	0.6052	0.5687	0.5533	0.5474	0.5404	0.5068

Table 19: Multi-surrogate assisted MOEA with LSTM (air quality regression problem).

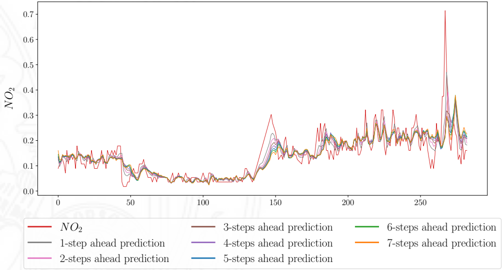


Figure 11: 7-steps ahead forecasting for NO_2 of the multi-surrogate assisted MOEA with LSTM for regression evaluated on test.

Results

Evaluation dataset	Performance metric	1-step ahead	2-steps ahead	3-steps ahead	4-steps ahead	5-steps ahead	6-steps ahead	7-steps ahead
R	BA	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	AUC	1.0	1.0	1.0	1.0	1.0	1.0	1.0
T	BA	0.3270	0.3209	0.3153	0.3158	0.3163	0.3195	0.3212
	AUC	0.8063	0.7911	0.7770	0.7779	0.7787	0.7862	0.7882

Table 20: Multi-surrogate assisted MOEA with RF (air quality classification problem).

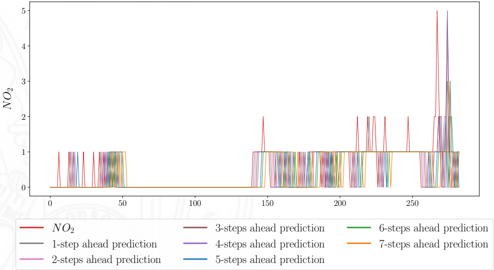
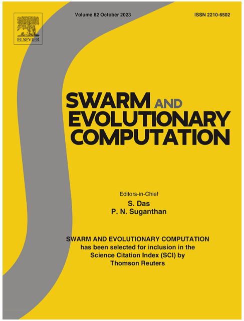


Figure 12: 7-steps ahead forecasting for NO_2 of the multi-surrogate assisted MOEA with RF for classification evaluated on test.

Surrogate-assisted multi-objective evolutionary feature selection of generation-based fixed evolution control for time series forecasting with LSTM networks

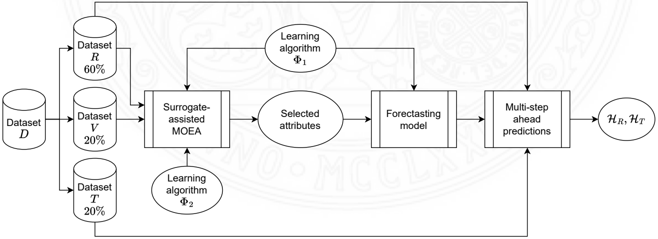


Title	Surrogate-assisted multi-objective evolutionary feature selection of generation-based fixed evolution control for time series forecasting with LSTM networks
Authors	Raquel Espinosa, Fernando Jiménez, José Palma
Journal	Swarm and Evolutionary Computation
Impact factor (2021)[4]	10
JCR Rank (2021)[4]	Computer science, artificial intelligence: (Q1)
Publisher	Elsevier
State	Under review
Contribution	Project administration, conceptualization, methodology, data curation, visualization, investigation, software, writing

[4] At the time of publication of this thesis, the impact factor and JCR rank data for the year 2023 were not available.

Surrogate-assisted MOEA of generation-based fixed evolution control for FS with DL

- Update of the surrogate model.
  - ➢ Updating the dataset and building a new surrogate model.
  - ➢ Incremental learning.
- Configurations of the generation-based fixed evolution control method adjusting the surrogate model update frequency: 25, 50, 100.
- Diebold Mariano test.
- Comparison with other surrogate-assisted approaches.



Results

Evaluation dataset	Performance metric	1-step ahead	2-steps ahead	3-steps ahead	4-steps ahead	5-steps ahead	6-steps ahead	7-steps ahead
	RMSE	0.0834	0.0820	0.0818	0.0819	0.0818	0.0818	0.0815
R	MAE	0.0572	0.0573	0.0575	0.0576	0.0574	0.0574	0.0573
	CC	0.8661	0.8681	0.8683	0.8668	0.8664	0.8655	0.8666
	RMSE	0.0943	0.0975	0.0979	0.0981	0.0983	0.0984	0.0985
T	MAE	0.0727	0.0760	0.0766	0.0769	0.0770	0.0771	0.0771
	CC	0.7535	0.7462	0.7469	0.7466	0.7465	0.7466	0.7467

Table 21: Results of the best model with RF for the air quality forecast problem, evaluated on the training set *R* and the test set *T*.

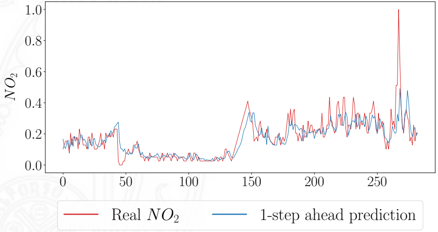


Figure 13: 1-step ahead time series predictions for NO_2 evaluated with RF in test set *T*.

Results

Method	$\mathcal{H}_R$	$\mathcal{H}_T$	Number of selected attributes	Run time (minutes)
Proposed method	0.0909	0.1421	14	25.17
M1	0.1246	0.1437	2	3.93
M2	0.1246	0.1437	2	4.09
O1O2	0.1234	0.1442	2	9.80
M4	0.1235	0.1876	6	22.60
M6	0.1701	0.2069	1	2.16
M3	0.1227	0.2243	14	5.45
M5	0.0602	0.2452	84	0.07
All attributes	0.0560	0.2763	84	0.01

Table 22: Comparison of the best generation-based fixed evolution control model with other FS methods and other surrogate-assisted approach for the air quality problem.

- ***O1O2***. Multi-objective optimization model proposed in *Surrogate-assisted and filter-based multi-objective evolutionary FS for DL*.
- ***M1***. Hybrid filter-wrapper based on correlation and LSTM with deterministic search.
- ***M2***. Hybrid filter-wrapper based on ReliefF and LSTM with deterministic search.
- ***M3***. Wrapper multi-objective evolutionary FS method based on LR.
- ***M4***. Wrapper multi-objective evolutionary FS method based on RF.
- ***M5***. CancelOut.

Results

Method	$\mathcal{H}_R$	$\mathcal{H}_T$	Number of selected attributes	Run time (minutes)
Proposed method	**0.0909**	**0.1421**	**14**	**25.17**
M1	0.1246	0.1437	2	3.93
M2	0.1246	0.1437	2	4.09
O1O2	0.1234	0.1442	2	9.80
M4	0.1235	0.1876	6	22.60
M6	0.1701	0.2069	1	2.16
M3	0.1227	0.2243	14	5.45
M5	0.0602	0.2452	84	0.07
All attributes	0.0560	0.2763	84	0.01

Table 22: Comparison of the best generation-based fixed evolution control model with other FS methods and other surrogate-assisted approach for the air quality problem.

- ***O1O2***. Multi-objective optimization model proposed in *Surrogate-assisted and filter-based multi-objective evolutionary FS for DL*.
- ***M1***. Hybrid filter-wrapper based on correlation and LSTM with deterministic search.
- ***M2***. Hybrid filter-wrapper based on ReliefF and LSTM with deterministic search.
- ***M3***. Wrapper multi-objective evolutionary FS method based on LR.
- ***M4***. Wrapper multi-objective evolutionary FS method based on RF.
- ***M5***. CancelOut.

Time series classification and clustering

Source: Siemens AG

Classification and clustering

Classification

- Time series aggregated for three granularities:
  - Day.
  - Week.
  - Month.
- Time series classification techniques:
  - kNN.
  - Learning shapelets.
  - SAX-VSM.
  - BOSSVS.
  - TimeSeriesForest.

Clustering

- Energy time series grouped by ID.
- 133 time series in total (7 days $\times$ 24 hours).
- Clustering techniques:
  - Silhouette index.
  - k-means.

Classification

	precision	recall	f1-score	support
Energy	0.98	0.97	0.98	24271
Heating	0.72	0.88	0.79	8675
Cold water	0.73	0.72	0.72	26667
Warm water	0.75	0.71	0.73	25747
accuracy			0.80	85360
macro avg	0.79	0.82	0.80	85360
weighted avg	0.80	0.80	0.80	85360

Table 23: Classification report for TimeSeriesForest with day granularity.

	precision	recall	f1-score	support
Energy	0.99	0.99	0.99	2797
Heating	0.78	0.87	0.82	1359
Cold water	0.79	0.78	0.79	3242
Warm water	0.83	0.79	0.81	3264
accuracy			0.85	10662
macro avg	0.85	0.86	0.85	10662
weighted avg	0.85	0.85	0.85	10662

Table 24: Classification report for TimeSeriesForest with week granularity.

	precision	recall	f1-score	support
Energy	0.99	0.98	0.98	825
Heating	0.82	0.87	0.84	521
Cold water	0.78	0.76	0.77	944
Warm water	0.82	0.81	0.81	945
accuracy			0.85	3235
macro avg	0.85	0.86	0.85	3235
weighted avg	0.85	0.85	0.85	3235

Table 25: Classification report for TimeSeriesForest with month granularity.

Clustering

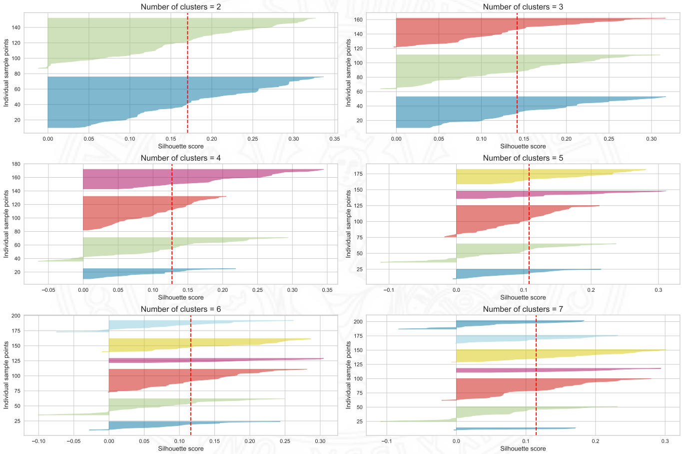


Figure 14: Silhouette score for 2 to 7 clusters for averaged hour of days of the week.

Clustering

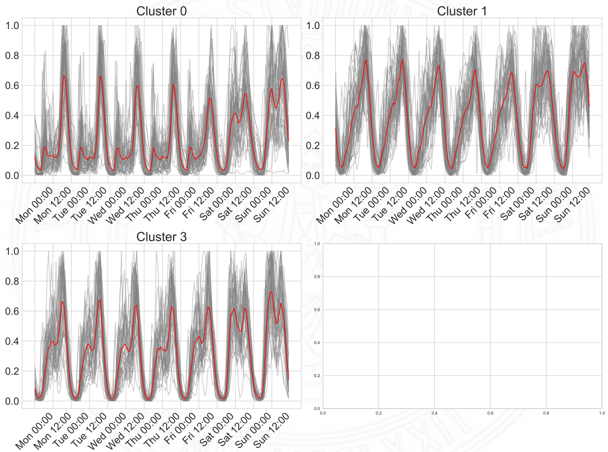
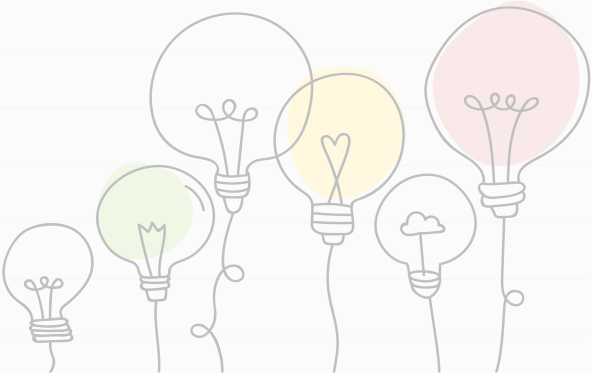


Figure 15: 3 clusters k-means for energy IDs with mean per hour of days of the week.

Overview

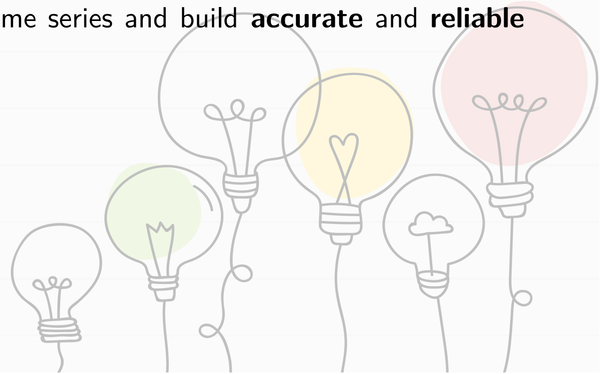
Conclusions I

- The **methodology** for the evaluation and comparison of learning algorithms obtained an **unified** and **adapted results** in order to solve any prediction problem with time series.



Conclusions I

- The **methodology** for the evaluation and comparison of learning algorithms obtained an **unified** and **adapted results** in order to solve any prediction problem with time series.
- **RNN** can capture the complexity of time series and build **accurate** and **reliable** predictive models.



Conclusions I

- The **methodology** for the evaluation and comparison of learning algorithms obtained an **unified** and **adapted results** in order to solve any prediction problem with time series.
- **RNN** can capture the complexity of time series and build **accurate** and **reliable** predictive models.
- **RF** presented a **satisfactory** performance when applied to time series forecasting.

Conclusions I

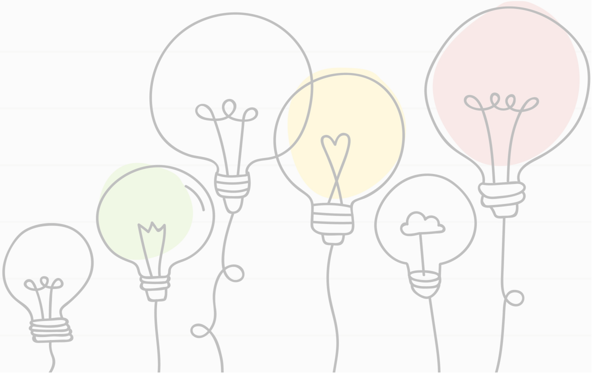
- The **methodology** for the evaluation and comparison of learning algorithms obtained an **unified** and **adapted results** in order to solve any prediction problem with time series.
- **RNN** can capture the complexity of time series and build **accurate** and **reliable** predictive models.
- **RF** presented a **satisfactory** performance when applied to time series forecasting.
- A **multi-criteria decision-making** process pooled several **performance metrics** and established an appropriate comparison between **different learning algorithms**.

Conclusions I

- The **methodology** for the evaluation and comparison of learning algorithms obtained an **unified** and **adapted results** in order to solve any prediction problem with time series.
- **RNN** can capture the complexity of time series and build **accurate** and **reliable** predictive models.
- **RF** presented a **satisfactory** performance when applied to time series forecasting.
- A **multi-criteria decision-making** process pooled several **performance metrics** and established an appropriate comparison between **different learning algorithms**.
- For **air quality forecasting** with time series in an area for which **no information** is available, the prediction has been approximated with **MOEAs** using forecasts from other geographically nearby areas.

Conclusions II

- **Surrogate-assisted MOEAs** have allowed **FS** in expensive problems such as time series forecasting based on **DL**.



Conclusions II

- **Surrogate-assisted MOEAs** have allowed **FS** in expensive problems such as time series forecasting based on **DL**.
- The use of a **surrogate-assisted MOEAs** with a **DL** algorithm for FS has managed to find a satisfactory subset of features in a **shorter computational time** compared to a conventional wrapper-type FS method.

Conclusions II

- **Surrogate-assisted MOEAs** have allowed **FS** in expensive problems such as time series forecasting based on **DL**.
- The use of a **surrogate-assisted MOEAs** with a **DL** algorithm for FS has managed to find a satisfactory subset of features in a **shorter computational time** compared to a conventional wrapper-type FS method.
- **NSGA-II** is the MOEA that has obtained the best results in terms of **hypervolume**, compared to other MOEAs of the state of the art.

Conclusions II

- **Surrogate-assisted MOEAs** have allowed **FS** in expensive problems such as time series forecasting based on **DL**.
- The use of a **surrogate-assisted MOEAs** with a **DL** algorithm for FS has managed to find a satisfactory subset of features in a **shorter computational time** compared to a conventional wrapper-type FS method.
- **NSGA-II** is the MOEA that has obtained the best results in terms of **hypervolume**, compared to other MOEAs of the state of the art.
- **Generation-based fixed evolution control** approach allows **information** to be **efficiently added** to surrogate models within the FS process.

Conclusions II

- **Surrogate-assisted MOEAs** have allowed **FS** in expensive problems such as time series forecasting based on **DL**.
- The use of a **surrogate-assisted MOEAs** with a **DL** algorithm for FS has managed to find a satisfactory subset of features in a **shorter computational time** compared to a conventional wrapper-type FS method.
- **NSGA-II** is the MOEA that has obtained the best results in terms of **hypervolume**, compared to other MOEAs of the state of the art.
- **Generation-based fixed evolution control** approach allows **information** to be **efficiently added** to surrogate models within the FS process.
- **Prediction models** have been identified in various **real contexts** that potentially allow **forecasting** in the near future and that can help institutions to make decisions on environmental issues.

Future work

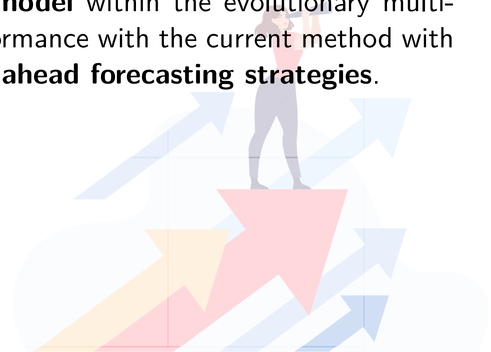
- Include the **FS** process within the **spatio-temporal** approach with LR.

Future work

- Include the **FS** process within the **spatio-temporal** approach with LR.
- Apply the **surrogate** and **multi-surrogate assisted** FS methods for forecasting other air quality time series and compare its performance with NO_2 results.

Future work

- Include the **FS** process within the **spatio-temporal** approach with LR.
- Apply the **surrogate** and **multi-surrogate assisted** FS methods for forecasting other air quality time series and compare its performance with NO_2 results.
- Use other **DL algorithms** as a **surrogate model** within the evolutionary multi-objective FS method and compare their performance with the current method with LSTM. Analyze the use of other **multi-step ahead forecasting strategies**.



Future work

- Include the **FS** process within the **spatio-temporal** approach with LR.
- Apply the **surrogate** and **multi-surrogate assisted** FS methods for forecasting other air quality time series and compare its performance with NO_2 results.
- Use other **DL algorithms** as a **surrogate model** within the evolutionary multi-objective FS method and compare their performance with the current method with LSTM. Analyze the use of other **multi-step ahead forecasting strategies**.
- Combine **architecture search** with **FS** for regression and classification problems with MOEAs.

Future work

- Include the **FS** process within the **spatio-temporal** approach with LR.
- Apply the **surrogate** and **multi-surrogate assisted** FS methods for forecasting other air quality time series and compare its performance with NO_2 results.
- Use other **DL algorithms** as a **surrogate model** within the evolutionary multi-objective FS method and compare their performance with the current method with LSTM. Analyze the use of other **multi-step ahead forecasting strategies**.
- Combine **architecture search** with **FS** for regression and classification problems with MOEAs.
- Apply the proposed FS methods in other fields.

Future work

- Include the **FS** process within the **spatio-temporal** approach with LR.
- Apply the **surrogate** and **multi-surrogate assisted** FS methods for forecasting other air quality time series and compare its performance with NO_2 results.
- Use other **DL algorithms** as a **surrogate model** within the evolutionary multi-objective FS method and compare their performance with the current method with LSTM. Analyze the use of other **multi-step ahead forecasting strategies**.
- Combine **architecture search** with **FS** for regression and classification problems with MOEAs.
- Apply the proposed FS methods in other fields.
- **Embed** FS in **LSTM** networks with multi-objective evolutionary **ensemble learning** for time series forecasting.

thank you!!