Università Politecnica delle Marche
Scuola di Dottorato di Ricerca in Scienze dell'Ingegneria
Curriculum in Ingegneria Biomedica, Elettronica e delle Telecomunicazioni

Advanced Computational Intelligence for Smart Water and Gas Grids

Ph.D. Dissertation of:
Marco Fagiani

Advisor:
Prof. Francesco Piazza

Curriculum Supervisor:
Prof. Franco Chiaraluce

XV edition - new series

Università Politecnica delle Marche
Scuola di Dottorato di Ricerca in Scienze dell'Ingegneria
Curriculum in Ingegneria Biomedica, Elettronica e delle Telecomunicazioni

Advanced Computational Intelligence for Smart Water and Gas Grids

Ph.D. Dissertation of:
Marco Fagiani

Advisor:
Prof. Francesco Piazza

Curriculum Supervisor:
Prof. Franco Chiaraluce

XV edition - new series

Università Politecnica delle Marche
Scuola di Dottorato di Ricerca in Scienze dell'Ingegneria
Facoltà di Ingegneria
Via Brecce Bianche – 60131 Ancona (AN), Italy

Abstract

In the last few years, due to the technological improvement of advanced metering infrastructures, water and natural gas grids can be regarded as Smart Grids, similarly to power ones. However, considering the role played by Computational Intelligence solutions, the number of studies related to applications for distribution grids, and the availability of suitable datasets, the gap between power grids and water/gas ones is notably wide.

At first, in this work the state-of-the-art techniques for water and natural gas grids is presented and discussed, focusing on load forecasting and leakage detection applications. The result of an extensive search of used datasets is also presented, and thus made available to the research community.

Later, experiments concerning the prediction and the leakage detection based on the analysis of water and natural gas consumption records are presented focusing on how to exploit data heterogeneity to get a reliable outcome.

With regard to forecasting experiments, the tests are performed according to two key aspects: homogeneous evaluation criteria and application of heterogeneous data. The evaluations are performed for different grid scenarios, from the residential to the national one.

Concerning leakage detection experiments, an algorithm based on the novelty detection paradigm is presented. Different Computational Intelligence approaches to model the normality condition are investigated, as well as the composition of an optimal set of features, including innovative temporal information and pressure ones. Due to the lack of suitable datasets, the target scenarios are limited to the residential and the office building one.

Finally, in accordance with the methodologies applied above, advancements for two particular power grid applications are also proposed. Specifically, a proper micro-grid energy manager with the support of a pricing profile forecaster, and improvements for non-intrusive load monitoring based on statistical and deep neural network solutions, are presented and evaluated.

Sommario

Negli ultimi anni, il miglioramento delle infrastrutture di contabilizzazione ha permesso di assimilare le reti di distribuzione di acqua e gas alle smart grids ideate per la distribuzione di energia elettrica. Tuttavia, considerando il ruolo delle soluzioni di Computational Intelligence, gli studi relativi ad applicazioni per reti di distribuzione e la disponibilità di dataset adeguati, il divario tra il livello di sviluppo delle reti elettriche e quello delle reti di distribuzione di acqua/gas è particolarmente ampio.

All'inizio di questo lavoro lo stato dell'arte delle tecniche applicate alle reti di acqua e gas viene presentato e discusso, con attenzione alle applicazioni di previsione e rilevamento delle perdite. Viene inoltre presentato il risultato della ricerca dei dataset utilizzati, mettendoli a disposizione per la comunità scientifica.

Successivamente, gli esperimenti di previsione dei consumi di acqua e gas e di rilevamento delle perdite ad essi connesse, sono presentati con attenzione all'utilizzo di dati eterogenei per ottenere un risultato affidabile.

Le previsioni sono eseguite garantendo sia criteri di valutazione omogenei sia l'applicazione di informazioni eterogenee. Le valutazioni sono effettuate per differenti scenari, dal caso residenziale a quello nazionale.

Nel rilevamento delle perdite, gli esperimenti sono eseguiti mediante un algoritmo basato sul paradigma della novelty detection. Insieme alla valutazione di differenti approcci di Computational Intelligence per la modellazione della condizione di normalità, viene selezionato anche un set di features adeguate, comprensivo di informazioni temporali e di pressione. La mancanza di dataset adeguati ha limitato gli scenari valutati a quello residenziale e a quello di edificio adibito ad uso ufficio.

Infine, in linea con le metodologie già applicate, si sono presentate innovazioni per due temi specifici della Smart Grid elettrica. Tali innovazioni riguardano un manager energetico per micro-grid che tenga conto della previsione del prezzo della risorsa e il monitoraggio non intrusivo del carico basato su approccio statistico e su rete neurale.

Contents

List of Figures

List of Tables

List of Tables

Chapter 1

Introduction

In the last years, the rapid spread and exploitation of renewable energy resources, such as solar irradiation, resulted in great efforts of the scientific community from a multidisciplinary perspective. The Smart Grid paradigm has been advanced and it is now a reality in most countries worldwide. One of the key concept in Smart Grid studies is represented by *metering*, i.e. the capability to measure the amount of resources being consumed at a specific grid point. The metering process is of the utmost importance since it allows to monitor the performance of the environment, also in real-time.

The collected data, resulting from the monitoring process, is also useful to extract key information, beneficial to prediction of future demand and the optimal exploitation of energy resources, which represent undoubtedly very hot research topics for utilities and consumers. On the one hand, service providers need to know in advance the demand trend in order to apply efficient management strategies, e.g. to optimize the resource storage, as well as monitoring the losses along distribution pipes, to maximize the quality of the service provided through the grid itself. On the other hand, end users may need to promptly identify undetected leakages in the household network to contain, or avoid altogether, the waste of billed resources and the damages to their property as well.

More and more Computational Intelligence techniques have been developed to overcome the hurdles for a complete implementation of the Smart Grid paradigm in water and natural gas fields. Many techniques, such as those for the demand forecasting, have been adapted from the electric energy research field, where the situation is more advanced as most of the methods are already widely applied in real-case applications [1, 2, 3]. Moreover, technology advancements in the metering and monitoring systems, achieved in both energy efficiency and sensor placement aspects [4, 5, 6, 7, 8], have allowed and facilitated the collection of large amount of data. In particular, concerning wireless sensor networks (WSNs), several crucial results have been achieved, such as the investigation of new low-power communication protocols, i.e. nodes based on Wireless Metering Bus (WM-Bus) protocol [4], and the development of ad-

vanced monitoring system, [6], adopting innovative metering approach [8]. Furthermore, the event known as "the Battle of the Water Sensor Networks", [5], highlighted the importance of the sensors positioning along the network. This awareness has led to the development of specific software for sensor placement [7].

Nevertheless, the literature regarding water and natural gas seems to lack of appropriate databases and studies for the development and/or improvement of suitable Computational Intelligence techniques, with special reference to load forecasting and leakage detection applications, which can have an immediate impact on the quality of service provided by the service utilities.

In fact, the water is a common good and it is important to preserve such valuable resource, at the basis of human life. Whereas, concerning the natural gas, the International Energy Outlook 2013 (IEO2013) [9] confirmed that, with an increase by $2,5\%$ per year, nuclear power and renewable energy are the world's fastest-growing energy sources. However, even now the natural gas remains an essential resource. Indeed, until 2040, the 80% of world energy will continue to be supplied by fossil fuels and, with an increase in global consumption of $1,7\%$ per year, natural gas is the fastest-growing fossil fuel in the outlook. Increasing supplies of tight gas, shale gas, and coalbed methane support growth in projected worldwide natural gas use.

In addition to the scarcity of studies on both smart water and gas grid, especially if compared with the electrical energy literature, the two grids have been jointly addressed because it is believed that achievements in each field can easily cross-fertilize the other. This issue is surely enhanced by the similar characteristics of the grids themselves and related common nature of metering devices used in both grids.

This dissertation is foremost devoted to illustrate and analyse the solutions based on Computational Intelligence approaches for the fulfilment of smart water and natural gas grids with particular focus on forecasting and leakage detection applications. For the former an exhaustive comparison among the state-of-the-art forecasting techniques is conducted evaluating different operational contexts, from the residential scenario to the national one. For the latter a leakage detection approach based on novelty detection, meant for residential and office building distribution, is proposed and properly evaluated.

Additionally, in accordance with the methodologies followed for the Smart Water and Natural Grids applications, for two specific topics of Smart Electrical Grids advancements to current solutions are proposed. A micro-grid energy management, with the support of dynamic pricing strategies, and state-of-the-art Non-Intrusive Load Monitoring (NILM) applications are, respectively, investigated.

The outline of the dissertation is the following.

In order to support research and to propose more and more performing solutions in load forecasting and leakage detection for the smart water and gas grid case studies, it is undoubtedly relevant having a survey that can serve as useful starting point. This motivated the author to produce the survey presented in Chapter 2, where a comprehensive collection of the recent state-of-the-art works and related databases, from 2009 to date, is presented.

A complete overview of the Computational Intelligence techniques adopted for the consumption prediction and the leakages identification is presented in Chapter 3 and it concerns the following methods: Autoregressive integrated moving average (ARIMA) models, Artificial Neural Networks (ANN), Deep Belief Networks (DBN), Echo State Networks (ESN), Autoencoders (AE), Support Vector Machines (SVM), Support Vector Regression (SVR), One-Class Support Vector Machines (OC-SVM), Genetic Programming (GP), Extended Kalman Filter (EKF), Extreme Learning Machines, Gaussian Mixture Models (GMM), and Hidden Markov Models (HMM).

ARIMA, ANN, DBN, ESN, SVR, GP, EKF-GP, and ELM are adopted for the for the consumption forecasting and the achieved performance is presented in Chapter 4. Specifically, the availabler datasets for the residential, office building, city, and national scenarios are evaluated with each technique by investigating all the different configurations and input variables. Moreover, a set of common evaluation criteria is selected and adopted to perform reliable and reproducible evaluations.

A novelty detection based approach for the identification of leakage occurrence is presented in Chapter 5, and HMM, GMM, One-Class SVM, and Autoencoder ANN, are exploited to characterize the normality condition, or background. The validity of the approach is confirmed by the achieved results, proving its suitability for both domestic and office building scenarios. A set of features, actually missing in the related literature, is proposed and an ad hoc selection is performed in order to obtain the best features combination. Temporal and pressure information is also exploited as features, the former ones are evaluated for all the available scenarios, whereas the latter ones are adopted only in the water consumption of the residential scenario by exploiting a synthetic data creation tool.

Applications for the power grid are also evaluated in Chapter 6. The effects of price forecasting over the scheduling process, since forecasting errors may hinder the management activity, are investigated. An energy management is implemented by revising and extending the framework [10]. As such, dynamic pricing is included, and both hourly and daily scheduling approaches are inspected.

Moreover, concerning the unsupervised NILM, improvements for both Additive Factorial Approximate MAP (AFAMAP) and Deep Neural Network (DNN)

approaches are proposed together with an experimental campaign.

Finally, Chapter 7 concludes the dissertation.

Chapter 2

Smart Water and Gas Grids

2.1 Load Forecasting

2.1.1 Problem Statement

Load forecasting aims to provide an estimate for future consumption, at the highest achievable accuracy. In particular, the focus is to develop approaches able to identify specific behaviours in the past data, $x_n \in \mathbb{R}$ with $n = 1, \ldots, N$, to produce models able to generate accurate predictions, $y_j \in \mathbb{R}$ with $j = 1, \ldots, J$, using only the information on recent consumption, x_r with $r = 1, \ldots, R$. Where N is the number of available past data, R denotes the number of inputs assumed for the prediction with $R < N$, and J is the maximum number of predictable values, limited to $N - R$.

Generally speaking, any load forecasting problem in smart grid contexts can be categorized by the prediction horizon, i.e. how far in time, and the target dimension, i.e. the number of consumers determining the load to be predicted. The former can be distinguished in: short-term, medium-term and long-term prediction. It is a loose classification, and slight changes are often encountered among different contributions. Therefore, to avoid confusion and misunderstanding, the classification adopted in this work assumes a *short-term* prediction with a time horizon lower than $48\,h$. The forecast featuring an horizon up to one to six months is addressed as *medium-term*, and from six months onwards, it is *long-term*.

The target dimension has become more and more important in recent years, with the spreading of the *smart-meter* paradigm. At the beginning, due to the difficulty of collecting the consumption of single consumers with an adequate rate, the forecast techniques have been developed and evaluated only with large amounts of aggregate consumption, e.g. cities, districts or distribution lines. Nowadays, the increasing use of *smart-meters* allows the collection of large amount of data, thanks to automatic readings and network connections, from end users (houses, shops or factories) as well as from primary points of the distribution network. So, the novel prediction approaches have to deal

with different origins for the data. In particular, working with city or district consumption reduces the short-term variability of the data, but causes the need to achieve a more precise forecasting, because even a small error can effect dramatically the predicted consumption, and cause wrong decisions. On the contrary, household consumption data presents a high short-term variability, thus, great changes could occur in short times, but larger errors are tolerable due to the reduced resource usage.

The consumption depends on the environmental and climatic conditions. Being able to consider these heterogeneous data is of major concern. Suitable data are, e.g, temperature, humidity, cloudiness values, compatible with the consumption datasets. However, higher complexity models are needed to correctly handle the heterogeneous information, therefore the application of heterogeneous data must be carefully considered. Unfortunately, higher complexity does not always mean better performance.

For these reasons, finding a suitable dataset is crucial to perform a proper analysis for a given scenario. Moreover, the collected data, possibly heterogeneous ones, must have a temporal length adequate to the prediction horizon and suitable to perform a valuable prediction: training and validation data have to be sufficiently populated.

2.1.2 State of the Art

Among load forecasting techniques, the SVM has been widely exploited to deal with forecasting problems. Among its variations, the least squares SVM (LS-SVM) has raised a strong interest, and recently Ji *et al.* [13] and Zhu and Chen [15] have proposed the ameliorate teaching-learning-based optimization (ATLBO) and the quantum particle swarm optimization based on phase encoding (PQPSO), for the selection of the optimum parameters in urban water consumption forecasting procedure.

Specifically, the ATLBO [13] has improved the standard teaching-learning-based optimization (TLBO). By selecting and keeping the best elite individuals (learners) for the next generation, ATLBO enhances the convergence ability of its predecessor. An adaptive teaching factor, also, has been introduced in the teaching phase and, additionally, the self-monitoring mechanism in the learner phase has been improved. The proposed enhancements have allowed to gain a better regression precision than TLBO, particle swarm optimization (PSO) and grid search. The water dataset has been collected in Shanghai (Qingcaosha water supply system, line 2) from June 1, 2012 to May 31, 2013, and also heterogeneous information have been used, i.e., the maximum and minimum temperatures, the precipitations, and the holiday information. The overall results for each technique have been evaluated in terms of mean relative error

Table 2.1: Summary of the existing contributions in the load forecasting field and their relative performance.

Water contr.:	Technique	Evaluation Criteria	Performance
[11]	EMD-ANN	RMSE - MAE - R^2	$2.71, 1.86, 0.54$
[12]	ARIMA-RC	RAE - MAPE - RMSE - PE	$-, 8.70\%, -, -$
[13]	ATLBO-LSSVM	MRE	0.41%
[14]*	Adaptive	MAPE - RRMSE - R^2	$2.91\%, 4.35\%, 0.750$***
[14]**	Adaptive	MAPE - RRMSE - R^2	$6.19\%, 9.63\%, 0.957$***
[15]	PQPSO-LSSVM	RE - MRE	1.76%
[16]	QPSO-RBF ANN	RE - MSRE	2.14%
[17]	EKF-GP	NMSE - R^2	0.23, 0.76
[18]	GM(1,1)	Max RE - MRE	$-19.50\%, 13.85\%$
[18]	RBF ANN	Max RE - MRE	$2.29\%, 2.01\%$
[19]	Fuzzy	MSE - NMSE - R^2 - MAPE	$0.042, 0.465, 0.760, 7.63\%$
[19]	Neuro-Fuzzy	MSE - NMSE - R^2 - MAPE	$0.008, 0.069, 0.931, 2.85\%$

*: 24-h time step. **: 15-min time step. ***: average values.

Gas contr.:	Technique	Evaluation Criteria	Performance
[20]	GM-Mc	Raw values and their difference	$-, -$
[21]	Step-DE-GM	PE - MAPE	$6.35, 2.28$
[22]	GM-BPNN	PE - MAPE	$2.7, 0.55$
[23]	GD-FNN	RE	–
[24]$^{+}$	ANN	RE	$1.5\% - 6.8\%$
[24]$^{++}$	ANN	RE	$0.33\% - 1.86\%$
[25]	MWS	RMSE - MAPE - wMAPE	–

$^{+}$: daily prediction. $^{++}$: monthly prediction.

(MRE) and mean square error (MSE).

The novel method PQPSO [15] has improved the quantum particle swarm optimization (QPSO) approach by expressing the state information as phases of a qubit, rather than as a linear combination of 0 and 1, and by introducing a method of adaptive adjustment of the inertia factor. The proposed method has demonstrated better prediction accuracy and computing speed than the approaches based on SVM and LS-SVM. The system performance has been evaluated on a database composed of water consumption and meteorological data from April 1 to July 10, 2010. The adopted evaluation criteria have been the relative error (RE) and the MRE.

Recently, also due to the promising improvements accomplished by deep belief network (DBN) and recurrent neural network (RNN), new attention has been paid to neural network applications.

Concerning urban water prediction, a recent innovation has been proposed by Zhu and Xu [16]. Specifically, the authors have combined the quantum particle swarm optimization (QPSO) algorithm with a radial basis function (RBF) neural network, in order to calculate the parameters (weights) of the net and to achieve an higher accuracy level of the prediction. The effectiveness of the approach has been confirmed by experimental results, as well as the higher convergence speed respect to the standard RBF network. The urban water consumption and meteorological data form April 1, 2010 to July 10,

2010 have been adopted as dataset. The evaluation criteria are the RE and the mean square relative error (MSRE).

A neural network, with a logistic transfer function, has been also adopted by Azari and Shariaty-Niassar [24], in order to predict the daily and monthly gas consumption. The approach has been tested by adopting heterogeneous data. Specifically, for the daily prediction the meteorological parameters, the gas consumption data for the previous five days, and the meteorological parameters forecasting for prediction day, have been adopted as input information. Whereas, only the monthly effective temperature for the previous and the predicted month, and the gas consumption for the previous month, have been used for the monthly prediction. The tests has been performed with gas consumption and meteorological data of Tehran from March 21, 2001 to August 8, 2005, and the only evaluation criteria adopted is the RE.

Shabri and Samsudin [11] have exploited a neural network in order to compose an hybrid integrating empirical mode decomposition (EMD) to forecast monthly water demand series. In the EMD method the original data is decomposed into a sum of intrinsic mode function (IMF) components with individual intrinsic time scale properties. Experimental results have confirmed the better performance of the EMD method combined with the neural network (EMD-ANN), over the use of the neural network only. Moreover, an autoregressive integrated moving average (ARIMA) model has been also adopted in the EMD, but both EMD-ARIMA and ARIMA approaches have provided worst results respect to the EMD-ANN. The Batu Pahat city water consumption from January 1995 to December 2011 have been used to test the approaches, and the root mean square error (RMSE), the mean absolute error (MAE), and the coefficient of correlation (R) have been adopted as evaluation criteria.

Another well-known method, adopted to address the prediction problem, is the regressive model theory and its variations. Recently, contributions in that field have been provided by Quevedo *et al.* [26], Broen *et al.* [27], and Akpinar and Yumusak [12].

ARIMA, basic structural and exponential smoothing models have been used by Quevedo *et al.* [26], in order to produce hourly water prediction. The results of the models have been compared by means of the explained variance (EV), the MAE, the MSE, and the mean absolute percentage error (MAPE). The dataset has been composed of sampled values from 100 pressure sensing points in the Barcelona Water Network.

As pointed out by Brown *et al.* [27], many factors need to be taken into account in order to obtain a proper prediction. Therefore, in order to produce more accurate long-term predictions of natural gas, the authors have proposed a novel de-trending algorithm. The de-trending approach, based on linear regression, exploits temperature information, specifically the heating degree

day, the change in heating degree day, and the cooling degree day. A comparison against state-of-the-art de-trending approaches, has proven the performance improvement of the proposed approach in the forecast of daily gas consumption. The adopted dataset, composed of natural gas consumption series from a U.S.-based local distribution company and temperature data, spans over 15 years, and the RMSE and the weighted mean average percentage error (WMAPE) have been adopted as evaluation criteria.

As with the previous contribution, in order to achieve an higher accuracy with an ARIMA model, Akpinar and Yumusak [12] have proposed the removal of the "*cycling component*" from the data series. The better performance of the models without "*cycling component*" have been proven in the experimental results, evaluated by means of the relative absolute error (RAE), the MAPE, the RMSE, and the standard percent error (PE). The tests have been performed by using the daily gas consumption from 2009 to 2012 of Sakarya city, Turkey.

Bakker *et al.*[14] have presented an adaptive forecasting model to predict the short-term water consumption. Heterogeneous information, i.e., static calendar data, have been adopted to improve the accuracy of the prediction. The model has been composed as a regression model in which the main contributions are updated at each new input data. These main contributions are the average forecasted demand for the next 48 hours, the normal forecasted demand for the 15 minutes step, and the extra sprinkle forecasted demand for the 15 minutes step. The datasets of urban water demand are collected over six different areas in the period 2006-2011 with a sample rate of 15 minutes. The model performance have been evaluated by using the RE, the MAPE, the relative root mean square error (RRMSE) and the determination coefficient (R^2).

The grey system theory has been widely adopted, as well, in combination with other well-known techniques, as recently presented by Wan *et al.* [20] and Wang *et al.* [21, 22].

For instance, Wan *et al.* [20] have proposed a combination of grey model and Markov chain to predict the annual natural gas demand in Chongqing. The results have been evaluated in terms of predicted value and difference between predicted and actual value.

In their first contribution, Wang *et al.* [21], have adopted the grey theory to enhance the differential evolution (DE) algorithm, by producing a novel model called Step-DE-GM. Whereas, in the latter contribution, [22], the grey model has been combined with back propagation neural network in order to forecast the urban water consumption and to deal with an insufficient amount of data, required for a robust optimization, or train, of the network. The used dataset is composed of historical annual data from 2001 to 2012 compose the dataset adopted in both the contribution, and the performance have been evaluated in term of percentage error (PE) and mean absolute percentage error (MAPE).

A non linear combination forecasting model, based on generalized dynamic fuzzy neural network (GD-FNN), has been proposed by Chen *et al.* [23] to predict daily gas demand. The fuzzy rule, that defines the FNN structure, is not predetermined, and it may change during the learning process. Moreover, the elliptic basis function has been adopted in order to allow more flexibility and a wider range of non-linear transformation. The daily gas load of Hardin, China, from April 17, 2008 to June 29, 2008 has been used to evaluate the performance, and the RE has been selected as evaluation criterion.

2.2 Leakage Detection

2.2.1 Problem Statement

Leakage detection is meant to recognize and to report, to the end user or the utility, and with minimal error, the presence of anomalous consumption levels in the monitored grid, e.g., due to pipe break.

Anomalies detection, clearly, requires a model of the normal condition of the system and, also, the ability to point out the outliers in the acquired data. To achieve this goal, a suitable set of features is mandatory, to manipulate and manage the data in an efficient manner. Indeed, ad hoc features can lower the computational burden, as well as support latent information extraction.

It should be noted, however, that a proper model is strictly bound to the data variability, thus it strongly depends on the network size. In fact, city or district aggregated consumption records show low short-term variability of data, thus, unexpected fast events are easily recognizable. On the contrary, household or building consumption data present a high short-term variability, thus, great changes in water consumption data are the norm, therefore, the unexpected, fast events are masked and hardly detectable. For these reasons, finding suitable datasets and features is crucial to perform a proper analysis for each given scenario. Moreover, in residential and building scenarios, the consumption levels, thus flow data, is collected in a single sensing point.

Other than data variability, even the leaks characteristics need to be modelled. For instance, the spilled amount of water or gas differs greatly when district sized grids and residential plumbings are considered.

Even the detection time is crucial. In fact, when a leakage occurs, the spill goes on until the piping is repaired.

2.2.2 State of the Art

In literature many contributions have been produced to address the leakage/fault detection problem in the industrial environment. But, being mainly

Table 2.2: Summary of the existing contributions in the leakage detection field and their characteristics. The "Target" columns report the target scenario, thus Residential (R) or District (D). The "Data" column reports the used data type, Flow (F) or Pressure (P), while the resource types, Water (W) or Gas (G), are reported in the column "Resource". MNF, Minimum Night Flow.

Contr.	**Target**	**Data**	**Resource**	**Technique**	**Main features**
[28]	D	P	W	SVM	44 junction sensors No novelty approach
[29]	D	F	W	ICI-CDT	High delay Change detection approach
[30]	D	F & P	W	MNF-MLR	Indirect detection 2 sensors No novelty approach
[31]	-	P	W	C4.5	Laboratory circuit 15 sensors No novelty approach
[32]	D	F & P	W	SVR	Anomalies detection
[33]	R	F & P	W	SVM and ANN	No real data 6 sensors No novelty approach
[34]	R	F	W	Heuristic	Ad hoc constraints No validation
[35]	D	F	W	Fuzzy logic	2 sensors No novelty approach
[36]	-	F & P	G	SVM	5 acoustic and pressure sensors High-pressure pipe No novelty approach

aimed at oil and natural gas pipelines, they are based on data collected at high sampling rates and depend on intrusive and/or manual (i.e., operated by person) detection techniques. Furthermore, the data could be also collected by multiple sensing points arranged along the pipeline. In the urban distribution network of the utilities however, sampling rates are usually low, whereas sensing points cannot be arranged along each branch of the piping. In fact, they are usually placed at very specific points and, of course at the end users' home. Because of these conditions, techniques aimed at the industrial environment are not suited to address the living one. Since these methods are too invasive for residential environments, only the state of the art for civil resources' distribution is presented below. Studies based on district network data have been also included to evaluate the suitability of the proposed technique for office buildings, scenarios that have not been addressed, as of yet, in the current literature, as shown in Table 2.2.

Common computational intelligence techniques, such as ANN and support vector regression (SVR), have been exploited by Nasir *et al.* [33]. To achieve the estimated position and size of water leakages, an EPANET [37] simulation of a residential network has been performed, and the raw data, acquired by 2 pressure sensors, 2 differential pressure sensors, and 2 flow sensors, have been

used to predict the leakages parameters. The performance has been evaluated in terms of MSE and squared correlation error coefficient (R-square). The proposed quasi-static analysis confirmed the good behaviour of the SVM and its resilience to sensor measurement errors.

SVM has been also exploited by Salam *et al.* [28], in order to analyse the pressure change pattern when a leakage occurs, therefore to identify the leakage position and size. A real network system of a district area has been reproduced with EPANET [37], and all the pressure data collected at the junctions have been used. The overall performance have been reported in terms of RMSE.

In order to generate a set of rules to categorise the features data, as *leakage* or *leakage-free*, Gamboa-Medina *et al.* [31] have used the algorithm C4.5. Water pressures data, at high sample rate, have been collected from a controlled experimental laboratory circuit, and a set of 4 features have been extracted. The single features and their combinations have been evaluated in terms of receiver operating characteristic (ROC) and area under curve (AUC).

Among the leakage detection techniques applied to large networks, the monitoring of the minimum night flow (MNF) is widely used. Recently, Alkasseh *et al.* [30] have exploited a multiple linear regressions method to correlate the overall loss and the number of connections, the total length of pipe, and the weighted mean age pipe of the network. The difference between the actual MNF and the estimated one, that could allow to establish the presence of a leakage, has been evaluated by using the R and the R-square.

Even fuzzy logic has been applied to both detection and localization of leakages in water networks by Sanz *et al.* [35]. The proposed Fuzzy Inductive Reasoning (FIR) approach has been applied to data collected by two pressure sensors located in a district network, whereas the leakage data have been synthetically created by means of the EPANET [37] software. The detection performance have been evaluated by pointing out the total amount of leakages detected in the various experiments.

The only contribution that reports an approach developed directly for domestic water systems, has been presented by Oren and Stroch [34]. The authors have proposed a mathematical model, based on the definition of threshold values regulated by means of average domestic water usages. The approach exploits the data acquired in one sensing point, but neither exhaustive experiments nor evaluation criteria have been presented.

Even the novel change detection test (CDT), developed by Boracchi and Roveri [29], has been developed adopting only the information collected in a single flow sensing point in the Barcelona water distribution network. The approach allows to detect structural changes in the time series, and tests with different type of manipulations have been carried out, i.e., leakage, sensor degradation, source change, and stack-at. The evaluation criteria have been directly

derived from the ones applied in novelty detection: False Positive Rate (FPR), False Negative Rate (FNR) and Detection Delay (DD).

2.2.3 Novelty Detection

Novelty detection techniques raised deep interest in many application fields, and a large corpus of literature has been produced so far [38, 39, 40]. Among these, non-intrusive techniques [41, 42] have been deemed of particular interest since they do not require additional tools, other than the water meter, to locate the leakage.

The state-of-the-art methods present many shortcomings, in addition to the number of data sources, even from the computational approach standpoint. In particular, the experimental methodology adopted in [30], [31], [33], [35] and [36] is incompatible with the novelty detection paradigm [43, 44], in which only the normal behaviour is known. In [31], the binary classifier has been trained to discriminate two classes: *leak* and *no-leak*. In [33], ANN and SVM have been trained to discriminate the leakage location and size; whereas [30] do not discriminate the presence of a leakage, but a linear correlation between the MNF (Minimum Night Flow) loss and the number of connections, the total length of the pipe network and the pipes' weighted mean age has been explicated.

On the other hand, in [32], [34], and [29], no faulty instances have been used to train the system, and the leakage scenario has been adopted only during the validation stage, in accordance with the novelty detection concept.

Unfortunately, these approaches show some limitations, as well. A CDT approach with no temporal limitation between the beginning of the anomalies and their detection, causing long detection delays in the identification, has been proposed in [29]. Whereas, [34] and [32] did not present a standard evaluation criterion.

2.3 Databases

In this section, the published state-of-the-art databases are reported. Table 2.3 reports the resource type, the reference period of the data collected, the number of data samples and the database public availability for each contribution. The authors have contacted researchers responsible for the databases creation and maintenance, in order to receive further information about the collected data and their public availability.

To evaluate two fault diagnosis approaches, Quevedo *et al.* [45] used the data acquired from the Barcelona water transport network. The database is composed of the received real-time data from 200 control points, which mainly

Table 2.3: Summary of the datasets for water (W) and natural gas (G) in the literature. The "Contr." column reports the evaluated contributions and corresponding addressed resource is reported in the "Res." column.

Contr.	Res.	Time Period	Samples	Availability
[29]	W	82 days per 10 time series	118,080	**publicly avail.**
[13]	W	Jun 2012 – May 2013	8,736	–
[45]	W	–	–	not public
[46]	W	35,000 days	–	not public
[47]	WGE	Apr 2012 – Mar 2013	$524\,k$*	**publicly avail.**
[14]	W	2006 – 2011	210,336	not public
[48]	W	11 Nov – 22 Dec, 2012	–	not public
[16, 15]	W	Apr – Jul, 2010	–	not public
[17]	W	May 1992 – Dec 2002	125	**publicly avail.**
[18]	W	1990 – 2000	10	not public
[19]	W	1991 – 2003	$\sim 4,380$	not public
[49]	WG	2011 – today	$\sim 10\,k$	**publicly avail.**
[50]	G	1973/1989 – 2013	492/300	**publicly avail.**
[24]	G	Mar 2001 – Aug 2005	$\sim 1,611$	not public
[25]	G	Jun 1996 – Mar 2012	$\sim 5,752$	not public

*: samples for each measured resource.

include flow meters and also some pressure sensors. No further information have been provided by the authors and the database is not publicly available.

In the study proposed by Cardell-Oliver [46], the water consumption data have been collected by a wireless sensor network. The network reaches about 11,000 smart meters (one per household) and the database consists of some selected information over 35,000 days. Unfortunately, the database is not publicly available, nor is it likely to be released, as confirmed by the author.

Makonin *et al.* [51] presented AMPds (Almanac of Minutely Power dataset), "a dataset that contains detailed measurements not seen in other databases". The dataset is composed of power, water and natural gas meter data. The energy consumption of a single house has been recorded using 21 sub-meters for an entire year (from April 1, 2012 to March 31, 2013) at one minute read intervals. In detail, for natural gas metering there were two meters: the whole-house meter (WHG) and the gas furnace meter (FRG). For the water metering there were also two meters: the whole-house meter (WHW) and the hot water meter (HTW). Nevertheless, being the dataset composed of values recorded from a single residence, it is not suitable for load forecasting and leakage detection purposes.

Bakker *et al.* [14] tested their forecasting method on a database composed of

the water demand, over 6 years, of several areas in the Southern region of The Netherlands. The reference period goes from 2006 to 2011 with a 15-min time step in m^3/h, i.e. a total of 210,336 values. For these reasons, the database is really suitable for research purposes, but it is not publicly available.

Boracchi *et al.* [48] used flow and pressure data collected in two inlets as well as five pressure values, measured by monitoring sensors right inside the pipelines. The information have been recorded from 11^{th} November to 22^{nd} December 2012. The authors have not requested the database since the recorded period is too short to test a forecasting system.

Zhu and Xu [16] tested their forecasting approach using the urban water consumption from April to July, 2010, of a Chinese city. Unfortunately, no further information have been provided.

Data on the daily water consumption of Tehran, from 1991 to 2003, have been used by Tabesh and Dini [19]. Presently, the database provided by the Tehran Water and Wastewater Company is not publicly available and no additional information have been supplied.

The dataset used by Nasseri *et al.* [17] is available and it is composed of the monthly water demand of Tehran, from May 1992 to December 2002.

Liu and Chang [18] have employed the annual water demand from 1990 to 2000, including urban domestic water, rural domestic water, industrial water and agricultural water. Despite presenting a wide time interval, both databases ([17, 18]) have been created from data collected over 10 years, but a monthly or an annual water demand can not completely fulfill the needs for a research study in smart metering and load forecasting. For this reasons, no further information have been requested to the authors.

A suitable database of natural gas consumption has been used by Azari *et al.* [24]. The daily consumption of gas has been recorded from 21 March 2001 to 8 August 2005, for a total of 1611 collected data. So far, the database is not publicly available.

In the thesis by Pang [25], the tests have been executed using a dataset of daily natural gas demand from January 1996 to March 2012. The data have been supplied by the GasDay Lab at Marquette University, and, presently, are not publicly available.

For grids larger than the residential, a valid candidate, that includes both water and natural gas recordings, is the dataset of monthly consumption, released in the "live data page for energy and water consumption" of the Department of International Development [49]. The records report energy, natural gas, and water consumption of the headquarters and Abercrombie House buildings. Each DFID subset has a sampling time of 30 minutes, and the last data that have been recorded dates back to March 31$^{\text{st}}$, 2015. Two subsets report gas consumption values, specifically the overall DFID and the Abercrombie House

sets; whereas, the water consumption values are reported in the Abercrombie House and Whitehall sets. Moreover, the records for the Abercrombie House water and gas sets started at August 8th, 2013, and July 28st, 2010, respectively. At August 1st, 2010, for the overall gas set, and at April 20th, 2013, for the Whitehall set.

The monthly natural gas consumption of the U.S.A. are available on the EIA site [50]. The U.S. consumption data are available starting from 1973, whereas the monthly gas consumption data of the single states are reported from 1989. Also, for each state the temperature data are available at the National Climatic Data Center [52].

2.4 Remarks

One of the most relevant result of this overview, is represented by the dearth of suitable databases in the literature, which acts as a serious bottleneck for the development of innovative computational intelligence techniques and approaches for load forecasting and leakage detection. Indeed, as cleared above, some of the databases show short data recording period or very low sample rate, and most are unfortunately unavailable, being therefore not useful for scientific exploitation. This yields to the presence of non-standard evaluation criteria and the inability to perform a comparison between different approaches, since each new method has been tested with a different database.

A great variety of methods and optimization techniques have been examined in the state of the art. For both water and natural gas, the methods have been designed to fulfil only medium or long-term forecasts, and only city consumption have been adopted. The only contribution that has been developed for short-term prediction [14] uses data collected over six different large areas. No one proposed evaluations and/or approaches applied to small targets, namely home and building environments. Additionally, the information available, about the performance of the approaches, are in many aspects inadequate to supply a clear understanding of the achieved performance. Moreover, it is impossible to discuss an objective comparison since each contribution presents a different set of evaluation criteria, making the reported results unsuitable to be compared with those of other contributions. The lack of common evaluation criteria is more evident for contributions [16, 15]. Despite the contributions present the same main author, the RE and the MRE are used to evaluate the approach in the first contribution, whereas the RE and the MSRE are adopted for the latter one.

As seen so far, many contributions have presented approaches aiming to provide forecasts with high accuracy levels for urban demands of water and natural gas. Actually, none of the contributions targets the consumption forecast of

a single house, that is, none can be used with the data collected by domestic metering systems and, thus, representing the demands of a single household. As reported in Section 2.3 this is partially due to the lack of suitable databases, because either they are not available, as well as not publicly available, or they do not provide suitable data, due to low time resolution and/or too short time series.

Moreover, in literature are often presented very promising techniques and results, but in many aspects the information supplied are inadequate to provide a comprehensive and objective comparison among the contributions. First of all, for various contributions it is difficult to have a clear understanding of the achieved performance, and homogeneous evaluation criteria among the contributions are not provided. Secondly, most of the contributions have not adopted common databases, for the same problem of availability cited above.

As seen so far, in the recent literature on leakage detection aimed at residential and district networks, none of the contributions has taken into account the natural gas. Moreover, most of the contributions, [30, 31, 33, 35], have developed approaches by assuming that the input data were composed of flow and/or pressure data collected in multiple sensing points. Therefore, these approaches result unsuitable for residential application, where only one sensing point is available. In addition, even the suitable ones, [29, 34, 32], show a few shortcomings from the computational approach standpoint.

As reported in Sections 2.1.1 and 2.2.1, datasets could be classified according to types of recorded consumption. Four main categories may be defined for the examined datasets: domestic, office building, district and city consumption. Most of them belong to the last category, and only one dataset has been found for the district consumption, as well as only one for domestic consumption and office building. Unfortunately, many datasets were not available, or not publicly available, as shown in Table 2.3. In addition, the few available ones do not provide suitable data, due to low time resolution and/or too short time series.

Chapter 3

Background

3.1 Neural Networks

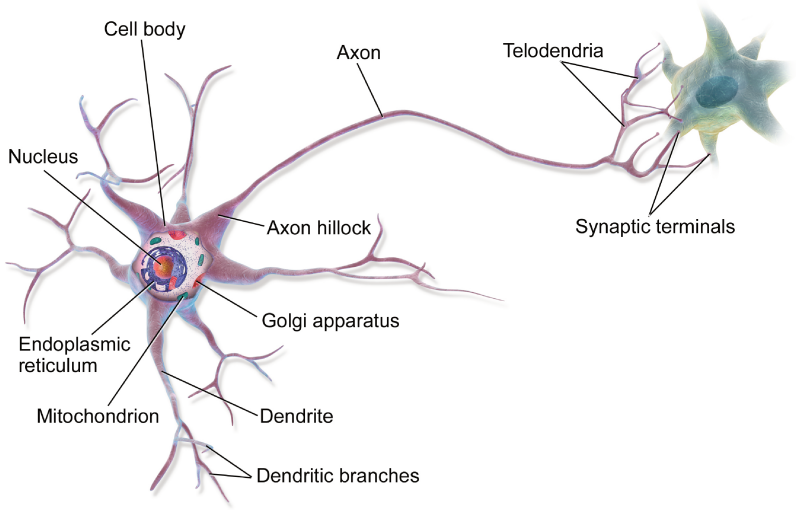


Figure 3.1: Biological multipolar neuron [53].

In order to mimic the brain capabilities a first model of artificial neuron, namely Threshold Logic Unit (TLU), has been proposed by McCulloch and Pitts *et al.* [54]. After many years, inspired by the biological neuron, Figure 3.1, the original idea has been improved to accomplish the nowadays artificial neuron model illustrated in Figure 3.2. As the biological neuron, the artificial one presents the same fundamental parts:

- dendrites: propagate the electrochemical stimulation received from other neural cells to the cell body, or soma;
- cell body (soma): sum all the received electrochemical stimulations;

- axon: reached a certain potential conducts the electrical impulses away from the neuron's cell body.

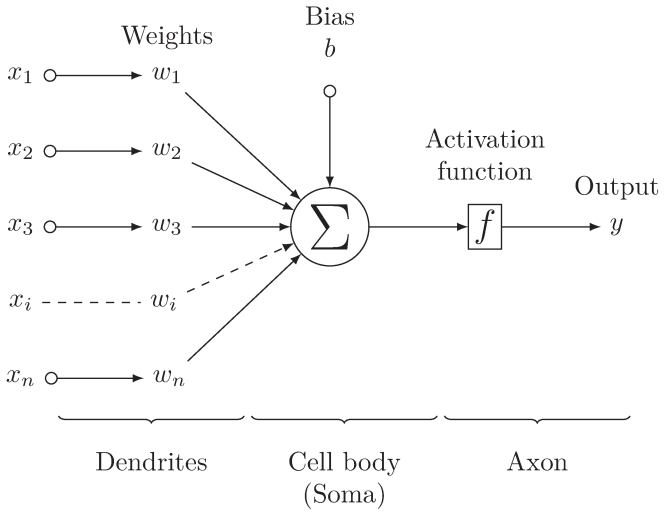


Figure 3.2: Artificial neuron.

Therefore, considering the general representation of an artificial neuron, Figure 3.2, with N_c input connection, its output state, y, is given as:

$$y = f\left(\sum_{j=1}^{N_c} \mathbf{W}_j \cdot \mathbf{X}_j - b\right), \tag{3.1}$$

assuming $W_0 = -b$ and $x_0 = 1$, is obtained the generic formulation below: N_c input connection, its output state, y, is given as:

$$y = f\left(\sum_{j=0}^{N_c} \mathbf{W}_j \cdot \mathbf{X}_j\right), \tag{3.2}$$

where $\mathbf{W}$ and $\mathbf{X}$ are the weight and input vectors, respectively, and f is the *activation function* (e.g., linear, sigmoid, etc.).

In a typical Neural Network (NN), without feedback connections (feedforward), the neurons are organized in layers, characterized by a fixed number of parallel neurons. Specifically, in each layer the neurons share the same inputs, but not the weights, and these are connected with the outputs of the previous layer neurons. In the first layer, the inputs correspond to the system inputs. In the same way, for the last layer the number of outputs, thus the neurons number, is set by the system outputs. Therefore, assuming a network with M

layers and N_l neurons in the l-th layer, with $l = 1, \ldots, M$, the description of the output state for the k-th neuron in the l-th layer, $y_k^{(l)}$, is obtained reformulating the equation (3.2) in:

$$y_k^{(l)} = f\left(\sum_{j=0}^{N_{l-1}} \mathbf{W}_{kj}^{(l)} \cdot \mathbf{X}_j^{(l)}\right), \tag{3.3}$$

where $\mathbf{W}$ and $\mathbf{X}$ respectively become the weights and the inputs matrices. In particular, $\mathbf{W}_{kj}^{(l)}$ represents the weight adopted by k-th neuron into the l-th layer and applied at the j-th input. $\mathbf{X}_j^{(l)}$ represents the j-th input value for the neurons in the l-th layer. Consequently, for the first and the last layer it is assumed that:

$$\mathbf{x} = \left[\mathbf{X}_0^{(1)}, \ldots, \mathbf{X}_{N_1}^{(1)}\right], \qquad \mathbf{y} = \left[y_0^{(M)}, \ldots, y_{N_M}^{(M)}\right], \tag{3.4}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the input and output data vectors, respectively.

Different activation functions can be chosen for the neurons: hyperbolic tangent, unipolar and bipolar sigmoid, and a set of radial basis functions.

3.1.1 Deep Belief Networks

The Deep Belief Networks have been proposed by Hinton and Salakhutdinov [55] stacking restricted Boltzmann machines (RBMs), which are artificial neural networks with only two layers, one hidden and one visible. The neurons of each layer are connected to all the neurons of the other layer, but no connections between neurons of the same layer is allowed (otherwise it becomes an "unrestricted" Boltzmann machine). Unlike ANNs, the DBNs have only one visible layer composed of input and output nodes, and all the remaining layers are hidden. They model the relation between the observation vector $\mathbf{x}$ and the variables $\mathbf{h}^l$ of the l-th hidden layer, with $l = 1, \ldots, M$, as the following joint distribution:

$$P(\mathbf{x}, \mathbf{h}^1, \mathbf{h}^2, \ldots, \mathbf{h}^M) = \left(\prod_{j=0}^{M-2} P(\mathbf{h}^j | \mathbf{h}^{j+1})\right) P(\mathbf{h}^{M-1} | \mathbf{h}^M), \tag{3.5}$$

where $\mathbf{x} = \mathbf{h}^0$, $P(\mathbf{h}^{M-1}|\mathbf{h}^M)$ is the top layer joint distribution in the top-level RBM and the computation of probability for the conditional distribution $P(\mathbf{h}^j|\mathbf{h}^{j+1})$ (for each j-th layer) is easy. The general structure of a DBN is depicted in Figure 3.3.

Due to the high complexity of the net, the training phase cannot be approached by adopting the same strategies applied for the ANNs, e.g., gradient

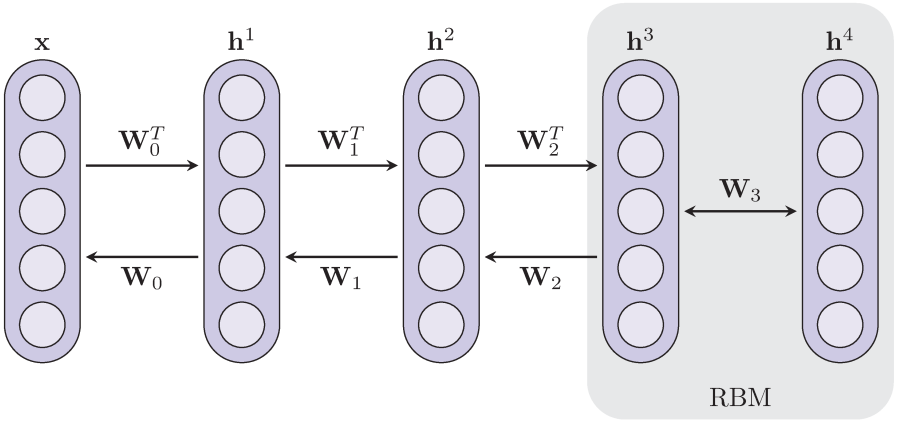


Figure 3.3: Deep Belief Network general structure.

descent or Levenberg - Marquardt algorithm. So, a fast training method has been developed: the "greedy algorithm" [56, 57]. The algorithm plans to train each hidden layer as an RBM that models the output data of the previous layer as its visible layer. For the first hidden layer, $\mathbf{h}^1$, the input vector $\mathbf{x} = \mathbf{h}^0$ is assumed as a visible layer, and used to tune the equivalent weight matrix $\mathbf{W}_0$. The representation of the input data for the training of the second hidden layer, $\mathbf{h}^2$, is obtained by mapping the input vector through the weight matrix $\mathbf{W}_0^T$. The procedure is repeated for each layer, each time propagating the mapped input vector.

3.1.2 Echo State Networks

Neural Networks wherein the neurons present feedback connections between them are known as Recurrent Neural Networks (RNNs). The complexity found in their use has led to the development of Echo State Networks (ESNs) [58, 59]: in short, an RNN (called *reservoir*) without a full adaptation of all the network weights. The typical update equations for an ESN are given as [60]:

$$\tilde{\mathbf{u}}(n) = \tanh\left(\mathbf{W}^{\text{in}} \cdot [1; \mathbf{x}(n)] + \mathbf{W} \cdot \mathbf{u}(n-1)\right), \tag{3.6}$$

$$\mathbf{u}(n) = (1-\alpha) \cdot \mathbf{u}(n-1) + \alpha \cdot \tilde{\mathbf{u}}(n), \tag{3.7}$$

where $\mathbf{x}(n) \in \mathbb{R}^{N_\mathbf{x}}$ is a given training input signal and $\mathbf{u}(n) \in \mathbb{R}^{N_\mathbf{u}}$ is a vector of reservoir neuron activations, at time step n, and its update $\tilde{\mathbf{u}}(n)$. $\mathbf{W}^{\text{in}} \in \mathbb{R}^{N_\mathbf{u} \times (1+N_\mathbf{x})}$ and $\mathbf{W} \in \mathbb{R}^{N_\mathbf{u} \times N_\mathbf{u}}$ are the input and recurrent weight matrices respectively, and $\alpha \in (0, 1]$ is the leaking rate. The linear readout layer is

defined as:

$$y(n) = \mathbf{W}^{\text{out}} \cdot [1; \mathbf{x}(n); \mathbf{u}(n)]\,, \tag{3.8}$$

where $y(n) \in \mathbb{R}^{N_\mathbf{y}}$ is the network output and $\mathbf{W}^{\text{out}} \in \mathbb{R}^{N_\mathbf{y} \times (1+N_\mathbf{x}+N_\mathbf{u})}$ the output weight matrix.

3.1.3 Autoencoders

Generally speaking, an autoencoder [61] is any neural network trained to get target values equal to input ones: $\mathbf{x} = \mathbf{y}$. Specifically, given a generic input $\mathbf{x} \in [0,1]^d$, the goal of the autoencoder is to first map the input (with an encoder) to an hidden representation $\mathbf{z} \in [0,1]^{d'}$, as:

$$\mathbf{z} = f(\mathbf{W}\mathbf{x} + \mathbf{b}), \tag{3.9}$$

where f is a non-linear function (e.g. tanh), $\mathbf{W}$ the weight matrix, and $\mathbf{b}$ teh bias vector. This level of representation, $\mathbf{y}$, also called *code*, is then mapped back into a reconstruction $\mathbf{y}$ of the same shape of $\mathbf{x}$, as:

$$\mathbf{y} = f(\mathbf{W}'\mathbf{z} + \mathbf{b}'), \tag{3.10}$$

where $\mathbf{W}'$ and $\mathbf{b}'$ denote the weight matrix and the bias vector of the reverse mapping, respectively. $\mathbf{y}$ is seen as a prediction of $\mathbf{x}$, give the code $\mathbf{z}$.

Different type of configurations can be obtained by varying the size and the dimension of the hidden representation, thus layers and neurons number. Among the state-of-the-art applications, the Denoising Autoencoder (DAE) [62] and the Compression Autoencoder (CAE) are ones of the most adopted configurations.

Denoising autoencoders are characterized by the use of a corrupted training sequence, in order to force the hidden layer to learn robust features and prevent it from learning the identity. Typically, a dAE is composed of multiple hidden layers that have a number of neurons greater than the input one.

On the other hand, if the number of hidden units is smaller than the number of input ones, a Compression Autoencoder is given. This network is forced to learn a compressed representation of the input: for example, if some of the input features are correlated, then the CAE is able to learn those correlations and reconstruct the input data from a compressed representation.

3.2 Support Vector Machines

Support Vector Machines [64] are binary classifiers that discriminate whether an input vector $\mathbf{x}$ belongs to class $+1$ or to class -1 based on the following

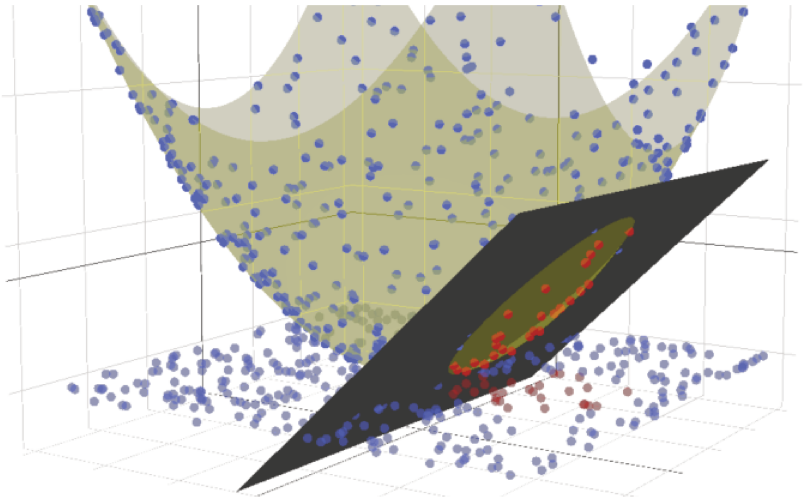


Figure 3.4: Example of SVM separation hyperplane with non-linear kernel [63].

discriminant function:

$$f(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i t_i K(\mathbf{x}_i, \mathbf{x}) + d, \tag{3.11}$$

where $t_i \in \{+1, -1\}$, $\alpha_i > 0$ and $\sum_{i=1}^{N} \alpha_i t_i = 0$. The terms $\mathbf{x}_i$ are the "support vectors" and d is a bias term that together with the α_i are determined during the training process of the SVM. The input vector $\mathbf{x}$ is classified as $+1$ if $f(\mathbf{x}) \geq 0$ and -1 if $f(\mathbf{x}) < 0$, given by the separation hyperplane, as depicted in Figure 3.4, defined by the "support vectors".

The kernel function $K(\cdot, \cdot)$ can assume different forms [65], common ones are the linear, the polynomial, and the Gaussian. Specifically, the linear one is defined as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = {\mathbf{x}_i}^T \mathbf{x}_j + c \ , \tag{3.12}$$

that gives the inner product between $\mathbf{x}_i$ and $\mathbf{x}_j$ plus an optional constant c.

On the other hand, the polynomial kernel is defined as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \left(\alpha {\mathbf{x}_i}^T \mathbf{x}_j + c\right)^d \ , \tag{3.13}$$

where are adjustable parameters the slope α, the constant term c, and the polynomial degree d.

Finally, the Gaussian kernel is an example of Radial Basis Function (RBF) kernel, which is defined as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i + \mathbf{x}_j\|^2}{2\sigma^2}\right), \tag{3.14}$$

where σ plays a major role in the performance of the kernel. If underestimated, the function will lack regularization and the decision boundary will be highly sensitive to noise in training data. On the other hand, if overestimated, the exponential will behave almost linearly and the higher-dimensional projection will start to lose its non-linear power.

3.2.1 Support Vector Regression

The Support Vector Regression (SVR) [66] approach, or SVM regression, is derived from the SVM [64] technique. Its goal is to find the function $f(x)$ that represents the target outputs with a maximum deviation ϵ, and, at the same time, is as flat as possible. Therefore, deviations lower than ϵ are not considered as errors, however, higher deviations are not allowed.

Differently from SVM, the solution of a linear model (in the feature space) for the SVR is obtained starting from a dual optimization problem, given as:

$$\begin{aligned} \text{maximize} \quad & \begin{cases} \frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)\left(\alpha_j + \alpha_j^*\right)K(\mathbf{x}_i, \mathbf{x}_j)\ , \\ -\epsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\ , \end{cases} \\ \text{subject to} \quad & \begin{cases} \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) & = 0\ , \\ \alpha_i, \alpha_i^* & \in [0, C]\ , \end{cases} \end{aligned} \tag{3.15}$$

to obtain:

$$f(\mathbf{x}) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) \cdot K(\mathbf{x}_i, \mathbf{x}) + d, \tag{3.16}$$

where α_i, α_i^* are determined during the training process, and $C > 0$ represents the trade-off between the flatness of f and the amount up to which deviations larger than ϵ are tolerated.

3.2.2 One-Class Support Vector Machine

In the One-Class Support Vector Machine (OC-SVM) theory [67], in order to separate the data set form the origin, the following quadratic problem has to be solved:

$$\begin{aligned} \min_{w\in\mathbf{F},\xi\in\mathbb{R}^l,\rho\in\mathbb{R}} \quad & \frac{1}{2}\|w\|^2 + \frac{1}{vl}\sum_i \xi_i - \rho\ , \\ \text{subject to} \quad & (w \cdot \Phi(\mathbf{x}_i)) \geq \rho - \xi_i, \xi_i \geq 0\ , \end{aligned} \tag{3.17}$$

to obtain the decision function:

$$f(\mathbf{x}) = \operatorname{sgn}\left(\sum_i \alpha_i \cdot k(\mathbf{x}_i, \mathbf{x}) - \rho\right) , \tag{3.18}$$

where $\nu \in (0, 1)$, $\mathbf{x}_i$ denotes the i-th Support Vector (SV) and $k(\cdot, \cdot)$ represents one of the kernel presented above.

The function f returns $+1$ in the "small" region where most of the data points fall and -1 everywhere else. The strategy is to map the data into a features space corresponding to the kernel and to separate the data from the origin with the maximum margin. For a new point $\mathbf{x}$, the value of $f(\mathbf{x})$ is determined by evaluating on which side of the hyperplane of the features space $\mathbf{x}$ falls, as depicted in Figure 3.4.

3.3 Auto-Regressive models

The autoregressive integrated moving average (ARIMA) model is the generalization of the autoregressive moving average (ARMA) model. The ARMA model is composed of two part (polynomials terms) describing the autoregressive (AR) part of order p and the moving average (MA) part of order q, respectively. For the ARMA model the output value at time step n can be approximated, $\tilde{y}(n)$, by the following linear combination, which depends on the previous p output values:

$$\tilde{y}(n) = \sum_{i=1}^{p} \phi(i) \cdot y(n-i) + a(n) + \sum_{j=1}^{q} \theta(j) \cdot a(n-j), \tag{3.19}$$

where $\phi(i)$ and $\theta(j)$ are the i-th and j-th coefficients of AR and MA models, respectively, and a is a input series of random *shocks*, according to [68]. Reformulating the ARMA expression introducing the backward shift operator, $B^i \cdot y(n) = y(n-i)$, the (3.19) becomes:

$$\left(1 - \sum_{i=1}^{p} \phi(i) \cdot B^i\right) \cdot y(n) = \left(1 - \sum_{j=1}^{q} \theta(j) \cdot B^j\right) \cdot a(n), \tag{3.20}$$

Now, considering the previous model and adding the d-th difference of the process, defined as $y(n) - y(n-d) = (1-B)^d \cdot y(n)$, that represents the non-stationary behaviour, the univariate ARIMA model can be expressed as:

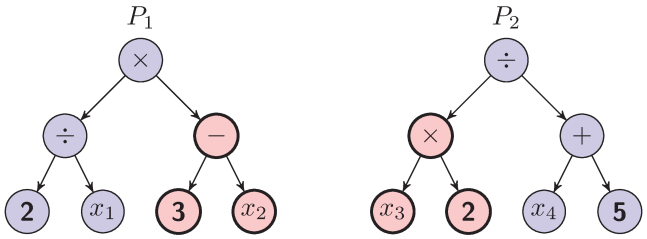
$$\left(1 - \sum_{i=1}^{p} \phi(i) \cdot B^i\right) \cdot (1-B)^d \cdot y(n) = \left(1 - \sum_{j=1}^{q} \theta(j) \cdot B^j\right) \cdot a(n), \tag{3.21}$$

where, besides the AR and MA parameters, the d defines the order of the integrated part.

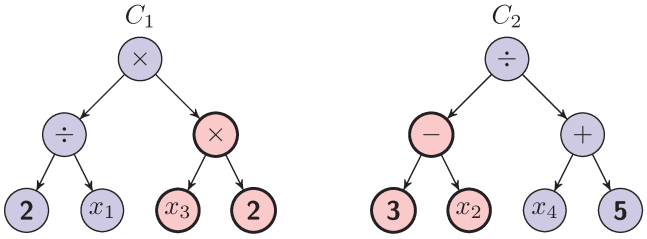
3.4 Genetic Programming

The evolution process is based on three main criteria [69]: heredity, variability and fecundity. This way, an individual of a population will be able to adapt to an environment. The genetic programming has been inspired by the concept of evolution and through the following steps the algorithm mimics the evolution criterion:

1. **Initialization:** the parameters are set and the first generation is randomly created.
2. **Selection:** the best individual is selected using the sum of the absolute errors (SAE) and it is properly used for the creation of the new generation.
3. **Control:** the second step is repeated for the new generation until either a stop condition or the maximum generation is reached.



(a) Parents P_1 and P_2 before cross-over.



(b) Children C_1 and C_2 created by cross-over.

Figure 3.5: Example of cross-over between parents to create children. In red the sub-trees selected for the operation.

At the end, the algorithm returns the best individual, chosen according to a pre-defined standard, found since the beginning of the simulation. In the second

phase, the operations designed to create new individuals, children, from the best previous one, parents, are cross-over and mutation, which basic examples are depicted in Figure 3.5 and Figure 3.6, respectively. The cross-over operation is performed swapping parents sub-trees. As a result of the transformation in Figure 3.5, the children equation, y_{c1} and y_{c2}, obtained from the parents one, y_{p1} and y_{p2}, are:

$$\begin{cases} y_{p1} = (2 \div x_1) \times (3 - x_2), \\ y_{p2} = (x_3 \times 2) \div (x_4 + 5), \end{cases} \implies \begin{cases} y_{c1} = (2 \div x_1) \times (x_3 \times 2), \\ y_{c2} = (3 - x_2) \times (x_4 + 5), \end{cases} \tag{3.22}$$

where x_n denotes the n-th time-lag input, with $n = 1, \ldots, L$. L is the number of time-lagged observations assumed as input data, introduced in Section 4.1, and is set to 5 in the examples.

The mutation operation concerns the substitution of terminal, function, or sub-trees with new ones. In Figure 3.6, a function mutation is presented, the parent equation y_p is muted in the children one, y_c.

$$y_p = (2 \div x_1) \times (3 - x_5), \qquad \implies \qquad y_c = (2 \times x_1) \times (3 - x_5). \tag{3.23}$$

The output model function depends on the mathematical operators chosen and the maximum depth of the node.

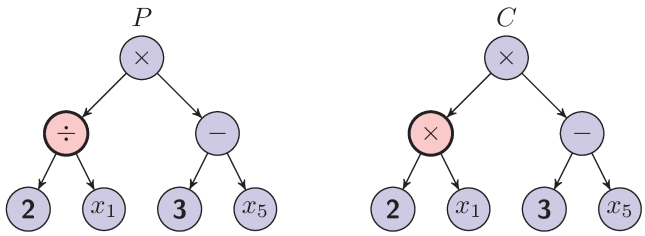


Figure 3.6: Example of parent P (left) mutation for the child C (right) creation. In red the selected function.

3.5 Extended Kalman Filter

The extended Kalman filter (EKF) model produces a suboptimal solution using a first order linear approximation of the filter proposed by Kalman [70]. The algorithm consists of successive uses of *predict* and *update* equations, for each input/output relation. The EKF block diagram is shown in Figure 3.7.

The algorithm can be summarized in two main steps: initialization and iteration phases. At the beginning, only the first input/output relation is available, and it is used to generate a first estimate. Following the minimization of MSE,

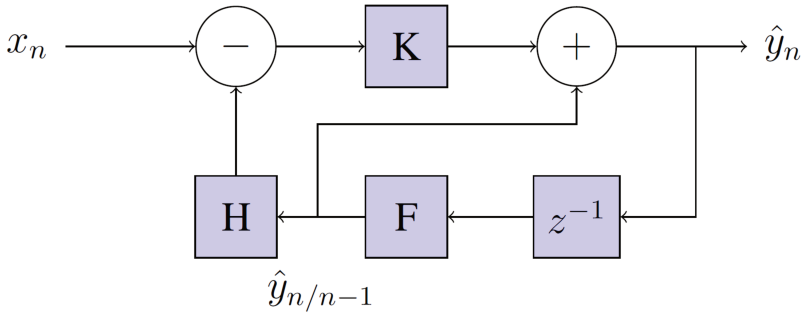


Figure 3.7: Extended Kalman filter block diagram.

the new estimate is corrected using the previous value and stored for the next input value.

The *predict* equations at time step n, for the *a priori* estimation of the state (output) $\hat{\mathbf{y}}_{n/n-1}$ and the error covariance matrix $\mathbf{P}_{n/n-1}$, from the information of the $(n-1)$-th step, are defined as:

$$\hat{\mathbf{y}}_{n/n-1} = \mathbf{F}_{n/n-1} \cdot \hat{\mathbf{y}}_{n-1/n-1}, \tag{3.24}$$

$$\mathbf{P}_{n/n-1} = \mathbf{A}_{n-1} \cdot \mathbf{P}_{n-1/n-1} \cdot \mathbf{A}_{n-1}^T + \mathbf{Q}_{n-1}, \tag{3.25}$$

where, for the extended version of the filter, $\mathbf{F}$ denotes the transition matrix, and $\mathbf{A}$ is the Jacobian matrix of partial derivatives of $\mathbf{F}$.

Assuming that $\mathbf{H}$ is the observation model and $\mathbf{x}$ is the measurement vector, the *update* equations are given as:

$$\mathbf{K}_n = \mathbf{P}_{n/n-1} \cdot \mathbf{H}_n^T \cdot (\mathbf{H}_n \cdot \mathbf{P}_{n/n-1} \cdot \mathbf{H}_n^T + \mathbf{R}_n)^{-1}, \tag{3.26}$$

$$\hat{\mathbf{y}}_{n/n} = \hat{\mathbf{y}}_{n/n-1} + \mathbf{K}_n \cdot (\mathbf{x}_n - \mathbf{H}_n \cdot \hat{\mathbf{y}}_{n/n-1}), \tag{3.27}$$

$$\mathbf{P}_{n/n} = \mathbf{P}_{n/n-1} - \mathbf{K}_n \cdot \mathbf{H}_n \cdot \mathbf{P}_{n/n-1}. \tag{3.28}$$

where $\mathbf{Q}$ and $\mathbf{R}$ are the process and measurement covariances matrices, respectively, and $\mathbf{K}$ is the Kalman gain.

3.5.1 EKF-GP

Nasseri *et al.* [17] proposed for the first time the combination of EKF and GP. The method involves the use of the best GP model obtained within the EKF parameters. In particular:

- the best GP model is used as transition matrix $\mathbf{F}$;
- the *NMSE* of the GP model is adopted as noise covariances $\mathbf{Q}_n$ and $\mathbf{R}_n$;

- the observation model $\mathbf{H}$ and the state variable $\mathbf{Y}$ are vectors of dimension $m \times 1$, with m equal to the lags number.

After the computation of the parameters, the EKF predicted states are recalculated using the same GP model. Therefore, the n-th EKF predicted states, whose number depends on the lags assumed, are used as input variables for the selected GP model, that returns the equivalent output value of the model equation.

3.6 Extreme Learning Machine

Extreme Learning Machine (ELM) has been presented by Huang *et al.* [71] as a fast learning algorithm for single hidden layer feedforward neural networks (SLFNs). Differently from standard ANNs approaches, the input weights are randomly generated and the output ones are tuned by a least-square method. In later work [72], an unified solution for regression, binary, and multi-class classification has been presented. Specifically, given a set of pairs $(\mathbf{x}_i, t_i)$, $i = 1, \cdots, N$, where $\mathbf{x}_i \in \mathbb{R}^L$ is the training data, and $t_i \in \{-1, 1\}$ denotes the corresponding label, the output function of the ELM for regression is computed as:

$$f(\mathbf{x}) = \mathbf{h}(\mathbf{x})\mathbf{H}^{\mathbf{T}} \left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^{\mathbf{T}} \right)^{-1} \mathbf{T}, \tag{3.29}$$

where $\mathbf{h}(\mathbf{x})$, namely the feature mapping, denotes the hidden layer output for the corresponding input $\mathbf{x}$, and C is the regularization coefficient, whereas the matrices $\mathbf{H}$ and $\mathbf{T}$ denote the hidden layer output matrix and the input labels matrix, respectively.

Moreover, a kernel-based approach has been also proposed [71] if a feature mapping $\mathbf{h}(\mathbf{x})$ is unknown. Differently from the standard ELM, the number of hidden neurons must not be known in advance, and the output relation (3.29) becomes:

$$f(\mathbf{x}) = \begin{bmatrix} K(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ K(\mathbf{x}, \mathbf{x}_N) \end{bmatrix}^T \left(\frac{\mathbf{I}}{C} + \mathbf{\Omega} \right)^{-1} \mathbf{T}, \tag{3.30}$$

where $K(\cdot, \cdot)$ denotes the chosen kernel function, and $\mathbf{\Omega}$ defines the kernel matrix, so that $\Omega_{i,j} = h(\mathbf{x}_i) \cdot h(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$. As for the SVM in Section 3.2, many types of kernels can be adopted.

3.7 Gaussian Mixture Model

In probability theory [65] the expression:

$$\mathcal{N}(x|\mu,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{(1/2)}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}, \tag{3.31}$$

represents a Gaussian distribution for a single real-valued variable x, where μ and σ^2 represent the mean and the variance, respectively. In particular, the distribution satisfies:

$$\mathcal{N}(x|\mu,\sigma^2) > 0, \tag{3.32}$$

$$\int_{-\infty}^{+\infty} \mathcal{N}(x|\mu,\sigma^2)dx = 1. \tag{3.33}$$

Moving to the D-dimensional case, the Gaussian distribution can be redefined as:

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}, \tag{3.34}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ denote the D-dimensional mean vector and the $D \times D$-dimensional covariance matrix, and $|\Sigma|$ is the determinant of Σ.

The Gaussian Mixture Model (GMM) is obtained as linear superposition of instances, Ng, of multidimensional Gaussian distributions, (3.34):

$$p(\mathbf{x}|\boldsymbol{\lambda}) = \sum_{g=1}^{Ng} w_g \, \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g), \tag{3.35}$$

where $w_g \in [0,1]$, $\sum_{g=1}^{Ng} w_g = 1$, $\lambda = \left\{ \boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g, w_g \right\}_{g=1}^{Ng}$, and Ng denotes the number of Gaussians adopted for the GMM. The training process of a GMM follows two steps: initialization and update/finalization. In the first step, the parameters of the Ng Gaussian components are computed by using the k-means algorithm [73]; in the second step, the Expectation-Maximization (EM) algorithm [74] re-estimates the parameters, guaranteeing a monotonic increase of the likelihood and reaching its maximum value. In this study, all GMMs have a diagonal covariance matrix.

3.8 Hidden Markov Model

The Hidden Markov Model (HMM) is characterized by a transition probability matrix $\mathbf{A}$, a sequence of emission probabilities $\mathbf{B}$, also called observation likelihoods, and a set of states $\mathbf{Q} = \{q_1, q_2, \ldots, q_{Ns}\}$. In the matrix $\mathbf{A}$, each element $a_{n,s}$ represents the probability of moving from a state n to a state s, assuming that $\sum_{n=1}^{Ns} a_{n,s} = 1 \ , \forall s$; whereas, letting $\mathbf{O} = \{o_1, o_2, \ldots o_T\}$ be a

given observation sequence, the generation probability of the observation o_t by the state n is expressed as $b_n(o_t)$.

A typical HMM structure used in speech recognition to model the temporal processes [75] is the left-to-right structure (also called Bakis) depicted in Figure 3.8.

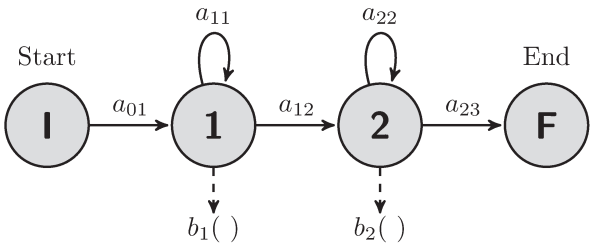


Figure 3.8: Example of HMM left-to-right structure.

The Markov assumptions affirm that the probability of a particular state is dependent only on the previous state. Furthermore, the probability of an output observation o_t is dependent only on the state q_n that produced the observation and not on any other states or on any other observations. In this case, it is possible to compute the output likelihood, for a given model, using the iterative procedure [76]:

$$\begin{aligned} &\alpha_1(n) = a_{0,n} b_n(o_1) \ , && 1 \leq n \leq Ns \\ &\alpha_t(n) = \sum_{s=1}^{Ns} \alpha_{t-1}(s) a_{s,n} b_n(o_t) \ , && 1 \leq n \leq Ns, 1 \leq t \leq T \\ &P(O|\lambda) = \alpha_T(q_F) = \sum_{s=1}^{Ns} \alpha_T(s) a_{s,F} \ , \end{aligned} \tag{3.36}$$

where $\lambda = (A, B)$ (automaton), T denotes the number of observations in $\mathbf{O}$ and Ns denotes the number of states. Furthermore, $a_{0,n}$ and $a_{s,F}$ denote the transition probabilities out of the start state and into the last (or end) state, which are not associated with the observations.

In order to create a continuous density model, each observation probability distribution, $b_n(o_t)$, is represented with a mixture Gaussian density. The final model is created by adopting a special case of the EM algorithm, the forward-backward or Baum–Welch algorithm [77], which allows one to train both the transition probabilities $\mathbf{A}$ and the emission probabilities $\mathbf{B}$ of the HMM.

Chapter 4

Short/Medium-Term Load Forecasting

4.1 Experimental Setup

The forecasting experiments have been conducted adopting the publicly available datasets presented in Section 2.3: AMPds [51] and DIFD [49] with both water and natural gas resources, Tehran dataset [17] with water consumption, and EIA datasets [50] with gas consumption. In particular, two different grid scenarios are represented by the datasets: the AMPds and the DFID refers to the micro-grid case (to be more precise, domestic and office building level, respectively), while both EIA and Tehran point out the macro-grid case. In order to have a better insight of the prediction performance and properly evaluate possible correlation between different input data, the author applied the algorithms to each possible combination of input data and forecast resolution.

Regarding the AMPds, the available input information have been merged creating eight different combinations of data, for both natural gas and water prediction, and forecasts with a resolution time of $1\,h$, $6\,h$, $12\,h$ and $24\,h$ have been evaluated. Predictions with short time step, even below the hour, are well suited for detailed control [14]. However, daily forecasts can be adopted for "flow control for treatment plant", and hourly predictions are useful for "detailed distribution pump scheduling". In order to evaluate the use of time steps below the hours, preliminary evaluations have been performed with the AMPds. However, setting time steps shorter than one hour produced poor performance, unsuitable for proper consideration of the approaches under study.

For the U.S. natural gas consumption the monthly data of 3 states have been chosen in order to represent a cold, a mild and a hot climate. The U.S. average temperature, from 1989 to 2014, is 11.3 °C, with a standard deviation of 8.7, and maximum and minimum values of 24.9 °C and -2.5 °C, respectively. Therefore, with an average of 4.6 °C, Maine has been selected as cold state, whereas Illinois and Louisiana have been chosen as mild and hot climate, having average temperatures of 10.9 °C and 19.2 °C, respectively. Specifically, Maine reached

a maximum and a minimum temperature of 20.9 ℃ and -15.5 ℃, respectively, with a standard deviation of 9.7. Illinois presented maximum and minimum temperatures of 27.7 ℃ and -8 ℃, respectively, with a standard deviation of 9.6. Finally, Louisiana reached a maximum and a minimum temperature of 30.4 ℃ and 5.6 ℃, respectively, with a standard deviation of 6.7.

For each state, being available the monthly average temperature too, two diverse input datasets have been evaluated: one composed of both gas consumption and temperature data, and the other composed only of gas consumption data.

Instead, not being available other heterogeneous information, for the DFID and the Tehran datasets the forecasting experiments have been conducted by assuming only the raw consumption data.

The forecasting experiments, presented below, have been conducted for the following approaches: ANN, DBN, ESN, SVR, ELM with linear kernel (ELM-LIN), ELM with RBF kernel (ELM-RBF), GP, and EKF-GP. Specifically, for all the proposed approaches the tests have been performed by assuming 2, 3, and 5 time-lagged observations as input data. The total number of input variables is given as the product between the number of lags and the number of input resources. For example, four input variables are assumed if the input data are composed of water and temperature information for two time-lags. So, for the ANNs, DBNs and ESNs, the number of nodes for the input layers has been set equal to the number of input variables. The training sets have been composed randomly selecting the 70% of the total data; the remaining data, 30%, have been used for the testing phase.

Additional prediction evaluations have been conducted using the univariate ARIMA representation (without heterogeneous data), which is a generalization of the ARMA model, in order to have a comparison with a linear model, as in [17].

In the following Tables 4.3, 4.4, and 4.5, the "Parameters" column reports for each approach the value of the parameters that have been used to get the corresponding best result. For each parameters sequence the first value (to the left) is the number of lags. The second parameter for the ANN denotes the number of neurons. In the DBN, the parameters are the number of nodes, the maximum iterations number and the "mini-batches" size, respectively. The reservoir size, the initial transient, the leaking rate and the regularization coefficient are reported for ESN. For SVR and ELM with RBF kernel, number of lags, C and γ values are indicated, respectively, while for ELM with linear kernel only number of lags and C values are reported. The maximum depth is reported for both GP and EKF-GP. Moreover, for the stochastic algorithms, which are GP, ANN, DBN, and ESN, each set of tested parameters has been executed 10 times and the overall best result has been selected as result for

the set. In order to guarantee the same initial conditions for each tested set, a fixed seed has been used to set the random generator at the beginning of the iterations. All the approaches are clearly reproducible, and the necessary information are given in this Chapter and in Chapter 3.

Finally, each input data presented in the datasets has been normalized by means of the Min-Max normalization, in order to have an input $x_n^N \in [0, 1]$. The normalization is based on the following rule:

$$x_n^N = \frac{x_n - x_{min}}{x_{max} - x_{min}}, \tag{4.1}$$

where x_n is the n-th original observed value, x_{max} and x_{min} are the maximum and minimum values for the considered dataset, respectively.

The experiments have been performed on two computers with 32GB of RAM and an Intel(R) Core(TM) i7-4930K CPU @ 3.40GHz, and an Intel(R) Core(TM) i7-5820K CPU @ 3.30GHz, respectively.

4.1.1 Parameters

ANN

In the tests, one of the radial basis function has been chosen as activation function, specifically the one defined as $f(a) = \exp(- \parallel a - c \parallel^2 / 2\sigma^2)$. Here, c is the mean, and σ is the standard deviation of the Gaussian function. The standard structure has been selected for the experiment, using 1 input layer, 1 hidden layer and 1 output layer. The number of hidden layer nodes has been varied, from 5 to 15 in order to identify the configuration that maximizes the prediction performance. The simulations have been performed by using the MATLAB® Neural Network Toolbox.

DBN

The experiments have been conducted using the DBN MATLAB® code [78] and according to the directives of Hinton [79] about the RBMs training. So, as suggested by the guide, the training set has been split in "mini-batches" of 10 or 100 cases. The other parameters are the maximum number of iterations and the number of nodes, within the following ranges:

$$\begin{aligned} \textit{Nodes} &= \{2, 3, \ldots, 10\}, \\ \textit{Max Iter} &= \{100, 300, 500\}. \end{aligned}$$

The number of nodes indicates the actual nodes used for the hidden layer, while the number of input and output nodes depends on the number of input and output variables, respectively. The introduction of more hidden layers has been addressed during a preliminary test resulting in over-fitting problems and poor results.

ESN

The ESN, adapting the ESN MATLAB® code [80], have been tested by changing the *reservoir* dimension, the value of the initial transient, the leaking rate (α) and the regularization coefficient (β). Each parameter has been changed according to the Lukoševičius guide [60], and specifically the adopted values are:

$$\begin{aligned} \textit{reservoir size} &= \{2, 5, 10, 15, 20, 25, 30, 35, 40\}, \\ \textit{initial transient} &= \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}, \\ \alpha &= \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}, \frac{1}{9}, \frac{1}{10}\right\}, \\ \beta &= \left\{1, 10^{-1}, 10^{-2}, 10^{-3}, \ldots, 10^{-15}, 0\right\}. \end{aligned}$$

SVR

All the SVR simulations have been performed using LIBSVM [81] (a library for Support Vector Machines) and the ν-SVR mode. The radial basis function has been chosen as kernel, that is given as $K(\mathbf{u}, \mathbf{v}) = \exp(-\gamma\|\mathbf{u} - \mathbf{v}\|^2)$. The γ and C parameters, and the chosen kernel are crucial to obtain the best performance. The datasets have been tested using the C and γ parameters, performing the grid search approach, within the following ranges:

$$\begin{aligned} C &= \left\{2^{-5}, 2^{-3}, \ldots, 2^{15}\right\}, \\ \gamma &= \left\{2^{-15}, 2^{-12}, \ldots, 2^{3}\right\}. \end{aligned}$$

ARIMA

In the experimental phase, conducted using the corresponding MATLAB® class, the input data x have been assigned to the previous values of $y(n)$ in the autoregressive model. For the d and q parameters, the advices in Box *et al.* [68] have been accounted for: "d is usually 0, 1, or at most 2" and "it is frequently true that adequate representation of [...] time series can be obtained,[...] in which p and q are not greater than 2 and often less than 2". Therefore, the tests have been executed using a value of p equal to the lags number, and for

the remaining parameters the following ranges have been considered:

$$d = \{0, 1, 2\}, \qquad q = \{0, \ldots, p\}.$$

GP & EKF-GP

GP simulations have been executed using the GPLAB Toolbox [82], testing a maximum tree depth of 20, 15 and 10. The selected operators are: plus (+), minus (−), product (×), division (÷), power (x^n), sine (sin) and cosine (cos). Finally, the best GP model has been sought by varying the maximum depth of the node.

After the computation of the GP parameters, the EKF predicted states are recalculated using the same GP model. Therefore, the n-th EKF predicted states, whose number depends on the lags assumed, are used as input variables for the selected GP model, that returns the equivalent output value of the model equation. The simulations have been executed using the EKF/UKF Toolbox [83].

ELM

The simulations have been performed for linear and RBF kernels. In both the approaches the regularization coefficient C has been varied within the range:

$$C = \left\{2^{-5}, 2^{-3}, \ldots, 2^{15}\right\}.$$

In case of RBF kernel, the γ has been also varied within the following range:

$$\sigma = \left\{2^{15}, 2^{12}, \ldots, 2^{-3}\right\}.$$

4.2 Evaluation Methods

A crucial point for the comprehension of the performance achieved by the prediction methods is the usage of suitable evaluation criteria. As discussed in Section 2.1.2, the state-of-the-art contributions fail to produce a clear and understandable comparison between different prediction techniques.

For that reason, a set of evaluation criteria to adopt for each experiment has been selected, in order to facilitate the comparison and the analysis of the methods. With reference to Bennett *et al.* [84] report, the criteria taken into

account are the normalized mean square error (NMSE), the determination coefficient (R^2), the mean square error (MSE), the mean absolute percentage error (MAPE) and the relative root mean square error (RRMSE). The corresponding formulas are:

$$\begin{aligned}
&\text{NMSE} = \frac{\frac{1}{N}\sum^{N}(y_n - \tilde{y}_n)^2}{\sigma_y^2} \ , && \text{R}^2 = 1 - \frac{\frac{1}{N}\sum^{N}(y_n - \tilde{y}_n)^2}{\frac{1}{N}\sum^{N}(y_n - \bar{y})^2} \ , \\
&\text{MSE} = \frac{1}{N}\sum^{N}(y_n - \tilde{y}_n)^2 \ , && \text{MAPE} = \frac{\frac{100}{N}\sum^{N} \mid y_n - \tilde{y}_n \mid}{\bar{y}} \ , \\
&\text{RRMSE} = 100 \cdot \frac{\sqrt{\frac{1}{N}\sum^{N}(y_n - \tilde{y}_n)^2}}{\bar{y}} \ .
\end{aligned} \tag{4.2}$$

where y_n is the n-th observed value, $\tilde{y}_n$ the corresponding n-th forecast value, $\bar{y}$ and σ_y^2 the average and the variance of the N observed values, respectively.

In accordance to the classification criterion proposed by Bennett *et al.* [84], the selected performance measures can be discerned in two distinct categories. The MSE, the MAPE, and the RRMSE belong to the methods that allow to couple observed/real and modelled values, whereas the R^2 and the NMSE preserve the data pattern thus allowing correlation and model efficiency performance measures.

Specifically, the MSE is classified as key residual methods, which calculate the difference between observed and modelled data point. The simple bias (that is the mean of the residuals) allows to understand if the model under- or overestimate the measured data, but is affected by the cancellation issue: positive and negative errors cancel each other out. In order to avoid this cancellation problem, the MSE squares the residuals, making all contributions positive. However, squaring the data may cause bias towards large events.

The MAPE and the RRMSE belong to the approaches used to obtain relative error and error transformation, in which the metric is weighted to focus on the aspects of interest. The adopted methods, frequently adopted in hydrologic modelling, compare the mean absolute error (MAE) and the root mean square error (RMSE), respectively, to the total measurement record. Similar to bias measurements, a low values does not mean low error, just balanced errors. The RMSE simply adopt the square root of the MSE to express the error metric in the same units as the original data, which is useful for interpretation.

Among the evaluation methods that preserve the data pattern, the NMSE and the Coefficient of Determination, commonly known as the Nash-Sutcliffe efficiency coefficient, are also widely used in hydrologic modelling. The for-

Table 4.1: Best results achieved for each technique applied to Natural Gas of AMPds database released in 2013 and 2014, with a length of 1 year and 2 years, respectively. The column marked with "Comb." reports the resources combination that achieves the best result.

AMPds	1 *h*			6 *h*			12 *h*			24 *h*		
	NMSE	RRMSE	Comb.	NMSE	RRMSE	Comb.	NMSE	RRMSE	Comb.	NMSE	RRMSE	Comb.
	Natural Gas Prediction - 1 year											
ARIMA	0.833	143.0	G	0.528	75.5	G	0.388	56.7	G	0.327	46.6	G
ANN	**0.686**	**127.3**	**WGET**	0.276	54.115	GET	0.200	39.3	WGET	**0.176**	**37.1**	**G**
DBN	0.785	138.2	WGT	0.340	60.6	WGT	0.201	40.8	WGET	0.199	38.0	GT
ESN	0.740	137.7	GT	0.333	60.0	GET	0.213	42.0	WGET	0.231	39.2	GT
SVR	0.733	134.1	GET	0.269	54.0	WGET	**0.191**	**39.7**	**GT**	0.250	42.6	GT
ELM-LIN	0.804	143.5	WGET	0.367	62.9	WGET	0.237	44.3	WGET	0.283	45.3	GT
ELM-RBF	0.700	131.1	WGET	**0.266**	**53.8**	**WGET**	0.194	40.1	WGET	0.241	41.8	GT
GP	0.781	141.4	GET	0.338	60.6	GET	0.255	43.9	WGT	0.242	41.9	GT
EKF-GP	0.718	132.7	WGE	0.456	70.1	GE	0.283	48.4	GT	0.299	46.6	WGT
	Natural Gas Prediction - 2 years											
ARIMA	0.872	135.2	G	0.953	96.7	G	0.374	51.8	G	0.232	34.4	G
ANN	**0.619**	**116.5**	**WGT**	0.182	42.7	GET	0.170	25.5	GET	**0.153**	**28.9**	**GT**
DBN	0.723	125.2	WGET	0.238	48.3	WGET	0.185	38.3	GET	0.188	31.9	GET
ESN	0.695	122.8	WGET	0.247	49.2	GET	0.176	37.3	GET	0.188	31.0	GT
SVR	0.644	118.1	WGET	0.182	42.3	WGET	**0.145**	**33.9**	**WGT**	0.178	30.2	GT
ELM-LIN	0.828	127.2	WGET	0.405	50.0	WET	0.579	39.0	WGT	0.428	33.2	WT
ELM-RBF	0.620	115.9	WGET	**0.180**	**42.0**	**WGET**	0.147	34.1	WGT	0.176	30.0	GT
GP	0.740	126.6	GT	0.250	49.5	WGT	0.177	37.5	GET	0.190	31.1	WGT
W = Water			G = Natural Gas			E = Electric Power			T = Temperature			

mer criterion represents the traditional MSE method weighted by the standard deviation of the observed value. A lower value indicates the better performance. The latter one indicates how well the model explains the variance in the observations, compared with using their mean as the prediction. A value of unity indicates a perfect model, while a value below zero indicates performance worse than simply using the mean. A negative value indicates even worse performance. Generally speaking, both the models compare the performance of the model to a model that only uses the mean of the observed data.

4.3 Experiments

4.3.1 AMP Dataset

For the sake of clarity, in Table 4.1 and 4.2 only the overall best results achieved for each approach, with the corresponding resources combination and the NMSE and R^2 values, are reported. Tables with the complete best results achieved for each approach, resources combination, and time resolution, are provided in Appendix 1 and 2.

The overall natural gas best result, reported in Table 4.1, is reached by the SVR approach with 2 years data at 12 *h* resolution. Specifically, the heterogeneous input composed of gas, water, and temperature (WGT) achieves this result, with a NMSE of 0.145, a R^2 of 0.854, a MSE of 0.0067, a MAPE of 23.690%, and a RRMSE of 33.933%. Moreover, the best result is achieved for 2 lags with C and γ values of 2^{-1} and 2^3, respectively. Whereas, for 1 year

Table 4.2: Best results achieved for each technique applied to Water of AMPds database released in 2013 and 2014, with a length of 1 year and 2 years, respectively. The column marked with "Comb." reports the resources combination that achieves the best result.

AMPds	1 *h*			6 *h*			12 *h*			24 *h*		
	NMSE	RRMSE	Comb.	NMSE	RRMSE	Comb.	NMSE	RRMSE	Comb.	NMSE	RRMSE	Comb.
	Water Prediction - 1 year											
ARIMA	0.939	158.7	W	0.990	89.6	W	0.766	48.5	W	0.403	28.4	W
ANN	**0.745**	**142.7**	**WET**	0.343	58.8	WGT	**0.408**	**34.2**	**WE**	**0.345**	**25.2**	**W**
DBN	0.876	145.9	WGET	0.401	61.2	WGET	0.573	45.7	WT	0.415	30.2	WGET
ESN	0.815	140.8	WET	0.367	58.6	W	0.522	43.6	W	0.334	25.8	WGT
SVR	0.817	140.3	WGET	**0.299**	**52.9**	**WET**	0.484	38.5	WET	0.352	26.5	WGT
ELM-LIN	0.829	142.0	WGET	0.405	61.5	WET	0.579	45.9	WGT	0.428	29.2	WT
ELM-RBF	0.769	136.7	WGET	0.310	53.8	WGET	0.490	38.8	WT	0.351	26.5	WGET
GP	0.900	155.4	WGE	0.357	57.7	W	0.505	39.4	WGT	0.359	26.8	WGE
EKF-GP	0.883	146.5	WE	0.383	59.8	WET	0.608	47.0	WG	0.454	30.1	WE
	Water Prediction - 2 years											
ARIMA	0.908	158.8	W	1.003	94.0	W	0.632	57.5	W	0.896	29.7	W
ANN	0.790	152.9	WET	0.319	54.1	W	**0.223**	**34.2**	**W**	**0.550**	**23.6**	**W**
DBN	0.867	149.8	WGET	0.407	57.8	WGET	0.317	40.8	WGET	0.607	26.3	WE
ESN	0.818	141.9	WGET	0.413	58.1	WGET	0.308	41.5	W	0.581	25.7	W
SVR	0.806	140.8	WGE	**0.316**	**50.9**	**WGET**	0.269	36.0	W	0.581	25.7	W
ELM-LIN	0.837	143.5	WGET	0.419	58.6	WGET	0.314	40.5	WGT	0.628	26.7	WE
ELM-RBF	**0.765**	**140.7**	**WGET**	0.319	51.1	WGET	0.279	38.2	WET	0.575	25.6	W
GP	0.890	151.8	WGET	0.361	54.3	WT	0.272	36.2	W	0.567	25.4	W
W = Water	G = Natural Gas			E = Electric Power			T = Temperature					

data, the best result is achieved at 24 *h* resolution with ANN, without using heterogeneous data. Specifically, a NMSE of 0.176, a R^2 of 0.823, a MSE of 0.0086, a MAPE of 24.98%, and a RRMSE of 37.12% are obtained with 2 lags and a hidden layer of 8 neurons.

In accordance to ELM theory [71, 72], SVR and ELM with RBF kernel show very close performance. Moreover, the major improvements of the 2 year predictions with respect to 1 year ones are achieved for 6 *h* and 12 *h* resolution, with a relative improvement of the NMSE of 48% and 32%, respectively. For both sets, a best improvement in the results is performed passing from 1 *h* to 6 *h* of resolution.

The results confirm that the heterogeneous data are essential information for the prediction, even if it seems reducing the data resolution decreases the number of heterogeneous components. Specifically, in the 1 year set, the best results at 1 *h* and 6 *h* resolution are obtained merging all the heterogeneous components, whereas for 12 *h* and 24 *h* resolution the best results are achieved with the pair gas-temperature and with the gas consumption only, respectively.

Finally, for the 2 years set, the reduction of the heterogeneous components does not appear as marked as in the 1 year set. Specifically, the best results at 1 *h* and 12 *h* resolution are achieved with three heterogeneous components (WGT). At 6 *h* resolution the best result is achieved with four components (WGET), and only two components (GT) are used as input for the best result at 24 *h* resolution.

Differently from the natural gas, water best results for the 2 years set at 1 *h*, 6 *h* and 24 *h* of resolution experienced a severe deterioration with respect

to the 1 year ones, reported in Table 4.2. Despite this, the ANN obtained the overall best result with the 2 years dataset at 12 h resolution, denoting a NMSE improvement of 83% with respect to the 1 year result. For this point, the prediction shows a NMSE of 0.223, a R^2 of 0.777, a MSE of 0.0037, a MAPE of 25.59%, and a RRMSE of 34.16%. Specifically, the input is composed of only the water information with 2 lags, and a network with 13 hidden neurons.

For the 1 year data, the overall best performance is achieved by the SVR at 6 h resolution with 5 lags and an heterogeneous input composed of water, energy, and temperature data (WET). This result is performed for C and γ values of 2^{-1} and 2, respectively, with a NMSE of 0.299, a R^2 of 0.700, a MSE of 0.0040, a MAPE of 37.34%, and a RRMSE of 52.85%.

For the 1 year set, a general reduction of the heterogeneous components is shown only for the results at 12 h resolution. For the 2 year set, the heterogeneous data do not provide a clear improvement at the prediction, and a clear reduction in the number of heterogeneous components is shown when the resolution decreases. Specifically, at 12 h and 24 h resolution, the best performance are achieved with only water consumption information as input. Noteworthy, for 1 year consumption, with the exception of ESN, all the approaches achieve their best results at 6 h resolution, then the performance show a slight degradation for 12 and 24 h resolution.

Finally, for both water and gas prediction, ESN, DBN, and GP approach have confirmed to be unsuitable for short-term prediction in domestic environment. On the contrary, ANN, SVR, and ELM have shown to be effective for short-term forecasting on this dataset.

Considering the natural gas prediction, the resources combinations that have achieved the best performance in each resolution and for each years, or close to the best as for DBN in 24 h resolution, always include the temperature data, denoted by a "T" marker in the column "Comb." of Table 4.1. For this reason, it is reasonable to assume the existence of a correlation, due to the resident habits, that links natural gas consumption and temperature. Since the central furnace for the house heating is activated only when it is too cold (starting from 2 ℃)[1], it is conceivable that intrinsic correlation can bind the temperature with the consumption of natural gas. Actually, the instant hot water unit is powered by natural gas and the use of hot water is strongly correlated to the temperature. So, supposing that no correlation between gas and water consumption exists, during the cold period the consumption of hot water increases, at the expense of the "cold" (not-heated) water, causing an increase in natural gas consumption.

The water results do not shown evident correlation, in fact the best results for 12 h and 24 h have been obtained using only the water consumption as

[1] AMPds FAQ: `http://ampds.org/`

input. On the contrary, the best results at $1\,h$ and $6\,h$ have been achieved using electrical energy, gas, and temperature information as input. Moreover, decreasing the resolution from $6\,h$ to $1\,h$, all the approaches exhibit a performance reduction stronger than the one obtained for any other resolution step.

4.3.2 DFID Dataset

Table 4.3: Best results achieved for each technique applied to DFID dataset. The column marked with "Parameters" reports the parameters combination that achieves the best result for the corresponding approach.

DFID	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Parameters
	Overall Gas					
ARIMA	0.354	0.646	4.000E-3	73.986	119.99	2-0-1
ANN	**0.266**	**0.734**	**3.041E-3**	**48.868**	**103.980**	**5-13**
DBN	0.509	0.491	5.766E-3	82.521	145.530	3-1-2-500-10-8
ESN	0.395	0.605	4.391E-3	65.309	128.750	5-25-1-0.5-0.01
SVR	0.274	0.726	3.048E-3	47.551	107.270	5-0.5-8
ELM-LIN	0.510	0.490	5.784E-3	66.951	145.770	3-2
ELM-RBF	0.267	0.733	2.969E-3	49.088	105.860	5-8-0.125
GP	0.457	0.543	5.080E-3	53.283	138.480	5-20-1
	Abercrombie House Gas					
ARIMA	**0.255**	**0.745**	**4.100E-3**	**40.911**	**98.335**	**3-0-1**
ANN	0.275	0.725	4.295E-3	50.809	103.310	5-13
DBN	0.493	0.507	7.557E-3	77.664	138.770	5-1-2-500-10-8
ESN	0.396	0.604	6.072E-3	65.935	124.390	5-25-1-0.5-0.01
SVR	0.298	0.702	4.575E-3	47.631	107.970	5-0.5-8
ELM-LIN	0.513	0.487	7.862E-3	67.763	141.540	5-128
ELM-RBF	0.286	0.714	4.378E-3	49.881	105.630	5-8-0.125
GP	0.535	0.465	8.201E-3	56.699	144.560	5-20-6
	Abercrombie House Water					
ARIMA	0.902	0.098	3.170E-3	130.439	172.404	5-0-1
ANN	**0.676**	**0.324**	**3.581E-3**	**98.505**	**149.360**	**5-13**
DBN	0.844	0.156	4.687E-3	123.42	164.260	2-1-2-100-100-7
ESN	0.750	0.250	4.118E-3	103.41	156.070	5-40-5-0.5-0.01
SVR	0.753	0.247	4.131E-3	94.931	156.320	5-512-2
ELM-LIN	0.843	0.156	4.631E-3	98.464	165.500	5-0.125
ELM-RBF	0.686	0.314	3.766E-3	98.358	149.250	5-128-0.125
GP	0.891	0.108	4.894E-3	83.909	170.140	5-10-10
	Whitehall Water					
ARIMA	0.233	0.767	5.800E-3	54.960	81.352	3-0-2
ANN	**0.121**	**0.879**	**8.345E-05**	**32.254**	**59.428**	**5-6**
DBN	0.578	0.422	4.598E-4	83.898	136.990	2-1-2-100-100-4
ESN	0.367	0.633	2.505E-4	56.398	103.730	5-40-12-0.25-1E-13
SVR	0.150	0.850	1.021E-4	31.783	66.214	5-32-8
ELM-LIN	0.489	0.511	3.335E-4	60.912	119.680	5-0.5
ELM-RBF	0.143	0.857	9.783E-5	32.824	64.820	5-8192-0.125
GP	0.343	0.657	2.342E-4	42.226	100.280	5-20-4

For all the DFID sub-sets, with the exception of the Abercrombie House gas, the ANN approach achieves the best performance, as depicted in Table 4.3.

However, the result achieved with the ANN for the Abercrombie House gas is close to the best one. The longest sub-sets are the gas ones, about 5 years of records, and they achieve close results.

Specifically, for the overall gas sub-set the best result shows a NMSE of 0.266, a R^2 of 0.734, a MSE of 0.0030, a MAPE of 48.87%, and a RRMSE of 103.98%, with 13 hidden neurons. For the Abercrombie House water sub-set, the best result achieves a NMSE of 0.676, a R^2 of 0.324, a MSE of 0.0036, a MAPE of 98.51%, and a RRMSE of 149.36%, with 13 hidden neurons. Whereas, the best result for the Whitehall sub-set achieves a NMSE of 0.121, a R^2 of 0.879, a MSE of $8.34E-5$, a MAPE of 32.25%, and a RRMSE of 59.43%, with 6 hidden neurons.

In the case of Abercrombie House gas sub-set, the best result, achieved by the ARIMA approach, reaches a NMSE of 0.255, a R^2 of 0.745, a MSE of $4.100E-3$, a MAPE of 40.911%, and a RRMSE of 98.335%, with an AR of order 3 (or lags) and a MA of order 1.

Considering the ANN results, for all the sub-sets the best performance is achieved using 5 lags, and both SVR and ELM-RBF confirmed their good behaviour for short-term prediction. As for the AMPds, ESN, DBN, and GP are unsuitable for this case scenario. Noteworthy, all the approaches show a marked performance deterioration for the Abercombrie House water sub-set.

4.3.3 EIA Datasets

For natural gas consumption of three U.S. States, provided by the EIA datasets, the achieved results are reported in Table 4.4. With the exception of the Illinois consumption, for all the techniques implemented the Louisiana and Maine forecasts have achieved better results when combining both consumption and temperature as input data.

The best results for each State have been obtained with the ANN or the EKF-GP. In particular, for Illinois the best result is obtained using the EKF-GP model for 5 lags and a depth of 10, without temperature information, by achieving a NMSE of 0.044, a R^2 of 0.955, a MSE of $3.374e-3$, a MAPE of 13.251%, and a RRMSE of 20.780%.

Concerning Louisiana, a NMSE of 0.061 and an R^2 of 0.938 have been reached, working with the GP model obtained for 3 lags and a depth of 20 and adopting the temperature data. The values achieved by the others evaluation criteria are $3.087E-3$ for the MSE, 15.724% for the MAPE, and 23.245% for the RRMSE.

Finally, the prediction for Maine, by unsing the temperature information, achieved a NMSE of 0.030 with an R^2 of 0.970, using the EKF-GP model obtained for 3 lags and depth 20. Specifically, the MSE is $1.282E-3$, the

Table 4.4: Best results achieved for each technique applied to EIA datasets. The column marked with "Param." reports the parameters combination that achieves the best result for the corresponding approach.

EIA	NMSE	RRMSE	Parameters	NMSE	RRMSE	Parameters
	Illinois			**Illinois&T**		
ARIMA	0.199	39.5	2-0-1	-	-	-
ANN	0.053	22.7	5-5	0.067	26.1	2-6
DBN	0.058	23.9	5-5-500-10	**0.055**	**23.1**	**5-7-500-10**
ESN	0.097	30.6	5-40-0-1/2-0	0.069	25.9	5-20-0-1/2-1E-15
SVR	0.066	25.4	$5\text{-}2^{11}\text{-}2^{-3}$	0.066	25.4	$5\ 2^{5}\text{-}2^{-3}$
GP	0.081	28.0	5-10	0.092	27.0	3-10
EKF-GP	**0.044**	**20.8**	**5-10**	0.073	24.1	3-10
ELM-LIN	0.203	39.8	2-2	0.090	29.5	$2\text{-}2^{15}$
ELM-RBF	0.062	24.6	$5\text{-}2^{5}\text{-}2^{-1}$	0.064	24.9	$5\text{-}2^{3}\text{-}2$
	Louisiana			**Louisiana&T**		
ARIMA	0.307	51.6	2-0-1	-	-	-
ANN	**0.085**	**30.6**	**2-6**	0.066	23.5	3-7
DBN	0.148	36.3	3-2-500-10	0.072	25.2	5-8-500-10
ESN	0.132	33.9	2-15-14-1-0	0.097	29.0	2-20-1-1/2-1E-4
SVR	0.096	29.0	5-2-23	0.070	24.9	5-2-2
GP	0.132	34.2	3-10	0.095	29.1	3-20
EKF-GP	0.086	27.6	3-10	**0.061**	**23.2**	**3-20**
ELM-LIN	0.292	50.4	$2\text{-}2^{15}$	0.155	37.0	5-2
ELM-RBF	0.103	30.0	$5\text{-}2^{5}\text{-}2^{-3}$	0.066	24.1	$5\text{-}2^{7}\text{-}2$
	Maine			**Maine&T**		
ARIMA	0.158	34.4	2-0-2	-	-	-
ANN	**0.051**	**21.5**	**5-5**	0.065	23.3	2-6
DBN	0.111	29.6	3-4-100-100	0.092	27.0	3-3-500-10
ESN	0.074	24.2	3-40-10-1/2-1e-14	0.079	24.3	2-35-10-1/3-1-4
SVR	0.066	22.9	$3\text{-}2^{13}\text{-}2^{-3}$	0.045	18.8	$3\text{-}2^{5}\text{-}2^{-1}$
GP	0.095	27.5	3-15	0.062	22.1	3-20
EKF-GP	0.052	20.4	3-15	**0.030**	**15.4**	**3-20**
ELM-LIN	0.147	34.2	$3\text{-}2^{15}$	0.110	28.7	$2\text{-}2^{9}$
ELM-RBF	0.072	23.9	$3\text{-}2^{11}\text{-}2$	0.048	19.5	$3\text{-}2^{9}\text{-}2$

MAPE is 11.267% and RRMSE is 15.353%. The Maine prediction results represent the overall best ones.

The DBN shows good performance for Illinois, with a slight improvement, NMSE goes from 0.058 to 0.055, by adding temperature data. For Maine, without the temperature information the best result is achieved by ANN, and for both Louisiana and Maine the SVR shows good performance. It is noteworthy that for Louisiana and Maine, where the best result has been achieved using the temperature information, all the approaches, except ANN, show an improvement when using the temperature. Conversely, for the Illinois case, where the best result has been achieved without temperature information, not all the approaches reach their best performance without the temperature data. Actually, DBN, ESN and, albeit slightly, SVR provided better results using the temperature too.

The correlation between gas consumption and temperature can be explained by the extensive use of natural gas for residential heating. So, further investiga-

tions are needed to understand the correlation of gas consumption and temperature for the Louisiana and Maine cases, and the apparently missing correlation among them for the Illinois case, also considering that Illinois presents a mild climate, whereas Maine and Louisiana have cold and hot climate, respectively.

4.3.4 Tehran Dataset

Table 4.5: Best results achieved for each technique applied to Tehran water consumption. The column marked with "Parameters" reports the parameters combination that achieves the best result for the corresponding approach.

Tehran	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Parameters
ARIMA	0.414	0.574	1.730E-2	26.690	31.150	2-0-1
ANN	**0.136**	**0.860**	**8.615E-3**	**12.638**	**18.540**	**5-7**
DBN	0.266	0.727	1.384E-2	17.620	21.352	3-3-500-100
ESN	0.209	0.785	1.119E-2	17.675	21.260	5-30-10-1-1E-2
SVR	0.148	0.849	7.893E-3	15.131	17.854	$5\text{-}2^{13}\text{-}2^{-3}$
GP	0.208	0.786	1.112E-2	17.430	21.192	5-20
EKF-GP	0.240	0.753	1.285E-2	19.402	22.780	5-15
old ELM-LIN	0.293	0.699	1.564E-2	20.804	25.129	$5\text{-}2^{3}$
ELM-RBF	0.152	0.843	8.139E-3	14.729	18.130	$5\text{-}2^{15}\text{-}2^{3}$

The overall best performance has been achieved with the ANN approach for 5 lags with 7 neurons in the hidden layer. The result reaches a NMSE of 0.136, a R^2 of 0.860, a MSE of $8.615E-5$, a MAPE of 12.639, and a RRMSE of 18.540. Noteworthy, the SVR and the ELM-RBF approaches for 5 lags show slightly worse performance but achieves a lower MSE. For SVR the results has been achieved with C and γ values of 2^{13} and 2^{-3}, respectively. For ELM-RBF the results has been achieved with C and σ values of 2^{15} and 2^{3}, respectively.

The introduction of the EKF did not produce further improvement for the GP, and, unfortunately, Nasseri *et al.* [17] used the EKF approach only to provide an on-line forecasting, so no comparative results have been provided in their paper.

4.4 Remarks

In the case of short-term forecasting, AMPds and DFID datasets, ANN, SVR, and ELM-RBF confirmed their good behaviour. On the contrary, ESN, DBN, and GP approaches have confirmed to be unsuitable for the domestic and office building environments.

On the other hand, for medium-term forecast, EIA and Tehran datasets, EKF-GP, ANN, SVR, and ELM-RBF confirmed their good behaviour.

The application of heterogeneous information, AMDds and EIA datasets, results in performance improvement for almost the approaches. In the case of AMPds, a correlation between gas consumption and temperature information has been highlighted in the natural gas forecasting, and the overall good results confirm the advantage of heterogeneous data utilization. In the case of EIA datasets, for Louisiana and Maine the combination of consumption data and temperature information achieved better performance than the use of the water consumption information only.

Chapter 5

Automatic Leakage Detection

The method adopted for the automatic leakage detection has been inspired by the works of [85] and [86], where an audio novelty detection is performed by exploiting a statistical model of the normal condition of the system. The statistical approach in novelty detection has never been used for leakage identification, and within the state of the art, Section 2, none of the contributions seem aimed at residential or building environments. Specifically, to the best of the author's knowledge, among recent studies, no one has targeted the natural gas grid in a residential scenario and water and gas grids in a office building scenario.

5.1 Framework

The proposed framework is composed of three main parts, which are the feature extraction, the feature selection and the leakage detection. In Figure 5.1, the constitutive blocks for the feature extraction and the leakage detection are depicted.

As stated above, the leakage identification is based on the normality model, which captures the statistical features of the data without recorded leakages. For this reason, a two-stage framework has been proposed. In the normality model creation stage, the features' information is processed to create and train the statistical model of the normal behaviour. After this training phase, the models are stored to be used in the novelty detection stage. In the validation phase, all of the decisions are based on the log-likelihood/distance values that are compared against a set of thresholds. For each frame, the corresponding log-likelihood/distance value is obtained from the constitutive equation of the model, with the normality parameters obtained during the training phase and the new input features' data. The set of thresholds depends on the overall log-likelihoods/distances computed for the validation set. Being that the dataset is split into frames, the lowest resolution of the decision is represented by the frame length. For both stages, the input datasets are firstly processed in the

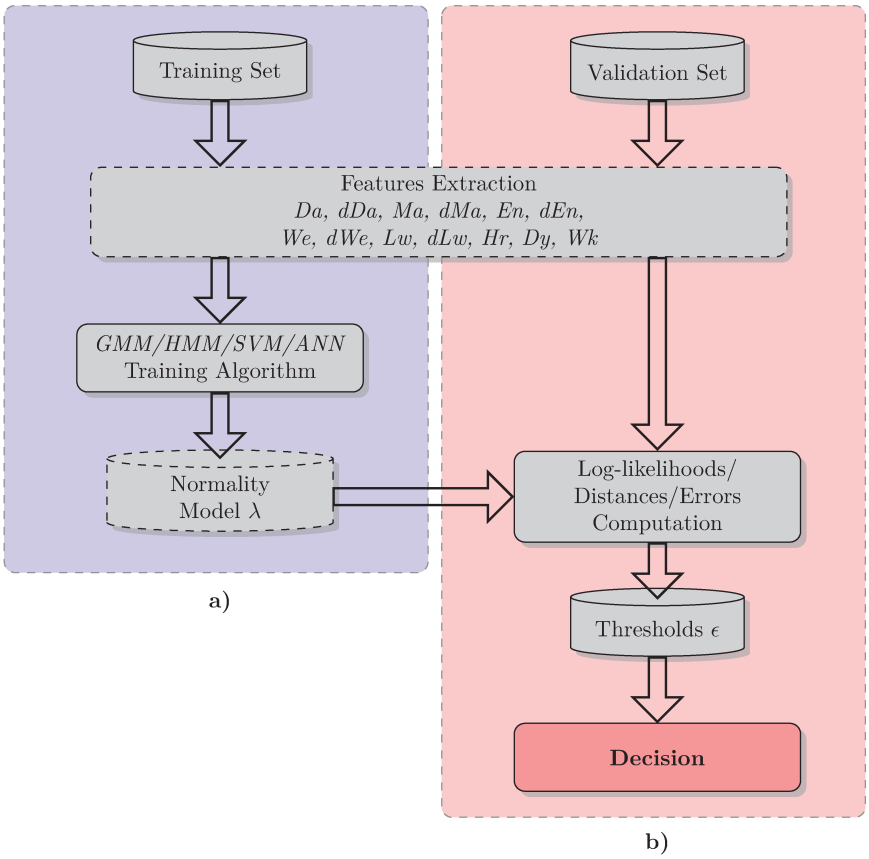


Figure 5.1: Algorithm stages: (a) normality models creation, (b) novelty detection.

features' extraction block, in order to split the dataset into frames and to extract the selected features.

Three different models, extensively described in the Chapter 3, have been adopted to compute the normality model: the GMM, the HMM and OC-SVM. Furthermore, preliminary evaluations have been also performed adopting the AE structure for the ANN.

The features' selection stage operates directly on the set of extracted features based on the detection performance. For each new selected feature, the whole feature extraction and leakage detection operations are repeated. The proposed selection algorithm is extensively described in the following section.

5.1.1 Decision

The decision stage, which carries out the detection of an event as being normal or abnormal, is depicted in the novelty detection stage in Figure 5.1, and is

composed of four blocks, strictly interconnected among themselves: *Normality Model* λ, *Log-likelihoods/Distances/Errors Computation*, *Thresholds* ϵ and *Decision*. For each frame, in the GMM approach, or for each sequence of frames, in the observation sequence for the HMM approach, the vector of extracted features is used to compute the corresponding log-likelihood value, by means of Equation (3.35) or (3.36), respectively, and by adopting the trained parameters of the background model.

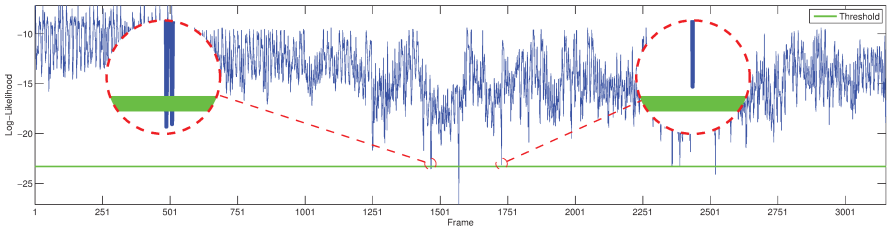


Figure 5.2: Threshold example: the log-likelihood exceeds the threshold, the abnormal event detected (**left circle**); the log-likelihood does not exceed the threshold, normal event (**right circle**). The y-axis represents the log-likelihood value computed for each frame, while the x-axis represents the frame number.

At first, the validation data are processed and the achieved log-likelihoods stored. Then, a set of thresholds, to compare against the achieved log-likelihoods, is computed, and a non-uniform distribution of the threshold values is adopted, by arranging a higher number of thresholds, wherever the log-likelihood presents a denser distribution.

In the decision block, each threshold is compared against the log-likelihoods computed for the frame (GMM case) or the sequence of frames (HMM case). As depicted in Figure 5.2, if the log-likelihood value exceeds the threshold, the event is marked as a leakage/abnormal event, otherwise it is considered a normal event. Being that the HMM is based on a sequence of observations, consecutive frames are dragged in the decision process. In particular, the number of frames, involved in the observation, is equal to the number of states assumed for the normality background. Since the value of the likelihood returned by the algorithm is a logarithm, the data that match the statistical model will produce a probability value close to one, which corresponds to a value close to zero to the logarithm. On the other hand, the data not matching the statistical model will generate a probability value close to zero; thus, the logarithm value will be negative, while its absolute value will be great, as shown in Figure 5.2.

With respect to the block diagram Figure 5.1, the adoption of the OC-SVM affects the novelty detection stage. Instead of using the log-likelihood values

returned by the GMM or HMM, the decision is based on the distances of each frame from the separation hyperplane of the feature points, obtained during the training phase. The set of thresholds used in the validation process is chosen consequently.

Again, also the adoption of the AE affects the novelty detection stage. In this case, the RMSE, computed from the input and output values of the network, is adopted in the thresholding process. The set of thresholds used in the validation process is chosen consequently.

5.2 Features

In the literature, a well-defined set of features suitable for this application is yet to be proposed, therefore a collection of both known and new features has been used to fill this gap. In particular, in order to provide a comparative evaluation, all of the suitable features available in the current literature, presented in Section 2.2, have been considered.

The energy of signals and of their wavelet decomposition sub-band, first presented in [31], have been used. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a vector composed of the data of the frame, with $N \in \mathbb{Z}$ samples; the energy is defined as:

$$En = \sum_{i=1}^{N} {x_i}^2 \ . \tag{5.1}$$

Moving to the wavelet components, the decomposition of order three is performed by applying the Daubechies 2 (*db2*) wavelet, and overall, four sequences are obtained: three detail sequences (D_1,D_2,D_3) and one approximation sequence (A_3). For each sequence, (5.1) holds true, as well; thus, the energy is computed by assuming N as the number of elements in the corresponding sequence. The feature, called *We*, is evaluated as a vector, composed of four elements and representing the percentage of energy distribution between the wavelet components. Let $C = \{En_{A_3}, En_{D_3}, En_{D_2}, En_{D_1}\}$ be the vector containing the computed energy for each decomposition sequence, for each j-th elements of the C; the corresponding *We* values is given as:

$$We_j = 100 \cdot \frac{C_j}{En_T} \ , \tag{5.2}$$

where En_T denotes the overall energy of the four components of the decomposition, given as the L^1-norm of C, $En_T = \|C\|_1$.

Based on this method, the author investigated the application of the logarithm operator to the energy values of the wavelet components; thus, the j-th

element of the feature Lw is defined as:

$$Lw_j = log(C_j + 1) \ . \tag{5.3}$$

In order to provide a better insight into the features' performance, the frame samples have been also adopted as a feature (Da), as well as the frame average (Ma), corresponding to the moving average of the set, for a window length equal to the frame length.

For the above features, the corresponding first order positive difference has been proposed as a feature, as well. For instance, a first order positive difference is denoted by the letter d followed by the name, F, of the related feature. Thus, being $F(n)$ the feature extracted from the n-th frame, the derived feature $dF(n)$ is computed as:

$$dF(n) = \begin{cases} F(n) \ , & if \ n \leq 2 \\ F(n) - F(n-2) \ . & otherwise \end{cases} \tag{5.4}$$

where $n \in [1, Nf]$ and Nf denotes the number of frames in the set. In the case of features with more components, an element-wise difference is performed. The value of two, which appears in the difference definition, is the required value that guarantees a marked dissimilarity between the frames, as well as their features, at the leakages' beginning.

5.2.1 Temporal Features

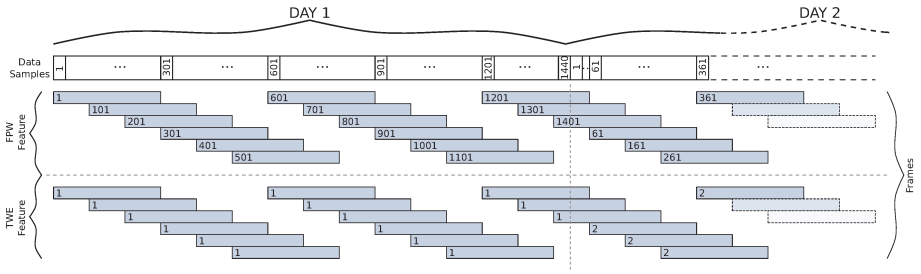


Figure 5.3: Example of daily temporal features. The pale blue blocks represents the overlapped frames, and the numbers in "Data Samples" indicate the samples enumeration within the reference window. In the frame blocks are reported the value assigned to the corresponding FPW and TWE feature.

To assess the usefulness of temporal information, said information has been introduced in the form of two alternative sets of features, namely *frame position in temporal window* (FPW) and *temporal window enumeration* (TWE). The

sets are composed of three features, computed for three time window lengths, respectively of one hour, one day and one week. Regarding these windows, each one is repeated over time without overlapping itself.

Specifically, the hourly window starts at the first minutes of the hour and ends at the last minute of the same hour. The daily window goes from the first minute (higher resolution for the adopted dataset) after the midnight, 00:01, to the midnight of the next day, 00:00. Finally, the weekly window starts at the first minute after the midnight of Sunday, and ends at the midnight of Sunday.

According to the position of the frame first sample, the FPW features provides the frame position within the temporal window. In particular, all the samples within a temporal window are enumerated and for each frame the assigned feature corresponds to the value assigned to its first sample, thus it represents the starting position of the frame with respect to the overall sample set within the temporal window.

On the other hand, regarding the TWE features, given a default sequence of temporal windows, the information regarding the belonging of the frame to a temporal window is provided as feature information. Specifically, the sequence of temporal windows are enumerated in accordance with the sequence span. About the hourly window the evaluated window sequence spans over a day; for the daily window the evaluated window sequence spans over a week; finally, for the weekly window the evaluated window sequence spans over a year. Specifically, for each frame, the assigned feature value corresponds to the position, within the sequence of temporal windows, of its first sample. Thus, according to the selected temporal window, it represents the starting position expressed in terms of hour of the day, day of the week, or week of the year. An example of the values assigned to the frames for both the temporal features with a daily window is represented in Figure 5.3.

Table 5.1: Feature details.

Index	**Name**	**Feature Size**	**Acronym**
1	Data	Number of samples	*Da*
2	Energy	1	*En*
3	Moving Average	1	*Ma*
4	Wavelet Decomposition Energy	4	*We*
5	Logarithmic Wavelet Energy	4	*Lw*
6	Hourly Window	1	*Hr*
7	Daily Window	1	*Dy*
8	Weekly Window	1	*Wk*

Except for *Da* and its first order positive difference (*dDa*), the lengths of which depend on the number of frame samples (300, 30, 10 samples at 1, 10 and 30 min of resolution, respectively), the remaining features have a fixed length, independently from dataset resolution and frame length. Specifically,

Ma, *En*, *We*, *Lw*, *Hr*, *Dy* and *Wk* have a feature length of 1, 1, 4, 4, 1, 1 and 1, respectively, as summarized in Table 5.1.

5.2.2 Features Selection

As anticipated in Section 5.2, in the literature, a well-defined set of standard features is yet to be defined. Furthermore, an exhaustive research requires testing a number of combinations that can be computed as the binomial coefficient: $m!/l!\,(m-l)!$, where m denotes the number of available features and l the features for each vector combination. In other words, the overall number of combinations is given as $\sum_{l=1}^{m} m!/l!\,(m-l)!$, with $m = 10$, for a total of 1023 combinations, meaning that an investigation of all of the possible combinations is unaffordable from a computational standpoint. To overcome this issue, a modified version of the Sequential Forward Selection (SFS) algorithm [87] has been adopted to perform the feature selection.

The SFS is a suboptimal method that allows reducing the number of tested combinations to $lm - l(l-1)/2$, by lowering the overall number of combinations to $\sum_{l=1}^{m} lm - l(l-1)/2 = m(m+1)/2$. For instance, given $m = 10$ available features, a total of 55 combinations is tested by means of the SFS, against the 1023 of the full-search.

The SFS consists of l sequential steps; in each step, a winning combination of features is selected from a pool of features, and with respect to the previous step, the number of features in the winning combination is increased by one, removing the selected feature from the pool. Further details on the SFS algorithm are described in Algorithm 1, where the Area Under the Curve (AUC), defined in Section 5.4, has been used as the criterion value.

Algorithm 1 Sequential forward selection.

procedure INITIALIZATION
    - Create a set of m feature vectors, each composed of a single feature.
end procedure
procedure COMBINATION
    for $i = 1 \rightarrow (m-1)$ **do**
        - Compute the criterion value for each vector of the set.
        - Select the winner vector, i.e., the one whose criterion value is greatest.
        - Create a new set of $m - i$ feature vectors, each of them composed of the winning vector and one of the excluded features.
    end for
end procedure
- Select the feature vector whose criterion value is greatest among all of the created feature vectors.

As depicted in Figures 5.4 and 5.5, the low performance achieved in the first iteration, in particular with the GMM, has suggested that the chances of missing a good combination are considerably high; this is the reason why a slight variation in the first and the second step of the SFS algorithm has been introduced.

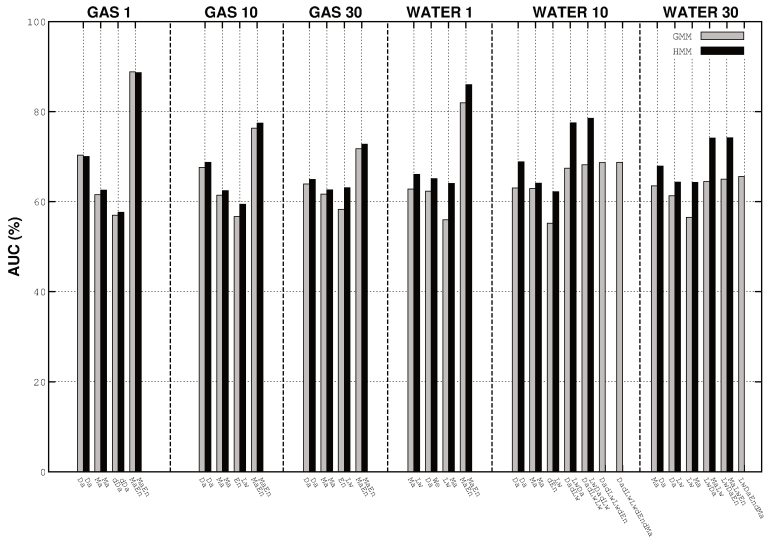


Figure 5.4: Feature selection steps and corresponding features' combination until the achievement of the best result for the AMPds datasets. In the x-axis are reported the corresponding best features' combinations for both GMM and HMM.

In our implementation of the SFS algorithm, as shown in Table 5.2, at the first step, instead of choosing a single winner, three winners are selected, that is the three features that have the best results. Then, in the second step, for each winner, a new features set, as reported in Algorithm 1, is created and tested. Among all of these sets, the winner to be used in the next step is selected, and the SFS default procedure is followed from now on. While improving the search performance, the proposed selection method remains a suboptimal search approach, keeping an overall number of required iterations, 73 for $m = 10$, still lower than an exhaustive search.

5.3 Experimental Setup

5.3.1 Offset Leakage

As discussed above, in the novelty detection approach, no faulty instance is adopted to train the normality model, and none is present in the datasets. As

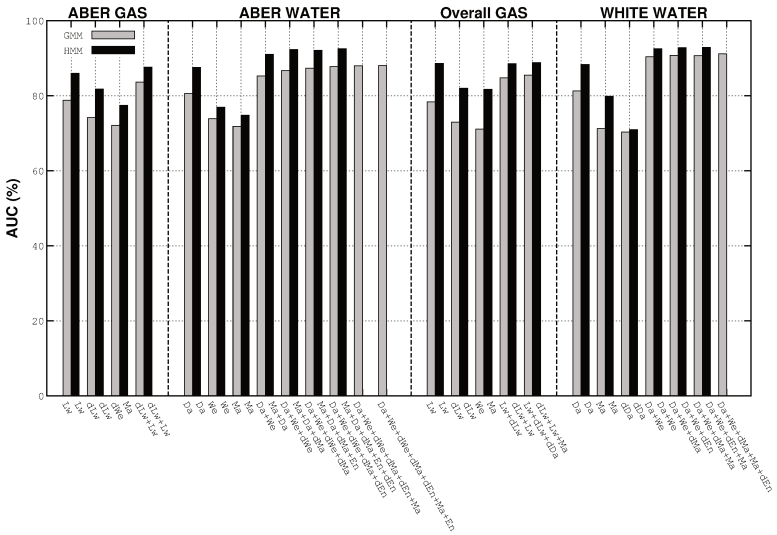


Figure 5.5: Feature selection steps and corresponding features' combination until the achievement of the best result for the DFID datasets. In the x-axis are reported the corresponding best features' combinations for both GMM and HMM.

Standard SFS

1st	2nd	3rd	n-th
F_1	F_2, F_1	F_2, F_3, F_1	...
F_2	F_2, F_3	⋮	
F_3	⋮	F_2, F_3, F_n	
⋮	F_2, F_n		
F_n			

Proposed SFS

1st	2nd	3rd	n-th
F_1	F_2, F_1 ⋮ F_2, F_n	F_3, F_1, F_2	...
F_2	F_3, F_1 ⋮ F_3, F_n	⋮	
F_3		F_3, F_1, F_n	
⋮			
F_n	F_n, F_1 ⋮ F_n, F_{n-1}		

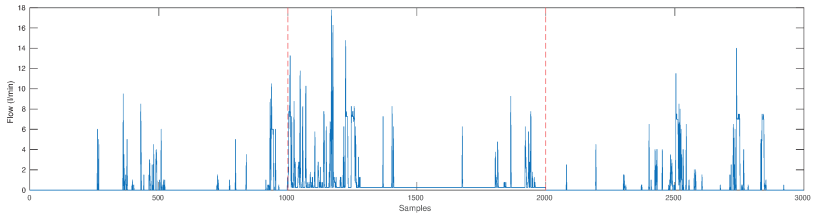
Table 5.2: Comparison between the first steps of the standard Sequential Feature Selection (SFS) algorithm and its revised counterpart. For each step, the red text denotes the winner feature or the features' combination. The red boxes denote the 3 different features sets generated starting from the 3 winner features of the previous step.

said, also, to the author's knowledge, in addition to the suitable datasets being limited in number, none of them records the occurrence of faults. Therefore, the validation datasets have been manipulated to include faulty instances, namely *offset* leakage, and to perform the validation of the system. Each validation

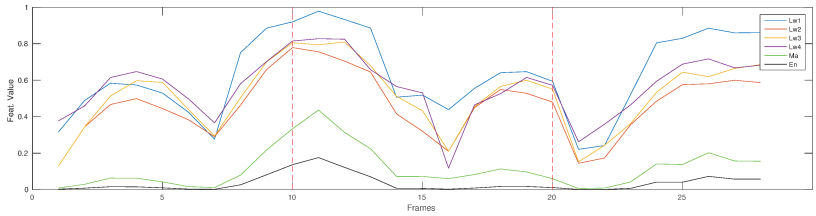
sequence, v, has been revised as:

$$v(i_v) = \begin{cases} v(i_v) + \beta \cdot l_k, & if\ {i_v}^{(i)} \geq n \geq {i_v}^{(f)} \\ v(i_v), & otherwise \end{cases} \tag{5.5}$$

where $\beta \in [0.25, 0.50]$, l_k is the average consumption computed from the training sequence, while ${i_v}^{(i)}$ and ${i_v}^{(f)}$ denote the leakage start and stop samples.



(a)



(b)

Figure 5.6: Example of flow data with induced leakage (**a**) and the corresponding set of extracted features (**b**). The dashed lines define the leakage start and end. For each frame, the reported features are the *Ma*, the *En* and the four components of the *Lw*.

The correctness of the modelled leakages records has been confirmed in a comparison against the ones collected in a real-life case environment [88]. In a residential scenario, and based on AMPds values, a leakage value spanning from 25% to 50% of the average consumption levels, corresponds to $7.76\,L/h$ to $15.53\,L/h$. In a real life environment, the most common leakage rates are $20\,L/h$ and $10\,L/h$, whereas the 49% of leakages are at most equal to $20\,L/h$, thus proving the correctness of the modelling criteria. To further improve the model, the leakage size, as well as the duration and the starting point, have been randomly selected. Specifically, a limited duration of the leakages has been chosen in order to simulate the correct detection of the anomaly and consequent restoration of standard operating conditions within the same observation window.

The leakage duration is randomly selected in a range between a minimum of 5 and a maximum of $10\,h$, which corresponds to 300 and 600 samples, respec-

tively, for 1 minute resolution, to 30 and 60 samples, respectively, for 10 *min* resolution, and 10 and 20 samples, respectively, for 30 *min* resolution. Furthermore, the starting point of the leakage is randomly selected over a span between the 10% and 90% of the validation set length. An example of leakage induced in the data, together with a subset of features extracted from the corresponding data, is shown in Figure 5.6.

5.3.2 Simulated Leakage

A further typology of faulty instance, namely *simulated* leakage, has been also taken into account. Differently from the previous definition, the leakages have been modelled in a simulated circuit, and both flow and pressure information have been computed. The circuit has been designed in order to represent a typical domestic network by exploiting the EPANET tool [37]. The network has a length of 30 *m*, with an overall elevation of 3 *m*. The water supplier is simulated by means of a reservoir with an elevation of 30 *m*, that is defined in the EPANET Users Manual [89] as "an infinite external source or sink of water to the network", and "its head (elevation) cannot be affected by what happens within the network". The available data pieces, about water consumption from the AMPds, are used as consumption patterns in specific junctions or nodes. The simulations have been performed manipulating the EPANET project file (text version) with MATLAB.

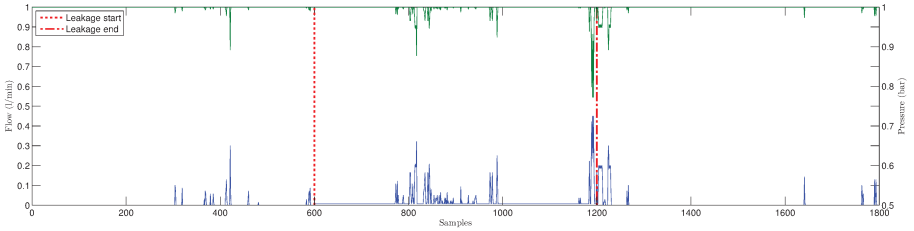


Figure 5.7: Example of simulated flow (blue) and pressure (green) data (normalized values) with induced leakage. The red dotted lines delimit the leakage.

Moreover, the network assumes a 20 *mm* diameter standard piping, and a roughness (C-factor) of 100. The output pressure and flow values are computed from the first junction of the network, connected to the reservoir, and from the corresponding connection pipe, respectively.

In particular, the *simulated* leakages have been modelled by means of an *emitter* in the simulated circuit. In this way, the leakage output flow (leakage size) depends on flow and pressure in the reference node. As done in Nasir *et al.* [33], the leakage size has been varied by manipulating the discharge coefficient

of the emitter. According to the EPANET Users Manual [89], "the emitter models the flow through a nozzle or orifice that discharges to the atmosphere. The flow rate through the emitter varies as a function of the pressure available at the node":

$$Q = C \cdot P^{g}, \tag{5.6}$$

where Q is the flow rate, P is the pressure, C the discharge coefficient, and g is the pressure exponent, usually equal to 0.5 for nozzles and sprinkler heads. For the simulated circuit both flow and pressure data are collected, measuring them from the input (first) pipe and node, respectively. Therefore, for the samples involved, the flow and pressure data, q_n and p_n, have been replaced with new values given by the simulation. A random discharge coefficient $C \in [0.024, 0.047]$ has been assumed, in order to evaluate leakages compliant with a real-case scenario [88], as discussed in Section 5.3.1. An examples of simulated flow and pressure data is depicted in Figure 5.7.

5.3.3 Computer Simulation Setup

The datasets have been split into two parts; 70% of the data are used for training purposes (background model creation), and the remaining 30% is used for testing (detection performance evaluation). Furthermore, each subset has been split in overlapping frames with a fixed length of $5\,h$, thus 300 samples at $1\,min$ resolution, 30 samples at $10\,min$ and 10 samples at $30\,min$. The overlap is equal to 2/3 of the frame length. A label vector has been created for the validations set, by marking with "0" (zero) the normal frames, and "1" (one) the abnormal ones.

The background model creation includes an initial random condition, which can cause slight changes in the model from one run to another. Therefore, based on the same training set, 10 background models have been created. Each model has been adopted to evaluate 10 leakages, the parameters of which are randomly selected.

The GMMs have been tested by varying the number of adopted Gaussian components as: $N_g = \{2, 4, 8, 16, 32, 64, 128, 256\}$. For the HMMs, a typical structure used in speech recognition to model the temporal processes [75] has been adopted, and it is based on the left-to-right structure (also called Bakis), depicted in Figure 3.8. In addition to the Gaussian components, the number of states has been varied as well, from one to four (plus the start and end states, which are always present). For the OC-SVM only, the γ parameter has been varied based on the following pool of values: $\left\{2^{-15}, 2^{-13}, \ldots 2, 2^{3}\right\}$.

The data preparation, the feature extraction and the decision stages have been developed in MATLAB®: MathWorks, Natick, MA, USA. The GMM and HMM training algorithms, as well as the likelihood computation, have

been implemented in C++ using the Torch3 library [90]. The OC-SVM has been implemented using the LIBSVM library [81]. The experiments have been performed on a computer with Intel(R) Xeon(R) CPU E5620 @ 2.40GHz, and 32 Gb RAM.

In order to produce reliable and consistent results, the random leakages parameters have been produced by initializing the MATLAB® random generator with the seed of 42. Moreover, the Min-Max normalization has been also applied to each feature set. In particular, let $\mathbf{x_f}$ be the feature set for each frame; the corresponding normalized set $\mathbf{x_f}^* \in [0,1]^L$, with L denoting the number of elements in the vector $\mathbf{x_f}$, is given performing the following element-wise operation:

$$\mathbf{x_f}^* = \frac{\mathbf{x_f} - \mathbf{x_f}^m}{\mathbf{x_f}^M - \mathbf{x_f}^m} \ , \tag{5.7}$$

where $\mathbf{x_f}^M$ and $\mathbf{x_f}^m$ denote the vectors of maximum and minimum values, respectively. For each component of $\mathbf{x_f}$, the maximum and the minimum values are computed from the training set.

5.4 Evaluation Methods

In order to evaluate the performance of the leakage detection system, the methods classified as concurrent comparison [84], that compare real and modelled values, have been chosen. Specifically, the probability to correctly detect an anomaly, namely true detection rate (TDR), and the probability to erroneously detect an anomaly, namely false detection rate (FDR), have been adopted, and, according to [91], are given as follows:

$$\begin{aligned} TDR = & \frac{no.\ of\ abnormal\ events\ detected\ as\ abnormal}{no.\ of\ abnormal\ events}, \\ FDR = & \frac{no.\ of\ normal\ events\ detected\ as\ abnormal}{no.\ of\ normal\ events}. \end{aligned} \tag{5.8}$$

Assuming a set of thresholds, a pair TDR-FDR is computed from each threshold, therefore a corresponding set of TDR-FDR pairs is obtained. This set is used to plot the Receiver Operating Characteristic (ROC) curve [92], which shows how the threshold affects the result.

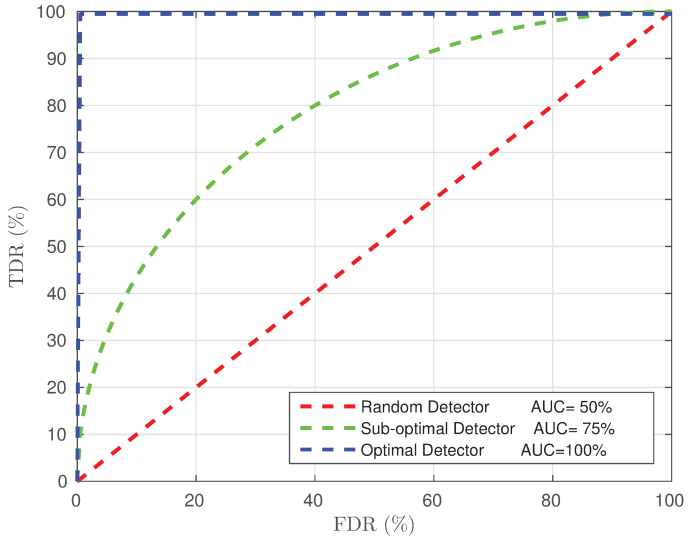


Figure 5.8: Examples of the ROC and AUC relations.

On the one hand, if the threshold is set to a value lower than the minimum likelihood, the system achieves a TDR and a FDR equal to 0%. Each normal event is correctly detected, but all of the abnormal events are classified as normal, as well. On the other hand, if the threshold is set to a value higher than the maximum likelihood, both the TDR and the FDR reach a 100% rate. In the latter case, then, each sequence is recognized always as abnormal notwithstanding its actual nature.

As last step, for a given ROC, the goodness of the classifier performance is evaluated in terms of AUC, as done in [85, 86, 91]. Specifically, the plot relating to a random guess detector and an ideal detector is reported, as a reference, in Figure 5.8. It is possible to observe that a completely random guess will produce a set of points along the diagonal of the ROC, producing an AUC of 50%. On the other hand, the optimal detector will produce a 100% AUC, and the best threshold will be the one that produces a TDR of 100% and a FDR of 0%. Thus, a novelty detector, to be of particular use, should have a plot above the diagonal of the ROC. Furthermore, the closer to the top-left corner of the ROC, the better the detector performance will be.

5.5 Experiments: Residential Scenario

In the following tables, the best results and the corresponding features' combination achieved for each resource and resolution (**Res.**) of the AMPds datasets, are reported. The column "**Parameters**" reports the number of Gaussians, the

states and Gaussians number and the γ value for GMM, HMM and OC-SVM, respectively.

5.5.1 Offset Leakage

The results achieved by each approach, based on the proposed SFS procedure, without temporal features, and adopting the *offset* leakage typology, are presented and discussed in this section. The best AUC results, achieved for the residential consumption and, thus, based on the AMPds datasets are reported in Tables 5.3 and 5.4. In Figure 5.4 for each dataset and resolution, the AUCs achieved by the winners of each selection step of the proposed SFS process, with both GMM and HMM, but without temporal features, are reported.

Table 5.3: *Offset* leakage: AMPds, Water resource, best results without temporal features.

Resource	Res.	AUC	SD	Model	Par.	Features Combination
Water	1	81.94	10.21	GMM	128	*Ma+En*
Water	1	86.36	10.54	HMM	4–64	*Ma+En*
Water	1	54.50	16.66	OC-SVM	2^3	*dWe+dEn*
Water	10	68.70	12.83	GMM	256	*Da+dLw+Lw+dEn+dMa*
Water	10	78.55	14.98	HMM	4–256	*Lw+Da+dLw*
Water	10	62.90	14.30	OC-SVM	2^3	*dWe+dEn*
Water	30	66.94	15.29	GMM	256	*Lw+Ma+En+dMa+dLw+Da+dEn*
Water	30	74.21	17.79	HMM	4–256	*Ma+Lw+En*
Water	30	65.48	12.24	OC-SVM	2^3	*We+dEn*

In the residential scenario, the best results are achieved, for almost every resolution level and resource type, by the HMM models, although with a higher standard deviation with respect to the results of the GMM models. The only exception is the case of natural gas at 1-min time resolution, in which the GMM models outperform the HMM counterpart. The OC-SVM performance is significantly worse than the GMM and HMM ones.

Regarding the features, it can be observed that, in the case of the gas results, the moving average, *Ma*, and the energy, *En*, features are always present, in both the GMM and HMM evaluations, and they seem to be a suitable set of features for the natural gas leakage detection. The overall best results for GMM and HMM, 88.82% and 88.68%, respectively, are achieved at 1-min resolution.

In the case of water records, HMM achieved the best performance notwithstanding the time resolution, although the best overall results, 86.36%, are obtained at 1-min time resolution.

In the water case, also, longer features' combinations have been obtained, especially at 10 and 30 min of resolution, and it is difficult to identify common features' combinations.

It is noteworthy that, with both water and natural gas data, decreasing the

Table 5.4: *Offset* leakage: AMPds, Gas resource, best results without temporal features.

Resource	Res.	AUC (%)	SD	Model	Par.	Features Combination
Gas	1	88.82	9.83	GMM	256	*Ma+En*
Gas	1	88.68	9.72	HMM	1–256	*Ma+En*
Gas	1	54.97	19.27	OC-SVM	2^3	*dWe+We*
Gas	10	76.33	14.07	GMM	128	*Ma+En*
Gas	10	77.45	15.56	HMM	3–256	*Ma+En*
Gas	10	57.29	17.55	OC-SVM	2	*dWe+dLw*
Gas	30	71.75	15.03	GMM	128	*Ma+En*
Gas	30	72.76	17.13	HMM	3–256	*Ma+En*
Gas	30	60.11	15.65	OC-SVM	2	*dWe+dLw*

resolution produces worse results with both GMM and HMM models, whereas the opposite trend is shown by OC-SVM.

It also should be noted that to apply novelty detection to metering systems based on low power devices, in the natural gas case, suitable performance levels are achieved only at 1-min resolution, whereas in the water case, valuable results have been obtained at every resolution value. With respect to the computational burden, on the other hand, in the gas case, the GMM model achieves performance close to the HMM one at all resolutions; therefore, the high computational burden due to the sampling rate can be partially lowered by choosing the GMM model instead of the HMM one.

5.5.2 Offset Leakage and Temporal Features

In this section, the experiments performed by adopting FPW and TWE temporal features, and adopting the *offset* leakage typology, are presented in Table 5.5 and 5.7, and Table 5.6 and 5.8, respectively.

Table 5.5: *Offset* leakage: AMPds, Water resource, best results with FPW features.

Resource	Res.	AUC (%)	SD	Model	Par.	Features Combination
Water	1	82.00	10.23	GMM	64	*Ma+En*
Water	1	86.95	10.88	HMM	3–256	*Ma+En+Dy*
Water	1	54.88	16.77	OC-SVM	2^3	*dWe+Hr*
Water	10	70.00	16.00	GMM	256	*Ma+Wk+Dy*
Water	10	79.45	15.67	HMM	4–256	*Lw+Da+Dy+dLw*
Water	10	63.05	14.27	OC-SVM	2^3	*dWe+Hr*
Water	30	70.91	15.10	GMM	256	*Ma+Wk*
Water	30	79.12	13.18	HMM	4–256	*Da+Dy+dEn*
Water	30	65.90	12.11	OC-SVM	2^3	*dWe+Hr*

In the case of the gas results, the temporal features do not provide a boost in performance.

In the case of the water results, the introduction of the temporal information has led to a performance improvement at 10 and 30 min of resolution with both

GMM and HMM. Concerning GMM, the improvements are about 1.9% and 4.96%, for 10 and 30 min, respectively. In the case of HMM, at the same time resolutions, the respective improvements are about 1.66% and 5.86%.

Table 5.6: *Offset* leakage: AMPds, Water resource, best results with TWE features.

Resource	Res.	AUC (%)	SD	Model	Par.	Features Combination
Water	1	81.94	10.07	GMM	128	*Ma+En*
Water	1	86.55	10.51	HMM	4–64	*Ma+En*
Water	1	54.50	16.66	OC-SVM	2^3	*dWe+dEn*
Water	10	70.60	16.13	GMM	256	*Ma+Hr+dMa*
Water	10	80.21	15.13	HMM	4–256	*Lw+Da+Hr+dLw+Wk+Ma+dEn*
Water	10	62.90	14.30	OC-SVM	2^3	*dWe+dEn*
Water	30	71.90	13.91	GMM	256	*Da+Hr+dEn+dMa+En+Ma*
Water	30	80.07	13.07	HMM	4–256	*Da+Hr*
Water	30	65.48	12.24	OC-SVM	2^3	*dWe+dEn*

Moreover, in the GMM case, the only temporal feature is the one obtained with the hourly window, *Hr*. In the case of HMM, at 10 min, two temporal features are present, hourly and weekly, whereas at 30 min, only the hourly one is present.

As noted in the the previous experiments, with both water and natural gas data, decreasing the resolution produces worse results with both GMM and HMM models, whereas the opposite trend is shown by OC-SVM.

Table 5.7: *Offset* leakage: AMPds, Gas resource, best results with FPW features.

Resource	Res.	AUC (%)	SD	Model	Par.	Features Combination
Gas	1	88.68	9.79	GMM	256	*Ma+En*
Gas	1	88.62	9.47	HMM	3–256	*Ma+En*
Gas	1	56.93	16.84	OC-SVM	2^3	*dWe+Hr*
Gas	10	76.26	14.64	GMM	256	*Ma+En*
Gas	10	77.39	15.72	HMM	3–256	*Ma+En*
Gas	10	58.36	16.27	OC-SVM	2^3	*dWe+Hr*
Gas	30	71.79	15.29	GMM	128	*En+Ma*
Gas	30	73.11	16.98	HMM	3–256	*Ma+En*
Gas	30	60.12	15.57	OC-SVM	2	*dWe+dLw+Hr*

Concerning the application of the novelty detection to metering systems based on low power devices, in the water case, the temporal features at 30-min resolution allow one to further reduce the computational burden of the system. Specifically, the features' vector *Ma+Lw+En* turns into *Da+Hr*, thus losing the heavy computation required by the wavelet decomposition. Thus, if the aim is to implement the algorithm on a concentrator or similar low power computation devices, there is room to optimize the procedure with little to no loss in terms of performance.

Table 5.8: *Offset* leakage: AMPds, Gas resource, best results without TWE features.

Resource	Res.	AUC (%)	SD	Model	Par.	Features Combination
Gas	1	88.78	9.84	GMM	256	*Ma+En*
Gas	1	88.85	9.69	HMM	3–256	*Ma+En*
Gas	1	54.97	19.27	OC-SVM	2^3	*dWe+We*
Gas	10	76.35	14.48	GMM	256	*Ma+En*
Gas	10	77.50	15.54	HMM	3–256	*Ma+En*
Gas	10	57.29	17.55	OC-SVM	2	*dWe+dLw*
Gas	30	71.83	15.49	GMM	256	*En+Ma*
Gas	30	72.83	17.19	HMM	3–256	*Ma+En*
Gas	30	60.11	16.65	OC-SVM	2	*dWe+dLw*

5.5.3 Simulated Leakage

In this section, the experiments performed without adopting the temporal features, and adopting the *simulated* leakage typology, are presented in Table 5.9. When jointly addressing the features extracted from flow and pressure data, the flow based feature will include an additional **F**, and the pressure based features will include a **P**.

First of all, in Table 5.9 the results achieved without the application of the temporal features are reported. Generally speaking, the HMM confirmed to be the best approach, followed by the GMM. With the exception of the results achieved operating on the pressure data, the OC-SVM seems to be unsuitable for the leakage detection approach. The results based on flow data present a reverse trend respect to the pressure ones, increasing the resolution achieves a lower performance.

5.5.4 Simulated Leakage and Temporal Features

In this section, the experiments performed by adopting FPW and TWE temporal features, and adopting the *simulated* leakage typology, are presented in Table 5.10 and 5.11, respectively. When jointly addressing the features extracted from flow and pressure data, the flow based feature will include an additional **F**, the pressure based features will include a **P**.

Concerning the FPW features in Table 5.10, as for the previous experiment, all the best performance are achieved with HMM, followed by GMM, with the exception of those based on pressure data at 30 minutes resolution where OC-SVM perform slightly better that HMM. Even if the pressure-based results show the lowest performance, with a very high standard deviation, they present a reverse trend with respect to flow-based ones. Decreasing the resolution produces lower results for the flow data, and better results, instead, for the pressure ones. Moreover, by combining the features extracted from flow and pressure, the trend exhibited by the flow is mitigated, and reducing the

Table 5.9: *Simulated* leakage: AMPds, Water resource, best results without temporal features.

Res.	AUC (%)	STD	Model	Par.	Features comb.
Pressure Data					
1	54.90	18.22	GMM	2	*dEn+Ma+dMa*
1	**54.91**	18.23	HMM	1-2	*dEn+Ma+dMa*
1	52.50	16.22	OC-SVM	2^{-5}	*dDa+We*
10	54.85	19.79	GMM	2	*dEn+dMa+Ma+dLw*
10	**55.34**	23.20	HMM	4-128	*We+En*
10	53.21	22.06	OC-SVM	2−15	*Da+We+dLw+Lw+En*
30	54.76	19.91	GMM	2	*En+dEn+dMa*
30	**55.78**	20.40	HMM	3-64	*Ma+En*
30	54.17	24.08	OC-SVM	2^{3}	*Da+We*
Flow Data					
1	81.03	10.51	GMM	128	*Ma+En*
1	**85.89**	11.39	HMM	4-64	*Ma+En+Lw*
1	60.43	19.02	OC-SVM	2^{-7}	*Ma+We*
10	66.99	12.45	GMM	256	*Da+dLw*
10	**78.55**	15.01	HMM	4-256	*Da+Lw+dLw*
10	65.55	27.76	OC-SVM	2^{-9}	*Da+Ma*
30	63.80	13.18	GMM	256	*Lw+Ma*
30	**73.69**	17.89	HMM	4-256	*Ma+Lw*
30	68.39	20.86	OC-SVM	2^{3}	*Da+dWe*
Flow&Pressure Data					
1	86.03	8.40	GMM	32	*MaF+EnP+MaP*
1	**87.87**	9.35	HMM	3-256	*MaF+EnF+EnP*
1	60.43	19.02	OC-SVM	2^{-7}	*MaF+WeF*
10	86.12	8.40	GMM	32	*MaF+MaP+EnP+dMaP*
10	**87.47**	8.81	HMM	3-64	*MaF+EnP+MaP*
10	65.55	27.26	OC-SVM	2^{-9}	*DaF+MaF*
30	86.34	9.14	GMM	32	*MaF+EnP+MaP+dEnP*
30	**88.51**	9.20	HMM	3-64	*MaF+EnP+MaP*
30	68.39	20.86	OC-SVM	2^{3}	*DaF+dWeF*

resolution produces slightly better results. With respect to experiments without the temporal features, the introduction of the FPW features produces a slight increase in the pressure data performance, but high standard deviations are obtained. Even in the case of flow data, the temporal features produce a slight improvement of the results at 1 and 10 minutes for HMM, and at 10 and 30 minutes for GMM. Unfortunately, the temporal features do not entail performance improvement in the combinations of flow and pressure features.

About the experiment with TWE features, the achieved results are reported in Table 5.11, and the same trends shown for FPW features are confirmed. Regarding the pressure data, the exploitation of TWE features produces a significant improvement with respect to the FPW one for both GMM and HMM. On the other hand, when based on the flow data, the TWE features at 1 minute resolution do not produce better performance, instead at 10 and 30 minutes of resolution, the TWE temporal features allow to reach better results than the FPW ones. Finally, about the combination of flow and pressure data,

Table 5.10: *Simulated* leakage: AMPds, Water resource, best results with FPW features.

Res.	AUC (%)	STD	Model	Par.	Features comb.
Pressure Data					
1	56.14	21.91	GMM	2	$dLw+Wk+Dy+dMa+dEn$
1	**56.21**	24.83	HMM	4-4	$Ma+Wk+dWe$
1	55.70	26.73	OC-SVM	2	$dMa+Dy+Wk+dEn$
10	56.44	19.07	GMM	2	$dEn+Dy+Wk+Ma+Hr+En$
10	56.78	24.92	HMM	4-32	$dWe+Wk+We$
10	**57.28**	28.28	OC-SVM	2^{-1}	$dDa+Wk$
30	56.33	19.74	GMM	4	$Lw+Wk+Ma$
30	**57.44**	25.49	HMM	3-8	$Lw+Dy+Wk+En+We$
30	57.25	28.29	OC-SVM	2^{-5}	$Lw+Wk$
Flow Data					
1	80.88	10.56	GMM	128	$Ma+En$
1	**86.14**	11.18	HMM	3-128	$Ma+En+Dy$
1	60.44	19.06	OC-SVM	2^{-7}	$Ma+We+En$
10	70.11	15.08	GMM	256	$Ma+Dy+Wk$
10	**78.56**	15.50	HMM	4-256	$Lw+Da+dLw+Dy$
10	65.55	27.26	OC-SVM	2^{-7}	$Ma+Da$
30	69.98	15.46	GMM	256	$Ma+Wk$
30	**78.72**	12.78	HMM	4-256	$Da+Dy$
30	67.52	20.75	OC-SVM	2^{3}	$Da+dWe$
Flow&Pressure Data					
1	85.95	8.62	GMM	32	$MaF+EnP+MaP+dMaP+dEnP$
1	**87.70**	9.54	HMM	3-256	$MaF+EnF+MaP$
1	61.29	18.75	OC-SVM	2^{-15}	$WeF+dLwP$
10	86.25	8.55	GMM	32	$MaF+EnP+MaP+dEnP+dMaP$
10	**87.64**	8.94	HMM	3-64	$MaF+EnP+MaP$
10	65.59	27.23	OC-SVM	2^{-7}	$MaF+DaF+dMaP$
30	86.61	8.86	GMM	32	$MaF+MaP+EnP+dEnP$
30	**88.06**	9.35	HMM	3-64	$MaF+EnP+MaP$
30	67.52	20.75	OC-SVM	2^{3}	$DaF+dWeF$

even the TWE features do not produce positive contribution, achieving a result similar to the previous experiments.

The overall best result for pressure data, 63.64%, is achieved with HMM at 30 minutes of resolution, by using as features Lw, Wk (TWE), and Ma. The introduction of the TWE temporal features produced a performance improvement of 6.57%, 5.75%, and 7.86% at 1, 10, and 30 minutes of resolution, respectively, for HMM.

For the flow data, the best results is 86.14% for the combination of Ma, En, and Dy (FPW) with HMM at 1 minute resolution, with an improvement of 0.25%. Whereas, for 10 and 30 minutes, the introduction of the TWE temporal features lead to a performance improvement of 0.85% and 5.95%, respectively, for HMM.

Finally, the combination of flow and pressure information obtained a relevant performance improvement, in terms of AUC increment and standard deviation reduction, that even the introduction of any temporal features fails to produce

Table 5.11: *Simulated* leakage: AMPds, Water resource, best results with TWE features.

Res.	AUC (%)	STD	Model	Par.	Features comb.
Pressure Data					
1	59.78	23.97	GMM	2	*dEn+Wk+dMa+Ma+Dy+Hr*
1	**61.48**	24.32	HMM	3-32	*dEn+Wk+dMa*
1	56.88	27.58	OC-SVM	2^{-9}	*dEn+Dy*
10	59.27	20.61	GMM	2	*dEn+Wk+dWe*
10	**61.09**	24.10	HMM	3-32	*dEn+Wk*
10	57.46	27.02	OC-SVM	2^{-3}	*Da+Dy*
30	58.93	21.13	GMM	64	*dDa+Wk+Dy+dMa*
30	**63.64**	21.88	HMM	4-16	*Lw+Wk+Ma*
30	56.78	26.47	OC-SVM	2^{-13}	*dDa+Dy*
Flow Data					
1	80.83	10.43	GMM	64	*Ma+En*
1	**85.54**	11.74	HMM	4-128	*Ma+En+Lw*
1	60.43	19.02	OC-SVM	2^{-7}	*Ma+We*
10	69.82	16.48	GMM	256	*Ma+Hr+dMa*
10	**79.40**	15.11	HMM	4-256	*Da+Lw+Hr+dLw+Wk+Ma*
10	65.55	27.26	OC-SVM	2^{-9}	*Da+Ma*
30	71.94	14.44	GMM	256	*Da+Hr+dMa+Ma*
30	**79.64**	14.85	HMM	4-256	*Da+Hr+Wk*
30	68.39	20.86	OC-SVM	2^{3}	*Da+dWe*
Flow&Pressure Data					
1	86.04	8.80	GMM	32	*MaF+MaP+EnP+dMaP*
1	**87.73**	9.34	HMM	3-64	*MaF+EnF+EnP+MaP*
1	60.43	19.02	OC-SVM	2^{-7}	*MaF+WeF*
10	86.26	8.58	GMM	32	*MaF+MaP+EnP+dEnP+dMaP*
10	**87.29**	9.36	HMM	3-128	*MaF+EnP+MaP*
10	65.55	27.26	OC-SVM	2^{-9}	*DaF+MaF*
30	86.68	9.07	GMM	32	*MaF+EnP+MaP+dMaP*
30	**87.34**	8.90	HMM	1-128	*MaF+EnP+MaP*
30	68.39	20.86	OC-SVM	2^{3}	*DaF+dWeF*

a further improvement, as it can be concluded by comparing the results reported in Table 5.9, 5.10, and 5.11. The overall best result reaches an AUC of 88.51% with HMM at 30 minutes of resolution by using MaF, EnP, and MaP. Specifically, with respect to the best results achieved with flow features and temporal information, the improvement is of about 2%, 9%, and 15% at 1, 10, and 30 minutes of resolution, respectively. The combination of flow and pressure data reduced considerably the performance gap between GMM and HMM, the average difference is of 1.5%, improving the usability of the GMM, that presents lower computational burden than HMM.

Among all the experiments, the best results achieved with the only pressure data, often involve the Da and/or dDa features, that have the greatest dimensions, Table 5.1. Therefore, even omitting the achieved performance, the computational burden for these combinations is certainly higher than the one of other combinations. Furthermore, except for some OC-SVM experiments, the test performed with the flow and pressure data, that reach very good per-

formance, does not present the *Da* features.

In the flow and pressure data experiments, the randomness of the trained models, together with the achievement of very close performance among many feature combinations, leaded to the selection of slight different best combinations for same experimental conditions. Specifically, for GMM at 10 minute resolution, without temporal features, the combination $MaF + MaP + EnP + dMaP + dEnP$ (selected in in both FPW and TWE evaluations) reached 86.09%. For GMM at 30 minutes, with TWE temporal features, the combination $MaF + EnP + MaP + dEnP$ reached 86.66%. Similar consideration apply to GMM and HMM at 1 minute resolution.

5.6 Experiments: Office Building Scenario

In the following tables, the best results and the corresponding features' combination achieved for each resource and resolution (**Res.**) of the DFID datasets, are reported. The column "**Par.**" reports the number of Gaussians, the states and Gaussians number and the γ value for GMM, HMM and One-Class (OC)-SVM, respectively. The Abercrombie and the Whitehall sets are reported as "Aber." and "White.", respectively.

5.6.1 Offset Leakage

The results achieved by each approach, based on the proposed SFS procedure, without temporal features, and adopting the *offset* leakage typology, are presented and discussed in this section. The best AUC results, achieved for the building consumption, based on the DFID dataset, are reported in Table 5.12. In Figure 5.5, for each dataset and resolution, the AUCs achieved by the winners of each selection step of the proposed SFS process, with both GMM and HMM, but without temporal features, are reported.

Table 5.12: *Offset* leakage: DFID best results without temporal features.

Resource	Res.	AUC (%)	SD	Model	Par.	Features Combination
Aber. Gas	30	83.61	9.30	GMM	256	*dLw+Lw*
Aber. Gas	30	87.59	10.15	HMM	4–64	*dLw+Lw*
Aber. Gas	30	75.30	25.91	OC-SVM	2^{-7}	*Da+dMa+Ma+En*
Aber. Water	30	87.77	5.93	GMM	256	*Da+We+dWe+dMa+dEn*
Aber. Water	30	92.52	5.58	HMM	3–256	*Ma+Da+dMa+En+dEn*
Aber. Water	30	67.06	33.50	OC-SVM	2^{-7}	*Da+Ma+En*
Overall Gas	30	85.47	10.85	GMM	256	*Lw+dLw+dMa*
Overall Gas	30	88.85	11.22	HMM	4–256	*dLw+Lw+Ma*
Overall Gas	30	79.56	23.95	OC-SVM	2^{-3}	*Da+dDa+dMa+dEn*
White. Water	30	91.21	5.83	GMM	256	*Da+We+dMa+Ma+dEn*
White. Water	30	92.88	4.53	HMM	4–128	*Da+We+dEn+Ma*
White. Water	30	65.39	39.03	OC-SVM	2^{-5}	*Da+Ma*

In the case of the building scenario, the best performance is achieved by the HMM model notwithstanding the time resolution of the samples. OC-SVM in this case, as well, results in being the worst performer.

About the Abercrombie gas subset, the best result, 87.59%, is achieved by the HMM model, by using wavelet fetaures *Lw* and *dLw*.

Concerning the Abercrombie water set, the HMM reaches the best result. Specifically, an AUC of 92.52% is achieved by adopting *Ma+Da+ dMa+En+dEn* features.

The best result performed with the overall gas set is 88.85%, by adopting *dLw+Lw+Ma* features.

In the case of the Whitehall set, GMM and HMM achieve very high and close performance, 91.21% nad 92.88%, respectively.

The results of the experiments show that the proposed solution is compatible with low power metering devices, and the data precessing may also be suitable for low power computational devices.

5.6.2 Offset Leakage and Temporal Features

In this section, the experiments performed by adopting FPW and TWE temporal features, and adopting the *offset* leakage typology, are presented in Table 5.13 and 5.14, respectively.

Table 5.13: *Offset* leakage: DFID best results with FPW temporal features.

Resource	Res.	AUC (%)	SD	Model	Par.	Features Combination
Aber. Gas	30	84.64	10.31	GMM	256	*Lw+dLw+Dy+dMa+Ma*
Aber. Gas	30	87.86	10.12	HMM	4-256	*Lw+dLw+Wk*
Aber. Gas	30	75.35	25.91	OC-SVM	2^{-7}	*Da+dMa+Hr+Ma+En*
Aber. Water	30	89.91	7.61	GMM	256	*Da+Dy+We+dWe+Ma+Hr*
Aber. Water	30	94.76	5.65	HMM	4–256	*Da+Dy+Ma+dMa*
Aber. Water	30	67.06	33.56	OC-SVM	2^{-7}	*Da+Ma+En*
Overall Gas	30	86.64	11.81	GMM	256	*Lw+dLw+dMa+Dy+Wk+Ma*
Overall Gas	30	89.12	11.89	HMM	4–256	*Lw+Wk+dLw+Ma+Dy*
Overall Gas	30	79.58	23.96	OC-SVM	2^{-7}	*Da+dDa+dMa+dEn+Hr*
White. Water	30	93.42	5.89	GMM	256	*Da+Dy+Hr+Wk*
White. Water	30	95.41	5.40	HMM	2–256	*Da+Wk+Dy+dEn*
White. Water	30	65.39	39.03	OC-SVM	2^{-5}	*Da+Ma*

The introduction of temporal features has provided additional performance improvement to both the GMM and HMM models, whereas OC-SVM has not shown significant gain.

About the Abercrombie gas subset, the best result, 87.80%, is achieved by the HMM model, introducing the hourly and weekly FPW temporal features, with a performance improvement of about 0.21%. On the contrary, in the case of GMM, the introduction of the hourly TWE temporal features produces a greater performance improvement, about 1.39%, thus, allowing to reach an overall AUC of 85.00%.

Concerning the Abercrombie water set, in the case of GMM, 2.30% of improvements is obtained by introducing TWE temporal features. Whereas, an improvement of 2.24% is reached in case of HMM by adopting FPW features. Specifically, the HMM best result, 94.76%, is achieved by adopting the daily temporal feature.

The best result performed with the overall gas set is 89.56%, by introducing both the daily and the weekly TWE temporal features, and obtaining enhancements of 0.71% for HMM.

In the case of the Whitehall set, the introduction of temporal features produces an improvement of 2.21% and 2.53, for GMM and HMM, respectively, in the case of FPW features. Whereas, an improvement of 2.44% for both GMM and HMM, in the case of TWE features, is achieved. The best result is 95.41%, and the improvements are achieved by introducing the hourly and the daily features for HMM.

As stated above, the results of the experiments show that the proposed solution is compatible with low power metering devices, and the data precessing may also be suitable for low power computational devices. Specifically, for some scenarios, the application of temporal features, together with the general performance improvement, allows one to reduce the computational burden by effecting the features' combinations. In the case of HMM model, the features' set obtained for the Abercrombie water data is reduced from *Ma+Da+dMa+En+dEn* to *Da+Dy+Ma+dMa*; thus from five to four features and from 14 to 13 overall components. Whereas, for the White water set, the features go from *Da+We+dEn+Ma* to *Da+Wk+Dy+dEn*, thus removing the need for the heavy computational power required for the wavelet decomposition. On the other hand, in the case of the GMM model, for the White water data, the features set *Da+We+dMa+Ma+dEn* is reduced to *Da+Hr+Dy*, going from five to three features and from 14 to 12 overall components.

Table 5.14: *Offset* leakage: DFID best results with TWE temporal features.

Resource	Res.	AUC (%)	SD	Model	Par.	Features Combination
Aber. Gas	30	85.00	10.29	GMM	256	*Lw+dLw+Hr+Ma*
Aber. Gas	30	87.80	10.55	HMM	4-256	*dLw+Lw+Hr+Wk+Ma*
Aber. Gas	30	75.35	25.91	OC-SVM	2^{-7}	*Da+dMa+Ma+En*
Aber. Water	30	90.07	6.64	GMM	256	*Da+Hr+We+dWe+Ma+dEn+dMa*
Aber. Water	30	94.73	5.62	HMM	4–256	*Da+Hr+Ma+dMa*
Aber. Water	30	67.06	33.56	OC-SVM	2^{-7}	*Da+Ma+En*
Overall Gas	30	86.60	10.63	GMM	256	*dLw+Lw+Hr+dMa+Ma*
Overall Gas	30	89.56	12.04	HMM	4–32	*Lw+dLw+dMa+Wk+Dy+Ma*
Overall Gas	30	79.56	23.95	OC-SVM	2^{-7}	*Da+dDa+dMa+dEn*
White. Water	30	93.65	6.74	GMM	256	*Da+Hr+Dy*
White. Water	30	95.32	5.67	HMM	2–256	*Da+Hr+Dy+dEn*
White. Water	30	65.39	39.03	OC-SVM	2^{-5}	*Da+Ma*

5.7 Further Evaluations

5.7.1 Best Results Details

In order to have a better insight into the detection performance for the best results achieved adopting the *offset* leakage typology, in Table 5.15, together with the best results achieved (column "AVG"), are also reported the best AUCs achieved among the background models (column "BEST BG"), and the overall best AUC (column "BEST"). The former refers to the best result achieved as the average of the AUC results, obtained for all of the tested leakages with the same background model. The latter shows the best AUC achieved among the detected leakages, regardless of the adopted background model. As discussed in Section 5.3.3, for each experiment, 10 background models are trained, and for each of them, 10 random leakages are tested. Therefore, the background number, for the "BEST BG", and the background and the leakage number, for the "BEST", that have achieved the reported result, are indicated in brackets. The ROCs concerning the average AUCs achieved for each background model, among all of the tested leakages and the overall averages, are depicted in Figure 5.8.

Table 5.15: Best results' details. TDR, True Detection Rate; FDR, False Detection Rate; BG, Background. The Whitehall set is reported as "White.".

Resource	AUC (%)			TDR (%)	FDR (%)		
	AVG	BEST BG	BEST		AVG	BEST BG	BEST
AMPds Gas	88.82	$93.27_{(6)}$	$99.84_{(5/3)}$	100	25.32	14.65	0.26
AMPds Water	86.36	$91.46_{(8)}$	$99.76_{(7/6)}$	100	29.43	22.33	0.57
Overall Gas	89.56	$94.68_{(8)}$	$99.99_{(3/5)}$	100	18.17	10.26	0.02
White. Water	95.32	$96.97_{(8)}$	$99.97_{(10/9)}$	100	9.36	6.14	0.06

Furthermore, also the TDRs and FDRs are reported in Table 5.15. The values are obtained from the corresponding ROCs. Each FDR has been computed for the lowest likelihood value that achieves a TDR of 100%. Note that the FDRs reported in Table 5.15 cannot be compared to the points for the TDRs of 100% in the curves depicted in Figure 5.9. Differently from the ROCs, where the average curves have been computed considering both TDR and FDR points, the results in Table 5.15 have been achieved performing the average of the FDR components only.

5.7.2 Performance Variability Analysis

In order to have a better insight into the system performance, in particular of the standard deviation shown in the results achieved adopting the *offset* leak-

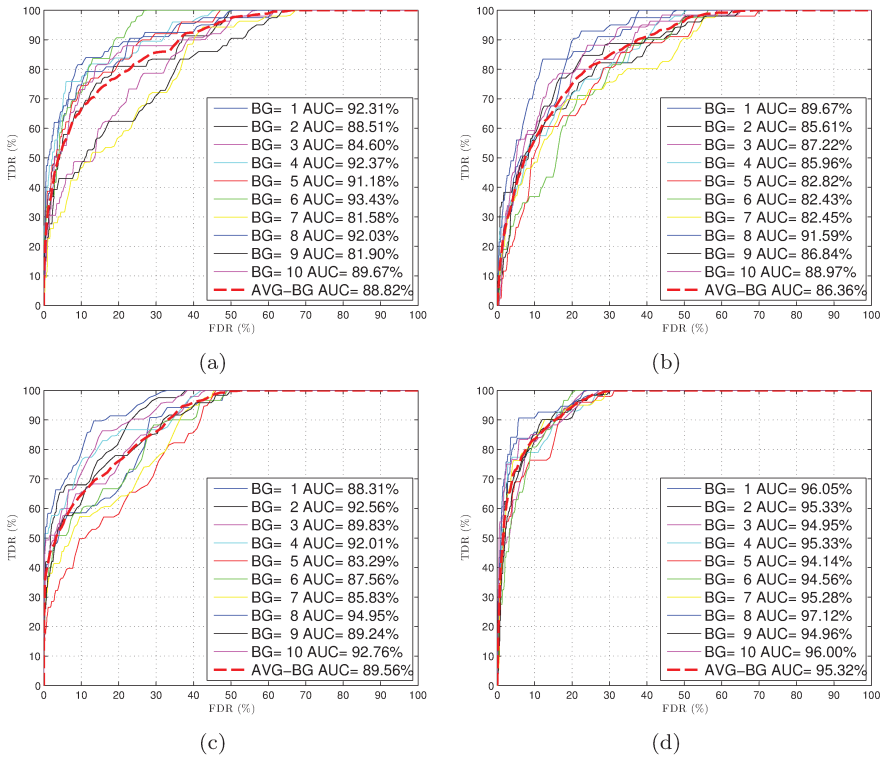


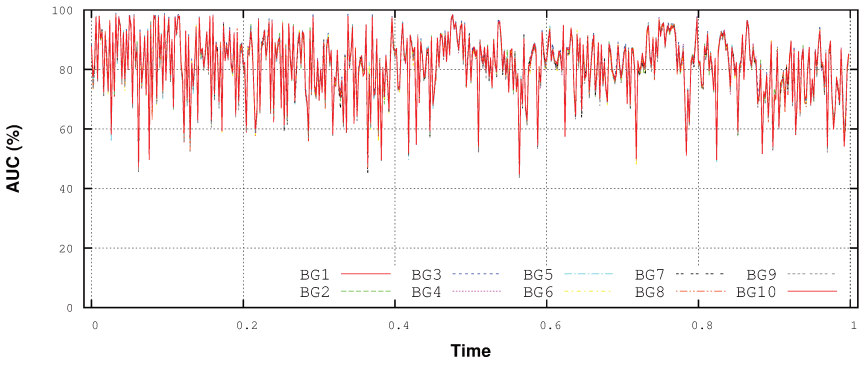
Figure 5.9: ROCs for the best achieved results with each background model. The red dotted line represents the overall mean curve. (**a**) AMPds gas; (**b**) AMPds water; (**c**) overall gas; (**d**) Whitehall House water.

age typology, a sensitivity analysis has been performed. As for the experiments discussed above, 10 leakages for 10 different background models have been considered, thus 100 evaluations for each condition. The tests have been performed for the higher resolution, 1 min, on both water and natural gas consumption of the AMPds, by using the best features' combinations. Moreover, for each dataset and technique, the evaluations have been performed three times, and each time, two out of three leakage parameters have been kept constant, in order to perform the analysis of the leakage starting point, size and duration.

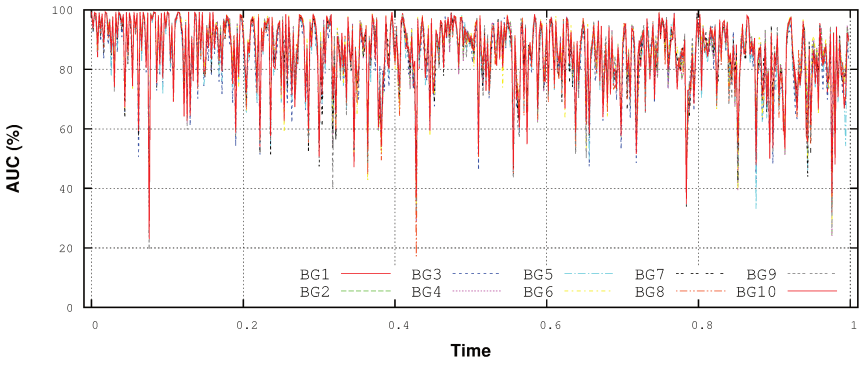
For the starting point sensitivity analysis, the leakage size is set to $\beta = 0.375$, and the leakage duration is fixed to 450 samples. For the leakage size test, the starting point is kept to half of the set length, and the β has been varied from 0.25 to 0.50, as for the previous experiments. Finally, for the leakage duration sensitivity analysis, the leakage duration has been varied from 300 to 600 samples. The results, depicted in Table 5.16, show that the bottleneck is due to the leakage position in the set, and a strong correlation between the leakage

Table 5.16: Performance of the variability tests.

Resource	Model	Leakage Starting Point		Leakage Size		Leakage Duration	
		AUC (%)	SD	AUC (%)	SD	AUC (%)	SD
Gas	GMM	89.09	9.48	87.88	3.39	88.57	1.41
Gas	HMM	89.10	9.67	87.49	3.64	88.20	1.44
Gas	OC-SVM	54.75	20.07	69.20	0.16	70.75	1.76
Water	GMM	83.36	8.33	90.13	7.11	89.68	2.15
Water	HMM	89.90	9.58	94.46	4.88	94.84	2.85
Water	OC-SVM	54.99	16.10	50.09	0.16	53.17	2.17



(a) GMM



(b) HMM

Figure 5.10: AUCs achieved by GMM and HMM backgrounds in the starting point analysis. (**a**) GMM; (**b**) HMM.

position and the achieved performance is revealed, as shown in Figure 5.10. In particular, for each background model, the AUCs evaluated for 500 leakage positions, uniformly distributed along the dataset length, are reported. All of the backgrounds present the same trend, and this confirms, for both GMM and

HMM, that the standard deviation does not depend on the variability of the initial conditions in the training procedure, but it is entirely due to the leakage position in the dataset instead; thus, the backgrounds present the same envelope.

The results achieved by means of these evaluations, together with the evaluations performed in Section 5.7.5, have allowed to understand the nature of the high variability of the system results. Therefore, it has been deemed necessary to evaluate the introduction of temporal information in the automatic leakage detection algorithm, as presented in Section 5.2.1, in order to attempt to mitigate the variability effect due to the time position of the leakage.

5.7.3 Cross-Validation Results

Table 5.17: *Offset* leakage: AMPds, Natural Gas and Water resources best results with k-fold cross-validation.

Res.	AUC (%)	Model	Par.	Feat. comb.	TDRs / FDRs			
Natural Gas								
1	91.47	GMM	256	*Ma+En*	67.29 4.71	74.79 8.98	77.92 14.51	82.92 18.7
1	92.27	HMM	3-256	*Ma+En*	68.96 4.65	78.13 5.96	64.38 9.02	64.38 9.78
10	68.91	GMM	128	*Ma+En*	74.38 39.21	87.71 54.43	75.21 39.25	91.25 51.53
10	68.96	HMM	1-128	*Ma+En*	70.21 41.13	87.71 56.54	62.92 36.64	91.25 49.03
30	65.14	GMM	128	*Ma+En*	61.67 42.18	86.25 54.68	66.88 37.44	86.67 0.22
30	65.16	HMM	1-128	*Ma+En*	75.00 45.14	86.25 56.28	65.83 39.26	88.75 51.24
Water								
1	83.89	GMM	128	*Ma+En*	59.79 16.35	71.46 29.35	63.13 20.35	69.38 31.37
1	86.59	HMM	3-256	*Ma+En+Dy* (FPW)	71.46 28.24	83.75 35.32	45.63 27.69	53.75 31.6
10	65.77	GMM	256	*Ma+Hr+dMa* (TWE)	74.38 59.66	91.88 78.44	41.88 37.28	56.88 47.89
10	77.23	HMM	4-256	*Lw+Da+Hr+dLw+Wk+* *Ma+dEn* (TWE)	33.75 14.84	58.75 22.15	41.25 29.15	46.88 32.48
30	70.26	GMM	256	*Da+Hr+dEn+dMa+* *En+Ma* (TWE)	42.29 41.19	66.88 54.38	45 40.92	64.17 51.41
30	78.36	HMM	4-256	*Da+Hr* (TWE)	44.17 35.82	62.08 44.8	27.92 28.61	41.25 33.64

In order to validate the best results and to provide a hint of the general performance, the three-way data split approach [93] combined with the K-fold cross-validation method [87], has been executed. Each dataset has been split into $K = 8$ folds, and for each parameters combination, two different folds are chosen iteratively as validation and test sets. The remaining $K - 2$ folds form

the training set, and the AUC has been determined on the validation fold. The final system performance, expressed in terms of TDR and FDR, has been evaluated by using one fold as test, and the remaining $K-1$ folds as training. All the evaluations have been computed on the best features combination achieved in the experiments reported in Section 5.5 and Section 5.6.

The experiments have been performed for both GMM and HMM models, and the test performance have been evaluated by adopting two sets of threshold types, each composed of two different thresholds. With reference to the ROC representation, the former set is composed of the thresholds that allow to achieve the lowest L^2 distance between the actual TDR-FDR values and the point with TDR 100% and FDR 0%. On the other hand, the latter set is composed of the first thresholds that achieve a TDR of 100%, regardless of the FDR value. Regarding the two different thresholds of each set, the first one is the threshold "as is" obtained by the log-likelihood returned by the model. The second one, instead, is the normalized threshold with respect to the maximum and minimum log-likelihood values of the validation set.

Moreover, for each parameters combination, the reported results express the AUC and the TDR/FDR values achieved for all the validation folds and for all the test folds, respectively, and averaged over each three-way data split.

The results achieved for the residential scenario, adopting the *offset* leakage typology, are reported in Table 5.17, whereas 5.18 reports the results for the *simulated* leakage typology. The results achieved for the office building scenario, thus *offset* leakage typology, are reported in Table 5.19.

Table 5.18: *Simulated* leakage: AMPds, Water synthetic resource best results with k-fold cross-validation.

Res.	AUC (%)	Model	Par.	Feat. comb.	TDRs / FDRs			
Water								
1	86.38	GMM	32	*MaF+EnP+MaP*	59.79 12.26	72.29 22.79	61.46 17.1	72.92 25.93
1	86.88	HMM	2-256	*MaF+EnF+EnP*	65.00 14.22	70.00 22.03	53.33 22.10	55.42 26.27
10	86.57	GMM	32	*MaF+MaP+EnP+dMaP*	81.25 13.57	89.38 22.87	34.38 12.68	34.38 14.26
10	87.27	HMM	1-128	*MaF+EnP+MaP*	78.75 12.08	91.88 22.65	75.63 12.12	86.25 20.69
30	87.29	GMM	32	*MaF+EnP+MaP+dEnP*	79.38 11.58	85.00 19.36	37.50 15.12	37.50 16.77
30	87.97	HMM	1-128	*MaF+EnP+MaP*	82.50 11.76	85.00 20.51	80.00 13.74	84.58 21.78

In the case of residential scenario, without pressure information, the validation AUC presents the same trend of the previous evaluations, Section 5.5: increasing the resolution achieves a lower performance. Moreover, for both water and natural gas, the HMM model achieves slightly better performance

than the GMM in terms of both AUC and TDR-FDR values. Specifically, with respect to the TDR and FDR values, the performance degradation affects the difference between TDR and FDR values, therefore, either the TDR values are decreased or the FDR values are increased.

The cross-validation tests with the introduction of the simulated flow and pressure data (*simulated* leakage typology), confirms the performance achieved in the previous experiments, Section 5.5.4. Increase the resolution no longer affects the system performance, and even switching from the GMM model to the HMM one, produces a negligible performance improvement.

Table 5.19: *Offset* leakage: DFID best results with k-fold cross-validation.

Res.	AUC (%)	Model	Par.	Feat. comb.	TDRs / FDRs			
Abercrombie House - Natural Gas								
30	83.81	GMM	256	*Lw+dLw+Hr+Ma* (TWE)	78.13 35.80	91.88 47.63	68.13 37.42	76.25 42.11
30	84.93	HMM	4-256	*Lw+dLw+Wk* (FPW)	38.33 15.59	58.33 20.45	67.5 20.94	67.5 25.12
Abercrombie House - Water								
30	88.93	GMM	256	*Da+Hr+We+dWe+Ma+ dEn+dMa* (TWE)	81.04 20.12	89.38 26.65	86.67 25.82	92.29 30.32
30	94.96	HMM	4-256	*Da+Dy+Da+dMa* (FPW)	93.33 19.80	95.42 21.66	97.92 24.20	97.92 25.61
Overall DFID - Natural Gas								
30	81.94	GMM	256	*Lw+dLw+dDa+Dy+Wk+ Ma* (FPW)	68.33 36.41	81.67 48.56	77.71 43.78	82.71 52.10
30	84.98	HMM	4-64	*Lw+dLw+dDa+Wk+Dy+ Ma* (TWE)	53.13 24.56	56.25 28.60	57.50 23.98	66.88 27.42
Whitehall - Water								
30	90.72	GMM	256	*Da+Hr+Dy* (TWE)	68.75 25.76	70.83 27.76	75.00 36.21	80.00 37.48
30	93.76	HMM	3-256	*Da+Hr+Dy+dEn* (TWE)	73.13 21.31	73.13 22.38	87.50 36.77	87.5 37.28

In the office building scenario, the high performance achieved in the previous experiments are confirmed, Section 5.6. The HMM model allows to achieve better results with all the datasets, in particular, lower FDR values are reached with respect to the one achieved with the GMM model.

5.7.4 Autoencoder Neural Network

In order to explore a fast growing computational intelligence approach, preliminary evaluations have been also performed adopting the Autoencoder Neural Network (AE). The evaluation algorithm has been adapted as stated in Section 5.1.1, and different parameters combinations have been evaluated. Specifically, the number of hidden layer, and their dimensions, have been varied by assuming the tanh as activation function with a constant learning rate of $1e-3$ and 2000 training epochs.

The evaluations have been conducted by adopting the best features combination achieved in the above Sections 5.5 and 5.6, without temporal features and for the *offset* leakage typology. The best results achieved, and the corresponding parameters, for the residential scenario are reported in Table 5.20, whereas the results for the office building scenario are reported in Table 5.21.

In the Tables, the column "Param." reports the multiplier to apply to each hidden layer. In this regard, the number of neurons for a specific hidden layer, is defined as the multiplication of the neurons number of the input layer (given by the feature size) and the hidden layer multiplier value. For example, let "2-1" be the parameters reported for a network, and let *Lw* be the feature adopted, then the reference network has two hidden layers with 8 and 4 neurons each, respectively, together with omitted input and output layers of size 4.

Table 5.20: AMPds, Natural Gas and Water resource best results with AE.

Res.	AUC (%)	STD	Features comb.	Param.
Natural Gas				
1	83.38	13.71	*Ma+En*	2-1
10	69.98	13.05	*Ma+En*	2-1
30	66.92	13.17	*Ma+En*	2-1
Water				
1	74.18	18.80	*Ma+En*	2-1
10	70.45	14.23	*Da+dLw+Lw+dEn+dMa*	2-2
30	63.29	11.46	*Ma+Lw+En*	2-2

In the case of residential scenario, the results achieved are lower than the one achieved with GMM and HMM models. On the other hand, with 10 and 30 minutes of resolution, the results show a slightly lower standard deviation.

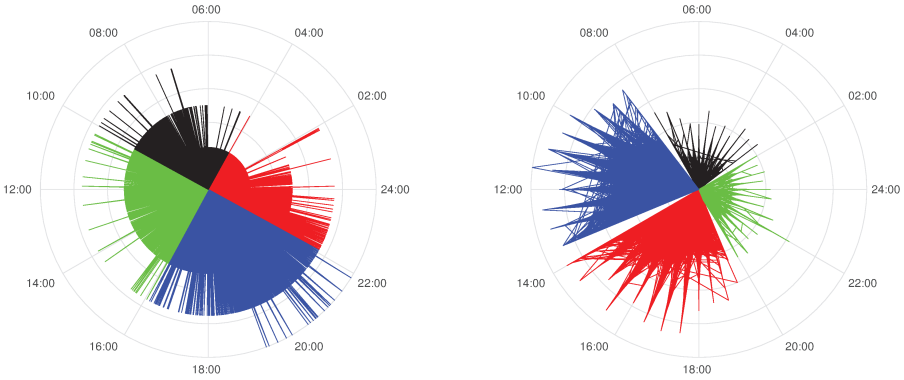
In the case of office building scenario, the result performed with the Abercrombie House Water dataset is close to the HMM one, reported in Table 5.12, and a lower standard deviation is achieved. All the remaining datasets have achieved worse performance respect to the GMM and HMM ones.

Table 5.21: DFID best results with AE.

Res.	AUC (%)	STD	Features comb.	Param.
Aber. Gas	77.68	11.09	*dLw+Lw*	2
Aber. Water	91.56	4.35	*Ma+Da+dMa+En+dEn*	2
Over. Gas	80.37	12.29	*Ma+En*	2-2
White. Water	84.75	10.34	*Da+We+dMa+Ma+dEn*	2

5.7.5 Temporal Clustering

In order to have a better insight of how time and consumption affect the leakage detection, further experiments have been conducted on the evaluated datasets. Specifically, for each datasets, four temporal clusters, on daily basis, have been



(a) Residential. (b) Office building.

Figure 5.11: Example of consumption temporal clustering for residential (**a**) and office building (**b**) scenarios. The blue data represents the "High" consumption cluster, the black one the "Low" consumption cluster.

identified, each characterizing a particular consumption level within a specific daily span, as shown in Figure 5.11.

The target dataset have been processed by means of the *k-means* [73] algorithm, that allowed to define four temporal slots. The two slots containing the lower and the higher average consumption have been classified as "Low" and "High" clusters, respectively. The remaining slots, representing average consumption, but at different daily times, have been generally classified as "C1" and "C2" clusters.

In order to perform the *k-mean* algorithm on the daily cycle, the consumption data have been rearranged in Cartesian coordinates (x, y) by exploiting the polar coordinate system (ρ, θ) using $24\,h$ as period. Therefore, let (x_i, i) be the pair representing the consumption level at the time instant $i = 1, \cdots, N$, respectively, the corresponding (ρ, θ) coordinates are obtained as:

$$\begin{aligned} \rho_i &= x_i \ , \\ \theta_i &= \frac{2\pi i}{T_{24h}} \ , \end{aligned} \tag{5.9}$$

where T_{24h} denotes the length of the $24\,h$ period expressed in the adopted resolution. For example, in the case of AMPds dataset at 1 minute resolution, each i-th value corresponds to the elapsed minutes from the data acquisition start, and T_{24h} is equal to 1440.

Therefore, in order to obtain the transformed coordinates (x_i^*, y_i^*), for the

application of the *k-means* algorithm, the following further transformation is performed:

$$\begin{aligned} x_i^* &= \rho\cos(\theta) = x_i \cos\left(\frac{2\pi i}{T_{24h}}\right), \\ y_i^* &= \rho\sin(\theta) = y_i \sin\left(\frac{2\pi i}{T_{24h}}\right). \end{aligned} \tag{5.10}$$

In the evaluation phase, the data that belong to the same cluster are joined together, and the leakage detection approach is applied to each new subsets.

For the AMPds and the DFID datasets, the time boundaries obtained for each cluster, and adopted in the dedicated experimental evaluations, are reported in Table 5.22.

Table 5.22: AMPds and DFID temporal clusters.

	Cluster			
Dataset	**Low**	**High**	**C1**	**C2**
AMPds Gas	23:30 - 05:30	18:00 - 23:30	12:00 - 17:30	06:00 - 11:30
AMPds Water	00:30 - 06:00	19:00 - 24:00	06:30 - 12:30	13:00 - 18:30
Aber. Gas	12:30 - 16:00	08:30 - 12:00	00:30 - 08:00	16:00 - 00:30
Aber. Water	06:00 - 12:00	12:30 - 15:30	20:30 - 05:30	16:00 - 20:30
Overall Gas	06:30 - 11:00	16:00 - 23:00	11:30 - 15:30	23:00 - 06:00
White. Water	21:30 - 04:30	05:00 - 10:30	16:00 - 21:30	11:00 - 15:30

Experiments

The leakage detection algorithm has been executed for each data cluster. The detailed results fo AMPds and DFID datasets are reported in Tables 16, 17, and 18, respectively.

For both AMPds and DFID datasets, the results reported in Table 5.23 validate the assertion made in Section 5.7.2: the standard deviation (shown by the results reported in Tables 5.3, 5.4, and 5.12) does not depend on the variability of the initial conditions in the training procedure, but it is entirely due to the leakage position in the dataset, thus the occurrence time, instead. In fact, the system produces remarkable different results between "High" and "Low" clusters. For the AMPds datasets, a considerable gap is shown for each resolution and model. Whereas, in the case fo the DFID datasets, for the Abercrombie House datasets, the gap between 'High" and "Low" clusters became clear for the HMM model.

Table 5.23: Best results achieved for temporal clustering. The column "**Par.**" reports the number of Gaussians, the states and Gaussians number for GMM and HMM models, respectively.

	Clusters							
	Low		**High**		**C1**		**C2**	
Resource	AUC(%)	Par.	AUC(%)	Par.	AUC(%)	Par.	AUC(%)	Par.
	AMPds							
Gas 1 GMM	74.75	32	65.94	256	80.19	256	74.83	128
Gas 10 GMM	81.86	128	63.04	256	69.64	64	75.87	256
Gas 30 GMM	69.36	8	64.65	256	73.79	128	64.44	64
Gas 1 HMM	91.64	1-256	67.25	2-128	80.34	1-256	87.64	2-64
Gas 10 HMM	75.36	1-256	64.82	2-128	71.19	1.256	78.67	1-128
Gas 30 HMM	71.42	1-8	67.09	2-64	76.21	1-64	66.39	4-4
Water 1 GMM	91.47	16	57.05	256	65.52	256	82.32	256
Water 10 GMM	93.36	16	58.13	256	73.10	256	69.31	256
Water 30 GMM	83.45	8	57.28	32	63.79	256	59.90	256
Water 1 HMM	92.48	4-64	61.90	2-64	82.06	4-16	92.90	3-128
Water 10 HMM	99.10	3-8	61.50	2-256	80.93	4-256	73.16	1-256
Water 30 HMM	95.50	2-8	64.01	4-2	73.07	4-256	68.44	3-128
	DFID							
Aber. Gas GMM	38.50	2	57.91	256	49.4	128	60.59	64
Aber. Gas HMM	94.17	3-2	68.23	4-64	88.56	2-4	70.74	2-32
Aber. Water GMM	70.45	8	71.23	256	84.64	16	75.23	256
Aber. Water HMM	93.20	4-16	79.03	4-64	84.64	3-16	75.23	4-128
Overall Gas GMM	70.75	2	56.81	256	61.56	32	73.69	2
Overall Gas HMM	93.82	3-2	64.05	3-32	71.31	3-16	93.63	4-4
White. Water GMM	99.29	256	78.41	128	89.47	16	84.51	256
White. Water HMM	99.37	1-256	85.74	4-16	96.23	2-8	92.67	3-128

The results achieved by means of this evaluation, together with the evaluations performed in Section 5.7.2, have allowed to understand the nature of the high variability of the system results. Therefore, it has been deemed necessary to evaluate the introduction of temporal information in the automatic leakage detection algorithm, as presented in Section 5.2.1, in order to attempt to mitigate the variability effect due to the time position of the leakage.

5.8 Remarks

A novelty detection based approach for the identification of leakage occurrence in smart water and natural gas grids has been presented. The validity of the statistical approaches, GMM and HMM, has been confirmed by the achieved results, and their suitability for both domestic and office building consumption has been proven using the AMPds and the DFID dataset.

The performance variability analysis have highlighted the correlation between the detection capability and the time position of the leakage. For this reason, temporal features have been also introduced in order to improve the

performance of the automatic leakage detection approach.

In the experiments performed with the *simulated* leakage typology, by means of the EPANET tool [37], realistic leakages have been simulated by introducing an emitter in the network, that produces an output flow (leakage size) dependent on flow and pressure in the reference node. Moreover, the features extracted from the pressure information have been also evaluated in combination with the flow features and the temporal ones.

For both GMM and HMM the overall best results are achieved by addressing both flow and pressure, for all time resolutions, and the GMM performance are close to the HMM ones. The introduction of pressure information allows to adopt an approach with lower computational burden, and to detect leakages even for low resolution (30 minutes), as well. Therefore, the usability of the proposed leakage detection approach in low-power devices/meters is improved.

Differently for the state-of-the-art methods, the proposed approach executes a detection for each frame, for the GMM, or for each frames sequence, for the HMM. Therefore, the system executes an on-line tracking and detects both abnormal and normal events. In this way, the assumption of time-limited leakages models has proven the detection time to be reasonably short. Moreover, the usefulness of temporal information, especially for low sampling data rates, has been confirmed as well.

Chapter 6

Advancements in the Smart Electrical Grid Field

6.1 Energy Management

Smart grids are meant to replace nowadays power grids, to provide efficient and flexible power distribution. Since the transition will take time, many techniques are being proposed, and adopted, to prevent the current technology from reaching its limit. Demand side energy production is one of them. However, the lack of coordination among demand side power plants has limited the achievable benefits [94, 95]. Demand Side Management schemes [96], such as dynamic pricing [97], is also a promising approach [98, 99] but, at times, it may conflict with local energy production [100]. On purpose, energy and resource scheduling have proven to be effective in complementing local demand and production, thus overcoming conflicts [101, 102, 103].

Nonetheless, to properly plan in advance the residential demand and production, the use of forecasts is mandatory. For this reason, in [10], an energy management approach also including solar energy forecasts has been proposed.

In this work, also, the effects of price forecasting over the scheduling process has been investigated, since forecasting errors may hinder the management activity. For instance, the energy management has been implemented by revising and extending the framework proposed in [10]. As such, dynamic pricing has been included, and both hourly and daily scheduling approaches have been inspected.

The price forecaster of choice has been based on Extreme Learning Machine technology [72, 104], proven to be particularly effective and robust as a price classifier [105]. In fact, although many forecasting techniques are available [106, 107, 108] in literature, some of them may be not suited for residential users [109, 110], due to data input requirements, i.e. the forecaster requires data not available to the user, or because they may not target the needs of residential user, i.e. the forecaster is able to predict the market prices while not accounting additional costs and taxes.

On purpose, in the present work, both input and target data sets have been intentionally restricted to the pricing profile applied to the customer. However, since Italian utility services do not provide any dynamic pricing based tariff, yet, the energy market hourly prices, the ancillary costs and the taxes, over a three years time period, have been used to compute a pricing profile that is consistent with the proposed scenario.

The experimentation has been aimed to hourly and daily scheduling, and has accounted both two tiered and dynamic tariffs. The achieved results have revealed that the improvement of the energy management efficiency depends, mostly, on the environmental setup, whereas the pricing profile scheme has a major effect on the number of degrees of freedom available to achieve savings. On the other hand, the price forecasting error limits the ability to lower the energy demand, although this aspect can be improved with a forecasting technique tailored to support the scheduling.

6.1.1 Real Life Based Model

The micro grid residential environment, namely Leaf House, is the reference used to implement the proposed framework. The Leaf House is one of the six international case studies selected by the IEA Task 40/ECBCS Annex 52 "Towards Net Zero Energy Solar Buildings" [111]. Built in 2008, it is located in Angeli di Rosora, Ancona, Italy (latitude 43°28'43.16 N, longitude 13°04'03.65 E, altitude 130 m above sea level).

By monitoring said environment, a data set has been compiled by collecting data regarding both local demand, of electricity and heat, and local solar based production, of electricity, heat and hot water. The integrated subsystems, also, have been used as a reference to model the various elements of the proposed framework [10]. Therefore, a modular design has been achieved, to model several variation over the original environmental structure. In particular, 8 different management architectures, based on the same building blocks, have been proposed to evaluate the energy management abilities of different residential environment design.

Many of the details regarding both the environment, and the framework, have been left out for conciseness sake, but they can be found in our previous work [10].

Modelling Approach

The scheduling and simulation framework has been implemented to model the energy distribution within the environment. Thus energy production, consumption, storage and transport have been modelled. The energy costs, then,

have been included to weight the different contributions, as to lower the energy bill, as well as to improve the efficiency of the residential environment.

Since the modelling process has required only linear equations to represent the energy balances, the framework has been implemented as a Mixed Integer Linear Programming (MILP) problem. Rearranging these equations, 8 configurations with increasing degree of integration among the resources have been proposed, based on the original structure.

In configuration 1, electricity management, hot water management and HVAC (heating ventilation and air conditioning) management have been accounted by means of separated subsystems, without the support of the storage. With respect to configuration 1, in configuration 2 the thermal energy storage has been also included, in configuration 3 the electrical energy storage, rather than the thermal energy storage, has been included whereas, in configuration 4, both thermal and electrical storages have been included, thus replicating the management implemented in the Leaf House. The common trait of these configurations is the low degree of integration among the subsystems. In fact each subsystem operates independently, meaning that both electrical energy surplus and thermal energy surplus are not routed from a subsystem to another.

The remaining configurations add incremental changes to configuration 4. In configuration 5 the management policy of the electrical storage has been removed, thereby in the case of configuration 5 the charge and the discharge of the electrical storage are not bound to a residual energy threshold. Configuration 6 adds to configuration 5 the integration of the hot water and the heating subsystems. Configuration 7 adds to configuration 6 the integration of the electrical and thermal subsystems. Configuration 8 differs from configuration 7 in that the three electrical subsystems have been merged in a single phase circuit. Configuration 8 carries the highest level of integration among the resources within the environment. All the proposed configurations are briefly listed as follows.

Configurations under test

1. storage devices not used.
2. thermal storage
3. electrical storage with a conservative electrical storage management policy
4. thermal and electrical storage with a conservative electrical storage management policy
5. thermal and electrical storage

6. thermal and electrical storage with the integrated production of hot water and heat
7. thermal and electrical storage with the integration of thermal and electrical management
8. thermal and electrical storage with the integration of thermal and electrical management at building level

---

With respect to the previous work [10], the structure of the framework and the simulation algorithm remain unchanged, therefore those aspects will not be discussed in the present manuscript. The current manuscript will focus on the changes applied, and their implications. On purpose, the previous version of the framework has been both revised and extended.

Specifically, the original framework has been revised to improve the cost function, to exert better control over the scheduling activity and to achieve a more conservative management of the storages. At the same time, the coding has been revised to improve its robustness. The revised framework, then, has been extended to operate on a 24 hour time frame, that is to manage an entire day at once. The overall structure of the framework, as well as the theory at the base of the model, have not been altered.

The MILP problem, hence, assume the generic form:

Minimize the overall electrical and thermal energy costs.

Subject to:

- electrical energy balance for the entire structure
- electrical energy balance for the first floor
- electrical energy balance for the second floor
- electrical energy balance for the third floor
- thermal energy balance for the entire structure
- hot water balance for the entire structure

In the following, the variables used to model the energy fluxes are recalled:

- EP_a: amount of energy purchased to supply the entire structure;
- ES_a: amount of energy surplus resulting from the entire structure;
- ESp_a: amount of solar thermal energy production ;
- Ch_a: amount of energy being stored;

- E_p: the energy purchase price;
- EP_i: energy required by the $i-th$ electrical block with $i = 1, 2, 3$.

In the previous implementation of the framework, the objective function, that computes the overall electrical and thermal energy costs, has been defined as:

$$Q = (EP_a + ES_a) \cdot E_p, \tag{6.1}$$

where the energy purchased EP_a is the sum of the energy purchased by each block, that is:

$$EP_a = EP_1 + EP_2 + EP_3 + Eg_1 + Eg_2 + EP_{hw} + EP_{he} + EP_{re}, \tag{6.2}$$

while the energy surplus ES_a is the sum of the surplus generated by each of the three photovoltaic panels:

$$ES_a = ES_1 + ES_2 + ES_3. \tag{6.3}$$

Therefore 6.1 has been used to represents the monetary "value" of the energy within the system.

From a broader perspective, additional aspects such as comfort or storage activity can be accounted to exert additional control over the optimization process.

On purpose, the objective function has been rewritten as:

$$Q = (EP_a) \cdot E_p + ES_a \cdot 10^{-6} - ESp_a \cdot 10^{-6} + Ch_a \cdot 10^{-7}\ , \tag{6.4}$$

where ES_a represents the overall energy surplus, to be demoted, ESp_a represents the solar thermal energy production, to be promoted, and Ch_a denotes the energy routed towards the storage, which is accounted to price the activity of the storages to be restricted.

The quantities 10^{-6} and 10^{-7} are arbitrary values acting as scaling factors. Said values have been introduced to manage the activity of the system correctly. They have been chosen based on the fact that energy is accounted in Wh, and the price E_p is accounted in € per Wh, that is the price of a Wh of purchased energy has an order of magnitude equal to 10^{-5}.

In fact, since the energy being stored has a value (10^{-7}), the manager is required to limit the activity of the storage. Similarly, since the energy surplus and the solar thermal energy have a value, the manager is required to account their amount.

On the one hand, removing these amounts from the objective function, allows the manager to ignore solar production, leaving the energy unused, or to dissipate the energy surplus through the storage to lower the unused surplus.

On the other hand, including said amounts in the objective function while assuming a scaling factor equal to 1, forces the energy manager to minimize both the amount of energy surplus and the storage activity, rather than the energy bill.

Thereby, the scaling factors are used to weight the optimization priority of some aspects of the system activity with respect to the management of the energy purchase. As such they are part of the energy management policy, and can be adjusted based on the user needs.

In the revised framework, the objective function presented in eq. (6.4) is the main (and noteworthy) novelty introduced by the revision process. Additional changes have also been applied to extend the time frame from 1 to 24 hours. However, they only concern the algebraic representation of the objective function, that has been rewritten as the sum of the objective functions associated to the 24 hours of the extended time frame. Regarding the constraints, not even their algebraic form required any modification. The only difference is that in place of one single constraint, 24 of them have been assigned, one for every hour of the time window.

6.1.2 Case Study Setup

The aim of the current work is to evaluate the compatibility between energy and resource scheduling and price forecasting. Specifically, the evaluation is aimed at daily scheduling, the hourly manager has been used as a reference. On purpose, the daily manager has been derived from the hourly manager, as to avoid evaluation biases due to differences in the implementation. In fact the only difference between the two is the operating time window. The evaluation covers different cases, that is the performance of both manager with respect to the original implementation, the performance difference between hourly and daily manager, and the performance of both in presence of data uncertainty.

Set up of the simulation scenario

The framework runs on the MATLAB® computation environment. It is based on the SCIP solver through the Opti Toolbox 2.11 interface [112]. The computational environment is hosted on a Laptop PC, based on the Intel Core i7 CPU series, with 8 GB of RAM, and running on the Microsoft Windows 8.1 64-bit OS.

The data set required by the framework includes both historical and forecast data. Historical data has been collected directly from the residential environment, through its management system. It includes hourly local energy production and demand, of both the electrical and thermal kind, on a per apartment basis. Forecast data includes both solar production forecasts and

price forecasts. Solar production forecasts have been generated by means of a Radial Basis Function Networks (RBNFs) based forecaster. The details are presented in our previous work [10]. The price forecasting will be discussed shortly thereafter.

To provide an insight on the scheduling problem while avoiding biases due to seasonality and short term fluctuation of the forecasting error, the results being evaluated represent the cumulative energy costs over a year time span. For instance, the reference time frame spans from 1st November 2012 to 31st October 2013. The entire data set, part of which has been used to train the forecasters, starts from January 2010.

Dynamic pricing

In the present work, along with the flat rate pricing scheme presented and discussed in the previous work, a dynamic tariff is proposed to investigate the compatibility between dynamic pricing and energy and resource scheduling.

Due to the lack of dynamic pricing offers from Italian service utilities, a dynamic profile has been computed based on the Italian energy market prices[1]. Dispatch costs, ancillary costs and taxes have been accounted based on the information provided by the Italian Regulatory Authority for Electricity Gas and Water[2].

In particular, since the residential environment used as a reference has a single point of common coupling (PCC), despite the six apartments, dispatch costs, ancillary costs and taxes applied to business users have been used. This choice has been also deemed consistent with the dynamic pricing scenario, since to business users single tiered costs are applied, whereas to residential user three tiered costs are applied.

Over the reference time frame the average of the two tiered tariff amounts to 0.00013354€ per Wh, whereas the average of the dynamic price amounts to 0.00012988€ per Wh.

To generate the dynamic price profile forecast, a day-ahead forecast has been used. For instance, the value corresponding to the current hour of the current day, has been forecast based on the values corresponding to the same hour of the previous seven days in order to better catch the weekly periodicity. Specifically, the price profile forecast, over the entire reference set, has been performed on hourly basis.

The forecaster of choice is a Single Layer FeedForward network (SLFN). Its training has been carried out by mean the kernel based ELM fast learning algorithm, with a radial basis function (RBF) kernel. The number of hidden

[1] https://www.mercatoelettrico.org/EN/Download/DownloadDati.aspx
[2] http://www.autorita.energia.it/it/inglese/index.htm

neurons does not need to be known in advance. The implementation is based on the code[3] provided by Hong-Ming Zhou and Guang-Bin Huang[4].

The two day ahead forecasters used as reference have been implemented as Multilayer Perceptron (MLP) and NARX neural network respectively. Both have been implemented, by means of the MatLab Neural Network Toolbox, as a 2 layer networks, the hyperbolic tangent has been adopted as activation function of the neurons, and the Levenberg-Marquardt algorithm has been used in the training phase.

To setup the forecasters, several test runs have been made, with each technique, in order to find the most performing configuration. For instance, since the prepared data sets date back to 2010, a price profile from 1st November 2010 to 31st October 2011 has been used to train the neural networks, whereas a price profile from 1st November 2011 to 31st October 2012 has been used to validate the neural networks. Then the price profile from 1st November 2012 to 31st October 2013 has been used to generate the price forecasts and to evaluate performance of the neural networks.

In the case of the SLFN forecaster, the best performance has been achieved assuming 0.082 as the value of the kernel parameter, which is referred to as *kernel parameter* σ, in [72]. In regards to the MLP forecaster, in the best performing configuration the neurons have amounted to 15 in the hidden layer and to 1 in the output layer. In the case of the NARX neural network, the best performing configuration has been obtained with 10 neurons in the hidden layer and 1 neuron in the output layer. In this case, also, the delay time steps amount to 0 in the case of the input, and to 24 in the case of the feedback.

6.1.3 Results Evaluation

Table 6.1: Forecasting indices: back to back comparison.

		MLP model	NARX model	Kernel based ELM model
MAPE	Mean:	5.57	5.57	5.45
	Variance:	0.002	0.01	0.00
RMSE	Mean:	9.33	9.49	8.71
	Variance:	0.02	0.04	0.00

In the current section the results of both forecasting techniques and energy and resource scheduling are proposed.

[3] http://www.ntu.edu.sg/home/egbhuang/elm_kernel.html
[4] http://www.ntu.edu.sg/eee/icis/cv/egbhuang.htm

Table 6.2: Energy cost from the original framework [10], based on historical data and two tiered tariff, for all addressed system configurations. The identification number is based on the entries listed in Section 6.1.1.

			Configurations							
			1	2	3	4	5	6	7	8
Hourly management	Cost due to electrical blocks	(€)	618.85	618.85	488.50	488.50	493.42	493.42	493.03	441.08
	Cost due to thermal blocks	(€)	1356.69	1318.76	1356.69	1318.76	1318.76	830.59	585.25	594.30
	Overall cost	(€)	1975.54	1937.61	1845.15	1807.26	1812.18	1324.01	1078.28	1035.38

Table 6.3: Energy cost based on historical data and two tiered tariff (ideal scenario with no prediction error), for all addressed system configurations. The identification number is based on the entries listed in Section 6.1.1.

			Configurations							
			1	2	3	4	5	6	7	8
Hourly management	Cost due to electrical blocks	(€)	618.85	618.85	488.50	488.50	429.27	429.27	599.02	562.96
	Cost due to thermal blocks	(€)	1356.69	1349.79	1356.69	1349.79	1349.79	830.59	430.56	443.86
	Overall cost	(€)	1975.54	1968.64	1845.19	1838.29	1779.06	1259.86	1029.58	1006.82
Daily management	Cost due to electrical blocks	(€)	618.85	618.85	474.38	474.41	426.51	426.51	479.04	427.89
	Cost due to thermal blocks	(€)	1356.69	1354.73	1356.69	1306.28	1306.28	804.05	377.04	390.41
	Overall cost	(€)	1975.54	1973.58	1830.57	1780.69	1732.79	1230.56	856.08	818.30
	Overall cost difference	(€)	-	-4.94	14.62	57.60	46.27	29.30	173.5	188.52

Forecasting techniques

To compare the forecasters performance, 10 forecasts per technique have been generated. Being the most common indices, the percentage MAPE (Mean Absolute Percentage Error) and percentage RMSE (Root Mean Squared Error) have been computed and their mean value and variance are proposed in Table 6.1.

For instance, percentage MAPE and percentage RMSE have been computed as follows:

$$\begin{aligned} \text{MAPE} =& 100 \cdot \frac{1}{n} \sum_{t=1}^{n} \left| \frac{P_t - F_t}{P} \right| , \\ \text{RMSE} =& 100 \cdot \frac{1}{n} \sqrt{\sum_{t=1}^{n} \left| \frac{P_t - F_t}{P_t} \right|^2} , \end{aligned} \tag{6.5}$$

where, within a year, n is the n-th hour, P_t is the $t-th$ historical hourly price value, F_t is the $t-th$ forecast hourly price value and P is the average of the historical hourly price.

The MLP and NARX forecasters show similar results, although the lower variance of the MLP based forecasts reveals an increased robustness of this method with respect to the NARX network. Most likely, the NARX model has been affected by the lack of information about the weekly and seasonal

correlation of the prices. It must be remarked that, omitting the day of the week and the day of the year, to adopt a simpler input set, may have resulted in sub-optimal working conditions for the NARX model. The results of the ELM based forecaster show that, with respect to MLP and NARX based forecasters, the percentage MAPE index mean value is 2.15% lower, the percentage RMSE index mean value is, respectively, 6.64% and 9.22% lower. The variability, also, is null. Thus, the ELM based forecaster appears to be more accurate, and more robust against the network random initialization.

Hourly management improvements

The first evaluation compares the hourly energy management results against the performance of the previous implementation of the framework, based on the overall cost of the purchased energy. Therefore a lower energy cost represents a better performance.

The scheduling results of the previous framework are reported in Table 6.2, whereas the results of the new implementations are proposed in Table 6.3. As stated above (Section 6.1.1), the setup and parameters of the new implementation remain unchanged from the previous work and the two-tiered pricing profile, used to run the scheduling, is also the same.

Regarding the hourly manager, the comparison of the results, reported in Tables 6.2 and 6.3 and labelled: *Overall cost*, reveals that with the exception of configuration 2 and 4, the revised cost function improves the efficiency of the energy management. If the thermal and the electrical contributions are accounted, respectively shown by the entries labelled: *Cost due to the thermal blocks* and *Cost due to the electrical blocks*, it can be observed that the revised cost function does not interfere with the policy of the electrical storage (configurations 3 and 4), it lowers the electrical needs when electrical and thermal subsystems are apart (configurations 5 and 6), whereas it promotes the use of the electrical storage to support the thermal subsystem. It should be reminded that only the electrical storages of configuration 3 and 4 have a management policy applied.

In the case of configuration 2 and 4, as said, the performance has worsened. However, the performance drop affects only the thermal management, whereas the electrical management remains unaffected. This behaviour is due to the value assigned to the scaling factors, which is close to the difference between the off peak price and the peak price. Due to that, whenever the heat pump is used, due to the COP, shifting the thermal load is regarded as "not convenient" by the energy manager.

In fact, by shifting the electrical load from peak hours to off-peak hours, a saving can be achieved. The saving per electrical Wh amounts to the difference between the peak-hour tariff and the off-peak tariff (about 9.73e-6 €). Because

Table 6.4: Energy cost based on historical data and dynamic tariff (ideal scenario with no prediction error), for all addressed system configurations. The identification number is based on the entries listed in Section 6.1.1.

			Configurations							
			1	2	3	4	5	6	7	8
Hourly management	Cost due to electrical blocks	(€)	631.08	631.08	486.05	486.05	424.13	424.13	596.73	557.45
	Cost due to thermal blocks	(€)	1367.62	1361.32	1367.62	1361.32	1361.32	824.26	425.30	438.78
	Overall cost	(€)	1998.70	1992.40	1853.67	1847.37	1785.45	1248.39	1022.03	996.23
Daily management	Cost due to electrical blocks	(€)	631.08	631.08	470.35	470.42	403.41	403.41	449.53	388.59
	Cost due to thermal blocks	(€)	1367.62	1161.62	1367.62	1160.71	1160.71	717.23	335.58	350.63
	Overall cost	(€)	1998.70	1792.70	1837.97	1631.13	1564.12	1120.64	785.11	739.22
	Overall cost difference	(€)	-	199.70	15.70	216.24	221.33	127.75	236.92	257.01

of the COP of the heat pump, however, the saving per thermal Wh is roughly one third of the saving per electrical Wh. As such, since the storage activity has a "value" per Wh, represented by the scaling factor, higher than the saving, the storage activity is regarded as unnecessary and avoided by the manager. That is to say that, within the framework, the user can define the threshold below which the storage activity is "less convenient" than the savings it produces.

Table 6.5: Energy cost based on forecast data and dynamic tariff (real life scenario with prediction error), for all addressed system configurations. The identification number is based on the entries listed in Section 6.1.1.

			Configurations							
			1	2	3	4	5	6	7	8
Hourly management	Cost due to electrical blocks	(€)	631.08	631.08	623.65	622.57	576.55	576.55	653.97	605.57
	Cost due to thermal blocks	(€)	1367.59	1361.14	1367.59	1361.14	1361.14	824.22	505.12	522.23
	Overall cost	(€)	1998.67	1992.22	1991.14	1983.71	1937.69	1400.77	1159.09	1127.80
Daily management	Cost due to electrical blocks	(€)	631.08	631.08	553.10	550.30	544.01	543.80	572.24	365.04
	Cost due to thermal blocks	(€)	1367.59	1174.83	1367.59	1189.34	1189.34	726.64	410.74	431.43
	Overall cost	(€)	1998.67	1805.91	1920.69	1739.64	1733.35	1270.44	982.98	796.47
	Overall cost difference	(€)	-	186.31	70.45	244.07	204.34	130.33	176.11	331.33

Table 6.6: Difference of energy cost between historical data based management (Table 6.4) and forecast data based management (Table 6.5), for all addressed system configurations. The identification number is based on the entries listed in Section 6.1.1.

		Configurations							
		1	2	3	4	5	6	7	8
Hourly management	(€)	0.03	0.18	-137.47	-136.34	-152.24	-152.38	-137.07	-131.57
Daily management	(€)	0.03	-13.21	-82.72	-108.51	-169.23	-149.80	-197.87	-57.25

Since a different choice of the scaling factors would change the presented behaviour, the framework can be used to evaluate different policies, depending on the user needs.

Daily management and hourly management

Regarding the daily energy management, comparing its performance, based on the results presented in Table 6.3, against the performance of the previous implementation (Table 6.2), shows that the configuration 1 does not provide the means to improve the system efficiency. In the case of configuration 2 the unexpected results are even worse than the revised hourly manager (shown above in Table 6.3). On the other hand, in the case of configuration 4, the daily manager achieves a lower cost even in the case of the thermal management, which leads to the conclusion that a daily management allows a better coordination of both thermal and electrical management. In fact, in every configuration from 3 onwards, the daily manager outperforms both the old and the revised hourly managers.

While the already mentioned Table 6.3 refers to a scenario based on historical data and a two tiered tariff, Table 6.4 presents the results corresponding to a scenario based on historical data and the historical dynamic pricing profile. An evaluation of the results reported in Table 6.4, for both configuration 2 and 4, reveals that, since the price changes every hour, and also the price difference is usually higher with respect to the two tiered tariff, the assigned policy does not lead to an unexpected result. Hence, it is possible to conclude that the daily manager outperforms the hourly one in every configurations, and that the improvement grows with the increased degree of integration among the subsystems.

Data uncertainty

Last but not least, the effect of data uncertainty over the energy management has been investigated. In particular, Table 6.5 presents the results based on a scenario that accounts forecast data for both solar production and dynamic pricing. In this case, for each time step, the forecast data is used to compute the schedule of the system activity, the activity of the residential environment is then simulated by means of the same framework, and the environment performance is computed. The simulation uses the historical data and the schedule previously computed as inputs. In each Table the difference between the performance of the hourly and the daily manager, for each configuration, is reported as well.

In this case, the results presented in Table 6.5 are to be compared against the ones presented in Table 6.4. The case of the hourly manager and the case of the daily manager have been addressed separately. The difference between the overall energy costs, for both cases, has been presented in Table 6.6.

Because of data uncertainty, the performance of the manager is sub optimal, i.e. while it is possible to improve the efficiency of the residential environment,

it is usually not possible to achieve the minimum operating cost.

At a glance, with respect to the hourly management, the results proposed in Table 6.6 reveal that the losses due to the forecasting errors are consistent, spanning from about 131€ to about 152€. Also, the fact that the losses affect configuration 3, but not configuration 2, leads to the conclusion that the electrical energy management is more exposed to the forecasting errors. In this scenario, configuration 8 is the only non trivial setup that suffers the least drop in performance. In term of percentage with respect to the optimal scheduling, the drop in saving ranges from 7.38%, of configuration 4, to 13.41%, corresponding to configuration 7, suggesting that the high variability comes from the increased saving of the configuration, rather than the variability of the drop in performance.

Similar considerations, also, hold true in the case of the daily management, although the saving losses have a wider spread, from about 57€ to about 198€. In term of percentage the value spans from 6.65% to 25.20%. Configuration 8 appears to be the most robust non trivial setup, since the loss in performance, is about 57€, which is half the drop with respect to configuration 4. That is to say that highly integrated systems appear to be less affected by the forecasting errors.

From a general point of view, in the case of the hourly management, the storage activity is planned depending on the expected local production and the local demand in the next hour. Once the activity is planned, the amount of energy routed to or from the storage cannot be adjusted. Hence, if the energy production exceeds the forecast value, the extra energy is not collected, ending wasted. If the production is lower tha. expected, the energy to be stored will be purchased.

In the case of the daily manager, on the other hand, the storage activity is planned over a time window of 24 hours. Thus local demand and local expected production are accounted for every hour within the time frame. Differently from the hourly manager, however, the expected energy prices are also accounted. In both cases, the cost function reaches its minimum when the storage activity is absent and, at the same time, the energy cost is at its minimum. However, the hourly manager does no plan the energy purchase accounting the prices. On the other hand, the daily manager, by means of the price forecast, can compare the prices, and plan the purchase as to when the price is expected hits its minimum as to lower the cost function. This behaviour, also, has been observed in our earlier work [102].

In fact, the daily manager is affected by the position of the local minima of the price forecasts, which in turn are affected by the forecasting error. In addition, since the relative position of the minima does not depend on the absolute value of the forecasting error, it is not possible to relate the scheduling performance

to the forecast accuracy. For instance, the results reported in Table 6.6 do not reveal any obvious relationship.

Clearly, the results shown in Table 6.6 prove that, although sub-optimal, the improvement of the efficiency of the environment is consistent. Nonetheless, in order to assess the impact of the price forecasting error over the scheduling, an index that takes into account the displacement of local minima may be more effective than the widespread MAPE or RMSE.

6.1.4 Remarks

The evaluated results have shown that the proposed frameworks can be used to evaluate different management policies, as well as different residential environment configurations. In fact, its ability to plan ahead, in the case of a daily planning, can greatly improve the efficiency of the environment. With respect to the use of forecasts, in the case of local production forecasts, accuracy is a fundamental requirement to achieve the optimal planning of the environmental activity. In the case of price forecasting, however, accuracy may be a mandatory requirement, although not sufficient, in order to accurately estimate the energy cost. To ensure a correct scheduling over a 24 hour time windows, on the other hand, the proper placement of the local minima rather than forecasting their exact value, may be more effective.

Nonetheless the scheduling algorithm has proven to be quite robust, improving the efficiency of the environment despite the forecasting error. Also, the price forecaster has proven to be accurate enough to support the management.

6.2 Non-Intrusive Load Monitoring

The issues relating to the energy conservation and efficiency have gained a role of great importance, both from the point of view of the consumer and the energy provider. Furthermore, over the years, the infrastructures for energy distribution have undergone an ageing process, which have led to the study of the possibility in smart grids implementation, in which a set of information from detection and network management systems can be transmitted in addition to energy [113, 114].

Useful information, about the characteristics and operating behaviour of an electrical system, can be obtained by means of the power consumption analysis, in order to predict the power demand (load forecasting), to apply management policies and to avoid overloading or blackouts over the energy network. Similarly, from the user perspective, the lifestyle of the people in a house can be predicted by the energy consumption analysis, allowing to implement policies for advantageous time tariffs [115].

Over the years, several studies have demonstrated that the energy consumption awareness (i.e., which appliances are operating at a certain time instant and how much electrical power they are consuming) influences the user behaviour [116]. Specifically, the awareness conducts to moderate energy consumption, resulting in monetary savings and reduction of the energy required to the provider. Furthermore, applying this consideration to commercial or industrial environments, it may provide larger energy saving [117].

Load monitoring has become a challenging problem, and several techniques have been studied to solve it. This work has been focused on Non-Intrusive Load Monitoring (NILM) algorithms, introduced by [118], which aim to separate the aggregated energy consumption signal, measured in a single centralized point, in the individual signals from each appliance, using a simple hardware but smart software algorithms. This solution replace a distributed smart socket grid inside the house, resulting in lower implementation costs and less invasive solutions for the end user.

Specifically, here the attention has been granted to unsupervised NILM techniques. These approaches do not require individual appliance data and the models information is captured only using the aggregated load, without the user intervention. Furthermore, the unsupervised approaches are independent from the number of the appliances forming the aggregated load and capable of dynamically adapt to the power system changes over time (i.e., addition, removal or substitution of appliance).

6.2.1 State of the Art

A widely used structure for modelling appliances consumption behaviour is the HMM: it defines a number of hidden state in which the model can be moved, representing the operating condition of the appliance (e.g., ON, OFF and possible intermediate states), and an observable output,depending from the actual state, representing the consumption data under analysis. Different variation of HMM are used in the considered contributions: Kolter and Jaakkola [119] introduces a method based on Additive Factorial HMMs (FHMMs), in which each chain evolves with independent dynamics and produces its own emissions, but the observations are defined as combinations of the independent emissions; a convex quadratic optimization problem called Additive Factorial Approximate Maximum A-Posteriori (AFAMAP) is used.

Johnson and Willsky [120] exploits Hierarchical Dirichlet Process HMM (HDP-HMM) structure to provided a powerful framework to infer arbitrarily large state complexity from data, and explicit-duration semi-Markov modelling (HSMM) to define the Markov transition with more accuracy; a Gibbs sampling inference method is used and multiple different power modes are modelled for

each appliance.

Also Kim *et al.* [121] model the electric power system with the composition of a factorial Hidden semi-Markov model (FHSMM) to better define the state occupancy durations of the appliances, but introduce a Conditional FHMM (CFHMM) to integrate additional features related to when and how appliances are used in the house; Expectation Maximization (EM) algorithm and Maximum Likelihood Estimation (MLE) are used for the parameters inference and several features are employed, like ON-duration distribution, OFF-duration shape, dependency between appliances and daily schedule of the occupants.

Pattem [122] uses HMM with segmented Viterbi to identify and separate as many appliance signatures as possible by matching magnitudes of ON/OFF power transitions for each appliance, employing a residual analysis to handle lower power signatures.

Parson *et al.* [123] introduce HMM as Bayesian network and use this approach to build a general model for each appliance type, which is tuned to the specific appliance instance in a household given only aggregate data, using a sliding window technique and extracting only the signal useful for the model tuning.

Following the same idea, Figueiredo *et al.* [124] provide an efficient technique for disaggregation after model tuning by Expectation-Maximization (EM); the power demand of each appliance is posteriorly adjusted with an optimization step, under the constraint of correspondence between the sum of all appliances consumption and the measured aggregated data.

A method based on Dynamic Time Warping (DTW) for template matching is proposed by Liao *et al.* [125]: after a signature detection within the aggregate power consumption, DTW performs a non-linear mapping of the detected event to the previously recognized in order to reproduce a clustering via dynamic programming, while the classification is performed calculating the minimum distance between the signature detected and the matching cluster.

Barsim *et al.* [126] perform an event-based matching, using a first stage of signature detection over the reactive and real power signal (transient and steady-state), and finally using a non-parametric clustering algorithm, named Mean-Shift Clustering Scheme (MSCS), after a logarithmic normalization for grouping the events detected.

More simpler linear discrimination is used by Liu *et al.* [127], in which statistical features are extracted from active and reactive power and a combination of steady and transient state analysis is used in the classification stage; in this contribution the appliance activity recognition also exploits the correlation between different appliances usage.

Figueiredo *et al.* [128] presents a method for energy disaggregation interpreted as a single-channel source separation problem, which is based on a Non

Negative Matrix Factorization (NMF) method; in order to improve the performance, additional information about dependencies between different appliance are used, introducing the technique called Source Separation Via Tensor and Matrix Factorizations (STMF): under the assumption of linear mixing of the source power signal, a tensor representation and associated PARAFAC factorization extract the dependencies between the sources, which are then used in the disaggregation phase.

Also Wytock and Kolter [129] exploit the correlation between different appliances in a convex optimization problem for source separation: using a representation of each individual signal as a linear function of some component bases, a contextually supervised method based on temperature time series is performed and the non-linear dependence, represented with radial-basis functions (RBFs), is defined as basis for the energy consumption; theoretical analysis shows that linear independence between features for different signals is the requirement for an accurate separation.

Instead, Goncalves *et al.* [130] presents a Blind Source Separation (BSS) approach for the definition of the problem: from the observation of steady-state event in real and reactive power profile in a linear mixing schema, a Genetic k-means clustering procedure is used to obtain a representation of the appliances consumption state and the cluster information are then used in the source reconstruction algorithm, based on Matching Pursuit (MP), to characterize each appliance.

Recently, more challenging techniques, such as the use of Deep Neural Networks (DNN), have been investigated by Kelly and Knottembelt [131]. Specifically, three DNN architectures have been adapted to energy disaggregation: recurrent neural networks, denoising autoencoders, and a network which regress the start time, end time and average power demand of each appliance. For each architecture, it has been trained one network for each target appliance.

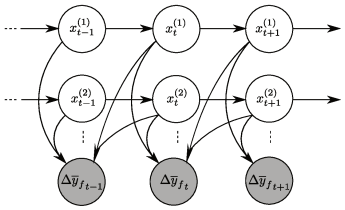


Figure 6.1: The Forward Differential FHMM.

6.2.2 AFAMAP

The theoretical approach towards disaggregation is based on the work of Kolter and Jaakkola [119]. In this work the system is modelled relying on Additive Factorial HMM (AFHMM), for which the value of each aggregated power sample corresponds to a combination of working states of the appliances into the system.

Also, in this approach, the assumption that at most one HMM may change its state at any given time is made, which holds true if the sampling time is reasonably short. In this case, the transition on the aggregate power, when moving from a sample to the next, corresponds to the state change of a particular HMM.

Because of that, the differential signal, built from the aggregated power, can be modelled as the result of a Differential Factorial HMM (DFHMM), which relies on the same HMM models composing the AFHMM.

In the reference approach, the DFHMMs are obtained as the difference, in term of power consumption, between the current and the previous sample (referred to as *backward transition*), so that a change in the state of an HMM can be evaluated against the change in the aggregated power consumption. Similarly, an additional evaluation, based on the next against the current sample (referred to as *forward transition*), is carried out. Furthermore, a smarter employment of the solver boundaries is evaluated, starting from a more accurate analysis of the aggregated power or using heterogeneous information, as the reactive power consumption of the electrical system.

Forward Differential Factorial Hidden Markov Model

Since the AFAMAP algorithm operates offline, it is possible to further extend the model by taking into account the transition from the current to the next state. The original DFHMM [119] is computed by looking backward from the current sample to the previous one, and thus it can be addressed to as Backward DFHMM. The new differential FHMM is computed by looking forward, as showed in Figure 6.1, and thus is referred to as Forward FHMM.

The formulation of the new model, also, differs from the original one, only in the index order. The new variables define the problem, as follow:

$$\mathcal{Q} = \left\{ \boldsymbol{Q(x_t^{(i)})} \in \mathbb{R}^{m_i}, \boldsymbol{Q(x_{t+1}^{(i)}, x_t^{(i)})} \in \mathbb{R}^{m_i \times m_i} \right\},$$

where the variables are indicators of the transition from the next to the current state: $Q(x_t^{(i)})_j = 1 \Leftrightarrow x_t^{(i)} = j$, and $Q(x_{t+1}^{(i)}, x_t^{(i)})_{j,k} = 1 \Leftrightarrow x_{t+1}^{(i)} = j, x_t^{(i)} = k$.

The consistent constraints between the state variables and transition variables need to be satisfied:

$$\mathcal{L} = \left\{ \mathcal{Q} : \begin{array}{c} \sum_{j=1}^{m_i} Q(x_t^{(i)})_j = 1 \\ \sum_{k=1}^{m_i} Q(x_{t+1}^{(i)}, x_t^{(i)})_{j,k} = Q(x_{t+1}^{(i)})_j \\ \sum_{k=1}^{m_i} Q(x_{t+1}^{(i)}, x_t^{(i)})_{k,j} = Q(x_t^{(i)})_j \\ 0 \leq Q(x_t^{(i)})_j, Q(x_{t+1}^{(i)}, x_t^{(i)})_{k,j} \leq 1 \end{array} \right\}. \tag{6.6}$$

Therefore, the new cost function is derived for the Forward DFHMM, based on the forward differential aggregated signal $\Delta\overline{y}_{f_t} = \overline{y}_t - \overline{y}_{t+1}$, as follow:

$$\begin{aligned} &\frac{1}{2{\sigma_3}^2} \sum_{t=1}^{T-1} E_t^{(fc)} + \frac{1}{2} \sum_{t=1}^{T-1} E_t^{(fnc)} + \\ &+ \sum_{t=1}^{T-1} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ k=1}}^{m_i} \left\{ Q(x_{t+1}^{(i)}, x_t^{(i)})_{j,k} \left(-\log P_{f\,k,j}^{(i)} \right) \right\} + \\ &+ \sum_{i=1}^{N} \sum_{j=1}^{m_i} \left\{ Q(x_T^{(i)})_j (-\log \phi_{f\,j}^{(i)}) \right\}, \end{aligned} \tag{6.7}$$

where the error terms in (6.7) are defined as:

$$E_t^{(fc)} = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ k=1 \\ k \neq j}}^{m_i} \left\{ \left(\Delta\overline{y}_{f_t} - \Delta\mu_{k,j}^{(i)} \right)^2 Q(x_{t+1}^{(i)}, x_t^{(i)})_{j,k} \right\}, \tag{6.8}$$

$$E_t^{(fnc)} = D\left(\frac{\Delta\overline{y}_{f_t}}{\sigma_3}, \lambda \right) \left(1 - \sum_{i=1}^{N} \sum_{\substack{j=1 \\ k=1 \\ k \neq j}}^{m_i} Q(x_{t+1}^{(i)}, x_t^{(i)})_{j,k} \right). \tag{6.9}$$

The transition matrix $\boldsymbol{P_f^{(i)}}$ represents the probability of state change from the next to the current time instant: this parameter is equivalent to the typical representation of the transition matrix (i.e., the probability of state change from the previous time instant to the actual) evaluated after flipping the signal,

thus it can be derived by using the available algorithm for HMM training. The parameter $\boldsymbol{\phi_f^{(i)}}$ represents the final state distribution, that is the initial state distribution starting from the end of the signal.

Since the duality in the forward and backward representation of the AFHMM (i.e., it is derived from the same observed signal, but in opposite directions), the problem definition using only one of the two versions of the DFHMM leads to the already known performance. Considering simultaneously both versions of DFHMM may lead to performance improvements: for this reason the forward differential function (6.7) is added to the reference formulation, thus leading to a new optimization problem.

The variable vector $\boldsymbol{x}$ in the QP problem accounts for the new terms, following the same structure introduced in [119]:

$$\boldsymbol{\Psi_t^{(i)}} = \begin{bmatrix} \boldsymbol{\xi_t^{(i)}} \\ \boldsymbol{\beta_t^{(i)}} \\ \boldsymbol{\phi_t^{(i)}} \end{bmatrix}, \qquad \boldsymbol{\phi_t^{(i)}} = \begin{bmatrix} Q(x_{t+1}^{(i)}, x_t^{(i)})_{1,1} \\ \vdots \\ Q(x_{t+1}^{(i)}, x_t^{(i)})_{1,m_i} \\ \vdots \\ Q(x_{t+1}^{(i)}, x_t^{(i)})_{m_i,1} \\ \vdots \\ Q(x_{t+1}^{(i)}, x_t^{(i)})_{m_i,m_i} \end{bmatrix},$$

where the new term $\boldsymbol{\phi_t^{(i)}}$ represents the variables for the forward transition.

The introduction of the new variables leads to an alteration of the problem constraints, represented by the parameters $\boldsymbol{A_{eq}}$ and $\boldsymbol{b_{eq}}$, and the variable boundaries $\boldsymbol{lb}$ and $\boldsymbol{ub}$.

Profile-based selection of the boundaries

In order to solve the optimization problem, different solutions, which satisfy the constraints, need to be evaluated before the solver finds the optimal one. As such, the values of $\boldsymbol{x}$ that are not compatible with the given set of samples can be discarded, to restrict the search domain and improve the search efficiency.

On purpose, the lower and upper boundaries of the variable $\boldsymbol{x}$ are selected beforehand in order to prevent that the solver investigates those combinations of states that do not match the value of the aggregated power consumption. The selection method is similar to the one proposed in [132].

If several runs of a single appliance are evaluated, although the same working states are identified, the signature tends to differ from a run to the other. For this reason, the appliance power consumption can be modelled as a stochastic process and, therefore, the output value $y_t^{(i)}$, relative to a working state $x_t^{(i)}$

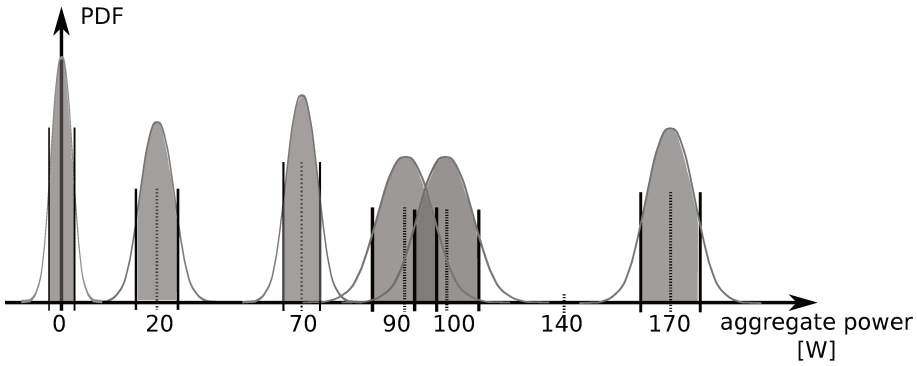


Figure 6.2: A sketch of the different probability density functions (PDF) for each aggregated power value produced by the combination of all appliances states power levels.

of an appliance, can be modelled as a gaussian variable, described by a mean value and a variance value:

$$y_t^{(i)} | x_t^{(i)} \sim \mathcal{N}\left(\mu_{x_t^{(i)}}^{(i)}, {\sigma_{x_t^{(i)}}^{(i)}}^2\right). \tag{6.10}$$

Regard to this, the power signal is replaced by a simplified model that presents a constant power consumption, corresponding to the mean value of the working state power value, with a superimposed noisy contribution, described by the variance value in the working state.

Since the aggregated data $\overline{y}_t$ is assumed to correspond with the sum of the power consumption of each appliance, it can be modelled as a gaussian variable, described by a mean value and a variance value equivalent to the sum of the corresponding values of each appliance, under the assumption of statistical independence between the appliances:

$$\overline{y}_t | x_t^{(1:N)} \sim \mathcal{N}\left(\sum_{i=1}^{N} \mu_{x_t^{(i)}}^{(i)}, \sum_{i=1}^{N} {\sigma_{x_t^{(i)}}^{(i)}}^2\right). \tag{6.11}$$

This simplified model results in a number of admissible combinations of working states equal to $\prod_{i=1}^{N} m_i$. It allows to evaluate which combination of working states fit the power value for each sample of the aggregated data, thus discarding the incompatible ones. The effectiveness interval for each combination is centred in mean value, and its width is twice the value of the standard deviation. For some combinations, which have similar mean value or great variance, the effectiveness intervals result overlapped: for those cases, if the power value falls in this region, both the combinations are considered valid.

Based on this observation, it is possible to manipulate the boundaries of the

QP problem domain. For instance, if 2 HMMs are considered, $M1$ and $M2$, whose power levels are, $M1 = \{70, 0\}$ and $M2 = \{100, 20, 0\}$ respectively, the different combined power levels are $\{0, 20, 70, 90, 100, 170\}$, each one with its own variance value. This example is represented in Figure 6.2. Considering a few different values of aggregated power, e.g., $\overline{y}_t = \{20, 95, 140\}$, it can be observed that $\overline{y}_t = 20$ is obtained as the combination $(x_t^{(1)} = 2, x_t^{(2)} = 2)$, therefore the allowed constraints are defined as:

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \leq \boldsymbol{\xi_t^{(1)}} \leq \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \leq \boldsymbol{\xi_t^{(2)}} \leq \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

If $\overline{y}_t = 95$, the value falls in an overlapped interval, belonging to the combinations $(x_t^{(1)} = 2, x_t^{(2)} = 1)$ and $(x_t^{(1)} = 1, x_t^{(2)} = 2)$, thus, the allowed constraints are defined as:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \leq \boldsymbol{\xi_t^{(1)}} \leq \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \leq \boldsymbol{\xi_t^{(2)}} \leq \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Whereas, if $\overline{y}_t = 140$, no combination is corresponding, thus the boundaries remain as default.

Clearly, the same process can be applied to bound the $\boldsymbol{\beta_t^{(i)}}$ and $\boldsymbol{\phi_t^{(i)}}$. In regard to this, however, since transitions are related to the steady states, the evaluation of the steady states is enough to bound both kinds of variables.

Problem constraints through reactive power disaggregation

Even though disaggregation is aimed for the aggregated power consumption, in most cases the focus is centred on the active power alone. Nonetheless, given the generality of the AFAMAP algorithm, targeting the reactive aggregated power is also possible. In regard to this, in the present work, the application of the AFAMAP algorithm to the aggregated reactive power has been investigated as well, based on the fact that reactive power is a common trait of the power signature of a residential appliances subset.

In the current scenario, the disaggregation of the reactive power samples is carried out, in order to collect additional information about the activity states of the appliances. This information, in turn, is used to further define the lower and the upper boundaries of the states in the active power disaggregation. Similarly to the active power case, the HMMs are modelled for each appliances starting from the signature in the reactive power and the AFAMAP algorithm is run by using the aggregated reactive power signal as input.

Following the basic knowledge in circuit theory, an electrical load with a

reactive component (i.e., an appliance) which has a reactive power consumption greater than 0 is necessary turned on, therefore the boundaries of the problem in active power disaggregation are assigned as follows:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \leq \boldsymbol{\xi_t^{(i)}} \leq \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Although, when the reactive power consumption is 0, the active component could be both null or greater than 0, depending on whether the appliance is turned off or only the load passive component is working. Therefore, the boundaries of the problem in active power disaggregation are setted as default.

AFAMAP Experiments

In this section different aspects of the experimentation are proposed. Among these are the metrics used in the performance evaluation of the proposed disaggregation approach. Also, the simulation scenario and the parameter settings used in the experiments are discussed, as well as the performance of the disaggregation algorithm.

Metrics

Two metrics, that take into account different aspects of disaggregation, are chosen. The first one, namely *state based*, considers the ability of the system to infer the exact state of evolution of each HMM in the model: for the i-th appliance, the multi-class confusion matrix is built by comparing, for each time instant $t = 1, 2, \ldots, T$, the disaggregation variables $\boldsymbol{\xi_t^{(i)}}$ value assumed in the problem solution, with the exact evolution state $x_t^{(i)}$, defined as the *ground truth*. Each class corresponds to a state $j = 1, \ldots, m_i$ of the i-th HMM. Since that the values in $\boldsymbol{\xi_t^{(i)}}$ are not-integral, the computed confusion matrix is soft weighted, similar to the fuzzy-logic [133]. For each class, the Precision $P_j^{(i)}$ and Recall $R_j^{(i)}$ are computed, then the average between the classes evaluates the medium performance for each HMM:

$$P_{S_f}^{(i)} = \frac{1}{m_i} \sum_{j=1}^{m_i} P_j^{(i)}, \quad R_{S_f}^{(i)} = \frac{1}{m_i} \sum_{j=1}^{m_i} R_j^{(i)}. \tag{6.12}$$

The second metric, namely *energy based*, deals with the comparison between the disaggregated profiles and the power consumption at appliance level: as defined by Kolter [119], Recall measures what part of the power consumption has been correctly classified. Precision, on the other hand, measures how much of the power assigned to an appliance truly belonged to that appliance. For

the i-th appliance, the output of the algorithm $\hat{y}^{(i)}_{1:T}$, and the true output $y^{(i)}_{1:T}$, are used to compute the Precision and Recall as:

$$P^{(i)}_{E_f} = \frac{\sum_{t=1}^{T} \min\left(\hat{y}^{(i)}_t, y^{(i)}_t\right)}{\sum_{t=1}^{T} \hat{y}^{(i)}_t}, R^{(i)}_{E_f} = \frac{\sum_{t=1}^{T} \min\left(\hat{y}^{(i)}_t, y^{(i)}_t\right)}{\sum_{t=1}^{T} y^{(i)}_t}. \tag{6.13}$$

Both metrics are evaluated for a signal window w_f with $f = 1, 2, \ldots, F$, therefore to evaluate the performance over the entire dataset the metrics are averaged over the windows:

$$P^{(i)}_{\{S,E\}} = \frac{1}{F}\sum_{f=1}^{F} P^{(i)}_{\{S_f,E_f\}}, R^{(i)}_{\{S,E\}} = \frac{1}{F}\sum_{f=1}^{F} R^{(i)}_{\{S_f,E_f\}}. \tag{6.14}$$

Finally, in order to consider the total performance of the disaggregation system, the average between the appliances is computed:

$$P_{\{S,E\}} = \frac{1}{N}\sum_{i=1}^{N} P^{(i)}_{\{S,E\}}, \quad R_{\{S,E\}} = \frac{1}{N}\sum_{i=1}^{N} R^{(i)}_{\{S,E\}}. \tag{6.15}$$

As unique evaluation metric, the F_1 is chosen and it is calculated as the geometric mean between Precision and Recall:

$$F_{1\{S,E\}} = 2\frac{P_{\{S,E\}} R_{\{S,E\}}}{P_{\{S,E\}} + R_{\{S,E\}}}. \tag{6.16}$$

Set up of the simulation scenario

The dataset used for the experiments is the Almanac of Minutely Power dataset (AMPds) [51]: it contains recordings of consumption profiles belonging to a single home in Canada for a period of two years at 1 minute sampling rate. It provides active and reactive power at appliance level, unlike most of the dataset in which only the active power is provided at appliance level: this information is crucial to test the new approach based on the reactive power disaggregation as constraint.

Analysing the contents of the dataset, it can be noticed that the usage of the appliances is homogeneous throughout the entire period, therefore the experiments are evaluated on 6 months of data, which can be considered representative of the entire dataset.

To create the HMM models of the appliances, the training requires at least one signature per appliance, although multiple signatures lead to a more general model. In the proposed work, a subset of the data, spanning over 14 days, has been deemed sufficient to collect all the signatures required to train all the HMMs.

The HMM are trained in accordance to the Baum-Welch algorithm, after determining the ground truth state over the time: those are obtained through a clustering procedure, in which every cluster represents a power consumption level of the appliance, thus a state of the HMM. This process is achieved using the k-means algorithm, in which the number of the cluster is imposed in a supervised manner, starting from the knowledge of the operating states of the appliance. The power level mean and the variance values are achieved by means of a gaussian variable fitting procedure over the samples belonging to each cluster.

To satisfy the condition of *denoised* system, the aggregated data is synthetically composed by summing the appliance level power signals. The experiments are conducted by using the appliances at higher contribution, therefore 6 appliances have been chosen: dryer, washing machine, dishwasher, fridge, oven, and heat pump.

The simulations are conducted in Matlab environment and the CPLEX solver is used to solve the QP problem.

The value of starting probability $\boldsymbol{\phi_b^{(i)}}$ of the i-th HMM is imposed to assume the certainty for the OFF state for $f = 1$, whereas for the consecutive windows, $1 < f \leq F$, it is imposed to assume the value of the last sample $\boldsymbol{\xi_T^{(i)}}$ of the previous window, in order to ensure the contiguity of the solution on the window border. The value of the ending probability $\boldsymbol{\phi_f^{(i)}}$, instead, is uniformly imposed in every state, since no information from the consecutive window is available.

Different experiments are conducted varying the size of the windows between the values $T \in \{10, 30, 60, 90, 120\}$ minutes, and the effectiveness of the innovative aspect is evaluated: the introduction of the forward term in the cost function, the selection of the boundaries related to the aggregated power level and to the disaggregation output of the reactive power.

The variance parameters are defined with ${\sigma_1}^2 = {\sigma_2}^2 = {\sigma_3}^2 = 0.01$ according to the variance of the experimental data and the regularization parameter $\lambda = 1$.

Results evaluation

The results of the experiments, based on the scenario described in Section 6.2.2, are presented in the current section.

In Figure 6.3, the AFAMAP disaggregated power consumption profiles of the appliances are compared against the corresponding true outputs, provided by the dataset: in the figure a time span of 10 hours, corresponding to 600 samples, is considered. At the bottom, the energy distribution over the same period, expressed among the appliances in terms of percent value, is compared between the reconstructed and the true appliances consumption.

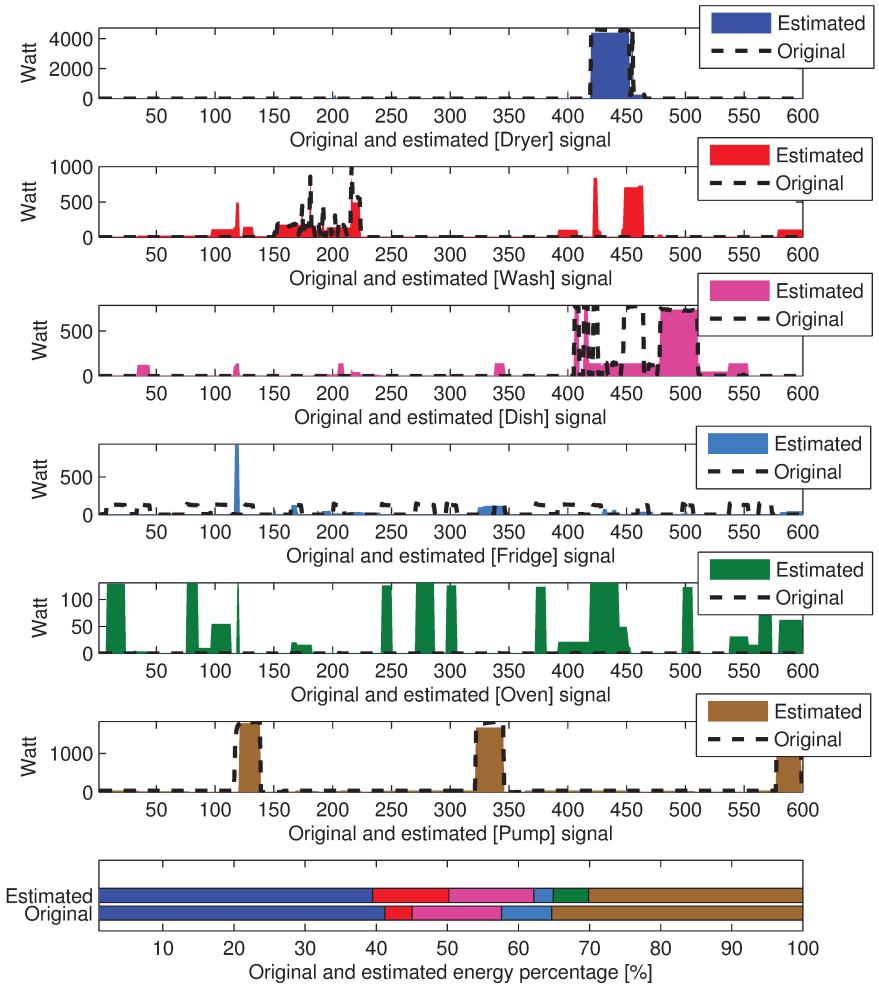
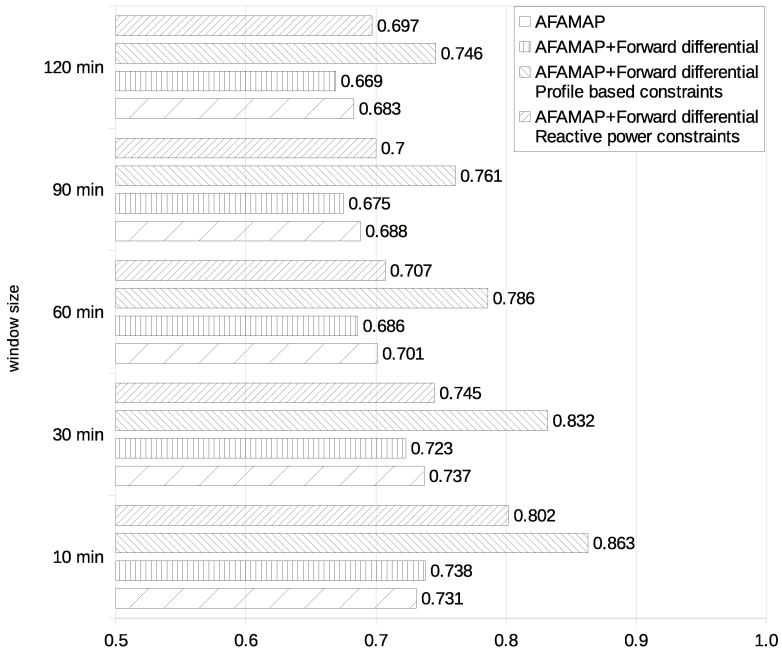
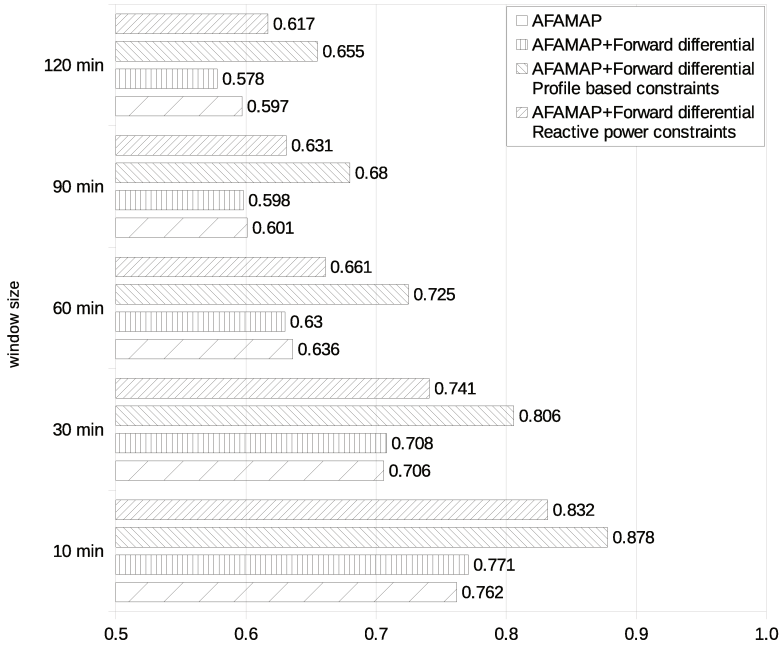


Figure 6.3: Appliances consumption: estimated AFAMAP disaggregation output against original signals.

The signals reveal that the appliances which show an high steady power consumption are easily recognized, whereas the appliances with complex working cycles, or with several power levels, are more difficult to detect. Indeed, whenever several appliances present similar consumption levels, many combinations may satisfy the problem constraints, thus additional information is required to identify the active appliances. For instance, in Figure 6.3, the oven and the fridge are seldom recognised, whereas the detection of the dryer and the washing machine are partially more successful.



(a) State based metric: F_{1S}.



(b) Energy based metric: F_{1E}.

Figure 6.4: Disaggregation performance on AMPds dataset using 6 appliances, with different algorithm configuration.

The evaluation of the algorithm performance is carried out by means of the metrics proposed in Section 6.2.2. Although the focus of the present work is the AFAMAP algorithm, the dataset being used and the proposed training method are different with respect to [119], therefore a direct comparison against the results proposed in the reference work is not possible. To overcome this shortcoming, the baseline has been created anew, by means of the AFAMAP algorithm, the AMPds dataset and the proposed training method.

The disaggregation results computed by means of the metrics are reported in Figure 6.4: in Figure 6.4(a) the state based metric is presented, whereas the energy based metric is proposed in Figure 6.4(b). The results are shown for different values of the time window length. Clearly, since all the results exceed 0.5, the plots have been drawn from 0.5 onwards.

Both plots show that the best results are achieved using the shortest time window. On a different note, however, not every configuration improves in the same way.

Focusing on the state based metrics, it is possible to observe that the AFAMAP baseline shows a significant performance improvement with the decreasing of the window length, except when passing from the 30 to 10 minutes window size. On the contrary, the forward differential model gives an improvements at the shorter window size, resulting in the best performance in the unbounded problem solution, with a F_{1S} of 0.738 and an improvement of 1% respect to the baseline.

Fixing the boundaries of the problem, in every simulation case, gives the benefit on the disaggregation results: the profile based method gives a considerable performance improvements with every window size, but the highest relative improvement can be noted at the smallest size, resulting to a F_{1S} of 0.863 and a relative improvement of 18%.

Alternatively, the boundaries can be setted based on the reactive power disaggregation feedback: the results, showed in Table 6.7, demonstrate that the reactive power reaches high performance in disaggregation. This is due to the high difference in the reactive components of each appliance, which involves a strong distinction in the creation of the HMM, therefore allowing an highly reliable disaggregation. The usage of this information results in a performance improvement for every window size, more considerable at the smallest size: in general, the usage of the reactive power feedback gives benefits to the disaggregation, with a F_{1S} of 0.802 and a relative improvement of 9.7%, therefore less than the profile based constraints.

Clearly, the same trends presented about the state based metrics still hold true when evaluating with the energy based metrics. The most notably difference between the two plot, in fact, is that the rate of improvement of the algorithms when decreasing the time window length:

indeed, the forward differential model introduction results to a F_{1E} of 0.771 and a relative improvement of 1.2% respect to the baseline, whereas the profile based setting of the boundaries results to a F_{1E} of 0.878 with a relative improvement of 15.2% and the reactive power based method to a F_{1E} of 0.832 with an improvement of 9.2%.

The forward differential model seems to be beneficial only with the shortest time window: it may be a direct consequence of the problem formulation alteration. Indeed, the introduction of additional variables increases the size of the problem, therefore the computational burden, for which the solver demonstrates worst performance, as it happens for the baseline approach with larger window size.

Table 6.7: Disaggregation results on reactive power. The configuration used is: AFAMAP + Forward differential.

	Window size				
Metric	10 min	30 min	60 min	90 min	120 min
State based: F_{1S}	0.922	0.877	0.869	0.867	0.865
Energy based: F_{1E}	0.935	0.883	0.877	0.875	0.874

Despite this, the improvements achieved adding the differential forward information to the model are restricted to the application scenario: since the algorithm operates on a per-sample basis, for each appliance behaviour two state changes unlikely happen across three contiguous samples, thus the forward difference cannot provide a substantial support to the inference of the actual working state.

The errors in the disaggregation phase are caused by the multiplicity of states combinations which can correspond to the same value of the aggregated data: for this reason the use of boundaries allows to exclude some solutions that are not eligible, therefore facilitates the solver to find the exact solution to the problem. Nevertheless, the variation over time of the power consumption associated to a specific appliance working state, causes an unwanted variability, i.e., a noise component, in the achieved solution.

6.2.3 DNN

In this section, the method proposed by Kelly and Knottenbelt [131] has been re-evaluated by adopting the Energy-based metric proposed in Section 6.2.2, then, further experiments have been performed by integrating the early stopping regularization and exploring a wider parameters range, respect to original one.

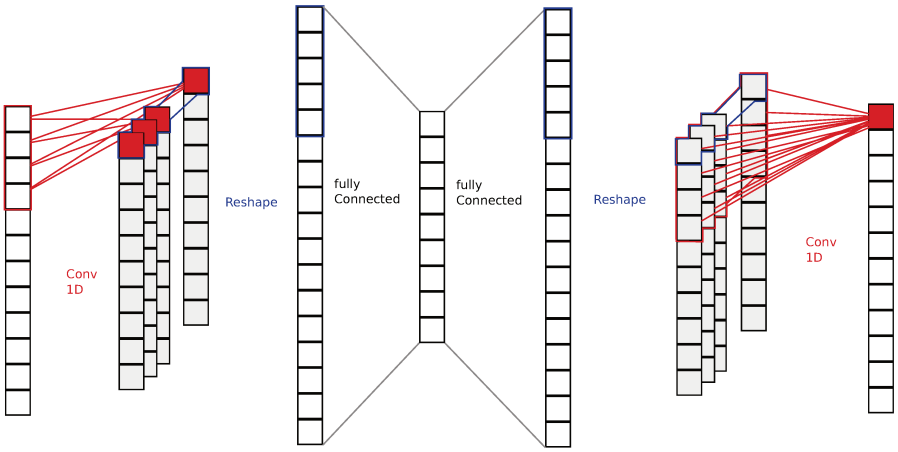


Figure 6.5: Adopted DNN architecture.

In the method proposed by Kelly and Knottenbelt [131], the energy disaggregation task is framed as a "denoising" problem. In particular, the energy disaggregation can be viewed as an attempt to extract the power signal of the target appliance (clean signal) from the aggregate signal (noisy background) resulting from the sum of other appliances contribution. Among the state-of-the-art approaches, the Denoising Autoencoder [62] has been chosen to carry out the "denoising" task.

Among the proposed net typologies, the one that includes the Convolutional Neural Network (CNNs) [134, 135, 136] seems to achieve notable performance. The basic idea is "to extract a small number of low level features with a small receptive fields across the entire input. All of our nets use at least one 1D convolutional layer at the input."

The aims of the proposed approach is to allow the network to learn "low level feature detectors". On purpose, the first and last layers of the proposed DAE are 1D convolutional layers. Specifically, the input layer equality applies the learned detection rule across the entire input window, whereas the output one operates as a "deconvolution". The whole network structure is the following:

1. Input (dimension=L).

2. 1D conv (number of filters=N, filter size=S, activation function=linear, border mode=valid).

3. Fully connected (number of neurons=(L - (S-1)) $\times$ N, activation function=ReLU).

4. Hidden layers fully connected (number of neurons=H; activation function=ReLU).

5. Fully connected (number of neurons=(L - (S-1)) $\times$ N, activation function=ReLU).

6. 1D conv (number of filters=1, filter size=S, activation function=linear, border mode=valid).

where L denotes the length determined by appliance duration, and N, S, H represent the number of filters (Kernels number), the filter size (Kernel size), and the number of neurons of the hidden layer, respectively.

As stated in Section 3.1.3, the code layer is represented by the layer 4.

In the training phase, the power demand of the appliance has been adopted as the target data, whereas the aggregate power has been adopted as input data for all the appliances. Moreover, for each appliance a different input/output window size has been chosen. Specifically, a window width able to capture the whole appliance working cycle, and minimize the useless data, has been chosen.

In the test phase, the aggregated data is split in overlapping window, by means of a sliding window, thus each network produces multiple window outputs, one for each time step, with multiple overlapped data characterized by the same timestamp. In the original approach, for each appliance, the output signal is recreated as average of the values marked with the same timestamp for each overlapped window.

In the proposed contribution, the appliance output signal is reconstructed by means of a median filter applied to the set of values marked with the same timestamp for each overlapped window. Specifically, perform the average of the set values can produce a dangerous attenuation or intensification of the recreated power demand of the specific appliance, especially with high variability outputs. Therefore, the 1D filter returns the median value that actually corresponds with a real power demand, and not an attenuated/intensified one.

DNN Experiments

The UK-DALE dataset [137] has been adopted to test the system performance, and, from the data collected in five houses, five target appliances have been chosen: the fridge, the washing machine, the dish washer, the kettle, and the

microwave. The track for each appliance has been extracted by means of the tool NILMTK [138].

The nets have been trained with both synthetic aggregate data and real aggregate data in a 50:50 ratio, and each network has been trained on one appliance.

With respect to [131], a wider experimental campaign has been also carried out. First of all, further network parameters have been explored, by assuming the following ranges:

$$\begin{aligned} \textit{Kernels number } (N) &= \{2, 4, 8, 16, 32\}, \\ \textit{Kernels size } (S) &= \{2, 4, 8, 16, 32\}, \\ \textit{Hidden layer size } (H) &= \{8, 16, 32, 128, 256, 512, 1024, 2048, 4096\}. \end{aligned}$$

Moreover, the early stopping approach has been applied in the training phase and the F score has been assumed as evaluation criteria for the validation set. After preliminary evaluations, the early stopping epochs have been set to 10000, hence the training phase ends if no improvement is obtained for 10000 consecutive epochs. The net that achieves the last best performance is saved and adopted in the test phase.

Table 6.8: Results for "seen" data evaluations, in comparison with reference net [131]. The "Best net" reports the triple N-S-H, representing the number of kernels, the kernel size, and the hidden layer size, respectively.

Baseline evaluations						
Metric	Kettle	Wash	Dish	Fridge	Microwave	AVG
Precision	0,911	0,627	0,788	0,912	0,628	0,773
Recall	0,715	0,243	0,462	0,494	0,565	0,496
F score	0,801	0,350	**0,582**	0,641	**0,595**	**0,604**
Best net	8-4-128					

New evaluations						
Metric	Kettle	Wash	Dish	Fridge	Microwave	AVG
Precision	0,748	0,431	0,666	0,798	0,362	0,601
Recall	0,865	0,375	0,431	0,604	0,587	0,572
F score	**0,802**	**0,401**	0,523	**0,688**	0,448	0,572
Best net	32-4-512	16-16-256	16-16-1024	32-16-2048	64-16-1024	

A slight variation of the Energy-based version of Precision and Recall, have been adopted as evaluation criteria. Differently from Section 6.2.2, Precision and Recall values have been computed by applying the definitions in (6.13) to the whole sequence of the reconstructed signal, without perform the average

among the windows, as defined in (6.14).

All the experiments have been performed on the CINECA HPC, Galileo [5], by using about 40 NVIDIA k80 GPUs with 24 GB GDDR5.

As carried out in the original contribution, the tests have been executed for two different scenarios: "seen" and "unseen". In the "seen" scenario, the tests have been performed against the appliances of a house seen during the training, and the corresponding results are reported in Table 6.8. On the contrary, in the "unseen" scenario the tests have been executed on appliances of a house unseen during the training, and the corresponding results are reported in Table 6.9.

First of all, the best nets achieved in the original contribution have been re-tested adopting the proposed evaluation criteria, in order to obtain a baseline reference. The achieved results are marked as "Baseline evaluations" in tables.

In the "seen" evaluations no better overall result has been achieved, with respect to the reference contribution, by introducing the proposed improvements. However, better results have been obtained for three appliances.

On the contrary, in the "unseen" case better overall results have been achieved by introducing the proposed improvements. Specifically, four appliances have reached higher results and all the best results have been achieved with different network parameters, namely larger networks.

Table 6.9: Results for "unseen" data evaluations, in comparison with reference net [131]. The "Best net" reports the triple N-S-H, representing the number of kernels, the kernel size, and the hidden layer size, respectively.

Baseline evaluations						
Metric	Kettle	Wash	Dish	Fridge	Microwave	AVG
Precision	0,919	0,247	0,909	0,901	0,448	0,685
Recall	0,694	0,221	0,250	0,508	0,053	0,345
F score	0,791	0,233	**0,392**	0,650	0,095	0,459
Best net	8-4-128					

New evaluations						
Metric	Kettle	Wash	Dish	Fridge	Microwave	AVG
Precision	0,777	0,304	0,507	0,805	0,134	0,505
Recall	0,865	0,351	0,314	0,642	0,965	0,627
F score	**0,819**	**0,326**	0,388	**0,714**	**0,235**	**0,496**
Best net	16-8-128	8-8-256	16-16-1024	32-16-2048	16-64-4096	

On the one hand, in the "seen" case, the washing machine and the fridge have performed significantly better with respect to baseline approach. Whereas, for both dish washer and microwave the introduction of the early stopping seems

[5] `http://www.hpc.cineca.it/hardware/galileo`

to have caused some issue. The performance degradation mainly affects the precision criterion.

On the other hand, for the "unseen" evaluations, remarkable improvements have been achieved for all the appliances, with the exception of the dish washer.

6.2.4 Remarks

The AFAMAP algorithm [119], resulting one of the most performing and computationally efficient among the evaluated approaches, has been revised in order to improve its performance through a more exhaustive exploitation of the information pertaining the appliance activity.

The proposed algorithm exploits both additive and differential FHMM to model the activity of the appliances. At each time step, the best combination of appliances working state is chosen to represent the actual aggregated consumption: as result of the optimisation process, a set of coefficients are returned to weight the appliances working state and compose the own disaggregated consumption.

The revised algorithm, however, has taken into account additional elements. In regards to the FHMM model, a forward differential model has been paired to the reference backward differential FHMM: as enhancement to this technique, an observation window of longer duration could be introduced in the differential model.

In addition, the use of solver boundaries has been explored: firstly, the setting has been related to the admissible state combination of the aggregated power; alternatively, the reactive power disaggregation output has been used to select the boundaries, endorsing the heterogeneous data usage effectiveness.

Concerning the DNN, an initial assessment to the application of neural network approach for the NILM problem has been proposed. Promising results have been achieved by adopting the autoencoder configuration that combines Multilayer Perceptron and Convolutional Neural Network. This contribution is an ongoing project, and further investigations are deemed necessary.

Chapter 7

Conclusions

In this dissertation, the Computational Intelligence approaches for the improvement of the water and natural gas grids, to be regarded as Smart Grids, have been studied. Specifically, the approaches have concerned the consumption forecasting and the leakage detection in different grid scenarios. The forecasting tests have been focused on homogeneous evaluation criteria among the different techniques, and the application of heterogeneous information. The evaluated scenarios have gone from the residential scale grids to their national counterparts. Concerning the leakage detection experiments, the novelty detection paradigm has been exploited, and evaluations with different techniques have been executed. In addition, in order to identify a suitable set of features, a SFS has been proposed and performed during the evaluation. The target scenarios have been limited to the residential and the office building one.

In addition, advancements in two specific applications for Smart Electrical Grids have been also proposed in accordance with the methodologies adopted in the investigations conducted for smart water and natural gas grids solutions. Specifically, an energy manager for the micro-grid scenario has been evaluated focusing of different price forecasting techniques, and a set of improvements have been proposed for state-of-the-art NILM solutions.

The state of the art of Computational Intelligence techniques, specifically for load forecasting, leakage detection, and datasets for water and natural gas grids, have been discussed in Chapter 2. Moreover, the problem statement concerning these applications has been also presented.

Concerning the techniques, it has been difficult to make an objective comparison among the methods presented in literature. Firstly, the approaches have to be tested using a common database. Secondly, for both water and gas references, there is a significant dissimilarity among the evaluation criteria adopted for each proposed method.

Regarding the datasets, further information have been requested to the authors, for both the gathered data in the dataset (e.g. sample rate and type) and the dataset availability. Moreover, the survey has highlighted two main shortcomings: the datasets either are not publicly available or lack of appropriate

data.

The Computation Intelligence approaches, adopted in both load forecasting and leakage detection algorithms, have been introduced in Chapter 3.

Chapter 4 has been devoted to evaluate the performance of the following techniques: ARIMA, ANN, DBN, ESN, SVR, GP, EKF-GP, and ELM. Each available dataset has been tested with each technique for different configurations and input variables. Furthermore, common evaluation criteria have been selected.

Concerning the residential scenario, within the AMPds dataset [51], a correlation between gas consumption and temperature information has been highlighted in the natural gas forecasting, and the results have confirmed the advantage of heterogeneous data utilization. Together with SVR, ELM and ANN, have confirmed to be suitable for shot-term prediction.

The ANN has achieved the best results for all the office building sub-sets, DFID dataset [49], and similar results have been performed by ELM and SVR. Unfortunately, the actual sets of the DFID has not allowed the evaluation with heterogeneous information.

The ANN and EKF-GP have proved to be the promising techniques to be applied to the national consumption levels recorded in the EIA datasets [50], and for two states out of three the combination of consumption data and temperature information have achieved better performance than the only use of water consumption information.

For the city scenario relating to the Tehran dataset [17], the best performance has been achieved again with SVR.

In Chapter 5 a novelty detection based approach has been proposed for the identification of leakage occurrences in smart water and gas grids. The validity of the presented algorithm has been confirmed by the achieved results, and its suitability for both domestic and office building consumption has been proven using the AMPds and the DFID datasets [51, 49].

GMM, HMM, and OC-SVM have been adopted, under a comparative perspective, to model the normality background by assuming different parameters. An innovative set of features, actually missing in the related literature and specifically defined for the application scenarios under study, has been evaluated. In order to identity the best features combinations, a slight variation of the SFS has been proposed and performed.

In addition, among the proposed features, two types of temporal information have been also introduced, namely FPW and TWE, and the usefulness of these features, especially for low sampling data rates, has been confirmed.

Later, pressure information has been exploited to improve the performance of the automatic leakage detection approach. The introduction of pressure information has allowed to adopt an approach with lower computational burden,

and to detect leakages even for low resolution. Therefore, the usability of the proposed leakage detection approach in low-power devices/meters [4] has been improved. Noteworthy, preliminary evaluations have been also performed by exploiting the Autoencoder Neural Network.

In the first part of Chapter 6, the performance of a micro-grid energy management, with the support of dynamic pricing strategies, has been investigated, based on a real-life environment which served as a reference model. In fact, although not available in Italy, yet, dynamic pricing has much potential to improve the flexibility and responsiveness of the grids, as well as to lower energy costs to the user. To investigate such a scenario, the support to energy management by means of forecasts has been evaluated, and the results have confirmed that hourly planning cannot take real advantage of price forecasting, whereas it can exploit local energy production forecasting. The daily planning, on the other hand, is more effective in increasing the energy efficiency of the system, although the drop in performance is more variable depending on the configuration. An analysis of the scheduling process has led to the conclusion that the price forecasting error affects the scheduling process through the displacement of the price minima within the profile.

The second part of Chapter 6 has been devoted to study and improve some state-of-the-art solutions for the NILM.

At first, a review of unsupervised NILM algorithms has been presented, together with an updated list of available datasets which are typically used for parameter tuning and evaluation purposes. The motivation behind this review is given by the fact that unsupervised solutions are better suited, with respect to their supervised counterparts, to deal with real applicative contexts where the user intervention has to be minimized.

Later the AFAMAP algorithm [119], one of the most performing and computationally efficient in literature, has been revised in order to improve its performance through a more exhaustive exploitation of the information pertaining the appliance activity. The proposed algorithm exploits both additive and differential FHMM to model the activity of the appliances.

Finally, improvements have been introduced in the DNN approach [131], and a wider experimental campaign has been performed with respect to the original one. The introduction of the early stopping in the training process and the application of the median filter in the output signal reconstruction, have allowed to improve the algorithm overall performance.

7.1 Future Research Topics

The author is confident that the spread of innovative monitoring systems, which are more and more often based on low-power wireless devices [139] [140], will

ensure a facilitation for collecting large amount of data, as partially shown in [46], and creating suitable and publicly available databases. This will enable a fair and comprehensive evaluation of already proposed Computational Intelligence solutions for smart water and gas grids, as well as the development of new ideas from the research community.

For this reason, future works could be oriented toward the collection of more suitable data in order to create a suitable dataset of water and natural gas for heterogeneity evaluations.

Concerning the forecasting for water and natural gas, a proper study of the seasonality issue could be conducted for all the dataset longer the few years. Additionally, a further in-depth analysis is needed to find optimal common technique and parameters.

Moreover, concerning the leakage detection, additional efforts could be oriented into the evaluation of smaller leakages and other types of anomalies, such metering system issues and faults, along with the employment of new features and TDR/FDR sensitivity analysis from noise into sensor data.

In reference to the energy grid, as future works concerning the AFAMAP solution, a more reliable appliance model will be considered in order to improve the representation of additional working states, e.g., the usage of Gaussian Mixture Model (GMM) within the HMM.

Furthermore, additional information about the working states duration will be introduced, to fully exploit the differential model.

Moving to the energy manager, the price forecasting error affects the scheduling process through the displacement of the price minima within the profile. For this reason, in the near future, this aspect of the problem will be investigated. Therefore, an index aimed to characterize the forecasting error focusing on the minima displacement will be devised, and also the scheduling performance based on different dynamic pricing profiles will be evaluated.

1 AMPds Results - 1 year

Table 1: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the ANN approach is reported. The "Par." column indicates the corresponding parameters: lags number and neurons number of the hidden layer, respectively.

ANN	NMSE	R^2	MSE	MAPE	RRMSE	Par.	NMSE	R^2	MSE	MAPE	RRMSE	Par.
	Natural Gas Prediction											
	1 *h*						6 *h*					
G	0.740	0.260	0.0077	93.453	134.660	3-6	0.316	0.684	0.0102	39.451	58.815	5-8
GT	0.687	0.312	0.0074	87.037	129.090	5-8	0.280	0.719	0.0091	39.421	55.667	3-9
GE	0.726	0.274	0.0078	89.382	130.860	3-5	0.311	0.688	0.0101	41.965	58.698	5-9
GET	0.693	0.307	0.0076	86.729	129.980	3-6	**0.276**	**0.723**	**0.0095**	**37.792**	**54.115**	**5-5**
WG	0.713	0.286	0.0077	87.331	131.170	3-10	0.354	0.645	0.0120	39.583	58.046	5-5
WGT	0.693	0.307	0.0071	88.026	130.010	2-5	0.292	0.707	0.0094	39.355	53.406	3-6
WGE	0.717	0.282	0.0074	89.745	132.760	5-5	0.351	0.648	0.0122	41.829	59.928	5-5
WGET	**0.686**	**0.314**	**0.0071**	**84.971**	**127.300**	**2-6**	0.284	0.715	0.0074	40.798	54.964	5-5
	12 *h*						24 *h*					
G	0.262	0.737	0.0124	31.315	42.878	2-12	**0.176**	**0.823**	**0.0086**	**24.979**	**37.116**	**2-8**
GT	0.203	0.796	0.0117	28.398	38.398	3-6	0.187	0.811	0.0097	25.273	39.577	2-6
GE	0.284	0.715	0.0175	33.247	48.300	2-9	0.248	0.750	0.0126	29.467	39.884	2-6
GET	0.225	0.774	0.0120	28.429	39.750	2-6	0.238	0.759	0.0127	29.491	44.450	3-5
WG	0.233	0.766	0.0131	31.392	42.768	2-8	0.222	0.776	0.0115	27.780	44.186	2-6
WGT	0.205	0.794	0.0103	24.478	40.401	2-5	0.241	0.757	0.0137	29.289	45.126	2-6
WGE	0.286	0.713	0.0138	33.619	47.687	2-6	0.339	0.658	0.0193	31.484	46.133	3-5
WGET	**0.200**	**0.799**	**0.0103**	**28.977**	**39.302**	**2-5**	0.305	0.692	0.0171	32.382	42.238	3-5
	Water Prediction											
	1 *h*						6 *h*					
W	0.802	0.198	0.0031	87.656	142.930	5-10	0.357	0.643	0.0051	39.078	56.075	5-5
WT	0.768	0.232	0.0030	91.108	140.900	2-9	0.352	0.647	0.0050	41.574	59.395	5-6
WE	0.798	0.202	0.0030	91.030	139.750	3-6	0.406	0.594	0.0058	43.142	60.654	5-5
WET	**0.747**	**0.253**	**0.0029**	**90.803**	**142.660**	**3-8**	0.356	0.643	0.0054	38.677	56.914	5-5
WG	0.789	0.210	0.0031	91.240	142.840	2-9	0.414	0.585	0.0061	43.086	60.282	5-5
WGT	0.791	0.208	0.0032	87.921	142.900	5-6	**0.344**	**0.656**	**0.0051**	**42.280**	**58.767**	**5-5**
WGE	0.787	0.213	0.0029	91.855	140.890	3-6	0.405	0.594	0.0055	41.422	56.859	5-6
WGET	0.793	0.206	0.0030	89.450	139.600	2-9	0.395	0.604	0.0056	43.110	61.760	5-5
	12 *h*						24 *h*					
W	0.426	0.572	0.0057	27.048	34.335	3-8	**0.345**	**0.652**	**0.0052**	**19.388**	**25.185**	**2-5**
WT	0.462	0.536	0.0063	30.908	39.709	2-10	0.462	0.534	0.0081	25.385	33.498	2-6
WE	**0.408**	**0.590**	**0.0059**	**27.404**	**34.153**	**5-5**	0.464	0.531	0.0082	26.915	33.582	2-6
WET	0.462	0.536	0.0063	28.853	37.336	2-5	0.546	0.449	0.0083	24.928	31.749	2-5
WG	0.456	0.541	0.0068	30.087	39.283	3-7	0.421	0.575	0.0074	24.718	31.992	2-6
WGT	0.514	0.484	0.0080	30.465	40.607	2-5	0.519	0.476	0.0085	24.248	30.324	2-6
WGE	0.551	0.447	0.0080	31.681	40.524	2-6	0.503	0.492	0.0083	23.256	29.843	2-6
WGET	0.477	0.521	0.0068	30.581	39.142	2-5	0.642	0.352	0.0104	27.012	34.767	2-7

W = Water G = Natural Gas E = Electric Power T = Temperature

Table 2: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the DBN approach is reported. The "Par." column indicates the corresponding best parameters: lags, neurons, max. iteration and batch size, respectively.

DBN	NMSE	R^2	MSE	MAPE	RRMSE	Par.	NMSE	R^2	MSE	MAPE	RRMSE	Par.
	Natural Gas Prediction											
	1 h						*6 h*					
G	0.857	0.143	0.0098	108.720	148.160	2-2-300-100	0.466	0.533	0.0151	54.207	70.913	5-4-300-100
GT	0.792	0.208	0.0090	102.760	142.400	2-2-300-100	0.381	0.618	0.0123	49.026	64.125	5-9-500-100
GE	0.853	0.147	0.0097	108.050	147.790	2-2-300-100	0.407	0.592	0.0132	51.142	66.273	5-2-300-10
GET	0.792	0.207	0.0090	102.930	142.460	2-2-300-100	**0.340**	**0.660**	**0.0110**	**44.432**	**60.558**	**5-2-300-10**
WG	0.846	0.154	0.0097	108.100	147.180	2-5-100-100	0.432	0.567	0.0140	53.263	68.339	5-7-100-100
WGT	**0.785**	**0.215**	**0.0080**	**96.882**	**138.240**	**5-5-300-100**	0.379	0.620	0.0123	47.727	64.016	5-9-500-100
WGE	0.835	0.164	0.0095	104.970	146.290	2-5-300-100	0.406	0.593	0.0132	50.987	66.251	5-9-300-100
WGET	0.786	0.214	0.0090	101.510	141.900	2-2-300-100	0.359	0.640	0.0116	45.138	62.285	5-2-300-10
	12 h						*24 h*					
G	0.278	0.720	0.0157	34.182	48.017	5-9-300-100	0.263	0.734	0.0102	28.062	43.706	5-2-500-10
GT	0.210	0.789	0.0118	30.778	41.677	5-4-300-10	**0.199**	**0.800**	**0.0077**	**26.352**	**37.954**	**5-8-500-10**
GE	0.264	0.735	0.0149	34.232	46.765	5-10-100-100	0.269	0.729	0.0104	27.727	44.148	5-7-100-10
GET	0.202	0.797	0.0114	30.649	40.936	5-2-500-10	0.257	0.740	0.0099	28.475	43.222	5-7-500-10
WG	0.262	0.737	0.0148	34.320	46.577	5-10-100-10	0.276	0.721	0.0107	27.899	44.756	5-9-300-10
WGT	0.209	0.790	0.0118	30.772	41.583	5-3-300-10	0.236	0.762	0.0091	26.988	41.360	5-5-500-10
WGE	0.252	0.747	0.0142	33.232	45.688	5-3-300-10	0.276	0.721	0.0107	30.527	44.786	5-6-300-100
WGET	**0.201**	**0.798**	**0.0113**	**30.207**	**40.802**	**5-4-300-10**	0.266	0.732	0.0103	28.476	43.905	5-6-300-10
	Water Prediction											
	1 h						*6 h*					
W	0.922	0.078	0.0035	104.600	157.240	3-2-300-100	0.457	0.542	0.0062	50.062	65.320	5-3-100-10
WT	0.903	0.097	0.0035	103.510	155.620	3-2-500-100	0.440	0.559	0.0060	49.574	64.131	5-9-100-100
WE	0.891	0.109	0.0032	103.710	147.150	2-4-100-100	0.448	0.551	0.0061	49.593	64.678	5-10-100-10
WET	0.881	0.119	0.0031	103.430	146.350	2-7-100-100	0.433	0.566	0.0059	49.243	63.578	5-10-300-100
WG	0.913	0.087	0.0033	103.290	148.970	2-5-100-100	0.427	0.572	0.0058	49.118	63.185	5-2-300-10
WGT	0.900	0.099	0.0032	97.901	147.270	5-9-300-100	0.422	0.577	0.0057	46.624	62.774	5-2-500-10
WGE	0.876	0.124	0.0031	101.330	145.950	2-7-100-100	0.429	0.570	0.0058	49.103	63.332	5-2-300-10
WGET	**0.876**	**0.124**	**0.0031**	**102.420**	**145.930**	**2-9-100-100**	**0.401**	**0.598**	**0.0055**	**45.885**	**61.243**	**5-2-500-10**
	12 h						*24 h*					
W	0.585	0.413	0.0094	33.395	46.114	5-3-300-100	0.465	0.531	0.0085	24.068	30.477	2-2-100-10
WT	**0.573**	**0.424**	**0.0093**	**32.570**	**45.668**	**5-7-500-10**	0.469	0.527	0.0089	25.146	32.101	5-3-500-100
WE	0.594	0.404	0.0096	33.214	46.465	5-10-300-100	0.457	0.538	0.0087	25.104	31.699	5-4-500-10
WET	0.584	0.414	0.0094	32.611	46.069	5-8-500-100	0.425	0.571	0.0081	23.658	30.558	5-5-500-10
WG	0.583	0.415	0.0094	32.984	46.033	5-9-300-100	0.451	0.545	0.0082	23.082	30.016	2-6 -100-100
WGT	0.578	0.420	0.0093	32.272	45.839	5-8-500-100	0.431	0.565	0.0082	24.132	30.764	5-2-500-100
WGE	0.592	0.406	0.0095	33.073	46.390	5-10-300-100	0.417	0.579	0.0076	21.783	28.877	2-5-500-10
WGET	0.583	0.414	0.0094	32.349	46.042	5-10-500-100	**0.415**	**0.581**	**0.0079**	**24.056**	**30.212**	**5-5-500-10**

W = Water G = Natural Gas E = Electric Power T = Temperature

Table 3: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the ESN approach is reported. The "Par." column indicates the corresponding best parameters: lags, reservoir size, initial transient, α and β, respectively.

ESN	NMSE	R^2	MSE	MAPE	RRMSE	Par.	NMSE	R^2	MSE	MAPE	RRMSE	Par.
	Natural Gas Prediction											
	1 *h*						6 *h*					
G	0.769	0.231	0.0088	97.529	140.320	2-25-15-1/4-1e-4	0.413	0.587	0.0134	48.024	66.747	5-40-4-1-1e-2
GT	**0.740**	**0.260**	**0.0084**	**94.901**	**137.680**	**2-30-8-1/3-1e-3**	0.351	0.648	0.0114	44.504	61.578	5-30-10-1/2-1e-14
GE	0.771	0.229	0.0088	98.683	140.510	2-25-15-1/10-1e-13	0.369	0.630	0.0119	46.123	63.135	5-25-12-1/2-1e-2
GET	0.744	0.256	0.0085	95.173	138.050	2-40-0-1/3-1e-3	**0.333**	**0.666**	**0.0108**	**43.967**	**59.987**	**5-25-0-1-1e-2**
WG	0.762	0.238	0.0087	94.846	139.700	2-20-8-1/4-1e-12	0.386	0.613	0.0125	4 7.752	64.534	5-40-0-1/3-1e-2
WGT	0.742	0.257	0.0085	94.955	137.890	2-40-2-1/2-1e-2	0.353	0.646	0.0114	44.360	61.740	5-25-6-1/2-1e-1
WGE	0.773	0.227	0.0088	96.192	140.680	2-30-15-1/10-1e-15	0.361	0.638	0.0117	44.781	62.447	5-25-9-1/2-1e-15
WGET	0.744	0.256	0.0085	96.509	138.050	2-15-12-1/3-1e-2	0.337	0.662	0.0109	43.429	60.345	5-35-9-1/2-1e-1
	12 *h*						24 *h*					
G	0.262	0.736	0.0148	33.016	46.615	5-15-0-1/2-0	0.247	0.751	0.0095	28.442	42.315	5-25-0-1/4-1e-15
GT	0.229	0.770	0.0129	31.953	43.587	5-20-0-1-1e-15	**0.231**	**0.767**	**0.0099**	**26.675**	**39.157**	**3-25-3-1/6-1e-6**
GE	0.240	0.759	0.0136	31.903	44.607	5-35-5-1/2-1e-2	0.244	0.753	0.0094	29.233	42.105	5-20-1-1/4-1e-3
GET	0.221	0.778	0.0125	31.046	42.817	5-40-5-1-1	0.247	0.751	0.0095	28.892	42.343	5-30-7-1/5-1e-3
WG	0.239	0.759	0.0135	33.041	44.535	5-25-3-1/4-1e-2	0.257	0.740	0.0099	27.353	43.208	5-30-6-1-1
WGT	0.219	0.780	0.0124	31.731	42.598	5-30-7-1/2-1e-1	0.250	0.748	0.0096	27.938	42.552	5-30-6-1/4-1e-2
WGE	0.234	0.765	0.0132	33.215	44.052	5-15-4-1/2-1e-2	0.246	0.752	0.0095	30.274	42.228	5-10-6-1/2-1e-3
WGET	**0.213**	**0.786**	**0.0120**	**31.457**	**41.999**	**5-25-1-1/2-1e-13**	0.250	0.747	0.0097	28.478	42.618	5-40-2-1/2-1e-2
	Water Prediction											
	1 *h*						6 *h*					
W	0.831	0.169	0.0030	94.671	142.150	2-20-10-1-1e-2	**0.367**	**0.632**	**0.0050**	**42.800**	**58.584**	**5-35-8-1/2-1e-4**
WT	0.830	0.170	0.0030	92.302	142.010	2-40-7-1/2-1e-3	0.392	0.607	0.0053	44.035	60.508	5-10-0-1/4-1e-14
WE	0.820	0.179	0.0029	93.703	141.220	2-10-9-1-1e-2	0.379	0.620	0.0052	43.386	59.533	5-25-12-1/2-1e-2
WET	**0.815**	**0.184**	**0.0029**	**93.042**	**140.780**	**2-2-0-1-1e-3**	0.381	0.618	0.0052	43.736	59.646	5-25-1-1/9-11
WG	0.823	0.177	0.0029	94.753	141.460	2-10-10-1/2-1e-13	0.397	0.602	0.0054	43.914	60.917	5-40-12-1/3-1e-1
WGT	0.832	0.168	0.0030	92.937	142.190	2-30-15-1/2-0	0.402	0.598	0.0055	44.766	61.250	5-30-10-1/10-1
WGE	0.820	0.180	0.0029	94.423	141.210	2-25-13-1-1e-12	0.384	0.615	0.0052	42.602	59.912	5-25-10-1/10-1e-2
WGET	0.821	0.179	0.0029	94.378	141.250	2-30-2-1/2-1e-1	0.385	0.614	0.0052	42.582	59.998	5-40-14-1/10-1
	12 *h*						24 *h*					
W	**0.522**	**0.475**	**0.0084**	**29.638**	**43.586**	**5-40-10-1/2-1e-4**	0.371	0.626	0.0068	19.984	27.234	2-35-13-1/4-1e-5
WT	0.538	0.459	0.0087	30.634	44.252	5-20-12-1/2-1e-2	0.344	0.653	0.0063	20.514	26.235	2-30-15-1/2-1e-3
WE	0.534	0.464	0.0086	30.552	44.063	5-30-4-1/2-1e-2	0.350	0.646	0.0064	20.018	26.466	2-25-10-1/6-1e-4
WET	0.547	0.451	0.0088	31.131	44.583	5-20-12-1-1e-1	0.358	0.639	0.0066	20.532	26.759	2-15-0-1/4-1e-4
WG	0.547	0.450	0.0088	31.470	44.609	5-25-2-1/2-1e-1	0.385	0.611	0.0071	20.895	27.759	2-20-2-1/2-1e-13
WGT	0.552	0.445	0.0089	31.019	44.810	5-25-14-1/2-1	**0.334**	**0.663**	**0.0061**	**19.572**	**25.830**	**2-15-2-1/4-1e-14**
WGE	0.549	0.448	0.0089	31.103	44.694	5-25-5-1/2-1e-2	0.356	0.640	0.0065	20.066	26.691	2-30-10-1/3-1e-3
WGET	0.548	0.449	0.0088	31.058	44.646	5-30-10-1/2-1e-2	0.363	0.633	0.0066	20.015	26.947	2-25-13-1/3-1e-2

W = Water G = Natural Gas E = Electric Power T = Temperature

Table 4: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the SVR approach is reported. The "Par." column indicates the corresponding best parameters: lags, C and γ, respectively.

SVR	NMSE	R^2	MSE	MAPE	RRMSE	Par.	NMSE	R^2	MSE	MAPE	RRMSE	Par.
	Natural Gas Prediction											
	1 *h*						6 *h*					
G	0.815	0.185	0.0089	93.538	141.420	$3\text{-}2^{-1}\text{-}2^{3}$	0.402	0.597	0.0130	44.846	65.929	$5\text{-}2^{-3}\text{-}2$
GT	0.746	0.254	0.0082	87.084	135.320	$3\text{-}2^{13}\text{-}2^{-3}$	0.326	0.673	0.0106	39.753	59.352	$5\text{-}2^{3}\text{-}2^{-1}$
GE	0.776	0.223	0.0085	91.730	138.050	$3\text{-}2^{15}\text{-}2^{-3}$	0.364	0.635	0.0118	42.951	62.730	$5\text{-}2\text{-}2^{-1}$
GET	**0.733**	**0.267**	**0.0080**	**88.219**	**134.090**	$\mathbf{3\text{-}2^{15}\text{-}2^{-5}}$	0.286	0.713	0.0093	37.681	55.576	$5\text{-}2^{-1}\text{-}2$
WG	0.787	0.213	0.0086	90.402	139.000	$3\text{-}2^{-3}\text{-}2^{3}$	0.385	0.614	0.0125	43.739	64.508	$5\text{-}2^{-3}\text{-}2$
WGT	0.743	0.257	0.0081	85.688	135.020	$3\text{-}2^{15}\text{-}2^{-5}$	0.295	0.704	0.0091	41.431	56.622	$3\text{-}2^{-1}\text{-}2^{3}$
WGE	0.755	0.244	0.0083	87.481	136.160	$3\text{-}2^{-3}\text{-}2^{3}$	0.358	0.641	0.0116	42.687	62.217	$5\text{-}2^{-1}\text{-}2^{-1}$
WGET	0.736	0.264	0.0084	87.585	137.270	$2\text{-}2^{15}\text{-}2^{-5}$	**0.269**	**0.731**	**0.0083**	**39.178**	**54.037**	**3-2-2**
	12 *h*						24 *h*					
G	0.278	0.721	0.0157	33.452	47.994	$5\text{-}2^{3}\text{-}2^{-7}$	0.281	0.716	0.0109	27.858	45.153	$5\text{-}2^{7}\text{-}2^{-11}$
GT	**0.191**	**0.809**	**0.0107**	**27.882**	**39.729**	**5-2-2**	**0.250**	**0.748**	**0.0097**	**28.678**	**42.600**	$\mathbf{5\text{-}2^{3}\text{-}2^{-1}}$
GE	0.245	0.754	0.0138	31.274	45.063	$5\text{-}2\text{-}2^{-1}$	0.282	0.715	0.0109	29.905	45.270	$5\text{-}2^{-1}\text{-}2^{-5}$
GET	0.196	0.803	0.0111	28.974	40.332	$5\text{-}2^{3}\text{-}2^{-3}$	0.254	0.744	0.0109	28.252	41.055	$3\text{-}2^{3}\text{-}2$
WG	0.255	0.744	0.0144	31.630	45.932	$5\text{-}2^{5}\text{-}2^{-1}$	0.282	0.716	0.0109	29.229	45.198	$5\text{-}2^{-3}\text{-}2^{-3}$
WGT	0.209	0.790	0.0118	29.035	41.582	$5\text{-}2^{-1}\text{-}2$	0.270	0.727	0.0124	28.478	44.040	$2\text{-}2^{15}\text{-}2^{-5}$
WGE	0.232	0.767	0.0131	30.922	43.866	$5\text{-}2\text{-}2^{-1}$	0.282	0.715	0.0109	29.822	45.253	$5\text{-}2^{-1}\text{-}2^{-5}$
WGET	0.192	0.807	0.0108	28.834	39.856	$5\text{-}2\text{-}2^{-1}$	0.257	0.741	0.0118	29.513	42.921	$2\text{-}2^{9}\text{-}2^{-1}$
	Water Prediction											
	1 *h*						6 *h*					
W	0.884	0.116	0.0031	87.026	145.880	$5\text{-}2^{-1}\text{-}2^{3}$	0.320	0.679	0.0043	37.609	54.669	$5\text{-}2^{-1}\text{-}2^{3}$
WT	0.850	0.150	0.0033	92.534	151.000	$3\text{-}2^{15}\text{-}2^{-3}$	0.311	0.688	0.0042	38.039	53.898	$5\text{-}2^{-3}\text{-}2^{3}$
WE	0.850	0.150	0.0030	87.881	143.750	$2\text{-}2^{15}\text{-}2^{-11}$	0.318	0.681	0.0043	38.114	54.543	$5\text{-}2^{-1}\text{-}2$
WET	0.840	0.160	0.0030	85.489	142.890	$2\text{-}2^{-1}\text{-}2^{3}$	**0.299**	**0.700**	**0.0041**	**37.337**	**52.852**	$\mathbf{5\text{-}2^{-1}\text{-}2}$
WG	0.854	0.146	0.0030	85.383	143.400	$5\text{-}2^{-3}\text{-}2^{3}$	0.347	0.652	0.0047	39.717	56.919	$5\text{-}2^{-3}\text{-}2^{3}$
WGT	0.827	0.173	0.0029	86.522	141.150	$5\text{-}2^{15}\text{-}2^{-9}$	0.331	0.668	0.0045	39.223	55.598	$5\text{-}2^{-3}\text{-}2$
WGE	0.819	0.181	0.0029	87.237	140.440	$5\text{-}2^{15}\text{-}2^{-9}$	0.319	0.680	0.0043	38.015	54.591	$5\text{-}2^{-3}\text{-}2$
WGET	**0.817**	**0.182**	**0.0029**	**86.547**	**140.300**	$\mathbf{5\text{-}2^{15}\text{-}2^{-9}}$	0.305	0.694	0.0042	37.598	53.415	$5\text{-}2^{-3}\text{-}2$
	12 *h*						24 *h*					
W	0.491	0.506	0.0056	30.084	38.822	$2\text{-}2^{-1}\text{-}2^{3}$	0.395	0.601	0.0072	20.778	28.106	$2\text{-}2^{15}\text{-}2^{-7}$
WT	0.495	0.502	0.0080	28.319	42.444	$5\text{-}2^{-3}\text{-}2^{3}$	0.385	0.611	0.0070	21.726	27.741	$2\text{-}2^{-15}\text{-}2^{-3}$
WE	0.512	0.485	0.0058	31.335	39.642	$2\text{-}2^{3}\text{-}2^{-1}$	0.404	0.592	0.0074	21.194	28.427	$2\text{-}2^{9}\text{-}2^{-5}$
WET	**0.484**	**0.514**	**0.0055**	**30.790**	**38.533**	$\mathbf{2\text{-}2^{-3}\text{-}2}$	0.407	0.589	0.0075	21.447	28.532	$2\text{-}2^{9}\text{-}2^{-3}$
WG	0.509	0.489	0.0058	31.083	39.500	$2\text{-}2^{-3}\text{-}2^{3}$	0.375	0.621	0.0069	19.367	27.384	$2\text{-}2\text{-}2$
WGT	0.502	0.495	0.0081	28.325	42.736	$5\text{-}2^{-3}\text{-}2^{3}$	**0.352**	**0.645**	**0.0064**	**20.062**	**26.517**	$\mathbf{2\text{-}2^{5}\text{-}2}$
WGE	0.524	0.474	0.0059	31.216	40.080	$2\text{-}2^{-3}\text{-}2^{3}$	0.411	0.585	0.0075	21.231	28.673	$2\text{-}2^{13}\text{-}2^{-7}$
WGET	0.491	0.506	0.0056	30.894	38.821	$2\text{-}2^{-3}\text{-}2$	0.419	0.577	0.0077	21.544	28.940	$2\text{-}2\text{-}2^{-1}$

W = water G = natural gas E = electric power T = temperature

Table 5: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the GP approach is reported. The "Par." column indicates the corresponding best parameters: lags and maximum depth, respectively.

GP	NMSE	R^2	MSE	MAPE	RRMSE	Par.	NMSE	R^2	MSE	MAPE	RRMSE	Par.
	Natural Gas Prediction											
	1 *h*						6 *h*					
G	0.810	0.190	0.0088615	90.201	141.01	3-10	0.436	0.563	0.014128	44.737	68.65	5-20
GT	0.797	0.203	0.0087179	80.867	139.86	3-20	0.361	0.638	0.011155	42.291	62.666	3-15
GE	0.851	0.149	0.0093062	82.833	144.51	3-15	0.467	0.532	0.015108	45.671	70.993	5-15
GET	**0.781**	**0.219**	**0.0089124**	**80.65**	**141.42**	**2-20**	**0.338**	**0.661**	**0.010433**	**41.489**	**60.605**	**3-20**
WG	0.837	0.162	0.009158	82.119	143.35	3-10	0.424	0.575	0.01374	44.608	67.703	5-15
WGT	0.799	0.201	0.0087347	79.762	140	3-20	0.363	0.636	0.011218	43.588	62.844	3-20
WGE	0.825	0.175	0.009025	82.383	142.31	3-10	0.451	0.548	0.013925	48.86	70.017	3-20
WGET	0.800	0.200	0.0091306	80.286	143.14	2-15	0.402	0.597	0.012411	46.846	66.102	3-15
	12 *h*						24 *h*					
G	0.287	0.712	0.016196	32.574	48.768	5-10	0.273	0.724	0.010548	28.041	44.514	5-20
GT	0.257	0.742	0.0145111	31.107	46.162	5-15	**0.242**	**0.755**	**0.0093663**	**26.027**	**41.947**	**5-10**
GE	0.287	0.711	0.016213	33.249	48.794	5-15	0.292	0.705	0.011276	28.788	46.024	5-10
GET	0.285	0.714	0.016046	33.048	48.542	5-10	0.282	0.715	0.010901	30.045	45.252	5-10
WG	0.284	0.715	0.01599	33.209	48.457	5-10	0.298	0.699	0.011515	28.214	46.509	5-10
WGT	**0.255**	**0.744**	**0.013399**	**31.299**	**43.856**	**3-10**	0.269	0.729	0.012349	26.787	43.905	2-10
WGE	0.306	0.692	0.017277	34.244	50.37	5-10	0.285	0.712	0.011011	27.609	45.481	5-15
WGET	0.274	0.724	0.014396	31.68	45.457	3-20	0.276	0.721	0.012697	28.651	44.518	2-15
	Water Prediction											
	1 *h*						6 *h*					
W	0.926	0.073	0.0032589	82.243	149.38	5-15	**0.357**	**0.642**	**0.0048518**	**39.555**	**57.741**	**5-10**
WT	0.927	0.072	0.0035679	83.409	157.73	3-20	0.374	0.625	0.0050829	39.737	59.101	5-20
WE	0.914	0.085	0.003518	83.653	156.62	3-10	0.380	0.619	0.0051704	39.687	59.607	5-15
WET	0.911	0.089	0.0035041	83.419	156.31	3-20	0.381	0.618	0.0051798	39.745	59.661	5-15
WG	0.913	0.087	0.0032528	82.188	148.97	2-10	0.396	0.603	0.0053821	40.552	60.815	5-15
WGT	0.916	0.083	0.0032238	80.951	148.57	5-20	0.373	0.626	0.005076	40.176	59.06	5-20
WGE	**0.900**	**0.100**	**0.0034619**	**82.549**	**155.36**	**3-20**	0.396	0.603	0.0053877	40.425	60.847	5-10
WGET	0.906	0.093	0.0034862	82.375	155.91	3-20	0.391	0.609	0.0053102	40.727	60.407	5-10
	12 *h*						24 *h*					
W	0.511	0.487	0.0057894	30.943	39.588	2-10	0.388	0.608	0.0071046	20.113	27.857	2-20
WT	0.510	0.488	0.0057787	30.967	39.552	2-10	0.367	0.630	0.0067097	20.039	27.071	2-10
WE	0.530	0.467	0.0060072	31.593	40.326	2-20	0.363	0.634	0.0066378	20.686	26.926	2-20
WET	0.511	0.487	0.00579	30.584	39.591	2-20	0.390	0.606	0.0071449	20.703	27.936	2-20
WG	0.524	0.473	0.0059378	31.36	40.093	2-10	0.379	0.618	0.0069289	20.455	27.51	2-15
WGT	**0.505**	**0.493**	**0.0057221**	**30.718**	**39.358**	**2-20**	0.379	0.618	0.0069339	19.895	27.52	2-10
WGE	0.534	0.463	0.0060532	31.48	40.48	2-10	**0.359**	**0.637**	**0.0065756**	**19.604**	**26.799**	**2-20**
WGET	0.532	0.466	0.006026	31.364	40.389	2-20	0.403	0.593	0.0073855	21.153	28.402	2-20

W = water G = natural gas E = electric power T = temperature

Table 6: For each resources combination, and time resolution, the higher performance achieved by testing al l available parameters combinations for the EKF-GP approach is reported. The "Par." column indicates the corresponding best parameters: lags and maximum depth, respectively.

EKF-GP	NMSE	R^2	MSE	MAPE	RRMSE	Par.	NMSE	R^2	MSE	MAPE	RRMSE	Par.
	Natural Gas Prediction											
	1 *h*						6 *h*					
G	0.891	0.109	0.0097	94.386	147.853	3-10	0.462	0.537	0.0150	45.813	70.649	5-15
GT	0.919	0.081	0.0101	87.053	150.181	3-15	-	-	-	-	-	3-15
GE	-	-	-	-	-	3-20	**0.456**	**0.543**	**0.0147**	**46.218**	**70.140**	**5-15**
GET	0.907	0.093	0.0104	86.477	152.438	2-10	0.498	0.501	0.0154	48.078	73.539	3-15
WG	-	-	-	-	-	3-10	0.507	0.491	0.0164	47.258	74.027	5-10
WGT	-	-	-	-	-	2-20	0.501	0.498	0.0162	46.957	73.557	5-15
WGE	**0.718**	**0.282**	**0.0079**	**74.526**	**132.749**	**3-15**	0.478	0.521	0.0155	45.999	71.820	5-10
WGET	0.790	0.209	0.0086	80.654	139.280	3-15	0.472	0.527	0.0153	46.444	71.408	5-10
	12 *h*						24 *h*					
G	0.461	0.537	0.0260	37.098	61.767	5-10	0.308	0.689	0.0119	30.509	47.262	5-20
GT	**0.283**	**0.715**	**0.0160**	**32.664**	**48.431**	**5-15**	0.330	0.666	0.0128	30.182	48.974	5-15
GE	0.373	0.626	0.0210	37.084	55.549	5-20	0.320	0.677	0.0124	31.067	48.196	5-20
GET	0.331	0.668	0.0187	34.099	52.350	5-15	0.390	0.606	0.0168	31.687	50.913	3-20
WG	0.306	0.693	0.0173	34.372	50.341	5-20	0.310	0.687	0.0120	29.458	47.438	5-10
WGT	0.309	0.690	0.0174	33.910	50.567	5-10	**0.299**	**0.698**	**0.0115**	**29.840**	**46.560**	**5-10**
WGE	0.322	0.676	0.0169	35.551	49.284	3-10	0.311	0.686	0.0120	29.545	47.490	5-10
WGET	0.326	0.672	0.0184	34.588	51.993	5-20	0.309	0.689	0.0119	30.528	47.321	5-10
	Water Prediction											
	1 *h*						6 *h*					
W	0.940	0.060	0.0036	83.992	158.756	3-15	0.397	0.602	0.0054	40.686	60.896	5-15
WT	0.942	0.058	0.0036	84.805	158.920	3-10	0.410	0.589	0.0056	41.101	61.928	5-20
WE	**0.883**	**0.117**	**0.0031**	**80.439**	**146.532**	**2-15**	0.731	0.267	0.0099	44.277	82.670	5-10
WET	0.950	0.049	0.0037	85.364	159.658	3-20	**0.383**	**0.616**	**0.0052**	**40.697**	**59.817**	**5-20**
WG	0.936	0.064	0.0036	84.144	158.413	3-15	0.479	0.520	0.0065	43.854	66.877	5-15
WGT	0.944	0.056	0.0036	84.032	159.107	3-20	0.413	0.586	0.0056	41.704	62.098	5-15
WGE	1.017	-0.018	0.0036	85.574	157.260	2-20	0.407	0.592	0.0055	41.036	61.669	5-20
WGET	0.928	0.072	0.0033	81.570	149.508	5-20	0.406	0.593	0.0055	41.170	61.564	5-10
	12 *h*						24 *h*					
W	1.677	-0.685	0.0190	48.897	71.721	2-20	0.567	0.428	0.0104	24.327	33.664	2-10
WT	1.157	-0.162	0.0131	45.970	59.565	2-10	0.466	0.530	0.0085	22.559	30.525	2-20
WE	1.036	-0.041	0.0117	44.261	56.361	2-15	**0.454**	**0.541**	**0.0083**	**22.164**	**30.140**	**2-15**
WET	1.029	-0.033	0.0117	43.314	56.167	2-15	0.642	0.352	0.0117	26.512	35.814	2-10
WG	**0.608**	**0.389**	**0.0098**	**32.131**	**47.030**	**5-15**	0.634	0.360	0.0116	26.721	35.607	2-20
WGT	1.089	-0.094	0.0123	44.021	57.803	2-15	0.550	0.445	0.0101	25.156	33.158	2-15
WGE	1.456	-0.462	0.0165	47.074	66.817	2-15	0.471	0.525	0.0086	22.705	30.689	2-10
WGET	1.277	-0.283	0.0145	48.813	62.592	2-15	0.564	0.431	0.0103	24.830	33.583	2-10

W = water G = natural gas E = electric power T = temperature

Table 7: For each resources combination, and time resolution, the higher performance achieved by testing al l available parameters combinations for the ELM-LIN approach is reported. The "Par." column indicates the corresponding best parameters: lags and maximum depth, respectively.

ELM-LIN	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.
Natural Gas Prediction												
	1 *h*						6 *h*					
G	0.868	0.132	9.911E-3	87.833	149.12	$2\text{-}2^{15}$	0.443	0.556	0.014345	48.475	69.176	5-2
GT	0.841	0.159	9.602E-3	94.742	146.78	$2\text{-}2^{15}$	0.401	0.598	0.012976	46.909	65.794	$5\text{-}2^{15}$
GE	0.819	0.181	9.347E-3	95.061	144.82	$2\text{-}2^{15}$	0.393	0.606	0.012733	46.81	65.174	$5\text{-}2^{5}$
GET	0.818	0.182	9.339E-3	95.308	144.76	$2\text{-}2^{15}$	0.369	0.630	0.011949	44.635	63.135	$5\text{-}2^{5}$
WG	0.865	0.135	9.870E-3	87.835	148.82	$2\text{-}2^{11}$	0.411	0.588	0.013309	47.554	66.631	5-2
WGT	0.834	0.166	9.516E-3	94.67	146.12	$2\text{-}2^{3}$	0.392	0.607	0.012689	45.978	65.062	5-2
WGE	0.806	0.194	9.199E-3	95.022	143.67	2-2	0.383	0.616	0.012397	46.256	64.309	5-2
WGET	**0.804**	**0.196**	**9.181E-3**	**95.37**	**143.53**	$\mathbf{2\text{-}2^{3}}$	**0.367**	**0.632**	**0.011871**	**44.426**	**62.93**	**5-2**
	12 *h*						24 *h*					
G	0.281	0.718	0.015844	33.961	48.235	$5\text{-}2^{3}$	0.294	0.704	0.011342	28.625	46.158	$2\text{-}2^{15}$
GT	0.251	0.747	0.014176	33.401	45.626	$5\text{-}2^{3}$	**0.283**	**0.715**	**0.01092**	**28.153**	**45.293**	$2\text{-}2^{15}$
GE	0.262	0.736	0.014802	33.471	46.622	$5\text{-}2^{15}$	0.297	0.700	0.011483	30.726	46.445	$5\text{-}2^{3}$
GET	0.241	0.758	0.013583	33.109	44.661	$5\text{-}2^{3}$	0.291	0.707	0.011224	28.577	45.919	$2\text{-}2^{15}$
WG	0.261	0.737	0.014743	33.711	46.53	$5\text{-}2^{15}$	0.294	0.704	0.011347	29.414	46.17	$2\text{-}2^{15}$
WGT	0.241	0.758	0.013572	33.122	44.643	$5\text{-}2^{3}$	0.285	0.713	0.010997	29.047	45.452	$2\text{-}2^{15}$
WGE	0.256	0.743	0.014445	33.564	46.057	$5\text{-}2^{15}$	0.297	0.700	0.01148	30.728	46.44	5-2
WGET	**0.237**	**0.762**	**0.01338**	**33.061**	**44.326**	$\mathbf{5\text{-}2^{3}}$	0.290	0.707	0.011217	30.281	45.904	$2\text{-}2^{15}$
Water Prediction												
	1 *h*						6 *h*					
W	0.949	0.051	0.0033804	85.431	151.87	$2\text{-}2^{1/2}$	0.440	0.559	0.0059839	44.3	64.125	$5\text{-}2^{15}$
WT	0.879	0.121	0.0031319	92.839	146.18	2-2	0.435	0.564	0.0059147	46.408	63.753	$5\text{-}2^{3}$
WE	0.836	0.164	0.0029774	91.839	142.53	$2\text{-}2^{1/2}$	0.413	0.586	0.0056156	43.477	62.12	$5\text{-}2^{15}$
WET	0.829	0.171	0.0029532	93.048	141.95	2-2	**0.405**	**0.594**	**0.0055089**	**44.404**	**61.527**	**5-2**
WG	0.939	0.061	0.0033449	86.369	151.07	$2\text{-}2^{1/2}$	0.454	0.545	0.0061753	45.051	65.142	$5\text{-}2^{15}$
WGT	0.873	0.127	0.0031094	93.661	145.65	2-1	0.436	0.563	0.0059336	46.237	63.855	$5\text{-}2^{3}$
WGE	0.836	0.164	0.002978	91.943	142.54	$2\text{-}2^{1/2}$	0.423	0.576	0.0057556	43.901	62.89	$5\text{-}2^{15}$
WGET	**0.829**	**0.170**	**0.0029549**	**93.918**	**141.99**	$\mathbf{2\text{-}2^{1/2}}$	0.408	0.592	0.005541	44.382	61.706	5-2
	12 *h*						24 *h*					
W	0.599	0.398	0.0096642	32.473	46.67	$5\text{-}2^{5}$	0.442	0.554	0.0080876	22.639	29.721	$2\text{-}2^{15}$
WT	0.588	0.409	0.0094946	32.004	46.258	5-2	**0.428**	**0.568**	**0.007826**	**22.579**	**29.237**	$\mathbf{2\text{-}2^{15}}$
WE	0.606	0.391	0.0097847	32.925	46.959	$5\text{-}2^{15}$	0.453	0.543	0.0086166	24.59	31.55	$5\text{-}2^{3}$
WET	0.591	0.406	0.0095347	32.443	46.356	$5\text{-}2^{3}$	0.453	0.543	0.0082879	22.874	30.087	$2\text{-}2^{15}$
WG	0.596	0.401	0.0096241	32.49	46.573	5-2	0.447	0.549	0.0081866	22.524	29.903	$2\text{-}2^{15}$
WGT	**0.579**	**0.419**	**0.0093365**	**32.067**	**45.871**	**5-1**	0.435	0.561	0.0079629	22.956	29.491	$2\text{-}2^{15}$
WGE	0.606	0.391	0.0097765	32.802	46.94	5-1	0.457	0.539	0.0086951	24.751	31.693	5-2
WGET	0.585	0.413	0.0094329	32.646	46.108	5-1	0.455	0.540	0.0083354	23.002	30.173	$2\text{-}2^{15}$

W = water G = natural gas E = electric power T = temperature

Table 8: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the ELM-RBF approach is reported. The "Par." column indicates the corresponding best parameters: lags and maximum depth, respectively.

ELM-RBF	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.
Natural Gas Prediction												
	1 *h*						6 *h*					
G	0.760	0.239	0.0083188	92.836	136.63	$3\text{-}2^{-1}\_2^{-3}$	0.395	0.604	0.012779	44.698	65.292	$5\text{-}1\text{-}2^{-1}$
GT	0.717	0.283	0.0078435	88.769	132.66	$3\text{-}2^{-1}\_2^{-3}$	0.315	0.684	0.010193	39.564	58.312	$5\text{-}2^{7}\_2$
GE	0.741	0.259	0.008107	90.477	134.87	$3\text{-}2^{-1}\_2^{-3}$	0.339	0.660	0.010977	42.071	60.513	$5\text{-}2^{3}\_2$
GET	0.704	0.295	0.0077036	87.775	131.48	$3\text{-}2\text{-}2^{-1}$	0.285	0.715	0.0092101	38.052	55.43	$5\text{-}2^{5}\_2$
WG	0.740	0.259	0.0080987	90.669	134.81	$3\text{-}2^{-1}\_2^{-3}$	0.376	0.623	0.012185	44.16	63.756	$5\text{-}1\text{-}2^{-1}$
WGT	0.711	0.289	0.007223	89.067	131.54	$5\text{-}1\text{-}2^{-1}$	0.287	0.712	0.0088604	40.709	55.851	$3\text{-}2\text{-}2^{-3}$
WGE	0.722	0.278	0.0078942	88.828	133.09	$3\text{-}2^{-1}\_2^{-3}$	0.344	0.655	0.01113	42.237	60.934	$5\text{-}2^{3}\_2$
WGET	**0.700**	**0.299**	**0.0076619**	**87.213**	**131.12**	$\mathbf{3}\text{-}2\text{-}2^{-1}$	**0.266**	**0.733**	**0.0082166**	**39.926**	**53.784**	$\mathbf{3}\text{-}2^{3}\text{-}2^{-1}$
	12 *h*						24 *h*					
G	0.277	0.722	0.015594	33.7	47.853	$5\text{-}2^{13}\_2^{9}$	0.272	0.726	0.010491	30.386	44.394	$2\text{-}2^{13}\_2^{3}$
GT	0.199	0.800	0.011218	28.138	40.588	$5\text{-}2^{3}\_2^{-1}$	**0.241**	**0.757**	**0.0093065**	**28.689**	**41.813**	$\mathbf{2}\text{-}2^{5}\text{-}2^{-1}$
GE	0.249	0.750	0.014028	32.47	45.387	$5\text{-}2^{5}\_2$	0.289	0.708	0.013293	29.498	45.552	$2\text{-}2^{15}\_2^{5}$
GET	0.196	0.803	0.011039	28.884	40.262	$5\text{-}2^{3}\_2$	0.245	0.752	0.011288	28.8	41.976	$2\text{-}2^{5}\_2$
WG	0.250	0.749	0.014088	31.685	45.483	$5\text{-}2^{7}\_2$	0.285	0.712	0.011012	30.723	45.483	$2\text{-}2^{5}\text{-}2^{-1}$
WGT	0.203	0.796	0.011436	28.656	40.979	$5\text{-}2^{3}\_2^{-1}$	0.253	0.744	0.0097914	29.404	42.888	$2\text{-}2^{5}\text{-}2^{-1}$
WGE	0.236	0.763	0.013295	31.487	44.185	$5\text{-}2^{5}\_2$	0.296	0.701	0.01145	31.981	46.378	$2\text{-}2^{5}\_2$
WGET	**0.194**	**0.805**	**0.010948**	**28.86**	**40.097**	$\mathbf{5}\text{-}2^{3}\_2$	0.252	0.746	0.01159	28.977	42.535	$2\text{-}2^{9}\_2^{3}$
Water Prediction												
	1 *h*						6 *h*					
W	0.812	0.187	0.002895	92.653	140.54	$2\text{-}2^{-1}\_2^{-3}$	0.318	0.681	0.0043263	38.165	54.524	$5\text{-}2\text{-}2^{-3}$
WT	0.803	0.197	0.0028606	92.055	139.71	$2\text{-}2^{-3}\_2^{-3}$	0.321	0.678	0.0043635	39.175	54.758	$5\text{-}1\text{-}2^{-3}$
WE	0.784	0.216	0.0027944	90.884	138.08	$2\text{-}2^{-1}\_2^{-3}$	0.318	0.682	0.0043195	38.804	54.482	$5\text{-}2\text{-}2^{-1}$
WET	0.777	0.223	0.0027686	89.901	137.44	$2\text{-}2^{-1}\_2^{-3}$	0.310	0.689	0.0042185	39.445	53.841	$5\text{-}2\text{-}2^{-1}$
WG	0.804	0.195	0.0028661	92.271	139.84	$2\text{-}2^{-1}\_2^{-1}$	0.349	0.651	0.0047396	41.014	57.069	$5\text{-}2\text{-}2^{-1}$
WGT	0.789	0.210	0.0027769	90.59	137.89	$2\text{-}2^{5}\_2$	0.337	0.662	0.0045873	41.14	56.145	$5\text{-}2\text{-}2^{-1}$
WGE	0.772	0.228	0.0027508	89.281	137	$2\text{-}2^{-1}\_2^{-3}$	0.312	0.687	0.0042479	38.637	54.028	$5\text{-}2\text{-}2^{-1}$
WGET	**0.769**	**0.231**	**0.0027403**	**88.576**	**136.73**	$\mathbf{2}\text{-}2^{-1}\_2^{-3}$	**0.310**	**0.689**	**0.0042137**	**39.117**	**53.81**	$\mathbf{5}\text{-}2\text{-}2^{-1}$
	12 *h*						24 *h*					
W	0.505	0.493	0.0081425	29.147	42.838	$5\text{-}1\text{-}2^{-3}$	0.386	0.610	0.0070688	20.668	27.786	$2\text{-}2^{13}\_2^{3}$
WT	**0.490**	**0.507**	**0.0055541**	**30.462**	**38.776**	$\mathbf{2}\text{-}3\text{-}2^{-3}$	0.368	0.628	0.006744	21.556	27.14	$2\text{-}2^{5}\text{-}2^{-1}$
WE	0.524	0.474	0.0059338	31.881	40.079	$2\text{-}2\text{-}2^{-1}$	0.386	0.611	0.0070635	20.968	27.776	$2\text{-}2^{15}\_2^{5}$
WET	0.503	0.495	0.0056933	31.393	39.258	$2\text{-}2^{-1}\_2^{-1}$	0.389	0.608	0.0071125	21.213	27.872	$2\text{-}2^{5}\_2$
WG	0.520	0.478	0.0083888	29.007	43.481	$5\text{-}2\text{-}2^{-1}$	0.364	0.632	0.0066677	20.002	26.986	$2\text{-}2^{5}\text{-}2^{-1}$
WGT	0.507	0.490	0.0081839	29.398	42.947	$5\text{-}2\text{-}2^{-1}$	**0.351**	**0.646**	**0.006417**	**20.644**	**26.474**	$\mathbf{2}\text{-}2^{5}\text{-}2^{-1}$
WGE	0.528	0.469	0.0085242	29.541	43.831	$5\text{-}2\text{-}2^{-1}$	0.391	0.605	0.007159	21.459	27.963	$2\text{-}2^{5}\text{-}2$
WGET	0.516	0.482	0.0058405	31.736	39.763	$2\text{-}2^{-1}\_2^{-1}$	0.385	0.611	0.0070514	21.13	27.752	$2\text{-}2^{9}\text{-}2^{3}$

W = water G = natural gas E = electric power T = temperature

2 AMPds Results - 2 year

Table 9: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the ANN approach is reported. The "Par." column indicates the corresponding parameters: lags number and neurons number of the hidden layer, respectively.

ANN	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.
	Natural Gas Prediction											
				1 *h*					6 *h*			
G	0.730	0.270	0.0082704	85.212	124.59	3-8	0.230	0.770	0.0086808	33.805	48.252	5-10
GT	0.632	0.368	0.0071542	77.823	115.27	5-8	0.203	0.797	0.0075724	31.624	44.511	5-12
GE	0.697	0.303	0.0079114	83.037	122.85	3-6	0.219	0.781	0.0083897	35.325	48.538	5-7
GET	0.624	0.376	0.0070199	78.96	117.15	5-11	**0.182**	**0.818**	**0.0063478**	**30.482**	**42.659**	**5-5**
WG	0.678	0.322	0.0075719	80.893	118.8	5-8	0.222	0.778	0.0079604	32.436	45.422	5-6
WGT	**0.619**	**0.381**	**0.00686**	**78.752**	**116.52**	**5-12**	0.191	0.808	0.0073155	31.303	43.745	5-8
WGE	0.661	0.339	0.0070898	80.974	118.95	5-5	0.217	0.783	0.0081738	33.22	46.686	5-6
WGET	0.619	0.381	0.0073228	77.212	116.52	3-13	0.195	0.805	0.007053	32.862	45.514	5-6
				12 *h*					24 *h*			
G	0.220	0.780	0.0099782	27.577	38.692	2-8	0.208	0.791	0.0080546	23.199	34.163	3 - 5
GT	0.171	0.829	0.0082771	24.758	34.657	2-5	**0.153**	**0.847**	**0.0066739**	**20.225**	**28.901**	**2-5**
GE	0.203	0.797	0.0089565	25.585	36.272	2-13	0.231	0.767	0.010066	24.981	35.058	2-5
GET	**0.170**	**0.830**	**0.0086205**	**25.524**	**35.192**	**2-6**	0.174	0.825	0.0078086	23.026	31.51	2-7
WG	0.203	0.796	0.0098542	26.725	37.815	2-5	0.201	0.798	0.0082558	22.896	33.358	2-5
WGT	0.172	0.828	0.00873	25.882	35.414	2-6	0.157	0.842	0.0064898	21.349	29.038	2-10
WGE	0.203	0.797	0.0089689	27.031	38.726	2-7	0.211	0.788	0.0087503	23.432	33.509	2-6
WGET	0.189	0.811	0.0081233	26.646	37.48	2-7	0.201	0.798	0.0087247	24.156	34.403	2-7
	Water Prediction											
				1 *h*					6 *h*			
W	0.790	0.209	0.0034303	92.31	146.53	5-8	**0.319**	**0.680**	**0.0046829**	**34.619**	**54.083**	**5-10**
WT	0.762	0.238	0.0033159	88.812	140.27	5-7	0.330	0.670	0.0047485	35.949	52.888	5-7
WE	0.783	0.217	0.0031675	90.021	138.85	3-8	0.327	0.672	0.0051105	35.068	54.186	5-6
WET	**0.747**	**0.253**	**0.0029371**	**86.688**	**132.95**	**5-6**	0.353	0.646	0.0053979	37.32	55.649	5-5
WG	0.779	0.221	0.0034992	89.714	141.46	5-6	0.327	0.673	0.0051055	35.931	54.16	5-6
WGT	0.766	0.234	0.0033029	89.634	139.81	5-7	0.344	0.656	0.0052836	34.491	53.656	5-8
WGE	0.761	0.239	0.003284	87.713	139.41	5-7	0.341	0.658	0.0048476	38.466	54.437	5-5
WGET	0.757	0.242	0.0031452	90.255	139.21	5-5	0.371	0.629	0.005136	40.308	57.663	5-7
				12 *h*					24 *h*			
W	**0.223**	**0.777**	**0.0037**	**25.5930**	**34.1630**	**2-13**	**0.550**	**0.448**	**0.0058943**	**18.687**	**23.592**	**2-5**
WT	0.263	0.736	0.0041394	26.577	36.879	2-7	0.551	0.447	0.008975	20.9	28.506	2-7
WE	0.265	0.735	0.0043602	25.564	36.363	2-6	0.611	0.386	0.0073871	20.739	26.325	2-6
WET	0.275	0.725	0.0046434	27.503	39.265	3-7	0.582	0.415	0.0098821	22.671	29.785	2-9
WG	0.252	0.747	0.0039091	26.519	37.008	5-5	0.619	0.378	0.009182	22.457	29.752	2-7
WGT	0.257	0.742	0.0040498	26.304	36.478	2-5	0.660	0.337	0.0083737	19.31	27.509	2-6
WGE	0.262	0.737	0.0041325	26.504	36.849	2-5	0.618	0.379	0.0078452	19.439	26.627	2-6
WGET	0.278	0.722	0.0042738	26.105	35.981	3-5	0.560	0.438	0.0082001	19.657	26.885	2-5

W = water G = natural gas E = electric power T = temperature

Table 10: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the DBN approach is reported. The "Par." column indicates the corresponding best parameters: lags, neurons, max. iteration and batch size, respectively.

DBN	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.
	Natural Gas Prediction											
	1 *h*								6 *h*			
G	0.824	0.175	0.0091422	98.363	133.66	3-4-100-100	0.329	0.671	0.011996	42.199	56.773	5-3-100-100
GT	0.753	0.247	0.008348	91.638	127.73	3-6-100-100	0.270	0.730	0.0098479	38.284	51.439	5-4-500-100
GE	0.820	0.180	0.0090912	97.807	133.29	3-7-100-100	0.285	0.715	0.010379	39.823	52.808	5-8-500-100
GET	0.746	0.254	0.0082743	91.685	127.16	3-7-100-100	0.245	0.754	0.0089472	37.501	49.03	5-2-100-10
WG	0.790	0.210	0.0087605	95.352	130.84	3-7-100-100	0.287	0.713	0.010453	39.607	52.997	5-8-300-100
WGT	0.735	0.265	0.0081547	89.923	126.24	3-5-100-100	0.256	0.744	0.009339	36.875	50.092	5-2-100-10
WGE	0.748	0.252	0.008295	91.827	127.32	3-2-500-100	0.261	0.738	0.0095373	37.548	50.621	5-2-100-10
WGET	**0.723**	**0.277**	**0.008021**	**89.907**	**125.2**	**3-2-300-100**	**0.238**	**0.762**	**0.0086657**	**36.099**	**48.252**	**5-2-100-10**
	12 *h*						24 *h*					
G	0.239	0.761	0.010575	30.769	41.41	5-9-500-100	0.233	0.765	0.0098504	23.97	34.533	2-2-500-100
GT	0.192	0.807	0.0089112	29.005	39.004	2-5-500-100	0.192	0.808	7.87E-03	24.441	32.198	3-9-500-100
GE	0.231	0.769	0.01022	30.512	40.71	5-6-100-10	0.223	0.776	9.17E-03	24.908	34.769	3-7-500-100
GET	**0.185**	**0.815**	**0.0085846**	**28.339**	**38.282**	**2-6-100-10**	**0.188**	**0.812**	**7.70E-03**	**23.934**	**31.86**	**3-7-300-10**
WG	0.226	0.774	0.0099992	30.178	40.267	5-4-100-10	0.224	0.775	0.0091816	24.972	34.787	3-7-300-100
WGT	0.191	0.809	0.0088368	28.699	38.841	2-2-100-10	0.189	0.810	0.007752	23.741	31.964	3-2-300-10
WGE	0.222	0.777	0.010312	31.631	41.957	2-2-100-10	0.206	0.793	8.70E-03	23.892	32.456	2-2-300-10
WGET	0.188	0.811	0.0087292	28.693	38.603	2-3-500-100	0.189	0.811	7.74E-03	23.952	31.949	3-7-300-10
	Water Prediction											
	1 *h*								6 *h*			
W	0.926	0.074	0.0038769	101.7	150.86	2-2-100-100	0.541	0.459	0.0070822	51.048	66.539	5-5-100-100
WT	0.918	0.082	0.0038445	104.35	150.23	2-5-100-100	0.490	0.510	0.006411	46.902	63.307	5-5-500-100
WE	0.887	0.113	0.0036975	99.831	151.47	5-2-500-100	0.507	0.492	0.0066393	49.729	64.425	5-9-100-100
WET	0.871	0.129	0.0036295	98.874	150.07	5-2-500-100	0.463	0.536	0.0060671	45.326	61.586	5-2-100-10
WG	0.920	0.079	0.0038552	102.05	150.44	2-4-100-100	0.458	0.542	0.0059958	45.836	61.223	5-2-300-10
WGT	0.901	0.098	0.0037756	99.534	148.88	2-2-300-100	0.426	0.574	0.0055766	43.089	59.044	5-2-300-10
WGE	0.881	0.119	0.0036717	99.39	150.94	5-2-500-100	0.443	0.557	0.005798	44.595	60.205	5-2-300-10
WGET	**0.867**	**0.133**	**0.0036158**	**98.823**	**149.79**	**5-2-500-100**	**0.407**	**0.592**	**0.0053355**	**42.414**	**57.753**	**5-2-300-10**
	12 *h*						24 *h*					
W	0.361	0.638	0.0063945	31.164	43.485	5-2-100-100	0.609	0.388	0.0074711	20.617	26.299	5-3-500-10
WT	0.345	0.654	0.0061095	31.019	42.505	5-3-100-10	0.630	0.367	0.0077327	20.983	26.755	5-5-500-100
WE	0.353	0.646	0.0062555	30.935	43.01	5-10-500-100	**0.607**	**0.390**	**0.0074474**	**20.578**	**26.257**	**5-3-300-10**
WET	0.341	0.658	0.0060368	30.313	42.251	5-5-100-10	0.625	0.372	0.0076738	20.711	26.653	5-4-300-10
WG	0.359	0.640	0.0063536	30.915	43.346	5-2-500-10	0.613	0.384	0.0075179	20.816	26.381	5-7-300-10
WGT	0.321	0.678	0.0056904	28.814	41.021	5-2-500-10	0.640	0.357	7.85E-03	21.127	26.961	5-9-300-100
WGE	0.351	0.649	0.0062072	29.72	42.844	5-2-300-10	0.625	0.372	0.0076725	20.762	26.651	5-7-300-10
WGET	**0.317**	**0.682**	**0.0056171**	**28.758**	**40.756**	**5-2-300-10**	0.634	0.363	0.0077762	20.975	26.83	5-10-300-100

W = water G = natural gas E = electric power T = temperature

Table 11: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the ESN approach is reported. The "Par." column indicates the corresponding best parameters: lags, reservoir size, initial transient, α and β, respectively.

ESN	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.
	Natural Gas Prediction											
	1 *h*						6 *h*					
G	0.730	0.270	0.0080983	88.732	125.8	3-40-4-1/3-1e-3	0.303	0.696	0.01106	39.931	54.512	5-25-5-1/2-1.e-15
GT	0.703	0.297	0.0077999	84.035	123.46	3-35-10-1/2-1e-2	0.268	0.732	0.0097645	37.564	51.221	5-40-6-1/2-1e-2
GE	0.738	0.262	0.0081829	90.297	126.46	3-35-11-1/3-1e-4	0.271	0.729	0.0098838	38.034	51.532	5-35-6-1-1
GET	0.709	0.290	0.0078672	88.463	123.99	3-20-2-1/2-1e-2	**0.247**	**0.753**	**0.0090027**	**36.46**	**49.182**	**5-25-6-1-1e-2**
WG	0.721	0.278	0.0080007	88.571	125.04	3-40-14-1/2-1e-3	0.276	0.723	0.01008	38.132	52.04	5-2-4-1/2-1e-15
WGT	0.696	0.304	0.0077201	85.661	122.83	3-30-13-1/2-1e-13	0.260	0.740	0.0094873	37.192	50.488	5-40-13-1/5-1e-2
WGE	0.723	0.277	0.008018	88.174	125.18	3-30-15-1/2-1e-2	0.264	0.736	0.0096133	37.584	50.822	5-20-14-1-1e-15
WGET	**0.695**	**0.305**	**0.0077108**	**85.144**	**122.75**	**3-40-1-1/2-1e-1**	0.240	0.760	0.0087491	36.209	48.484	5-35-5-1/2-1e-2
	12 *h*						24 *h*					
G	0.213	0.787	0.0094235	27.821	39.091	5-30-15-0.33333-0	0.210	0.789	0.0088754	23.36	32.779	2-30-3-1/4-1e-14
GT	0.184	0.816	0.0085254	27.516	38.15	2-30-3-1/2-1e-2	0.188	0.811	0.0079229	22.602	30.97	2-30-4-1-1e-14
GE	0.199	0.800	0.0092397	28.255	39.716	2-30-8-1/2-1e-3	**0.214**	**0.785**	**0.0090399**	**23.717**	**33.082**	**2-30-3-1-1e-15**
GET	**0.176**	**0.823**	**0.0081687**	**27.333**	**37.344**	**2-30-6-1/2-1e-13**	0.193	0.806	0.0079252	24.34	32.319	3-5-2-1-1e-3
WG	0.200	0.799	0.0088811	26.847	37.949	5-35-3-1/2-1e-2	0.208	0.791	0.0087654	22.432	32.575	2-30-9-1-1
WGT	0.181	0.819	0.0083858	27.694	37.836	2-25-0-1-1e-15	0.193	0.806	0.0079186	24.679	32.306	3-20-3-1-1e-15
WGE	0.202	0.797	0.0093869	28.339	40.031	2-25-7-1/3-1e-2	0.215	0.784	0.0088132	24.927	34.082	3-10-14-1-1e-14
WGET	0.176	0.823	0.0081834	26.697	37.377	2-30-2-1/2-1e-2	0.198	0.801	0.0081479	24.543	32.77	3-5-14-1-1e-14
	Water Prediction											
	1 *h*						6 *h*					
W	0.848	0.152	0.0035513	94.057	144.39	2-40-15-1/3-1e-13	0.429	0.570	0.0056199	42.65	59.273	5-40-14-1/2-1e-14
WT	0.852	0.148	0.0035699	95.426	144.76	2-30-7-1/3-1e-3	0.442	0.557	0.0057949	43.585	60.188	5-40-12-1/2-1e-13
WE	0.830	0.170	0.0034771	94.017	142.87	2-35-8-1/2-1e-4	0.440	0.559	0.0057645	42.979	60.03	5-10-12-1/2-1e-14
WET	0.828	0.171	0.0034699	93.716	142.72	2-15-1-1/3-1e-14	0.424	0.575	0.0055553	42.586	58.931	5-2-13-1-1e-13
WG	0.834	0.166	0.0034944	94.489	143.23	2-30-1-1/2-1e-12	0.457	0.543	0.0059793	44.331	61.139	5-10-3-1/2-1e-14
WGT	0.839	0.161	0.0035132	95.534	143.61	2-15-6-1/2-1e-13	0.430	0.569	0.0056337	42.692	59.345	5-40-13-1-1
WGE	0.827	0.173	0.0034632	91.823	142.59	2-35-8-1/2-1e-2	0.445	0.555	0.0058266	44.013	60.353	5-5-1-1-1
WGET	**0.818**	**0.181**	**0.0034283**	**92.713**	**141.86**	**2-35-12-1/2-1e-15**	**0.413**	**0.587**	**0.0054049**	**42.778**	**58.128**	**5-2-1-1-1e-13**
	12 *h*						24 *h*					
W	**0.308**	**0.691**	**0.0057251**	**29.036**	**41.507**	**3-40-7-1/3-1e-6**	**0.581**	**0.417**	**0.0071261**	**20.18**	**25.684**	**5-5-1-1-1e-3**
WT	0.311	0.689	0.0055021	28.915	40.337	5-20-10-1-0.01	0.604	0.393	0.0074103	20.46	26.191	5-5-11-1-1e-15
WE	0.314	0.685	0.0055622	29.038	40.557	5-5-15-1-1.00E-14	0.595	0.402	0.0072993	20.422	25.995	5-5-0-1-1e-1
WET	0.309	0.690	0.0054736	28.82	40.232	5-25-10-1-0.01	0.596	0.401	0.0073145	20.274	26.022	5-15-11-1-1
WG	0.319	0.680	0.0056562	29.397	40.898	5-20-10-1-0.1	0.606	0.392	0.0074302	20.676	26.227	5-5-0-1-1e-3
WGT	0.314	0.685	0.005566	28.964	40.571	5-5-9-1-1e-14	0.631	0.366	0.007744	20.966	26.775	5-15-11-1/2-1
WGE	0.317	0.682	0.0056139	29.311	40.745	5-30-14-1-0.1	0.604	0.393	0.0074126	20.606	26.196	5-30-0-1/2-1
WGET	0.309	0.690	0.0054733	28.666	40.231	5-5-15-1-1e-15	0.625	0.372	0.0076703	20.919	26.647	5-30-0-1-1

W = water G = natural gas E = electric power T = temperature

Table 12: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the SVR approach is reported. The "Par." column indicates the corresponding best parameters: lags, C and γ, respectively.

SVR	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.
Natural Gas Prediction												
	1 *h*						6 *h*					
G	0.775	0.225	0.0085909	86.797	129.57	$3\text{-}2^{3}\text{-}2^{3}$	0.246	0.754	0.0089807	33.882	49.122	$5\text{-}2^{-3}\text{-}2^{3}$
GT	0.675	0.325	0.00748	81.125	120.9	$3\text{-}2^{15}\text{-}2^{-5}$	0.201	0.798	0.0073436	31.089	44.419	$5\text{-}2\text{-}2$
GE	0.737	0.263	0.0081729	82.969	126.38	$3\text{-}2\text{-}2$	0.208	0.791	0.0076049	32.077	45.203	$5\text{-}2^{-1}\text{-}2$
GET	0.659	0.341	0.0073088	77.32	119.51	$3\text{-}2^{-}3\text{-}2^{3}$	0.184	0.816	0.0067146	30.688	42.475	$5\text{-}2^{-1}\text{-}2$
WG	0.725	0.275	0.0082748	81.85	127.28	$5\text{-}2^{-}3\text{-}2^{3}$	0.215	0.785	0.0078498	32.043	45.925	$5\text{-}2^{-1}\text{-}2$
WGT	0.658	0.342	0.0075019	77.234	121.19	$5\text{-}2^{-}3\text{-}2^{3}$	0.189	0.811	0.0068892	30.317	43.023	$5\text{-}2^{-1}\text{-}2$
WGE	0.700	0.300	0.0077646	80.346	123.18	$3\text{-}2^{-}3\text{-}2^{3}$	0.199	0.801	0.007243	31.576	44.114	$5\text{-}2^{-1}\text{-}2$
WGET	**0.644**	**0.356**	**0.0071406**	**76.158**	**118.13**	$\mathbf{3\text{-}2^{-}3\text{-}2^{3}}$	**0.182**	**0.818**	**0.0066456**	**30.675**	**42.256**	$\mathbf{5\text{-}2\text{-}2}$
	12 *h*						24 *h*					
G	0.204	0.795	0.0090536	27.354	38.316	$5\text{-}2^{3}\text{-}2^{-1}$	0.217	0.782	0.0089283	24.056	34.303	$5\text{-}2^{-3}\text{-}2^{3}$
GT	0.159	0.841	0.0073623	24.863	35.452	$2\text{-}2^{-1}\text{-}2^{3}$	**0.178**	**0.821**	**0.0075218**	**21.054**	**30.176**	$\mathbf{2\text{-}2^{11}\text{-}2^{-3}}$
GE	0.188	0.811	0.0087353	26.886	38.617	$2\text{-}2^{-1}\text{-}2^{3}$	0.211	0.788	0.0086734	23.792	33.81	$3\text{-}2^{3}\text{-}2^{-1}$
GET	0.156	0.844	0.0072259	24.256	35.122	$2\text{-}2^{-1}\text{-}2^{3}$	0.183	0.816	0.0075285	22.608	31.5	$3\text{-}2^{9}\text{-}2^{-5}$
WG	0.180	0.819	0.0083634	25.98	37.786	$2\text{-}2^{-1}\text{-}2^{3}$	0.205	0.794	0.0086651	22.002	32.389	$2\text{-}2^{9}\text{-}2^{-1}$
WGT	**0.145**	**0.854**	**0.0067**	**23.6900**	**33.9330**	$\mathbf{2\text{-}2^{-1}\text{-}2^{3}}$	0.182	0.817	0.0074748	22.624	31.387	$3\text{-}2^{5}\text{-}2^{-3}$
WGE	0.182	0.818	0.0084357	26.467	37.949	$2\text{-}2^{-3}\text{-}2^{3}$	0.205	0.794	0.0084021	24.148	33.277	$3\text{-}2^{5}\text{-}2^{-3}$
WGET	0.154	0.846	0.0071347	24.377	34.9	$2\text{-}2^{-3}\text{-}2^{3}$	0.181	0.818	0.0074213	22.833	31.275	$3\text{-}2^{9}\text{-}2^{-5}$
Water Prediction												
	1 *h*						6 *h*					
W	0.881	0.119	0.003673	88.456	150.97	$5\text{-}2\text{-}2^{3}$	0.325	0.675	0.0042516	35.348	51.555	$5\text{-}2^{-1}\text{-}2^{3}$
WT	0.831	0.169	0.0034659	90.756	146.65	$5\text{-}2^{15}\text{-}2^{-5}$	0.321	0.679	0.0042043	35.892	51.267	$5\text{-}2^{-3}\text{-}2^{3}$
WE	0.838	0.162	0.0034921	88.503	147.2	$5\text{-}2^{15}\text{-}2^{-9}$	0.328	0.672	0.004292	36.287	51.799	$5\text{-}2^{-3}\text{-}2^{3}$
WET	0.818	0.182	0.0034258	87.989	141.81	$2\text{-}2^{15}\text{-}2^{-5}$	0.323	0.677	0.0042288	36.123	51.416	$5\text{-}2^{-3}\text{-}2$
WG	0.853	0.146	0.0035582	86.478	148.59	$5\text{-}2^{-3}\text{-}2^{3}$	0.318	0.681	0.0041687	35.401	51.05	$5\text{-}2^{-3}\text{-}2^{3}$
WGT	0.827	0.173	0.0034652	86.529	142.63	$2\text{-}2^{15}\text{-}2^{-7}$	0.319	0.680	0.0041805	35.957	51.122	$5\text{-}2^{-3}\text{-}2^{3}$
WGE	**0.806**	**0.194**	**0.0033749**	**86.535**	**140.75**	$\mathbf{2\text{-}2^{15}\text{-}2^{-5}}$	0.322	0.678	0.0042111	35.674	51.308	$5\text{-}2^{-3}\text{-}2$
WGET	0.806	0.194	0.0033599	83.339	144.39	$5\text{-}2^{-3}\text{-}2^{3}$	**0.316**	**0.683**	**0.0041431**	**35.598**	**50.892**	$\mathbf{5\text{-}2^{-3}\text{-}2}$
	12 *h*						24 *h*					
W	**0.269**	**0.730**	**0.0039715**	**25.366**	**36.044**	$\mathbf{2\text{-}2^{9}\text{-}2^{3}}$	**0.581**	**0.416**	**0.0071328**	**20**	**25.696**	$\mathbf{5\text{-}2^{3}\text{-}2}$
WT	0.289	0.710	0.0042647	26.414	37.351	$2\text{-}2^{13}\text{-}2^{-1}$	0.619	0.378	0.0075925	20.699	26.512	$5\text{-}2^{15}\text{-}2^{-15}$
WE	0.275	0.724	0.0040623	26.174	36.454	$2\text{-}2^{7}\text{-}2$	0.605	0.392	0.0074234	20.286	26.215	$5\text{-}2^{15}\text{-}2^{11}$
WET	0.279	0.720	0.0041249	26.398	36.734	$2\text{-}2^{2}\text{-}2$	0.612	0.386	0.0075034	20.401	26.356	$5\text{-}2^{5}\text{-}2^{-7}$
WG	0.292	0.708	0.0051639	27.4	39.078	$5\text{-}2^{-3}\text{-}2$	0.615	0.382	0.0075484	20.481	26.434	$5\text{-}2^{-3}\text{-}2^{-1}$
WGT	0.295	0.705	0.0043504	26.368	37.725	$2\text{-}2^{-3}\text{-}2^{3}$	0.624	0.373	0.0076573	20.654	26.624	$5\text{-}2^{-1}\text{-}2^{-3}$
WGE	0.291	0.708	0.0042953	26.38	37.485	$2\text{-}2^{-3}\text{-}2^{3}$	0.600	0.397	0.0073644	20.022	26.11	$5\text{-}2^{15}\text{-}2^{-9}$
WGET	0.291	0.709	0.0042912	26.365	37.467	$2\text{-}2^{-3}\text{-}2^{3}$	0.620	0.377	0.0076053	20.671	26.534	$5\text{-}2^{15}\text{-}2^{-11}$

W = water G = natural gas E = electric power T = temperature

Table 13: For each resources combination, and time resolution, the higher performance achieved by testing all available parameters combinations for the GP approach is reported. The "Par." column indicates the corresponding best parameters: lags and maximum depth, respectively.

GP	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.
	Natural Gas Prediction											
	1 *h*						6 *h*					
G	0.782	0.218	0.0086742	84.936	130.2	3-20	0.279	0.721	0.010181	34.776	52.302	5-15
GT	**0.740**	**0.260**	**0.0082017**	**73.991**	**126.6**	**2-20**	0.263	0.737	0.0095843	33.968	50.746	5-20
GE	0.797	0.203	0.0088357	76.053	131.4	3-15	0.264	0.736	0.0096151	34.539	50.827	5-10
GET	0.743	0.257	0.0082417	74.009	126.91	3-10	0.279	0.721	0.0101165	34.742	52.26	5-20
WG	0.781	0.218	0.0086659	76.465	130.14	3-10	0.302	0.698	0.011007	36.089	54.381	5-15
WGT	0.754	0.246	0.0083638	73.285	127.85	3-20	**0.250**	**0.750**	**0.009121**	**33.274**	**49.504**	**5-20**
WGE	0.772	0.228	0.0085575	75.103	129.32	3-10	0.298	0.701	0.010888	35.761	54.087	5-10
WGET	0.747	0.253	0.0082834	73.917	127.23	3-15	0.267	0.733	0.0097375	33.598	51.15	5-20
	12 *h*						24 *h*					
G	0.217	0.782	0.00962	26.827	39.496	5-10	0.225	0.774	0.0092541	25.192	34.924	3-10
GT	0.184	0.815	0.0085525	26.308	38.211	2-15	0.201	0.798	0.0084863	21.996	32.053	2-15
s GE	0.197	0.803	0.0091381	26.701	39.497	2-20	0.227	0.772	0.0093044	25.277	35.019	3-15
GET	**0.177**	**0.822**	**0.0082163**	**25.259**	**37.452**	**2-20**	0.201	0.798	0.0082707	23.224	33.016	3-15
WG	0.208	0.791	0.0096533	27.557	40.595	2-20	0.221	0.778	0.0093121	22.925	33.576	2-15
WGT	0.181	0.819	0.0083917	25.495	37.85	2-15	**0.190**	**0.809**	**0.0080138**	**21.703**	**31.147**	**2-15**
WGE	0.213	0.787	0.0098618	27.612	41.031	2-10	0.205	0.794	0.0084148	23.999	33.302	3-15
WGET	0.188	0.812	0.0087126	25.698	38.567	2-15	0.212	0.787	0.0089291	22.635	32.878	2-10
	Water Prediction											
	1 *h*						6 *h*					
W	0.899	0.101	0.003766	88.697	148.69	2-10	0.370	0.629	0.0048517	36.899	55.073	5-20
WT	0.922	0.078	0.0041518	82.817	160.02	3-20	**0.361**	**0.639**	**0.0047232**	**36.517**	**54.338**	**5-20**
WE	0.912	0.088	0.0038182	81.573	149.71	2-20	0.378	0.621	0.0049517	36.928	55.637	5-20
WET	0.904	0.096	0.0037697	80.518	152.94	5-15	0.385	0.615	0.0050367	37.842	56.113	5-10
WG	0.911	0.089	0.0038144	82.101	149.64	2-10	0.385	0.614	0.0050453	37.718	56.161	5-10
WGT	0.918	0.081	0.0038472	81.913	150.28	2-15	0.386	0.613	0.0050571	37.804	56.227	5-20
WGE	0.906	0.093	0.0037966	81.713	149.29	2-20	0.376	0.623	0.0049274	36.911	55.501	5-20
WGET	**0.890**	**0.110**	**0.0037117**	**81.12**	**151.76**	**5-20**	0.364	0.636	0.0047659	36.636	54.584	5-15
	12 *h*						24 *h*					
W	**0.272**	**0.728**	**0.0040105**	**25.219**	**36.221**	**2-15**	**0.567**	**0.430**	**0.0069617**	**19.645**	**25.386**	**5-20**
WT	0.290	0.709	0.0042789	25.36	37.413	2-10	0.612	0.385	0.0075076	20.736	26.363	5-20
WE	0.277	0.722	0.0040941	26.611	36.597	2-20	0.616	0.381	0.0075551	20.302	26.446	5-20
WET	0.284	0.716	0.0041886	26.221	37.017	2-20	0.641	0.356	0.0078689	20.8	26.99	5-20
WG	0.292	0.708	0.0043068	26.003	37.535	2-20	0.606	0.391	0.0074327	20.381	26.231	5-15
WGT	0.294	0.705	0.004338	26.424	37.671	2-20	0.653	0.344	0.0080148	21.182	27.239	5-20
WGE	0.293	0.706	0.0043279	25.639	37.627	2-20	0.612	0.385	0.0075092	20.214	26.366	5-20
WGET	0.297	0.703	0.0043785	26.479	37.846	2-15	0.622	0.376	0.0076259	20.444	26.57	2-20

W = water G = natural gas E = electric power T = temperature

Table 14: For each resources combination, and time resolution, the higher performance achieved by testing al l available parameters combinations for the ELM-LIN approach is reported. The "Par." column indicates the corresponding best parameters: lags and maximum depth, respectively.

ELM-LIN	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Par.
Natural Gas Prediction												
	1 *h*						6 *h*					
G	0.822	0.178	0.0091127	82.511	133.45	$3\text{-}2^{-1}$	0.311	0.689	0.011334	39.351	55.184	$5\text{-}2^{-1}$
GT	0.796	0.204	0.0088285	88.944	131.35	$3\text{-}2^{-1}$	0.282	0.718	0.010282	37.606	52.56	5-2
GE	0.798	0.202	0.0088455	85.923	131.48	$3\text{-}2^{15}$	0.276	0.723	0.010086	38.238	52.056	5-1
GET	0.789	0.210	0.0087538	88.344	130.79	3-1	0.260	0.740	0.0094843	36.96	50.48	5-2
WG	0.810	0.190	0.0089841	81.734	132.5	$3\text{-}2^{15}$	0.288	0.712	0.010498	38.336	53.109	$5\text{-}2^{15}$
WGT	0.769	0.231	0.0085239	89.163	129.06	3-2	0.272	0.727	0.009936	37.272	51.668	5-2
WGE	0.762	0.238	0.0084462	84.827	128.48	$3\text{-}2^{15}$	0.269	0.731	0.00982	37.705	51.366	$5\text{-}2^{5}$
WGET	**0.747**	**0.253**	**0.0082835**	**88.517**	**127.23**	**3-2**	**0.255**	**0.745**	**0.0092892**	**36.722**	**49.958**	$\mathbf{5\text{-}2^{3}}$
	12 *h*						24 *h*					
G	0.236	0.763	0.010471	28.609	41.206	5-8	0.232	0.767	0.0095408	25.443	35.461	$3\text{-}2^{2}$
GT	0.204	0.795	0.0094777	28.69	40.224	$2\text{-}2^{15}$	0.210	0.789	0.0086117	24.8	33.69	$3\text{-}2^{9}$
GE	0.217	0.782	0.010086	29.971	41.495	$2\text{-}2^{-1}$	0.223	0.776	0.0094073	23.924	33.747	2-2
GET	0.193	0.807	0.0089418	28.025	39.071	$2\text{-}2^{15}$	0.206	0.793	0.0084628	25.067	33.397	$3\text{-}2^{15}$
WG	0.219	0.781	0.0096946	28.316	39.649	5-8	0.226	0.773	0.0092712	25.226	34.956	$3\text{-}2^{3}$
WGT	0.200	0.800	0.0092725	28.389	39.787	$2\text{-}2^{15}$	0.209	0.790	0.0085665	24.764	33.601	$3\text{-}2^{5}$
WGE	0.212	0.787	0.00984	29.489	40.986	2-1	0.220	0.779	0.0090519	25.215	34.54	$3\text{-}2^{7}$
WGET	**0.192**	**0.808**	**0.0088996**	**27.906**	**38.979**	$\mathbf{2\text{-}2^{15}}$	**0.204**	**0.795**	**0.0083769**	**24.917**	**33.227**	$\mathbf{3\text{-}2^{15}}$
Water Prediction												
	1 *h*						6 *h*					
W	0.960	0.040	0.004321	88.335	163.25	$3\text{-}2^{-1}$	0.535	0.465	0.007003	45.49	66.166	$5\text{-}2^{-1}$
WT	0.878	0.122	0.0036788	95.575	146.96	$2\text{-}2^{-1}$	0.471	0.528	0.0061682	43.468	62.097	$5\text{-}2^{-1}$
WE	0.871	0.129	0.0036486	89.071	146.35	$2\text{-}2^{-3}$	0.498	0.501	0.0065277	44.694	63.881	$5\text{-}2^{-1}$
WET	0.842	0.158	0.0035258	93.962	143.87	$2\text{-}2^{15}$	0.442	0.557	0.0057908	43.114	60.167	$5\text{-}2^{-1}$
WG	0.949	0.051	0.0042737	88.971	162.35	$3\text{-}2^{-3}$	0.513	0.487	0.0067152	45.457	64.792	$5\text{-}2^{-1}$
WGT	0.869	0.131	0.0036386	95.667	146.15	$2\text{-}2^{15}$	0.442	0.558	0.005788	43.436	60.153	$5\text{-}2^{-1}$
WGE	0.868	0.132	0.0036345	89.186	146.07	$2\text{-}2^{-1}$	0.480	0.519	0.00629	44.42	62.707	$5\text{-}2^{-1}$
WGET	**0.837**	**0.163**	**0.0035072**	**94.119**	**143.49**	$\mathbf{2\text{-}2^{15}}$	**0.419**	**0.580**	**0.0054936**	**42.642**	**58.603**	$\mathbf{5\text{-}2^{-1}}$
	12 *h*						24 *h*					
W	0.346	0.653	0.0061225	29.386	42.551	5-8	0.633	0.364	0.0077667	20.671	26.814	$5\text{-}2^{-1}$
WT	0.324	0.675	0.0057382	29.027	41.193	$5\text{-}2^{15}$	0.642	0.355	0.0078817	21.231	27.012	5-2
WE	0.349	0.650	0.0061748	29.732	42.732	$5\text{-}2^{15}$	**0.628**	**0.369**	**0.0077079O**	**20.709**	**26.712**	$\mathbf{5\text{-}2^{-1}}$
WET	0.325	0.674	0.0057506	29.212	41.238	$5\text{-}2^{15}$	0.637	0.361	0.0078098	21.217	26.888	5-1
WG	0.344	0.655	0.0060932	29.315	42.448	5-2	0.641	0.356	0.007865	20.866	26.983	$5\text{-}2^{-1}$
WGT	**0.314**	**0.686**	**0.0055546**	**28.518**	**40.529**	**5-2**	0.662	0.335	0.0081206	21.615	27.418	5-2
WGE	0.348	0.651	0.0061689	29.751	42.711	$5\text{-}2^{15}$	0.637	0.360	0.0078148	20.878	26.897	$5\text{-}2^{-1}$
WGET	0.316	0.683	0.0055987	28.819	40.69	5-128	0.658	0.339	0.0080731	21.451	27.338	5-1

W = water G = natural gas E = electric power T = temperature

Table 15: For each resources combination, and time resolution, the higher performance achieved by testing al l available parameters combinations for the ELM-RBF approach is reported. The "Par." column indicates the corresponding best parameters: lags and maximum depth, respectively.

ELM-RBF	NMSE	R^2	MSE	MAPE	RRMSE	Par.	NMSE	R^2	MSE	MAPE	RRMSE	Par.
	Natural Gas Prediction											
	1 *h*						6 *h*					
G	0.734	0.266	8.142E-3	86.364	126.14	$3\text{-}2\text{-}2^{-3}$	0.245	0.755	0.008925	33.957	48.969	$5\text{-}2^{3}\text{-}2^{-1}$
GT	0.651	0.349	0.007425	80.715	120.56	$5\text{-}1\text{-}2^{-3}$	0.198	0.801	0.0072398	31.052	44.104	$5\text{-}2^{-5}\text{-}2^{-1}$
GE	0.697	0.303	0.0077302	82.875	122.91	$3\text{-}1\text{-}2^{-3}$	0.210	0.790	0.0076542	32.875	45.349	$5\text{-}2^{-5}\text{-}2$
GET	0.631	0.369	0.0069994	78.28	116.95	$3\text{-}1\text{-}2^{-3}$	0.184	0.816	0.0066953	30.973	42.414	$5\text{-}2^{-5}\text{-}2$
WG	0.693	0.307	0.007688	82.236	122.57	$3\text{-}2^{3}\text{-}2^{-3}$	0.214	0.786	0.0078159	32.351	45.825	$5\text{-}2\text{-}2^{-1}$
WGT	0.632	0.368	0.0072097	79.04	118.8	$5\text{-}1\text{-}2^{-3}$	0.188	0.812	0.0068518	30.574	42.906	$5\text{-}2^{3}\text{-}2^{-1}$
WGE	0.666	0.334	0.007381	80.583	120.1	$3\text{-}1\text{-}2^{-3}$	0.197	0.803	0.0071859	31.852	43.94	$5\text{-}2^{-5}\text{-}2$
WGET	**0.620**	**0.380**	**0.0068781**	**77.109**	**115.94**	**3-1-2^{-3}**	**0.180**	**0.820**	**0.0065539**	**30.747**	**41.963**	**5-2^{-5}-2**
	12 *h*						24 *h*					
G	0.209	0.790	0.009712	28.01	40.717	$2\text{-}2^{3}\text{-}2^{-3}$	0.224	0.775	0.0094353	23.274	33.797	$2\text{-}2^{3}\text{-}2^{3}$
GT	0.159	0.840	0.0073842	24.887	35.505	$2\text{-}2^{3}\text{-}2^{-3}$	**0.176**	**0.823**	**0.0074151**	**21.375**	**29.961**	**2-2^{3}192-2^{3}**
GE	0.177	0.822	0.0082211	26.386	37.463	$2\text{-}2^{3}\text{-}2^{-3}$	0.219	0.780	0.0089799	24.923	34.403	$3\text{-}2^{15}\text{-}2^{7}$
GET	0.155	0.845	0.0071919	25.199	35.04	$2\text{-}2^{-5}\text{-}2^{-1}$	0.185	0.815	0.0075781	23.032	31.603	$3\text{-}2^{15}\text{-}2^{7}$
WG	0.176	0.824	0.0081535	26.004	37.309	$2\text{-}2^{3}\text{-}2^{-3}$	0.206	0.794	0.008672	23.098	32.401	$2\text{-}2^{11}\text{-}2$
WGT	**0.147**	**0.853**	**0.0068143**	**23.814**	**34.108**	**2-2^{3}-2^{-3}**	0.177	0.822	0.0074733	21.033	30.079	$2\text{-}2^{3}192\text{-}2^{3}$
WGE	0.176	0.823	0.0081803	26.412	37.37	$2\text{-}2\text{-}2^{-3}$	0.214	0.785	0.0087713	24.934	34	$3\text{-}2^{11}\text{-}2^{-5}$
WGET	0.151	0.848	0.0070217	24.957	34.623	$2\text{-}2^{-5}\text{-}2^{-1}$	0.187	0.812	0.0076811	23.346	31.818	$3\text{-}2^{15}\text{-}2^{7}$
	Water Prediction											
	1 *h*						6 *h*					
W	0.824	0.175	0.0034534	92.37	142.38	$2\text{-}2^{11}\text{-}2$	0.328	0.672	0.0042913	35.887	51.795	$5\text{-}1\text{-}2^{-3}$
WT	0.807	0.193	0.0033631	91.823	144.46	$5\text{-}1\text{-}2^{-3}$	0.324	0.676	0.0042388	36.348	51.477	$5\text{-}1\text{-}2^{-3}$
WE	0.802	0.198	0.0033595	90.219	140.43	$2\text{-}2^{-1}\text{-}2^{-3}$	0.339	0.661	0.0044403	36.405	52.686	$5\text{-}1\text{-}2^{-3}$
WET	0.782	0.218	0.0032602	89.384	142.23	$5\text{-}2\text{-}2^{-1}$	0.325	0.674	0.0042599	36.965	51.605	$5\text{-}2^{-1}\text{-}2^{-1}$
WG	0.806	0.194	0.0033616	91.374	144.43	$5\text{-}1\text{-}2^{-3}$	0.330	0.669	0.0043237	35.804	51.99	$5\text{-}2\text{-}2^{-3}$
WGT	0.786	0.214	0.0032753	90.329	142.56	$5\text{-}2^{-1}\text{-}2^{-3}$	0.332	0.667	0.0043534	37.138	52.168	$5\text{-}1\text{-}2^{-1}$
WGE	0.787	0.213	0.0032964	89.076	139.11	$2\text{-}2^{-1}\text{-}2^{-3}$	0.326	0.674	0.0042683	35.928	51.656	$5\text{-}1\text{-}2^{-1}$
WGET	**0.765**	**0.235**	**0.0031905**	**88.33**	**140.7**	**5-1-2^{-1}**	**0.319**	**0.681**	**0.0041728**	**36.059**	**51.074**	**5-2^{-1}-2^{-1}**
	12 *h*						24 *h*					
W	0.286	0.713	0.0042267	26.325	37.184	$2\text{-}2^{7}\text{-}2^{-3}$	**0.575**	**0.422**	**0.007061**	**19.823**	**25.567**	**5-2^{3}-2^{-1}**
WT	0.284	0.716	0.0050204	27.217	38.531	$5\text{-}1\text{-}2^{-1}$	0.622	0.375	0.0076354	20.865	26.586	$5\text{-}2\text{-}2^{3}$
WE	0.282	0.717	0.0050016	27.399	38.459	$5\text{-}1\text{-}2^{-1}$	0.581	0.416	0.0071341	19.985	25.699	$5\text{-}2^{-5}\text{-}2$
WET	**0.279**	**0.720**	**0.0049433**	**27.209**	**38.234**	**5-1-2^{-1}**	0.619	0.378	0.0075922	20.729	26.511	$5\text{-}2\text{-}2^{3}$
WG	0.284	0.716	0.0050256	27.34	38.551	$5\text{-}1\text{-}2^{-1}$	0.605	0.392	0.0074219	20.633	26.212	$5\text{-}2^{15}\text{-}2^{7}$
WGT	0.285	0.714	0.0050493	27.498	38.642	$5\text{-}1\text{-}2^{-1}$	0.640	0.357	0.0078577	21.106	26.971	$5\text{-}2\text{-}2^{3}$
WGE	0.284	0.715	0.0050333	27.399	38.58	$5\text{-}1\text{-}2^{-1}$	0.588	0.410	0.0072095	20.184	25.834	$5\text{-}2^{11}\text{-}2^{-5}$
WGET	0.285	0.715	0.0050376	27.473	38.597	$5\text{-}1\text{-}2^{-1}$	0.636	0.361	0.0078021	20.977	26.875	$5\text{-}2^{3}\text{-}2^{3}$

W = water G = natural gas E = electric power T = temperature

3 AMPds Results - Temporal Clustering

Table 16: Best clustering results and corresponding features combination achieved for the Natural Gas Consumption in residential scenario, AMPds dataset. The column "**Par.**" reports the number of Gaussians, the states and Gaussians number for GMM and HMM models, respectively.

Cluster	**Res.**	**AUC (%)**	**SD**	**Par.**	**Features Combination**
	GMM Model				
Low	1	74.75	7.73	32	*Da+En+Lw+dLw*
High	1	65.94	14.28	256	*Lw+Ma+En*
C1	1	80.19	13.83	256	*Lw+Ma+En*
C2	1	74.83	12.68	128	*Da+dEn+En*
Overall	1	73.93	12.13	-	-
Low	10	81.86	14.51	128	*En+Ma+dMa*
High	10	63.04	19.96	256	*Da+Lw+dMa*
C1	10	69.64	14.23	64	*Lw+Ma+En*
C2	10	75.87	11.86	256	*Lw+Da+dMa*
Overall	10	72.60	15.14	-	-
Low	30	69.36	14.30	8	*Da+Lw*
High	30	64.65	18.54	256	*Da+Ma*
C1	30	73.79	13.79	128	*Da*
C2	30	64.44	14.22	64	*Lw+Ma*
Overall	30	68.06	15.24	-	-
	HMM Model				
Low	1	91.64	1.98	1-256	*Ma+En+dMa*
High	1	67.25	18.42	2-128	*Ma+En+Lw*
C1	1	80.34	13.91	1-256	*Da+En+Lw*
C2	1	87.64	8.82	2-64	*Ma+En+Lw*
Overall	1	81.72	10.78	-	-
Low	10	75.36	16.78	1-256	*Da+dLw+Lw+dWe*
High	10	64.82	19.43	2-128	*Da+En+We*
C1	10	71.19	13.64	1-256	*Da+En+dMa+We*
C2	10	78.67	10.34	1-128	*En+Da+dLw*
Overall	10	72.51	15.05	-	-
Low	30	71.42	16.09	1-8	*Ma+En*
High	30	67.09	19.11	2-64	*Da+En+dMa*
C1	30	76.21	16.76	1-64	*En+Ma+dMa*
C2	30	66.39	22.71	4-4	*Ma+Lw*
Overall	30	70.28	18.67	-	-

Table 17: Best clustering results and corresponding features combination achieved for the Water Consumption in residential scenario, AMPds dataset. The column "**Par.**" reports the number of Gaussians, the states and Gaussians number for GMM and HMM models, respectively.

Cluster	**Res.**	**AUC (%)**	**SD**	**Par.**	**Features Combination**
	GMM Model				
Low	1	91.47	4.50	16	*Da+dWe+dLw+dMa*
High	1	57.07	16.70	256	*Da+dEn+dWe*
C1	1	65.52	14.83	256	*Da+dEn+En+Ma*
C2	1	82.32	11.69	256	*Lw+Ma+En*
Overall	1	74.10	11.93	-	-
Low	10	93.36	0.33	16	*da+dLw+Ma*
High	10	58.13	17.76	256	*Da+En+Ma+dMa*
C1	10	73.10	10.92	256	*Da+En+Lw*
C2	10	69.31	13.33	256	*Da+En+Lw*
Overall	10	73.48	10.58	-	-
Low	30	83.45	8.97	8	*Da+Lw+Ma*
High	30	57.28	20.52	32	*dDa+En*
C1	30	63.79	14.65	256	*Lw+Da*
C2	30	59.90	16.88	256	*Da*
Overall	30	66.11	15.25	-	-
	HMM Model				
Low	1	92.48	2.87	4-64	*Lw+Ma+dMa*
High	1	61.90	20.06	2-64	*Ma+dMa*
C1	1	82.06	12.98	4-16	*Ma+En+dMa*
C2	1	92.90	7.70	3-128	*Ma+En*
Overall	1	82.34	10.90	-	-
Low	10	99.10	0.49	3-8	*Da+dEn+dLw+Ma+dMa*
High	10	61.50	19.00	2-256	*Da+En+dMa*
C1	10	80.93	12.02	4-256	*En+Da+We*
C2	10	73.16	12.33	1-256	*Da+En+Lw*
Overall	10	78.67	10.96	-	-
Low	30	95.50	1.87	2-8	*Ma+dMa*
High	30	64.01	21.38	4-2	*Ma+En*
C1	30	73.07	11.86	4-256	*Da+En*
C2	30	68.44	17.34	3-128	*Da+En*
Overall	30	75.26	14.61	-	-

4 DFID Results - Temporal Clustering

Table 18: Best clustering results and corresponding features combination achieved for the office building scenario, DFID datasets. The column "**Par.**" reports the number of Gaussians, the states and Gaussians number for GMM and HMM models, respectively.

Cluster	Res.	AUC (%)	SD	Par.	Features Combination
	Aber. Gas GMM Model				
Low	30	38.50	21.74	2	*dLw+dWe*
High	30	57.91	16.67	256	*Lw+dLw+Ma*
C1	30	49.44	2.47	128	*dLw+Da+Lw*
C2	30	60.50	15.61	64	*dLw+Da+Lw+dWe+dMa*
Overall	30	51.59	14.12	-	-
	Aber. Gas HMM Model				
Low	30	94.17	0.87	3-2	*We+Da+Ma+En*
High	30	68.23	23.92	4-64	*we+Lw+dMa+Ma*
C1	30	88.56	4.11	2-4	*We+Da*
C2	30	70.74	23.93	2-32	*dWe+Ma*
Overall	30	80.43	13.21	-	-
	Aber. Water GMM Model				
Low	30	70.45	21.12	8	*dWe+Da*
High	30	71.23	16.05	256	*Da+dMa+Ma+En*
C1	30	84.64	14.38	16	*dWe+Da+Lw+Ma*
C2	30	75.23	11.21	256	*Da+dMa+Ma*
Overall	30	75.39	15.69	-	-
	Aber. Water HMM Model				
Low	30	93.20	7.67	4-16	*dWe+Da+dDa+Lw+Ma*
High	30	79.03	19.24	4-64	*Da+Ma+dMa*
C1	30	93.52	4.49	3-16	*Lw+dW+Da*
C2	30	85.36	13.53	4-128	*Ma+Da+dMa*
Overall	30	87.78	11.23	-	-
	Overall Gas GMM Model				
Low	30	70.75	20.83	2	*dLw+dWe+We*
High	30	56.81	18.08	256	*Lw+dLw+We+dMa+dWe*
C1	30	61.56	17.16	32	*Da+dLw+dWe+Lw+dMa*
C2	30	73.69	6.94	2	*Lw+En+Ma*
Overall	30	65.70	15.75	-	-
	Overall Gas HMM Model				
Low	30	93.82	1.81	3-2	*dMa+Ma*
High	30	64.05	19.78	3-32	*Lw+dMa+We*
C1	30	71.31	21.27	3-16	*dLw+Da+We*
C2	30	93.63	7.89	4-4	*We+En*
Overall	30	80.70	12.69	-	-
	White. Water GMM Model				
Low	30	99.29	0.52	256	*Da+dWe+We+dMa+Ma*
High	30	78.41	15.01	128	*Da+dWe+dEn*
C1	30	89.47	10.78	16	*Da+dMa*
C2	30	84.51	9.19	256	*Da+dEn+dMa*
Overall	30	87.92	8.88	-	-
	White. Water HMM Model				
Low	30	99.37	0.47	1-256	*Da+dWe+We+Ma+En*
High	30	85.74	115.44	4-16	*Da+dMa+Ma*
C1	30	96.23	5.36	2-8	*dMa+Da*
C2	30	92.67	7.04	3-128	*Da+dMa*
Overall	30	93.50	7.08	-	-

5 EIA Results

Table 19: Best results achieved for each technique applied to EIA datasets. The column marked with "Param." reports the parameters combination that achieves the best result for the corresponding approach.

	Illinois						**Illinois&T**					
	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Parameters	NMSE	R^2	MSE	MAPE(%)	RRMSE(%)	Parameters
ARIMA	0.199	0.799	1.620E-2	31.587	39.458	2 -0 -1						
ANN	0.053	0.946	4.692E-3	15.931	22.685	5-5	0.067	0.932	5.945E-3	18.278	26.118	2-6
DBN	0.058	0.940	4.453E-3	16.168	23.873	5-5-500-10	**0.055**	**0.944**	**4.167E-3**	**15.598**	**23.095**	**5-7-500-10**
ESN	0.097	0.902	7.338E-3	22.179	30.647	5-40-0-1/2-0	0.069	0.930	5.259E-3	19.106	25.944	5-20-0-1/2-1e-15
SVR	0.066	0.933	5.033E-3	17.056	25.380	$5\text{-}2^{11}\text{-}2^{-3}$	0.066	0.933	5.022E-3	16.755	25.354	$5\ 2^5\text{-}2^{-3}$
GP	0.081	0.918	6.131E-3	18.705	28.012	5-10	0.092	0.907	7.251E-3	19.046	27.008	3-10
EKF-GP	**0.044**	**0.955**	**3.374E-3**	**13.251**	**20.780**	**5-10**	0.073	0.926	5.794E-3	17.285	24.141	3-10
ELM-LIN	0.203	0.795	1.654E-2	29.814	39.819	2 - 2	0.090	0.909	9.794E-3	21.084	29.488	$2 - 2^{15}$
ELM-RBF	0.062	0.937	4.737E-3	16.247	24.623	$5\text{ -}2^5\text{ - }2^{-1}$	0.064	0.936	4.841E-3	16.668	24.893	$5 - 2^3\text{ -}2$
	Louisiana						**Louisiana&T**					
ARIMA	0.307	0.690	1.680E-2	34.953	51.588	2 - 0 -1						
ANN	**0.085**	**0.914**	**4.498E-3**	**21.400**	**30.604**	**2-6**	0.066	0.933	3.633E-3	15.718	23.518	3-7
DBN	0.148	0.850	7.515E-3	25.792	36.270	3-2-500-10	0.072	0.927	2.535E-3	15.753	25.171	5-8-500-10
ESN	0.132	0.866	7.233E-3	24.724	33.861	2-15-14-1-0	0.097	0.902	5.294E-3	20.514	28.968	2-20-1-1/2-1e-4
SVR	0.096	0.903	3.364E-3	19.722	28.994	5-2-23	0.070	0.929	2.478E-3	15.774	24.888	5-2-2
GP	0.132	0.867	6.682E-3	21.340	34.202	3-10	0.095	0.904	4.831E-3	19.293	29.080	3-20
EKF-GP	0.086	0.913	4.354E-3	15.855	27.607	3-10	**0.061**	**0.938**	**3.087E-3**	**15.724**	**23.245**	**3-20**
ELM-LIN	0.292	0.704	1.600E-2	32.592	50.360	$2\text{ -}2^{15}$	0.155	0.843	5.475E-3	26.292	36.992	5 - 2
ELM-RBF	0.103	0.896	3.611E-3	20.401	30.042	$5\text{ -}2^5\text{ -}2^{-3}$	0.066	0.934	2.315E-3	15.246	24.053	$5\text{ - }2^7\text{ -}2$
	Maine						**Maine&T**					
ARIMA	0.158	0.841	7.700E-3	25.725	34.351	2 - 0 - 2						
ANN	**0.051**	**0.948**	**2.517E-3**	**16.094**	**21.528**	**5-5**	0.065	0.934	2.763E-3	15.939	23.330	2-6
DBN	0.111	0.888	4.774E-3	22.820	29.642	3-4-100-100	0.092	0.907	3.946E-3	19.959	26.950	3-3-500-10
ESN	0.074	0.925	3.194E-3	19.029	24.244	3-40-10-1/2-1e-14	0.079	0.921	3.820E-3	18.837	24.250	2-35-10-1/3-1e-4
SVR	0.066	0.933	2.844E-3	17.464	22.878	$3\text{-}2^{13}\text{-}2^{-3}$	0.045	0.955	1.925E-3	14.467	18.822	$3\text{-}2^5\text{-}2^{-1}$
GP	0.095	0.903	4.099E-3	19.547	27.466	3-15	0.062	0.938	2.651E-3	15.634	22.090	3-20
EKF-GP	0.052	0.947	2.253E-3	14.692	20.365	3-15	**0.030**	**0.970**	**1.281E-3**	**11.267**	**15.353**	**3-20**
ELM-LIN	0.147	0.851	6.341E-3	25.214	34.163	$3 - 2^{15}$	0.110	0.889	5.332E-3	21.242	28.650	$2 - 2^9$
ELM-RBF	0.072	0.927	3.110E-3	18.376	23.924	$3 - 2^{11}\text{ -}2$	0.048	0.952	2.060E-3	15.025	19.472	$3\text{ -}2^9\text{ -}2$

List of Publications

[1] S. Spinsante, S. Squartini, P. Russo, A. De Santis, M. Severini, M. Fagiani, V. Di Mattia, and R. Minerva, "Enabling Technologies for IoT-Oriented Sustainable Smart Gas and Water Grids: Communication Protocols and Data Analysis," in *Internet of Things Concepts, Technologies, Applications, and Implementations*, CRC Press, Florida, USA. Minor revision.

[2] R. Bonfigli, E. Principi, S. Squartini, M. Fagiani, M. Severini, and F. Piazza, "User-aided Footprint Extraction for Appliance Modelling in Non-Intrusive Load Monitoring," in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, December 2016. To appear.

[3] M. Fagiani, S. Squartini, L. Gabrielli, M. Severini, and F. Piazza, "A statistical framework for automatic leakage detection in smart water and gas grids," *Energies*, vol. 9, no. 9, p. 665, 2016.

[4] M. Fagiani, S. Squartini, R. Bonfigli, M. Severini, and F. Piazza, "Exploiting Temporal Features and Pressure Data for Automatic Leakage Detection in Smart Water Grids," in *2016 IEEE World Congress on Computational Intelligence (WCCI)*, July 2016. To appear.

[5] R. Bonfigli, M. Severini, S. Squartini, M. Fagiani, and F. Piazza, "Improving the performance of the AFAMAP algorithm for Non-Intrusive Load Monitoring," in *2016 IEEE World Congress on Computational Intelligence (WCCI)*, July 2016. To appear.

[6] M. Severini, A. Scorrano, S. Squartini, M. Fagiani, and F. Piazza, "SW Framework for simulation and evaluation of partial shading effects in configurable PV systems," in *Environment and Electrical Engineering (EEEIC), 2016 IEEE 16th International Conference on*, June 2016. To appear.

[7] M. Fagiani, S. Squartini, M. Severini, and F. Piazza, "A novelty detection approach to identify the occurrence of leakage in smart gas and water grids," in *2015 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, July 2015.

[8] M. Severini, S. Squartini, M. Fagiani, and F. Piazza, "Energy management with the support of dynamic pricing strategies in real micro-grid scenarios," in *2015 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, July 2015.

[9] R. Bonfigli, S. Squartini, M. Fagiani, and F. Piazza, "Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview," in *Environment and Electrical Engineering (EEEIC), 2015 IEEE 15th International Conference on*, pp. 1175–1180, June 2015.

[10] M. Fagiani, S. Squartini, R. Bonfigli, and F. Piazza, "Short-term load forecasting for smart water and gas grids: A comparative evaluation," in *Environment and Electrical Engineering (EEEIC), 2015 IEEE 15th International Conference on*, pp. 1198–1203, June 2015.

[11] M. Fagiani, S. Squartini, L. Gabrielli, S. Spinsante, and F. Piazza, "A review of datasets and load forecasting techniques for smart natural gas and water grids: Analysis and experiments," *Neurocomputing*, vol. 170, pp. 448 – 465, 2015.

[12] M. Fagiani, E. Principi, S. Squartini, and F. Piazza, "Signer independent isolated italian sign recognition based on hidden markov models," *Pattern Analysis and Applications*, vol. 18, no. 2, pp. 385–402, 2015.

[13] M. Fagiani, S. Squartini, L. Gabrielli, S. Spinsante, and F. Piazza, "Domestic water and natural gas demand forecasting by using heterogeneous data: A preliminary study," *Smart Innovation, Systems and Technologies*, vol. 37, pp. 185–194, 2015.

[14] M. Fagiani, S. Squartini, L. Gabrielli, M. Pizzichini, and S. Spinsante, "Computational intelligence in smart water and gas grids: An up-to-date overview," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, pp. 921–926, July 2014.

Bibliography

[1] G.K. Venayagamoorthy, "Dynamic, Stochastic, Computational, and Scalable Technologies for Smart Grids," *Computational Intelligence Magazine, IEEE*, vol. 6, no. 3, pp. 22 – 35, 2011.

[2] Luca Ardito, Giuseppe Procaccianti, Giuseppe Menga, and Maurizio Morisio, "Smart grid technologies in europe: An overview," *Energies*, vol. 6, no. 1, pp. 251–281, 2013.

[3] Pierluigi Siano, "Demand Response and Smart Grids - A Survey," *Renewable and Sustainable Energy Reviews*, vol. 30, no. 0, pp. 461 – 478, 2014.

[4] Susanna Spinsante, Stefano Squartini, Leonardo Gabrielli, Mirco Pizzichini, Ennio Gambi, and Francesco Piazza, "Wireless m-bus sensor networks for smart water grids: analysis and results," *International Journal of Distributed Sensor Networks*, vol. 2014, Article ID 579271, pp. 16, 2014.

[5] Avi Ostfeld et al., "The Battle of Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms," *Journal of Water Resources Planning and Management*, vol. 134, no. 6, pp. 556–568, 2008.

[6] M. Allen, A. Preis, M. Iqbal, and A.J. Whittle, "Water Distribution System Monitoring and Decision Support Using a Wireless Sensor Network," in *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2013 14th ACIS International Conference on*, 2013, pp. 641 – 646.

[7] D.G. Eliades, M. Kyriakou, and M.M. Polycarpou, "Sensor Placement in Water Distribution Systems Using the S-PLACE Toolkit," *Procedia Engineering*, vol. 70, no. 0, pp. 602–611, 2014.

[8] S.G. Kashid and S.A. Pardeshi, "A Survey of Water Distribution System and New Approach to Intelligent Water Distribution System," in *Networks Soft Computing (ICNSC), 2014 First International Conference on*, Aug 2014, pp. 339–344.

[9] Energy Information Administration (U.S.), *International Energy Outlook 2013 With Projections to 2040*, U.S. Government Printing Office, 2013.

[10] Gabriele Comodi, Andrea Giantomassi, Marco Severini, Stefano Squartini, Francesco Ferracuti, Alessandro Fonti, Davide Nardi Cesarini, Matteo Morodo, and Fabio Polonara, "Multi-apartment residential microgrid with electrical and thermal storage devices: Experimental analysis and simulation of energy management strategies," *Applied Energy*, vol. 137, pp. 854 – 866, 2015.

[11] A. Shabri and R. Samsudin, "A new approach for water demand forecasting based on empirical mode decomposition," in *Software Engineering Conference (MySEC), 2014 8th Malaysian*, Sept 2014, pp. 284–288.

[12] M. Akpinar and N. Yumusak, "Forecasting household natural gas consumption with arima model: A case study of removing cycle," in *Application of Information and Communication Technologies (AICT), 2013 7th International Conference on*, Oct 2013, pp. 1–6.

[13] Gang Ji, Jingcheng Wang, Yang Ge, and Huajiang Liu, "Urban Water Demand Forecasting by LS-SVM with Tuning based on Elitist Teaching-Learning-based Optimization," in *Control and Decision Conference (2014 CCDC), The 26th Chinese*, May 2014, pp. 3997 – 4002.

[14] M. Bakker, J.H.G. Vreeburg, K.M. van Schagen, and L.C. Rietveld, "A Fully Adaptive Forecasting Model for Short-term Drinking Water Demand," *Environmental Modelling & Software*, vol. 48, no. 0, pp. 141 – 151, 2013.

[15] X. Zhu and J. Chen, "Urban Water Consumption Forecast based on PQPSO-LSSVM," in *Natural Computation (ICNC), 2013 Ninth International Conference on*, July 2013, pp. 834 – 837.

[16] X. Zhu and B. Xu, "Urban Water Consumption Forecast Based on QPSO-RBF Neural Network," in *Computational Intelligence and Security, Eighth International Conference on*, 2012, pp. 233 – 236.

[17] Mohsen Nasseri, Ali Moeini, and Massoud Tabesh, "Forecasting Monthly Urban Water Demand using Extended Kalman Filter and Genetic Programming," *Expert Systems with Applications*, vol. 8, no. 6, pp. 7387 – 7395, 2011.

[18] Junping Liu and Mingqi Chang, "Application of the Grey Theory and the Neural Network in Water Demand Forecast," in *Natural Computation (ICNC), 2010 Sixth International Conference on*, 2010, vol. 2, pp. 1070 – 1073.

[19] M. Tabesh and M. Dini, "Fuzzy and Neuro-fuzzy Models for Short-term Water Demand Forecasting in Tehran," *Iranian Journal of Science & Technology, Transaction B, Engineering*, vol. 33, no. B1, pp. 61 – 77, 2009.

[20] X. Wan, Q. Zhang, and G. Dai, "Research on forecasting method of natural gas demand based on gm (1,1) model and markov chain," in *Cognitive Informatics Cognitive Computing (ICCI*CC), 2014 IEEE 13th International Conference on*, Aug 2014, pp. 436–441.

[21] W. Wang, J. Jiang, and M. Fu, "An enhanced differential evolution based grey model for forecasting urban water consumption," in *Control Conference (CCC), 2014 33rd Chinese*, July 2014, pp. 7643–7648.

[22] W. Wang, J. Jiang, and M. Fu, "A grey theory based back propagation neural network model for forecasting urban water consumption," in *Control Conference (CCC), 2014 33rd Chinese*, July 2014, pp. 7654–7659.

[23] H. Chen, Z. Wang, and P. Yu, "Study on combination forecasting of gas daily load based on the generalized dynamic fuzzy neural network," in *Control Conference (CCC), 2014 33rd Chinese*, July 2014, pp. 6235–6239.

[24] A. Azari, M. Shariaty-Niassar, and M. Alborzi, "Shot-term and Medium-term Gas Demand Load Forecasting by Neural Networks," *Iranian Journal of Chemistry and Chemical Engineering*, vol. 31, no. 4, pp. 77 – 84, 2012.

[25] B. Pang, "The Impact of Additional Weather Inputs on Gas Load Forecasting," Master's theses, Electrical and Computer Engineering, Marquette University, August 2012, Provided by the SAO/NASA Astrophysics Data System.

[26] J. Quevedo, J. Saludes, V. Puig, and J. Blanch, "Short-term demand forecasting for real-time operational control of the barcelona water transport network," in *Control and Automation (MED), 2014 22nd Mediterranean Conference of*, June 2014, pp. 990–995.

[27] R. H. Brown, S. R. Vitullo, G. F. Corliss, M. Adya, P. E. Kaefer, and R. J. Povinelli, "Detrending daily natural gas consumption series to improve short-term forecasts," in *2015 IEEE Power Energy Society General Meeting*, July 2015, pp. 1–5.

[28] A. E. U. Salam, M. Tola, M. Selintung, and F. Maricar, "A leakage detection system on the water pipe network through support vector machine method," in *Electrical Engineering and Informatics (MICEEI), 2014 Makassar International Conference on*, Nov 2014, pp. 161–165.

[29] G. Boracchi and M. Roveri, "Exploiting Self-Similarity for Change Detection," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, Jul 2014, pp. 3339–3346.

[30] J. M. A. Alkasseh, M. N. Adlan, I. Abustan, H. A. Aziz, and A. B. M. Hanif, "Applying Minimum Night Flow to Estimate Water Loss Using Statistical Modeling: A Case Study in Kinta Valley, Malaysia," *Water Resources Management*, vol. 27, no. 5, pp. 1439–1455, 2013.

[31] M.M. Gamboa-Medina, L.F. Ribeiro Reis, and R. Capobianco Guido, "Feature Extraction in Pressure Signals for Leak Detection in Water Networks," *Procedia Engineering*, vol. 70, no. 0, pp. 688–697, 2014.

[32] Stephen R. Mounce, Richard B. Mounce, and Joby B. Boxall, "Novelty Detection for Time Series Data Analysis in Water Distribution Systems using Support Vector Machines," *Journal of Hydroinformatics*, vol. 13, no. 4, pp. 672–686, 2011.

[33] M.T. Nasir, M. Mysorewala, L. Cheded, B. Siddiqui, and M. Sabih, "Measurement Error Sensitivity Analysis for Detecting and Locating Leak in Pipeline using ANN and SVM," in *Multi-Conference on Systems, Signals Devices (SSD), 2014 11th International*, Feb 2014, pp. 1–4.

[34] G. Oren and N. Y. Stroh, "Mathematical Model for Detection of Leakage in Domestic Water Supply Systems by Reading Consumption from an Analogue Water Meter," *International Journal of Environmental Science and Development*, vol. 4, no. 4, pp. 386–389, 2013.

[35] G. Sanz, R. Perez, and A. Escobet, "Leakage Localization in Water Networks using Fuzzy Logic," in *Control Automation (MED), 20th Mediterranean Conference on*, 2012, pp. 646–651.

[36] Jiangwen Wan, Yang Yu, Yinfeng Wu, Renjian Feng, and Ning Yu, "Hierarchical leak detection and localization method in natural gas pipeline monitoring sensor networks," *Sensors*, vol. 12, no. 1, pp. 189, 2012.

[37] L. A. Rossman, *The EPANET Water Quality Model*, vol. 2, Research Studies Press Ltd., Somerset, England, Coulbeck B. edition, 1993, Software available at `www.epa.gov/nrmrl/wswrd/dw/epanet.html`.

[38] Markos Markou and Sameer Singh, "Novelty Detection: A Review-Part 1: Statistical Approaches," *Signal Processing*, vol. 83, no. 12, pp. 2481–2497, 2003.

[39] Markos Markou and Sameer Singh, "Novelty Detection: A Review-Part 2: Neural Network Based Approaches," *Signal Processing*, vol. 83, no. 12, pp. 2499–2521, 2003.

[40] Marco A.F. Pimentel, David A. Clifton, Lei Clifton, and Lionel Tarassenko, "A Review of Novelty Detection," *Signal Processing*, vol. 99, pp. 215–249, 2014.

[41] Timur Chis, "Pipeline Leak Detection Techniques," *CoRR*, vol. abs/0903.4283, 2009.

[42] Pal-Stefan Murvay and Ioan Silea, "A Survey on Gas Leak Detection and Localization Techniques," *Journal of Loss Prevention in the Process Industries*, vol. 25, no. 6, pp. 966 – 973, 2012.

[43] Vladimir N. Vapnik, *Statistical Learning Theory*, Wiley-Interscience, New York, U.S.A., Sep 1998.

[44] Petra Perner, "Concepts for Novelty Detection and Handling Based on a Case-based Reasoning Process Scheme," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 1, pp. 86–91, 2009.

[45] J. Quevedo, C. Alippi, M.A. Cuguero, S. Ntalampiras, V. Puig, M. Roveri, and D. Garcia, "Temporal/Spatial Model-based Fault Diagnosis vs. Hidden Markov Models Change Detection Method: Application to the Barcelona Water Network," in *Control Automation (MED), 21st Mediterranean Conference on*, 2013, pp. 394 – 400.

[46] R. Cardell-Oliver, "Discovering Water Use Activities for Smart Metering," in *Intelligent Sensors, Sensor Networks and Information Processing, IEEE Eighth International Conference on*, 2013, pp. 171 – 176.

[47] Stephen Makonin, Bradley Ellert, Ivan V. Bajic, and Fred Popowich, "Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014," *Scientific Data*, vol. 3, no. 160037, pp. 1–12, 2016.

[48] Giacomo Boracchi, Vicenç Puig, and Manuel Roveri, "A Hierarchy of Change-Point Methods for Estimating the Time Instant of Leakages in Water Distribution Networks," in *Artificial Intelligence Applications and Innovations*, Harris Papadopoulos, AndreasS. Andreou, Lazaros Iliadis, and Ilias Maglogiannis, Eds., vol. 412 of *IFIP Advances in Information and Communication Technology*, pp. 615 – 624. Springer Berlin Heidelberg, 2013.

[49] Department for International Development, "Live Data Page for Energy and Water Consumption," Online: accessed on 2015-10-30.

[50] U.S. Energy Information Administration (EIA), "Natural Gas Consumption by End Use," 2014, Online: accessed on 2014-01-30.

[51] Stephen Makonin, Fred Popowich, Lyn Bartram, Bob Gill, and Ivan V. Bajic, "AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research," in *Electrical Power and Energy Conference, IEEE*, 2013, pp. 1 – 6.

[52] National Climatic Data Center (NCDC), "U.S. Local Climatological Data," 2014, Online: accessed on 2014-01-30.

[53] Wikipedia, "Multipolar neuron — Wikipedia, the free encyclopedia," 2013, Online: accessed on 14-10-2016.

[54] Warren S. McCulloch and Walter Pitts, "A logical calculus of the ideas immanent in nervous activity," *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[55] G E Hinton and R R Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol. 313, no. 5786, pp. 504 – 507, July 2006.

[56] Geoffrey E. Hinton and Simon Osindero, "A Fast Learning Algorithm for Deep Belief Nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, July 2006.

[57] Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle, "Greedy Layer-Wise Training of Deep Networks," *Advances in Neural Information Processing Systems*, vol. 19, pp. 153 – 160, 2007.

[58] Herbert Jaeger, "The "Echo State" Approach to Analysing and Training Recurrent Neural Networks," GMD Technical Report 148, German National Research Center for Information Technology, Bonn, Germany, 2001.

[59] H. Jaeger, "Echo State Network," *Scholarpedia*, , no. 9, pp. 2330, 2007, revision 138672.

[60] Mantas Lukoševičius, "A Practical Guide to Applying Echo State Networks," in *Neural Networks: Tricks of the Trade*, Grégoire Montavon, GenevièveB. Orr, and Klaus-Robert Müller, Eds., vol. 7700 of *Lecture Notes in Computer Science*, pp. 659 – 686. Springer Berlin Heidelberg, 2012.

[61] Yoshua Bengio, "Learning deep architectures for ai," *Found. Trends Mach. Learn.*, vol. 2, no. 1, pp. 1–127, Jan. 2009.

[62] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proceedings of the 25th International Conference on Machine Learning*, New York, NY, USA, 2008, ICML '08, pp. 1096–1103, ACM.

[63] Cathy Yeh, "Support vector machines for classification," 2015, Online: accessed on 21-11-2016.

[64] Corinna Cortes and Vladimir Vapnik, "Support-Vector Networks," *Machine Learning*, vol. 20, no. 3, pp. 273 – 297, 1995.

[65] Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer, Secaucus, NJ, USA, 2006.

[66] Vladimir N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag New York, Inc., New York, NY, USA, 1995.

[67] Bernhard Schölkopf, John C. Platt, John C. Shawe-Taylor, Alex J. Smola, and Robert C. Williamson, "Estimating the Support of a High-Dimensional Distribution," *Neural Comput.*, vol. 13, no. 7, pp. 1443–1471, July 2001.

[68] George Edward Pelham Box, Gwilym Jenkins, and Gregory C. Reinsel, *Time Series Analysis, Forecasting and Control*, John Wiley & Sons, Hoboken, N. J., 4th edition edition, 2007.

[69] Vladan Babovic and Michael B. Abbott, "The Evolution of Equations from Hydraulic Data Part I: Theory," *Journal of Hydraulic Research*, vol. 35, no. 3, pp. 397 – 410, 1997.

[70] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME - Journal of Basic Engineering*, , no. 82 (Series D), pp. 35 – 45, 1960.

[71] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme Learning Machine: Theory and Applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.

[72] Guang-Bin Huang, Hongming Zhou, Xiaojian Ding, and Rui Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 42, no. 2, pp. 513–529, April 2012.

[73] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, Berkeley, Calif., 1967, pp. 281–297, University of California Press.

[74] Douglas A. Reynolds, Thomas F. Quatieri, and Robert B. Dunn, "Speaker Verification Using Adapted Gaussian Mixture Models," *Digital Signal Processing*, vol. 10, no. 1-3, pp. 19–41, 2000.

[75] L. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb 1989.

[76] Daniel Jurafsky and James H. Martin, *Speech and Language Processing*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2nd edition, 2009.

[77] L. Baum, "An inequality and associated Maximization Technique in Statistical Estimation for Probabilistic Functions of Markov Processes," *Inequalities*, vol. 3, pp. 1–8, 1972.

[78] Masayuki Tanaka, "Deep Neural Network," 2014, MATLAB Central File Exchange.

[79] GeoffreyE. Hinton, "A Practical Guide to Training Restricted Boltzmann Machines," in *Neural Networks: Tricks of the Trade*, Grégoire Montavon, GenevièveB. Orr, and Klaus-Robert Müller, Eds., vol. 7700 of *Lecture Notes in Computer Science*, pp. 599 – 619. Springer Berlin Heidelberg, 2012.

[80] Mantas Lukoševičius, "Educational Echo State Network implementations," 2014, Online: accessed on 14-10-2014.

[81] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.

[82] Sara Silva and Jonas Almeida, "Gplab - A Genetic Programming Toolbox for MATLAB," in *In Proc. of the Nordic MATLAB Conference (NMC-2003)*, 2005, pp. 273 – 278.

[83] Jouni Hartikainen, Arno Solin, and Simo Särkkä, *Optimal Filtering with Kalman Filters and Smoothers - A Manual for MATLAB Toolbox EKF/UKF Version 1.3*, Department of Biomedical Engineering and Computational Science, Aalto University School of Science, P.O.Box 1100, FI-00076 AALTO, Espoo, Finland, August 2011.

[84] Neil D. Bennett, Barry F.W. Croke, Giorgio Guariso, Joseph H.A. Guillaume, Serena H. Hamilton, Anthony J. Jakeman, Stefano Marsili-Libelli, Lachlan T.H. Newham, John P. Norton, Charles Perrin, Suzanne A. Pierce, Barbara Robson, Ralf Seppelt, Alexey A. Voinov, Brian D. Fath, and Vazken Andreassian, "Characterising Performance of Environmental Models," *Environmental Modelling & Software*, vol. 40, no. 0, pp. 1 – 20, 2013.

[85] Emanuele Principi, Stefano Squartini, Roberto Bonfigli, Giacomo Ferroni, and Francesco Piazza, "An Integrated System for Voice Command Recognition and Emergency Detection Based on Audio Signals," *Expert Systems with Applications*, vol. 42, no. 13, pp. 5668–5683, 2015.

[86] Ludmila I. Kuncheva, "Change Detection in Streaming Multivariate Data Using Likelihood Detectors," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 25, no. 5, pp. 1175–1180, May 2013.

[87] Sergios Theodoridis and Konstantinos Koutroumbas, *Pattern Recognition*, Academic Press, Burlington, U.S.A., 4th edition, 2008.

[88] Tracy C. Britton, Rodney A. Stewart, and Kelvin R. O'Halloran, "Smart Metering: Enabler for Rapid and Effective Post Meter Leakage Identification and Water Loss Management," *Journal of Cleaner Production*, vol. 54, no. 0, pp. 166–176, 2013.

[89] Lewis A. Rossman, *EPANET 2: USERS MANUAL*, Water Supply and Water Resources Division, National Risk Management Research Laboratory, Cincinnati, OH 45268, Sep. 2000.

[90] Ronan Collobert, Samy Bengio, and Johnny Mariéthoz, "Torch: a modular machine learning software library," Idiap-RR Idiap-RR-46-2002, IDIAP, 0 2002.

[91] S. Ntalampiras, I. Potamitis, and N. Fakotakis, "Probabilistic Novelty Detection for Acoustic Surveillance Under Real-World Conditions," *Multimedia, IEEE Transactions on*, vol. 13, no. 4, pp. 713–719, Aug 2011.

[92] Tom Fawcett, "An Introduction to ROC Analysis," *Pattern Recogn. Lett.*, vol. 27, no. 8, pp. 861–874, June 2006.

[93] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge, 1996.

[94] Elizabeth L. Ratnam, Steven R. Weller, and Christopher M. Kellett, "An optimization-based approach to scheduling residential battery storage with solar pv: Assessing customer benefit," *Renewable Energy*, vol. 75, pp. 123 – 134, 2015.

[95] Duško M. Tovilović and Nikola LJ. Rajaković, "The simultaneous impact of photovoltaic systems and plug-in electric vehicles on the daily load and voltage profiles and the harmonic voltage distortions in urban distribution systems," *Renewable Energy*, vol. 76, pp. 454 – 464, 2015.

[96] C. Bergaentzlé, C. Clastres, and H. Khalfallah, "Demand-side management and european environmental and energy goals: An optimal complementary approach," *Energy Policy*, vol. 67, pp. 858–869, 2014.

[97] Zheng Hu, Jin ho Kim, Jianhui Wang, and John Byrne, "Review of dynamic pricing programs in the u.s. and europe: Status quo and policy recommendations," *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 743 – 751, 2015.

[98] J. Upton, M. Murphy, L. Shalloo, P.W.G. Groot Koerkamp, and I.J.M. De Boer, "Assessing the impact of changes in the electricity price structure on dairy farm energy costs," *Applied Energy*, vol. 137, pp. 1–8, 2015.

[99] M. Yalcintas, W.T. Hagen, and A. Kaya, "An analysis of load reduction and load shifting techniques in commercial and industrial buildings under dynamic electricity pricing schedules," *Energy and Buildings*, vol. 88, pp. 15–24, 2015.

[100] Naïm R. Darghouth, Galen Barbose, and Ryan H. Wiser, "Customer-economics of residential photovoltaic systems (part 1): The impact of high renewable energy penetrations on electricity bill savings with net metering," *Energy Policy*, vol. 67, pp. 290 – 300, 2014.

[101] A. Parisio, E. Rikos, and L. Glielmo, "A model predictive control approach to microgrid operation optimization," *Control Systems Technology, IEEE Transactions on*, vol. 22, no. 5, pp. 1813–1827, Sept 2014.

[102] Marco Severini, Stefano Squartini, and Francesco Piazza, "Hybrid soft computing algorithmic framework for smart home energy management," *Soft Computing*, vol. 17, no. 11, pp. 1983–2005, 2013.

[103] M. Severini, S. Squartini, and F. Piazza, "Computational framework based on task and resource scheduling for micro grid design," in *Proceedings of the International Joint Conference on Neural Networks*, 2014, pp. 1695–1702.

[104] Guang-Bin Huang, "An insight into extreme learning machines: Random neurons, random features and kernels," *Cognitive Computation*, vol. 6, no. 3, pp. 376–390, 2014.

[105] NitinAnand Shrivastava, BijayaKetan Panigrahi, and Meng-Hiot Lim, "Electricity price classification using extreme learning machines," *Neural Computing and Applications*, pp. 1–10, 2014.

[106] Rafa? Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030 – 1081, 2014.

[107] Francesco Lisi and Fany Nan, "Component estimation for electricity prices: Procedures and comparisons," *Energy Economics*, vol. 44, pp. 143 – 159, 2014.

[108] Gao Huang, Guang-Bin Huang, Shiji Song, and Keyou You, "Trends in extreme learning machines: A review," *Neural Networks*, vol. 61, pp. 32 – 48, 2015.

[109] F. Ziel, R. Steinert, and S. Husmann, "Efficient modeling and forecasting of electricity spot prices," *Energy Economics*, vol. 47, pp. 98–111, 2015.

[110] V. Kekatos, Y. Zhang, and G.B. Giannakis, "Electricity market forecasting via low-rank multi-kernel learning," *IEEE Journal on Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1182–1193, 2014.

[111] Maurizio Cellura, Giuseppina Ciulla, Valerio Lo Brano, Aldo Orioli, Lucia Campanella, Francesco Guarino, and Davide Nardi Cesarini, "The redesign of an italian building to reach net zero energy performances: A case study of the shc task 40-ecbcs annex 52.," *ASHRAE transactions*, vol. 117, no. 2, 2011.

[112] Jonathan Currie and David I. Wilson, "OPTI: Lowering the Barrier Between Open Source Optimizers and the Industrial MATLAB User," in *Foundations of Computer-Aided Process Operations*, Nick Sahinidis and Jose Pinto, Eds., Savannah, Georgia, USA, 8–11 January 2012.

[113] Steven R Shaw, Steven B Leeb, Leslie K Norford, and Robert W Cox, "Nonintrusive load monitoring and diagnostics in power systems," *Instrumentation and Measurement, IEEE Transactions on*, vol. 57, no. 7, pp. 1445–1454, 2008.

[114] Jon Froehlich, Eric Larson, Sidhant Gupta, Gabe Cohn, Matthew Reynolds, and Shwetak Patel, "Disaggregated end-use energy sensing for the smart grid," *IEEE Pervasive Computing*, vol. 10, no. 1, pp. 28–39, 2011.

[115] Hesham K Alfares and Mohammad Nazeeruddin, "Electric load forecasting: literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, no. 1, pp. 23–34, 2002.

[116] Corinna Fischer, "Feedback on household electricity consumption: a tool for saving energy?," *Energy efficiency*, vol. 1, no. 1, pp. 79–104, 2008.

[117] B Neenan, J Robinson, and RN Boisvert, "Residential electricity use feedback: A research synthesis and economic framework," *Electric Power Research Institute*, 2009.

[118] George William Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.

[119] J. Zico Kolter and Tommi Jaakkola, "Approximate inference in additive factorial HMMs with application to energy disaggregation.," in *AISTATS*, Neil D. Lawrence and Mark Girolami, Eds. 2012, vol. 22 of *JMLR Proceedings*, pp. 1472–1482, JMLR.org.

[120] Matthew J. Johnson and Alan S. Willsky, "Bayesian nonparametric hidden semi-markov models," *J. Mach. Learn. Res.*, vol. 14, no. 1, pp. 673–701, Feb. 2013.

[121] Hyungsul Kim, Manish Marwah, Martin F. Arlitt, Geoff Lyon, and Jiawei Han, "Unsupervised disaggregation of low frequency power measurements.," in *SDM*. 2011, pp. 747–758, SIAM / Omnipress.

[122] Sundeep Pattem, "Unsupervised disaggregation for non-intrusive load monitoring," in *Machine Learning and Applications (ICMLA), 2012 11th International Conference on*. IEEE, 2012, vol. 2, pp. 515–520.

[123] Oliver Parson, Siddhartha Ghosh, Mark Weal, and Alex Rogers, "An unsupervised training method for non-intrusive appliance load monitoring," *Artificial Intelligence*, , no. 217, pp. 1–19, August 2014.

[124] Marisa B. Figueiredo, Bernardete Ribeiro, and Ana de Almeida, "On the optimization of appliance loads inferred by probabilistic models," in *Proceedings of the 2nd International Workshop on Non-Intrusive Load Monitoring*, 2014, `http://nilmworkshop.org/`.

[125] Liao Jing, Georgia Elafoudi, Lina Stankovic, and Vladimir Stankovic, "Power disaggregation for low-sampling rate data," in *Proceedings of the 2nd International Workshop on Non-Intrusive Load Monitoring*, 2014, `http://nilmworkshop.org/`.

[126] Karim Said Barsim, Roman Streubel, and Bin Yang, "An approach for unsupervised non-intrusive load monitoring of residential appliances," in *Proceedings of the 2nd International Workshop on Non-Intrusive Load Monitoring*, 2014, `http://nilmworkshop.org/`.

[127] Li Liu, Jiaman Ding, Jiayu Zhong, Xiaodong Fu, and Yang LV, "An unsupervised model for classification and recognition of household appliances," in *Journal of Computational Information Systems*, vol. 10, pp. 403–410. Binary Information Press, 2014.

[128] M. Figueiredo, B. Ribeiro, and A. de Almeida, "Electrical signal source separation via nonnegative tensor factorization using on site measurements in a smart home," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 2, pp. 364–373, Feb 2014.

[129] Matt Wytock and J. Zico Kolter, "Contextually supervised source separation with application to energy disaggregation.," *CoRR*, vol. abs/1312.5023, 2013.

[130] Hugo Goncalves, Adrian Ocneanu, and Mario Berges, "Unsupervised disaggregation of appliances using aggregated consumption data," in *1st KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, Aug. 2011.

[131] Jack Kelly and William Knottenbelt, "Neural NILM: Deep neural networks applied to energy disaggregation," in *Proceedings of the 2Nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, New York, NY, USA, 2015, BuildSys '15, pp. 55–64, ACM.

[132] Dominik Egarter, Manfred Pöchacker, and Wilfried Elmenreich, "Complexity of power draws for load disaggregation," *CoRR*, 2015.

[133] Claudia Beleites, Reiner Salzer, and Valter Sergo, "Validation of soft classification models using partial class memberships: An extended concept of sensitivity and co. applied to grading of astrocytoma tissues," *Chemometrics and Intelligent Laboratory Systems*, vol. 122, pp. 12 – 22, 2013.

[134] Kunihiko Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193–202, 1980.

[135] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 411–426, March 2007.

[136] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov 1998.

[137] Jack Kelly and William J. Knottenbelt, "UK-DALE: A dataset recording UK domestic appliance-level electricity demand and whole-house demand," *CoRR*, vol. abs/1404.0284, 2014.

[138] Nipun Batra, Jack Kelly, Oliver Parson, Haimonti Dutta, William Knottenbelt, Alex Rogers, Amarjeet Singh, and Mani Srivastava, "NILMTK: An open source toolkit for non-intrusive load monitoring," in *Proceedings of the 5th international conference on Future energy systems*. ACM, 2014, pp. 265–276.

[139] S. Spinsante, M. Pizzichini, M. Mencarelli, S. Squartini, and E. Gambi, "Evaluation of the Wireless M-Bus Standard for Future Smart Water Grids," in *Wireless Communications and Mobile Computing Conference, 9th International*, 2013, pp. 1382–1387.

[140] S. Squartini, L. Gabrielli, M. Mencarelli, M. Pizzichini, S. Spinsante, and F. Piazza, "Wireless M-Bus Sensor Nodes in Smart Water Grids: The Energy Issue," in *Intelligent Control and Information Processing, Fourth International Conference on*, 2013, pp. 614 – 619.