1 2 9 0

UNIVERSIDADE D
COIMBRA

Gustavo Valeriano Neves Luizon

Real time Energy Price Forecasting System in 5G Smart Grids Scenarios

Dissertation Proposal in the context of the Master in Data Science and Engineering, advised by Professor Bruno Sousa and presented to the Department of Informatics Engineering of the Faculty of Sciences and Technology of the University of Coimbra.

September 2023

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE D
COIMBRA

DEPARTMENT OF INFORMATICS ENGINEERING

Gustavo Valeriano Neves Luizon

Real time Energy Price Forecasting System in 5G Smart Grids Scenarios

Dissertation Proposal in the context of the Master in Data Science and Engineering, advised by Prof. Bruno Sousa and presented to the Department of Informatics Engineering of the Faculty of Sciences and Technology of the University of Coimbra.

September 2023

Acknowledgements

First, I would like to express my gratitude to my advisors Professor Bruno Sousa, Professor Tiago Cruz and Professor Paulo Simões who were always available for guidance during the development of this work.

Additionally, I would like to acknowledge support by the project Smart5Grid (co-funded by FEDER - Competitiveness and Internationalisation Operational Program (COMPETE 2020), Portugal 2020 framework).

I would also like to express my gratitude to Professor Cesar Teixeira. His important contribution during the data analysis and time series forecasting course inspired me to choose the theme of this work.

Especially, I would like to thank my family and friends for all the received support during the time in Portugal in search of my professional development.

Abstract

Smart grids are modernized electrical grids that use advanced technology to improve the efficiency and reliability for power supply. The capacity of Smart grids can be maximized with the use of 5G networks to support communication and data transfer between devices connected to the grid that allows real-time monitoring and control of the grid.

The integration of 5G networks in smart grids can also enable the use of advanced pricing and market mechanisms for energy. The time-of-use pricing can be used to incentive customers to use less energy during peak demand periods, while dynamic pricing can be used to respond to changes in supply and demand in real-time. This can help to reduce the need for expensive peak generation capacity and improve the overall efficiency of the energy market.

In future scenarios with intelligent and flexible energy networks, energy supply systems will demand a high degree of automation to ensure resilience, reliability and efficiency, in this context, real-time energy price information will be an important input for process optimizations. The present work evaluates the performance of real-time energy prices calculation and forecasting system in 5G high-speed communication networks.

Finally, the obtained results present good potential to achieve accurate real time forecasts of energy prices in Portugal in the short term through the use of statistical and neural networks based models. Complementarily, a 5G communication networks simulation system is presented to evaluate the performance of 5G networks for energy pricing updates.

Keywords

Smart Grids, 5G, Data Science, Energy Market, Energy Price, Forecast, SIMU5G, Real Time Emulation, Time Series, SARIMA, SARIMAX, LSTM, GRU.

Contents

Acronyms

5GC 5G Core Network.

ACF Autocorrelation Function.

ACS Autocorrelation Function/Sequence.

ADF Augmented Dickey–Fuller.

AMF Mobility Management Function.

AMI Advanced Metering Infrastructure.

AR Autoregressive.

ARIMA Auto Regressive Integrated Moving Average.

ARMA Auto Regressive Moving Average.

BGP Border Gateway Protocol.

DF Dickey-Fuller.

Dl Downlink.

DLMS Device Language Message Specification.

ERSE Entidade Reguladora dos Serviços Energéticos.

ETSI European Telecommunications Standards Institute.

FCS Frame check sequence.

FDD Frequency Division Duplexing.

FT Fourier Transform.

GNB GNodeB.

GPRS General Packet Radio Service.

GRU Gated Recurrent Unit.

GTP GPRS Tunnelling Protocol.

HCS Header Check Sequence.

HDLC High-level Data Link Control.

IEC International Electrotechnical Commission.

IEEE Institute of Electrical and Electronics Engineers.

INI Initialization.

IP Internet Protocol.

IPv4 Internet Protocol version 4.

IPv6 Internet Protocol version 6.

KDE Kernel Density Estimate.

LDP Label Distribution Protocol.

LOESS Locally Weighted Scatter Plot Smooth.

LOWESS Locally Weighted Scatter Plot Smooth.

LSTM Long Short-Term Memory.

LTE Long Term Evolution.

MA Moving Average.

MAE Mean Absolute Error.

MANET Mobile Ad hoc Network.

MAPE Mean Absolute Percentage Error.

MEC Multi-Acess Edge Computing.

MIBEL Iberian Electricity Market.

MPLS Multiprotocol Label Switching.

MSDU MAC Service Data Unit.

NG-RAN Radio Access Network.

NIC Network Interface Card.

NR New Radio.

OMIE Electricity Market Operator.

OSPF Open Shortest Path First.

PACS Partial Autocorrelation Function/Sequence.

PDU Protocol Data Units.

PPP Point-to-Point Protocol (PPP).

REN Redes Energéticas Nacionais.

RMSE Root Mean Squared Error.

RNN Recurrent Neural Networks.

RSVP-TE Resource Reservation Protocol - Traffic Engineering.

SARIMA Seasonal Autoregressive Integrated Moving Average.

SARIMAX Seasonal Autoregressive Integrated Moving-Average with Exogenous Variables.

SDAC Single Day-Ahead Coupling.

TCP Transmission Control Protocol.

TDD Time Division Duplexing.

TS Time Series.

UDP User Datagram Protocol.

UE User Equipment.

Ul Uplink.

UPF User Plane Function.

USIM Universal Subscriber Identity Module.

List of Figures

List of Tables

Chapter 1

Introduction

The Smart grids are modernized power systems that use advanced technology to monitor and control the flow of electricity, allowing more efficient and reliable energy distribution. The 5G networks, on the other hand, are the next generation of mobile networks that offer faster speeds, lower latency, increased capacity and security.

The integration of smart grid technology with 5G networks is expected to bring about significant transformation in the energy market, it enables real-time monitoring and control of energy distribution, allowing for more efficient use of resources and the ability to quickly respond to changes in demand.

This thesis aims to investigate the specific aspects of the energy/electricity market focusing on the price of energy and exploring the transformations that will come with the technological advances of smart networks in 5G.

Deeper transformations in the electric energy market sector are slow due to its characteristics, it is a market that requires a large physical structural volume for production and availability.

Despite the greater resistance to changes in the energy/electricity market, 5G Smart Grids enable the improvement of electrical systems for clean and friendly power generation, combination of multiple distributed clean power sources, integration of energy storage and electric vehicles, as well as balance of decentralization, reliability and load, with clean and low-carbon, grid-power source coordination, flexible and efficient features; safe and efficient power transmission/transformation, characterized by situation awareness, flexibility, reliability, and coordinated optimization; flexible and reliable power distribution, featuring controllable, compatible and economical nature; diverse and interactive power consumption, with the characteristics of diversity, two-way interaction, flexibility, energy saving and efficiency [2].

This can further drive the transformation of the energy market by allowing the fast integration and control of intermittent energy sources and increasing the participation of consumers as will be explored throughout this work.

1.1 Objectives

The main objective of this work is to obtain calculation and forecast of energy prices in real time exploring the potential of Smart Grids in combination with 5G communication networks. Services like this enables consumers to react to price fluctuations allowing better coupling of consumption according to the most favourable periods for generation. As a consequence, the entire operating system of the electrical networks would be favored through the immediate responsiveness of the energy supply.

To achieve this objective, mathematical models capable of making energy price predictions will be defined and operate in 5G Smart Grids simulation scenarios in order to attest the effective capacity of this system to collect, transmit and process all the data necessary to update the electricity price in real-time as well as providing the price forecast for the next 24 hours.

1.2 Contributions

This document makes contributions as follows:

1. The proposal of forecast models for the price of energy using statistical and machine learning-based methods that use time series of previously realized prices and real-time energy generation values of the various technologies present in the electrical system.

2. A specific analysis of the Portuguese energy system with a focus on the appropriate adjustment of forecasting models.

3. The proposal for a distributed energy generation data collection in 5G and processing system to enable model training and price forecast execution with subsequent sending of the obtained forecast values to end users.

4. Propose the topic for submission of scientific publications.

1.3 Structure

The remaining document is organized as follows:

- Chapter 2 - Background

  Provides an overview of the conceptual topics covered by this work such as Smart Grids, 5G Communication Networks, Time-Series analysis and forecast, the electricity sector and the energy/electricity market in Portugal.

- Chapter 3 - Related Work

  Reviews the related work regarding energy price forecast and 5G Networks simulation/emulation.

- Chapter 4 - Research Objectives and Approach

  Describes the work's research objectives and introduces the approach taken throughout the dissertation work, as well as the methodology used.

- Chapter 5 - System Architecture

  Presents the details of the architecture for the system implemented for the prototype simulation of a real time energy price forecasting system.

- Chapter 5 - Energy Price Analysis

  Presents an exploratory analysis of the composition of the energy matrix in Portugal and the necessary data analyzes for the energy price forecast models fitting.

- Chapter 6 - Results

  Provides the summary and discussion of experimental obtained results.

- Chapter 7 - Conclusion

  Provides the main conclusions, a summary of the results, and the future work that can be followed.

Chapter 2

Background

This chapter covers the conceptual topics necessary for the development of this work. The first section introduces the theoretical basement for the price prediction models explored, namely, SARIMA, SARIMAX, LSTM and GRU. Next, a contextualization of smart grids and the technical basis necessary for the development of the proposed work plan are provided.

Also present in this chapter is a contextual introduction to the electricity sector in Portugal, including current scenario of the electricity sector centred on a broad view of energy generation, transmission and distribution systems, as well as regulatory and economic aspects.

The objective is to provide the reader a perception of how the sector works, the main challenges associated with the energy transition, and a background basis for a better understand the presented work.

2.1 Energy Price Forecasting Models

The energy price issue is widely discussed in the context of energy systems, after all it is of broad interest, it affects the viability of investments for energy generation, operation of energy transmission and distribution systems and the final consumer.

The theme of forecasting energy prices with a focus on investment in production is recurrent in the available literature. Financial analyzes for investment decisions involve the need for accurate forecasts of production costs and availability of energy supplies. See Chapter 3.

The objective of this work is to analyze the price of energy from the perspective of the final consumer and the intelligent operation of transmission and distribution networks.

In this context, the price of energy should act as an indicator of the level of energy efficiency. The need for support for the energy supply would be indicated through price increases as an immediate incentive for available generating sources.

The topic of optimal operation of smart grids is of great importance among the ongoing discussions on the evolution of energy transmission and distribution systems currently in operation, as can be seen in publications such as "Optimal Multi-Operation Energy Management in Smart Microgrids in the Presence of RESs Based on Multi-Objective Improved DE Algorithm: Cost-Emission Based Optimization"[5].

The Real-time price updates are fundamentally important for the optimal operation of energy transmission systems. The indicative of the price increase made available to consumers by very short time scales communication networks could enable smart home operating systems by injecting power into the electrical system, generating mutual benefit. For consumers, who would be protected from high energy costs, and for Network Operators, which will receive support from distributed energy generated at the most critical times.

The energy price and generation data are available at hourly granularity, so very short-term prediction models will be evaluated for one-hour-ahead prediction. However, the evolution of 5G smart grids equipped with systems such as the one proposed in this work could make it possible to update the price of energy on scales of milliseconds.

The price of electricity is time-varying, therefore, statistical techniques for time series analysis are suitable for analyzing and interpreting the data. It involves analyzing the patterns, trends and seasonal behaviour of the energy price and other variables that directly or indirectly interfere with its formation.

Understanding the behaviour of the time series involves identifying not only the relationships between the current and previous instants of the series but also understanding the relationships between the current and past instants of the price time series in relation to the current and past instants of external variables.

The patterns and trends identified in the data can be used to make predictions about future values and develop mathematical models that can describe the behaviour of the time series.

The time series analysis and forecasting techniques uses as theoretical foundation the books "Time Series Analysis and Forecasting" [18], "Dive into Deep Learning" [27] and "Time series analysis: forecasting and control" [17].

2.1.1 Features and Parameters Selection

Feature selection is an important process for improving the quality of the input data when developing predictive models. It reduces the number of input variables by selecting the features that better represent the data for reduced computational cost and performance improvement.

In time series analyses, feature selection is applicable for multivariate models, in addition, many models used for evaluation and prediction in time series are autoregressive, which requires statistical analysis, data descriptive techniques and transformations for the correct adjustment of parameters of the models as will be detailed throughout this section.

Seasonal Decomposition

The time series seasonal decomposition method consists of separating data into trend, seasonal and irregular/erratic variation components for a better understanding of data patterns. Once the components have been identified and separated, they can be analyzed separately

1. Trend
   The trend represents the overall long-term direction of the series, a systematic change in the mean or general direction of the time series. The trend can be linear or nonlinear and can usually be modelled using polynomial functions of different orders by global model-based approaches or by localized data filtering/smoothing approaches like moving average (MA) filtering or Locally Weighted Scatter Plot Smooth (LOWESS/LOESS).

2. Seasonality
   The seasonal component represents the regular and repeating pattern. The presence of the seasonal component can be verified by making a frequency analysis of the time series performed by computing the Fourier Transform (FT) and observing if important seasonal components will appear with relevant magnitude.

   Once the main frequency components are identified, low pass frequency filters can be used to explicitly remove them.

3. Irregular/Erratic Component
   The irregular variation component represents the random fluctuation of the time series, it can be obtained by removing the trend and seasonality from the time series.

The removal of trend and seasonality components are important to obtain the stationarity of the time series. The popular methods for time series analysis such as Auto Regressive Moving Average (ARMA) require by definition that the series be stationary as described in detail in Section 2.1.2.

Besides the explicit way, popular approaches as the Box-Jenkins use to attain stationarity implicitly, by differencing.

Autocorrelation Function

The Autocorrelation Function/Sequence (ACF/ACS) is one of the main existing techniques to study the correlation between observations of a time series with their past values. It is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies.

The Autocorrelation Function is defined as the value of the autocorrelation coefficient ρ_k as a function of the lag k, the autocorrelation coefficient ρ_k is described as:

$$\rho_k = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sqrt{E[(z_t - \mu)]^2 E[(z_{t+k} - \mu)]^2}} = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sigma_z^2}$$

Under the stationarity assumption, the covariance between values z_t and z_{t+k}, separated by k intervals of time, or by lag k, must be the same for all t, it is called autocovariance $\gamma_k = E[(z_t - \mu)(z_{t+k} - \mu)]$ and the variance $\sigma_z^2 = \gamma_0$ is the same at time t + k as at time t. Thus, the autocorrelation at lag k, that is, the correlation between z_t and z_{t+k}, is:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

In particular, $\rho_0 = 1$. Figure 2.1 illustrates the autocorrelation matrix diagonals represented by the autocorrelation function.

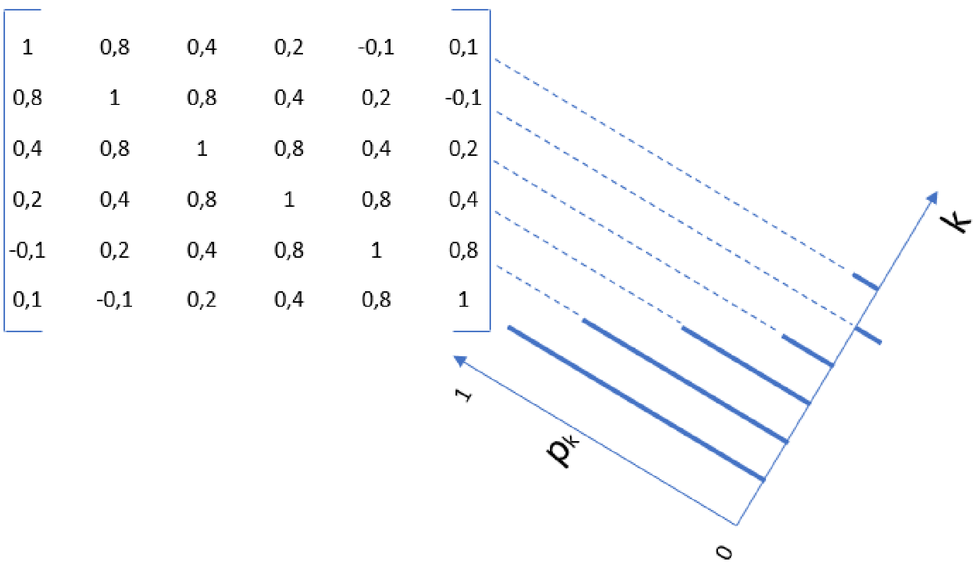


Figure 2.1: Autocorrelation matrix and corresponding autocorrelation function

The Partial Autocorrelation Function/Sequence (PACS) gives the partial correlation of a stationary time series with its own lagged values, regressed the values of the time series at all shorter lags. In data analysis, it aimed at identifying the extent of the lag in an Autoregressive (AR) model as part of the Box–Jenkins approach.

Cross Correlation Function

The Cross Correlation Function measures the similarity between two different time series overlapped in time. While the Correlation Function measures how the value of time t of the time series is related to the value t+k of the same time series, the objective of the Cross Correlation Function is to measure the similarity between the value of time t of one time series to the value t+k of another time series.

The autocovariance coefficients of each of the two series at lag k are defined by the usual formula:

$$\gamma_{xx}(k) = E[(x_t - \mu_x)(x_{t+k} - \mu_x)]$$

$$\gamma_{yy}(k) = E[(y_t - \mu_y)(y_{t+k} - \mu_y)]$$

While the cross-covariance coefficients between x_t and y_t series at lag +k is defined by:

$$\gamma_{xy}(k) = E[(x_t - \mu_x)(y_{t+k} - \mu_y)]$$

Similarly to the auto-correlation coefficient, the cross-correlation coefficient at lag k is provided by:

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y}$$

The Cross correlation function corresponds to the Cross correlation coefficients defined for $k = 0, \pm 1, \pm 2, ...$ In contrast to the autocorrelation function, the $\rho_{xy}(k)$ coefficient is not equal to $\rho_{xy}(-k)$, so the cross-correlation function is not symmetric about k = 0.

2.1.2 SARIMAX Model

The Seasonal Autoregressive Integrated Moving Average with Exogenous Variables, SARIMAX, is a statistical time series forecasting model capable of handling trend and seasonality components of the time series incorporating external variables that may have an influence on it.

The SARIMA and SARIMAX models are extensions of ARIMA models that arise through the combination of simpler models, namely: Autoregressive (AR), Integrated(I) and Moving Average (MA) models as described in the sequence.

- Moving Average (MA)

  The Moving Average is a linear model for time series, it is used to calculate the current value assuming it's linearly dependent on the current and past error terms. Moving Average models are based on the assumption that the underlying data-generating process is stationary.

  When the data is stationary, the models can effectively capture the short-term dependencies and provide accurate forecasts. If the time series data is non-stationary, it can result in unreliable and spurious results from MA models.

  A moving average model expresses a given random process, X(n), as:

  $$X(n) = \beta_0 Z(n) + \beta_1 Z(n-1) + ... + \beta_q Z(n-q)$$

  Defining the backward shift operator B as $B^j Z(n) = Z(n-j)$ we have:

  $$\theta(B) = \beta_0 + \beta_1 B + ... + \beta_q B^q$$

  Then,

  $$X(n) = \theta(B) Z(n)$$

- Auto Regressive (AR)

  An autoregressive (AR) process expresses actual values as a function of past process values and a random perturbation. The model is regressed on its own past values. Thus,

  $$X(n) = \alpha_1 X(n-1) + \alpha_2 X(n-2) + ... + \alpha_p X(n-p) + Z(n)$$

  Considering the backward shift operator B as $B^j Z(n) = Z(n-j)$ we have:

  $$\phi(B) = 1 - \alpha_1 B - \alpha_2 B^2 - ... - \alpha_p B^p$$

  Then,

  $$\phi(B) X(n) = Z(n)$$

- Auto Regressive Moving Average (ARMA)

  The ARMA models are composed by the junction of Auto Regressive and Moving Average models making it possible, in many cases, to achieve more flexibility in fitting the model to the data when compared of MA or AR alone.

  $$X(n) = \alpha_1 X(n-1) + \alpha_2 X(n-2) + ... + \alpha_p X(n-p) + \beta_0 Z(n) + \beta_1 Z(n-1) + ... + \beta_q Z(n-q)$$

Considering the backward shift operator B as $B^j Z(n) = Z(n-j)$ we have:

$$\phi(B) = 1 - \alpha_1 B - \alpha_2 B^2 - ... - \alpha_p B^p$$
$$\theta(B) = \beta_0 + \beta_1 B + \beta_2 B^2 + ... + \beta_q B^q$$

Then,

$$\phi(B)X(n) = \theta(B)Z(n)$$

- Auto Regressive Integrated Moving Average (ARIMA)

  The AR, MA and ARMA models require that the time series be stationary as previously described, one of the most effective ways to obtain the stationarity of a time series is through differentiation.

  The ARIMA models are capable of describing non-stationary processes as they obtain the stationarity of processes by differentiation and then use ARMA models to describe the process. To obtain the real time series values on the output the ARMA model must be integrated. Due to this integration step, the name Integrated(I) is added to the model.

  In ARIMA models, the mathematical representation uses the differential operator described by: $W(n) = \bigtriangledown^d X(n) = (1-B)^d X(n), d = 1, 2, ...$

$$W(n) = \alpha_1 W(n-1) + ... + \alpha_p W(n-p) + \beta_0 Z(n) + \beta_1 Z(n-1) + ... + \beta_q Z(n-q)$$

  Considering the backward shift operator B as $B^j Z(n) = Z(n-j)$ we have:

$$\phi(B) = 1 - \alpha_1 B - \alpha_2 B^2 - ... - \alpha_p B^p$$
$$\theta(B) = \beta_0 + \beta_1 B + \beta_2 B^2 + ... + \beta_q B^q$$

  Then,

$$\phi(B)W(n) = \theta(B)Z(n)$$

- Seasonal Auto Regressive Integrated Moving Average (SARIMA)

  The ARIMA model is not able to consider the seasonal component of the time series and considering that seasonality is a very common property, the SARIMA model appears as a variation of the ARIMA model capable of representing it. SARIMA model is composed of two ARIMA models, one that deals with short-term dependencies and the other that deals with seasonal dependencies.

  Considering the backward shift operator B as $B^j Z(n) = Z(n-j)$ we have:

$$\phi_p(B) = 1 - \alpha_1 B - \alpha_2 B^2 - ... - \alpha_p B^p$$
$$\Phi_P(B^S) = 1 - a_1 B^S - a_2 B^{S2} - ... - a_P B^{SP}$$
$$\theta_q(B) = \beta_0 + \beta_1 B + \beta_2 B^2 + ... + \beta_q B^q$$
$$\Theta_Q(B^S) = b_0 + b_1 B^S + b_2 B^{S2} + ... + b_Q B^{SQ}$$

$\bigtriangledown^d$: d order simple differencing as a trend removal mechanism

$\bigtriangledown_S^D$: D order seasonal differencing, removing seasonal patterns

Then,

$$\phi_p(B)\Phi_P(B^S) \bigtriangledown^d \bigtriangledown_S^D X(n) = \theta_q(B)\Theta_Q(B^S)Z(n)$$

Where ϕ_p defines the short-term AR parameters, Φ_P defines the seasonal AR parameters. Similarly, θ_q defines the short-term MA parameters, and Θ_Q defines the seasonal MA parameters.

Therefore, the correct fit of SARIMA models requires the following parameters to be specified:

$$SARIMA(p,d,q)(P,D,Q)s$$

- p: Autoregressive order (AR order)
- d: Degree of differencing (integration order)
- q: Moving average order (MA order)
- P: Seasonal autoregressive order (SAR order)
- D: Seasonal degree of differencing (seasonal integration order)
- Q: Seasonal moving average order (SMA order)
- s: Seasonal period (number of time steps in one season)

Finally, after obtaining the structure of the SARIMA model, it becomes possible to include exogenous variables in the model, which results in the SARIMAX model as defined below.

Considering $X_1(n), X_2(n), ..., X_M(n)$ as the exogenous variables and the backward shift operator B as $B^j Z(n) = Z(n-j)$ we have:

$$\gamma_i(B) = \xi_{i0} + \xi_{i1} B + \xi_{i2} B^2 + ... + \xi_{ip} B^p$$

$$\gamma_1(B)X_1(n) + \gamma_2(B)X_2(n) + ... + \gamma_M(B)X_M(n) = \sum_{i=1}^{M} \xi_i X_i(n)$$

And the SARIMAX model as an extension of SARIMA model can be expressed by:

$$\phi_p(B)\Phi_P(B^S) \bigtriangledown^d \bigtriangledown_S^D X(n) = \sum_{i=1}^{M} \xi_i X_i(n) + \theta_q(B)\Theta_Q(B^S)Z(n)$$

2.1.3 Neural Networks

Neural networks consist of interconnected artificial neurons, called nodes or units, organized in layers. Each node receives input signals, applies a mathematical operation to them, and produces an output signal. The connections between nodes are associated with weights that determine the strength of the input signal.

The performance of time series forecasting models based on neural network techniques can be extensively explored among the many existing possibilities, for this work, recognized efficient methods like Long Short-Term Memory Networks (LSTMs), Gated Recurrent Units (GRU) will be evaluated to obtain the desired results.

Recurrent Neural Networks - RNN

The RNNs process sequential data by incorporating feedback connections, each node can maintain a state that captures information from previous time steps giving the neural network the capacity to retain memory.

In a time-series data set, the values on successive time-stamps are closely related to one another. If the values of these time-stamps are used as independent features the key information about the relationships among the values is lost. For this reason, the capacity to consider dependencies among the attributes makes the RNNs well-suited for time-series data.

- Long Short-Term Memory - LSTM

  Long Short-Term Memory (LSTMs) are a type of recurrent neural network (RNN) architecture that are able to selectively remember and forget information from previous time steps, making them particularly effective for modelling long-term dependencies in time series data. See LSTM network structure in Figure 2.2

  The main component of an LSTM is its memory cell, it is responsible for storing and updating the memory state. The cell interacts with different gates that control the flow of information, including the input gate, forget gate, and output gate.

  The memory cell is equipped with an internal state and a number of multiplicative gates as described:

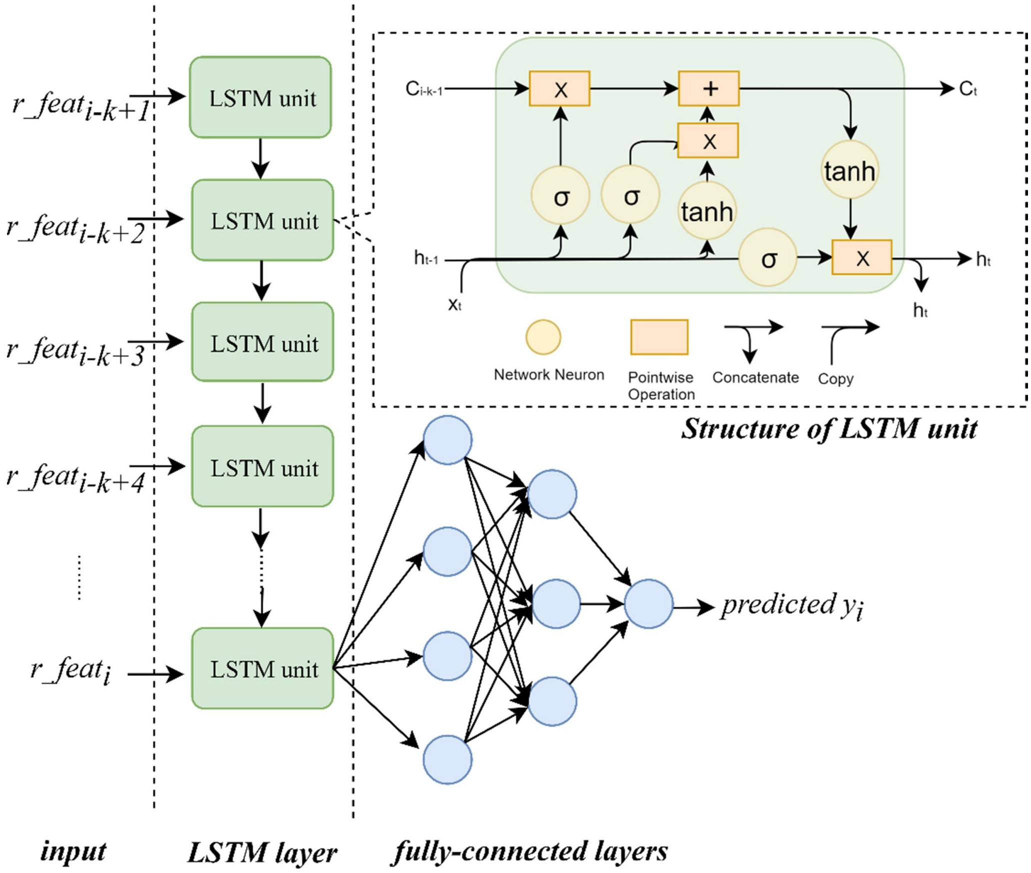


Figure 2.2: LSTM neural network structure [24]

1. Input gate determines whether a given input should impact the internal state;
2. Forget gate determine whether the internal state should be changed to 0;
3. Output gates determine whether the internal state of a given neuron should be allowed to impact the cell's output.

The LSTM gate receives the current time step and the hidden state of the previous time step as inputs. Three fully connected layers with sigmoid activation functions compute the values of the input, forget, and output gates, as a consequence the gate's output is limited to the range (0,1).

In Figure 2.3 is possible to observe that the forget gate determines whether to keep the current value of the memory or flush it through a product operator. The input node is typically computed with a hyperbolic tangent activation function (tanh) and it determines how much of the input node's value should be added to the current memory cell internal state and the output gate determines whether the memory cell should influence the output at the current time step.

Mathematically, the expressions for gate output calculations are:

$$I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i)$$
$$F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f)$$

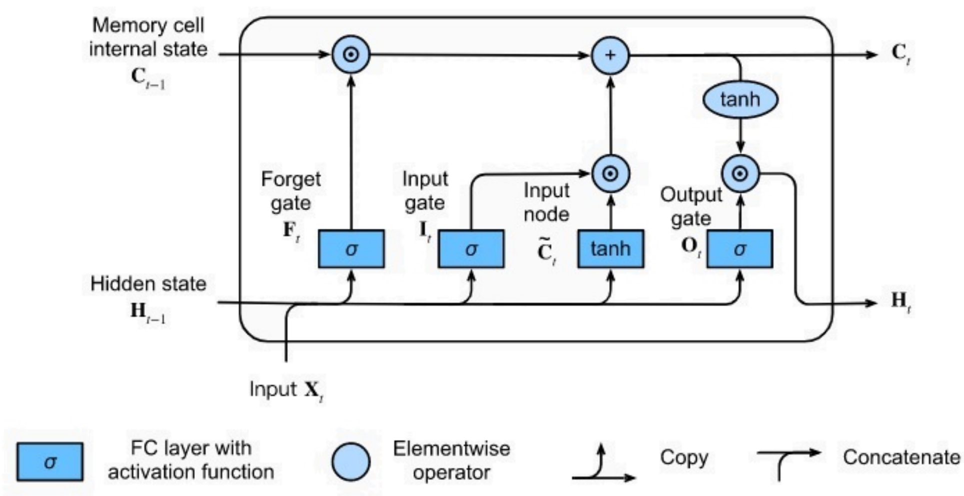


Figure 2.3: Structure of LSTM Unit [27]

$$O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o)$$

Where:
X_t is the input at step t
H_{t-1} is the hidden state at step t-1
I_t is the input gate at step t
F_t is the forget gate at step t
O_t is the output gate at step t
W are the weight parameters
b are the bias parameters

Similarly, the equation for the input node can be expressed as:

$$C'_t = tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c)$$

Consequently, the current memory cell state can be expressed as a function of the input and forget gates, the one step behind the memory cell state and the input node:

$$C_t = F_t \cdot C_{t-1} + I_t \cdot C'_t$$

Finally, the new hidden state is calculated by applying the product operator between the hyperbolic tangent (tanh) function of the memory cell's internal state and the output gate.

$$H_t = O_t \cdot tanh(C_t)$$

- Gated Recurrent Units - GRU

The Gated Recurrent Units (GRU) architecture is a variation of the LSTM, the LSTM's three gates are replaced by two sigmoid activations gates: the reset gate and the update gate. The LSTM and GRU methods achieve comparable performance but GRU has the advantage of being faster to compute.

Likewise, an update gate would allow us to control how much of the previous hidden state should be retained and how much of the new information should be added to the current hidden state. The reset gate decides how much of the previous hidden state should be forgotten.

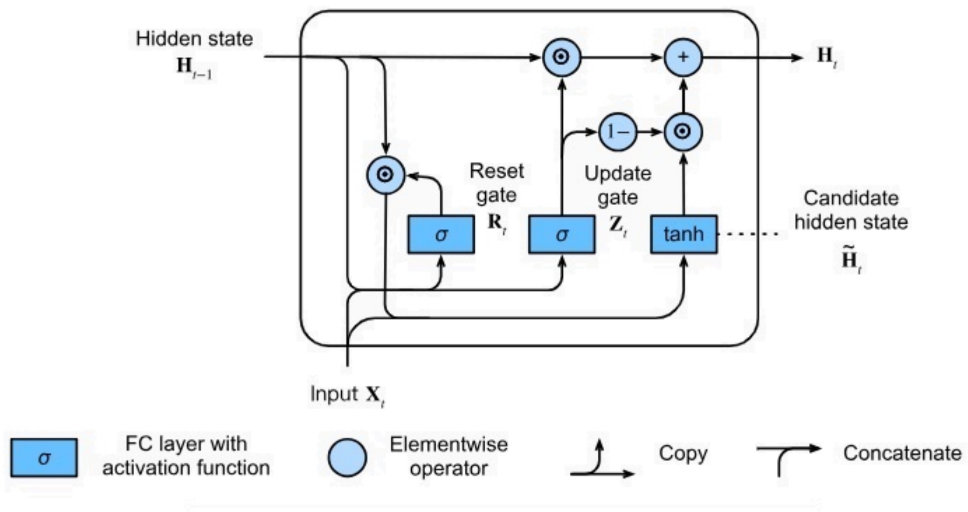


Figure 2.4: Structure of GRU Unit [27]

Mathematically, the expressions for gate output calculations are:

$$R_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r)$$
$$Z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z)$$

Where:
X_t is the input at step t
H_{t-1} is the hidden state at step t-1
R_t is the reset gate at step t
Z_t is the update gate at step t
W are the weight parameters
b are the bias parameters

The GRU's candidate Hidden State is computed like LSTM's input node:

$$H_t' = tanh(X_t W_{xh} + H_{t-1} W_{hh} + b_h)$$

And the final Hidden State update equation is described as:

$$H_t = Z_t \cdot H_{t-1} + (1 - Z_t) \cdot H'_t$$

2.1.4 Evaluation Metrics

In order to evaluate the performance of energy price forecasting methods, comparison metrics were used between the time series obtained through the prediction of the models with the time series with real prices realized in the same period.

The evaluation metrics used are described as follows, where Y represents the real energy price values and X represents the energy price prediction values:

- Mean Absolute Error (MAE)
  $$MAE = \frac{\sum_{i=1}^{N} |Y_i - X_i|}{N}$$
- Root Mean Squared Error (RMSE)
  $$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Y_i - X_i)^2}{N}}$$
- Mean Absolute Percentage Error (MAPE)
  $$MAPE = \frac{\sum_{i=1}^{N} |\frac{Y_i - X_i}{X_i}|}{N} * 100\%$$

2.2 Smart Grids

A smart grid incorporates digital technology, communication networks, and automation to enhance the efficiency, reliability, and sustainability of the electrical grid.

It represents a significant upgrade from traditional power grids, which are often characterized by one-way electricity flows and limited real-time data monitoring and control. Smart Grids enable bidirectional communication between various components of the grid, allowing for better coordination and management.

2.2.1 Smart Grids Challenges

Historically, electricity consumption was static and predictable, any unexpected change in its behaviour was reflected in a major challenge for operators of energy transmission and distribution systems.

With the entry of intermittent renewable sources in electrical systems, the complexity of forecasting demand increases, making static systems completely obsolete.

Considering the context presented, it is important to identify the main challenges to be faced by smart grids in the upcoming years.

1. **Growing amount of renewable energy sources**: The increasing amount of renewable energy sources in the system is especially challenging due to the intermittent nature of these sources since the volume of generation derives from climatic factors and therefore cannot be controlled or predicted with accuracy. The oscillation of generation from renewable sources introduces difficulty in controlling the voltage levels in the systems.

2. **Losses in the transmission of electricity**: It derives from the one-way characteristic of traditional transmission systems. All real systems will suffer losses, but for long transmission systems to serve consumers very far from the generation centres, electrical losses are significant and greatly reduce the efficiency of the transport process.

3. **Interruptions in power delivery** Interruptions in supply services are problems that get worse in systems with little mesh, where there are not many path options for the interconnection between nodes. Thus, any partial system unavailability that occurs along the entire radial circuit will affect blackouts from that point onwards.

4. **High Consumption Demand Operation**: Power transmission and distribution systems are known for their off-duty operation most of the time, this is because they are sized to meet times of high demand that generally only occur during a few hours of the day or atypical days and all the efforts of operators of systems are geared precisely to meet periods of high demand.

5. **Electromobility**: The electromobile transition will be responsible for a major change in the demand curve for electricity with the increasing replacement of vehicles powered by fossil fuels by electric ones, in addition to new possibilities for electric transport that begin to emerge with the advancement of batteries. In the context of public transport, electrical solutions are also taking up more and more space considering their efficiency and ecological acceptability.

6. **Network Modernization**: Energy networks modernization refers to the process of updating and upgrading the infrastructure that is used to transmit and distribute electricity.

   This can include replacing old equipment and technology, and implementing new control systems and automation.

   The goal of energy network modernization is to improve the efficiency, reliability, and sustainability of energy systems. It can also help to reduce greenhouse gas emissions and support the transition to a low-carbon economy. Examples of modernization initiatives include:

   - Smart grid technology, which allows for two-way communication between utilities and consumers.

- Advanced metering infrastructure (AMI), which uses smart meters to communicate real-time energy usage data to utilities, enabling them to better manage demand and improve grid operations.
- Distribution automation, which uses sensors, communication networks, and advanced control systems to improve the reliability and efficiency of the distribution grid.

7. **Threat of Cyber Attacks**: As systems become more automated and remotely manoeuvrable, they also become susceptible to cyber attacks and the consequences of this type of interference will be more devastating. In this context, cyber security is fundamental for the viability of intelligent networks.
8. **Threat of terrorist attacks**: Energy systems are fundamental to the entire operation of an entity, be it a country, state or even a company. Using this fragility, terrorist attacks will always consider them in their strategic action. It is important that the systems are less and less dependent on their regions in order to reduce the consequences of this type of vandalism.

The global energy scenario presents great challenges, besides the technical challenges to the immediate implementation of solutions, it also requires a large volume of investments, considering the necessary infrastructure for the entire existing process between energy production and consumption. [19]

Therefore, this thesis seeks to investigate possible solutions considering technical and economical aspects.

First, it is necessary to understand the economic context of the electricity market and the characteristics that influence the formation of the energy price that is made available to the final consumer. It is very important to bear in mind that the cost of generating and transmitting energy changes with time and depends on several factors, such as the availability of energy inputs, operation and maintenance of assets and the amount of energy demanded.

The periods of greatest energy demand are those that generate more difficulties and costs for the operation of transmission and distribution systems. Additionally, the power required to meet peak hours in many cases requires the use of more expensive energy sources to be available, so it seems plausible consumers receive financial incentives to allocate their consumption at times when the system is not overloaded.

To make this possible in a concrete way, it is necessary to carry out periodic measurements for each energy consumer in order to record the energy consumption at different times of the day, this means a huge volume of data to be transmitted and processed.

The evolution of technology for photovoltaic generation has reached the level of modular viability at the individual level: the costs of equipment are affordable and its efficiency is acceptable to be used by any individual and new system owners are already considering the possibility of commercialization

As mentioned earlier, the challenge of operating the distribution systems increases with the entry of distributed microgeneration, in addition, regulatory issues need to be reviewed to allow this type of operation. Actually, what exists for the individual consumer is the possibility of reducing the energy tariff through the injection of private generation. [22]

It's also important not to lose sight of the fact that the current energy market is controlled by a small group of large companies that would certainly offer resistance to this type of modification.

In this document, we will analyze the current energy market and investigate how smart grids with high performance in 5G communication and data processing can promote changes in current energy models.

2.2.2 Smart Grids Emulation

In order to achieve the objectives of this work and obtain a realistic prototype for simulating a system capable of providing real-time energy prices, it is essential that the 5G communication system is as close as possible to the real communication network. To meet this requirement, open-source software capable of simulating the communication between the smart meters and the Multi-Acess Edge Computing(MEC) servers will be used.

In this section, there is an introductory description of the Simu5G used to simulate the 5G network and the Gurux simulator, used to simulate the smart meters.

SIMU5G

Simu5G is a 5G New Radio (NR) developed by a research project carried out by Intel Corporation[12] and the Computer Networking Group of the University of Pisa, Italy and widely used by industry and academia.

Schematically, the 5G system uses a User Equipment (UE), itself composed of a Mobile Station and a Universal Subscriber Identity Module(USIM), the Radio Access Network (NG-RAN) and the Core Network (5GC).

The main entity of the NG-RAN is the GNB, where "g" stands for "5G" and "NB" for "Node B", which refers to the radio transmitter.

The 5GC is represented UPF entity: the User Plane Function (UPF), handling the user data and, in the signalling plane, the Access and Mobility Management Function (AMF) that accesses the UE and the NG-RAN.

Simu5G is based on the OMNeT++ simulation framework and provides a collection of models with well-defined interfaces, which can be instantiated and connected to build arbitrarily complex simulation scenarios.

For a better understanding of the full potential of the SIMU5G simulator, more detailed descriptions of the OMNet++ and INET frameworks are necessary. The information below are selected and transcribed parts from the original descriptions of the software and summarizes its main features:

- OMNet++ [11]

  The Discrete Event Simulator OMNeT++ is an extensible, modular, component-based C++ simulation library and framework, primarily for building network simulators.

  More specifically, OMNet++ includes wired and wireless communication network simulations, Domain-specific functionality such as support for sensor networks, wireless ad-hoc networks, Internet protocols, performance modeling, photonic networks, etc., is provided by model frameworks, developed as independent projects.

  OMNeT++ offers an Eclipse-based IDE, a graphical runtime environment, and a host of other tools. There are extensions for real-time simulation, network emulation, database integration, SystemC integration, and several other functions.

- INET [6]

  The INET Framework can be considered the standard protocol model library of OMNeT++. INET contains models for the Internet stack and many other protocols and components. Several other simulation frameworks take INET as a base and extend it into specific directions, such as vehicular networks (Veins, CoRE), overlay/peer-to-peer networks (OverSim), or LTE (SimuLTE).

  The INET Framework contains models for the Internet stack (TCP, UDP, IPv4, IPv6, OSPF, BGP, etc.), wired and wireless link layer protocols (Ethernet, Point-to-Point Protocol (PPP), Institute of Electrical and Electronics Engineers (IEEE) 802.11, etc), support for mobility, Mobile Ad hoc Network (MANET) protocols, DiffServ, Multiprotocol Label Switching(MPLS) with Label Distribution Protocol(LDP) and Resource Reservation Protocol - Traffic Engineering(RSVP-TE) signalling, several application models, and many other protocols and components.

  INET is built around the concept of modules that communicate by message passing. Agents and network protocols are represented by components, which can be freely combined to form hosts, routers, switches, and other networking devices.

The Simu5G simulator is able to simulate generic TCP/IP networks including 5G NR layer-2 interfaces through the models of the INET library, more specifically, Simu5G simulates the data plane of the 5G Radio Access Network (RAN) and core network in both Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD) modes, with heterogeneous GNBs cells (macro, micro, pico, etc.).

For the specific purposes of the development of this work, the simulator SIMU5G offers the functionality to run in real-time emulation mode, enabling interaction with real devices, thanks to OMNeT++'s real-time scheduling of events and

INET's capability to exchange IP packets between local applications or network interfaces and the simulator. These IP packets are processed by the simulator as if they were traversing the 5G cellular network. The above allows a user to run live networked applications having an emulated 5G network in the middle

Additionally, SIMU5G includes a model of ETSI MEC with models of MEC orchestrator, MEC platforms, MEC hosts and MEC services. In the latter, interfaces towards application endpoints (MEC app and Device app) are ETSI compliant, which means that one can also use real MEC-based applications and run them through a simulated 5G network, also in real time. The MEC model offered by Simu5G comes with MEC services, namely the Radio Network Information Service and the Location Service, which return information taken from the simulated 5G network. This way, a MEC developer can test real MEC applications in a realistic and fully controllable MEC-enabled 5G network which will be essential for future developments of the work presented here.

2.2.3 Smart Meters

In order to compose a realistic operational scenario in smart grids, it is necessary to consider the standard protocols for communication between devices. For this purpose, Gurux smart meter simulators are used in this work.

Gurux is a Finnish company specialized in Device Language Message Specification DLMS protocol used in smart meter communication, their products are licensed globally. The smart meter simulator provided by Gurux is an open-source software that simulates the existence of a real smart meter connected to the system, through which it is possible to perform measurement reading requests through the DLMS protocol, exactly the same way as with the real device.

Device Language Message Specification - DLMS

DLMS is a protocol standard that is used in electricity, water and gas meters globally.

Protocol standard is needed to read data from different meter types and manufacturers. DLMS is based on the following IEC standards:

- IEC 62056-21 Direct local data exchange
- IEC 62056-42 Physical Layer Services and Procedures for Connection-Oriented Asynchronous Data Exchange
- IEC 62056-46 Data link layer using HDLC protocol
- IEC 62056-47 COSEM transport layers for IPv4 networks
- IEC 62056-53 COSEM application layer
- IEC 62056-61 OBIS Object identification system

- IEC 62056-62 Interface objects

DLMS doesn't define what kind of functionality a meter must implement, only how to communicate with the meters allowing it to be used for several types of certified meters and their respective communication channels.

In DLMS there are different authentication levels, each authentication level gives different kinds of control for the meter becoming adjustable for various client-server interface configurations.

In closed systems, there is no need for authentication levels, but in DLMS one needs to establish communication with the meter before reading.

Secured connections are mandatory when data is sent Over The Air (OTA). DLMS supports three different ways to secure the data.

- Authentication
- Encryption
- Authentication and Encryption

In DLMS, three different kinds of encryption methods to secure a connection between the client and the meter are supported:

- AES-GCM-128 AES-GCM-128
- ECDH-ECDSAAES-GCM-128SHA-256
- ECDH-ECDSAAES-GCM-256SHA-384

Logical Name and interfaces

In DLMS there are interfaces that describe what kind of data you want to get from the meter. All meter manufacturers should use the same interfaces and Logical names. The standardization of logical names and interfaces makes it possible to replace meters with new ones when different meters are using the same interfaces and logical names. or even change the meter model and manufacturer and an old data collecting system can be used.

Client and Server addresses

Each authentication level has its own client address. So when the authentication level changes also client address changes. There is a client address defined in the DLMS standard only when a connection is made without authentication.

Each meter must have a unique server address. Using this address meter knows what messages to receive. Also, the client knows the message sender. The meter serial number can be usually used as the server address. This makes it possible that there are several meters operating in the same network (UDP, radio, RS-485).

PDU and frame size

The size of the Protocol Data Units PDU depends on the meter. If the meter doesn't have a lot of memory PDU size is smaller. Frame size depends on the communication channel.

Protocol specification for the MAC sublayer

The DLMS protocol uses the HDLC frame format type 3 as defined in Annex H.4 of ISO/IEC 13239 in MAC sublayer as shown in Figure 2.5.

Flag	Frame format	Dest. address	Src. address	Control	HCS	Information	FCS	Flag

Figure 2.5: HDLC Frame Format Type 3 [3]

A description of the frame can be found in the following clauses:

1. Flag field

   The length of the flag field is one byte and its value is 0x7E. When two or more frames are transmitted continuously, a single flag is used as both the closing flag of one frame and the opening flag of the next frame.

2. Frame format field

   The length of the frame format field is two bytes. It consists of one 4 bits Format type sub-field, one Segmentation bit S and the 11 bit frame length sub-field, represented by Figure 2.6:

MSB														LSB	
1	0	1	0	S	L	L	L	L	L	L	L	L	L	L	L
Format type					Frame length sub-field										

   Figure 2.6: Frame Format Field [3]

   The value of the format type sub-field is 1010 used to identify the frame format type 3.

   The segmentation bit indicates whether the data being transmitted is part of a larger message that needs to be segmented for transmission. It is used to indicate whether the data payload of the current frame is a complete message or if it is a segment of a larger message that has been divided into smaller parts for transmission efficiency.

   The value of the frame length subfield is the size of the frame in bytes excluding the opening and closing frame flag sequences.

3. Destination and source address fields

   Destination and address frames are used for usual destination and source addressing. Depending on the direction of the data transfer, both the client and the server addresses can be destination or source addresses.

4. Control field

   HDLC uses the control field to determine how to control the process of communication. The control field is different for different types of frames in the HDLC protocol. The types of frames can be Information frame (I-frame), Supervisory frame (S-frame), and Unnumbered frame (U-frame).

   The control field is a 1-2-byte segment depending on the type of the frame and it is generally required for flow and error control.

5. Header check sequence (HCS) field

   The HCS field is used to check the sequence applied only for the header and it has a two bytes length.

6. Information field

   The information field may be any sequence of bytes. In the case of data frames (I and UI frames), it carries the MAC Service Data Unit (MSDU).

7. Frame check sequence (FCS) field

   The length of the FCS field is two bytes. The frame checking sequence is calculated for the entire length of the frame, excluding the opening flag, the FCS and any start and stop elements (start/stop transmission).

2.3 Monopolistic Nature of Energy Market

A natural monopoly exists typically due to the high start-up costs or powerful economies of scale of conducting a business in a specific industry which can result in significant barriers to entry for potential competitors. The energy industry is a classic example of a natural monopoly, the bulk of power generation and power transmission entails large up-front fixed costs and once this investment is paid, each additional unit of generated energy has a very low cost, at the same time, it generates enormous economies of scale, the fixed costs will be spread between the retails customers providing increasingly reasonable final prices for more units sold. An important aspect of the energy industry that reinforces its permanence as a natural monopoly is the unavailability of energy resources for the different possible sources of energy.

Hydroelectric energy resources can be taken as an example, the water resource is considered a public good that must be made available to all members of society. Typically, these services are administered by governments and paid for collectively through taxation which usually results in exploitation through the concession regime. The point to be highlighted is that even if the high volume of investments was not an entry barrier for this specific market, the existence of hydropower potential depends on a set of geomorphological characteristics and there are few locations that can meet all the necessary requirements, the expected final result is the existence of one or few companies exploiting the totality of the hydroelectric potential. The same analogy can be applied to non-renewable sources such as nuclear power plants or thermoelectric power plants. In these cases, the competitive advantage can be established in companies that own the main sources of production resources, such as uranium, coal, natural gas, and oil.

As recognized natural monopolies, the electric utilities rates charged to customers were historically regulated by the governments in order to set at the supposed tariff modicity level and at the same time cover the operational costs and capital return. In the 1990s, there were many moves toward deregulation around the world based on the belief that competition would reduce the final energy price for consumers.

Deregulation has been orchestrated in different ways around the world, but mostly it has included competition between generators selling to distributors and competition between distributors selling to final customers. However, the regulation proponents underestimated the monopolist nature of the market and most households can't really choose their electricity supplier since the local power market is still dominated by one generator. The consequences of a deregulated market combined with the lack of genuine competition may imply market manipulation, intentionally reducing the amount of generated power in order to drive up prices.

The breaking of a monopoly brings the gains to consumers outweigh the loss to the producer, but it's not so clear whether a natural monopoly, one in which a large producer has lower average total costs than small producers, should be broken up because this would raise the average total cost.

The advancement of technologies has gradually changed the characteristics of the energy market through the growth of distributed microgeneration. New advances such as smart grids and smart homes equipped with battery energy storage systems will generate environments conducive to the free negotiation of energy on the electricity grid thanks to the new capacity for flexible operation in response to sudden variations in generation.

Therefore, the advancement of technology has allowed the opening of markets historically restricted to a few companies, allowing the entry of small companies and even individuals. The characterization of the different markets, namely perfect competition, monopoly and oligopoly is necessary so that it is possible to analyze the possibilities for the energy sector.

2.3.1 Perfect Competition

Consumers and producers are price-takers in a perfectly competitive market, which means that their individual decisions cannot affect the market price. Two main conditions must be satisfied to classify a market as perfectly competitive:

1. It must contain many producers, none of whom have a large market share.
2. Products from different producers are considered indifferent to consumers.

Most competitive industries can enter and leave the industry, there are no governmental regulation obstacles or limited access to key resources and there are no additional costs associated with shutting down a company.

In a perfectly competitive market, the price is defined by a horizontal line invariant with the quantity produced, the maximum profit point is defined at the point of intersection of the price line with the marginal cost curve, as can be seen in Figure 2.7.

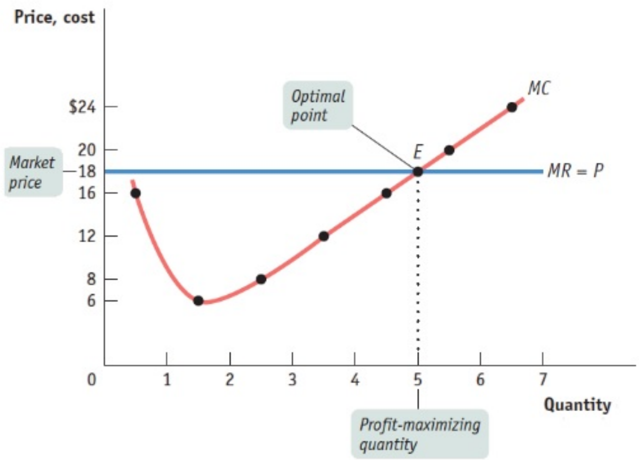


Figure 2.7: Price - Taking Firm's Profit Maximizing [23]

2.3.2 Monopoly

The fundamental characteristic of a monopoly is the capability to control the utility price by controlling the availability of a good on the market. In a perfect competitive market, the producer faces a fixed market price independent of the sold amount, a monopolist, though, can affect the price because he is the sole supplier of that good in the industry facing a downward-sloping demand curve where price decreases with quantity sold, so by reducing output, it raises the price as illustrated in Figure 2.8.

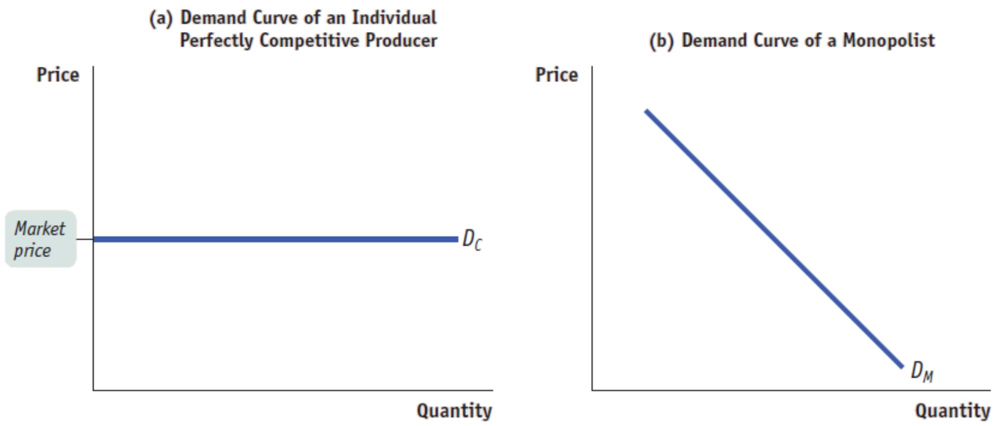


Figure 2.8: Perfectly Competitive and Monopolist Demand Curves[23]

The key to the existence of a monopolist market is the presence of strong barriers to entry, the five principal types of barriers are as follows:

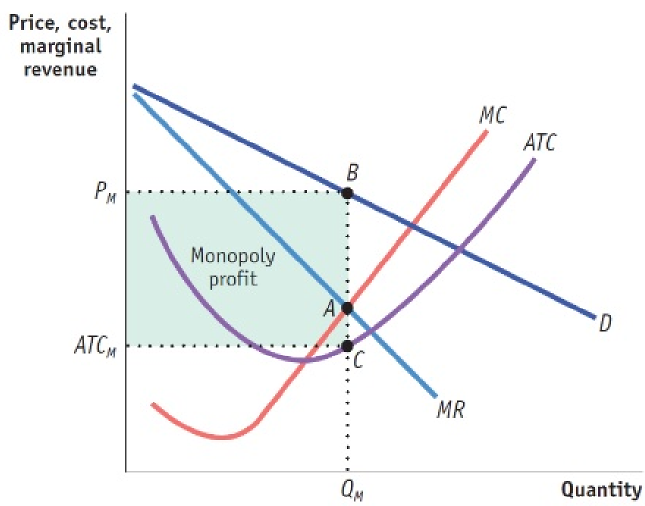
1. Control of a scarce resource or input Oil is a clear example of energy source control, oil reserves are limited and the extraction and distribution of oil is controlled by a small number of large companies and countries.
2. Increasing returns of scale Identified when long-run average total cost declines as output increases, the fixed cost is diluted with a larger number of units produced
3. Technological superiority

   The technological superiority held by a company can enable the complete dominance of the market in which it operates when no other competing product is comparable to this one.
4. Network externality

   The network externality is characterized whereby the value of a good or service to an individual is greater when many others use the same good or service. As examples of network externality we can cite energy distribution systems, the internet, and roads.
5. Government created barrier

   In some cases, the monopoly is purposely created with the intention of protecting acquired rights or creating incentives, as examples we can mention copyrights and patents.

A more general picture of monopoly shows us that the profit-maximizing level of output is the output at which marginal revenue equals marginal cost, indicated by point A assuming the usual marginal cost curve has a "swoosh" shape and the average total cost curve is U-shaped.



D: Demand; MC: Marginal Cost; MR: Marginal Revenue; ATC: Average Total Cost

Figure 2.9: Monopolist's Profit [23]

It is easy to see in Figure 2.9 that it is not interesting for the monopolist to produce a different amount of the profit-maximizing quantity (QM), the monopolist, having full control over the production of the good, will always adjust demand to operate at its optimal point.

2.3.3 Oligopoly

True monopolies are hard to find, partly because of legal obstacles, oligopolies on the other hand are much more common. Oligopoly originates in the same way as monopoly, but materializes in a weaker form, increasing returns to scale is one of the most important sources for the existence of oligopolies. An oligopoly will be formed when a small number of producers are the only suppliers of a good in a specific region. The oligopolists compete with one another, but they know that its decision about how much to produce would affect the market price, this type of competition is called "imperfect competition".

In oligopolistic markets, the unilateral choice of operating at the point of maximum profit by all companies generates oversupply which results in falling prices and consequent reduction in profits for all, therefore, oligopolistic markets will always tend to form collusion or cartels aiming to find the optimal point for everyone, in practice this is not so simple due to legal blockages and possibilities of non-compliance with agreements by the participants to increase their profit individually.

Among all types of markets, the oligopoly is the closest to the current model in Portugal and Spain, defined through the Iberian Electricity Market - MIBEL, the acquisition of electricity is done through auctions to ensure competitiveness and there is also the action of the Energy Services Regulatory Authority - ERSE for the regulation of natural monopolies.

2.4 The Electric Sector in Portugal

The electric sector in Portugal is configured in a similar way to the majority of the countries worldwide. The basic composition is represented by four classes of agents with well-defined functions:

1. Generators, responsible for the production of electric energy from available sources;
2. Transmitters: responsible for the operation and maintenance of large capacity transmission systems, these are generally systems that operate at very high voltage and transport energy over long distances;
3. Distributors: transport local energy, operate at medium/high voltage and carry out delivery to final consumers, divided into industrial, commercial and residential classes.
4. Auxiliary Agents: In addition to the main agents, the electricity sector also has important auxiliary agents, usually government agencies that take care of regulation and centralized operation of the system.

The regulation agency is important for quality control and reasonableness in tariffs, as already mentioned in the (refer to previous Section 2.3). The current electrical systems are still not well equipped with energy storage capability and therefore very dependent on an efficient control system to ensure good behavior in the face of consumption and generation fluctuations in real time, the system operator is then responsible for commanding the physical operation of the system through the various agents involved. [20]

2.4.1 Electric Sector Agents

In this section there is a brief description of the main agents in the electricity sector and their respective roles.

Entidade Reguladora dos Serviços Energéticos - ERSE

The "Entidade Reguladora dos Serviços Energéticos" (ERSE), is a public legal entity responsible for the regulation of energy services, its function is to guarantee the efficiency of the services provided, ERSE has a specific council to act on the regulation of prices and tariffs. [14]

Energy Generators

Power generators in Portugal are divided into two operating regimes named ordinary and special regimes. Producers who use renewable endogenous resources or combined heat and electricity production technologies are included in the special regime, all other producers are classified in the ordinary regime.

Redes Energéticas Nacionais - REN

The “Redes Energéticas Nacionais”(REN) is the company responsible for managing the electricity transmission systems in Portugal, it operates under a public service concession regime and its activity includes the planning, construction, operation and maintenance of the national transmission network. [16]

Energy Distributors

The Power distributors are responsible for the maintenance and operation of lines, substations, transformer units and disconnectors, these companies operate at voltage levels ranging from low to high and are responsible for the electrical systems that interconnect the transport systems to the final consumer.

Energy Trading Agents

In the context of liberalization of the energy market, energy traders have the function of acquiring energy from producers and offering it to final consumers, for which they have access to transmission and distribution systems through the payment of fees for the use of these assets.

Electricity Market Operator - OMIE

The operator of the Iberian energy market OMIE performs the integration of energy supply by producers for acquisition by traders, this activity is carried out through auctions for the purchase and sale of energy that aim to implement the necessary commercial relationships to meet demand at the lowest possible price for Portugal and Spain. The energy acquisition process is carried out in 2 principal market modalities, the Day-ahead market, the Intraday market:

1. Day-ahead market

   The day-ahead market, also called single day-ahead coupling (SDAC), is the environment where energy is acquired through purchase and sale transactions for the next twenty-four hours, the price and volume offered are defined through the intersection of forecast supply and demand curves.

2. Intraday market

   The intraday market is important for making adjustments to the generation programming defined in the Day-ahead market, this is done through the intraday auction market modality, which makes adjustments through auction sessions and through the intraday continuous market modality, which trades volumes smaller amounts of energy with greater liquidity, it can be carried out up to one hour in advance. [15]

2.4.2 The Electricity Market Model

The energy market in Portugal follows the liberalization process common to most European countries. The activities of energy generation, transport and distribution are objects of economic regulation as natural monopolies, therefore In the post liberalization model, the production and commercialization of electricity were open to competition in order to obtain greater efficiency in the management and operation of energy resources.

Within the scope of the commercialization of electric energy, energy producers make their production available through a wholesale market and the sale to the final consumer is made through the retail market. The commercialization activity is dissociated from the activities of generation, transmission and distribution in the liberal vertical model, which is why there is the figure of the energy trader who makes the commercial connection between producers and consumers.

Wholesale Market

The wholesale energy market is managed by the operator of the Iberian energy market - OMIE through market mechanisms called the Day-Ahead Market and the Intraday Market as defined in Ref. 2.4.1 - OMIE.

Retail Market

The liberalized energy market allows the free commercialization of electric energy, this happens through the retail market where each supplier offers the acquired energy and establishes supply contracts with final consumers. Ref. 2.4.1 - Energy Trading Agents.

Energy Price

The final energy price is the composition of the energy price obtained by the day-ahead market added to the price obtained with the adjustment mechanisms carried out by the intraday market. This price model is represented by the "Day ahead price", "Adjustment mechanism", "Additional cost constrains", "Additional cost intradaily m.", in addition to the adjustments carried out by the operator of the system, represented by the captions "Additional cost SO processes", "Interruptibility Service". See Figure 2.10.

It is possible to verify the predominance of the price obtained through the day-ahead market in the final price, the participation of the adjustment prices vary according to the adversities encountered.

On several occasions the amount of energy effectively demanded will be different from that foreseen at the time of contracting in the day-ahead market, the greater the need for correction, the greater the participation of the adjustment component.

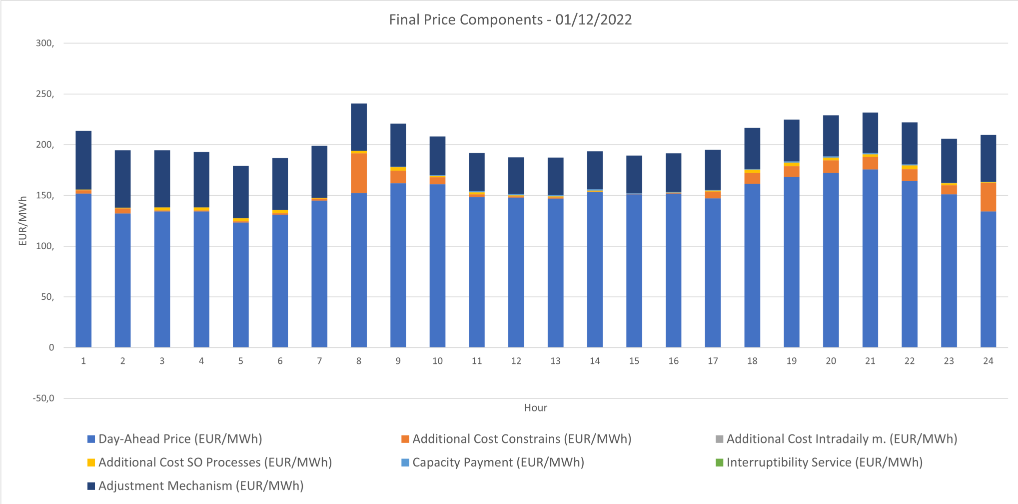


Figure 2.10: Final Price Components - Reference Retailers OMIE [15]

2.5 Benefits of a free competition energy market

The market of free competition in the energy sector presents itself as the main trend for the energy transition considering all the benefits pointed out, therefore this work aims to investigate the financial, economic and technical benefits of this scenario, as well as to evaluate the technological advances necessary for its implementation.

The energy transition is just starting around the world, there is much to be discussed and many possibilities to be explored, some of which have been addressed by this study. What can be accurately stated is that it will take the evolution of all areas that make up the current energy sector, in this context 5G Smart Grids assume a key position.

The 5G Smart Grids is flexible technology to the point of integrating all current technologies, meeting the performance requirements in terms of transmission speed, reliability in the delivery of information and security requirements that highly automated systems require.

An important aspect about the development of smart grids is that it is not an innovative invention that can revolutionize an entire industry in a short period of time, it is an ideological model that arises with the technological improvement of several areas that together bring the disruptive result, this feature makes the transitional process more complex and each current impediment to its implementation must be carefully studied and eliminated.

It is for this reason that, despite Smart Grids playing the main role in this context, the central theme of this work is the financial and economic context, in particular related to the price of energy and its impacts on the entire energy transition process.

Among the main obstacles to replacing current systems with smart grids is the large amount of capital needed to replace devices and mechanisms with others with communication capabilities and automatic action required to meet the objective functionalities. In this sense, the ever-increasing opening of the energy markets comes not only to the reduction of energy prices in general, but also to attract the necessary investment for this upgrade, as it is expected that the return on investment expected from the negotiation of energy justifies the necessary investments for it, investments that bring benefits to the energy chain as a whole.[19]

Below are some examples of how increasing the participation of producer consumers on the energy market can benefit the overcoming of a series of existing difficulties in the systems responsible for the generation, transmission and distribution of electric energy as mentioned in Section 2.2.1.

1. Growing amount of renewable energy sources:

   The great challenge for controlling voltage levels in a system with a large share of renewables comes from the fact that the variation in the amount of energy supplied varies very quickly on several occasions, reducing the time required to correct the voltage profiles. [26]

   Batteries are great allies of these systems because they are able to smooth the power variation curves, supplying the system when there is a lack of main sources or using the excess energy generated for storage.With the growing increase in renewable sources, storage possibilities demand large investments in batteries and automatic control systems.[21]

   For this reason, financial incentives for producing consumers to invest in their generation systems are extremely important, the use of batteries would optimize individual generation systems, storing the energy produced for use at times when the system most depends on it, leading to increased costs. Profits from individual systems and cutting edge prices, as a bonus, the entire system becomes more stable, flexible and reliable.

   It is important to remember that when we talk about storing the energy produced, the electrical system has many other technologies in addition to Electrochemical energy storage or well known batteries, energy stocks can be done through pumped storage hydro, compressed air energy storage, thermal energy storage, Hydrogen Chemical energy storage and others in development. [25]

   It may seem not applicable to the context, but we must remember that the entry of producer consumers into the energy market would not be limited to the residential consumer, large commercial and industrial consumers would have volume of generation and financial resources to invest in various storage systems. In addition, the widespread participation in the energy market would still subject everyone to the remuneration of the transport systems used for the power supplied, as is currently the case, this additional revenue from new generators would increase the purchasing power of the distributors, allowing for the necessary investments.

2. Interruptions in power delivery, losses in the transmission of electricity and terrorist attacks:

The rapid increase in distributed generation that can be expected with financial incentives would generate the sharp increase in microgrids, networks capable of self-sustaining becoming independent of the rest of the system. The emergence of microgrids sharply reduces system losses by preventing the circulation of energy over long distances and reduces the consequences of network outages.

3. High Consumption Demand Operation

   Currently, the need to develop energy awareness is widely publicized in order to guide people about the most appropriate times to consume energy. Financial stimulus is one of the most effective means of achieving this. Today, many people already know that high demand times are not appropriate for consuming energy, but they do not make great efforts to do so, since in the vast majority of cases energy supply contracts energy do not differentiate tariffs for different times. As an example, it would hardly be possible to convince a person to charge his electric car at 5 am instead of doing it the moment he arrives home, at 9 pm, however, if that same person knows that he is supplying the energy of his system batteries at a high value at 9 pm, the incentive to change the charging time of your car would be much greater and could even lead you to invest in automatic starters to charge at the scheduled time.

   In this way, both providing the current energy price and the possibility of direct participation in the energy market would give people the necessary incentive to intelligently manage energy consumption.

4. Electromobility and Network Modernization

   Electromobility and the modernization of networks are realities that have gained more and more space and that can be extremely accelerated through the increase in the volume of generation and private investment. Further studies could conclude that electric vehicles are very attractive with really low energy prices, which would lead many people to make an early exchange of vehicles.

5. Threat of cyber attacks

   A large part of the resources raised would inevitably need to be applied to network security, but the emergence of microgrids also comes to corroborate with a reduction in the consequences of cyber attacks since it reduces the interdependencies between different regions of the system.

   The issue of making the energy market open to everyone in a practical and quick way goes far beyond eliminating monopoly markets or making a profit from the sale of energy, although these are relevant issues, the most important thing is to achieve accessibility to the energy, as there are many people who still cannot have access to this basic commodity, it makes no sense to work to build so many technological facilities if we are unable to meet these much more basic needs.

Chapter 3

Related Work

This chapter provides an overview of the available literature related to the theme explored in this thesis.

3.1 Price models

The predictability of energy prices is a topic of great importance in the global economic context. Therefore, some studies already published were used as a background for the work presented in this document.

The "Energy Markets Forecasting. From Inferential Statistics to Machine Learning: The German Case"[13] investigates the performance of statistical methods and neural networks for forecasting the energy price in the German system. The referenced article uses the SARIMA and LSTM methods to forecast prices, however, it does not analyze the generation inclusion of the different existing technologies as exogenous variables. In addition, the work focuses on the medium and long term for forecasts and not on the short term.

The "Multi-Attribute Forecast of the Price in the Iberian Electricity Market" article [8] evaluates 24-hour multi-attribute energy price predictions for the Iberian market on the TIM ('Tangent Information Modeler') tool with AutoML ('Auto Machine Learning') capabilities. The referenced article uses data from Portugal and Spain to evaluate price forecasts, the work developed in this document focuses on the Portuguese system in order to obtain models that capture well the dynamics of the local electrical system. In addition, the work presented here brings comparisons between statistical methods and neural networks for the specific analysis of energy prices in the Portuguese system.

The "Short-Term Electricity Prices Forecasting Using Functional Time Series Analysis"[7] evaluates the Auto Regressive (AR) model and its variations for energy price forecasting in Italian energy system. The referenced work evaluates the short-term price forecasting performance in order to meet the needs of the competitive electricity market in the development of bidding strategies. The work presented in this document is similar to the referenced work in the sense of achieving accurate forecasts of the price of energy in the short term, however, the work presented in this document also seeks to evaluate methods based on neural networks and treatment of seasonal components through the model SARIMA.

The "Hybrid Wavelet-PSO-ANFIS Approach for Short-Term Wind Power Forecasting in Portugal"[1] paper proposes a novel hybrid approach, combining wavelet transform, particle swarm optimization, and an adaptive-network-based fuzzy inference system, is proposed in this for short-term wind power forecasting in Portugal. The referenced study evaluates an alternative method to those used in the work presented in this work for the treatment of time series. Furthermore, its main objective is to meet the needs of investors in the wind energy sector, while the study presented here seeks to forecast the price with a focus on the interests of network operators and end consumers.

Despite the different objectives and approaches, the methodologies proposed by the referenced studies can bring benefits to the models developed here in future work. The different characteristics of the Portuguese electrical system compared to the German, Italian and Spanish systems also bring additional contributions to the work presented in this document.

3.2 5G Communication Networks

To structure the 5G emulation system, published studies of simulation of 5G networks using the SIMU5G simulator were used as references.

The "Scalable Real-Time Emulation of 5G Networks With Simu5G"[10] presents an evaluation of the SIMU5G's emulation capabilities, showing that networks with hundreds of simulated users and tens of cells can be emulated on a single desktop machine. The 5G network emulation platform required in the work presented in this document is similar to the one used in the referenced study, however, this document brings a specific application for using the network, which is the transport of packets containing the electric power measurement energy generators installed in the electric system to the MEC server and transports the energy price prediction information back to users.

The "SIMU5G: A System-level Simulator for 5G Networks"[9] discusses the modelling of the protocol layers, network entities and functions, and validates our abstraction of the physical layer using 3GPP-based scenarios. The referenced paper aims to evaluate transition scenarios of communication systems to the 5G standard through the SIMU5G simulator, this is a different approach from the one used for this work, which uses only 5G networks in the simulations. However, the 5G Network Modeling implemented in the simulator are common in both scenarios and were used as a reference.

Chapter 4

Research Objectives and Approach

The integrating characteristic of smart grids allows it to connect technologies of different natures and also to act in the most diverse needs, whether those of technical and operational origin of the systems or from the public utility point of view.

The Smart Grids bring many possibilities for the offer of new services and products, however, this great integrative capacity brings great challenges for modelling the technology of new applications.

Therefore, this chapter describes the final objectives that this work seeks to achieve with the development of a real-time energy pricing system and the approach used to achieve them.

4.1 Research Objectives

The objective of this work is to test a real-time energy price prediction system through a 5G communication network capable of collecting data on the generation of electricity from smart meter devices applying models of energy price forecast and updating the end user with energy price values in the short and very short terms.

The research objectives are subdivided into:

1. Evaluate the performance of energy price prediction models applied to the Portuguese electricity system;
2. Evaluate the performance of the system for collecting generation and response data in 5G networks.

4.1.1 Energy Price Forecast

The objective of this work is to analyze the price of energy from the perspective of the final consumer and the intelligent operation of transmission and distribution networks, as described in Section 2.1.

Under this approach, the work seeks to verify the capacity of well-regarded forecast models for the treatment of time series aiming at predicting the price in the very short term.

Considering the hourly granularity of price and energy generation data, price predictions for one hour ahead will be made to analyze the effectiveness of the models in the very short term.

Energy price updates for very short periods are essential for energy transmission and distribution network operators, who require an automatic and immediate response from the system in case of contingency of important system elements e.g. high-power generators, voltage transmission lines, etc.

The work also seeks to evaluate the performance of models for 24-hour ahead forecasting. Precise 24-hour price forecast information enables automatic smart management systems in homes and industry to optimize energy consumption, providing individual financial gain and efficient global energy management [4].

The correct adjustment of energy price and generation forecast models is individual for each system. As an additional contribution, this work uses the Portuguese energy system as a case study, serving as complementary information to other similar publications in Europe, as per section 3.1.

4.1.2 Real Time Energy Price System

The simulation of a real-time energy pricing system developed in this work has the objective to evaluate the viability of this system in smart grid scenarios. With this purpose, the simulation of the system includes User Equipments(UEs) sending the measurement information of energy generation obtained from smart meters through an emulated 5G communications network, receiving this information in MEC server(s) and using it to obtain future energy price values and, finally, the sending of the prediction information in return to the User Equipments(UEs). More details about the development of the system are presented in Chapter 5.

To evaluate the obtained results in terms of the capacity of 5G networks to meet the requirements demanded by operation and control systems of energy transmission and distribution systems, global companies reference studies in the smart grids area were consulted.

Table 4.1 shows reference information on communication requirements for the four major 5G application scenarios in the 5G Smart Grids electricity industry in studies provided by the Deloitte and State Grid Companies [2].

The results obtained through the 5G network emulated through SIMU5G also allow evaluating aspects such as packet loss, throughput, and RLC delay in the 5G Network.

This study also seeks to evaluate the total time required to update energy prices in a prototype with limited processing capacity, which can be scaled in future

Table 4.1: 5G Smart Grids Electricity Industry Communication Requirements.

Major Application Scenarios of 5G Smart Grids electricity industry				
Scenario	**Bandwidth**	**Latency**	**Reliability**	**Connection density**
Precise load control	<256 kbps	<50ms	>99.999%	<1,000dev. /100km
Differential protection for distribution network	<10 Mbps	<10ms	>99.999%	<1,000dev. /100km
Electricity consumption information collection	US<2Mbps, DS<1Mbps	<200ms	>99.99%	<10,000 dev./km
Mobile inspection	100Mbps	<100ms	>99.999%	2 to 10 dev., local areas
Multi-station integration	100Mbps-1G	5ms-20ms		10-1000 dev., local areas

studies, making it possible to evaluate performance for larger networks.

4.2 Research Approach

This study is based on studies published in journals with Q1 Quality, and high reputable conferences (CORE ranking A).

The research of the scientific articles related price forecast and articles related to real-time 5G emulation was performed via IEEE Explore, ResearchGate, Elsevier, Academia, Google Scholar and specific journal sites like MDPI.

On these websites, there was a search around the keywords of "Energy", "Price, "Forecast", "Time Series","SARIMA","SARIMAX","LSTM","GRU", "SIMU5G".

Chapter 5

System Architecture

In this section, the emulation architecture of a real time energy price forecasting system prototype used for experimentation is described.

The core of the emulation system is provided by SIMU5G simulator operating as a real time emulator in order to obtain the performance evaluation of distributed applications running on 5G networks.

In emulation mode, SIMU5G is used as a network transport, having application endpoints exchange packets through it, in real time. The packets transported by the emulated network will suffer the same impairments (e.g., delay and losses) as if they were transported by the real network.

The SIMU5G functioning is based on OMNet++ and INET frameworks as described in Section2.2.2.

The OMNet++ modules exchange messages through connections between their gates and the behavior of a module is implemented by event handlers.

The Network description with gates, connections and parameter definitions is coded separately using Network Description Language (NED).

Last, the simulation parameter values are defined in the Initialization file (INI). The INI files are read in the runtime and initialize the model.

In a discrete event simulator the time advances because events are processed. However, OMNet++ allows the use of a real time event scheduler to flow the simulation time at the pace of wall clock. The real time emulation in SIMU5G is possible if simulated time flows faster than real time, therefore the density of events and their processing time has to be such as to not overload the system processing capacity.

The OMNet++ computer network elements like hosts, protocols, router/switches, and connections are provided by the INET library. The INET library includes many TCP/IP protocol models, such as TCP, UDP, IPv4, IPv6, OSPF, BGP and wired and wireless layer-2 protocols (ETHERNET, PPP, IEEE, 802.11, etc) allowing to simulate the communication between the endpoints of the network.

Moreover, the INET library provides the External Interface (*ExtInterface*) modules to interface the simulation environment with the host operating system.

The SIMU5G simulator core network allows users to instantiate a UPF and an arbitrary network topology where forwarding occurs using the GPRS tunnelling protocol (GTP).

For radio access, SIMU5G allows one to instantiate GNBs and UEs, which interact using a model of the New Radio protocol stack, GNBs can be connected to the core network directly.

The UEs and GNBs are modelled as compound OMNeT++ modules. Their architecture is shown in Figure[5.1]. UEs have all the protocol stack until the application layer, whereas GNBs only have communication functionalities. Both include an NR Network Interface Card (NIC), which models the NR protocol stack.

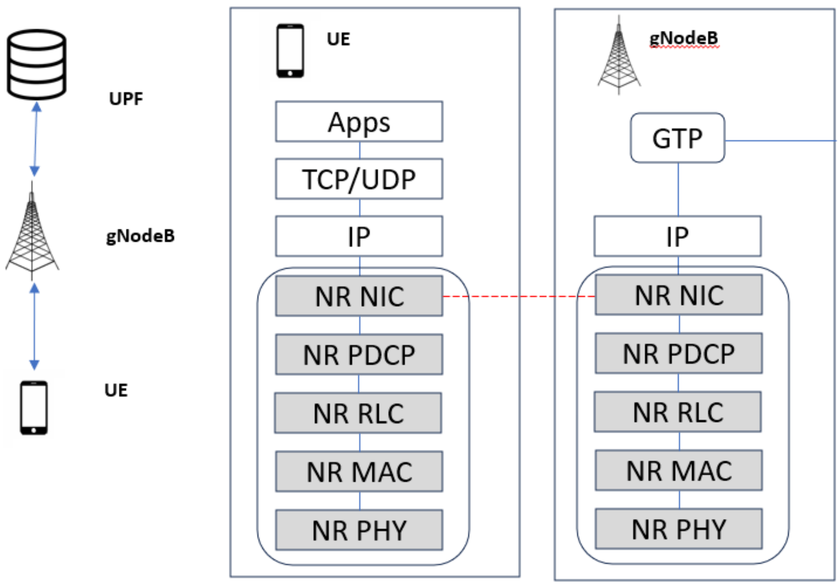


Figure 5.1: SIMU5G Modules

5.1 General Architecture

The general architecture of the real time energy price prediction system in 5G networks comprises the representation of several User Equipment (UE) devices distributed throughout the energy transmission and distribution system.

In the model defined for the system simulation, the UEs are responsible for making requests to the Smart Meters (SM) connected to the system's energy generators. Smart meters will be installed in every system power generators and will measure the power generated upon receiving the request from the UE.

The UEs are responsible for collecting measurements from all local smart meters in real time and sending the collected data to the nearest gNodeB, which is forwarded to the MEC server, afterwards.

The MEC Server will be responsible for aggregating all the information received by many UEs, processing the data to calculate price forecasts for the next 24 hours and sending the information back to the UEs.

Figure 5.2 represents the model developed to simulate the system, each separated box identifies a Linux namespace, creating an isolated system for the various modules necessary for the simulation. A detailed description of the Linux namespaces can be found in the next sections.

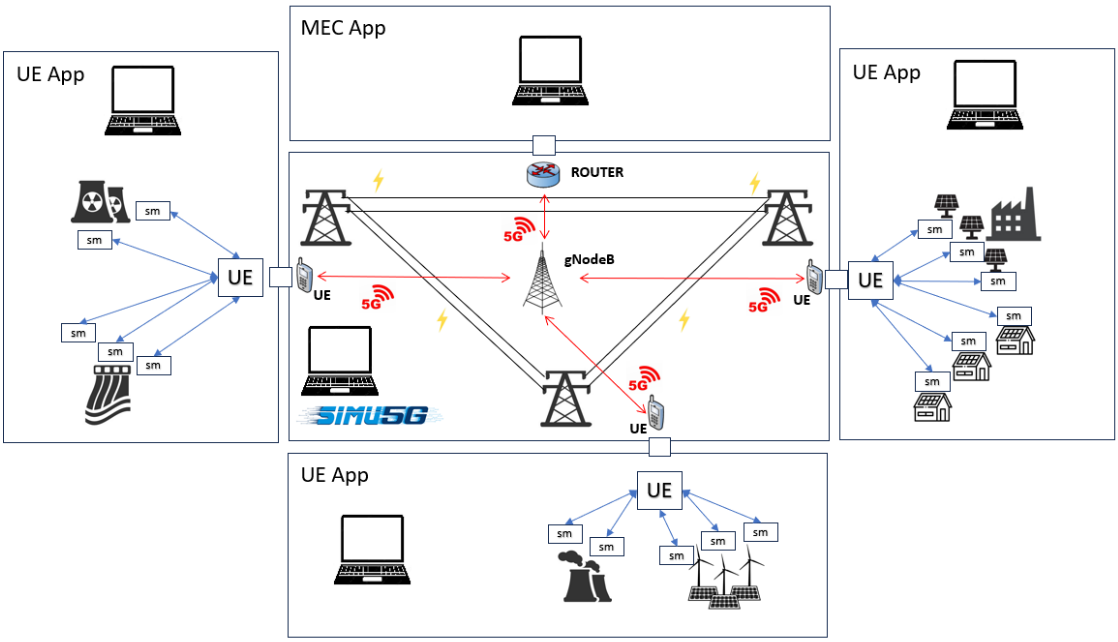


Figure 5.2: General Architecture

5.2 Network Topology

The network topology defined for the current project was composed of a MEC server, a User Plane Function (UPF) and gNodeB(GNB) modules connected as shown in the Figure 5.3.

Figure 5.3: SIMU5G Network Topology [12]

The number of UEs is variable and each UE is responsible for reading a variable number of smart meters and sending the reading information through the 5G network according to the simulation scenarios defined in Section 5.7. The UEs have fixed positions representing measurement base stations of a group of smart meters installed in power generators.

As described at the beginning of this section, the INET library provides SIMU5G with the ability to create interfaces with real applications through external interface modules.

Figure 5.4 details the emulation platform configured in 3 isolated environments defined as ns_mec0, SIMU5G and ns_ue0.

The SIMU5G is responsible for the network emulation, the ns_mec0 represents the MEC Server running the MEC applications and the ns_ue0 represents one single User Equipment (UE) running the user applications. For each UE defined in the emulation scenario, a new namespace must be created.

The isolated environments are created using Linux namespaces. The Linux namespaces make it possible to run multiple applications on a single real machine and ensure no two of them can interfere with each other.

For the communication between namespaces host OS takes care of forwarding their packets through Virtual Ethernet (*veth*) interfaces, as depicted in Figure 5.4.

The sender transmits data via a TCP/UDP socket by specifying the IP address of *veth* and the port number the receiver is listening to, the routing table of the host is configured to reroute arrived packets to the destination *veth* interface and vice-versa.

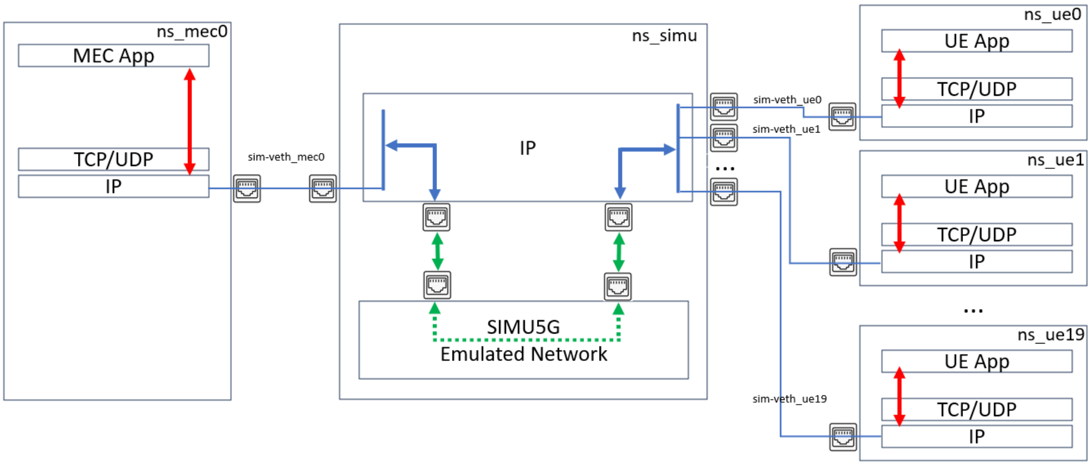


Figure 5.4: Simulation Environment

5.3 Network Configuration

The Network topology and configuration in SIMU5G are defined in the ".INI" and ".NED" files as mentioned in the introduction of this chapter.

In the ".NED" file the modular components of the network are instantiated and the connections between the modules are defined. The Network topology defined for the experiments includes a router, an UPF, a GNB and a vector of UE modules as shown in Figure 5.2.

In the "INI" file are defined the parameters and configurations of the emulation such as the number of UEs connected to the Network, the routing table configuration, and the external interface configuration.

The main parameters of the network are the ones shown in Table 5.1 defined by default in SIMU5G simulator.

Table 5.1: Main Network Parameters.

Parameter	Value
Carrier Frequency	2 GHz
GNB TX Power	46 dBm
GNB Antenna Gain	8 dBi
GNB Noise Figure	5 dB
UE Antenna gain	0 dBi
UE Noise Figure	7 dB
CQI reporting period	40 TTIs
Path Loss Model	15 UMa(Urban Macro)
Fading Model	Jakes
Shadowing Model	Log-normal distribution
UE Mobility	Static

The main ".INI" file configuration steps are transcribed below:

Step 1: Configure the routing table for the emulated network using ".mrt" routing files.

**.router.ipv4.routingTable.routingFile = "routing/router.mrt"*
**.upf.ipv4.routingTable.routingFile = "routing/upf.mrt"*
**.gnb.ipv4.routingTable.routingFile = "routing/gnb.mrt"*
...
**.ue[0].ipv4.routingTable.routingFile = "routing/ue0.mrt"*
...
**.ue[9].ipv4.routingTable.routingFile = "routing/ue9.mrt"*

The routing files define the destination address for packets arriving in the modules, see the routing files content in Chapter A. For the External Ethernet Interface configuration virtual ethernet links and external host IP address are indicated to simulation modules.

Step 2: Configure the External Ethernet Interface

**.router.numEthInterfaces = 1*
**.router.eth[0].typename = "ExtLowerEthernetInterface"*
**.router.eth[0].device = "sim-veth_mec0"*

**.ue[*].numEthInterfaces = 1*
**.ue[*].eth[0].typename = "ExtLowerEthernetInterface"*
**.ue[*].ipv4.forwarding = true*

**.ue[0].eth[0].device = "sim-veth_ue0"*
**.ue[1].eth[0].device = "sim-veth_ue1"*
...
**.ue[9].eth[0].device = "sim-veth_ue9"*
...

**.ue[0].extHostAddress = "192.168.3.2"*
**.ue[1].extHostAddress = "192.168.4.2"*
...
**.ue[9].extHostAddress = "192.168.12.2"*
...
**.router.extHostAddress = "192.168.2.2"*

5.4 Real Applications Description

The calculation and forecast of the price of energy in real time requires as input the measurement of the average power of all the energy generation available at the moment for which the price is to be calculated, therefore each generator connected to the system must be equipped with a smart meter responsible for carrying out this measurement and sending it to the server.

This section describes in detail the implementation of the applications used in the MEC server and in the UEs to achieve the objectives of the study.

5.4.1 User Equipment Applications

The User Equipments run 2 different applications denominated UE_App1 and UE_App2. The UE_App1 is responsible for requesting measurements to smart meters and sending information to the server, and the UE_App2 is responsible for receiving energy price forecast messages that are sent by the MEC server. The implementation of user applications independently is important since many users may not have installed power generation systems and will still use energy price information for consumption management.

The UE_App1 can be configured to send power measurement information from a group of smart meters, however, for system evaluations within the scope of the simulation, it is not possible to have real smart metering devices and the results would be unreliable if we do not consider the time required for requesting and receiving the reading of each of the smart meters connected to the UEs. To simulate the time delay required for this procedure, a time delay parameter is added before each packet is sent to the server. Through UE_App1's time delay parameter, it is possible to adjust the frequency of sending packets to the MEC server.

Considering the processing limitation of the host used to test the system, the criterion for defining the value of this time delay in the experimental simulations was specified as the shortest possible time that allows the receipt of all packets by the MEC server assuring it is greater than the average time required for the procedure of obtaining the smart meters measurement. The referred values was defined in 0.5 seconds.

The message sent to the MAC server is composed of 2 initial bytes containing information on the total length of the information to be sent followed by the information itself.

The sent information is structured as a string of characters composed of the smart meter identification number, reference timestamp of the measuring time, the measured value of the power of each of the energy sources connected to the smart meter and finalized by the timestamp of the moment of package shipping. The different fields of the message are separated by ";".

The format for sending the message was standardized to contain the measurement of all energy sources, considering that complex generation systems with different origins will exist in the system. Simpler systems such as residential solar generation systems, for example, would have just the information on the solar power generated and the other fields would be filled with 0.

The format of the message sent by the UE_App1 to the MEC Server is represented in Figure 5.5.

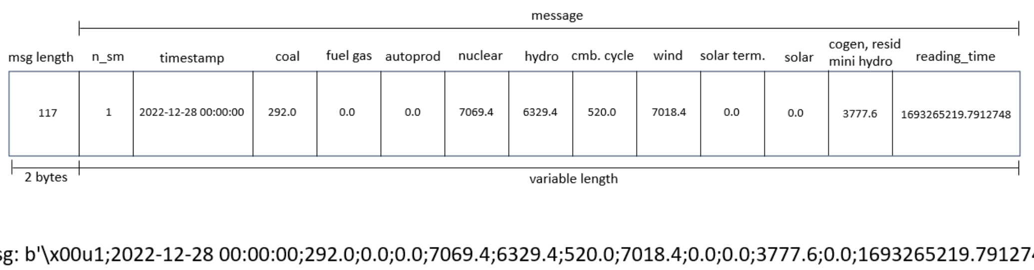


Figure 5.5: UE_App1 Message Structure

The UE_App2 is a simpler application that only receives packets sent by the MEC server with the energy price forecast. The implementation of possible automation and energy efficiency management systems that can make use of the information received is not part of the scope of this work.

5.4.2 MEC Server Application

The MEC Server application(MEC_App) running on the server must be prepared to aggregate the data received from each of the UEs connected to the network, pre-process the data to ensure the good performance of the forecast models and, finally, send the results of the calculations in return to end users who will be able to use the information in different ways as described in Section 4.1.1.

The server application was developed in Python and works through 2 Thread-based routines running in parallel:

1. **Thread 1 - Listen Messages:**

   The “Listen Messages” routine is responsible for keeping the server permanently able to receive messages sent by the UEs and only performs the function of storing received data.

   The receipt of packets is carried out according to the structure described in Figure[5.5]. The first 2 bytes are read, which contain the length of the information, and then the complete information is read according to its length.

   The Message storage is done in 2 main data structures. The first one is a Python "Dataframe" structure that acts as a buffer memory of the system. This structure stores the original information of each package received with the addition of the arrival time of the package, at the end of the emulation the data is saved for later analysis.

   The second structure is a Python "Dataframe" structure that will be effectively used to forecast future energy price values and therefore it uses a memory structure shared with Thread 2, see Figure 5.6. In this structure, the information contained is the sum of the measured values of the power of each packet received for every source. In this way, the prediction models are fed with energy generation information for the entire considered system, this stage also counts the number of the received packets for the same timestamp, the information can be used for packet loss detection and error identification systems.

2. **Thread 2 - File Operations:**

   The "File Operations" routine is responsible for forecasting the energy price for the next 24 hours. To this end, Thread 2 monitors the data received by Thread 1 and forecasts prices through machine learning models trained with historical information of the price and energy generation by source, as described in Section 4.1.2.

   The prediction routine is called whenever the receipt of packets for the current timestamp is complete, this control is done by counting the number of packets received and has a waiting time limit for the total receipt of all packets referring to each timestamp, this mechanism ensures that the routine proceed with its normal operation in case of packet losses.

   Whenever the time limit for receiving the current timestamp expires and the number of packets received does not correspond to the total number of smart meters in the system, a missing data pre-processing routine is called.

   The missing data treatment routine implemented for the test system fills in the missing data with the average of the other packets received for each timestamp. This procedure is not the most suitable for all occasions considering the differences in generation capacity between the different models of generators for each energy source, however, the ideal treatment for missing packets requires a much more in-depth study and should be carried out in future works.

   After receiving the data and the error treatment described, the prediction routine is called and its result is sent back to all user equipment connected to the network. The sending time record is attached to the message for further system performance analysis regarding time delay.

The Figure 5.6 illustrates the operation of external applications.

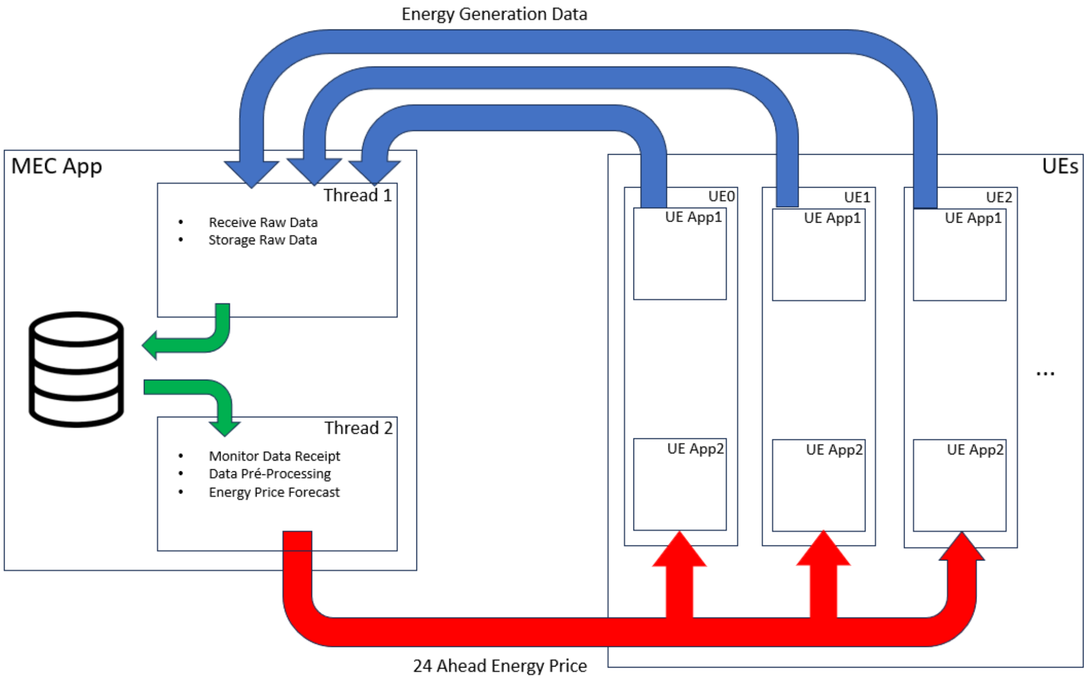


Figure 5.6: Applications Operation

5.5 Environment Configuration

The SIMU5G emulation modules use the INET external interface to send/receive IP datagrams or UDP segments between the network emulation and the real external applications as described in Section 2.2.2. In practical terms, it means that it is possible to configure the emulation in different and isolated environments as long as they can be referenced through an IP address.

The experimental configuration defined for the experiments carried out in this work uses Linux namespaces to create different isolated environments for the execution of user equipment and server applications that communicate through the 5G network emulated by SIMU5G as illustrated in Figure 5.4.

Therefore, the emulation environment must be configured to allow correct packet routing.

The Linux namespaces are created by running an executable file called "setup.sh". The following steps are necessary for the desired result:

Step1: Create namespaces for the MEC server and each one of the user equipment.

sudo ip netns add ns_mec0

sudo ip netns add ns_ue0
. . .
sudo ip netns add ns_ue9
. . .

Step2: Create virtual links between the namespaces created and the SIMU5G emulation environment.

sudo ip link add veth_mec0 netns ns_mec0 type veth peer name sim-veth_mec0
sudo ip link add veth_ue0 netns ns_ue0 type veth peer name sim-veth_ue0
. . .
sudo ip link add veth_ue9 netns ns_ue9 type veth peer name sim-veth_ue9
. . .

Step 3: Increase the namespaces maximum transmission unit (MTU) if necessary

sudo ip netns exec ns_mec0 ip link set dev veth_mec0 mtu 10000
sudo ip netns exec ns_ue0 ip link set dev veth_ue0 mtu 10000
. . .
sudo ip netns exec ns_ue9 ip link set dev veth_ue9 mtu 10000
. . .

Step 4: Bring interfaces up

sudo ip netns exec ns_mec0 ip link set dev lo up
sudo ip netns exec ns_ue0 ip link set dev lo up
. . .
sudo ip netns exec ns_ue9 ip link set dev lo up
. . .

Step 5: Bring virtual ethernet links up

sudo ip netns exec ns_mec0 ip link set veth_mec0 up
sudo ip netns exec ns_ue0 ip link set veth_ue0 up
. . .
sudo ip netns exec ns_ue9 ip link set veth_ue9 up
. . .

sudo ip link set sim-veth_mec0 up

sudo ip link set sim-veth_ue0 up
...
sudo ip link set sim-veth_ue9 up
...

Step 6: Assign the IP address with netmask 255.255.255.0 to veth

sudo ip netns exec ns_mec0 ip addr add 192.168.2.2/24 dev veth_mec0
sudo ip netns exec ns_ue0 ip addr add 192.168.3.2/24 dev veth_ue0
...
sudo ip netns exec ns_ue9 ip addr add 192.168.12.2/24 dev veth_ue9
...

Step 7: Add IP route to namespaces

sudo ip netns exec ns_mec0 route add default dev veth_mec0
sudo ip netns exec ns_ue0 route add default dev veth_ue0
...
sudo ip netns exec ns_ue9 route add default dev veth_ue9
...

Step 8: Disable TCP checksum offloading to make sure that TCP checksum is actually calculated

sudo ip netns exec ns_mec0 ethtool –offload veth_mec0 rx off tx off
sudo ip netns exec ns_ue0 ethtool –offload veth_ue0 rx off tx off
...
sudo ip netns exec ns_ue9 ethtool –offload veth_ue9 rx off tx off

5.6 System Configuration

The real time emulation in SIMU5G is only possible if simulated time flows faster than the real time, i.e., if the density of events and their processing time are not such as to overload the system processing capacity.

The above condition depends on the hardware/software system, on how a simulator is coded, and also on the scenario being run. Three processes within SIMU5G are particularly computation-intensive: MAC-level scheduling at the

GNBs; PHY-layer reporting at the UEs, and protocol stack traversal for a packet. [10]

The testbed for the experiment is composed of an Oracle Virtual Machine Ubuntu 64bit configured with 2 processors and 12 GB RAM running in an Intel(R) Core(TM) i7-9750H CPU 2.60GHz with 16 GB of RAM.

5.7 Simulation Scenarios

The emulation scenarios were defined considering the limits defined by the processing capacity of the system available for experimentation. Increasing the number of UEs or Smart Meters above the number specified in the scenarios generates inability to receive packets sent the MEC server and the consequent discarding of these packets by the network.

The dimension of the emulated system can be expanded in future studies through the use of exclusive servers of greater processing capacity to perform the network emulation functions and the MEC server function.

Each scenario has a number of UEs connected to the 5G network and each UE has a number of smart meters connected to them. During the initial tests of the system, the inability to process a quantity greater than 20 EU and 20 smart meters was verified, therefore the scenarios were created with arbitrary quantities below these. Reference studies obtained similar threshold values, reaching a maximum value of 25 UEs [10].

Five scenarios were defined for experimentation:

- Scenario 1: 10 UEs - 10 Smart Meters per UE
- Scenario 2: 10 UEs - 20 Smart Meters per UE
- Scenario 3: 15 UEs - 10 Smart Meters per UE
- Scenario 4: 20 UEs - 10 Smart Meters per UE
- Scenario 5: 20 UEs - 20 Smart Meters per UE

Chapter 6

Energy Price Analysis

This chapter presents an exploratory analysis of the composition of the energy matrix in Portugal and the necessary analyzes for the energy price forecast models fitting.

6.1 Energy Price Overview

The energy price and generation data evaluated describe the results in the Portuguese energy market from March 2022 to December 2022. The dataset is available in the OMIE [15] platform.

Section 2.4.1 details the energy auction process for price formation.

Figure 6.1 presents the electricity generation in Portugal by technology compared to the price of energy for the period evaluated in this study.

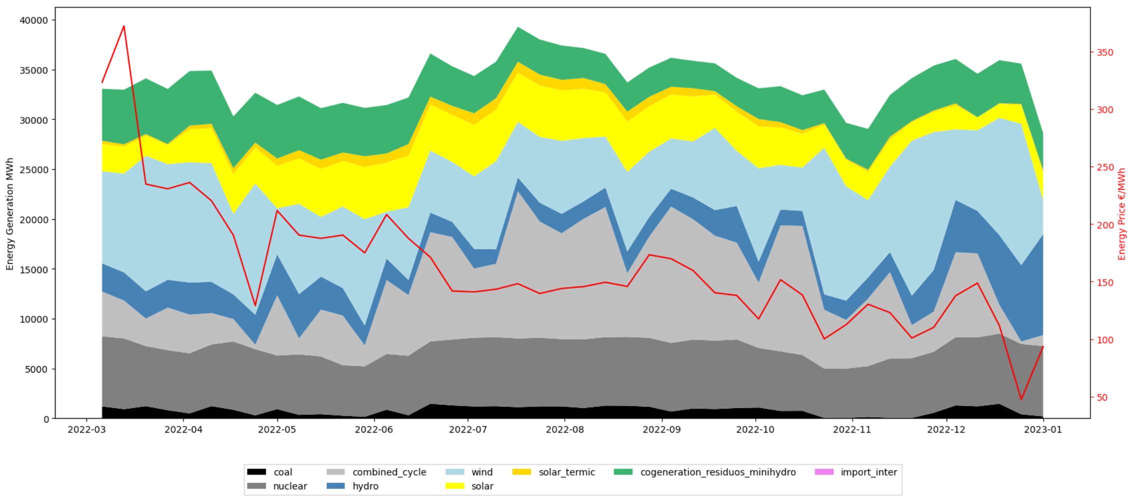


Figure 6.1: Hourly Energy Price and Energy Generation by Technology [15]

On March 8 2022, the energy price in Portugal reached 570 Euros €/MWh, a significantly higher level than the rest of the period analyzed. The increase in prices in this period is coincident with the beginning of the conflicts between Russia and Ukraine in February 2022, it is then an unusual variation in prices.

The availability of different renewable energy sources will vary depending on weather conditions and the season of the year. Non-renewable sources, in turn, have more stability in terms of availability, but their cost varies with their raw material, which is generally subject to variations in the international market.

The intrinsic characteristics of the energy composition of each system generate different energy dispatch strategies from the electrical system operators, however, it is common to identify complementarity between energy sources due to the fact that the total energy demand does not suffer large variations in the short term, but the availability of some energy sources does.

In Figure 6.2 it is possible to visualize the correlation between the different technologies and also the correlation of the price of energy with each of the technologies present in the Portuguese system from July 2022 to December 2022.

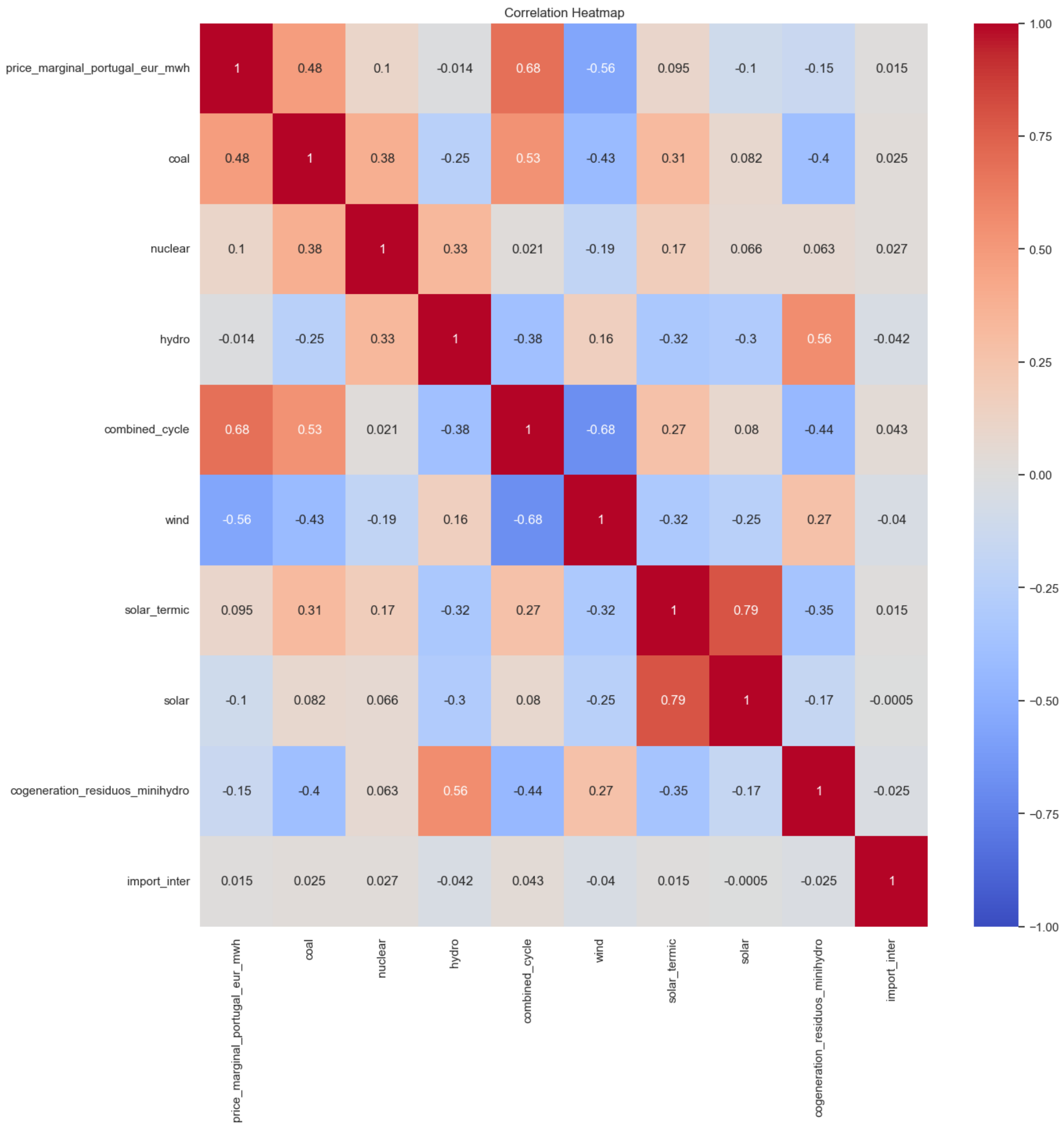


Figure 6.2: Energy Price and Generation Heat Map [15]

The atypical period of high prices in the system was removed in this analysis to avoid distortions in the real correlation between the price and the various related sources.

The final result shows a strong positive correlation between coal and combined cycle sources and a strong negative correlation between price and wind source.

Among generation technologies, it is possible to identify negative correlations between renewable and non-renewable sources, which is expected for a country undergoing an energy transition process.

For a more in-depth analysis of daily energy price, Figure 6.3 shows an hourly box plot from energy price data.

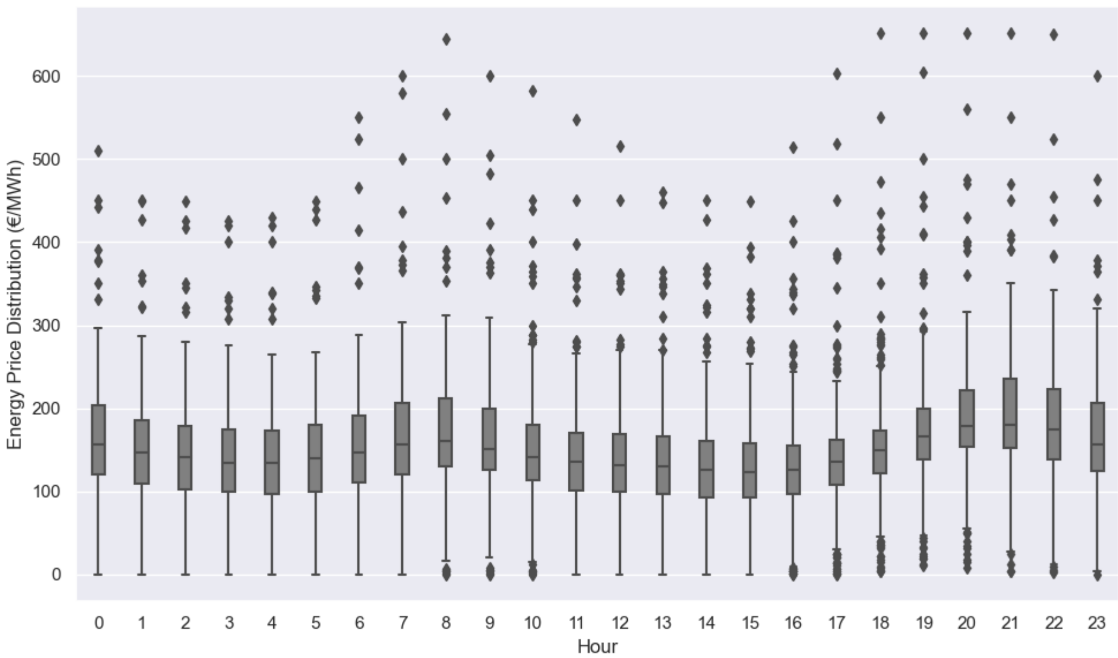


Figure 6.3: Hourly Energy Price Boxplot by Hour

The hourly price box plot throughout the analysed period points to the intervals of 7:00-9:00 and 20:00-23:00 as daily energy price peaks and a distribution of each generation technology is represented in Figure 6.4.

The formation of energy prices is a much more complex issue than it may seem in superficial analyses. Technologies with a greater amount of energy generated at times when the price is higher are not always the cause of price increases, there will be several occasions when the increase in generation from a specific technology comes precisely to avoid indiscriminate price increases.

The energy price is affected by the availability of energy sources. Every technology for power generation has its availability affected by many factors, which makes the final price prediction a complex matter.

Furthermore, the process of free competition between producers also generates upward and downward fluctuations according to the offers made available in daily auctions.

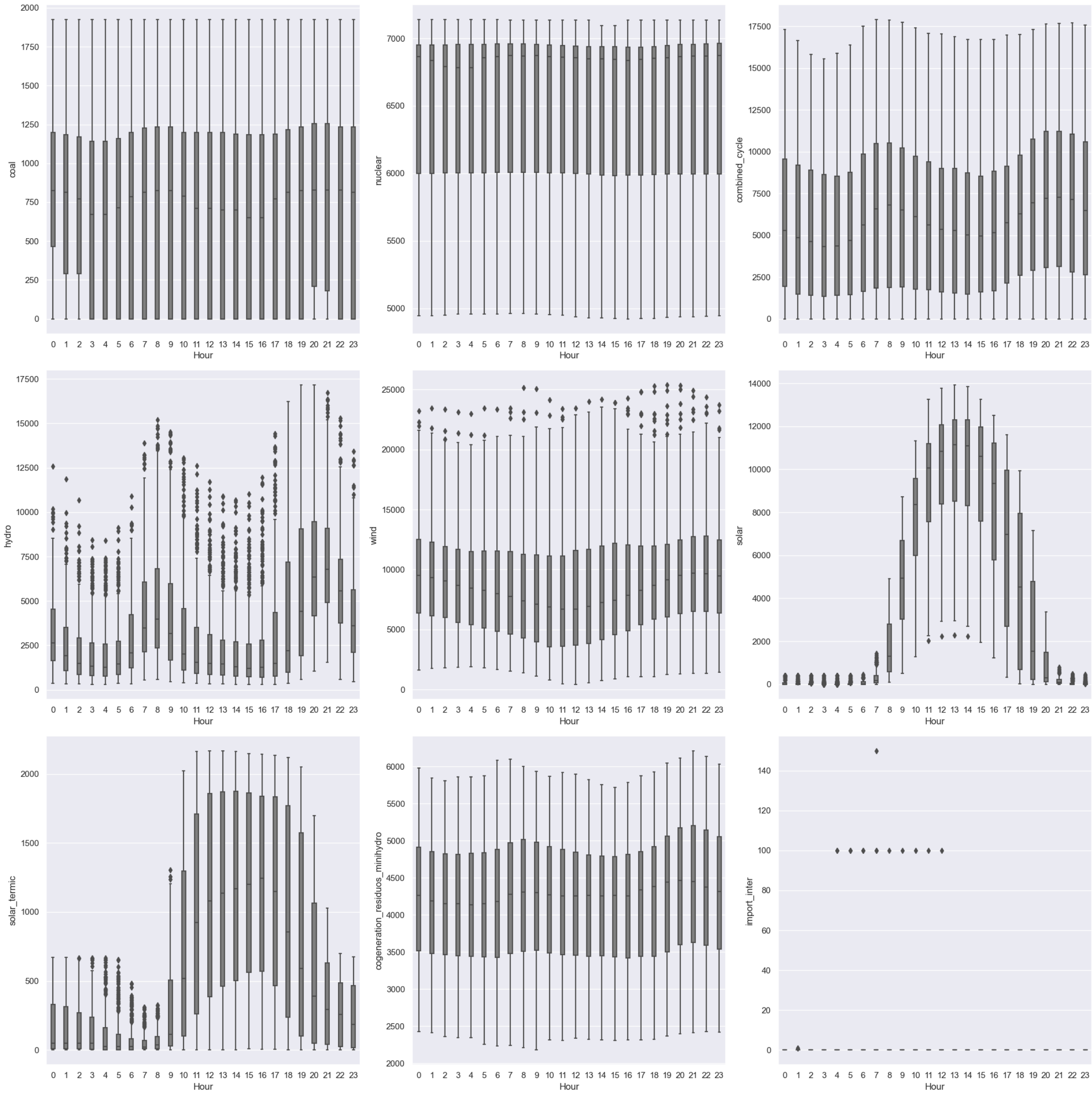


Figure 6.4: Hourly Energy Generation Boxplot by Hour

The objective of this study is not a detailed analysis of the factors that positively or negatively affect the price of energy, but rather to verify whether machine learning and statistical methods are capable of identifying existing patterns to produce reliable price forecasts.

6.2 Fitting Models

For the effective operation of a real-time energy pricing prediction system, an in-depth analysis of the data that will be used for training the prediction models is very important. Through data analysis, it is possible to adjust the forecast models in order to obtain reliable results.

The experimental analyzes carried out in this study involve the receipt of energy generation information obtained from the measurement of smart meters installed in the various generators of the system. The MEC server receives the measured values and uses them as input data in forecast models previously trained with price information and generation amounts from previous periods.

This study uses energy price and energy generation values obtained from the official OMIE website dataset [15]. The dataset used contains hourly values of energy prices and generation amounts for each source available in the Portuguese electrical system from March 2022 to December 2022.

The dataset was divided between training and test dataset in the proportion of 80% and 20% respectively as shown in Figure 6.5. The training dataset are used to train the forecast models while the test dataset are used as input for prediction.

The Energy power values are collected in real time by simulated smart meters connected to user equipment and sent to the MEC server for price prediction. For models that use past price and energy values as input, it is assumed that the server keeps this information in storage. The predicted values are then compared with price values contained in the test energy price dataset to evaluate the performance of the prediction models.

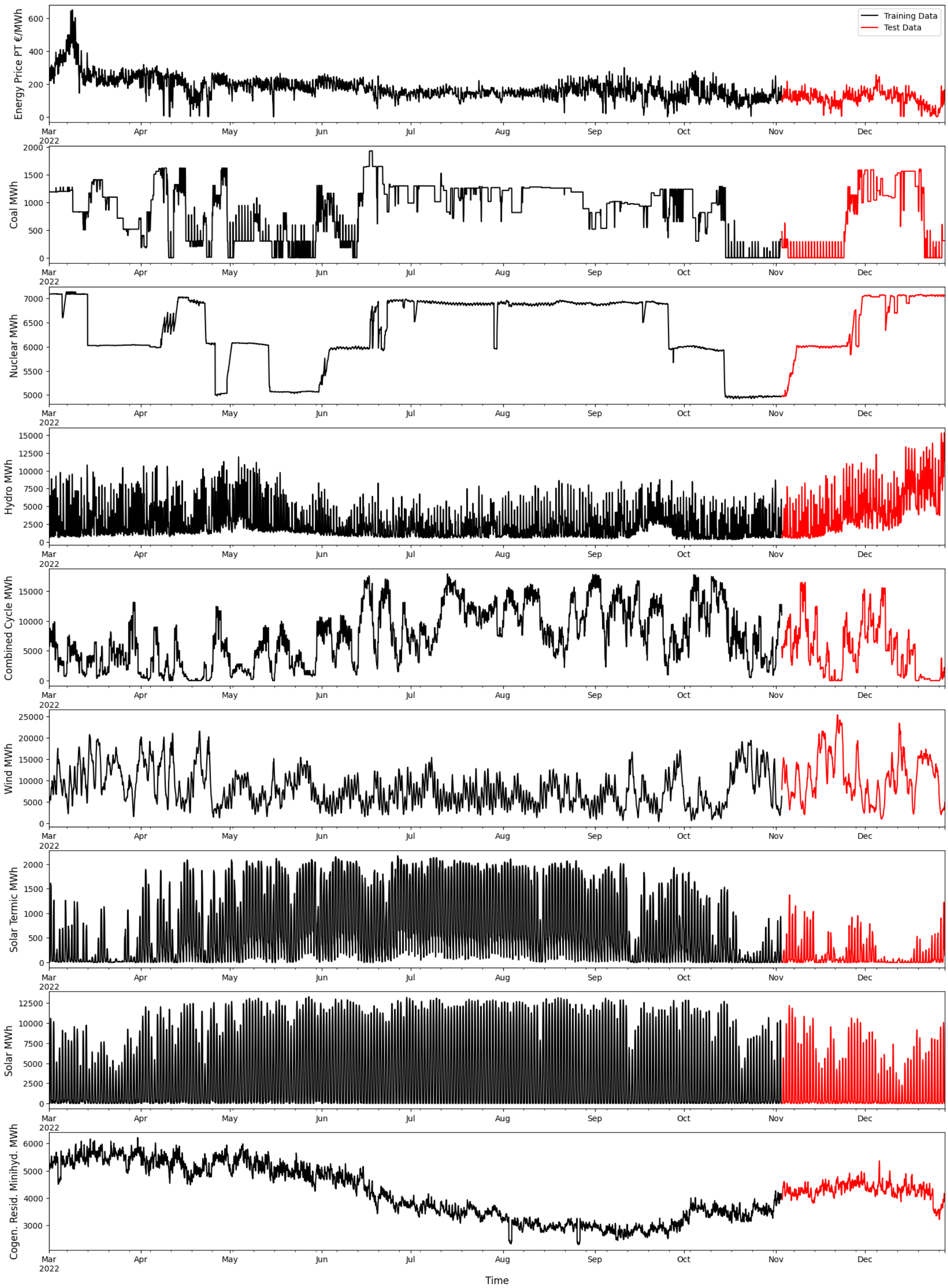


Figure 6.5: Energy Price and Generation Dataset

The SARIMA model fitting requires understanding the characteristics of the time series data like trends, seasonality, or irregular patterns for the specification of several parameters.

This section presents the step by step used to specify the parameters used in the SARIMA models and at the same time provides knowledge of the time series that can be used to adjust other forecast models.

The SARIMA model fitting requires the following parameter definition as described in Section 2.1.2:

$$SARIMA(p,d,q)(P,D,Q)s$$

- p: Autoregressive order (AR order)
- d: Degree of differencing (integration order)
- q: Moving average order (MA order)
- P: Seasonal autoregressive order (SAR order)
- D: Seasonal degree of differencing (seasonal integration order)
- Q: Seasonal moving average order (SMA order)
- s: Seasonal period (number of time steps in one season)

The autocorrelation function for the energy price time series (TS) can be used to identify trend and seasonality that must be removed to obtain the stationarity.

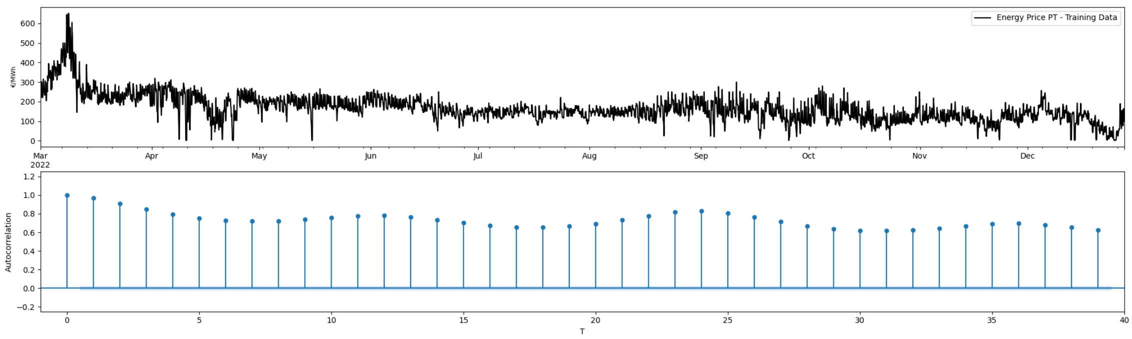


Figure 6.6: Energy Price TS and Autocorrelation Function

The Figure 6.6 shows clearly a non-stationary process with the presence of a trend. A first order simple differencing process can be used to remove the trend.

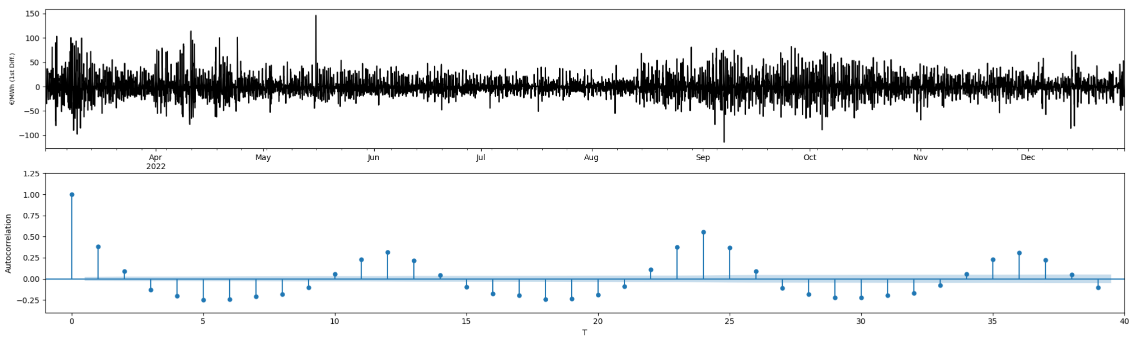


Figure 6.7: 1st Order Diff Energy Price TS and Autocorrelation Function

The Figure 6.7 shows that the result of a first order differencing process seems to be effective for the trend removal, so the degree of differencing (d) parameter can be fixed in 1.

$$d = 1$$

It is possible to observe in Figure 6.7 that seasonality is still present, so the next step is to try to remove the seasonality through a first-order seasonal differencing process. It can be noted that there are high positive correlation indices for lags T=12 and T=24, the respective results of the seasonal differentiation processes are shown in Figure 6.8 and Figure 6.9.

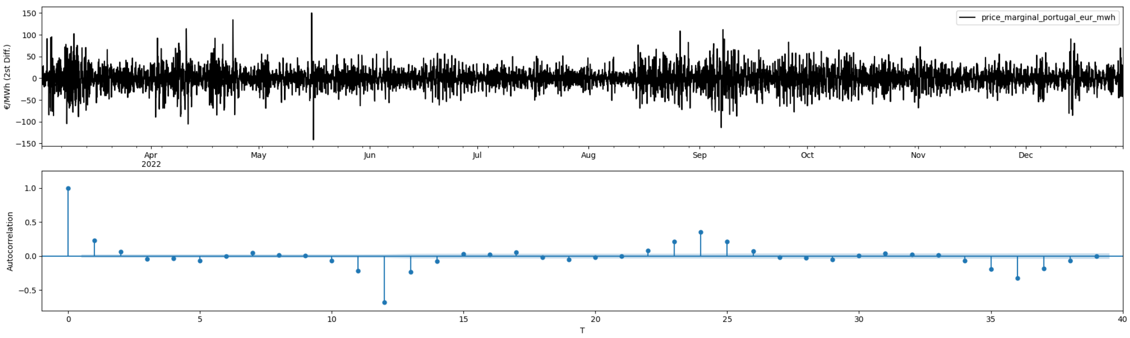


Figure 6.8: Seasonal Diff. Energy Price TS - T=12

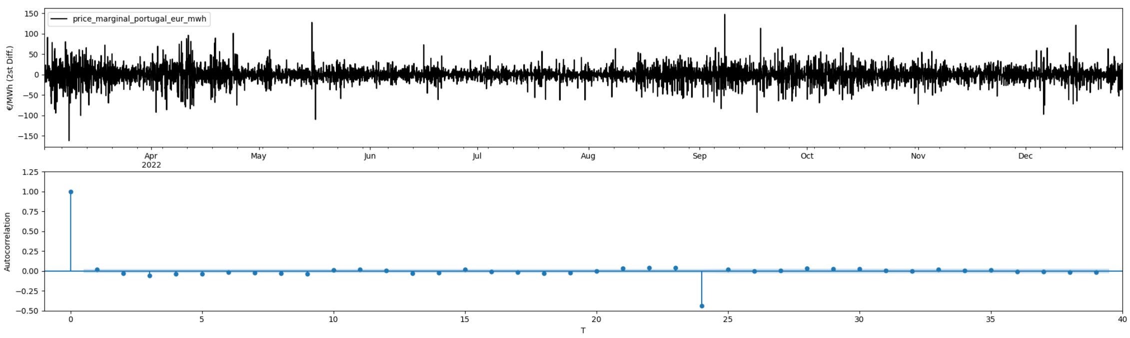


Figure 6.9: Seasonal Diff. Energy Price TS - T=24

The results achieved by the seasonal differentiation process at T=24 seen to be more effective for obtaining stationarity.

The only autocorrelation index with value distant from 0 is the T=24 lag itself, Second-order differentiations and other T values were tested in an attempt to eliminate it, but they proved to be less effective, therefore, the seasonal degree of differencing parameter (D) was set to 1 and the Seasonal period parameter(s) was set to 24.

$$D = 1$$

$$s = 24$$

The statistical Dickey-Fuller (DF) unit root test was applied to confirm the stationary assumption. As the p-value is close to zero and the augmented Dickey–Fuller (ADF) statistic is lower than the critical value at 1% we can assume the TS is stationary.

ADF Test - Energy Price Data
ADF Statistic: -4.328226
p-value: 0.000397
Critical Values:
1%: -3.431
5%: -2.862
10%: -2.567

ADF Test - Energy Price Data (1st Diff + Seasonal Diff 24)
ADF Statistic: -22.846005
p-value: 0.000000
Critical Values:
1%: -3.431
5%: -2.862
10%: -2.567

For defining the auto-regressive order parameter (p) and the moving average order parameter (q) the ACS and PACS were verified. See Figure 6.10.

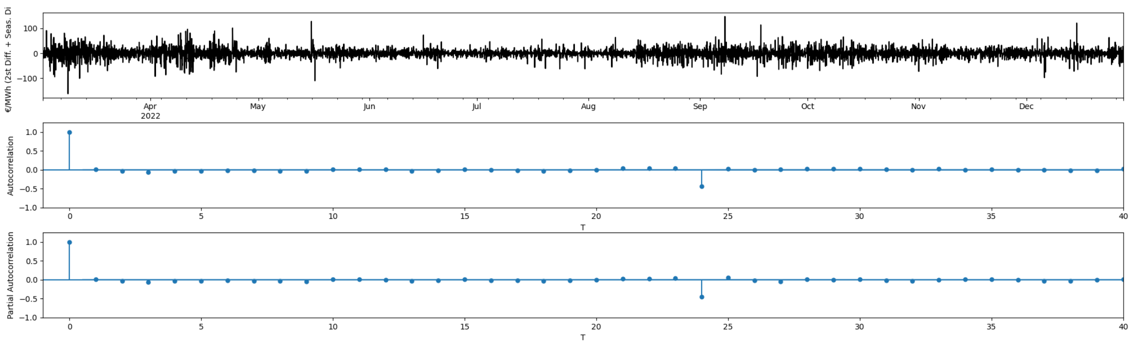


Figure 6.10: Stationary TS, ACS and PACS

Since in both correlograms, the ACS and the PACS, drop values at lag T=1, it is possible to assume that the regressive order parameter (p) and the moving average order parameter (q) can be set to 1.

$$p = 1$$

$$q = 1$$

The same procedure was executed to the definition of the Seasonal autoregressive order (P) and Seasonal moving average order (Q) for the correlograms at lags T=[0,12,24,36]. See Figure 6.11.

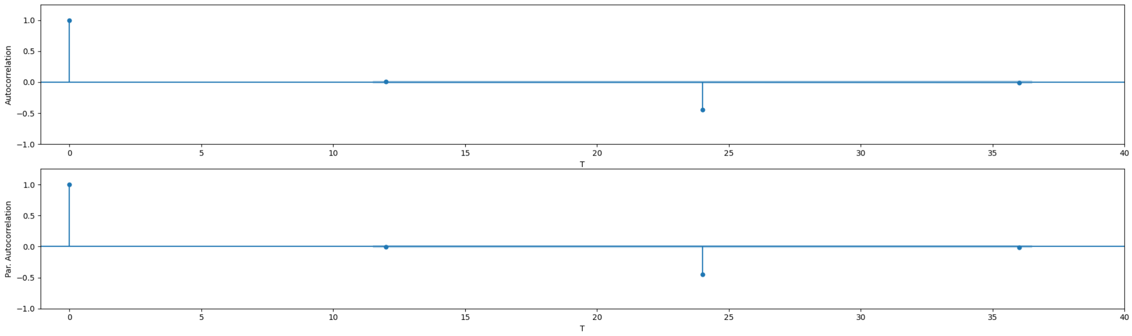


Figure 6.11: ACS and PACS for lags[0,12,24,36]

Once more, the presence of a component in lag T=24 for ACS and PACS raises doubts about the choice of parameters P and Q, to choose the best combination between P and Q the residuals evaluation was performed for $P, Q = [0, 0]$ and $P, Q = [1, 1]$.

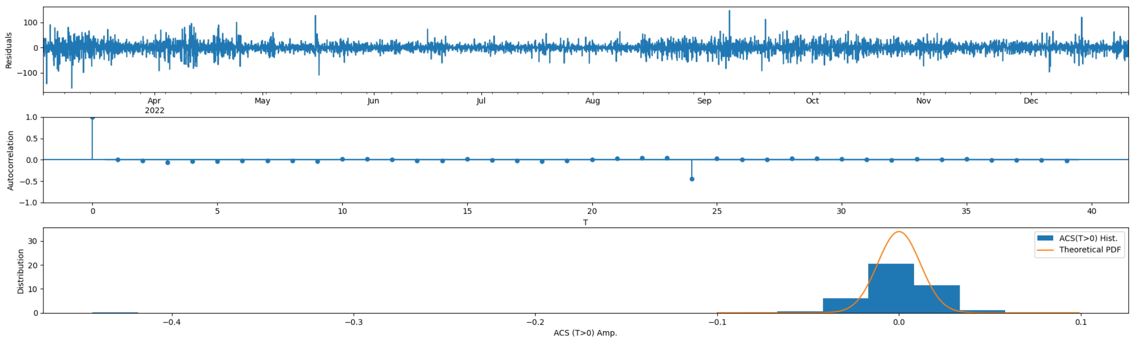


Figure 6.12: Residual Evaluation P,Q=[0,0]

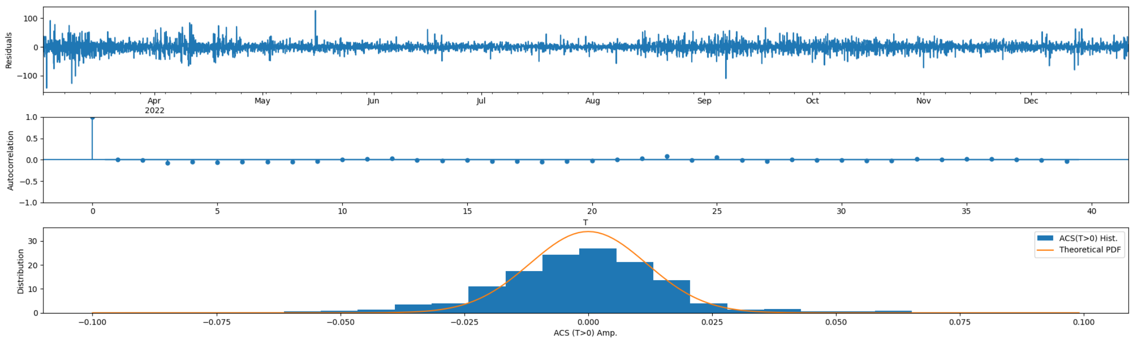


Figure 6.13: Residual Evaluation P,Q=[1,1].

As can be seen in Figure 6.12 and Figure 6.13, the residual ACS obtained using the parameters $P, Q = [1, 1]$ does not have components far from the value 0 and has an AIC index lower than that obtained using the parameters $P, Q = [0, 0]$. Therefore, it can be concluded that the model has a good fit for $P, Q = [1, 1]$.

$$P = 1$$

$$Q = 1$$

As a result of the evaluation of the time series, it is concluded that the most adequate parameters for model adjustment are p=1, d=1, q=1, P=1, D=1, Q=1 and S=24.

$$SARIMA(1,1,1)(1,1,1)24$$

Once the parameters of the SARIMA model have been adjusted, an assessment of the impacts of the different energy sources present in the system under study is relevant for the selection of the features that most influence price formation.

6.3 Features Selection

The energy price is obviously influenced by the variation of the different energy sources available in the system and their respective production costs. Therefore, an analysis of the influence of each of the energy sources is important to identify those relations.

The Figure 6.14 shows the energy price scatter plots in relation to the produced quantity of each of the energy sources available in the Portuguese system.

In the energy price plot in relation to imported energy, position (1,1) of the scatter grid, it is easy to notice that energy imports are zero for the vast majority of points, so this attribute is immediately discarded as it offers little influence on the price.

Renewable sources, namely solar(3,2), solar thermal(3,1) and wind(2,3), show a downward trend in energy prices at high production levels as expected, especially the wind source, which has a large share in the Portuguese energy matrix.

The thermal sources, coal(1,2) and combined cycle(2,2), also show behaviour within the expected range and have a positive correlation with the price of energy. It is understood that the biggest influencer of the increase is the combined cycle due to the large participation of this source in the system. Coal is generally used on an emergency basis to support the system and it is natural to have its production increased at times when the price is higher, as these are times when there is an energy shortage.

Co-generation, residuous and mini hydro sources(3,3) show a positive correlation with price. It is necessary to be more careful when analyzing these sources, as they are not sources that are always available, which means that they can be used in the system in emergency moments such as coal or simply have their dispatch aimed at greater financial gain for the producers, which explains higher generation values when the price of energy is also higher.

Hydro(2,1) and nuclear(1,3) sources do not present a neutral behaviour in relation to price in this analysis, they are typically cheaper energy sources and serve as the basis for the system, however, it is important to note that there are no high price points when hydro generation is high and there is also a large concentration of points with high prices when nuclear generation is high.

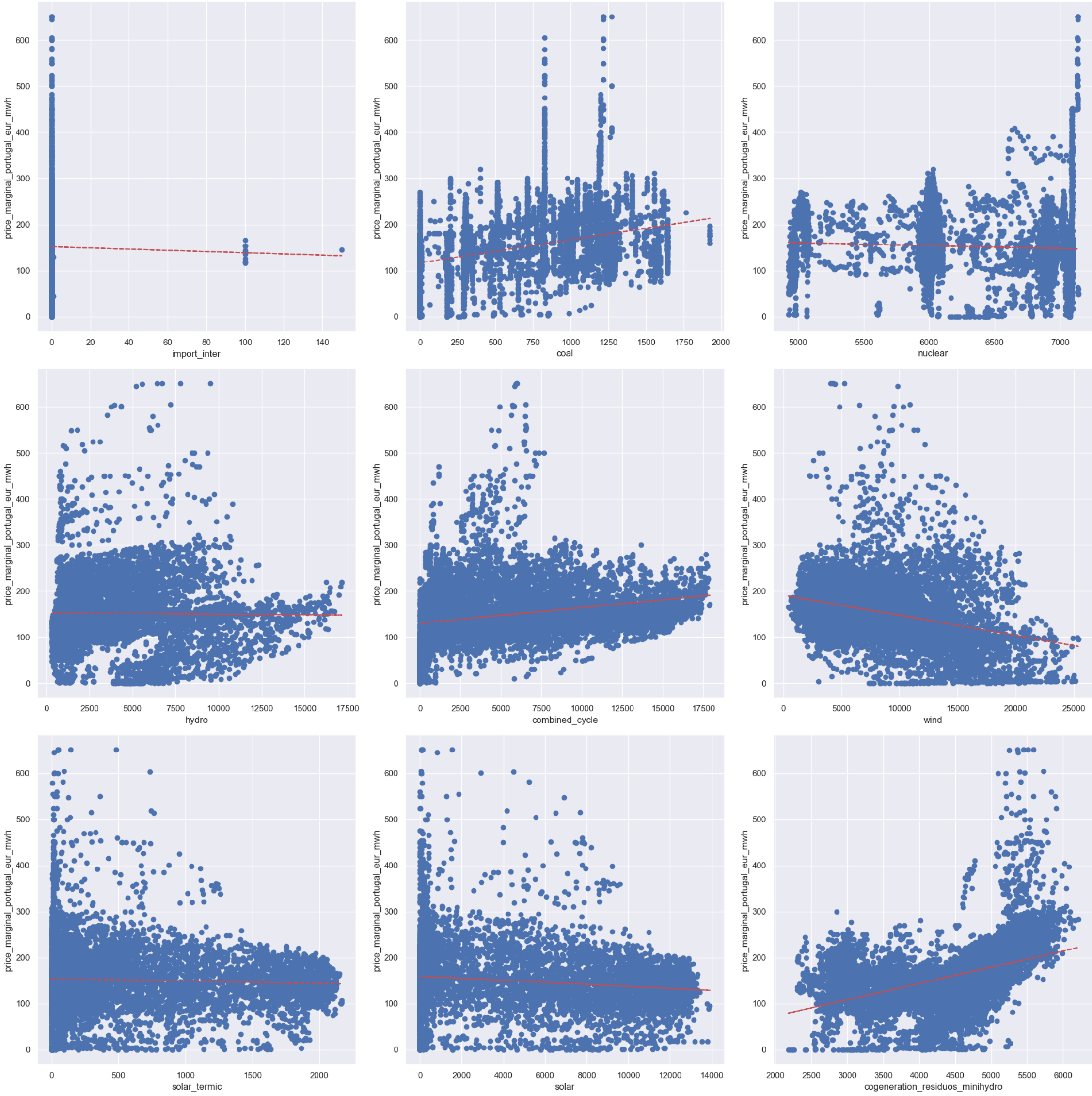


Figure 6.14: Energy Price vs Energy Generation Technologies Scatter Plots

Another important point to be considered is the influence of generation from different sources in the following hours of the day, particularly important for 24-hour ahead forecast. Figure 6.15 shows cross-correlation plots between energy price and generation for each of the system's energy sources for this analysis.

The cross-correlation graph shows the energy price behaviour in relation to the variation of each generation source for T lags from 0 to 25 hours. Positive cross-correlation indices indicate that the price of energy T hours ahead tends to increase when the generation of the analyzed source increases and negative indices indicate that there is an inverse relationship between the 2 variables for lag T.

In general, the results corroborate those presented in Figure 6.14.

Energy sources based on coal(1,2) and combined cycle(2,2) always have a positive correlation with the price of energy for any lag T considered, being a strong indicator for identifying the price in the next few hours with higher amounts of generation of these sources. An expected behaviour, as presented in the analysis

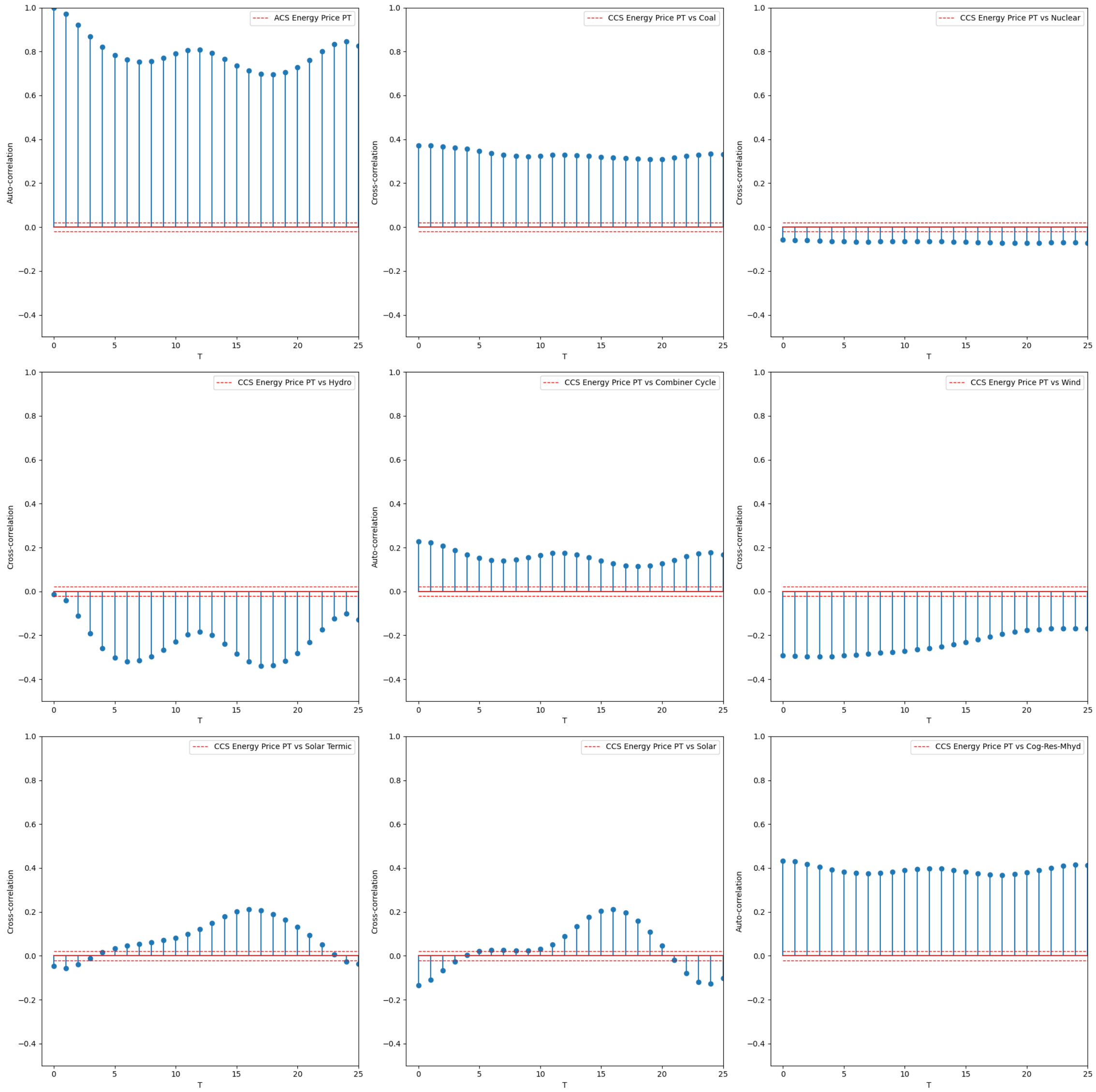


Figure 6.15: Energy Price vs Energy Generation Technologies Cross Correlation Plots

of scatter plots before.

The wind source, in turn, has a strong inverse correlation, being a good indicator to identify the price reduction whenever there is an increase in wind energy generation in the electrical system also corroborating with the analysis carried out earlier.

For water and solar sources, a cyclical behaviour is observed. The water source shows a neutral behaviour on average in the previous analysis carried out with scatter plots, however, the cross-correlation graph makes clear its action in the sense of lowering energy prices for later hours, as it is a base source it enters into supply the system whenever there is a shortage of renewable and/or less costly resources.

The solar source shows a different behaviour than initially expected since it is only imagined the reduction of energy prices with larger amounts of solar generation.

The cross-correlation for solar sources results clearly shows an upward trend in energy prices for T-lags between 10 and 20 hours.

Considering the daily cycles of the solar source, this behaviour can be explained by the need to replace this source at night, which is probably done with higher-cost sources.

Despite this effect, the influence of solar energy is clear in the sense of reducing the price by the negative index at lag T=0, confirming that higher values of solar generation promote a reduction in prices at the same time it is generated.

The nuclear energy results indicate price reductions in a constant way.

Finally, the cogeneration, residual and mini-hydraulic sources maintain the behaviour observed in the previous analysis, always showing an increase in energy prices for higher amounts of generation.

Based on the performed analyses, guidance was obtained for choosing the most appropriate set of variables to increase the performance of the forecast models. Even if the price formation is done by the composition of all energy sources, the inclusion of all variables can make it difficult to evaluate trends by forecast models and better results can be obtained with the correct selection of features.

Therefore, the hydraulic, combined cycle, wind and solar sources were selected initially because they are the sources that are very present in the systems, reaching values close to 20 GW of power generated in the system for some hours of the day, in addition to having a clear influence on the price of energy, as verified in previous analyses.

The nuclear source was rejected for its low participation and behaviour relatively close to the water source that is already contained in the chosen set.

The coal source was also rejected. Both, the coal source and the cogeneration, residual and mini hydro sources have little participation in the energy matrix and both seem to be used to support the system in the face of high energy demand or scarcity of other sources, despite the presence of cogeneration, residuous and mini hydro sources is a much stronger indicator of high energy prices than coal, as can be seen in the cross-correlation indices, so it was included in the selection set.

Energy sources selected as models features:

- Hydro
- Combined Cycle
- Wind
- Solar
- Cogeneration-Residuous-Mini Hydro

Other combinations of the feature set, including the set with all the variables, were tested in the performance evaluation of the models and did not show more satisfactory results. However, a much more in-depth analysis can be done in future works fully dedicated to improving energy price forecasting models through feature selection.

The number of layers used for the LSTM and GRU models were defined by exploratory analysis and those that obtained the best prediction results were chosen. A 3-layer configuration were selected as follows: The LSTM/GRU layer with 128 neurons, a hidden layer with 64 neurons and an output layer with 1 neuron, for 24-hour predictions the output layer has 24 neurons.

Chapter 7

Results

This chapter presents the results obtained through real-time energy price forecasting system simulations. The results are divided into 2 main parts, the first of which focuses on the performance of the evaluated price prediction models.

The second part focuses on the performance of the 5G network and the total response time of the systems from sending the generation data to receiving the response from the server with the price forecast information for the next 24 hours.

7.1 Energy Price Forecast

In this section, the results of the models evaluated for predicting the price of energy are presented.

The models were evaluated for 1-hour ahead and 24-hour ahead forecasts.

The performance evaluation metrics used were MAE, RMSE and MAPE 2.1.4 between the real price value obtained from the training time series and the value obtained as forecast by the model.

7.1.1 SARIMA

The Seasonal Autoregressive Integrated Moving Average (SARIMA) model was evaluated using the parameter setup defined in Section 6.2. As it is a univariate model, the exogenous variables of energy generation by source are not included.

The Figure 7.1 presents the time series obtained as 1 hour ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period.

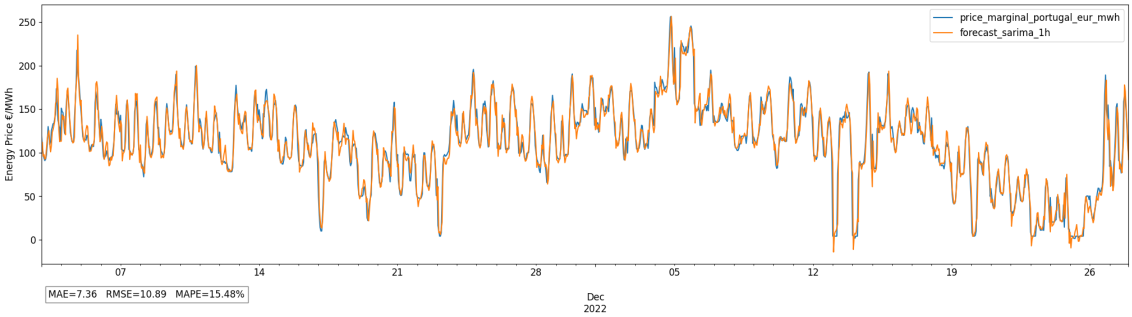


Figure 7.1: SARIMA - 1 Hour Ahead Forecast

The Figure 7.2 presents the time series obtained as 24 hour ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period.

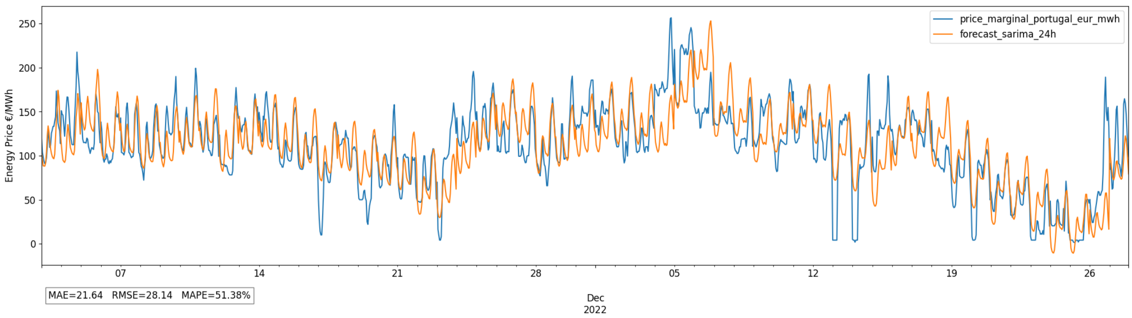


Figure 7.2: SARIMA - 24 Hours Ahead Forecast

7.1.2 SARIMAX

The Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX) model was evaluated using the parameter setup defined in Section 6.2. As it is a multivariate model, the exogenous variables of energy generation by source are included as defined in Section 6.3.

The Figure 7.3 presents the time series obtained as 1 hour ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period.

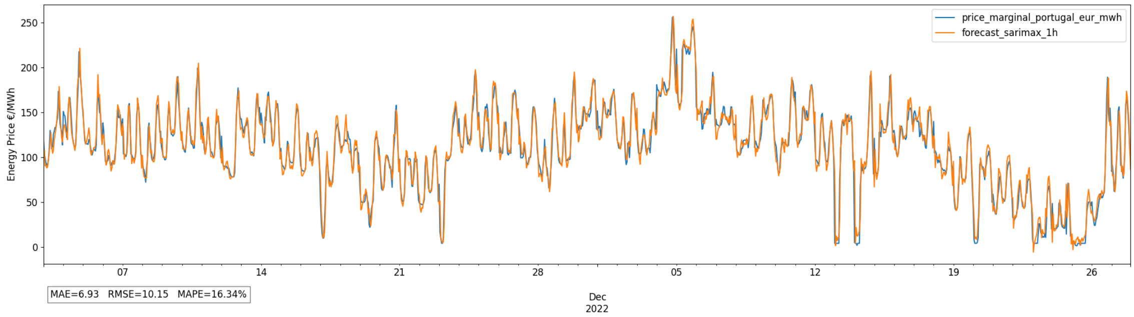


Figure 7.3: SARIMAX - 1 Hour Ahead Forecast

The Figure 7.4 presents the time series obtained as 24 hours ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period.

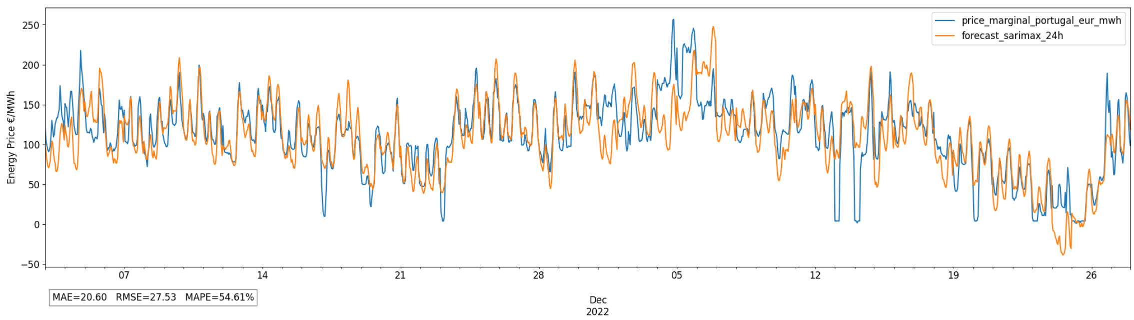


Figure 7.4: SARIMAX - 24 Hour Ahead Forecast

7.1.3 LSTM

The Long Short-Term Memory Recurrent Neural Networks (LSTM) model was evaluated using the parameter setup defined in Section6.2. The model was used in univariate and multivariate mode.

The Figure 7.5 presents the time series obtained as 1 hour ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period. For the presented results, the model was used in univariate mode, therefore only the energy price itself was used for training and as input data.

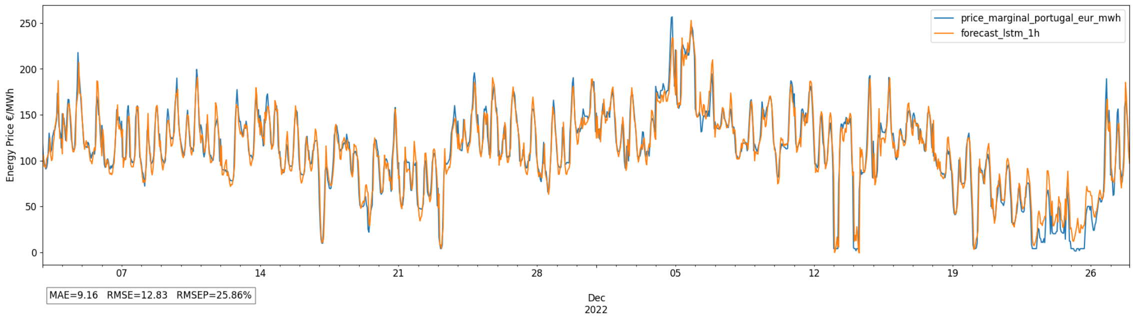


Figure 7.5: LSTM - 1 Hour Ahead Univariate Forecast

The Figure 7.6 presents the time series obtained as 24 hours ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period. For the presented results, the model was used in univariate mode, therefore only the energy price itself was used for training and as input data.

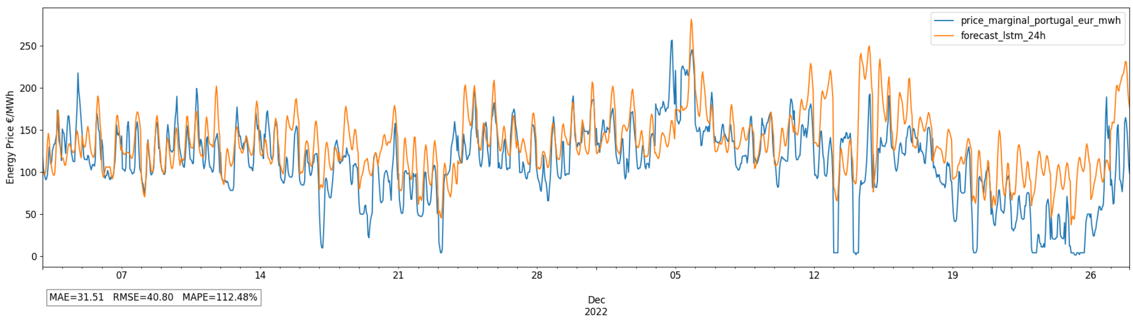


Figure 7.6: LSTM - 24 Hours Ahead Univariate Forecast

The Figure 7.7 presents the time series obtained as 1 hour ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period. For the presented results, the model was used in multivariate mode, therefore the energy price and the energy generation of the selected variables defined in Section 6.3 were used for training and as input data.

The Figure 7.8 presents the time series obtained as 24 hours ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period. For the presented results, the model was used in multivari-

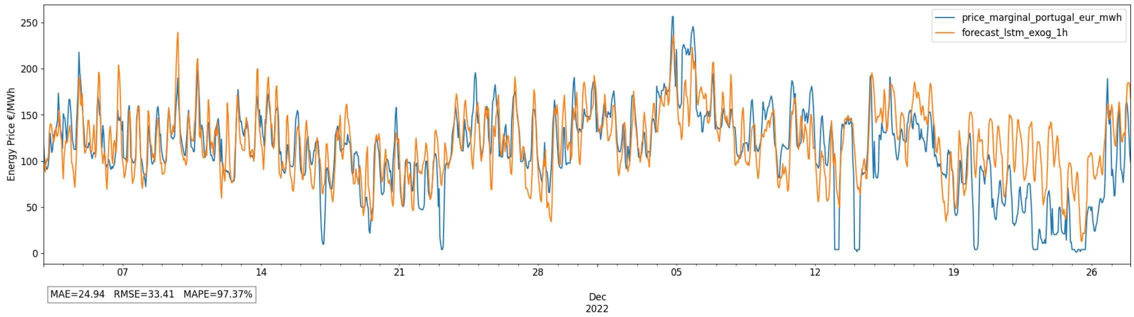


Figure 7.7: LSTM - 1 Hour Ahead Multivariate Forecast

ate mode, therefore the energy price and the energy generation of the selected variables defined in Section 6.3 were used for training and as input data.

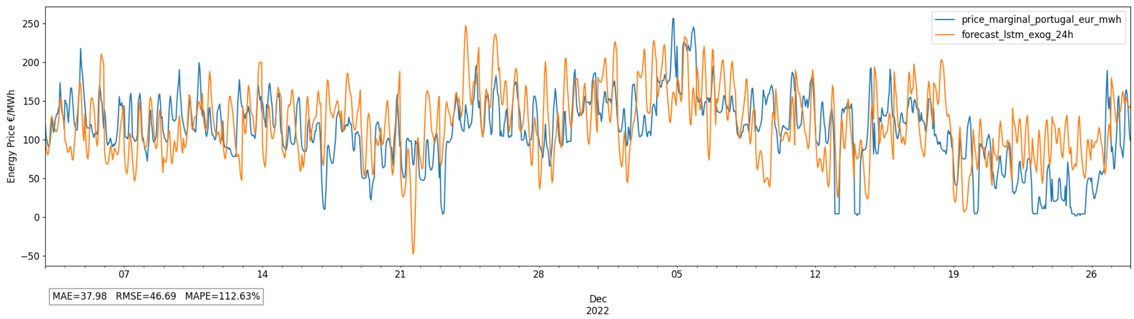


Figure 7.8: LSTM - 24 Hours Ahead Multivariate Forecast

7.1.4 GRU

The Gated Recurrent Unit Neural Networks (GRU) model was evaluated using the parameter setup defined in Section 6.2. The model was used in univariate and multivariate mode.

The Figure 7.9 presents the time series obtained as 1 hour ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period. For the presented results, the model was used in univariate mode, therefore only the energy price itself was used for training and as input data.

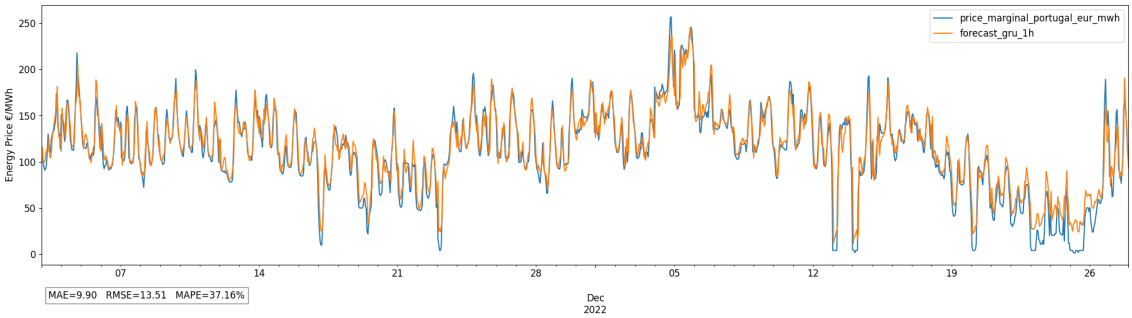


Figure 7.9: GRU - 1 Hour Ahead Univariate Forecast

The Figure 7.10 presents the time series obtained as 24 hours ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period. For the presented results, the model was used in univariate mode, therefore only the energy price itself was used for training and as input data.

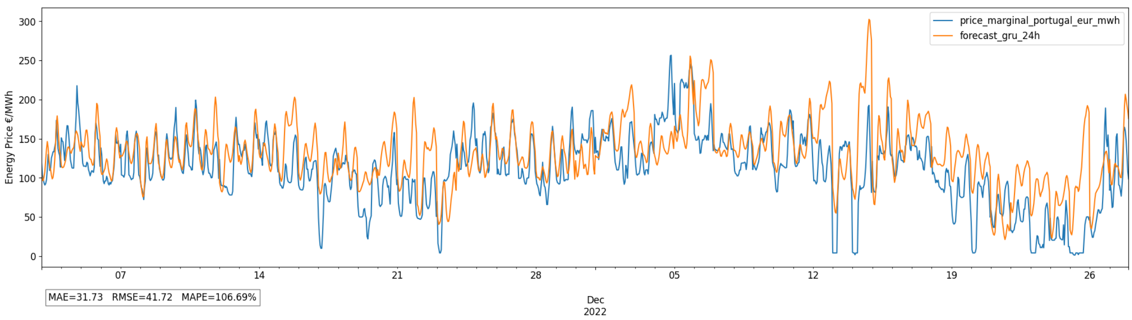


Figure 7.10: GRU - 24 Hours Ahead Univariate Forecast

The Figure 7.11 presents the time series obtained as 1 hour ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period. For the presented results, the model was used in multivariate mode, therefore the energy price and the energy generation of the selected variables defined in Section 6.3 were used for training and as input data.

The Figure 7.12 presents the time series obtained as 24 hours ahead forecast by the model in comparison with the real values verified for the Portuguese system in the same period. For the presented results, the model was used in multivari-

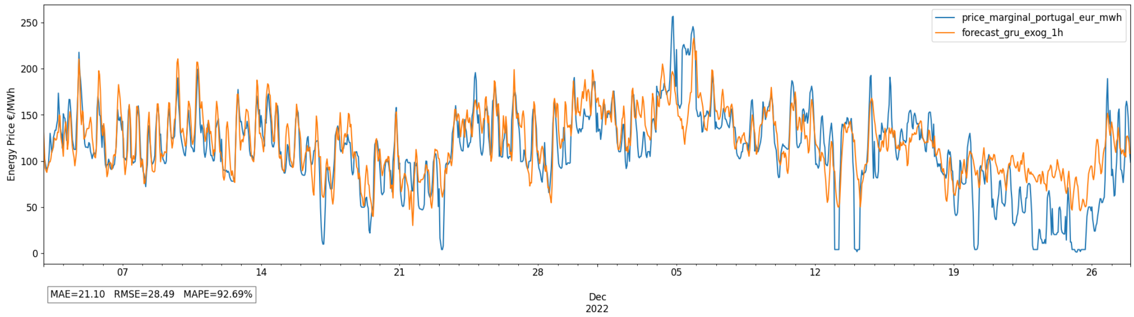


Figure 7.11: GRU - 1 Hour Ahead Multivariate Forecast

ate mode, therefore the energy price and the energy generation of the selected variables defined in Section 6.3 were used for training and as input data.

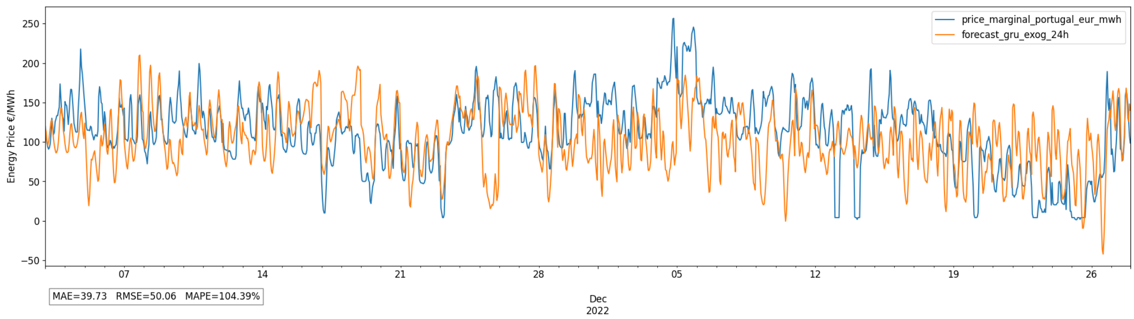


Figure 7.12: GRU - 24 Hours Ahead Multivariate Forecast

7.1.5 Model Results Analysis

The results presented for the various forecast models tested show great potential for their use in the energy price forecast function.

Prediction models and machine learning require very in-depth studies and refinement to achieve highly reliable results. As already mentioned throughout the body of this work, there are improvements to be made from the data selection and pre-processing to the final methodology that will be applied, since this goes beyond the adequate fitting of the model.

Several techniques already exists and others can be developed for models fitting that can prove to be effective in improving the accuracy and reliability of the final result. It is always possible, and even recurrent, that a single technique does not obtain adequate performance in all situations, therefore methodologies with varied techniques where each technique works only in the moments in which its performance is optimized, in general, bring great gain to the final result obtained.

This section presents the analysis of the results obtained in the analysis, seeking to understand the causes for the below-expected performances, in order to raise hypotheses for improvement. The Figure 7.13 shows the RMSE values and distribution of errors comparatively between the models.

The GRU model was removed from the comparison analysis of the results for being a variation of the LSTM model and having presented results similar to this one.

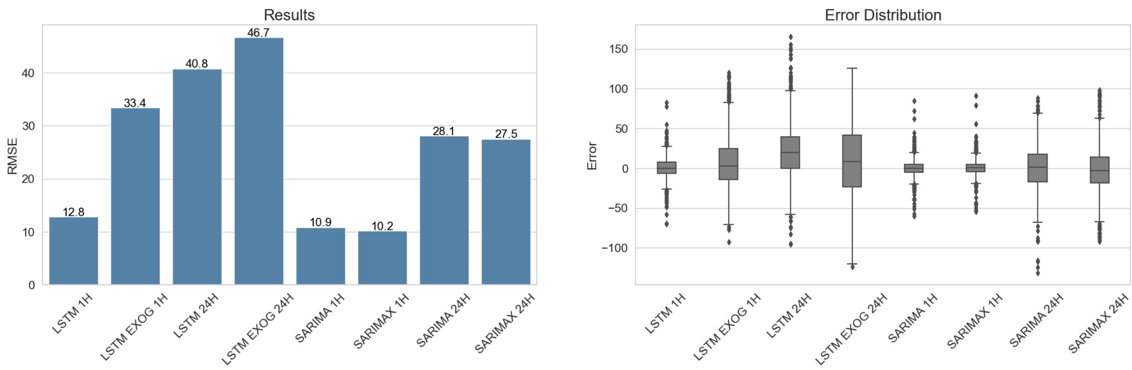


Figure 7.13: Energy Price Forecast Results

The models that presented the best performances for forecasting the energy price were "LSTM 1h", "SARIMA 1h" and "SARIMAX 1h" presenting RMSE around 10. This result shows that the models are able to capture the dynamics of the time series and produce good results for the very short term.

A lower performance for forecasts 24 hours ahead is expected since the models do not receive inputs indicative of abrupt changes in the behaviour of the time series that may happen throughout the day. This effect is evident if we observe the results of the "SARIMA 24h" in Figure 7.2 and "LSTM 24h" in Figure 7.6 models for the minimum values of the energy price observed on November 17th, 19th and 22nd and on December 13th, 14th and 20th.

In general, the models for 24 hours ahead have low accuracy in the local maximum and minimum points of the time series, this is most likely due to the influence of the minimum and maximum values obtained for the previous 24 hours, which are the last values received for training.

The superior performance of the SARIMA 24h models compared to the LSTM 24h model may indicate that the long-term influence is not being adequately captured by the model, but further analysis are needed to confirm this hypothesis.

Another important observation regarding the performance of the LSTM 24h model is the decoupling between the forecast curves and the actual value verified as of December 19th.

The month of December 2022 presents the lowest energy price values in the analyzed period, see Figure 6.1. In the time series used for training, there is no data for the months of December of previous years, which may be the main cause for this effect.

Good performances for LSTM models require a large volume of data that was not yet available during the development of this work. Therefore, using a longer time series for model training can greatly improve its performance.

The results of the multivariable models presented a substantially inferior performance to the univariate models for LSTM forecasts 1 hour ahead and for LSTM forecasts 24 hours ahead. This result was not expected considering the fact that the price of energy is a direct result of its composition among the generation sources available in the system, as can be observed in detail in the exploratory analysis available in Chapter 6.

A comparative analysis between the forecast results for the LSTM univariate 1-hour ahead model and the LSTM multivariate 1-hour ahead model can give a good indication of the cause of the worsening performance.

In the results presented for the multivariable LSTM 1h ahead model, Figure 7.7, it is possible to observe that the local maximum and minimum values are consistent with the moment in which they occur for real price values, however, they are overestimated or underestimated for the vast majority occurrences, which does not happen for the LSTM univariate forecasting model.

The presence of energy generation variables may condition the model to generate results similar to past moments that have the same energy generation composition. That is, compositions of similar generation sources generate similar price values.

The observed effect is correct behaviour and was even intentional when adding the exogenous variables. However, the final energy price value for each of the generation sources is not provided to the models and this may be the main cause for the observed effect.

The price of each energy source also varies over time, which means that the same amounts of energy generated from each source have a different final price and only the value of the final total price as an input to the models does not seem to be enough to differentiate them.

A possible approach to improve the performance of multivariable models would be the use of multi-univariate models. The univariate models proved to be effective at least in the very short term and it is plausible to expect the same behaviour for their individual use in each of the time series of sources of energy generation. Furthermore, the same can be done for time series of the prices of each of the energy sources, which are currently not available on the official data sources.

With the combination of precise forecasts of generation and price of each energy source, a more accurate result of the forecast of the final price of electric energy is also expected.

The SARIMAX 1h ahead model, in turn, performed slightly better than the SARIMA 1h ahead model, indicating that there are real possibilities for improving the results of the forecast models using the generation amounts of the different energy sources as exogenous variables.

The Kernel Density Estimate (KDE) plot in Figure 7.14 provides a clear visualization of the distribution of errors measured between the values predicted by each of the models and the target value.

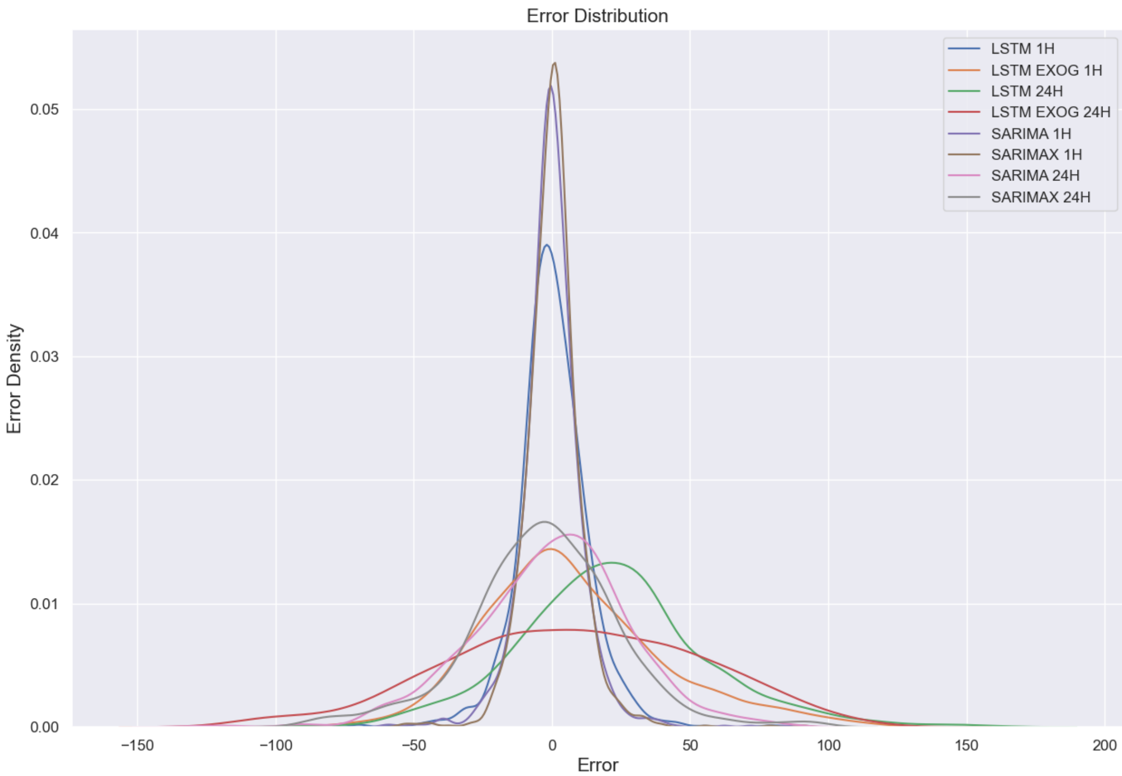


Figure 7.14: Models Error Density

7.2 5G Network Evaluation

This chapter is intended to evaluate the performance metrics of the 5G network and the response time obtained by the system to update the energy price for the simulation scenarios defined in Section 5.7.

The following scenarios have been defined:

- Scenario 1: 10 UEs - 10 Smart Meters per UE
- Scenario 2: 10 UEs - 20 Smart Meters per UE
- Scenario 3: 15 UEs - 10 Smart Meters per UE
- Scenario 4: 20 UEs - 10 Smart Meters per UE
- Scenario 5: 20 UEs - 20 Smart Meters per UE

7.2.1 General Performance

The 5G network performance metrics are obtained as output from network emulations performed by the SIMU5G simulator. The Packet losses in the network, throughput rate and delays in the RLC layer were evaluated. The Figure 7.15 presents the number of packets sent by the UEs to the MEC server.

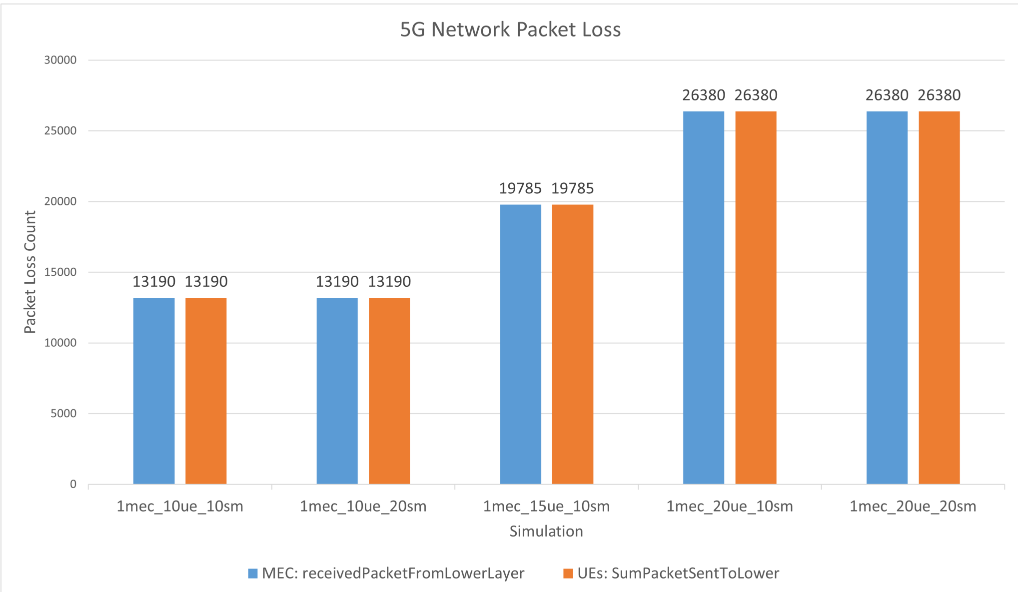


Figure 7.15: Packet Loss Count

No packet losses were identified for the simulated scenarios. The good performance of the simulations regarding packet loss is related to the low number of UEs due to host processing capacity restrictions. Simulations with a greater number of UEs and sending a greater number of packets are necessary to identify packet losses on the network.

The Figure 7.16 presents the Average User Equipments Throughput in terms of Downlink (Dl) and Uplink (Ul) measured in bps.

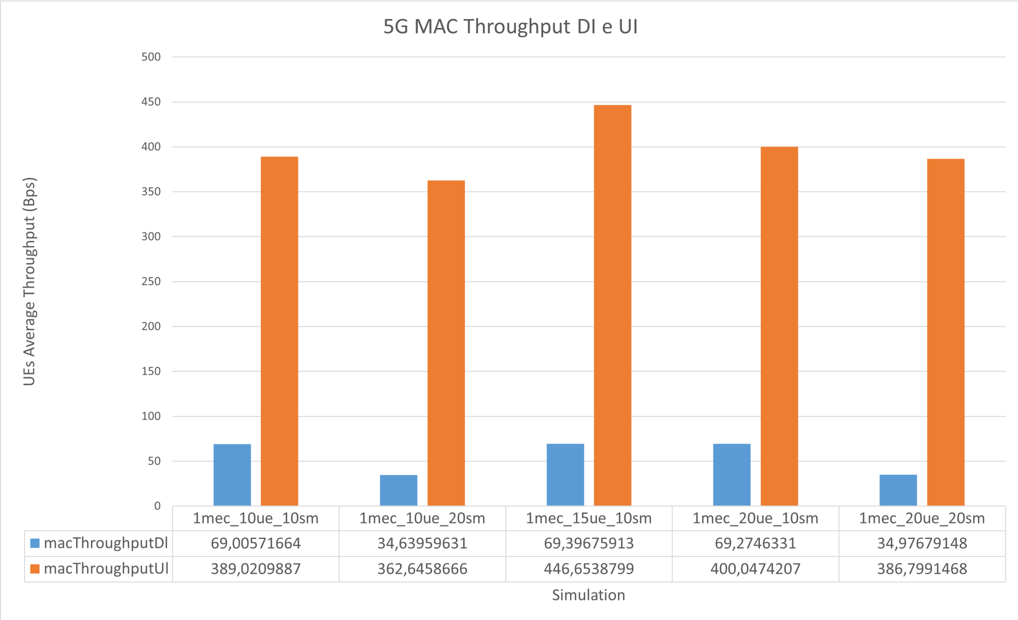


Figure 7.16: UEs Average Throughput Dl and Ul

The results obtained show a decrease in Throughput with the increase in the number of nodes in the network, in particular with the increase in the number of smart meters connected to each UE. The effect can be observed when comparing the results between simulations 10 UE/10SM and 10 UE/20SM and also between simulations 20UE/10SM and 20UE/20SM.

The Figure 7.17 shows the average delay value in the RLC layer of the UEs.

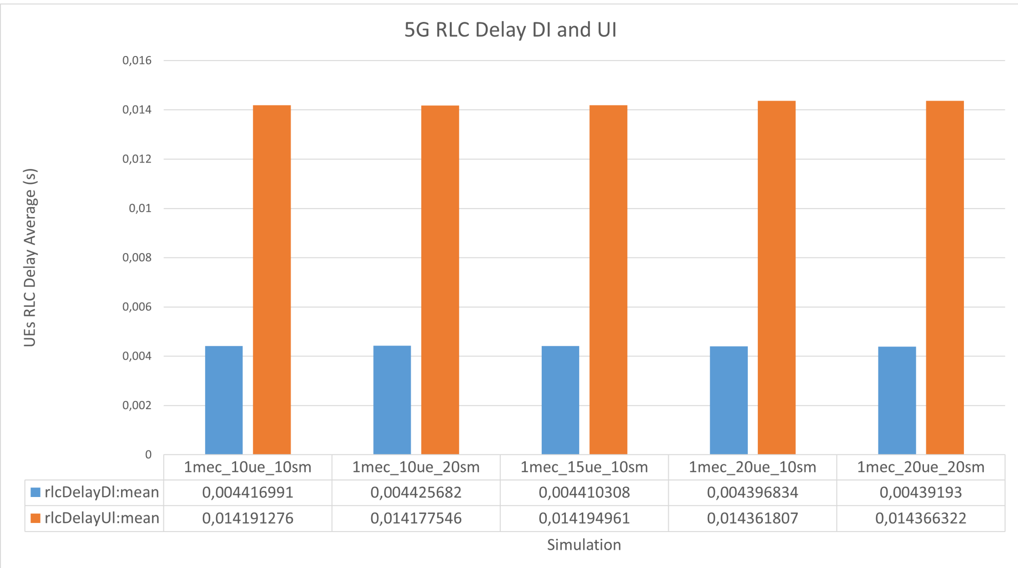


Figure 7.17: UEs Average Throughput Dl and Ul

The UE RLC layer Delay average values did not present significant variation between the simulated scenarios.

7.2.2 System Time Performance

This section presents the time performance results obtained by the energy price update system.

The objective of the simulations for time performance is to test the shortest time necessary for the system to update energy prices, therefore the delay time parameters between packets sent by the UEs were set to 0, as a consequence, the simulation scenario with 20 UEs and 20 Smart Meters per UE it was not feasible due to insufficient processing capacity of the simulation host.

The Figures 7.18,7.19,7.20 and 7.21 present histograms of packet latency observed in the transmission started at the UE and ended at the MEC server, for each simulation, respectively.

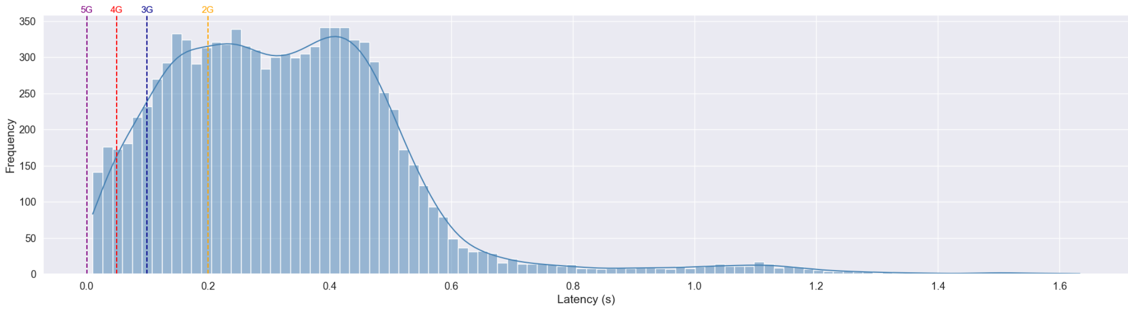


Figure 7.18: Packet Latency Histogram - 1MEC_10UE_10SM

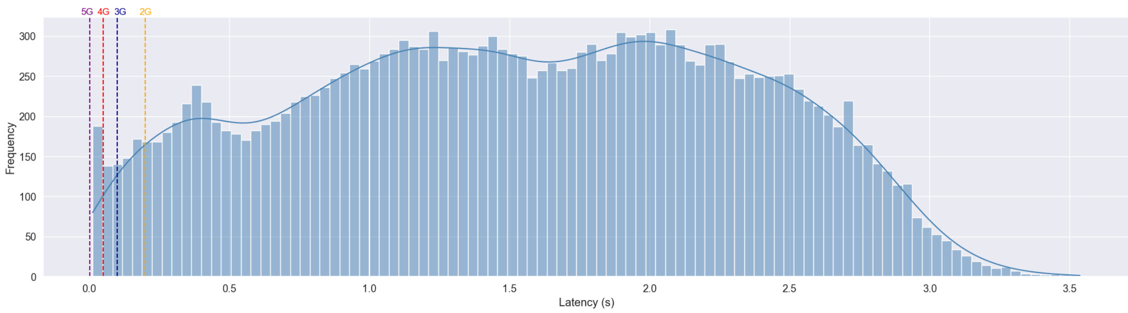


Figure 7.19: Packet Latency Histogram - 1MEC_10UE_20SM

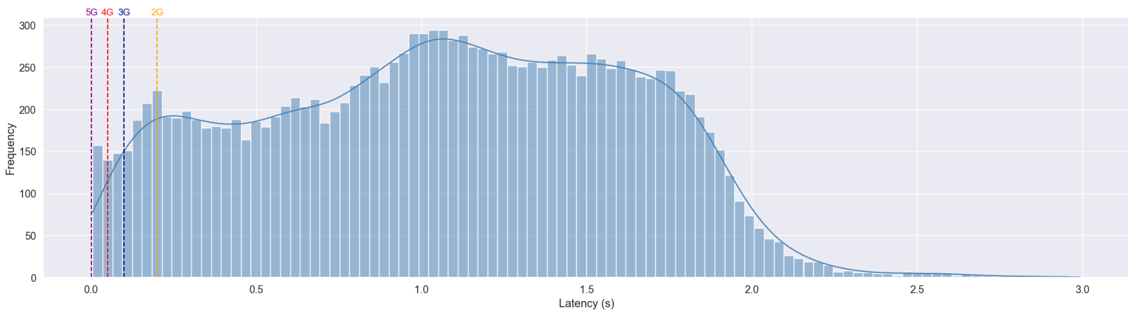


Figure 7.20: Packet Latency Histogram - 1MEC_15UE_10SM

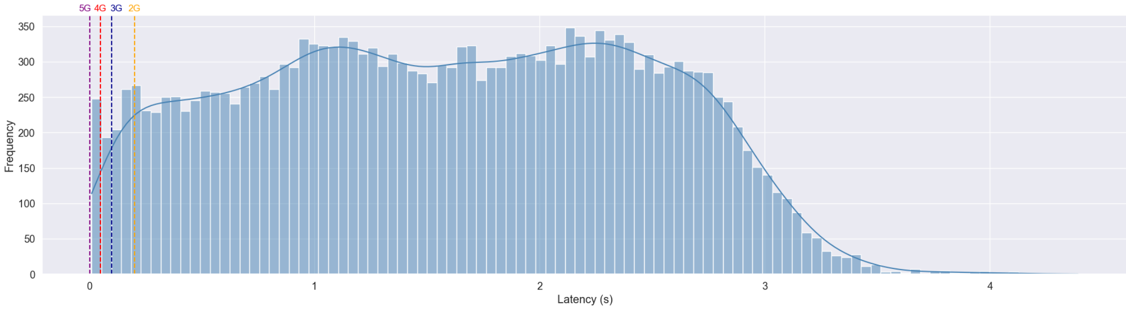


Figure 7.21: Packet Latency Histogram - 1MEC_20UE_10SM

The Figure 7.22 presents the Packet Latency KDE plots between simulations for comparison.

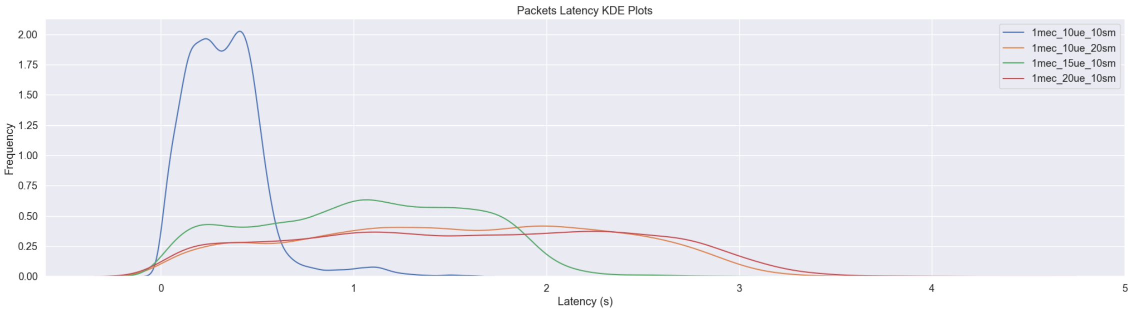


Figure 7.22: Packet Latency KDE Plots

The results show high latency for the vast majority of packets sent over the emulated 5G network. To achieve real-time emulation of 5G networks, the emulator must be able to process the events generated by the transmission of packets in a time shorter than the actual transmission latency. Scenarios with a number of UEs or SMs above 10 clearly present this effect in their distribution curves since they have a greater number of packets with high latency.

The Figure 7.23 shows the histogram of simulation latency results for 10 UE/10 SM focusing only on packets that had latency lower than 300 ms for further analysis.

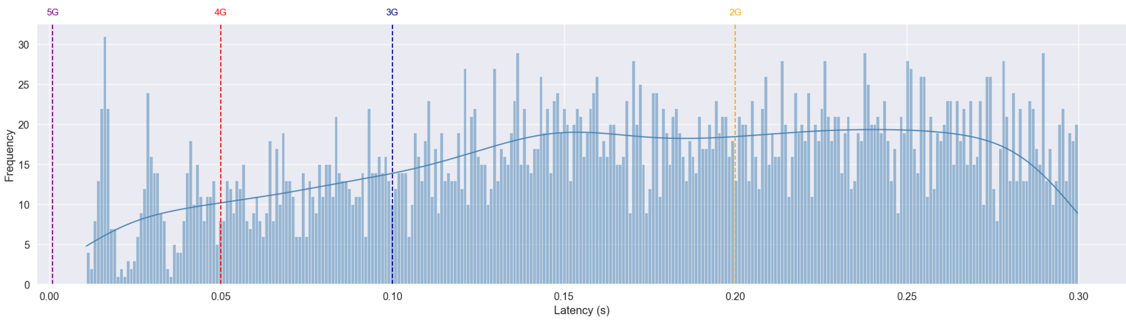


Figure 7.23: Packet Latency <0.3ms Histogram - 1MEC_10UE_10SM

It can be observed that some packets reached latency compatible with 4G transmission levels, but none of them achieved the expected performance for 5G.

Therefore, the system proved to be capable of performing simulations in high-speed communication networks as long as improvements in processing capacity are implemented.

The Figure 7.24 presents the total time histogram results for collecting data measured by smart meters and updating the energy price for each simulation.

The total time was obtained through the time elapsed between the sending of the first packet referring to a given timestamp and the arrival of the forecast results referring to the same timestamp, therefore corresponding to the delay in receiving the forecast for the next 24 hours in relation to the measurement time.

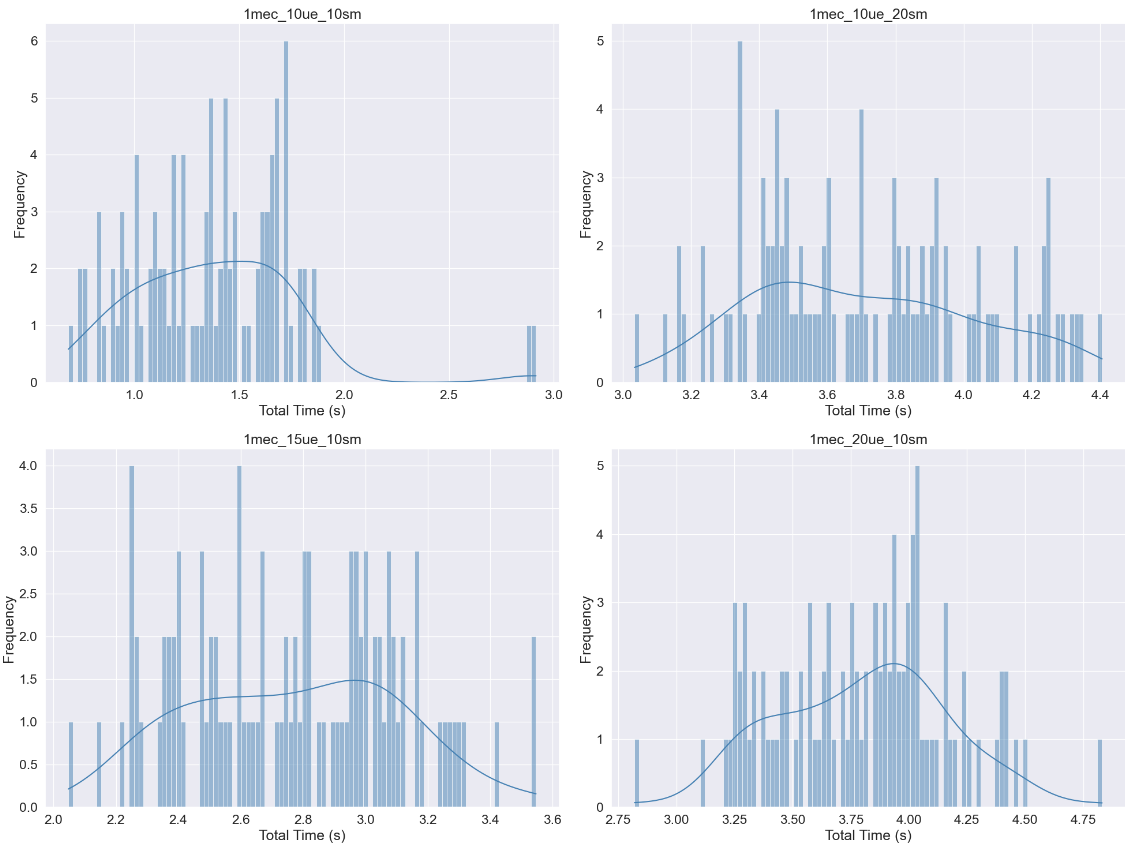


Figure 7.24: Total Time Histograms per Simulation

The Figure 7.25 compares the average total time and box plots for all timestamps evaluated between simulations.

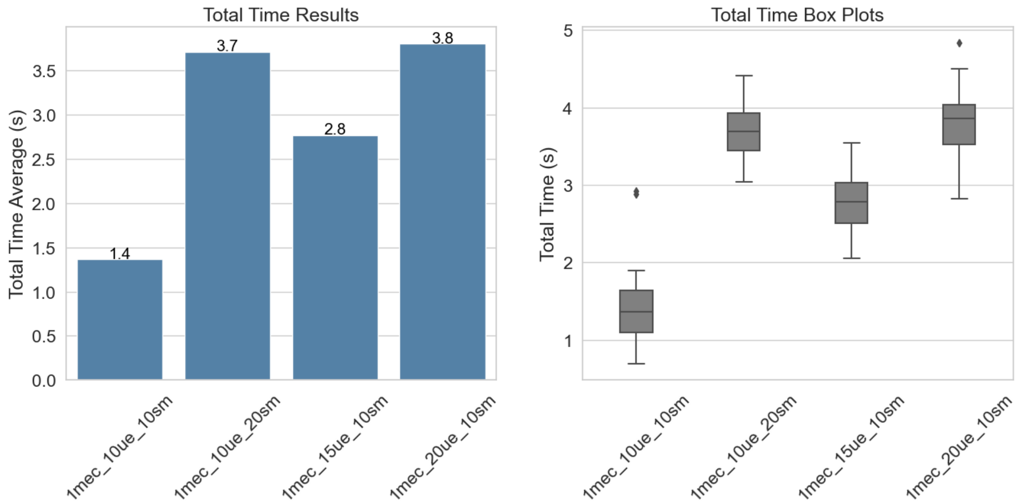


Figure 7.25: Total Time Simulations Average and Box Plots

The results of the total time to update the energy price verified for the implemented prototype system are insufficient to meet the needs for quick responses required by operational criteria in network operations described in Table 4.1. However, it shows that it is capable of updating the energy price every 4 seconds, which would be enough to bring great benefits in the context of the energy efficiency of smart homes.

Chapter 8

Conclusion and Future Works

This chapter includes conclusion and future works.

8.1 Conclusion

Smart grids represent a disruptive advancement in the field of electrical power distribution and management. There is too much to be explored regarding the innovation and automation possibilities that can be achieved with telecommunication networks with ultra-low latency, higher bandwidth, reliability and security.

Energy price information updated in real-time is just one of the many possibilities in terms of services and utilities. In fact, the price of energy in real-time will be a consequence of the evolution of smart grids in order to solve a real problem that already exists today.

Currently, within the scope of the operation of transmission and distribution networks, already exists the necessity to monitor the entire infrastructure of the system, in terms of generation, transmission and consumption. This work explores the idea of energy pricing service mainly aimed at the final consumer through its use as information to achieve greater energy efficiency in their homes and also in the sense of increasing the reliability of the electrical system since the increase in prices indicates the need to increase the supply of low cost energy in the system.

To achieve the objectives proposed in this work, some of the most well-regarded models for forecasting time series were evaluated. The results show great potential for predicting energy prices in the very short term.

The mathematical model requires improvement to achieve the specific purpose proposed. To this end, many methods have already proven to be effective and can be considered in improving energy price forecasting based on information about the energy being generated at the time the forecast is made.

This dynamic characteristic is essential to bring the desired reliability to electrical systems through the use of Smart Grids, as they are susceptible to unexpected events that require a quick response from the system. To this end, the 5G Networks are fundamental and becomes the technological advance that will effectively provide this capacity to electrical systems.

In order to evaluate the potential of this technology in smart grids, the work presented implements the architecture of a 5G network simulation system. The simulation prototype developed has the natural restrictions of any embryonic project, but it has already shown itself to be capable of reproducing the functioning of systems such as forecasting the price of energy in real-time.

The results obtained throughout this work show that the new generation radio networks can greatly reduce the times for measuring the necessary information to update prices, which currently can only be done at intervals of 1 hour.

Briefly, this work contributes with statistical and neural network-based models for the prediction of electricity prices in real-time, based on momentary energy generation and past energy prices, with specific analysis for the Portugal system as an additional contribution.

Furthermore, it is also an important contribution of this work, a structure for simulating communication networks in New Radio 5G capable of helping the development of important projects that will integrate the reality of smart networks in the coming years.

8.2 Future Works

Future work seeks to improve the real-time energy price prediction system through two main objectives: improving the performance of price prediction models and improving the processing capacity of system and simulation, as detailed below.

1. **Energy Price Forecast Models**

   The price forecast models evaluated in this study showed good results for very short-term forecasts, carried out for forecasts one hour ahead of the moment when measurements of the energy actually generated in the system are made. However, they still present average errors that can be very harmful for applications such as energy trading, which would bring significant financial losses to users.

   To this end, future work should consider more significant amounts of data for model training, in particular for models based on neural networks, where this characteristic is essential to achieve better performances.

   Furthermore, the reasons why the inclusion of the generation variables of the different system technologies did not contribute to more accurate model results, in most cases, must be thoroughly investigated. Changing price forecasts as a result of changes in generated energy is of fundamental importance so that smart grids can receive support from users in times of high demand and in cases of contingencies of important components for the operation of energy transmission networks.

To this end, more complex forecasting models that use different techniques for treating time series can be considered and select the best results among them. It must be considered that each energy generation technology has its own characteristics and the same method will hardly be ideal for all of them.

In summary, there is still a lot to explore regarding the methodology used to predict price values, from the data selection and processing stages and selection of features to the application of the methods.

2. **Energy Price System Capacity**

The performance of the price prediction system in 5G networks clearly presented processing limitations. The emulation processing of the 5G network must be increased so that it is capable of processing events related to the sending of packets in a time lower than the desired latency, which in the case of 5G networks can be quite challenging.

However, the simulation experiments in this work use a single host, with separate network components through isolated Linux operating system environments (Linux namespaces). Therefore, it is believed that the separation of the various network components into physically separate hosts with superior processing capabilities and dedicated to the network emulation and to the MEC server can bring great performance improvements to the simulations.

The correction measures would likely provide the desired latency performance for 5G networks and, as an additional benefit, it would also allow the simulation of broader networks, with greater amounts of connected UE devices, more Smart Meters and, possibly, additional MEC servers.

References

[1] J. p. s. catalão, h. m. i. pousinho and v. m. f. mendes, "hybrid wavelet-pso-anfis approach for short-term wind power forecasting in portugal," in ieee transactions on sustainable energy, vol. 2, no. 1, pp. 50-59, jan. 2011.10.1109/tste.2010.2076359. URL https://ieeexplore.ieee.org/document/5571038.

[2] 5g empowers the future of electricity. URL https://www2.deloitte.com/cn/en/pages/energy-and-resources/articles/5g-empowers-smart-power-and-promotes-innovation.html.

[3] Dlms user association - dlms/cosem architecture and protocols green book 10.

[4] Fensel, a.; gómez berbís, j.m. energy efficiency in smart homes and smart grids. energies 2021, 14, 2054. URL https://doi.org/10.3390/en14082054.

[5] Ghiasi, m.; niknam, t.; dehghani, m.; siano, p.; haes alhelou, h.; al-hinai, a. optimal multi-operation energy management in smart microgrids in the presence of ress based on multi-objective improved de algorithm: Cost-emission based optimization. appl. sci. 2021, 11, 3661. https://doi.org/10.3390/app11083661. URL https://www.mdpi.com/2076-3417/11/8/3661.

[6] Inet framework. URL https://inet.omnetpp.org/.

[7] Jan, f.; shah, i.; ali, s. short-term electricity prices forecasting using functional time series analysis. energies 2022, 15, 3423. URL https://doi.org/10.3390/en15093423.

[8] Monteiro, c.; ramirez-rosado, i.j.; fernandez-jimenez, l.a.; conde, p. short-term price forecasting models based on artificial neural networks for intraday sessions in the iberian electricity market. energies 2016, 9, 721. URL https://doi.org/10.3390/en9090721.

[9] Nardini, g.; stea, g.; virdis, a. and sabella, d. (2020). simu5g: A system-level simulator for 5g networks. in proceedings of the 10th international conference on simulation and modeling methodologies, technologies and applications - simultech; isbn 978-989-758-444-2; issn 2184-2841, scitepress, pages 68-80. doi: 10.5220/0009826400680080. . URL https://www.scitepress.org/Link.aspx?doi=10.5220/0009826400680080.

[10] G. nardini, g. stea and a. virdis, "scalable real-time emulation of 5g networks with simu5g," in ieee access, vol. 9, pp. 148504-148520, 2021, doi: 10.1109/access.2021.3123873. . URL https://ieeexplore.ieee.org/document/9591605.

[11] Omnet++ library. URL https://omnetpp.org/.

[12] Simu5g. URL http://simu5g.org/.

[13] Viviani, e.; di persio, l.; ehrhardt, m. energy markets forecasting. from inferential statistics to machine learning: The german case. energies 2021, 14, 364. URL https://doi.org/10.3390/en14020364.

[14] Entidade reguladora dos serviços energéticos. URL https://www.erse.pt/en/home/.

[15] Electricity market operator. URL https://www.omie.es/pt.

[16] Redes energéticas nacionais. URL https://www.ren.pt/.

[17] *Box, George E. P. Time series analysis : forecasting and control. – Fifth edition / George E.P. Box, Gwilym M. Jenkins,Gregory C. Reinsel, Greta M. Ljung.* 2016.

[18] Teixeira A. D. César. *Time Series Analysis and Forecasting*. 2021.

[19] Luciano De Castro and Joisa Dutra. Paying for the smart grid. URL https://www.sciencedirect.com/science/article/pii/S0140988313002119.

[20] Paula Ferreira, Madalena Araújo, and M.E.J. O'Kelly. An overview of the portuguese electricity market. URL https://www.sciencedirect.com/science/article/pii/S030142150600262X.

[21] Vimukthi Gamage, Nishan Withana, Channa Silva, and Rasara Samarasinghe. Battery energy storage based approach for grid voltage regulation in renewable rich distribution networks. In *2020 2nd IEEE International Conference on Industrial Electronics for Sustainable Energy Systems (IESES)*, 2020.

[22] Campos Inês, Pontes Luz Guilherme, Marín-González Esther, Gährs Swantje, Hall Stephen, and Holstenkamp Lars. Regulatory challenges and opportunities for collective renewable energy prosumers in the eu. URL https://www.sciencedirect.com/science/article/pii/S0301421519307943.

[23] Paul Krugman and Robin Wells. Economics. In *3th Edition*.

[24] Zhewei Liu, Wenzhong Shi, Yue Yu, Pengfei Chen, and Bi Yu Chen. A lstm-based approach for modelling the movement uncertainty of indoor trajectories with mobile sensing data. *International Journal of Applied Earth Observation and Geoinformation*, 108:102758, 2022. ISSN 1569-8432. doi: https://doi.org/10.1016/j.jag.2022.102758. URL https://www.sciencedirect.com/science/article/pii/S0303243422000848.

[25] INTERDISCIPLINARY MIT STUDY. *The Future of Energy Storage*. 2022. URL https://energy.mit.edu/wp-content/uploads/2022/05/The-Future-of-Energy-Storage.pdf.

[26] Bartosz Tarakan, Stanislaw Czapp, Krzysztof Dobrzynski, Ryszard Zajczyk, and Marcin Sarnicki. Voltage control in a power system with renewable sources of energy. In *2019 International Conference on Information and Digital Technologies (IDT)*, 2019.

[27] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola. *Dive into Deep Learning*. Cambridge University Press, 2023.

Appendices

Appendix A

A.1 UE Routing File

ifconfig:
 # interface to the external client
name: eth0
inet_addr: 192.168.3.1
Mask: 255.255.255.0
MTU: 1500
Metric: 1
POINTTOPOINT MULTICAST

name: cellular
inet_addr: 10.0.0.1
Mask: 255.255.255.0
MTU: 1500
POINTTOPOINT MULTICAST

ifconfigend.
route:
#Destination Gateway Genmask Flags Metric Iface
*192.168.2.0 * 255.255.255.0 H 0 cellular*
*192.168.3.0 * 255.255.255.0 H 0 eth0*
*0.0.0.0 * 0.0.0.0 G 0 cellular*
routeend.

A.2 gNodeB Routing File

ifconfig:

 # interface to the radio access network
name: cellular
inet_addr: 10.0.0.100
Mask: 255.255.255.0
MTU: 1500
POINTTOPOINT MULTICAST

name: pppIf
inet_addr: 10.0.1.2
Mask: 255.255.255.0
MTU: 1500
POINTTOPOINT MULTICAST
ifconfigend.

route:
#Destination Gateway Genmask Flags Metric Iface
*10.0.0.0 * 255.255.255.0 H 0 cellular*
*10.0.1.0 * 255.255.255.0 H 0 pppIf*
*10.0.2.0 * 255.255.255.0 H 0 pppIf*
*192.168.2.0 * 255.255.255.0 H 0 pppIf*
*192.168.3.0 * 255.255.255.0 H 0 cellular*
*0.0.0.0 * 0.0.0.0 G 0 pppIf*

routeend.

A.3 UPF Routing File

ifconfig:

 # interface 0 to the router
name: pppIf
inet_addr: 10.0.2.2
Mask: 255.255.255.0
MTU: 1500
POINTTOPOINT MULTICAST

 # interface 0 to the gnb
name: ppp0
inet_addr: 10.0.1.1
Mask: 255.255.255.0
MTU: 1500
POINTTOPOINT MULTICAST
ifconfigend.

route:

#Destination Gateway Genmask Flags Metric Iface
*10.0.1.0 * 255.255.255.0 H 0 ppp0*
*10.0.0.0 * 255.255.255.0 H 0 ppp0*
*10.0.2.0 * 255.255.255.0 H 0 pppIf*
*192.168.2.0 * 255.255.255.0 H 0 pppIf*
*192.168.3.0 * 255.255.255.0 H 0 ppp0*
*0.0.0.0 * 0.0.0.0 G 0 pppIf*

routeend.

A.4 Router Routing File

ifconfig:

 # interface to the external server
name: eth0
inet_addr: 192.168.2.1
Mask: 255.255.255.0
MTU: 1500
Metric: 1
POINTTOPOINT MULTICAST

 # interface to the nat router
name: ppp0
inet_addr: 10.0.2.1
Mask: 255.255.255.0
MTU: 1500
POINTTOPOINT MULTICAST
ifconfigend.

route:

#Destination Gateway Genmask Flags Metric Iface
*192.168.2.0 * 255.255.255.0 H 0 eth0*
*192.168.3.0 * 255.255.255.0 H 0 ppp0*
*0.0.0.0 * 0.0.0.0 G 0 ppp0*

routeend.