Comparative Analysis of Predictive Models for Subway Rolling Stock Energy Forecasting

Mircea Stefan Simoiu
Automation and Industrial Informatics
University "Politehnica" of Bucharest
Bucharest, Romania
mircea_stefan.simoiu@upb.ro

Grigore Stamatescu
Automation and Industrial Informatics
University "Politehnica" of Bucharest
Bucharest, Romania
grigore.stamatescu@upb.ro

Vasile Calofir
Automation and Industrial Informatics
University "Politehnica" of Bucharest
Bucharest, Romania
vasile.calofir@upb.ro

Ioana Fagarasan
Automation and Industrial Informatics
University "Politehnica" of Bucharest
Bucharest, Romania
ioana.fagarasan@upb.ro

Sergiu Stelian Iliescu
Automation and Industrial Informatics
University "Politehnica" of Bucharest
Bucharest, Romania
stelian.iliescu@upb.ro

***Abstract*—Devising robust energy management strategies for sustainable future urban transportation systems requires in depth understanding of the underlying generative data patterns. Once quality datasets are collected and pre-processed, various computational intelligence technniques can be applied to extract actionable forecasts in order to improve the overall system efficiency. We present a data science methodology for subway rolling stock energy forecasting based on real collected data from traction power transformers of two subway stations in the city of Bucharest, Romania. The effect of variable passenger loads is also investigated in a two-stage forecasting process involving exponential smoothing filtering and regressive forecasting. The main forecasting results yield a MAPE indicator under 0.2% for both evaluated scenarios, with an improvement achieved by integrating the passenger traffic as additional contextual explanatory variable for the model.**

***Index Terms*—energy forecasting, predictive models, exponential smoothing, machine learning, rolling stock, subway station**

I. Introduction

A key pillar of the Green Deal [1] strategy is the investment in energy efficient transportation systems. The role of light urban rail and subway transportation systems is becoming essential for lowering carbon emissions and traffic decongestion which directly impacts public health and quality of life. Reliable energy management strategies are thus required to quantify and control the power consumption of such systems in close correlation to the actual usage thereby increasing overall efficiency. Achieving a thorough understanding of the factors influencing subway transportation energy consumption is also highly relevant in order to suitably dimension renewable and sustainable energy supplies within localized energy microgrids.

Energy forecasting is a well established field which enables robust control of energy systems through good quality estimation of the future system variables and state. Two main concurrent approaches rely on either statistical time series models such as exponential smoothing (ES) and autoregressive, moving average, models such as ARMA/ARIMA or machine learning algorithms. In recent years a turning point has been achieved in time series forecasting strategies as, given the wider availability of rich, densely sampled datasets and increased computational power, new machine learning models have become competitive for out-of-sample/training set forecasting accuracy. This includes both highly optimized methods as linear regression, regression trees and support vector regression and deep neural network models as universal approximators for complex non-linear variable dependencies.

Alongside a typical time series autoregressive forecasting strategy, we study the passenger load effect on subway rolling stock energy consumption for energy management and forecasting purposes. As input data we use energy measurements from one subway station in the city of Bucharest, Romania. The measurements correspond to the traction power electrical subsystem of the station and are largely determined by the power consumption of the trains upon leaving the station, in accordance with predefined daily schedules. The results are replicable and extendable with minimal adjustments to other subway stations in the same network for similar types of rolling stock.

The contributions of this paper are argued to be two-fold:

- problem formulation and analysis of subway rolling stock energy prediction based on a two-stage methodology that combines exponential smoothing as input data filter with machine learning algorithms;
- investigation of contextual explanatory variables, such as time-of-day varying passenger loads on the prediction performance.

The rest of the paper is structured as follows. Section II presents a context of related work for both subway system energy modelling and energy forecasting as a current relevant topic for the scientific community. We present a detailed methodology for data preparation, feature engineering, model development and evaluation in Section III. Section IV illustrates and discusses the results obtained based on classical performance metrics such as MSE/RMSE and MAPE. We conclude the paper with the main lessons learned during this study and outlook on future work in Section V.

II. Related Work

Recent years proved to be very demanding in many sectors under the recent renewable energy trends, especially in research areas where a thorough need for new solutions has been repeatedly emphasised. Thus, researchers and government officials alike have turned their attention to important public domains where potential for energy efficient solutions could present a broad social, environmental and economic impact. The public underground transportation sector, for instance, was already regarded as an important power consumer [2] and researchers have already started to investigate and identify sustainable energy management solutions.

A natural step in this direction was to develop an in-depth analysis over the structure of a subway station and its power consumers and subsystems [3]. Such a study provided relevant information regarding the power distribution inside a station and consequently identified the traction power subsystem as the most important consumer. Moreover, this analysis paved the way for studies where both infrastructure related solutions (such as renewable energy generation and energy metering, for example [4]), as well as operational solutions (such as energy-efficient train driving) were analysed, along with the efficiency impact and implementation feasibility [5]. It was clear that more research was needed in this direction to explore these opportunities and analyse possible implementations.

From another point of view, studies involving modeling and simulation instruments also provide important insights and discussions [6] [7] [8], especially since they involve modeling techniques, possible microgrid architectures and suitable control solutions. However, to accurately assess if one such solution is suitable for implementation, a thorough load analysis must be conducted. This implies the usage of energy forecasting techniques, machine learning instruments and time series analysis that already have proved to be useful in other energy sectors.

Deep learning methods for energy forecasting are described in [9]. These are compared with classical methods and the trade-off between model explainability with feature engineering and increased accuracy through sequence models is discussed. Several open source libraries have currently become available which allow testing of such approaches at scale given availability of quality datasets and computational resources. Local energy forecasting models for *in situ* control are argumented by [10] where local inference to predict and classify future anomalous power consumption patterns is discussed. Cyber physical energy systems represent a sub field of cyberphysical systems for bridging. computation, communication and control, while relying on intelligent data processing algorithms with applications in the energy domain [11]. Modern neural networks for energy forecasting are evaluated by [12] from the transmission system operator (TSO) perspective in which large consumer, such as urban rail and subway system can play an important role for reliable energy management strategies such as load shaping or load shifting. In a similar way, several first steps were taken in estimating the consumption of a subway station with neural networks [13], however simpler and more robust approaches must be further investigated by taking into account the difficulties of the data-gathering process for such a complex infrastructure point.

III. Methodology

A. Load Forecast Proposed Procedure

To forecast the power consumption of the rolling stock based on *a priori* measured load data, a suitable methodology is proposed in Figure 1.

The first step is to analyse the available data in terms of sample rate, date, time, also with a focus on possible missing data and redundant parameters. The result expected here is a data series based on a date-time composed index.

The second step is to transform the available data by adding features for the model implementation. Feature variables represent attributes that describe a particular data point from a specific point of view. For example, common time-related features represent the day, hour or minute the measurement was taken. Moreover, it is important to adapt the features to a suitable format required by the prediction models.

The third step consists in applying an Exponential Smoothing filter [14] in order to remove unnecessary time periods (for example night-time consumption) and to obtain an adequate profile for further forecasting. The new power profile is given by the following Exponential Smoothing model:

$$\hat{y}_0 = y_0 \tag{1a}$$

$$\hat{y}_t = \alpha \cdot y_t + (1 - \alpha) \cdot \hat{y}_{t-1} \tag{1b}$$

where $\hat{y}_t$ is the estimated value at t, y represents the raw measured value and α represents a smoothing factor between 0 and 1.

The fourth step focuses on a randomized split of the new data into two sets:

- the training data set - used to determine the prediction model parameters
- the test data set - used for model and metrics evaluation.

To better illustrate this step, the independent variable is considered to be X and the dependent variable will be addressed as Y. In typical forecasting problems, the training and testing data sets are determined with respect to the length of the initial data set. More specifically:

$$dim(X_{train}) = n_{test}^{size} \cdot dim(X) \tag{2a}$$

$$dim(X_{test}) = n_{train}^{size} \cdot dim(X) \tag{2b}$$

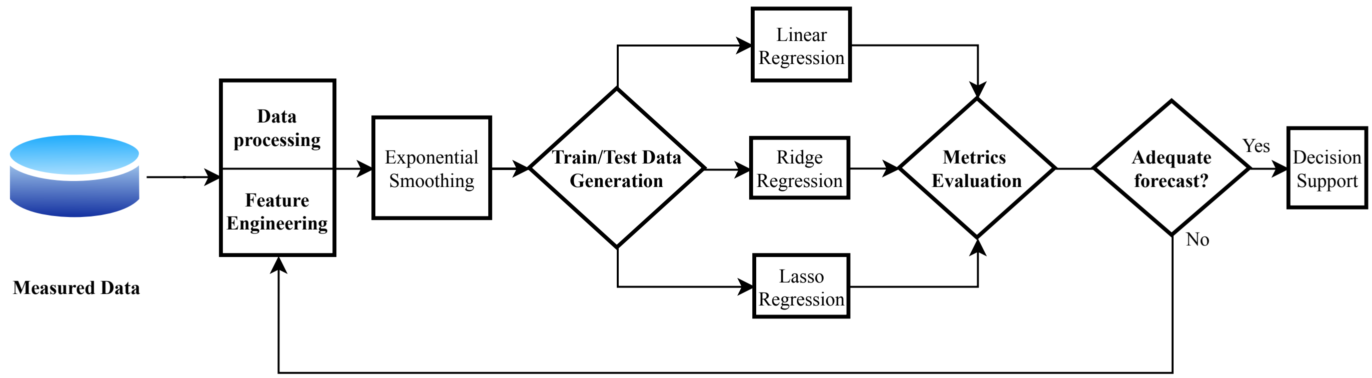

Fig. 1. Proposed procedure

where X_{train} represents the train data set and X_{test} represents the test data set. n_{test}^{size} and n_{train}^{size} represent the desired size of the train and test data sets.

To determine the best load forecast, the proposed procedure analyses three model types: linear regression, ridge regression and lasso regression.

1) Linear Regression: Linear regression models are developed based on the hypothesis that there is a linear dependence between the independent variables $X = (x_1, x_2, ..., x_n)^T$ representing the model features and the dependent variable Y. Consequently, the estimation $\hat{Y}$ can be determined using the following model:

$$\hat{Y} = a + B \cdot X \tag{3}$$

where a represents the intercept and B is the model slope.

The best estimation is achieved through minimisation of residual points representing the difference between the real observed values vector $Y = (y_1, y_2, ...y_n)^T$ and the estimation vector $\hat{Y} = (y_1, y_2, ...y_n)^T$ [15]. Thus, the objective function can be defined in vector form:

$$E = \sum_{i}^{n} (y_i - a - \sum_{j}^{p} x_{ij} b_j)^2 \tag{4}$$

There are also variations of the typical linear regression model that imply a penalty over the model coefficients. These variations are known in specialised literature as the Ridge Regression and Lasso Regression [15].

The Ridge method minimises the residuals function considering a regularization parameter λ, such as Eq. 4 becomes:

$$E^{ridge} = \sum_{i}^{n} (y_i - a - \sum_{j}^{p} x_{ij} b_j)^2 + \lambda \sum_{j}^{p} b_j^2 \tag{5}$$

The idea behind the Ridge Regression is to converge the coefficients towards zero. A similar form that aims to penalize the coefficients is represented by the Lasso Regression Model, where the residuals formula becomes Eq. 6.

$$E^{lasso} = \sum_{i}^{n} (y_i - a - \sum_{j}^{p} x_{ij} b_j)^2 + \lambda \sum_{j}^{p} |b_j| \tag{6}$$

The proposed methodology implements all three regression models and a discussion is made upon the most important evaluation metrics and whether a regression model is best suited for rolling stock load forecasting. Standard metrics are used for evaluating the regression models: Mean Squared Error (MSE) - sum of squared bias and variance, Root Mean Squared Error (RMSE) and the Mean Absolute Percentage Error (MAPE) as relative performance indication. The metrics are expressed as follows [16]:

$$MSE = \sum_{1}^{n} \frac{(y_i - \hat{y}_i)^2}{n} \tag{7}$$

$$RMSE = \sqrt{MSE} \tag{8}$$

$$MAPE = \frac{1}{n} \sum_{1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| 100 \tag{9}$$

y_i is the actual power value, $\hat{y}_i$ is the output prediction at time sample i and n is the sample size of the test set.

IV. Results

A. Subway Station Power Flow Analysis

Figure 2 illustrates the active power flow through a typical subway station from Bucharest, Romania. Power is provided from the grid through one main power supply and multiple power transformers are converting the energy to suitable voltage levels.

Inside the station, power systems can be grouped in two categories: traction subsystems and services subsystems. The traction power is used specifically to power the rolling stock, while the services power system have multiple purposes: to power the lightning system, the automatic ventilation system and all other auxiliary systems integrated in the station. These two categories are separated since the rolling stock is powered

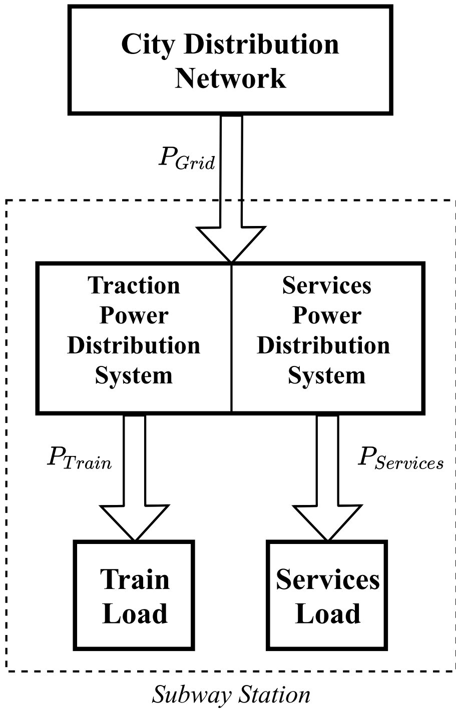

Fig. 2. Typical Subway Station Power Flow Diagram

at 850V DC voltage and services are power in 400V 3 phase voltage.

Another particular aspect regarding the subway station is that each power subsystem has at least one auxiliary power supply connected in case a system malfunction is present or power is cut off. Also, in case of major system failures, energy may be obtained from nearby subway stations thorough the city distribution grid energy feeder management.

Analyzing the subway station power flow, the balance equation can be written as:

$$P_{Grid} = P_{Train} + P_{Services} \tag{10}$$

where the three variables represent the respective active power for each power system. Since the main consumer of a subway station is the traction power subsystem, the paper will further focus on implementing the proposed methodology for the load forecast of the traction side.

B. Measured data

Available data has been obtained through measurements during a full working day, specifically on the power transformers responsible with converting 3 phase medium voltage power from the distribution grid to the 850V DC power for the rolling stock. The measuring device is a Fluke 434 - II Energy Analyzer and the sampling rate has been chosen at 30 seconds.

It can be noticed that the electric trains are functioning from 05:00 to 23:00 and two power spikes can be identified: one in the morning, between 08:00 and 09:00 and one in the afternoon between 17:00 and 18:00. Since there is no electric train passing between 23:30 and 04:30 (maintenance work is conducted on the rail so power is decoupled), it is clear that there is no electric load during that time. However, energy consumption increases in the morning and in the afternoon mainly because these are the intervals during which employees commute and passenger traffic has an important effect upon train energy consumption. This is a first important insight that will contribute to the forecast model development.

Also, passenger traffic in the station has been estimated from hourly data acquired from the station access points during a working day from the same month.

C. Data processing stage

In order to improve the forecasting model accuracy, data must be specifically adapted through a noise-filtering stage. Multiple variants of the exponential smoothing have been analyzed, as it can be observed in Fig. 4.

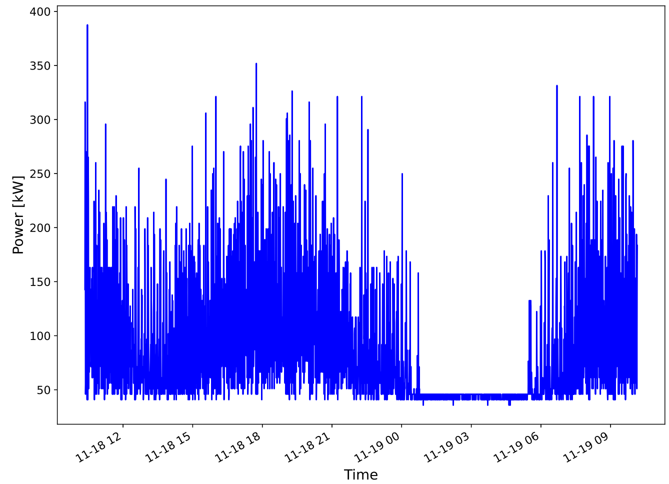

Fig. 3. Measured traction load profile - 1 working day

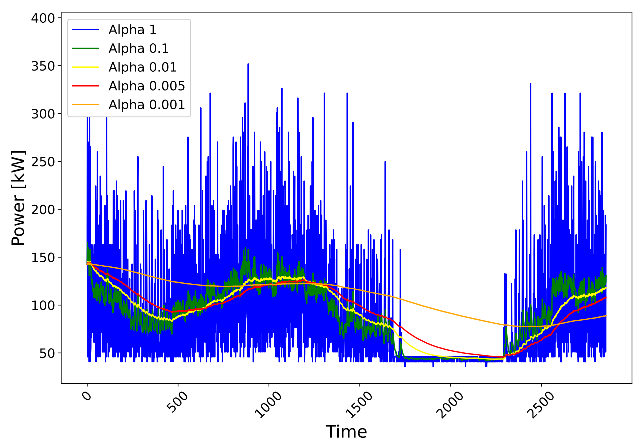

Fig. 4. Exponential Smoothing Comparison

As expected, decreasing the value of α in the filtering process visibly increases data smoothness and removes unwanted noise. However, a value of α that is too low may also remove relevant features of the data profile such as the increase in energy consumption during peak traffic hours. Also, it can be noticed that the exponential smoothing filtering also removes unnecessary consumption data acquired during the night.

For future model development, we considered that an α value of 0.005 offers a good balance between filtering and feature retain.

As for the feature engineering process, in order to correspondingly assess the passenger influence in load forecast, two scenarios are to be analysed: a scenario in which forecasting models are developed based only on previous measured samples and date/time related information and a scenario that additionally includes passenger related traffic.

1) Scenario 1 - lag and date/time features: In this scenario, the independent variable X has the features described according to Table I.

TABLE I
FEATURE SELECTION FOR THE FIRST SCENARIO

Feature	Description
y_{t-1}	Consumption at $t-1$
y_{t-2}	Consumption at $t-2$
hour	hour in day
time_of_day	Encoded between: morning, noon, afternoon, evening

Considering these features, data has been splitted between a training set and a testing set, with $n_{train}^{size} = 0.3$ and $n_{test}^{size} = 0.7$, according to Eq. 2.

Consequently, linear regression models have been trained on data. A comparison has been further established between metrics (Table II), as well as between the predicted load evolution and measured data from the testing set (Fig. 5).

It can be observed that, overall, linear regression models offer very good performances in terms of MAPE, MSE and RMSE, and also the forecast profiles are very similar in evolution.

TABLE II
METRICS EVALUATION ON SCENARIO 1

Model	Evaluation Metric	Value
Scaled Linear Regression	MAPE	0.192 %
	MSE	65921.13 W^2
	RMSE	256.75 W
Ridge Regression	MAPE	0.211 %
	MSE	72255.51 W^2
	RMSE	268.80 W
Lasso Regression	MAPE	0.209 %
	MSE	68432.46 W^2
	RMSE	261.60 W

2) Scenario 2 - lag and date/time features and passenger data: The second scenario additionally involves passenger traffic in the load forecast. In the same manner, forecast models have been developed based on available training data in a separate experimental case and in the end the models have been used to forecast the load based on the testing data. A comparison can be visualised in Fig. 6.

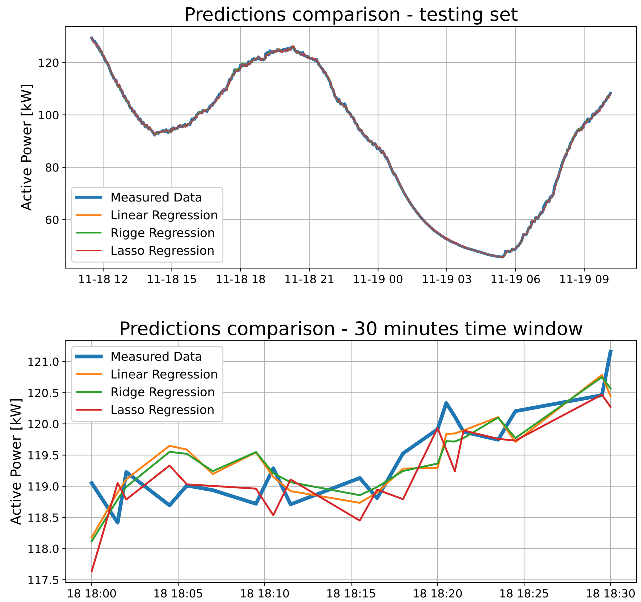

Fig. 5. Scenario 1 - regression model performance comparison

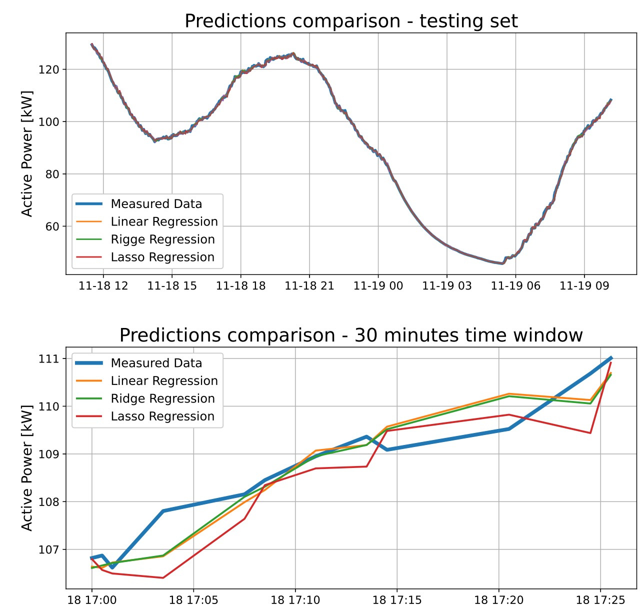

Fig. 6. Scenario 2 - regression model performance comparison

Metrics have been evaluated as well and a comparison between the three regression models can be analysed in Table III.

TABLE III
METRICS EVALUATION ON SCENARIO 2

Model	Evaluation Metric	Value
Scaled Linear Regression	MAPE	0.185 %
	MSE	64359.29 W^2
	RMSE	253.69 W
Ridge Regression	MAPE	0.200 %
	MSE	70167.52 W^2
	RMSE	264.89 W
Lasso Regression	MAPE	0.205 %
	MSE	64359.29 W^2
	RMSE	253.69 W

By analysing the second scenario, it can be observed that passenger traffic improves the forecast obtained by all models. This is highlighted especially by the metrics, where an error decrease can be observed especially in MSE and RMSE. Even if the MAPE error decreases as well, it is important to account for long term scenarios the MSE and RMSE errors. Both indicators suggest that the passenger traffic should be clearly included as a feature in forecast models. Also, the Lasso model performances indicate that in this context it should be regarded as a good alternative to the classic regression models, if it is provided with a corresponding tuning mechanism of the regularization parameter for long term simulations.

V. CONCLUSION

We presented a data science methodology for load forecasting applied to subway traction power in urban rail systems. A two stage data processing and modelling pipeline involves first filtering the input values with an exponential smoothing technique while subsequently applying robust regression techniques for accuracy metric enhancement.

Forecast models have been compared through metrics evaluation and time a time series evolution comparison has been conducted on a multiple scenario framework: a scenario with standard time-related features and a scenario involving passenger traffic.

Results indicated that monitored passenger traffic in the station increases the forecast accuracy for all the developed models and we concluded that it should be considered a required feature from a robustness point of view.

Future work is focused on leveraging the best quality forecasting in a data driven model predictive framework (DD-MPC) for improved energy management at the subway system level with increased usage of unreliable renewable energy sources and localized energy storage.

ACKNOWLEDGMENT

The support of the company Metrorex SA for providing the datasets used for developing the forecasting models presented in this paper is gratefully acknowledged.

REFERENCES

[1] E. Commission. The european green deal. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52019DC0640

[2] "Sustainable energy management for underground stations," CORDIS Project, May 31, 2017. [Online]. Available: https://cordis.europa.eu/project/id/285408

[3] M. Casals, M. Gangolells, N. Forcada, M. Macarulla, and A. Giretti, "A breakdown of energy consumption in an underground station," *Energy and Buildings*, vol. 78, pp. 89–97, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778814003284

[4] R. Faranda and S. Leva, "Energetic sustainable development of railway stations," *2007 IEEE Power Engineering Society General Meeting, PES*, pp. 1–6, 2007.

[5] A. González-Gil, R. Palacin, P. Batty, and J. P. Powell, "A systems approach to reduce urban rail energy consumption," *Energy Conversion and Management*, vol. 80, pp. 509–524, 2014.

[6] S. Siad, G. Damm, L. Galai Dol, and A. De Bernardinis, "Design and control of a DC grid for railway stations," *PCIM Europe 2017 - International Exhibition and Conference for Power Electronics, Intelligent Motion, Renewable Energy and Energy Management*, 2017.

[7] F. Perez, A. Iovine, G. Damm, L. Galai-Dol, and P. Ribeiro, "Regenerative Braking Control for Trains in a DC MicroGrid Using Dynamic Feedback Linearization Techniques," *IFAC-PapersOnLine*, vol. 52, no. 4, pp. 401–406, 2019. [Online]. Available: https://doi.org/10.1016/j.ifacol.2019.08.243

[8] M. S. Simoiu, V. Calofir, I. Fagarasan, N. Arghira, and S. Stelian Iliescu, "Towards Hybrid Microgrid Modelling and Control. A Case Study: Subway Station," *2020 24th International Conference on System Theory, Control and Computing, ICSTCC 2020 - Proceedings*, pp. 602–607, 2020.

[9] C. Nichiforov, G. Stamatescu, I. Stamatescu, V. Calofir, I. Fagarasan, and S. S. Iliescu, "Deep learning techniques for load forecasting in large commercial buildings," in *2018 22nd International Conference on System Theory, Control and Computing (ICSTCC)*, 2018, pp. 492–497.

[10] G. Stamatescu, R. Entezari, K. Römer, and O. Saukh, "Deep and efficient impact models for edge characterization and control of energy events," in *2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*, 2019, pp. 639–646.

[11] G. Stamatescu, I. Stamatescu, N. Arghira, V. Calofir, and I. Fagarasan, "Building Cyber-Physical Energy Systems," *arXiv e-prints*, p. arXiv:1605.06903, May 2016.

[12] A. M. Tudose, D. O. Sidea, I. I. Picioroaga, V. A. Boicea, and C. Bulac, "A cnn based model for short-term load forecasting: A real case study on the romanian power system," in *2020 55th International Universities Power Engineering Conference (UPEC)*, 2020, pp. 1–6.

[13] P. C. Leung and E. W. Lee, "Estimation of electrical power consumption in subway station design by intelligent approach," *Applied Energy*, vol. 101, pp. 634–643, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.apenergy.2012.07.017

[14] R. J. Hyndman, A. B. Koehler, J. K. Ord, and R. D. Snyder, *Forecasting with Exponential Smoothing*. Springer Series in Statistics, 2008.

[15] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. Springer Series in Statistics, 2009, vol. 27, no. 2.

[16] A. Zheng, *Evaluating Machine Learning Models*. O'Reilly Media, Inc., 2021.