Power grid load forecasting using ridge regression including weather forecasts, reanalyses, terrestrial and satellite weather data*

Franko Pandžić[1][0000−0002−3153−2419], Ivan Sudić[1][0000−0002−5703−0513], Amalija Božiček[1][0000−0003−2101−0155], Matko Mesar[1][0000−0001−6878−4401], Bojan Franc[1][0000−0003−0331−2707], Marija Žmire[2], and Ivan Šturlić[2]

[1] Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
[2] Croatian Transmission System Operator, Croatia

Abstract. Power system control requires information regarding network operations available hours, days and even months in advance. Load forecasting plays a pivotal role in system scheduling and decision making. Having accurate load forecasts can ensure uninterrupted and reliable power supply to customers along with potentially reducing operational costs for power system operators. This paper presents a method of short-term grid load forecasting for the Croatian power system usable for various power system control services, e.g. participation on the electricity day-ahead market. Outputs from multiple regression models, each trained for a specific hour of the forecast, were combined to get the complete grid load forecast. Multiple models with the same architecture but with different features were compared to prove that weather data can greatly improve load model performance. Models were trained and tested on data obtained from the Croatian Transmission System Operator and open source ERA5 data provided by Copernicus Climate Change Services.

Keywords: Regression · Time series · Power system · Grid load

1 Introduction

1.1 Motivation

Croatian power system operators are legally obligated by the Electricity Market Act [1] to manage distribution and transmission power systems in a certain way. Operators need to maintain system balance; i.e. have the difference between electricity fed to the grid and consumed in the grid be zero to have continuous power supply. This way the grid load dictates operators' decision making as they aim to deliver uninterrupted and reliable power supply to all customers by having a balanced system at all times. Thus, having precise load forecasts play a crucial role in stable and efficient system scheduling and management. Operators trade

* Supported by European Regional Development Fund (European Union)

energy to keep the system in balance and the exchange in which Croatian system operators mainly participate in is the Croatian power exchange (Cropex). Cropex day-ahead market requires hourly electricity bids for the next day to be submitted by noon of the current day, otherwise the market closes for new participants.
This paper focuses on short-term load forecasts which are available to use for trading strategies on the day-ahead electricity market. Load forecasts from one hour up to one week ahead are considered short-term [7]. This presents one of the main goals for this paper; having load forecasts of the next day (00:00-24:00 CET) [4] available before noon of the current day, i.e. at least 12 hours in advance.
The paper is structured as follows: in section *Methodology* the process of obtaining and transforming data, training time series models and producing load forecasts are explained. Section *Results and Discussion* describes models' performance on multiple measures on the test set while in section *Conclusion and Future Work* the paper is concluded and future research directions are proposed.

1.2 Data

TSO data The Croatian Transmission System Operator (TSO) provided historical load measurements (Jan 2018 - Mar 2022) along with numerical weather predictions (NWP) and meteorological measurements (May 2019 - Mar 2022) for the scope of this paper. Measurements are recorded using automatic weather stations installed throughout the network.
The NWP data are obtained using Weather Research and Forecasting (WRF) model. It is a next-generation mesoscale numerical weather prediction system provided by National Centre for Atmospheric Research (NCAR), USA. WRF has two dynamical solvers, the Advanced Research WRF (ARW) core and the Nonhydrostatic Mesoscale Model (NMM) core. The model is designed for both operational weather forecasting applications and atmospheric research. It can produce simulations of the ideal atmospheric conditions and simulations based on actual conditions. The application of WRF products varies from micro to global scale [5]. WRF data provided by TSO are simulated using ARW core.

Copernicus data In order to fully utilize the load measurements obtained from the TSO, 17 months of missing meteorological data (Jan 2018 - May 2019) needed to be prepended to already existing meteorological data. It was done with publicly available ERA5 data obtained from the Copernicus Climate Change Services (C3S) website [3]. ERA5 is a fifth-generation European Center for Medium-Range Weather Forecast's (ECMWF) reanalysis of meteorological and climatological data. Data are available from 1950. onwards. Reanalyses combine observations and model data to present a consistent and more complete global weather and climate overview. It is possible to obtain reanalysis data at 37 pressure levels, 16 potential temperature levels, and 1 potential vorticity level. Since it is an atmospheric model in ECMWF Integrated Forecast System

(IFS) combined with the land-surface model and ocean wave model, ERA5 also provides information about vegetation and ocean waves.

Observations used in the ERA5 are in-situ and satellite. In-situ data (e.g. SYNOP, METAR, AIREP, etc.) are provided by World Meteorological Organization (WMO) Information System (WIS). Satellite data are obtained with different satellites, and instruments and in different spectral bands by reliable satellite agencies such as NASA, ESA, EUMETSAT, etc. Before utilization in the ERA5, satellite data are preprocessed by space agencies or by relevant institutions (e.g. NOAA). Satellite measurements can provide information about air humidity, column water vapor, cloud liquid water, cloud temperature, precipitation, wind vector, etc [8, 2].

All mentioned data used in this paper has hourly resolution. Different WRF and ERA5 data resolutions are interpolated to be consistent with the weather stations' coordinates explained in the subsection *Meteorological data transformation.*

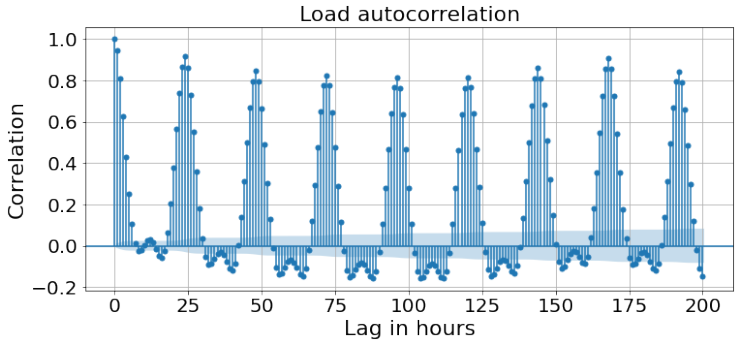
2 Methodology

2.1 Grid loads

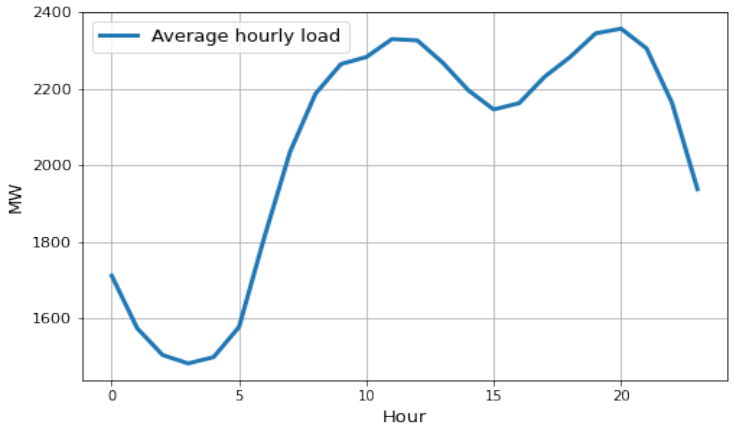
Total grid load is a seasonal time series, as shown in Fig. 1(a). Fig. 1(a) shows the load autocorrelation plot where it is obvious that daily periodicity along with weekly periodicity (higher correlation at the lag of 168 hours) are prominent. Each day has a characteristic camel-like curve shape with the lowest load being early in the morning/late at night and the highest load being from the 10th to the 22nd hour with two peaks as shown in Fig. 1(b). Since the load time series is seasonal, historical loads are the main feature to use for forecasting future loads as shown in [6, 9]. One of the aims of this paper is to show if and how meteorological data can improve load forecasts. As the goal was to produce load forecasts for the whole Croatian power grid, 4 locations around the 4 largest cities in Croatia were selected whose meteo data, specifically temperature, were taken into account. Locations were selected to represent the differences in climate throughout Croatia. The 4 locations are: AMP NDC station (Zagreb), AMP TS Vrboran station (Split), AMP TS Pehlin station (Rijeka) and AMP TS Osijek 1 station (Osijek). NWP data for these stations are provided by the TSO while the Copernicus data was in grid form and needed further transformation.

2.2 Meteorological data transformation

As stated, Copernicus data was in grid form (netCDF format). In other words, meteorological data such as temperature was available only on the intersections of meridians and parallels on the geographic grid system in steps of 0.5° (..., 16.25°, 16.75°, 17.25°,...). Given the station coordinates (x, y), 4 nearest intersections ($P_{11} = (x_1, y_1), P_{12} = (x_1, y_2), P_{21} = (x_2, y_1), P_{22} = (x_2, y_2)$) and values



(a)



(b)

Fig. 1. 1(a): Total grid load autocorrelation, 1(b): Average hourly daily-recurring load curve

on said intersections (v_1, v_2, v_3, v_4); approximated values $f(x, y)$ can be obtained via bilinear interpolation. First, interpolation in the x-direction is done:

$$\begin{aligned} f(x, y_1) &= \frac{x_2 - x}{x_2 - x_1} v_1 + \frac{x - x_1}{x_2 - x_1} v_3, \\ f(x, y_2) &= \frac{x_2 - x}{x_2 - x_1} v_2 + \frac{x - x_1}{x_2 - x_1} v_4 \end{aligned} \tag{1}$$

then the desired estimate is obtained by further interpolating in the y-direction:

$$f(x, y) = \frac{1}{(x_2 - x_1)(y_2 - y_1)} [x_2 - x \; x - x_1] \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix} \begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix} \tag{2}$$

Such interpolation was done for all hourly temperature values (Jan 2018 - May 2019) for all locations and the obtained interpolated temperature time series were prepended to NWP data (May 2019 - Mar 2022) available on the same locations. Fig. 2 shows the dependence between local temperature (x axis) and total grid loads (y axis). It can be seen that grid loads and local temperatures form a characteristic bimodal V-shaped correlation. This correlation can be explained by predictable human behaviour; consuming more energy when it is cold/hot while consuming less energy on heating or air conditioning around the comfortable 15-20°C. The bimodal nature of the correlations stems from different grid loads in different time periods throughout the day; the lower curve represents part of the night and early morning (23-06 h) when people are less active as seen on Fig. 1(b) while the upper V-shaped curve represents the state for all other hours (07-22 h).

2.3 Model inputs

Historical loads As stated, grid loads are seasonal as shown by the autocorrelation plot (Fig. 1(a)), which is why they are the base feature present in all implemented models. Four weeks of hourly historical loads were chosen as a feature (a vector consisting of 24 (hours in a day) * 7 (days in a week) * 4 (number of weeks) elements).

Date features Four features indicating is the observed day a holiday (true/false), which day of the week is it (0-6), is it a weekend (true/false) and which season it is (0-3) as loads are a date-dependent parameter.

Temperature data as is Temperature data in its raw form as a feature.

Temperature as a result of least squares quadratic function The goal is to find the least squares quadratic function that describes temperature and load dependence:

$$f_l(t) = w_{1,l} t^2 + w_{2,l} t + b_l \tag{3}$$

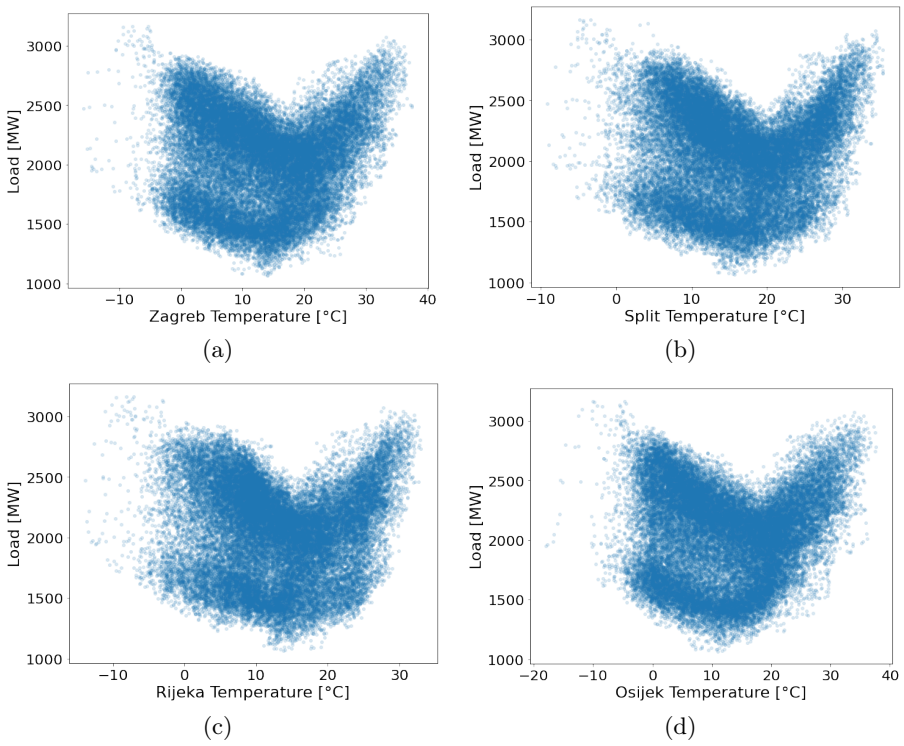


Fig. 2. Scatter plots showing the dependence between total grid loads and temperatures of selected locations: 2(a) - Zagreb, 2(b) - Split, 2(c) - Rijeka, 2(d) - Osijek

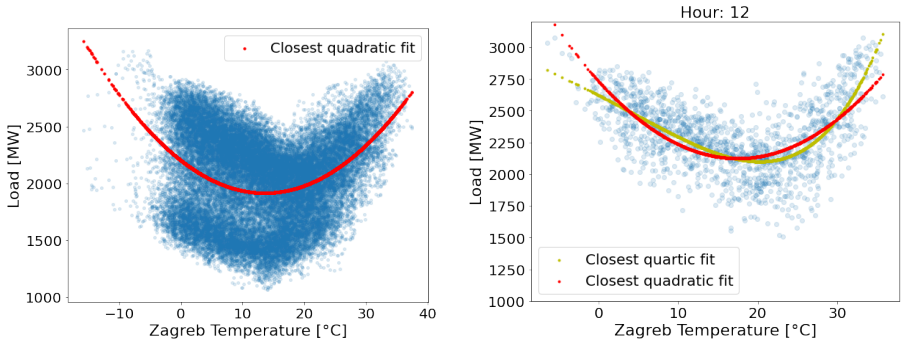
where t is the temperature value, $w_{1,l}$, $w_{2,l}$ and b_l are the least squares weights. Every transformed value per location l is then calculated by solving (3) and as such fed to the models for training and testing. Fig. 3(a) shows an example of such function.

Temperature as a result of least squares quadratic function for every hour in a day Expanding on the previous feature with fitting the least squares quadratic function for every hour separately (example shown in Fig. 3(b)). Again, every transformed value is calculated by solving (3) for a specific location and hour.

Temperature as a result of closest fit quartic function for every hour in a day Analogously to the previous feature, the goal is to find the least squares function for every hour separately; this time the quartic function:

$$f_l^h(t) = w_{1,l}^h t^4 + w_{2,l}^h t^3 + w_{3,l}^h t^2 + w_{4,l}^h t + b_l^h \tag{4}$$

where t is the temperature value, $w_{1,l}^h$, $w_{2,l}^h$, $w_{3,l}^h$, $w_{4,l}^h$ and b_l^h are the least squares weights. Feature values are calculated by solving (4) for a specific location l and hour h. An example of such function is shown in Fig. 3(b).



(a) (b)

Fig. 3. Plots showing closest fits for selected data (3(a): single quadratic fit for Zagreb temperature, 3(b): quadratic and quartic fits for Zagreb temperature for the 12th hour

2.4 Dataset generation

All aforementioned features were normalized separately using min-max scaling:

$$x_s = a + \frac{(x - min(X))(b - a)}{max(X) - min(X)} \tag{5}$$

where X is a time series vector containing all elements of a given feature, x is an element of X and x_s is the scaled x. a and b represent the lower and upper boundaries of scaling, respectively. All features were scaled with the same boundaries.

The i-th sample (X_i, Y_i) is constructed as follows: given i-th time of forecast T_i, X_i is constructed by combining a 4-week historical load vector in the time frame of $[T_i - 672\text{h}, T_i]$ with date features and temperature forecasts for all locations for the same time frame as Y_i: $[T_i + 1\text{h}, T_i + 48\text{h}]$.
Therefore, $T_0 - 672\text{h}$ is the first date of the available dataset (01.01.2018.). All following samples were generated by sliding the time of forecast T by one hour, i.e. $T_1 = T_0 + 1\text{h}$, $T_2 = T_0 + 2\text{h}$ and so on.

The first 80% of the total dataset was used for training models while the rest was used for testing.

2.5 Ridge regression

Ridge regression, also known as Tikhonov regularization, is a linear regression regularized in L2-norm (Euclidean norm). Regularization helps reduce overfitting by shrinking model parameters and thus simplifying the model. Model output is determined by:

$$F(\vec{x_j}) = \sum_{i=1}^{n} w_i f_i, f_i \in \vec{x_j}, w_i \in \vec{w_j}, \tag{6}$$

where $\vec{x_j}$ is the input vector, n is the number of features in $\vec{x_j}$, f_i is the i-th feature of the input vector and w_i is the weight associated with the i-th feature. The optimal fit (optimal weights $\vec{w_j}$) are found by solving:

$$\underset{\vec{w}}{\operatorname{argmin}} \left(\frac{1}{N} \sum_{i=1}^{N} (y_i - F(\vec{x_i}))^2 + \lambda \sum_{i=1}^{N} \vec{w_i}^2 \right), \tag{7}$$

N is the total number of samples in the train set and λ is a hyperparameter which determines the strength of the regularization. Finding optimal weights, i.e. solving (7) is done using averaged stohastic gradient descent.
For the purpose of this paper, 48 (one per forecast hour) such models per model type were fitted with their combined output being the total load forecast:

$$\hat{y_i} = [F_1(\vec{x_i}), F_2(\vec{x_i}), ..., F_{48}(\vec{x_i})] \tag{8}$$

Models are arranged in chain fashion, meaning that the output of every prior model becomes the input for the next one, cummulatively. As the idea is to improve load forecasts using meteorological data, we trained and compared ridge regressors with differentiating inputs as shown in Table 1.

Date column refers to date-dependent features. *Temp* column refers to temperature in raw form as a feature while *Single fit* refers to temperature as a result of least squares quadratic function as a feature. *Hour quadratic fit* refers to the same situation as *Single fit* but for every hour in a day while *Hour quartic fit* is temperature as a result of (4) as a feature.

Table 1. Overview of compared ridge regression models with their respective features

Model name	Historical loads	Date	Temp	Single fit	Hour quadratic fit	Hour quartic fit
ridge	Yes	Yes	No	No	No	No
ridge_meteo	Yes	Yes	Yes	No	No	No
ridge_singlefit	Yes	Yes	No	Yes	No	No
ridge_quad	Yes	Yes	No	No	Yes	No
ridge_quart	Yes	Yes	No	No	No	Yes

3 Results and Discussion

Table 1 lists all implemented models for comparison. The mentioned models were compared using measures explained in the next subsection.

3.1 Comparison measures

Hourly mean absolute error:

$$MAE_h = \frac{1}{N}\sum_{i=1}^{N}|y_{h_i} - \hat{y_{h_i}}|, h = 1, ..., 48 \tag{9}$$

Hourly mean absolute percentage error:

$$MAPE_h = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_{h_i} - \hat{y_{h_i}}}{y_{h_i}}\right| * 100, h = 1, ..., 48 \tag{10}$$

where N is the number of samples in the test set, $\hat{y_{h_i}}$ is the predicted and y_{h_i} is the target value for the h-th hour of i-th sample. MAE_h is expressed in megawatts (MW), while $MAPE_h$ in percentage (%) deviation.

3.2 Model performance

Table 2 shows the average MAE and MAPE for each model on 4 subsets of the train set. The subsets are differentiated by their time of forecast. These subsets were chosen as load forecasts generated at that time are usable for participation on the day-ahead electricity market, thus satisfying the first constraint of the paper. Even though each load forecast is for the next 48 hours, MAE and MAPE in Table 2 were determined only on the relevant hours, i.e. day-ahead forecasts:

- relevant hours for TOF_{8h}: $[TOF_{8h} + 16\text{h}, TOF_{8h} + 39\text{h}]$
- relevant hours for TOF_{9h}: $[TOF_{9h} + 15\text{h}, TOF_{9h} + 38\text{h}]$
- relevant hours for TOF_{10h}: $[TOF_{10h} + 14\text{h}, TOF_{10h} + 37\text{h}]$
- relevant hours for TOF_{11h}: $[TOF_{11h} + 13\text{h}, TOF_{11h} + 36\text{h}]$

Table 2. Overview of model performance for relevant hours on subsets of the test set considering time of forecast (TOF)

TOF	metric	ridge	ridge_meteo	ridge_singlefit	ridge_quad	ridge_quart
08:00 h	MAPE [%]	2,995	2,955	2,665	2,376	2,109
	MAE [MW]	65,176	64,242	57,242	51,035	44,695
09:00 h	MAPE [%]	2,85	2,811	2,579	2,332	2,086
	MAE [MW]	61,823	60,911	55,155	49,889	44,029
10:00 h	MAPE [%]	2,732	2,672	2,502	2,252	2,036
	MAE [MW]	58,937	57,788	53,324	48,099	42,955
11:00 h	MAPE [%]	2,64	2,579	2,454	2,225	2,038
	MAE [MW]	56,929	55,625	52,157	47,458	42,968

Table 2 shows that adding temperature in any form as a feature improves model accuracy. *ridge_meteo* has a slightly favorable MAE and MAPE while *ridge_singlefit*, *ridge_quad* and *ridge_quart* improve MAPE by 7-11%, 16-21% and 23-30% over *ridge*, respectively. Models rank the same based on both MAE and MAPE.

Fig. 4 displays the average MAE and MAPE of the models for each hour of the forecast. Fig. 4(a) shows model errors for the whole test set (TOF is any hour in a day) while fig. 4(b) shows model errors for forecasts with $TOF = 9h$. *ridge_quart* is the best performing model on average for every hour of the forecast by MAE and MAPE, followed by *ridge_quad* and *ridge_singlefit*, respectively.

4 Conclusion and Future Work

To conclude the paper, accurate load forecasting is of high importance to energy system operators. In order to improve said forecasts, the influence of temperature data on load forecasting was explored. It is shown in this paper that including temperature and temperature derivatives at specific locations can improve load forecast by over 29%.

One research direction to propose would be analysing the influence of other meteorological data on grid loads such as thermal comfort, wind speed, rainfall etc. Moreover, it could be beneficial to further engineer temperature data with characteristic days (holidays), days of the week and seasons of the year (spring, summer, autumn, winter). Furthermore, more advanced machine learning methods, such as a deep residual network [6] are worthwhile to apply for the given problem of load forecasting.

Acknowledgement

The project is co-financed by the European Union from the European Regional Development Fund and by the Environmental Protection and Energy Efficiency Fund under project KK.05.1.1.02.0002 RESdata (Solutions for power

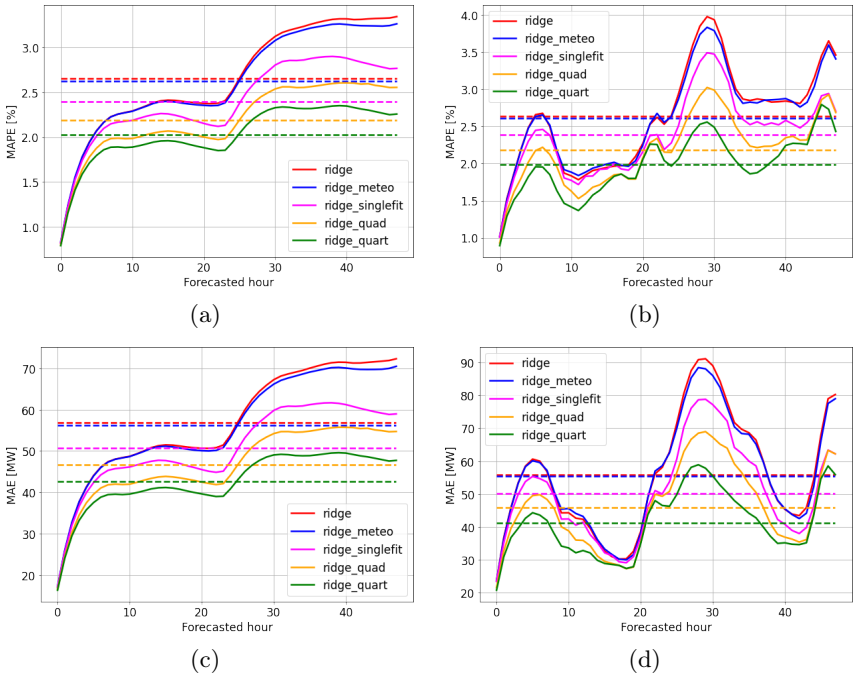


Fig. 4. Plots comparing average MAPE (4(a), 4(b)) and MAE (4(c), 4(d)) of each forecasted hour for every implemented model; Errors on 4(a) and 4(c) are calculated on the whole test set while 4(b), 4(d) are calculated on TOF_{9h} subset

systems adjustment to climate changes based on large quantity of data) and the European Regional Development Fund (European Union), grant number KK.01.2.1.02.0117, with the project title "Advanced power system management in conditions of uncertainty arising from climate changes".

References

1. Zakon o tržištu električne energije (*Electricity Market Act)* (2021), https://www.zakon.hr/z/377/Zakon-o-t%C5%BEi%C5%A1tu-elektri%C4%8Dne-energije
2. ERA5 documentation (2022), https://confluence.ecmwf.int/display/CKB/ERA5%3A+data+documentation
3. Copernicus climate data store (2022), https://cds.climate.copernicus.eu/cdsapp!/dataset/derived-near-surface-meteorological-variables
4. Croatian power exchange (2022), https://www.cropex.hr/en/documents.html
5. Weather research and forecasting model (2022), https://www.mmm.ucar.edu/weather-research-and-forecasting-model
6. Chen, K., Chen, K., Wang, Q., He, Z., Hu, J., He, J.: Short-term load forecasting with deep residual networks. IEEE Transactions on Smart Grid **10**(4), 3943–3952 (2019). https://doi.org/10.1109/TSG.2018.2844307
7. Hammad, M.A., Jereb, B., Rosi, B., Dragan, D.: Methods and models for electric load forecasting: a comprehensive review. Logistics, Supply Chain, Sustainability and Global Challenges **11**(1), 51–76 (2020)
8. Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The era5 global reanalysis. Quarterly Journal of the Royal Meteorological Society **146**(730), 1999–2049 (2020)
9. Kuster, C., Rezgui, Y., Mourshed, M.: Electrical load forecasting models: A critical systematic review. Sustainable cities and society **35**, 257–270 (2017)