Short-term forecasting of carbon emissions from electricity generation: the Italian case

Pierdomenico Duttilo and Francesco Lisi
Department of Statistcal Sciences
University of Padova
Padova, Italy
pierdomenico.duttilo@unipd.it, francesco.lisi@unipd.it

***Abstract*—This study investigates short-term forecasting of carbon emissions from electricity generation in Italy. The analysis considers a range of forecasting models, including linear parametric models, such as seasonal autoregressive integrated moving average (SARIMA); functional parametric models, including seasonal functional autoregressive (S)FAR and functional autoregressive with exogenous input (SFARX); and nonparametric (possibly nonlinear) models, such as generalised additive models (GAM) and the TBATS model. In addition, forecast combination techniques are explored, incorporating predictions from multiple models through simple averaging, the Bates and Granger (1969) method, and a selection-based approach. Forecast performance is evaluated using hourly root mean squared error (RMSE) over a one-year test period. The results indicate that the functional models consistently achieve the lowest relative RMSE during the early morning hours, whereas the GAM models outperform all others throughout the day, afternoon, and evening. These findings highlight the benefits of forecast combinations, particularly the selection-based approach, which effectively integrates functional and GAM models to improve predictive performance.**

***Index Terms*—CO_2 emissions forecasting, electricity generation, generalized additive models, functional autoregressive models, forecast combinations**

I. Introduction

Carbon dioxide (CO_2) emissions are a major contributor to global climate change, with electricity generation being one of the main sources [1]. Carbon emissions vary significantly based on the energy mix used for power production, which includes both renewable sources like solar and wind, and non-renewable sources like coal, oil and natural gas.

Reducing CO_2 emissions from electricity generation is essential to achieve international climate targets, such as those established by the Paris Agreement in 2015. This global effort aims to keep the increase in global temperatures well below 2°C above preindustrial levels, with a more ambitious target of 1.5°C. It also sets specific goals to reduce CO_2 emissions by 20%, increase the share of renewable energy by 20%, and improve energy efficiency by 20% worldwide [2].

In this context, CO_2 short-term forecasts play a critical role in enabling emission-based scheduling and optimising the operation of energy storage systems. Carbon emissions from energy use can be reduced by scheduling flexible energy demand [3]. This approach aligns high-energy operations with periods of peak renewable energy availability, reducing dependence on fossil fuels [4]. In addition, accurate short-term forecasting can enhance the efficiency of energy storage technologies. For example, battery systems can be programmed to charge when emissions are low and discharge during high-emission intervals [5]. Beyond their operational benefits, emission predictions play a crucial role for policymakers, enabling them to design taxes and incentives that encourage consumers to adopt flexible energy strategies.

Several forecasting methods have been developed to address the need for accurate CO_2 emission predictions. The comprehensive review by Jin et al. (2024) provides valuable insight into the progress made in this field over the past decade [6]. Carbon emission prediction models can be broadly classified into four main categories: statistical models, shallow intelligent models, neural networks, and hybrid approaches.

Statistical models are widely used for the prediction of carbon emissions, offering interpretable frameworks based on historical data. Among these, the grey model is the most frequently applied, especially with small sample sizes and limited information. Regression analysis, including panel regression and linear regression, has also been used to examine trade-related carbon emissions and forecast trends in different regions. Furthermore, univariate time series models such as seasonal autoregressive integrated moving average (SARIMA) and autoregressive distributed lag have been applied to capture temporal dependencies in emissions data, demonstrating reliable forecast accuracy in various case studies. These models capture trends and seasonal patterns, but can be affected by fluctuations when carbon emissions data are influenced by various factors.

Shallow intelligent models, including support vector machines and decision trees, enhance predictive performance by handling complex relationships in the data. Neural networks, both feedforward and feedback types, have gained prominence for their ability to model nonlinear relationships and temporal dependencies in emissions data.

The latest advances in this field are hybrid approaches that combine multiple methodologies to leverage their respective strengths. Hybrid methods combine one or more approaches

This study was funded by the European Union - NextGenerationEU, Mission 4, Component 2, in the framework of the GRINS - Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 – CUP C93C22005270001). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

979-8-3315-1278-1/25/$31.00 ©2025 IEEE

from the classes described so far and can be further grouped into statistical-statistical, statistical-intelligent, and intelligent-intelligent combinations, each offering unique advantages in terms of accuracy, explainability, and computational efficiency.

This study aims to extend the comparison of statistical models for short-term CO_2 emissions forecasting in Italy by considering classical and functional approaches, specifically: linear parametric models, including SARIMA models and its extension with exogenous variables (SARIMAX); functional parametric models, such as the seasonal functional autoregressive (S)FAR and the functional autoregressive with exogenous input (SFARX); and nonparametric (and possibly nonlinear) models, including generalized additive models (GAM) and the TBATS model (trigonometric seasonality, Box-Cox transformation, ARMA errors, trend, and seasonal components), which are designed to capture complex nonlinear patterns. Furthermore, both Naïve forecasts and forecast combinations that integrate predictions from different models are examined. The forecast combination methods considered include the simple average, the Bates and Granger (1969) method, and the selection-based combination approach.

The rest of the paper is structured as follows: CO_2 emissions are described in Section II; single model and forecast combination setups are introduced in Section III and IV; the forecasting experiment is described in Section V; Section VI presents the forecasting results.

II. Data

Carbon emissions from electricity generation, denoted as E_t, are calculated using the emission factor-based method [7], as follows:

$$\begin{aligned} E_t &= \sum_{i=1}^{I} E_{t,i}, \\ E_{t,i} &= G_{t,i} \times EF_i \times O_i \times M, \end{aligned} \tag{1}$$

where $E_{t,i}$ denotes carbon emissions at time t for fuel type i, $G_{t,i}$ represents the hourly electricity generation at time t by fuel type i, EF_i is the country-specific emission factor for fuel type i, O_i is the oxidation rate for fuel type i, and M is the ratio of the molecular weight of CO_2 to the atomic weight of carbon.

Carbon emissions from electricity generation are calculated for Italy's day-ahead electricity market. The electricity generation data are obtained from ENTSO-E [8] and cover a 3-year period from 1 January 2020 to 31 December 2023. The data is recorded with an hourly frequency over 24 hours. Carbon emission data are high-frequency time series with daily, weekly, and annual seasonality.

Figures 4-6 in the Appendix illustrate the daily, weekly, and monthly boxplot of Italian carbon emissions (tCO_2), highlighting higher emissions during peak demand hours.

III. Single model setup

To model the different components, the following calendar variables are considered: H_t corresponds to the daily seasonal component and serves as an indicator variable for the hour at time t; W_t denotes the weekly seasonal component and is an indicator variable for the day of the week at time t; B_t is a dummy variable for the bank holiday; Y_t represents the yearly seasonal component and is an indicator variable for the specific day of the year at time t; T_t represents the trend variable, which indicates a continuous time index for each time instant t.

A Naïve persistent model is used for benchmark purposes. It is expressed as follows:

$$E_t = E_{t-24}. \tag{2}$$

The SARIMAX$(p,d,q)(P,D,Q)S$ model is specified as follows:

$$\begin{aligned} E_t =& \mu + \sum_{i=1}^{p} \phi_i E_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{P} \Phi_k E_{t-kS} \\ &+ \sum_{l=1}^{Q} \Theta_l \epsilon_{t-lS} + \alpha_1 W_t + \alpha_2 Y_t + \alpha_3 B_t + \epsilon_t, \end{aligned} \tag{3}$$

where μ represent the intercept term; ϕ_i denote the non-seasonal autoregressive coefficients; θ_j the non-seasonal moving average coefficients; S the seasonal period (24 for daily seasonality and 7 for weekly seasonality); Φ_k the seasonal autoregressive coefficients; Θ_l the seasonal moving average coefficients; and α_1, α_2, and α_3 represent the effects of the day of the week, the day of the year, and bank holidays, respectively; ϵ_t denotes the error term.

CO_2 emissions from electricity generation can be naturally represented as curves, making it logical to capture the serial dependence among daily curves and their temporal dynamics [9]. Let $\mathcal{H}$ be a real separable Hilbert space and consider the continuous-time carbon emissions process E_t defined in $\mathcal{H}$. The seasonal functional autoregressive model with exogenous input SFARX(p) is specified as follows:

$$\begin{aligned} E_t =& \sum_{i=1}^{p} \rho_i(E_{t-i}) + \gamma(E_{t-7}) + \alpha_1(W_t) + \alpha_2(Y_t) \\ &+ \alpha_3(B_t) + \epsilon_t, \quad \text{for } t \in \mathbb{N}, \end{aligned} \tag{4}$$

where: ρ_i, $\gamma, \alpha_1, \alpha_2$ and α_3 are bounded linear operators in $\mathcal{H}$; W_t, Y_t and B_t are continuous exogenous variables defined in $\mathcal{H}$; ϵ_t is a process of white noise valued at $\mathcal{H}$. The linear operator ρ_i quantifies the influence of the lagged emission curve E_{t-i}, while γ quantifies the influence of the lagged emission curve of the previous week E_{t-7}. The operators α_1, α_2 and α_3 account for the effects of weekly and yearly seasonalities and bank holidays, respectively.

The GAM model is defined by

$$\begin{aligned} E_t =& + f_1(T_t) + f_2(Y_t) + \alpha_1 H_t + \alpha_2 W_t + \alpha_3 B_t \\ &+ \alpha_4 E_{t-24} + \alpha_5 E_{t-48} + \alpha_6 E_{t-72} + \alpha_7 E_{t-168} + \epsilon_t, \end{aligned} \tag{5}$$

where: $f_1(T_t)$ is a smoothing spline function that models the long-term behavior of carbon emissions over time; $f_2(Y_t)$ is a seasonal smoothing spline function that captures the seasonal

variation across the year; $\alpha_1 H_t$ captures the hourly seasonality; $\alpha_2 W_t$ captures the weekly seasonality; $\alpha_3 B_t$ models the effect of bank holidays; $\alpha_4 E_{t-24}$, $\alpha_5 E_{t-48}$, and $\alpha_6 E_{t-72}$ capture the hourly lag from one to three days earlier; $\alpha_7 E_{t-168}$ captures the hourly lag one week earlier; ϵ_t is the error term.

The TBATS model, introduced by De Livera et al. (2011), is a flexible time series forecasting method designed to handle multiple seasonalities, trend components, and short-term dependencies [10]. The TBATS model can be represented mathematically as follows:

$$\begin{aligned} E_t^{(\omega)} &= l_{t-1} + \phi b_{t-1} + \sum_{i=1}^{M} S_{t-m_i}^{(i)} + d_t \\ \text{with } E_t^{(\omega)} &= \begin{cases} \frac{E_t^{\omega}-1}{\omega}, & \text{if } \omega \neq 0 \\ \log(E_t), & \text{if } \omega = 0 \end{cases} \end{aligned} \tag{6}$$

where: $E_t^{(\omega)}$ is the Box-Cox transformation of E_t with parameter ω; $l_t = l_{t-1} + \phi b_{t-1} + \alpha d_t$ is the local level component; $b_t = (1-\phi)b + \phi b_{t-1} + \beta d_t$ is the short-term trend component; ϕ is the trend damping parameter; $S_t^{(i)}$ denotes the ith seasonal component at time t modeled with a Fourier-based trigonometric representation; d_t is the error term which follows an ARMA(p, q) process; m_i denotes the seasonal periods.

IV. Forecast combinations setup

Combined forecasts, denoted as $\tilde{E}$, are linear combination of individual forecasts expressed as follows

$$\tilde{E}_{T+\tau|T} = \mathbf{w}'_{T+\tau|T} \hat{\mathbf{E}}_{T+\tau|T}, \tag{7}$$

where $\mathbf{w}_{T+\tau|T} = (w_{T+\tau|T,1}, \ldots, w_{T+\tau|T,N})'$ is an N-dimensional vector of linear combination weights assigned to N individual forecasts [11]. The simple average combination method assigns equal weights to all individual forecasts.

Bates and Granger (1969) introduced a method to determine the optimal weights by minimising the variance of the combined forecast error [12], assigning weights inversely proportional to the forecast error variance, given by

$$\mathbf{w}^{\text{BG}}_{T+\tau|T} = \frac{\mathbf{\Sigma}^{-1}_{T+\tau|T}\mathbf{1}}{\mathbf{1}'\mathbf{\Sigma}^{-1}_{T+\tau|T}\mathbf{1}}, \tag{8}$$

where $\mathbf{\Sigma}_{T+\tau|T}$ is the covariance matrix $N \times N$ of the τ-step forecast errors.

In addition to conventional methods, a selection-based combination approach can be applied, where the model with the lowest hourly forecast error measure $\mathcal{E}_{h,i}$ is selected for each forecast horizon $T+\tau$. Then, a weight of 1 is assigned to the best-performing model and 0 to the others. The selection-based combination approach can be expressed as follows:

$$\begin{aligned} \mathbf{w}^{\text{SEL}}_{T+\tau|T} &= \mathbf{e}_{j^*} \\ j^* &= \arg \min_{i \in \{1,\ldots,N\}} \mathcal{E}_{h,i} \end{aligned} \tag{9}$$

where $\mathbf{e}_{j^*}$ is a binary $(0; 1)$ vector that selects the best model.

V. Forecasting experiment

The forecast setup is shown in Figure 1. The models are trained on data spanning from 1 January 2021 to 30 June 2022. Carbon emissions in the training set ranges from 4,034.37 to 26,145.38 tCO_2, with a mean of 13,443.26 tCO_2, and a standard deviation of 4,283.90 tCO_2, indicating notable variability. Once trained, forecasts are generated for a one-day-ahead horizon over the validation set, which covers the period from 1 July 2022 to 31 December 2022. The models are updated daily (every 24 observations/hours). By the validation set, the best-performing model is identified for each class. Then, forecast combinations among the best-performing models are evaluated in the test set with one-day-ahead forecasts from 1 January 2023 to 31 December 2023.

Forecast performance are evaluated using the root mean squared error (RMSE), computed for each hour as follows:

$$\text{RMSE}_h = \sqrt{\frac{1}{T_h}\sum_{t=1}^{T_h}(E_{t,h} - \hat{E}_{t,h})^2}, \quad h = 0, \ldots, 23 \tag{10}$$

where T_h denotes the number of observations for hour h. In addition, the relative RMSE and the hourly average RMSE are also used to assess performance.

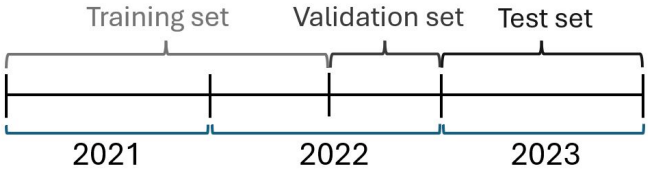

Fig. 1. Forecast setup.

VI. Forecasting results

A. Single model results

Table I presents the best model in each class based on the hourly average RMSE. In the validation set, the GAM model achieves the lowest RMSE of 1841.89, followed by FARX(2) with a value of 2192.22. The SARX(2)24 model has an RMSE of 2770.01. In comparison, the TBATS model performs the worst, with an RMSE of 3165.58, slightly higher than the RMSE of the Naïve model of 3151.45.

Figure 2, panel a, shows the hourly relative RMSE to the Naïve model for each selected model in the validation set. The FARX(2) model consistently produces the lowest relative RMSE during the early hours of the day, from 00:00 to 05:00. However, the GAM model outperforms all other models by producing lower relative RMSE values during the daytime, afternoon, and evening hours from 06:00 to 23:00. The SARX(2)24 and TBATS models exhibit a higher relative RMSE compared to the FARX(2) and GAM models throughout most hours of the day.

The test set confirms the results of the validation set. In Table I, the GAM model achieves the lowest RMSE at 1805.20, followed by FARX(2) with a value of 2186.14 (Table

I). In Figure 2, panel (b), the FARX(2) model demonstrates better prediction accuracy during the early hours of the day, while the GAM model outperforms it from morning to night.

TABLE I
HOURLY AVERAGE RMSE BY MODELS

Model	Validation set	Test set
Naïve	3151.45	2858.33
SARX(2)24	2770.01	2530.43
FARX(2)	2192.22	2186.14
GAM	1841.89	1805.20
TBATS	3165.58	2868.11

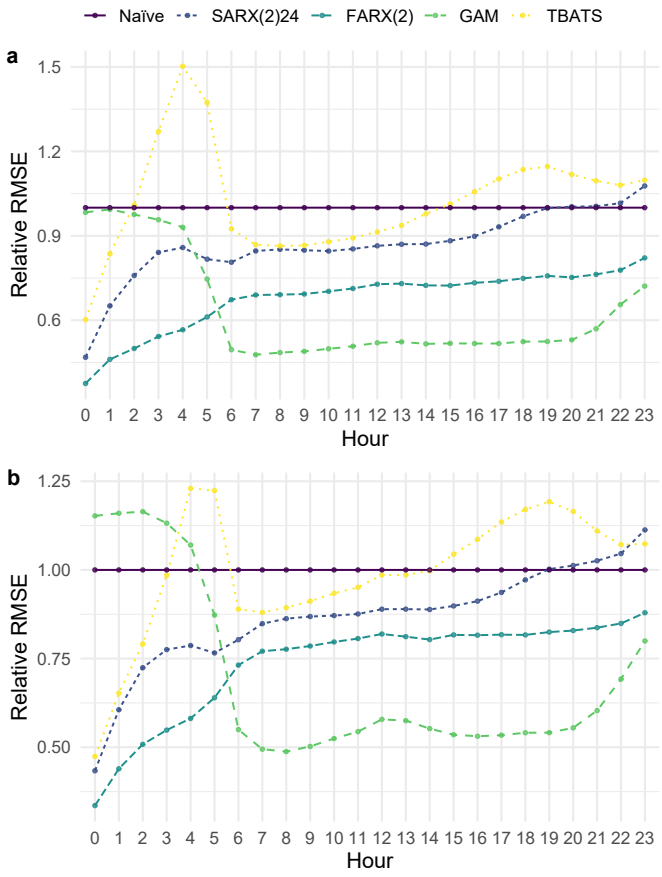

Fig. 2. Relative RMSE of the top-performing models by class for the validation set (panel a) and the test set (panel b). The relative RMSE is calculated with respect to the Naïve model.

B. Forecast combinations results

Since CO_2 emissions dynamics fluctuate throughout the day, different models capture specific patterns more effectively at different times. Rather than relying on a single best model, combining multiple forecasts leverages their strengths, enhancing overall reliability. Using the information available from the validation set, it is possible to build the following forecast combinations:

- COMB1.BG: SARX(2)24, FARX(2), GAM, TBATS;
- COMB2.SA: SARX(2)24, FARX(2), GAM, TBATS;
- COMB3.BG: FARX(2), GAM;
- COMB4.SA: FARX(2), GAM;
- COMB5.BG: SARX(2)24, GAM;
- COMB6.SA: SARX(2)24, GAM;
- COMB7.SEL: FARX(2), GAM.

Here, SA, BG, and SEL refer to the simple average, Bates and Granger (1969), and selection-based combination approaches, respectively. In particular, COMB7-SEL employs an hourly selection-based approach determined in the validation set. Specifically, for the early hours of the day (00:00 - 05:00), the forecasts from FARX(2) are used, while for the remaining hours (06:00 - 23:00), the predictions from GAM are applied.

Table II reports the hourly average RMSE of the forecast combinations for the test set. The simple average combination yields a lower RMSE than the Bates and Granger (1969) method, except for the combination of FARX(2) and GAM (COMB3.BG and COMB4.SA). The best performing combination, COMB7.SEL, achieves an RMSE of 1564.54, representing a 45.27% improvement over the Naïve model and a 15.06% improvement over the single GAM model (in Table I). As an alternative, COMB3.BG, the second-best forecast combination, achieves an RMSE of 1670.77, representing a 41.55% improvement over the Naïve model. Figure 3 presents the hourly average RMSE of the forecast combinations for the test set. COMB7.SEL is the best model in terms of relative RMSE for 19 out of 24 hours.

The results confirm that the forecast combinations improved the prediction of CO_2 emissions. In particular, the combination of functional models and GAM models highlights the complex nature of the data, which is challenging to capture with single models. These findings align with recent literature, which supports hybrid modelling approaches [6]- [13]. Several limitations should be acknowledged. The analysis relies on RMSE, which could be complemented by metrics such as mean absolute error or confidence intervals. Furthermore, regional differences in generation mix and policies may affect the models' applicability across Italian electricity market zones. Future research may explore other methodologies, expand electricity market comparison, and assess computational challenges or broader applicability in different contexts.

TABLE II
HOURLY AVERAGE RMSE BY FORECAST COMBINATIONS

Combination	RMSE
Naïve	2858.33
COMB1.BG	2410.49
COMB2.SA	1965.72
COMB3.BG	1670.77
COMB4.SA	1701.20
COMB5.BG	2360.74
COMB6.SA	1847.81
COMB7.SEL	1564.54

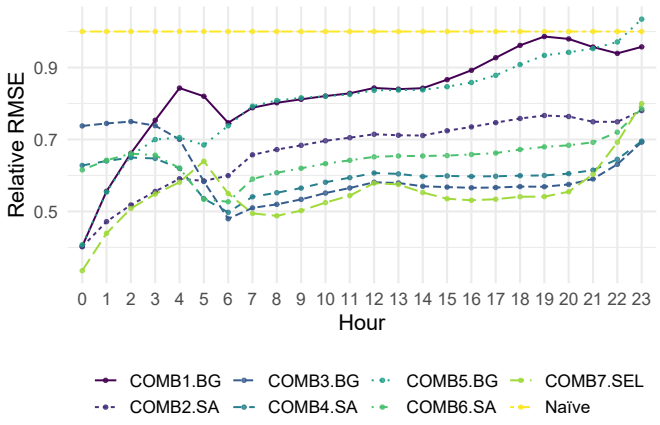

Fig. 3. Relative RMSE of the forecast combinations for the test set. The relative RMSE is calculated with respect to the Naïve model.

REFERENCES

[1] Global CO_2 emissions by sector, IEA, Paris [Online; accessed 4-Feb-2024]; 2017. https://www.iea.org/data-and-statistics/charts/global-co2-emissions-by-sector-2017

[2] The Paris Agreement, UNFCCC, Paris [Online, accessed: 04-Feb-2025]. https://unfccc.int/process-and-meetings/the-paris-agreement

[3] Leerbeck, K., Bacher, P., Junker, R. G., Goranović, G., Corradi, O., Ebrahimy, R., Tveit, A., and Madsen, H. (2020). "Short-term forecasting of CO2 emission intensity in power grids by machine learning", Appl. Energy, 277, 1-13

[4] Liang, Z., Liang, J., Wang, C., Dong, X., and Miao, X. (2016). "Short-term wind power combined forecasting based on error forecast correction", Energy Convers. Manag., 119, 215-226

[5] Bokde, N. D., Tranberg, B., and Andresen, G. B. (2021). "Short-term CO2 emissions forecasting based on decomposition approaches and its impact on electricity market scheduling", Appl. Energy, 281, 1-20

[6] Jin, Y., Sharifi, A., Li, Z., Chen, S., Zeng, S., and Zhao, S. (2024). "Carbon emission prediction models: A review", Sci. Total Environ., 927, 1-20

[7] M. Bertolini, P. Duttilo, and F. Lisi (2024) "Accounting carbon emissions from electricity generation: a review and comparison of emission factor-based methods", arXiv, 1-29, https://doi.org/10.48550/arXiv.2411.13663

[8] ENTSO-E. European Network of Transmission System Operators for Electricity. https://www.entsoe.eu/

[9] Chen, Y., Koch, T., Lim, K. G., Xu, X., and Zakiyeva, N. (2021). "A review study of functional autoregressive models with application to energy forecasting", WIREs Comput. Stat., 13(3), 15-25

[10] De Livera, A. M., Hyndman, R. J., and Snyder, R. D. (2011). "Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing", J. Am. Stat. Assoc., 106(496), 1513-1527

[11] Wang, X., Hyndman, R. J., Li, F., and Kang, Y. (2023). "Forecast combinations: An over 50-year review", Int. J. Forecast., 39(4), 1518-1547

[12] Bates, J. M., and Granger, C. W. J. (1969). "The Combination of Forecasts", OR, 20(4), 451-468

[13] Bordignon, S., Bunn, D. W., Lisi, F., and Nan, F. (2013). "Combining day-ahead forecasts for British electricity prices", Energy Econ., 35, 88-103

Data availability statement The authors do not have permission to share the data.

APPENDIX

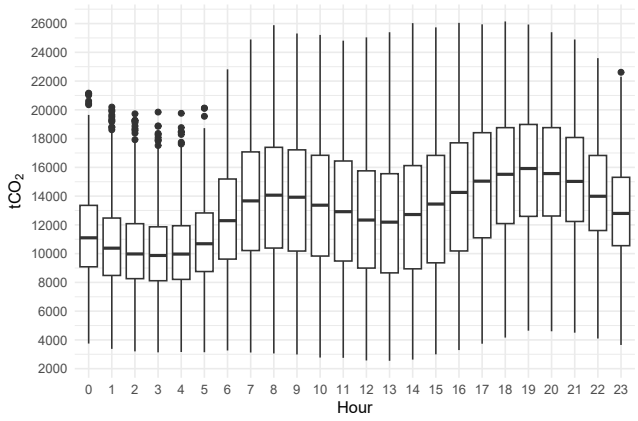

Fig. 4. Hourly boxplot of Italian carbon emissions (tCO_2), 2021-2023.

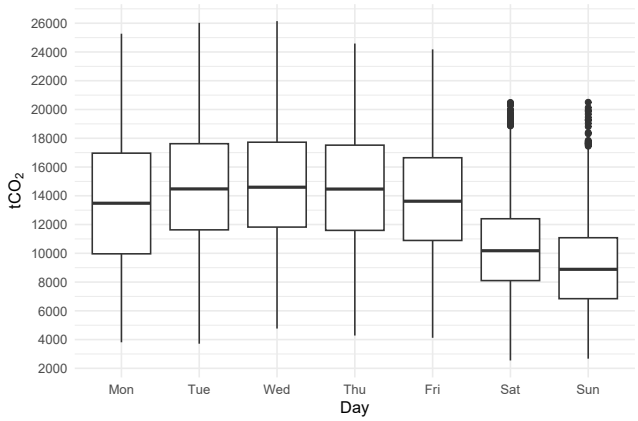

Fig. 5. Weekly boxplot of Italian carbon emissions (tCO_2), 2021-2023.

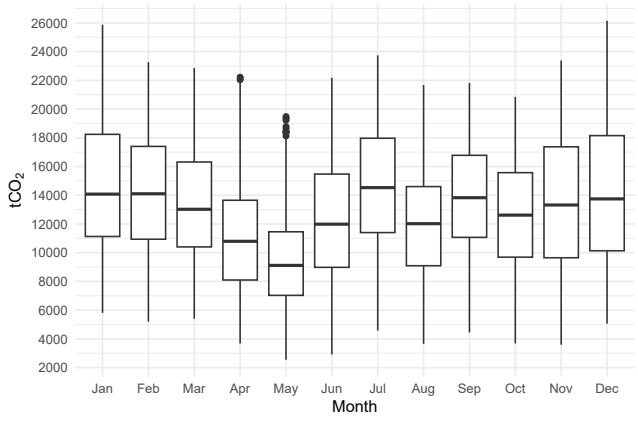

Fig. 6. Monthly boxplot of Italian carbon emissions (tCO_2), 2021-2023.