UNIVERSITÀ di VERONA

Department of ECONOMICS

Working Paper Series
Department of Economics
University of Verona

The zonal and seasonal CO2 marginal emissions factors for the Italian power market

Filippo Beltrami, Fulvio Fontini, Monica Giulietti, Luigi Grossi

WP Number: 1 January 2021

ISSN: 2036-2919 (paper), 2036-4679 (online)

The zonal and seasonal CO_2 marginal emissions factors for the Italian power market

Filippo Beltrami[‡], Fulvio Fontini[†], Monica Giulietti[⋆], and Luigi Grossi[‡]

[†]Department of Economics, University of Verona, Verona, Italy;
filippo.beltrami@univr.it

[†]Department of Economics, University of Padova, Padova, Italy;
fulvio.fontini@unipd.it

[⋆]School of Business and Economics, University of Loughborough, Loughborough, UK; M.Giulietti@lboro.ac.uk

[‡]Department of Economics, University of Verona, Verona, Italy;
luigi.grossi@univr.it

Abstract

This paper estimates the seasonal and zonal CO_2 marginal emissions factors (MEFs) from electricity production in the Italian electricity system. The inclusion of the zonal configuration of the Italian wholesale power market leads to a complete measurement of marginal emission factors which takes into account the heterogeneous distribution of RES power plants, their penetration rate and their variability within the zonal power generation mix. This article relies on the fractional cointegration methodology to incorporate the typical features of long-memory processes into the estimation of MEFs. We find high variability in annual MEFs estimated at the zonal level. Sardinia reports the highest MEF (0.7189 tCO_2/MWh), followed by the Center South (0.7022 tCO_2/MWh), the Center North (0.4236 tCO_2/MWh), the North

(0.2018 tCO_2/MWh) and Sicily (0.146 tCO_2/MWh). The seasonal analysis also shows a large variability of MEFs in each zone across time. The heterogeneity of results leads us to recommend that policymakers consider the zonal configuration of the power market and the large seasonal variability related to carbon emissions and electricity generation when designing incentives for Renewable Energy Sources (RES) expansion and for achieving emission reduction targets.

Keywords: *Decarbonization; Electricity Price; Fractional Cointegration; Marginal Emission Factor (MEF); Renewable Energy Sources (RES).*

Abbreviations: ACF (AutoCorrelation Function); AEF (Average Emission Factor); FCVAR (Fractional Cointegration Vector Auto Regressive); LR (Likelihood Ratio); MEF (Marginal Emission Factor); MISO (Midcontinent Independent System Operator); NRRES (Non Relevant Renewable Energy Sources); OLS (Ordinary Least Squares); PACF (Partial AutoCorrelation Function); PV (Photo Voltaic); RES (Renewable Energy Sources); TSO (Transmission System Operator).

JEL Codes: P18; Q41; Q42; Q51; C22; C32; C63

1 Introduction

Climate change mitigation policies call for the evaluation of the relationship between renewable energy generation and greenhouse gases reduction. For the electricity sector, this requires evaluating the impact that electricity generation from Renewable Energy Sources (RES) has on CO_2 emissions (Bretschger and Pittel, 2020).

The simplest way to address this issue is to measure the ratio between the emissions of a given power system and the corresponding power generation in a given period. This is the so-called Average Emission Factor (AEF), which however ignores the structure of the generation mix in the power system and neglects the dynamic changes in the merit-order of power supply. Indeed, electricity is produced by different technologies, each of them with a different rate of activity (the so called *capacity factor*) and a different contribution to the load in a given period (the *load factor*). For example, when load rises, the power plants that are producing at the margin (and beyond it) are progressively called to serve the extra load. This causes an increase in the level of emissions which is explained by the extra emissions of these marginal (and beyond marginal) plants, and not by the average emissions that are related to the production of the base-load plants which are unaffected by the load change. Thus, a proper measure of the sensitivity of carbon emissions to the change of the load would be given by the average emissions of the marginal plants, i.e., the plants that produce at the margin. This is the so-called Marginal Emission Factor (MEF).

From a policy point of view, the correct measurement of the MEF is a topical issue. An energy saving policy, for instance, would induce a change in marginal emissions, whose impact would be measured by the MEF. Similarly, any policy that would change the merit-order structure, for example, by increasing the amount of zero-emission RES that shift the merit-order supply curve right-ward, would have an impact on emissions captured by the MEF.

In a seminal paper, Hawkes (2010) calculated the MEF for the GB electricity system.[1] Since then, the emerging literature on MEF has taken into account several aspects of the analysis. A large body of literature has estimated MEFs in several countries and focused on the relationship between the MEF and AEF, including Siler-Evans et al. (2012) and Ryan et al. (2016), for several systems across the US, Bettle et al. (2006) and Jansen et al. (2018) for the British system, Voorspools and D'haeseleer (2000) for Belgium, Oliveira et al. (2019) for UK, France and Spain, Koffi et al. (2017), Climate Transparency (2017) and Beltrami et al. (2020a) for Italy.

Other scholars have discussed the role and importance of RES for the correct calculation of MEF. Indeed, Hawkes (2010, 2014) disregards RES contribution in its calculation on the basis of the assumption that they would not act as the marginal technology in the GB system, given their near-zero operational marginal cost and non-dispatchability. However, he admits that this is one of the main limitation of his approach. This criticism is shared by Li et al. (2017), who highlight the drawbacks of relying only on emitting sources (conventional non-renewable sources of energy) for the computation of MEFs. They also measure the difference in the evaluation of MEFs for the Midcontinent Independent System Operator (MISO) when RES are included in the analysis.

[1] In a subsequent paper, Hawkes (2014) further addresses the issue by distinguishing between the short-run MEFs, which rely on the assumption of a given structure of the power supply and the long-run MEFs, which include changes in the power supply structure.

Beltrami et al. (2020b) found evidence that, in the Italian power exchange market, RES can be marginal and that market operators often bid non-zero prices for electricity generated by RES, which cast doubts on the correctness of excluding RES from the calculation of MEFs in a system with a large penetration of renewables such as the Italian electricity market.

Another stream of literature has focused on the correct empirical strategies to evaluate MEF with regard to the specific nature of power supply and the consequences in terms of data analysis. Indeed, the day-ahead electricity market is composed of 24 hourly sub-auctions, in which power plants are dispatched on the basis of the merit-order and the equilibrium quantities and prices are determined for each hour of the day. Hawkes (2010) follows the so-called inter-day approach, for which 48 independent time series are constructed[2], each one indexed to a given half hour of the day. The inter-day approach is often adopted in the empirical literature on electricity markets, in particular in those analyses that explicitly focus on the different supply structure in each hour.[3]

However, the inter-day approach has the drawback of treating consecutive hours as independent. It is well-known in power system analysis that several dynamic constraints in power system operations give rise to non-convexities in production costs (see Cretì and Fontini, 2019, Ch. 5). Several types of power plants, such as thermal power plants with rump constraints or large hydropower or pumped-storage plants, have operational constraints such that their operations depend on what has happened in previous hours or is expected to happen in subsequent ones. Treating the hours as independent causes a loss of critical information on the dynamic nature of power production. For this reason, an intra-day approach has been developed (Beltrami et al., 2020a) to consider the time structure of power markets as part of a single time series. By doing so, the linkages between hours are fully respected within the time series.

The intra-day approach has the further advantage of increasing the data availability for the considered time period. This allows to measure the MEF for short periods of time, while the inter-day approach needs to consider longer time spell which can be problematic for the estimation of

[2]The GB power market is based on half-hourly settlement periods.

[3]For instance, this approach might be useful when studying the impact of PV (Photo Voltaic) generation on power markets as it is possible to consider the data at midday (the hour of the day with the highest solar irradiation) and compare it with data from night hours. Other hours are considered (such as hour 11 and 18) when the analysis is focused on the consequences of ramping up and down of RES, which is observed in the so called *duck-curve* phenomenon (see Lazar, 2016).

MEF, since structural changes in power supply can occur overtime. Nevertheless, the intra-day approach requires a complex data analysis such that simple econometric techniques (such as OLS, as in Hawkes 2010, 2014) cannot be employed. Power data have a periodic daily and weekly patterns that need to be depurated from the original time series. Moreover, the resulting data often show a non-stationary patterns with past lags influencing the observations over time. This requires the use of sophisticated econometric techniques to obtain reliable estimates.

In this paper, we estimate the Italian MEFs, including RES; we follow the intra-day approach, distinguishing across Italian market zones and further disaggregating the data into quarterly time period. Our contribution to the literature is multifold.

The first contribution is the calculation of the zonal MEFs of Italy and the comparison with the zonal AEF. Beltrami et al. (2020a) calculated the Italian MEFs without RES generation, at the national level. However, the Italian power market is divided into six zones, each with its specific power supply structure. The national MEF corresponds to a weighted average of the zonal MEFs. From the policy perspective, it becomes crucial to analyse the MEFs for each zone, in order to assess the relative impact of the different power supply structures on the local MEFs and to identify the areas where decarbonization effort would be most effective.

The second contribution of this paper is to include RES generation in the calculation of zonal MEFs. The impact of RES on emission reduction due to displacement of generation from polluting plants has been studied very recently (Di Cosmo and Valeri, 2018). As mentioned, the Italian power supply is quite different across zones, with a large penetration of hydro power in the North zone, a large contribution of PV and Wind in the South and Sicily, and no gas-fired plants in Sardinia. RES are marginal in different periods and across zones; moreover, power producers who rely on RES do not necessarily bid at zero prices. For these reasons we explicitly take into account RES generation in the calculations and compare the result at the aggregated (national) level with the one reported in Beltrami et al. (2020a) to reveal the overestimation error that is made when RES generation is omitted from the calculation.

A third contribution refers to the evaluation of MEF for each quarter of the year, that is possible only by adopting the intra-day approach. The matrix of quarterly MEFs for each zone represents a tool that can be useful in planning policies that have seasonal and regional impacts on load. For instance, energy saving policies aimed at the substitution of gas-fired heating systems with heat pumps would have the highest impact on winter load in the colder North zone. The first quarter MEF of the North zone provides a more accurate measure of its impact on emission than the yearly one. Similarly, a policy in support of PV generation would have the highest impact on quarter 3 (Q3) in southern Italian zones.

Finally, we contribute to the empirical literature on MEF estimation under the intra-day approach. We show that the time series of the seasonally adjusted data exhibit long memory, and this requires using a fractional cointegration approach (FCVAR), that extends the standard cointegration case by allowing for the linear cointegration of the variables with a fractional order of integration (i.e. less than one). This captures the long-memory nature of the series, and provides reliable estimates of the MEFs. To the best of our knowledge, this paper is the first that applies the FCVAR methodology to estimate MEFs.

The paper is structured as follows. Section 2 describes the methodology used. Section 3 describes the dataset. Results are illustrated in Section 4. Final remarks and policy implications are presented in Section 5.

2 Methodology. A Fractional Cointegration VAR model

The identification strategy adopted in this paper builds on previous works by Hawkes (2010), Callaway and Fowlie (2009), Carson and Novan (2013), Li et al. (2017) and Beltrami et al. (2020a).

The simplest model to estimate the relation between electricity generation (G) and carbon emissions (E) is a regression model with just one regressor with coefficients estimated by OLS. This basic model is illustrated in Equation (1):

$$\Delta E_h = \beta \Delta G_h + \varepsilon_h \tag{1}$$

where the subscript h is the hourly settlement period , $\Delta E_h = E_h - E_{h-1}$, $\Delta G_h = G_h - G_{h-1}$ and ε represents the error term of the regression, which includes all the possible sources of variation. The

β coefficient provides an estimate of the MEF. The first contribution of our analysis with respect to the existing literature is the introduction of additional information by studying the spatial (zonal) heterogeneity. Hence, we separate the data into zones.[4] Therefore, Equation (1) is expanded to the following formalisation:

$$\Delta E_{h,z} = \beta \Delta G_{h,z} + \varepsilon_{h,z} \tag{2}$$

with $z = 1, \ldots, Z$ and Z is the total number of zones in the market.[5]

We add RES to the calculation of $\Delta G_{h,z}$. Hence, we identify it as:

$$\Delta G_{h,z} = (G^{e}_{h,z} + G^{ne}_{h,z}) - (G^{e}_{h-1,z} + G^{ne}_{h-1,z}) \tag{3}$$

where G^e stands for the generation from emitting sources (i.e. thermoelectric power plants fueled by natural gas, coal or oil) whereas G^{ne} is the generation from RES (non-emitting).

We interpret the estimated MEF as the carbon intensity of the fuel that is mostly used at the margin in the day-ahead (DA) market, and this represents the technological structure of the power plants typically operating at the margin in the Italian system in a given zone.

The empirical approach to be used depends on the nature of the data. To identify possible order of integration of the processes generating the time series involved in the estimation of MEF, we preliminarily perform the following unit roots and stationarity tests: Augmented-Dickey-Fuller (ADF), Dickey-Fuller Generalized Least Squares (DF-GLS), Phillips-Perron (PP) and Zivot-Andrews (ZA) unit root tests, together with the Kwaitkowski-Phillips-Schmidt-Shin (KPSS) stationarity test.[6]

We then estimate the AutoCorrelation Function (ACF) and the Partial AutoCorrelation Function (PACF) on the seasonally-adjusted time series for each zone. Hourly and weekly seasonality has been removed in order to correctly identify the data generating process, making sure that the deterministic cycles do not affect the dynamics of the time series.

[4] We follow the characterization provided by the Italian TSO, Terna, with the identification of six physical zones (North, Center North, Center South, South, Sicily and Sardinia).

[5] In this analysis, the cross-zonal estimation of MEF is not included in order to keep the model as simple as possible.

[6] We also perform a robust version (RKPSS) of the KPSS test which was suggested by Pelagatti and Sen (2013).

As pointed out by Jones et al. (2014) and Beltrami et al. (2020a), the presence of fractional integration in a process is usually verified when both the null hypothesis of unit root and the null hypothesis of stationarity are rejected. When multiple time series are analysed, it becomes necessary to explore the potential presence of a fractional co-integration relationships between the single time series.

Cointegration is often defined as a stationary relationship between nonstationary variables. The classical cointegration approach dates back to the work by Engle and Granger (1987), where a basic OLS estimator is exploited for the detection of the cointegration parameter β, incorporating the co-movement between non-stationary variables.

The concept of fractional cointegration was originally presented in the same paper by Engle and Granger (1987) which generalised the traditional cointegration theory based on unit roots. As a matter of fact, the concept of standard cointegration is a specific case in which the order of cointegration is an integer value. However, the more general theoretical framework of fractional cointegration allows the memory parameter to take fractional values. Specifically - following the indications by Carlini and Santucci de Magistris (2019), in the framework of long-memory processes, fractional cointegration allows linear combinations of $I(d)$ processes to be $I(d-b)$, with $d, b \in R_+$with $0 < b \leq d$. This framework allows for the existence of common stochastic trends integrated of order d, with short-term perturbations from the long-run equilibrium integrated of order $d-b$.

We introduce the $FCVAR_{d,b}$ model for our variables set[7] X_t:

$$\Delta^d X_t = \alpha\beta' \Delta^{d-b} L_b X_t + \sum_{i=1}^{k} \Gamma_i \Delta^d L_b^i X_t + \varepsilon_t \tag{4}$$

where, in detail:

- d is the fractional parameter, or the order of fractional integration of an observable time series process which we allow to assume any real value[8];

[7]We deal with a two-dimensional vector of variables $X_t = (X_{1,t}; X_{2,t}) = (E_t; G_t)$, with $t = 1, \ldots, T$. Although we omit the subscript z, the model applies individually to each zone z of the Italian market.

[8]A generic stochastic process y_t is both stationary and invertible if all roots of AR and MA polynomials lie outside the unit circle and $|d| < 0.5$. The process is nonstationary for $|d| \geq 0.5$, as it possesses infinite variance; see Granger and Joyeux (1980).

- b is the the degree of fractional cointegration obtained by the linear combination of $I(d)$ variables, or the cointegration gap. We assume here[9] $d = b$;

- $\Delta^d = (1 - L)^d$ is the fractional difference operator;

- the operator $L_b = 1 - \Delta^b$ is the fractional cointegration lag operator;

- α and β are the long-run parameters which are $p \times r$ matrices with $0 \leq r \leq p$. p defines the dimension of the time series X_t, thus $p = 2$ in our case. The cofractional rank is defined by r;

- k is the number of lags in the system;

- $\mathit{\Gamma}$ is a $p \times p$ matrix leading short-run dynamics;

- ε_t is a p-dimensional vector of $i.i.d.(0, \Omega)$ disturbances.

Note that under the special assumption $d = b = 1$ the standard cointegration definition by Engle and Granger (1987) is obtained. Thus, the $FCVAR_{d,b}$ can be considered a general model comprising both fractional cointegration and classic cointegration. A final remark regards the rank of the long-run matrix $\Pi = \alpha\beta^{'}$ which is restricted to $r \leq p$.

A simple way to attenuate the influence of pre-sample values on the estimated coefficients of the model is to adopt the specification suggested by Johansen and Nielsen (2016), also applied in Jones et al. (2014). The basic idea is to de-mean the observed variables by including a restricted constant. This is a shift (or "truncation") of the variables by a level parameter, which can be used for empirical results. Therefore, Equation (4) becomes the following:

$$\mathit{\Delta}^d(X_t - \mu) = \alpha\beta'\mathit{\Delta}^{d-b}L_b(X_t - \mu) + \sum_{i=1}^{k} \mathit{\Gamma}_i \mathit{\Delta}^d L_b^i (X_t - \mu) + \varepsilon_t \tag{5}$$

The restricted constant μ can be retrieved from $-\rho' = \beta'\mu$, which is the expression denoting the mean level of the stationary long-run cointegration relationships. In practice, to obtain the

[9] The model built under this constraint is called *"baby model"*. This restriction avoids the identification issues which are well described by Carlini and Santucci de Magistris (2019). As a matter of fact, when the number of lags k is unknown, the FCVAR model is not globally identified and this results in a multiplicity of not-identified sub-models for any possible combination of the parameters d and b.

intercept of the long-run cointegrating relationship, we compute the vectorial product $\hat{\beta}'\hat{\mu}$, i.e. the product between the estimated beta coefficients and the estimated level parameters.
This specification is convenient because it allows to reduce the bias arising from the pre-sample observations of X_t, thus accommodating the fact that the first value of the process is different from zero.

The coefficients of the model defined in Equation (5) are estimated by maximum likelihood, conditional on N initial observations[10] by maximising the following log-likelihood function:

$$logL_T(\lambda) = -\frac{Tp}{2}\left[log(2\pi)+1\right] - \frac{T}{2}log\left\{det\left[T^{-1}\sum_{t=N+1}^{T+N}\varepsilon_t(\lambda)\varepsilon_t(\lambda)^{'}\right]\right\} \tag{6}$$

where *det* stands for determinant. For further details, see Johansen and Nielsen (2016). The maximisation of Equation (6) produces the optimal estimates of parameters $(\alpha, \beta, \Gamma, \rho, \xi)$.

The Matlab code used to estimate the coefficients of the FCVAR model[11] takes into account the issue discussed by Johansen and Nielsen (2012) regarding the distribution of the conditional maximum likelihood parameter estimates. This issue is relevant to obtain correct outcomes from the cointegration rank test, which has to be based on alternative critical values. Specifically, for the estimation of the model, the asymptotic theory is standard when $b < 0.5$ and the limit distributions of the estimated parameters follow either a Gaussian or a χ^2 distribution. For the case of $b > 0.5$, the asymptotic theory is non-standard and assumes that the limit distribution of the likelihood ratio test for the cointegration rank is a functional of fractional Brownian motion of type II[12].

Conditioning on the optimal selection of k derived from a lag selection test based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), we proceed by running the Likelihood Ratio (LR) trace test to identify the order of cointegration of our system. The Likelihood

[10] A conditional approach for the maximum likelihood estimation simplifies the computational procedure by reducing the numerical problem to least squares. Therefore, it is more convenient than the unconditional approach, which is computationally complex (see Johansen and Nielsen, 2012).

[11] The main functions of the Matlab code, which has been adapted to our case, have been provided by Carlini and Santucci De Magistris, who are kindly acknowledged by the authors.

[12] To obtain the simulated p-values of our cointegration rank test automatically, we made use of a separately installed program to be uploaded in Matlab, called **fdpval**. This is the C++ implementation of a Fortran program used to obtain simulated p-values from MacKinnon and Nielsen (2014), who simulated numerical cumulative distribution functions (CDFs) as functions of b_0.

Ratio test is sequentially repeated until it is not possible to reject the null hypothesis.

Once the matrix $\Pi = \alpha\beta^{'}$ is defined, we test the following hypothesis:

$$\begin{aligned} \mathscr{H}_r &: rank(\Pi) = r \\ \mathscr{H}_p &: rank(\Pi) = p \end{aligned} \tag{7}$$

In particular, the LR test starts from the lowest possible rank, i.e. 0, and evolves until the highest possible rank, i.e. p. The LR test stops once the null hypothesis cannot be rejected[13]. This corresponds to the correct rank of the matrix, i.e. the number of potential cointegrating relationships in the system.

To check for potential statistical endogeneity we can impose some restrictions motivated by the economic theory, especially on the adjustment parameters α and cointegration vector β. Most importantly, we are able to perform a strong robustness check of our model.

The idea is to test several restricted models (with restrictions on estimated parameters) and derive some conclusions about the goodness of the overall approach. We formulate the following structure of hypothesis tests:

$$\begin{aligned} R_{\psi}\psi &= r_{\psi} \\ R_{\alpha}vec(\alpha) &= 0 \\ R_{\beta}vec(\beta^*) &= r_{\beta} \end{aligned} \tag{8}$$

where ψ is the set of restrictions imposed on (d, b) and $\beta^* = (\beta', \rho')'$, where ρ represents the restricted constant term. Most importantly, the first hypothesis test is a valid check for the correct identification of our econometric model[14].

The restrictions formalised in Equation (8) enable to validate our model and are described here:

1. $\mathscr{H}_d^1$: the key hypothesis to test is whether it is correct to adopt a $FCVAR_{d,b}$ model against the alternative of using a standard CVAR (Cointegrated Vector Auto Regressive) model;

[13]This has to be read sequentially. Starting from $rank = 0$, we firstly test the hypothesis that the rank of the matrix is null against the alternative that the rank is equal to p. If this null hypothesis is rejected, then the second step is carried out by testing if $rank = 1$ against the alternative that is equal to p, and so on.

[14]We here use the same notation as from Nielsen and Popiel Ksawery (2018).

2. $\mathscr{H}_{\beta}^{1}$: are emissions part of the cointegrating relationship? We test the hypothesis on the long-run fractional cointegrating coefficient β to check that the estimated co-integrating relationship has the expected direction of causality, i.e. that electricity generation causes the changes in carbon emissions;[15]

3. $\mathscr{H}_{\alpha}^{1}$ and $\mathscr{H}_{\alpha}^{2}$: test hypothesis on α. This is the long-run exogeneity test which is carried out on each variable included in the model.[16]

The required restrictions obtained from the hypothesis tests are then incorporated in a second step. The model needs to be restricted and re-estimated to improve the efficiency of the estimated MEF coefficients and that of the estimated fractional parameters d and b. In this case, the same procedure described in this Section is repeated for the formalisation of a new correct long-term equilibrium relationship.

3 Data

3.1 Data description

We construct six independent zonal datasets, for the year 2018, that are described as follows.

1. **Estimated CO_2 emissions (tCO_2).** They have been obtained by applying the fuel consumption model to Italian power plants, as described by Beltrami et al. (2020b). Plant-level data have been aggregated to obtain the hourly amount of carbon emissions from electricity output in each zone. Data for zone South are missing because, due to its market configuration as specified by the Italian TSO, it was not possible to estimate the corresponding emissions. In the period under examination, the biggest thermal production plants physically located in the regions belonging to the South zone were not included in the market zone, but were kept separate in two limited production poles (Brindisi and Foggia), due to physical permanent

[15]This test was applied by Jones et al. (2014) to check the validity of including additional exogenous regressors in the main equation.

[16]The long-run exogeneity tests verify the magnitude of short-term adjustments of each variable to shocks in the long-run equilibrium relationship.

transmission congestion that impeded their merge into the physical South zone. This implies that the South zone lacks emission data for the largest power plants.

2. **Electricity generation (MWh).** Generation data have been obtained as the aggregation of hourly accepted bids of electricity supply in each physical zone in the Italian day-ahead market, by considering both thermoelectric polluting generation and carbon-neutral RES generation. The source of this information is GME (Gestore del Mercato Elettrico), the Italian Power Market Operator.

Table 1 displays summary statistics of emissions and generation data for the six Italian physical zones. They differ in load, emissions and electricity generated from thermoelectric and RES power plants. The North zone represents nearly 75% of total national load[17] and generates 58% of total national electricity. The largest share of RES production in this zone is from hydroelectric power plants (located in Trentino, Piedmont, Lombardy and the Aosta Valley) and non relevant RES (NRRES)[18] (particularly small PV plants). Center South is the second contributor in terms of load and electricity generation. Power demand is mostly met by wind turbines, NRRES and hydro (several pump storage hydro plants). The Center North zone strongly relies on geothermal power production plants (nearly 32% of total zonal power production). The residual renewable generation mix comprises NRRES and hydro. The South zone is similar to the Center South zone, except for its higher production from large-scale wind turbines, which - together with small-scale solar and wind units - provides nearly 73% of total zonal power generation. The amount of load and power generation in Sicily and Sardinia is relatively small compared to the other zones. Nevertheless, the power generation from large-scale wind farms and small-scale PV is significant and contributes in each zone to, respectively, 44% and 23% of total zonal power production.

[17]Data for the virtual production zones, the limited production poles and the interconnectors are not shown in Table 1.

[18]The definition of non relevant RES (NRRES) depends on the specific Italian encoding of plants and RES subsidy rules. Small-scale renewables, i.e. renewable energy plants smaller than 10 MW (mostly connected at distribution level), receive subsidies by means of a purpose-built public company, called GSE (Gestore del Sistema Energetico, in Italian - Energy System Manager). The individual supply of NRRES is collected by the GSE, aggregated and placed on the market at zero price. This category includes various small-scale RES plants. Depending on the zone, the majority of these plants are either small PV or run hydro. However, there are also other types, such as small-scale wind, or small old co-generation plants (even not renewables). Further disaggregation is not possible and therefore they are classified as a single category within the RES group.

Table 1: *Zonal configuration of the Italian power market: yearly load, yearly calculated emissions, accepted generation by RES type and Average Emission Factors (AEF).*

	North	Center North	Center South	South	Sicily	Sardinia	Total
Yearly load	164,518,420	31,081,787	45,943,518	23,633,923	17,680,705	8,974,450	291,832,803
Yearly emissions	36,625,956	2,291,111	11,723,255	-	799,100	3,340,594	54,780,226
Yearly accepted generation	128,956,528	18,560,878	28,740,029	18,808,488	10,817,844	11,130,253	217,014,020
Of which: (1) Solar	417,426	23,105	420,040	246,598	28,647	84,537	1,220,353
(2) Wind	34,502	167,446	2,508,066	8,209,909	2,899,634	1,634,064	15,453,621
(3) Biomass	1,210,190	71,489	93,618	1,667,322	129,618	315,496	3,487,733
(4) Waste	2,832,667	103,038	1,421,965	971,108	-	-	5,328,778
(5) Geothermal	-	5,718,643	-	-	-	-	5,718,643
(6) Non-Relevant RES	26,987,673	4,044,144	4,594,367	5,712,600	2,079,473	952,958	44,371,215
(7) Hydro	27,081,187	2,318,525	3,147,286	1,423,159	68,198	300,481	34,338,836
AEF	0.2840	0.1234	0.4078	-	0.0738	0.3001	0.2524

Note: Values of load and generation are in MWh (year 2018). Values of emissions are expressed in tCO_2. Values of AEFs are in tCO_2/MWh. Source: our processing of GME data.

Figure 1 displays the time dynamics of emissions and generation for the North zone (the plots for the other zones are available upon request). The figure shows the typical daily and weekly behaviour of power time series due to the seasonal patterns, which are related to consumption levels that drops at night and in the weekends. Therefore, we must perform a preliminary adjustment for each zone, by extracting and removing the deterministic component of time series variables. The preliminary adjustment is carried out by constructing 6 dummy variables to account for the day of the week (omitting the 7th dummy related to Sundays to avoid multi-collinearity) and 23 hourly dummies to account for the settlement periods. We then regress each time series on the chosen dummy variables and store the residuals. The seasonally-adjusted time series variable results from the sum of the residuals and the estimated constant[19].

3.2 Preliminary data analysis

The autocorrelation and partial autocorrelation functions (ACF and PACF) plots of seasonally adjusted time series by zone are shown in the Appendix A. They show evidence of long memory data generation processes.

Table 2 reports the unit root and stationarity tests for each physical zone, both for carbon

[19]From now on, we will focus only on the seasonally-adjusted time series, i.e. the stochastic component of the observed time series.

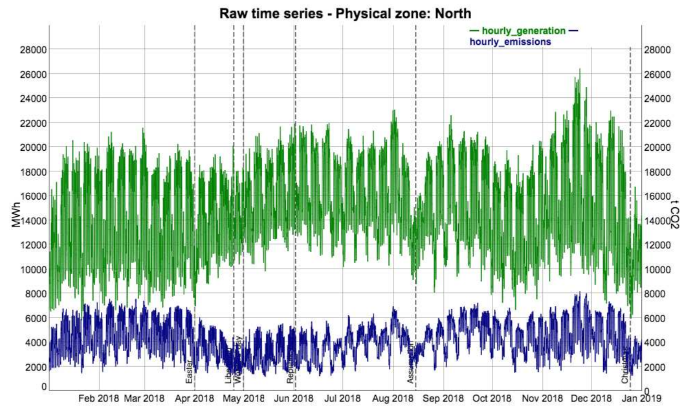


Figure 1: *Hourly emissions (tCO2) in green and hourly generation (MWh) in blue for the market zone North in 2018. Source: own elaboration.*

emissions and electricity generation. The comparison of the two types of tests leads to opposite conclusions regarding the nature of the time series we analyse. On one hand, the unit root tests mainly reject the null hypothesis of having a unit root in the variable. On the contrary, based the KPSS and RKPSS stationarity tests that it is not possible to accept the null of stationarity. This mixed outcome, similar to the results in Jones et al. (2014) and Beltrami et al. (2020a), is the main indicator of the presence of a fractional order of integration. We thus proceed to analyse the potential for fractional cointegration relationships between our emissions and generation.

The results on MEF estimates are based on the methodology set out in Section 2. The implementation of the model suggested in Equation (5) follows the procedure suggested by Johansen (2008) and Carlini and Santucci de Magistris (2019).

In the first step, the unrestricted FCVAR model is estimated. Preliminarily to the estimation, we conduct the lag-order selection based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). According to order selection results, we proceed with the Likelihood Ratio (LR) test to detect the cointegrating rank of the model. If the hypothesis $r = 0$ is accepted, the procedure is stopped and the MEF is estimated by using the same procedure as in Beltrami et al. (2020a).[20]

[20] The final model is a regression with AR(F)IMA errors. See Equation (4) and (5) of Beltrami et al. (2020a).

Table 2: *Unit root and stationarity tests by each physical zone for the seasonally adjusted emissions (E_t) and generation (G_t) time series. Sample period: 2018.*

	North		Center North		Center South		Sicily		Sardinia	
	E_t	G_t	E_t	G_t	E_t	G_t	E_t	G_t	E_t	G_t
ADF_{stat}	-6.98	-4.70	-14.01	-4.23	-13.27	-8.38	-20.38	-4.45	-5.56	-2.58
ADF_{cval}	-1.95	-1.95	-1.95	-1.95	-1.95	-1.95	-1.95	-1.95	-1.95	-1.95
$DFGLS_{stat}$	-12.19	-9.59	-16.16	-19.06	-15.75	-21.23	-23.24	-11.25	-6.27	-6.29
$DFGLS_{cval}$	-1.94	-1.94	-1.94	-1.94	-1.94	-1.94	-1.94	-1.94	-1.94	-1.94
PP_{stat}	-14.27	-15.95	-22.74	-19.35	-21.39	-20.45	-23.59	-14.35	-7.65	-10.05
PP_{cval}	-2.86	-2.86	-2.86	-2.86	-2.86	-2.86	-2.86	-2.86	-2.86	-2.86
$KPSS_{stat}$	4.29	4.35	2.08	2.51	1.91	1.09	0.26	0.22	2.94	2.73
$KPSS_{cval}$	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
$rKPSS_{stat}$	4.49	4.74	1.93	2.56	1.89	1.12	0.37	0.29	2.88	2.80
$rKPSS_{cval}$	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
ZA_{stat}	-17.79	-18.86	-26.64	-27.72	-24.03	-21.95	-26.42	-14.93	-11.50	-10.18
ZA_{cval}	-5.08	-5.08	-5.08	-5.08	-5.08	-5.08	-5.08	-5.08	-5.08	-5.08

Based on the above hypothesis testing procedure, we conduct several diagnostic checks on the FCVAR estimations to prove the validity of the estimated model. Firstly, we test for serial correlation in the residuals. Secondly, the adoption of the FCVAR framework is validated through the LR test. Thirdly, the direction of the long-run cointegrating relationship is verified by a significance test on β. Lastly, the long-run exogeneity tests are carried out on the α parameters and a restricted FCVAR model is finally estimated by taking into account the restrictions resulting from the previous hypothesis testing procedure.

For the full yearly sample, we do find evidence for the adoption of the FCVAR for all zones from the unit root and stationarity tests except for Center-South and South zones. For the former, the LR cointegrating test does not show any evidence of fractional cointegration relationship, and thus we estimate the MEF according to the methodology described by Beltrami et al. (2020a). For the South zone, as already mentioned, we do not have valid data for the computation of emissions and thus the corresponding MEF cannot be estimated. For zones North, Center North and Sardinia, the statistics (reported in Appendix A) show that the preferred model for the MEF estimation is the Restricted FCVAR, while the unrestricted FCVAR is selected for Sicily.

We then repeat the same analysis at quarterly frequency by zone [21]. In all cases, except for

[21] The preliminary unit root and stationarity tests by quarter and by zone are presented in Appendix A.

Sardinia, the unit root tests reject the null hypothesis, and so does the stationarity test[22]. Therefore, the adoption of the FCVAR approach is justified. Nevertheless, in the quarterly analysis, most cases does not show evidence of fractional (or standard) cointegrating relationship which is rejected by the cointegration rank test. Thus, the integrated econometric approach leads to the calculation of the MEF according to the methodology used in Beltrami et al. (2020a).

For Sardinia, the unit root and stationarity tests lead to more heterogeneous results. As regards emissions, the unit root and stationarity tests suggest the presence of fractional integration for all cases[23]. On the other hand, for the time series of generation, the ADF test accepts the null hypothesis of unit root, in all quarters. In spite of this, the other unit root tests reject the unit root hypothesis and therefore generation variable is kept in levels in all quarters.

4 Results

Table 3 summarises the estimated zonal MEFs for Italy in 2018, according to the integrated empirical methodology presented in Section 2. The full econometric calculations are shown in Appendix A.

Table 3: *Marginal emission factor ($tCO2/MWh$) in 2018 by market zone.*

Zone	MEF
North	0.2018
Center North	0.4236
Center South	0.7022
Sicily	0.1460
Sardinia	0.7189
Italy	**0.3921**

Estimated values of MEFs are highly variable among zones. The highest MEF (0.7189 tCO_2/MWh) is obtained in Sardinia, followed by Center South (0.7022 tCO_2/MWh), Center North (0.4236 tCO_2/MWh), North (0.2018 tCO_2/MWh) and Sicily (0.146 tCO_2/MWh).

[22] Although for Q1 in the Center North the ADF test accepts the null of unit root, the other unit root tests do not accept it. Hence, in this case, there is not sufficient evidence to use the variable in first differences and the variable is kept in levels.

[23] With the same argument of the case of Q1 for Center North, the rejection of the null hypothesis of the ADF test for Q4 does not give sufficient evidence to use first differences. Hence, emissions in Q4 for Sardinia are kept in levels.

Importantly, these results reflect the generation mix in each zone and the dynamic evolution of the carbon intensity of marginal power generators in the market. Sardinia strongly relies on thermoelectric coal-fired plants, which have the highest carbon intensity. Moreover, the result of MEF for Sardinia is more than double the annual AEF (0.3001 tCO_2/MWh), revealing a high carbon intensity of marginal generation in this area which is not reflected in the AEF calculation.

A similar outcome is obtained in the Center South. This zone is strongly relies on conventional power plants (representing more than 50% of the generation mix) which affect the overall carbon intensity of marginal generation. The estimated MEFs are nearly 75% higher than the annual AEF (0.4078 tCO_2/MWh)[24].

The MEF estimated in the Center North zone is roughly four times higher than the annual AEF (0.1234 tCO_2/MWh). In this case, the value of MEF mostly reflects the carbon intensity of gas-fired power plants, which represent nearly 33% of the generation mix. This result is also influenced by the significant contribution of geothermal production to marginal generation, which helps attenuating the overall carbon intensity of the area.

Interestingly, the North zone represents the only case in which the estimated annual MEF (0.2018 tCO_2/MWh) is lower than the computed AEF (0.284 tCO_2/MWh). A similar conclusion was drown in the study by Voorspools and D'haeseleer (2000) using Belgian data. This outcome seems to indicate that the improved dispatchability of RES and their integration with pumped hydroelectric storage (PHS) plants might reduce the actual carbon intensity of marginal generation in the area. Under the merit-order perspective, this result proves the effectiveness of carbon-neutral generation from RES - which has higher priority of dispatch due to its near-zero marginal costs - in the displacement of more expensive carbon-intensive production from thermoelectric units.

At the national level, the average MEF calculated by simply aggregating the data on generation and emissions of the six physical market zones amounts to 0.3921 $tCO2$/MWh. We can compare this figure with the findings of Beltrami et al. (2020a), who estimated the "conventional" Italian MEF, that is excluding RES generation data, ranging between 0.5 and 0.65 $tCO2$/MWh. We can

[24] We also estimate the AEF by removing the generation from RES, thus obtaining a "conventional" measure of the AEF. "Conventional" AEF in the Center South is 0.7106 tCO_2/MWh, thus revealing an almost perfect correspondence with the result of the MEF in Table 3.

see that the inclusion of RES lowers the estimate of the national MEF by nearly 32%.

The national MEF estimated in this paper is in line with the findings obtained by Koffi et al. (2017), who report for 2013 an Italian emission factor from electricity consumption equal to 0.343 $tCO2$/MWh. Similarly, the report by Climate Transparency (2017) for year 2017 indicates that the emission intensity of the Italian power sector ($tCO2$/MWh) is 0.331. Note, however, that our figure does not include the South zone. However, the latter two findings are directly comparable and consistent with our result of the Italian average MEF, given the inclusion of electricity generated by RES. Our approach also has the advantage of taking into account the role of RES in their increasing contribution to marginal generation, especially during peak hours.

Furthermore, we recognise that the measurement of the national MEF reported in Table 3 might be biased due to the omission of the impact of congestions across zones[25]. We thus compute a weighted measurement of the annual MEF at the national level starting from the zonal estimates of MEF. The weights are given by the ratio between each zonal load and the national annual load in 2018, with the omission of the data for South for which the MEF has not been calculated. The weighted MEF for Italy is 0.3268 $tCO2/MWh$.

The analysis is fully replicated for each zone of the market and quarter of 2018, by adopting the most efficient methodology based on the procedure discussed above.

Table 4 summarises the results of the quarterly MEFs by zone in 2018.

Table 4: *Quarterly MEFs in 2018 by zone in the Italian day-ahead electricity market.*

	North	Center North	Center South	Sicily	Sardinia
Q1	0.3937	0.6437	0.7358	0.1476	0.8596
Q2	0.3624	0.3313	0.7535	0.2960	0.8102
Q3	0.2738	0.3297	0.6616	0.3214	0.7953
Q4	0.3340	0.3649	0.6315	0.3015	0.7737

In general, the MEFs are lower during warmer months, due to the higher penetration of RES, especially from the contribution of PV panels at peak daytime hours. In the North zone, the MEF

[25]Intuitively, the simple aggregation of data of the 6 physical zones would create an hypothetical Italian macrozone, which is clearly not the case for a zonal power market that is constrained by a given transmission capacity and results in market splitting situations more than 50% of the times.

reaches its minimum value in Q3, at a level (0.2738 $tCO2/MWh$) lower than the annual AEF of 0.2840 $tCO2/MWh$ but far from the annual computation of the MEF of 0.2018 $tCO2/MWh$. In Center North zone, the MEF in Q3 is halved as compared to the MEF in Q1. It is worth pointing out that the quarterly analysis for this zone confirms the fact that the annual MEF is strongly affected by the high carbon intensity observed in Q1 of 2018. The same evidence is confirmed for the Center South zone, whose MEFs are nearly 50% higher than the AEF. Interestingly, only for Sicily the warmer months display higher MEFs compared to colder months. This result might depend on the role of congestion which often affects the transmission capacity between Sicily and the South zone. Lastly, Sardinia displays the highest MEFs, significantly larger than the estimated annual MEFs.

5 Conclusions

This paper contributes to the existing literature on MEFs by providing the first estimation of zonal CO_2 marginal emission factors for the Italian electricity system.

The motivation behind this work stems from the growing debate on the detection of the best strategies to meet the decarbonisation targets set by the European Union's Climate Law (2020) and the Clean Energy Package (2019) aimed at mitigating climate change, improving energy security and increasing RES.

Previous studies in this area enabled us to fully characterise MEFs, to identify the correct techniques to account for spatial and temporal heterogeneity, and eventually to produce reliable MEF estimations in the Italian context, which may guide zone-specific policy interventions by discriminating among different technological solutions.

The main contribution of this paper is the econometric characterisation of MEFs for each physical zone in the Italian power exchange market, by accounting for the heterogeneous distribution of power plants and their variable penetration across zones. The approach presented here can be extended to other electricity systems organised in market zones. Moreover, the inclusion of generation from both polluting and non-polluting (RES) generators allows us to fully account for the dynamic change in the marginal generation mix across time and space.

Using the yearly sample, we find support for the adoption of the FCVAR in all zones except the Center South, for which we do not identify any long-run fractional cointegrating relationship. We find high variability of MEFs calculations at the zonal level. Sardinia records the highest MEF (0.7189 tCO_2/MWh), followed by the Center South (0.7022 tCO_2/MWh), the Center North (0.4236 tCO_2/MWh), the North (0.2018 tCO_2/MWh) and Sicily (0.146 tCO_2/MWh). We also argue that the result for the North zone indicates the improved dispatchability of RES, their increasing contribution as marginal generators (especially hydroelectric generation) and their effectiveness in crowding out more expensive carbon-intensive thermoelectric generation. However, our quarterly analysis suggests that any well targeted policy intervention should not only consider the heterogeneity of power plants distribution across zones but also the seasonal patterns of consumption (and production) of electricity. The estimated quarterly MEFs enrich the findings for the annual MEFs and show that carbon intensity tends to be higher during the coldest months than during warmer periods. The only exception is represented by the Sicilian zone.

We are aware of some limitations in our analysis. Firstly, we did not include data about import and the carbon intensity of electricity imported through interconnectors. Indeed, the inclusion of this information would shed some light on the carbon footprint of neighbouring countries and its impact on zonal Italian MEFs.

Secondly, we do not provide an estimation of the MEF for the South zone, due to its unusual market design: in 2019, the South zone was redesigned by encompassing also the previously separated limited production poles of Brindisi and Foggia. Once sufficient data points are available in future, it should be possible to extend the analysis to this zone as well.

A potential further step would be to forecast future market scenarios and to calculate long-run MEFs. The largest Italian market operator, ENEL, has committed to mothballing coal power plants from 2022 onwards as the result of the capacity market auctions that took place at the end of 2019. This will clearly have an impact on the MEF of the zones where these plants are located, which is not accounted for in the present analysis.

Despite these limitations, our work sheds light on the spatial and temporal impact that changes in load and power supply have on CO_2 emission levels in Italy. Our approach allows researchers to produce more accurate measurements, compared to AEFs and aggregate MEFs, of the impact that energy and environmental policies have across time and geographical areas. The availability of these more precise measurement tools can support policymakers in the design of interventions which can more effectively achieve the challenging emission targets enshrined in the recent EU legislation, such as the 60% reduction in emissions compared to 1990 level by 2030 (Meles et al., 2020) and net-zero by 2050. Furthermore, the opportunity to assess the impact of policy measures across different geographical areas will become more important in the near future as energy systems are moving towards a more decentralised structure with increasing localised generation and the establishment of local markets for energy communities and individual consumers who will be equipped with generation and storage assets.

Acknowledgements

Filippo Beltrami gratefully acknowledges Paolo Santucci de Magistris and Federico Carlini from Luiss University for their overall econometric support and the implementation of the Matlab code. Monica Giulietti gratefully acknowledges funding from UKERC (grant FF3/3). Luigi Grossi acknowledges the financial support from the Italian Ministry of Education and University (MIUR: Ministero dell'Istruzione, dell'Università e della Ricerca), award code FFABR 2017.

References

Beltrami, F., Burlinson, A., Grossi, L., Giulietti, M., Rowley, P., and Wilson, G. (2020a). Where did the time (series) go? Estimation of marginal emission factors with autoregressive components. *Energy Economics*, 104905.

Beltrami, F., Fontini, F., and Grossi, L. (2020b). The value of carbon emission reduction induced by renewable energy sources in the Italian power market. *University of Padua mimeo.*

Bettle, R., Pout, C., and Hitchin, E. (2006). Interactions between electricity-saving measures and carbon emissions from power generation in England and Wales. *Energy Policy*, 34(18):3434–3446.

Bretschger, L. and Pittel, K. (2020). Twenty Key Challenges in Environmental and Resource Economics. *Environmental and Resource Economics*, 77(4):725–750.

Callaway, D. and Fowlie, M. (2009). Greenhouse Gas Emissions Reductions from Wind Energy: Location, Location, Location? *AERE Workshop, Washington DC.*

Carlini, F. and Santucci de Magistris, P. (2019). On the Identification of Fractionally Cointegrated VAR Models With the F(d) Condition. *Journal of Business and Economic Statistics*, 37(1):134–146.

Carson, R. T. and Novan, K. (2013). The private and social economics of bulk electricity storage. *Journal of Environmental Economics and Management*, 66(3):404–423.

Climate Transparency (2017). Brown to Green Report 2017 - The G20 Transition to a Low-Carbon Economy. Technical report.

Cretì, A. and Fontini, F. (2019). *Economics of Electricity: Markets, Competition and Rules.* Cambridge University Press, Cambridge.

Di Cosmo, V. and Valeri, L. M. (2018). How Much Does Wind Power Reduce CO2 Emissions? Evidence from the Irish Single Electricity Market. *Environmental and Resource Economics*, 71(3):645–669.

Engle, R. F. and Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2):251.

Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–29.

Hawkes, A. D. (2010). Estimating marginal CO2 emissions rates for national electricity systems. *Energy Policy*, 38(10):5977–5987.

Hawkes, A. D. (2014). Long-run marginal CO2 emissions factors in national electricity systems. *Applied Energy*, 125:197–205.

Jansen, M., Staffell, I., and Green, R. (2018). Daily marginal CO2 emissions reductions from wind and solar generation. In *International Conference on the European Energy Market, EEM*, volume 2018-June. IEEE Computer Society.

Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory*, 24(3):651–676.

Johansen, S. and Nielsen, M. Ø. (2012). Likelihood Inference for a Fractionally Cointegrated Vector Autoregressive Model. *Econometrica*, 80(6):2667–2732.

Johansen, S. and Nielsen, M. Ø. (2016). The role of initial values in conditional sum-of-squares estimation of nonstationary fractional time series models. *Econometric Theory*, 32(5):1095–1139.

Jones, M. E. C., Nielsen, M. Ø., and Popiel, M. K. (2014). A fractionally cointegrated VAR analysis of economic voting and political support. *Canadian Journal of Economics/Revue canadienne d'économique*, 47(4):1078–1130.

Koffi, B., Cerutti, A., Duerr, M., Iancu, A., Kona, A., and Janssens-Maenhout, G. (2017). Covenant of Mayors for Climate and Energy: Default emission factors for local emission inventories. *Joint Research Centre (JRC). Publications Office of the European Union.*

Lazar, J. (2016). Teaching the "Duck" to Fly. *The Regulatory Assistance Project.*

Li, M., Smith, T. M., Yang, Y., and Wilson, E. J. (2017). Marginal Emission Factors Considering Renewables: A Case Study of the U.S. Midcontinent Independent System Operator (MISO) System. *Environmental Science and Technology*, 51(19):11215–11223.

MacKinnon, J. G. and Nielsen, M. Ø. (2014). Numerical distribution functions of fractional unit root and cointegration tests. *Journal of Applied Econometrics*, 29(1):161–171.

Meles, T. H., Ryan, L., and Wheatley, J. (2020). COVID-19 and EU Climate Targets: Can We Now Go Further? *Environmental and Resource Economics*, 76(4):779–787.

Nielsen, M. Ø. and Popiel Ksawery, M. (2018). A Matlab program and user's guide for the fractionally cointegrated VAR model. *Queen's Economics Department Working Paper*, 1(1330):1–47.

Oliveira, T., Varum, C., and Botelho, A. (2019). Econometric modeling of CO2 emissions abatement: Comparing alternative approaches. *Renewable and Sustainable Energy Reviews*, 105:310–322.

Pelagatti, M. M. and Sen, P. K. (2013). Rank tests for short memory stationarity. *Journal of Econometrics*, 172(1):90–105.

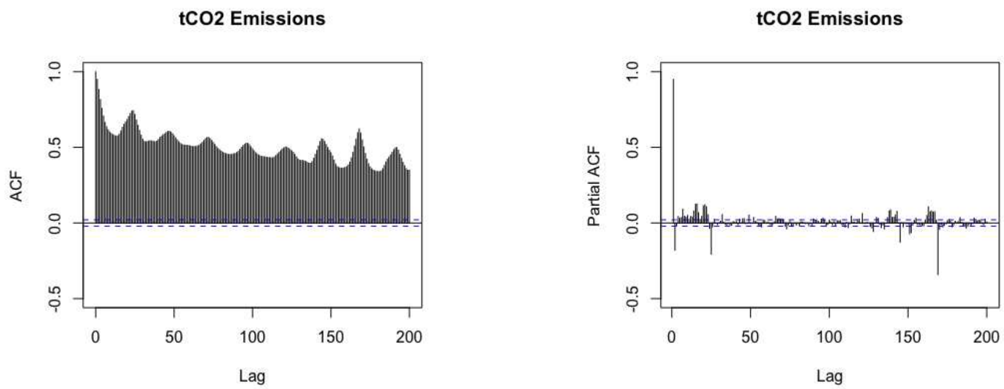
Ryan, N. A., Johnson, J. X., and Keoleian, G. A. (2016). Comparative Assessment of Models and Methods to Calculate Grid Electricity Emissions. *Environmental Science and Technology*, 50(17):8937–8953.

Siler-Evans, K., Azevedo, I. L., and Morgan, M. G. (2012). Marginal Emissions Factors for the U.S. Electricity System. *Environmental Science and Technology*, 46(9):4742–4748.

Voorspools, K. R. and D'haeseleer, W. D. (2000). An evaluation method for calculating the emission responsibility of specific electric applications. *Energy Policy*, 28(13):967–980.

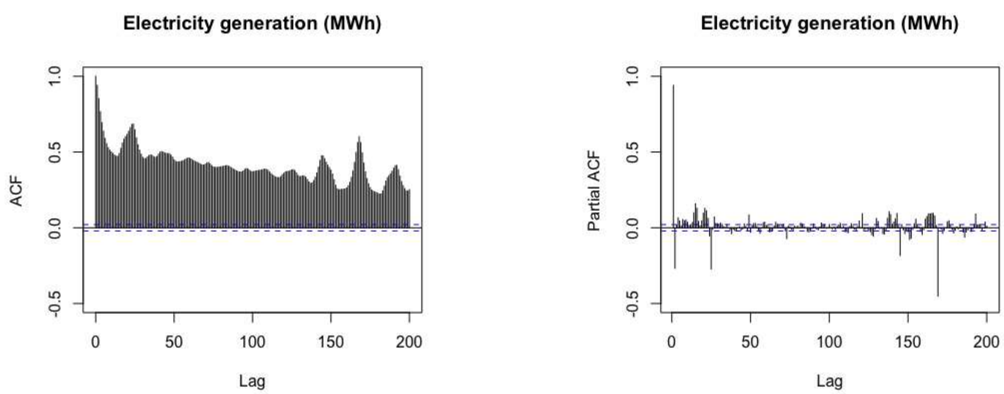
A Appendix

A.1 Zone North



(a) ACF (b) PACF

Figure A1: Zone North, ACF (panel a) and PACF (panel b) for seasonally adjusted emissions.



(a) ACF (b) PACF

Figure A2: Zone North, ACF (panel a) and PACF (panel b) for seasonally adjusted generation.

Table A1: *Unit root and stationarity tests for North for the seasonally adjusted emissions and generation time series. Sample period: quarters of 2018.*

	quarter	ADF_teststat	ADF_cval	DFGLS_teststat	DFGLS_cval	PP_teststat	PP_cval	KPSS_teststat	KPSS_cval	rKPSS_teststat	rKPSS_cval	ZA_teststat	ZA_cval	variable	zone
1	1	-3.46	-1.95	-6.03	-1.94	-10.63	-2.86	1.34	0.15	1.25	0.15	-11.94	-5.08	Emissions	North
2	2	-5.76	-1.95	-8.98	-1.94	-11.32	-2.86	1.42	0.15	1.45	0.15	-12.55	-5.08	Emissions	North
3	3	-2.94	-1.95	-9.67	-1.94	-8.42	-2.86	1.39	0.15	1.30	0.15	-10.61	-5.08	Emissions	North
4	4	-3.12	-1.95	-7.95	-1.94	-8.62	-2.86	1.74	0.15	1.68	0.15	-10.08	-5.08	Emissions	North
5	1	-2.88	-1.95	-7.54	-1.94	-11.06	-2.86	0.98	0.15	0.86	0.15	-12.21	-5.08	Generation	North
6	2	-1.99	-1.95	-8.30	-1.94	-9.45	-2.86	1.05	0.15	1.08	0.15	-12.40	-5.08	Generation	North
7	3	-2.19	-1.95	-6.16	-1.94	-10.43	-2.86	0.67	0.15	0.56	0.15	-10.99	-5.08	Generation	North
8	4	-2.49	-1.95	-6.16	-1.94	-8.16	-2.86	3.18	0.15	3.31	0.15	-10.36	-5.08	Generation	North

Table A2: *Summary findings for the zone North - FCVAR results. Note: k = 4 from the lag selection test.*

Rank tests:

Rank	$\hat{d}$	Log-likelihood	LR statistic	p-value
0	0.569	-128603.562	37.341	0.000
1	0.542	-128586.063	2.344	0.333
2	0.537	-128584.891	—	—

Unrestricted model:

$$\Delta^{\hat{d}}(\begin{bmatrix} E_t \\ G_t \end{bmatrix} - \begin{bmatrix} 1605.680 \\ 5516.842 \end{bmatrix}) = L_{\hat{d}} \begin{bmatrix} -0.44 \\ -0.167 \end{bmatrix} v_t + \sum_{i=1}^{4} \widehat{\Gamma}_i \Delta^{\hat{d}} L_{\hat{d}}^i (X_t - \hat{\mu}) + \hat{\epsilon}_t$$

$$\hat{d} = \underset{(0.011)}{0.206}, Q_{\hat{\epsilon}}(24) = \underset{(pv=0.000)}{2456.580}, log(\mathscr{L}) = -128553.739$$

Long-term equilibrium relationship:

$$E_t = 502.31 + 0.2G_t + v_t \tag{9}$$

Hypothesis tests:

	$\mathscr{H}_d^1$	$\mathscr{H}_\beta^1$	$\mathscr{H}_\alpha^1$	$\mathscr{H}_\alpha^2$
df	1	1	1	1
LR	866.252	51.327	18.842	1.042
p-value	0.000	0.000	0.000	0.307

Restricted model:

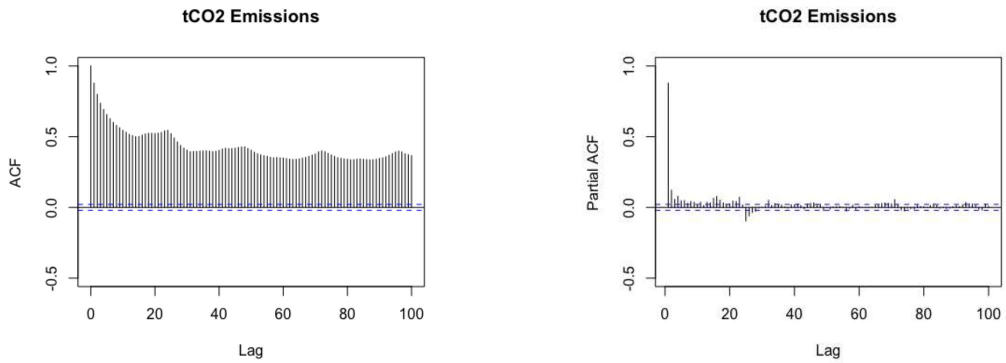
$$\Delta^{\hat{d}}(\begin{bmatrix} E_t \\ G_t \end{bmatrix} - \begin{bmatrix} 1583.383 \\ 5485.098 \end{bmatrix}) = L_{\hat{d}} \begin{bmatrix} -0.3973 \\ 0.000 \end{bmatrix} v_t + \sum_{i=1}^{4} \widehat{\Gamma}_i \Delta^{\hat{d}} L_{\hat{d}}^i (X_t - \hat{\mu}) + \hat{\epsilon}_t$$

$$\hat{d} = \underset{(0.004)}{0.203}, Q_{\hat{\epsilon}}(24) = \underset{(pv=0.000)}{2456.559}, log(\mathscr{L}) = -128554.260$$

Long-term equilibrium relationship:

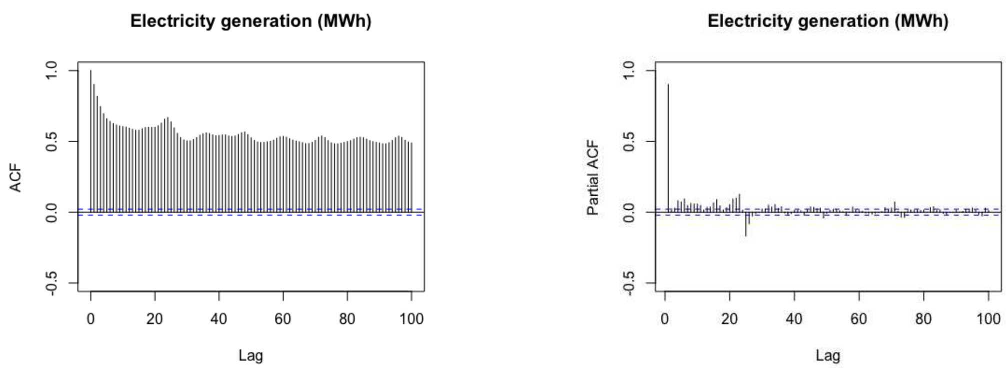
$$E_t = 476.49 + 0.2018G_t + v_t \tag{10}$$

A.2 Zone Center North



(a) ACF (b) PACF

Figure A3: Zone Center North, ACF (panel a) and PACF (panel b) for seasonally adjusted emissions.



(a) ACF (b) PACF

Figure A4: Zone Center North, ACF (panel a) and PACF (panel b) for seasonally adjusted generation.

Table A3: *Unit root and stationarity tests for Center North for the seasonally adjusted emissions and generation time series. Sample period: quarters of 2018.*

	quarter	ADF_teststat	ADF_cval	DFGLS_teststat	DFGLS_cval	PP_teststat	PP_cval	KPSS_teststat	KPSS_cval	rKPSS_teststat	rKPSS_cval	ZA_teststat	ZA_cval	variable	zone
1	1	-5.90	-1.95	-7.03	-1.94	-14.63	-2.86	0.33	0.15	0.32	0.15	-16.42	-5.08	Emissions	Center North
2	2	-8.15	-1.95	-9.90	-1.94	-12.96	-2.86	0.34	0.15	0.31	0.15	-15.08	-5.08	Emissions	Center North
3	3	-7.32	-1.95	-8.26	-1.94	-10.73	-2.86	1.05	0.15	1.04	0.15	-17.35	-5.08	Emissions	Center North
4	4	-7.21	-1.95	-5.79	-1.94	-9.16	-2.86	1.51	0.15	1.43	0.15	-14.61	-5.08	Emissions	Center North
5	1	-1.66	-1.95	-7.80	-1.94	-13.14	-2.86	0.79	0.15	0.84	0.15	-15.61	-5.08	Generation	Center North
6	2	-1.95	-1.95	-11.90	-1.94	-13.59	-2.86	0.96	0.15	0.97	0.15	-15.86	-5.08	Generation	Center North
7	3	-2.32	-1.95	-6.35	-1.94	-11.30	-2.86	0.49	0.15	0.50	0.15	-16.90	-5.08	Generation	Center North
8	4	-2.77	-1.95	-8.47	-1.94	-10.71	-2.86	1.05	0.15	1.03	0.15	-15.95	-5.08	Generation	Center North

Table A4: *Summary findings for the zone Center North - FCVAR results. Note: k = 4 from the lag selection test.*

Rank tests:

Rank	$\hat{d}$	Log-likelihood	LR statistic	p-value
0	0.512	-98563.369	30.31	0.000
1	0.482	-98548.358	0.288	0.591
2	0.481	-98548.214	—	—

Unrestricted model:

$$\Delta^{\hat{d}}(\begin{bmatrix} E_t \\ G_t \end{bmatrix} - \begin{bmatrix} 103.942 \\ 1478.090 \end{bmatrix}) = L_{\hat{d}} \begin{bmatrix} -0.092 \\ 0.673 \end{bmatrix} v_t + \sum_{i=1}^{4} \widehat{\Gamma}_i \Delta^{\hat{d}} L_{\hat{d}}^i (X_t - \hat{\mu}) + \hat{\epsilon}_t$$

$$\hat{d} = \underset{(0.01)}{0.231}, Q_{\hat{\epsilon}}(24) = \underset{(pv=0.000)}{1369.419}, log(\mathscr{L}) = -98532.904$$

Long-term equilibrium relationship:

$$E_t = -488.77 + 0.401 G_t + v_t \tag{11}$$

Hypothesis tests:

	$\mathscr{H}_d^1$	$\mathscr{H}_\beta^1$	$\mathscr{H}_\alpha^1$	$\mathscr{H}_\alpha^2$
df	1	1	1	1
LR	1068.479	36.906	1.353	15.016
p-value	0.000	0.000	0.245	0.000

Restricted model:

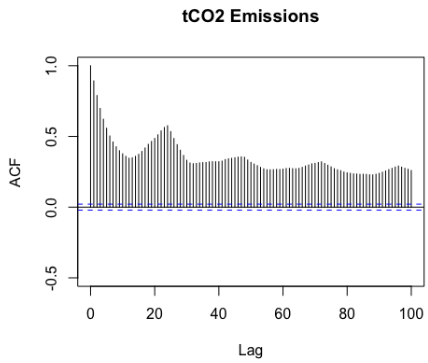
$$\Delta^{\hat{d}}(\begin{bmatrix} E_t \\ G_t \end{bmatrix} - \begin{bmatrix} 119.326 \\ 1531.562 \end{bmatrix}) = L_{\hat{d}} \begin{bmatrix} 0.000 \\ 0.7022 \end{bmatrix} v_t + \sum_{i=1}^{4} \widehat{\Gamma}_i \Delta^{\hat{d}} L_{\hat{d}}^i (X_t - \hat{\mu}) + \hat{\epsilon}_t$$

$$\hat{d} = \underset{(0.013)}{0.236}, Q_{\hat{\epsilon}}(24) = \underset{(pv=0.000)}{1367.562}, log(\mathscr{L}) = -98540.412$$

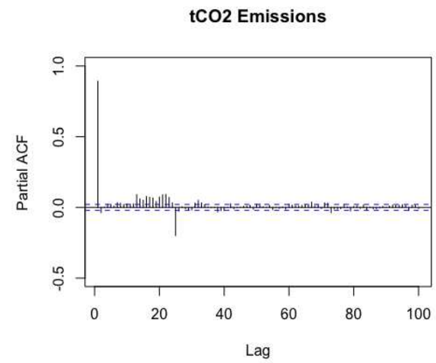
Long-term equilibrium relationship:

$$E_t = -529.44 + 0.4236 G_t + v_t \tag{12}$$

A.3 Zone Center South

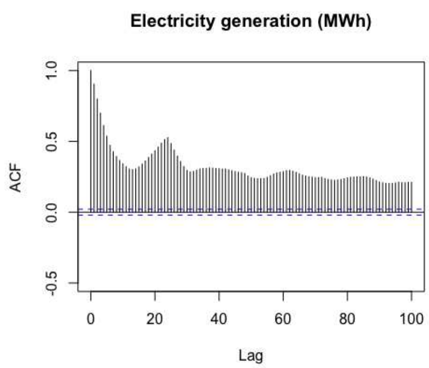


(a) ACF

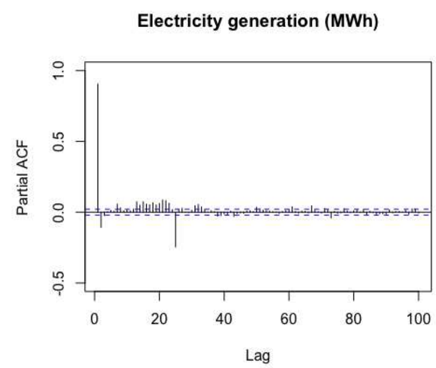


(b) PACF

Figure A5: Zone Center South, ACF (panel a) and PACF (panel b) for seasonally adjusted emissions.



(a) ACF



(b) PACF

Figure A6: Zone Center South, ACF (panel a) and PACF (panel b) for seasonally adjusted generation.

Table A5: *Unit root and stationarity tests for Center South for the seasonally adjusted emissions and generation time series. Sample period: quarters of 2018.*

	quarter	ADF_teststat	ADF_cval	DFGLS_teststat	DFGLS_cval	PP_teststat	PP_cval	KPSS_teststat	KPSS_cval	rKPSS_teststat	rKPSS_cval	ZA_teststat	ZA_cval	variable	zone
1	1	-6.14	-1.95	-7.16	-1.94	-12.84	-2.86	1.12	0.15	0.90	0.15	-14.28	-5.08	Emissions	Center South
2	2	-9.84	-1.95	-11.33	-1.94	-12.90	-2.86	0.86	0.15	0.86	0.15	-15.42	-5.08	Emissions	Center South
3	3	-5.73	-1.95	-7.68	-1.94	-11.83	-2.86	0.42	0.15	0.45	0.15	-12.03	-5.08	Emissions	Center South
4	4	-6.52	-1.95	-11.31	-1.94	-11.00	-2.86	0.63	0.15	0.69	0.15	-12.33	-5.08	Emissions	Center South
5	1	-3.85	-1.95	-11.28	-1.94	-12.76	-2.86	0.66	0.15	0.59	0.15	-13.70	-5.08	Generation	Center South
6	2	-5.29	-1.95	-5.86	-1.94	-11.95	-2.86	0.96	0.15	0.92	0.15	-14.03	-5.08	Generation	Center South
7	3	-4.05	-1.95	-7.71	-1.94	-10.68	-2.86	0.86	0.15	0.89	0.15	-10.94	-5.08	Generation	Center South
8	4	-4.09	-1.95	-8.77	-1.94	-10.52	-2.86	0.57	0.15	0.54	0.15	-12.01	-5.08	Generation	Center South

Table A6: *Summary findings for the zone Center South. Note: $k = 4$ from the lag selection test.*

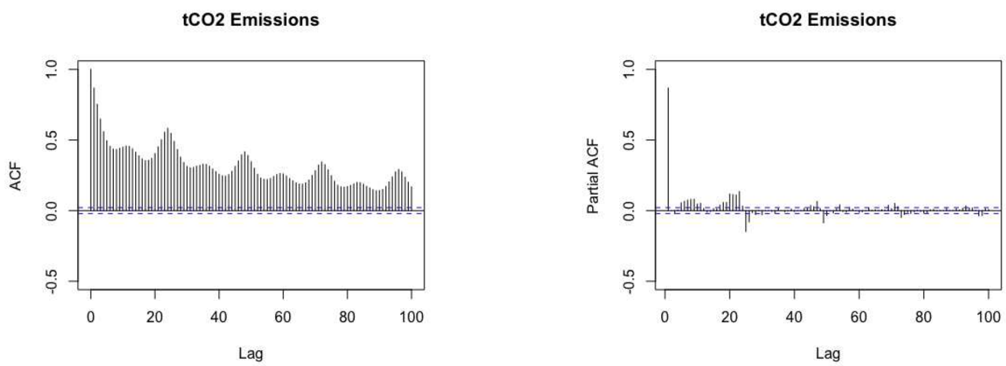
Rank tests:

Rank	$\hat{d}$	Log-likelihood	LR statistic	p-value
0	0.483	-118789.866	6.408	0.171
1	0.479	-118787.066	0.809	0.368
2	0.484	-118786.662	—	—

The rank test shows the absence of cointegrating relationships. Therefore, we estimate the MEF by following the methodology by Beltrami et al. (2020a). Note that the s is the selected order of the autoregressive (AR) component of the process, while q is the selected order of the moving average (MA) component.

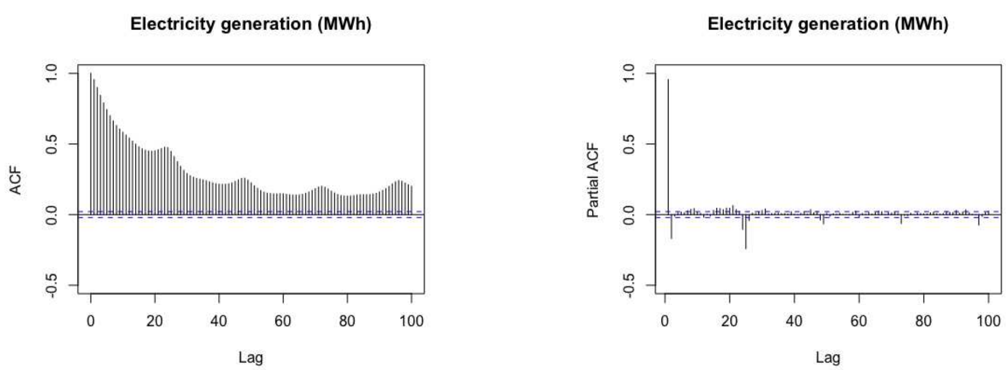
	Hawkes	Hawkes FE	US-FE	ARIMA-FE	ARFIMA-FE
$\hat{\beta}$	0.6967821	0.6967676	0.6639039	0.7030515	0.7022071
S.E. $\hat{\beta}$	0.0042	0.0042	0.0038		0.0043
t value	165.4015	165.1048	172.4515		
p-value	0.00	0.00	0.00		
AIC	111349.2	111349.2	124338.7	110723.2	85861.78
BIC	111370.5	111631.4	124565.2	110970.9	86123.66
s				1	1
q				3	3
$\hat{d}$					0.2823146
S.E. $\hat{d}$					0.04467662
Confidence interval of $\hat{d}$					$[0.1947485; 0.3698808]$

A.4 Zone Sicily



(a) ACF (b) PACF

Figure A7: Zone Sicily, ACF (panel a) and PACF (panel b) for seasonally adjusted emissions.



(a) ACF (b) PACF

Figure A8: Zone Sicily, ACF (panel a) and PACF (panel b) for seasonally adjusted generation.

Table A7: *Unit root and stationarity tests for Sicily for the seasonally adjusted emissions and generation time series. Sample period: quarters of 2018.*

	quarter	ADF_teststat	ADF_cval	DFGLS_teststat	DFGLS_cval	PP_teststat	PP_cval	KPSS_teststat	KPSS_cval	rKPSS_teststat	rKPSS_cval	ZA_teststat	ZA_cval	variable	zone
1	1	-10.45	-1.95	-15.93	-1.94	-14.01	-2.86	0.27	0.15	0.36	0.15	-14.89	-5.08	Emissions	Sicily
2	2	-8.60	-1.95	-10.68	-1.94	-9.26	-2.86	0.68	0.15	0.53	0.15	-12.63	-5.08	Emissions	Sicily
3	3	-10.86	-1.95	-9.05	-1.94	-13.10	-2.86	0.52	0.15	0.78	0.15	-16.81	-5.08	Emissions	Sicily
4	4	-11.54	-1.95	-11.38	-1.94	-12.59	-2.86	0.69	0.15	0.76	0.15	-14.19	-5.08	Emissions	Sicily
5	1	-2.11	-1.95	-4.99	-1.94	-8.00	-2.86	0.45	0.15	0.42	0.15	-7.70	-5.08	Generation	Sicily
6	2	-2.40	-1.95	-2.23	-1.94	-7.86	-2.86	0.39	0.15	0.50	0.15	-8.29	-5.08	Generation	Sicily
7	3	-2.43	-1.95	-8.45	-1.94	-9.26	-2.86	0.82	0.15	1.13	0.15	-10.56	-5.08	Generation	Sicily
8	4	-2.06	-1.95	-5.70	-1.94	-7.33	-2.86	0.63	0.15	0.83	0.15	-7.28	-5.08	Generation	Sicily

Table A8: *Summary findings for the zone Sicily - FCVAR results. Note: $k = 4$ from the lag selection test.*

Rank tests:

Rank	$\hat{d}$	Log-likelihood	LR statistic	p-value
0	0.569	-95156.442	16.007	0.022
1	0.591	-95148.837	0.796	0.71
2	0.598	-95148.439	—	—

Unrestricted model:

$$\Delta^{\hat{d}}(\begin{bmatrix} E_t \\ G_t \end{bmatrix} - \begin{bmatrix} 25.905 \\ 774.52 \end{bmatrix}) = L_{\hat{d}} \begin{bmatrix} -0.282 \\ -0.256 \end{bmatrix} v_t + \sum_{i=1}^{4} \widehat{\Gamma}_i \Delta^{\hat{d}} L^i_{\hat{d}}(X_t - \hat{\mu}) + \hat{\epsilon}_t$$

$$\hat{d} = \underset{(0.011)}{0.289}, Q_{\hat{\epsilon}}(24) = \underset{(pv=0.000)}{1232.508}, log(\mathscr{L}) = -95151.957$$

Long-term equilibrium relationship:

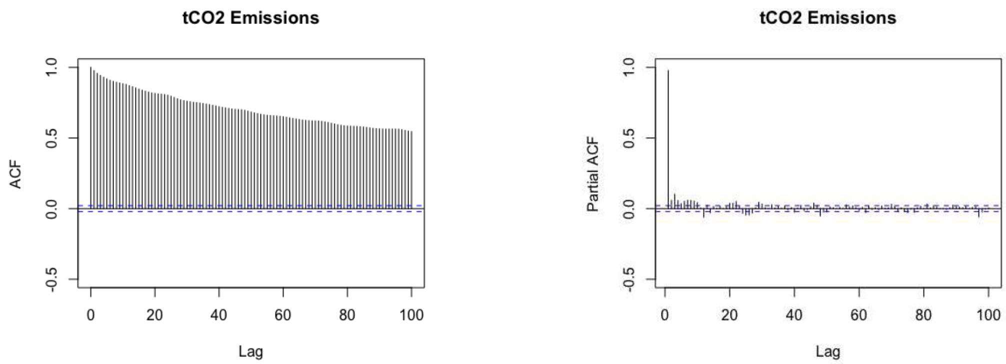
$$E_t = -87.17 + 0.146 G_t + v_t \tag{13}$$

Hypothesis tests:

	$\mathscr{H}^1_d$	$\mathscr{H}^1_\beta$	$\mathscr{H}^1_\alpha$	$\mathscr{H}^2_\alpha$
df	1	1	1	1
LR	513.896	19.63	19.194	8.157
p-value	0.000	0.000	0.000	0.004

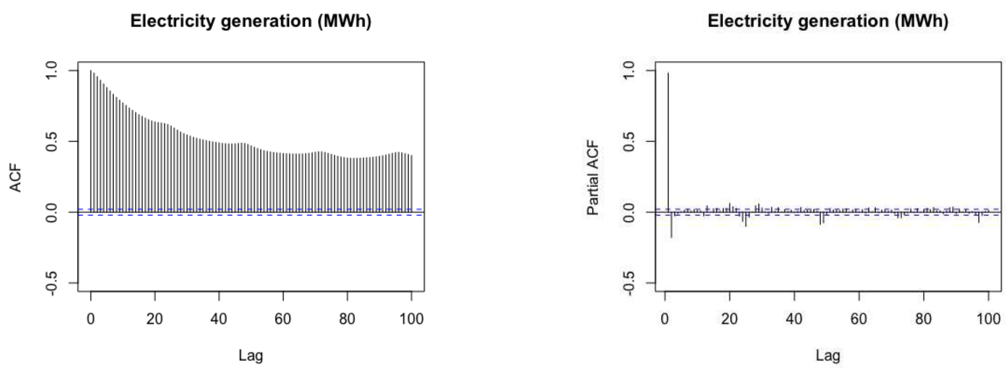
The unrestricted model is the correct specification.

A.5 Zone Sardinia



(a) ACF (b) PACF

Figure A9: Zone Sardinia, ACF (panel a) and PACF (panel b) for seasonally adjusted emissions.



(a) ACF (b) PACF

Figure A10: Zone Sardinia, ACF (panel a) and PACF (panel b) for seasonally adjusted generation.

Table A9: *Unit root and stationarity tests for Sardinia for the seasonally adjusted emissions and generation time series. Sample period: quarters of 2018.*

	quarter	ADF_teststat	ADF_cval	DFGLS_teststat	DFGLS_cval	PP_teststat	PP_cval	KPSS_teststat	KPSS_cval	rKPSS_teststat	rKPSS_cval	ZA_teststat	ZA_cval	variable	zone
1	1	-2.70	-1.95	-5.62	-1.94	-5.84	-2.86	0.85	0.15	0.47	0.15	-8.26	-5.08	Emissions	Sardinia
2	2	-5.56	-1.95	-5.23	-1.94	-5.73	-2.86	0.58	0.15	0.62	0.15	-8.59	-5.08	Emissions	Sardinia
3	3	-2.66	-1.95	-2.34	-1.94	-5.21	-2.86	1.95	0.15	2.25	0.15	-8.61	-5.08	Emissions	Sardinia
4	4	-1.72	-1.95	-2.21	-1.94	-2.73	-2.86	2.27	0.15	2.25	0.15	-8.02	-5.08	Emissions	Sardinia
5	1	-1.32	-1.95	-4.54	-1.94	-6.53	-2.86	0.59	0.15	0.70	0.15	-6.15	-5.08	Generation	Sardinia
6	2	-1.81	-1.95	-2.35	-1.94	-6.41	-2.86	0.57	0.15	0.60	0.15	-7.80	-5.08	Generation	Sardinia
7	3	-1.08	-1.95	-3.82	-1.94	-5.70	-2.86	1.44	0.15	1.69	0.15	-6.18	-5.08	Generation	Sardinia
8	4	-1.05	-1.95	-3.12	-1.94	-5.35	-2.86	0.98	0.15	0.83	0.15	-6.42	-5.08	Generation	Sardinia

Table A10: *Summary findings for the zone Sardinia - FCVAR results. Note: $k = 3$ from the lag selection test.*

Rank tests:

Rank	$\hat{d}$	Log-likelihood	LR statistic	p-value
0	0.586	-88001.107	20.038	0.006
1	0.64	-87991.149	0.122	0.955
2	0.639	-87991.088	—	—

Unrestricted model:

$$\Delta^{\hat{d}}\left(\begin{bmatrix} E_t \\ G_t \end{bmatrix} - \begin{bmatrix} 513.503 \\ 1438.099 \end{bmatrix}\right) = L_{\hat{d}} \begin{bmatrix} 0.003 \\ 0.017 \end{bmatrix} v_t + \sum_{i=1}^{3} \widehat{\Gamma}_i \Delta^{\hat{d}} L^i_{\hat{d}}(X_t - \hat{\mu}) + \hat{\epsilon}_t$$

$$\hat{d} = \underset{(0.024)}{0.64}, Q_{\hat{\epsilon}}(24) = \underset{(pv=0.000)}{784.022}, log(\mathscr{L}) = -87991.149$$

Long-term equilibrium relationship:

$$E_t = 1540.30 + 0.714G_t + v_t \tag{14}$$

Hypothesis tests:

	$\mathscr{H}^1_d$	$\mathscr{H}^1_\beta$	$\mathscr{H}^1_\alpha$	$\mathscr{H}^2_\alpha$
df	1	1	1	1
LR	194.160	19.884	0.73	14.185
p-value	0.000	0.000	0.393	0.000

Restricted model:

$$\Delta^{\hat{d}}\left(\begin{bmatrix} E_t \\ G_t \end{bmatrix} - \begin{bmatrix} 466.937 \\ 1433.553 \end{bmatrix}\right) = L_{\hat{d}} \begin{bmatrix} 0.000 \\ 0.0143 \end{bmatrix} v_t + \sum_{i=1}^{3} \widehat{\Gamma}_i \Delta^{\hat{d}} L^i_{\hat{d}}(X_t - \hat{\mu}) + \hat{\epsilon}_t$$

$$\hat{d} = \underset{(0.024)}{0.609}, Q_{\hat{\epsilon}}(24) = \underset{(pv=0.000)}{795.844}, log(\mathscr{L}) = -87998.242$$

Long-term equilibrium relationship:

$$E_t = -563.644 + 0.7189G_t + v_t \tag{15}$$