Spatio-temporal clustering and reconciliation for regional electricity demand forecasting

Margaux Brégère[a], Raffaele Mattera[b]

[a]*LPSM, Sorbonne Université, Paris, France and EDF R&D, Lab Paris-Saclay, France,*
[b]*Department of Mathematics and Physics, University of Campania "Luigi Vanvitelli", Caserta, Italy*

Abstract

This paper proposes a three-stage approach to forecasting the electricity demand time series of several areas belonging to the same region. First, time series are aggregated on the basis of a spatio-temporal clustering approach, and a three-level hierarchy is build. Second, benchmark forecasts are generated for all series using generalized additive models. Finally, the forecasts are optimally projected onto a coherent subspace to ensure that the final forecasts satisfy the hierarchical constraints. Our approach is tested on Alberta electricity demand data; experimental results suggest that successive clustering and projection steps improve the benchmark forecasts both at the aggregate level and at the disaggregated level.

Keywords: Spatio-temporal clustering, Reconciliation forecasting, electricity demand forecasting

1. Introduction

1.1. Motivation: Electricity Demand Forecasting

To balance the supply and demand of electricity and keep the grid operating 24 hours a day, it is crucial to predict the electricity demand as accurately as possible at various levels of aggregation. Predicting the overall electricity demand is necessary in order to inject the right amount of electricity into the grid at any given time. It is also necessary to forecast demand at a local level in order to dispatch correctly the electricity into the grid. Thus, forecasts at various aggregated levels (bottom and top) are useful for an efficient management of the electrical grid. In this work, we consider local electricity demand time series that we aim to forecast (bottom-level), as well as the sum of them (top-level). To do so, we first build a three-level hierarchy by aggregating bottom-level time series based on a spatio-temporal clustering approach. Clustering is adopted to account for similarities in consumption patterns under the assumption of an unknown hierarchical structure, that is, there is no prior knowledge of how local electricity demands should be clustered to improve forecasting. Second, benchmark forecasts are independently generated for all series (bottom-level, clusters and top-level) using generalized additive models, which are powerful electricity demand forecasting models - see, among others, Pierrot and Goude (2011) for further details. Noticing that these time series may be correlated (they highly depend on quite similar weather condition, e.g.) and are linked through summation constraints, our setting can be situated within the broader field of reconciliation time series forecasting. Therefore, the forecasts are optimally reconciled using the Minimum Trace reconciliation approach (MinT) introduced by Hyndman et al. (2011). This ensures that the final forecasts satisfy the hierarchy. In short, our approach consists of three stages:

- Building a hierarchy based on a spatio-temporal clustering of the bottom level times series
- Training a forecasting model for each time series
- Reconciliating optimally the forecasts.

Our approach is tested on Alberta electricity demand data; experimental results suggest that successive clustering and reconciliation steps improve the benchmark forecasts both at bottom and top levels.

1.2. Related works

This work takes inspiration from literature in different fields. In particular, we consider previous studies discussing forecasting methods for electricity demand predictions, but also papers using clustering techniques in this context. Moreover, it also relates to spatio-temporal clustering literature and to more recent works employing forecast reconciliation to enhance the forecasting accuracy of electricity demand.

Forecasting Electricity Demand. Electricity demand strongly depends on calendar and weather exogenous variables, so a lot of performing forecasting models are based on regression rather than time series methods. Moreover, from an operational point of view, it is generally preferable to use models based on exogenous variables only, as they are easier to interpret and are not dependent on the quality of recent data. This work focuses on offline regression methods. However, it should be noted that when there is an access to recent data, classical time series methods like ARIMA (see, ammong others Fard and Akbari-Zadeh 2014) as well as recurrent neural networks based models (e.g., Kong et al. 2017) are efficient. Moreover, these recent data can be used to update models trained in a regression framework with online or transfer learning methods (see, among others, Gaillard et al. 2016 and Obst et al. 2021). The relationship between electricity demand and exogenous variables is rarely linear, which is why Generalized Additive Models (GAMs) are particularly effective. Studied by Fan and Hyndman (2010) and Pierrot and Goude (2011), they model the expected demand as a sum of independent exogenous variable effects, which are approached with smooth functions. Load forecasting has not escaped the current deep learning trend (see, among others Keisler et al. 2024) and many applications in the electrical field, based on neural networks have been successfully carried out (see Massaoudi et al. 2021 for a quite recent review). Finally, Chen et al. (2004), Taieb and Hyndman (2014) and Dudek (2015) proposed efficient load forecasting methods based on and support vector machine, gradient boosting and random forests, respectively.

Clustering Electricity Demand. Clustering belongs to unsupervised learning and has been frequently used for classifying electricity demand profiles. Historically, it has been used for dealing with large datasets of single energy users' demand profiles - see Chicco (2012) for a comprehensive review of clustering techniques used for this aim. Both hierarchical clustering (e.g. Alonso et al., 2020) and k-means (e.g. Dong et al., 2022, Syed et al., 2021) methods are widely used for this aim. However, clustering techniques have also been employed to classify time series of electricity demand of more aggregated levels. For example, Chévez et al. (2017) proposed to use k-means for clustering areas with similar consumption patterns in La Plata region, Argentina, while Zhang et al. (2021) adopted a similar approach for clustering districts of Zhejiang Province, China. Our paper is in line with this second strand of literature because we are interested in clustering spatial areas – and not individual customers – with similar demand patterns.

Spatio-Temporal Clustering. Spatio-temporal clustering is the process of grouping units based on both their spatial and temporal similarities (Kisilevich et al., 2010) and it is a relatively new field of temporal data mining. A recent review on the spatio-temporal clustering approach can be found in Ansari et al. (2020). While the temporal dimension is used to assess the similarity of the spatial units' evolution over time, the spatial dimension describes how these are localized in the space. In this paper, we adopt a spatio-temporal clustering approach as we conjecture that electricity demand patterns depend also on unobservable geographical factors. We expect forecasts obtained from models using spatio-temporal clustering to be more accurate compared to those using time series clustering if this assumption holds. In this paper, following the majority of studies dealing with electricity demand data, we consider a distance-based clustering approach. In sum, distance-based approaches transform the complex temporal patterns of the units – spatial aggregates in our instance – into feature vectors. In doing so, standard clustering algorithms, such as the k-means or a hierarchical clustering procedure, can be used. Once dealing with spatial dimension, however, spatially penalized clustering algorithms are commonly considered. The penalization serves as a way of enhancing spatial clustering in the data. In the context of partitional algorithms, Pham (2001) proposed a k-means algorithm with spatial constraint, which has been extended to the spatio-temporal setting by D'Urso et al. (2019). From the hierarchical clustering perspective, Chavent et al. (2018) recently proposed a Ward Jr (1963)-like algorithm, based on a modified dissimilarity matrix. In Chavent et al. (2018), the dissimilarity is defined as a convex combination of geographical distances and the dissimilarity matrix computed from attribute variables. Mattera and Franses (2023) have extended the algorithm of Chavent et al. (2018) to the case of spatio-temporal data.

Forecast Reconciliation. Over the past few decades, there has been a notable expansion in the field of reconciliation forecasting, accompanied by the advent of methods that guarantee coherent forecasts and enhance forecast precision - Athanasopoulos et al. (2023) offers an exhaustive review of research on the subject. In this work, we use the approach proposed by Wickramasuriya et al. (2019), which relies on orthogonal or oblique projections of the vector of forecasts over a predefined constraint sub-space. In an electricity demand forecasting framework, Goehry et al. (2019) and Brégère and Huard (2022) employ reconciliation methods, namely bottom-up and projection, respectively, with a view to enhancing forecasts at various aggregated levels. Recent developments in forecast reconciliation literature (Li et al., 2019, Mattera et al., 2024, Zhang et al., 2024) also consider that the hierarchical structure is unknown. We further differentiate from these studies by proposing to estimate the hierarchical structure with spatio-temporal clustering.

1.3. Main contribution

In sum, the paper's contribution is threefold. First, we contribute to the literature on electricity demand forecasting by adding positive evidence on the usefulness of forecast reconciliation in this context. Moreover, we adopt a forecast reconciliation with cluster structure, given that the hierarchy is unknown. Therefore, we show that forecast reconciliation with cluster structure – as discussed in Mattera et al. (2024) and Zhang et al. (2024) – is useful in the field of electricity demand forecasting. Consequently, the second contribution of this paper is to the forecast reconciliation literature, in that we propose a MinT forecast reconciliation procedure based on spatio-temporal clustering. Previous papers dealing with unknown hierarchies considered temporal dimension only. Our results indicate that, in the case of electricity demand data, reconciled forecasts obtained from spatio-temporal clustering are more accurate than those obtained from time series clustering. In other words, we show the existence of benefits from the forecasting perspective in including spatial dimension into the clustering problem while working with a hierarchical structure of unknown form. Finally, we propose a simple approach to extend the Chavent et al. (2018) algorithm to the case of hourly time series. The consideration of hourly time series introduces a degree of complexity to the clustering model, because the commonly used temporal distances considered by previous studies (for an overview, see Díaz and Vilar, 2010) are not well-suited in this case. Therefore, to measure temporal dissimilarity across the spatial units, we propose the following procedure. We first divide the electricity consumption time series into 24-hour patterns (e.g. see Voulis et al., 2018, Durante et al., 2023) and obtain 24 multivariate time series for each spatial unit, one for each hour of the day. We then compute 24 distance matrices and calculate a consensus temporal distance matrix using the DISTASIS algorithm (Abdi et al., 2005, 2012). In doing so, we optimally weight the differences between the units observed at different hours of the day. The rest of the paper is structured as follows. Section 2 discusses the adopted methodology in detail, describing the forecast reconciliation with cluster structure and the adopted spatio-temporal clustering approach. Section 3 presents the data, while the forecasting experiment design and the results are shown in Section 3.3. Section 4 concludes with final remarks and future research directions.

2. Methodology

Our intention is to forecast the time series of local (bottom-level) electricity demand in m areas $\{(\mathrm{b}_t^i)_{t>0},\ i=1,\ldots,m\}$ as well as the global demand $(\mathrm{d}_t)_{t>0}$, namely the sum of local time series: $\mathrm{d}_t = \sum_{i=1}^m \mathrm{b}_t^i$, for any time step t. To do so, we first cluster the space-time series $(\mathrm{b}_t^i)_{t>0}$ into k clusters. By re-indenting the time series according the clustering such that the first n_1 belong to Cluster 1, the following n_2 to Cluster 2 and so one, and we define $m_0 = 1$, and, for $\ell = 1,\ldots,k$ $m_\ell = \sum_{j=1}^{\ell} n_j$, and we create the k middle-level time series $\{(\mathrm{c}_t^\ell)_{t>0},\ \ell = 1,\ldots,k\}$ with $\mathrm{c}_t^\ell = \sum_{i=m_{\ell-1}}^{m_\ell} \mathrm{b}_t^i$, for any time step t. This induces the hierarchy represented in Figure 1. For any time

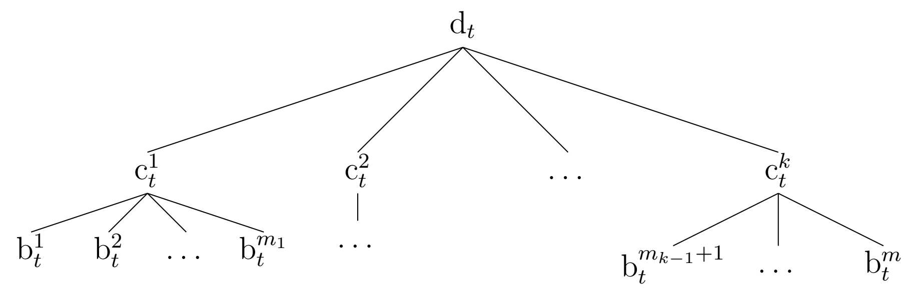

Figure 1: Representation of the three-level hierarchy induced by a clustering of the space-time series $(b_t^i)_{t>0}$ into k clusters.

step t, with $\mathbf{b}_t$ be the vector of the m bottom-level time series at time t, and $n = m + k + 1$, we define the n-dimensional vector containing all the time series $\mathbf{y}_t$:

$$\mathbf{y}_t = \begin{bmatrix} \mathrm{d}_t & \mathrm{c}_t^1 & \cdots & \mathrm{c}_t^k & \mathbf{b}_t^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} = \mathbf{S}\mathbf{b}_t\,,$$

where $\mathbf{S}$ is the $n \times m$ "summing" or "structural" matrix defined thanks to the clustering. It is given by

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{I}_m \end{bmatrix} \text{ with } \boldsymbol{A} = \left[\begin{array}{cccccccccc} 1 & \dots & 1 & 1 & \dots & 1 & \dots & 1 & \dots & 1 \\ 1 & \dots & 1 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \multicolumn{3}{c}{\underbrace{0 \quad \dots \quad 0}_{n_1}} & \multicolumn{3}{c}{\underbrace{1 \quad \dots \quad 1}_{n_2}} & \dots & \multicolumn{3}{c}{\underbrace{1 \quad \dots \quad 1}_{n_k}} \end{array}\right] \begin{array}{l} \leftarrow \text{Top-level} \\ \leftarrow \text{Cluster } 1 \\ \leftarrow \text{Cluster } 2 \\ \vdots \\ \leftarrow \text{Cluster } k \end{array} .$$

Once the time series vector $(\mathbf{y}_t)_{t>0}$ has been obtained, using some exogenous variables $(\mathbf{x}_t)_{t>0}$, we train n statistical or machine learning models to forecast each of its component. For $i = 1, \dots, n$ and any time step t, we denote by $\hat{f}_i$ the model which outputs the prediction $\widehat{\mathrm{y}}_t^i = \hat{f}_i(\mathbf{x}_t)$ of y_t^i. The forecasts are then gathered into a vector $\widehat{\mathbf{y}}_t$. Finally, we update them using a linear reconciliation method, which is based on the estimation of a matrix $\mathbf{G}$, so that the final forecasts are

$$\widetilde{\mathbf{y}}_t = \mathbf{S}\mathbf{G}\widehat{\mathbf{y}}_t \,.$$

Similarly to Wickramasuriya et al. (2019), we consider linear reconciliations such that the reconciliation matrix ee qual to

$$\mathbf{G} = \left(\mathbf{S}'\boldsymbol{W}^{-1}\mathbf{S}\right)^{-1} \mathbf{S}'\boldsymbol{W}^{-1}, \tag{1}$$

with $\boldsymbol{W}$, a $n \times n$ definite positive matrix. To sum-up, our three-steps approach relies on the estimations of the summing matrix $\mathbf{S}$ (obtained with the clustering), the forecasting models $\left(\hat{f}_i\right)_{1 \leq i \leq n}$ and $\boldsymbol{W}$, the matrix needed to define the reconciliation matrix $\mathbf{G}$. To validate our procedure, we cut the data into three sub-data sets: the training data set is used for the estimation of the clustering structure and of the models, the calibration data set is for the estimation of the reconciliation matrix $\mathbf{G}$. Finally, we assess the relevance of our approach on a testing test. In what follows, we detail each step of our procedure, illustrated in the diagram in Figure 2.

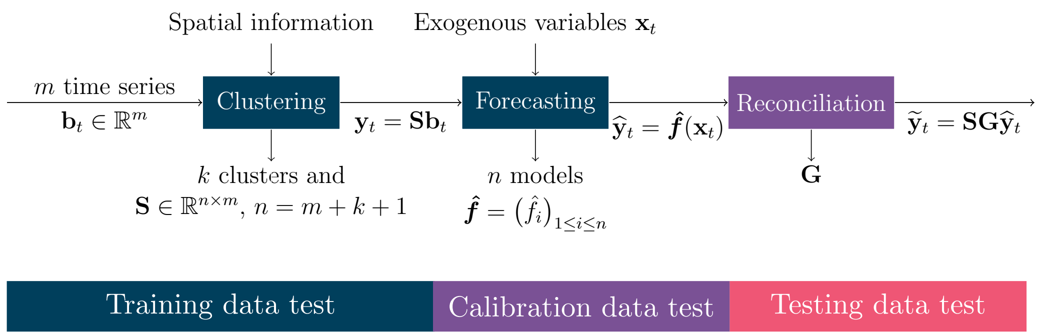

Figure 2: Clustering, Forecasting, and Reconciliation Process Flow.

2.1. Hierarchical Spatio-Temporal Clustering of Hourly Data

We highlight that, along with $\boldsymbol{W}$, also the knowledge of the summation matrix $\boldsymbol{S}$ is required to optimally reconcile base forecasts. In particular, we note that the summation matrix $\boldsymbol{S}$ depends on the aggregation matrix $\boldsymbol{A}$, which is unknown in our setting as it depends on the clustering we set. Therefore, to implement optimal reconciliation via the reconciliation procedure defined in Equation (1), we need to define a suitable approach for estimating the cluster structure of the data. We propose to estimate such a structure considering the similarity in both time and spatial domains. In doing so, we conjecture that electricity demand time series depend also on unobservable geographical factors so that better clustering can be achieved if the spatial dimension is taken into account. It is thus anticipated that forecasts reconciled with spatio-temporal clustering will prove more accurate than those reconciled with temporal clustering. For this aim, we consider the weighted Ward-like hierarchical clustering algorithm as discussed in Chavent et al. (2018) and Mattera and Franses (2023), where the temporal dimension is suitably weighted with the spatial dimension. In general, any Ward-like hierarchical clustering approach starts with an initial partition with m clusters of singletons. Then, at each step, the algorithm aggregates two clusters according to an objective function related to the within-cluster inertia. More precisely, let us consider the set of the m bottom-level time series and

$\mathcal{P}_k = \{\mathcal{C}^1, \ldots, \mathcal{C}^k\} \in \mathcal{P}\big(\{1, \ldots, m\}\big)$, any partition of the time series into k clusters. For any $\ell = 1, \ldots, k$, the weighted within-clusters inertia $I\left(\mathcal{C}^\ell\right)$ is defined as

$$\begin{aligned} I\left(\mathcal{C}^\ell\right) = (1-\alpha) \sum_{(i,i') \in \mathcal{C}^\ell \times \mathcal{C}^\ell} & \frac{w_i w_j}{2 \sum_{j \in \mathcal{C}^\ell} w_j} d_t(i,i')^2 \\ + \; \alpha & \sum_{(i,i') \in \mathcal{C}^\ell \times \mathcal{C}^\ell} \frac{w_i w_{i'}}{2 \sum_{j \in \mathcal{C}^\ell} w_j} d_g(i,i')^2 \,, \end{aligned} \tag{2}$$

where w_i is the weight associated to the i^{th} time series, $d_t(i,i')^2$ and $d_g(i,i')^2$ are squared distances between times series i and i' defined over some temporal and geographical spaces, respectively. The parameter α balances these two distances. In the absence of any a prior information, we assume all time series to be equally important and set $w_i = 1/m, \forall i = 1, \ldots, m$. We note that, the smaller the inertia $I\left(\mathcal{C}^\ell\right)$ is, the more homogeneous are the observations in cluster $\mathcal{C}^\ell$. The key idea in Ward's method is to minimize the increase in within-cluster inertia defined in Equation (2) when merging clusters. The construction of the partitions follows a Ward-like hierarchical procedure, which is greedy. Indeed, the clustering algorithm starts with the initial partition formed by m distinct clusters of all singletons: $\mathcal{P}_m = \big\{\{1\}, \ldots, \{m\}\big\}$. Then, to decide which clusters to merge at each step, it considers all possible pairs of clusters and calculates the total within-cluster inertia that would result from merging each pair. Those minimizing the total within-cluster inertia are merged. After merging two clusters $\mathcal{C}^\ell$ and $\mathcal{C}^{\ell'}$, the algorithm computes the inertia between the newly merged cluster and all the remaining ones. The process of merging clusters continues iteratively until all the m time series are in one single cluster $\mathcal{P}_1 = \{1, \ldots, m\}$.

Setting of the number of cluster k. The main benefit of a hierarchical clustering algorithm compared to the k-means is that the first does not require the selection of the number of clusters k in advance. The selection of k is done ex-post, through the automatic elbow criterion (see Morgado et al., 2023).

Setting of the balance between temporal and geographical information: α. The parameter α can be either fixed by the user or selected automatically. To select the optimal value of α, Chavent et al. (2018) proposed a heuristic method aimed to balance temporal and geographical homogeneity in the clustering process. The approach evaluates how the trade-off between these two criteria changes with α. Two metrics, Q_t for temporal cohesion and Q_g for geographical cohesion, are calculated for different values of α. The optimal α is identified either at the point where the two metrics intersect or by minimizing the squared difference between them. In this paper, we estimate α automatically following this procedure. For further details, interested readers can refer to Chavent et al. (2018) and related papers (e.g. see Bucci et al., 2023, Morelli et al., 2024, Mattera and Franses, 2025).

Setting of the temporal and geographical distances. It remains to define the temporal and geographical distances $d_t(\cdot,\cdot)$ and $d_g(\cdot,\cdot)$, respectively. Following Mattera and Franses (2023), we define the $d_g(\cdot,\cdot)$ as the Euclidean distance in the geographical coordinates, namely, for any time series m, the latitude and the longitude of the barycenter of the area under consideration. Although other approaches can also be considered (e.g. see Fouedjio, 2016, Mattera, 2022), this choice is enough to provide a soft spatial constraint to the partition. About the temporal distance, $d_t(\cdot,\cdot)$, we must remark that the electricity demand data is available at hourly frequency. This high frequency introduces a higher degree of complexity into the clustering problem, because the commonly used temporal distances considered by previous studies are not well-suited in this setting. Following a common idea in previous research (e.g. see Voulis et al., 2018, Durante et al., 2023), for each electricity consumption time series b_t^i, we split it into 24-hour patterns, obtaining 24 multivariate time series $\mathbf{b}_t^i = \left[b_{1,t}^i, b_{2,t}^i, \ldots, b_{24,t}^i\right]^{\mathrm{T}}$ with daily frequency. We define the temporal distance as the average correlation-based distances over the 24 hours:

$$d_t(i, i') = \frac{1}{24} \sum_{h=1}^{24} \sqrt{2[1 - \rho_{i,i'}(h)]}\,, \quad \text{with} \quad \rho_{i,i'}(h) = \mathrm{cor}(b^i_{h,t}, b^{i'}_{h,t})$$

the pairwise correlation coefficient between the series $b^i_{h,t}$ and $b^j_{h,t}$ in the h^{th} hour of the day (in practice, we estimate these coefficients on the training data set). Therefore, the more correlated demand patterns are for all h^{th} hour of the day, the more similar the two areas i and i' are. We highlight that instead of simply averaging all the distances, it could have been appropriate to assign λ_h weights to each h^{th} hour and therefore compute a compromise (or consensus) distance. We address the task of the compromise computation with the DISTATIS algorithm (Abdi et al., 2005, 2012). We finally emphasize that, to make the temporal and spatial dimensions comparable, we consider representative load patterns (RLPs) by normalizing the two distances defined above:

$$\widetilde{d}(i, i') = \frac{d(i, i')}{\max_{(j,j') \in \{1,\dots,m\}^2} d(j, j')}, \quad \text{for} \quad d = d_t \quad \text{and} \quad d = d_g\,.$$

Once the clustering $\mathcal{P}_k = \{\mathcal{C}^1, \dots, \mathcal{C}^k\}$ defined, time series are re-intended so that the first ones belong to $\mathcal{C}^1$ and so on; and the middle-level time series $c^1_t, \dots c^\ell_t$ are computed.

2.2. Electricity Demand Forecasting Models

As previously discussed in the literature review (Subsection 1.2), Generalized Additive Models (GAMs, see Wood 2017 for a comprehensive introduction) are an effective semi-parametric approach for forecasting electricity demand. They model the demand with a sum of independent, exogenous (potentially non-linear) variable effects. In the experiments of Section 3, we consider local meteorological and calendar exogenous variables. The data set includes temperature readings at hourly intervals from a number of weather stations. To every area $i = 1, \dots, m$ (bottom level), we associate a single weather station (either a station in the area under consideration, or the station geographical ly closest to the area) and denote $(\tau_t^i)_{t>0}$ its temperature time series. For the aggregated levels, the temperature time series of the areas included in the aggregation are averaged. Therefore, for each $i = 1, \ldots, n$, we have access to a temperature time series $(\tau_t^i)_{t>0}$. As the incorporation of an exponential smoothed temperature, which models the thermal inertia of buildings, is likely to enhance the quality of forecasts (see, e.g. Goude et al., 2013), we create the n smoothed temperature times series defined by

$$\begin{cases} \overline{\tau}_1^i = \tau_1^i \\ \overline{\tau}_t^i = \theta \overline{\tau}_{t-1}^i + \big(1-\theta\big)\tau_t^i \,, \quad \text{for any } t \leq 2 \quad \text{and} \quad \theta \in [0,1]\,. \end{cases}$$

For a time step t, the following calendar variables are also considered: the day of the week d_t (equal to 1 for Monday, 2 for Tuesday, etc.), the hour of the day $h_t \in \{1, \ldots, 24\}$. We emphasize that we implicitly assume weather variables τ_t^i, for any $i = 1, \ldots, n$, to be available (and deterministic) at time t. In an operational context, it will obviously not be the case and to compute the load forecasts, it is common to use weather predictions. Finally, to model a potential change in the demand during and after the covid pandemic, we denote by t_0 the time of the beginning of the lockdown imposed as a consequence of the propagation of the disease (May 1$^{\text{st}}$, 2020). All the variables: t, d_t, h_t, τ_t^i, $\overline{\tau}_t^i$, for all $i = 1, \ldots, n$ are gathered in the vector of exogenous variables $\mathbf{x}_t$. The more aggregated level the time series, the smoother the demand of electricity and the easier it is to forecast. Therefore, in the case of middle and top-level time series forecasting, a greater number of exogenous variables may be considered than in the case of bottom-level time series. The additive models for the electrical demand time series that have been subjected to consideration are as follows:

- For the bottom-level time series:

$$\widehat{f^i}(\mathbf{x}_t) = \sum_{d=1}^{7} \sum_{h=1}^{24} \widehat{\alpha}_{d,h}^i \mathbf{1}_{d_t=d} \mathbf{1}_{h_t=h} + \widehat{\alpha}_{\text{covid}}^i \mathbf{1}_{t>t_0}\,.$$

- For the middle-level time series:

$$\widehat{f^i}(\mathbf{x}_t) = \widehat{s}^i_{\overline{\tau}}\left(\overline{\tau}^i_t\right) + \sum_{d=1}^{7}\sum_{h=1}^{24} \widehat{\alpha}^i_{d,h}\mathbf{1}_{d_t=d}\mathbf{1}_{h_t=h} + \widehat{\alpha}^i_{\text{covid}}\mathbf{1}_{t>t_0}\,.$$

- For the top-level time series:

$$\widehat{f^i}(\mathbf{x}_t) = \widehat{s}^i_{\tau}\left(\tau^i_t\right) + \widehat{s}^i_{\overline{\tau}}\left(\overline{\tau}^i_t\right) + \sum_{d=1}^{7}\sum_{h=1}^{24} \widehat{\alpha}^i_{d,h}\mathbf{1}_{d_t=d}\mathbf{1}_{h_t=h} + \widehat{\alpha}^i_{\text{covid}}\mathbf{1}_{t>t_0}\,.$$

The $\widehat{s}^i_{\tau}$ and $\widehat{s}^i_{\overline{\tau}}$ functions catch the effect of the temperatures. They are cubic splines: C^2-smooth functions made up of sections of cubic polynomials joined together. The coefficients $\widehat{\alpha}^i_{d,h}$, $\widehat{\alpha}^i_{\text{covid}}$ model the influence of the hour, the day of the week and the break after the covid pandemic. To estimate each model, we use the Penalized Iterative Re-Weighted Least Square (P-IRLS) method of Wood (2017), implemented in the R-package `mgcv`, on a training data set.

2.3. Optimal reconciliation

We recall that choices for the matrix $\boldsymbol{W}$ in the reconciliation formula

$$\mathbf{G} = \left(\mathbf{S}'\boldsymbol{W}^{-1}\mathbf{S}\right)^{-1}\mathbf{S}'\boldsymbol{W}^{-1},$$

lead to different reconciliation approaches. Indeed, according to Wickramasuriya et al. (2019), the optimal minimum trace reconciliation is obtained by setting $\boldsymbol{W} = \widehat{\boldsymbol{\Sigma}}$, the variance-covariance matrix of the corresponding base forecast errors. With $\boldsymbol{W} = \boldsymbol{I}_n$ we obtain an orthogonal projection and get the solution of the Ordinary Least Squares (OLS) reconciliation (Hyndman et al., 2011). Following Figure 2, we estimate the covariance matrix $\widehat{\boldsymbol{\Sigma}}$ on the calibration data set.

3. Forecasting electricity demand in Alberta, Canada

In what follows, we discuss the empirical experiment to data about the electricity demand in Alberta, Canada. We presents the data first, and then discuss the cluster structure obtained considering the procedure explained in Section 2. We compare the results with the same clustering approach but without considering geographical information, that is setting $\alpha = 0$ in Equation (2). Codes and data are provided on a public Github repository[1].

3.1. The underlying data sets

The Alberta Electric System Operator (AESO) is dedicated to assuming a position in facilitating the transformation of the province's electricity sector, while guaranteeing the uninterrupted provision of power to Albertans. It provides open source electricity demand data[2] for the 42 areas in the region at hourly intervals from January 1$^{\text{st}}$, 2011 to October 31$^{\text{st}}$, 2023. Temperature data come from the NOAA[3] data base: we selected several weather stations (with records available over the considered period) in Alberta region. Smoothed temperatures are computed with $\theta = 0.98$. Both observed and smoothed temperatures are plotted in Figure 3 for one of the weather stations. Table 1 sums up the variables of the data set and gives their range. The electricity demand, ranging from 0 to 2326 MW, highlights the diversity in power consumption across different times and areas, reflecting variations likely influenced by AESO areas-specific factors. Temperature, a critical variable for modeling energy demand, spans an extensive range from -46°C to 40°C, capturing extreme climatic conditions that significantly impact heating and cooling requirements. The smoothed version of the temperature, with a narrower range of -36°C to 30°C, likely represents a processed version to remove noise or account for lag effects in temperature-related energy use.

[1] `https://anonymous.4open.science/r/clustering_and_reconciliation-7647`

[2] `https://www.aeso.ca/market/market-and-system-reporting/data-requests/hourly-load-by-area-and-region/`

[3] National Oceanic and Atmospheric Administration, `https://www.noaa.gov/`

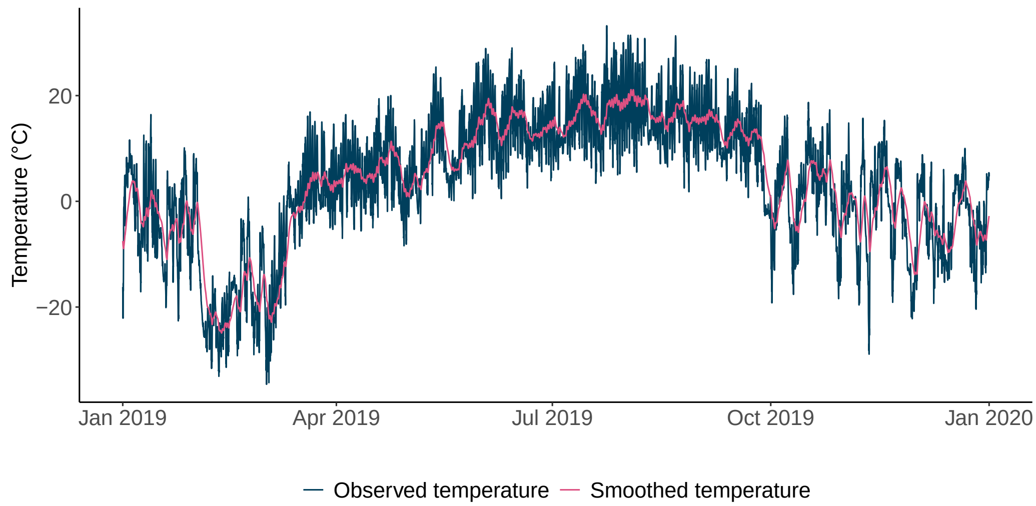

Figure 3: Observed and smoothed temperatures for the weather station “CLARESHOLM” through 2019.

Variable and description	Range / Value
Date	From January 1, 2019 to October 31, 2023 (hourly)
Area	42 numbers
Weather station	22 numbers
Electricity demand	From 0 to 2326 MW
Temperature	From −46°C to 40°C
Smoothed temperature	From −36°C to 30°C
Hour (integer)	From 1 to 24
Day of the week (integer)	From 1 (Monday) to 7 (Sunday)

Table 1: Summary of the variables provided and created for each area of the data set.

The training data set contains two full years of data: from January 1st, 2019 to December 31st, 2020. The calibration data set contains a full year of data: from January 1st, 2021 to December 31st, 2022. The testing data set contains the data available from the end of the calibration set: from January 1st, 2022 to October 31st, 2023. Figure 4 shows the distribution of the electricity demand time series for the 42 AESO areas, distinguishing between training, calibration and testing datasets. With few exceptions (i.e. areas 25, 33

Figure 4: Box-plots of the 42 Alberta electricity demand time series.

and 35), for the vast majority of AESO areas we find that the distributions of electricity demand between the different data sets are quite similar, suggesting that the information captured by the models during training should generalize well to both calibration and testing data sets. These similarities are particularly important as it also ensures that the cluster structure defined during the training phase remains consistent across calibration and testing phases, supporting the robustness of the clustering approach. Such consistency is crucial because it maintains the interpretability and applicability of the clusters, ensuring they reflect the same underlying relationships in all phases of the analysis. Moreover, the similarity between the calibration and testing datasets is even more critical, as calibration data plays a pivotal role in determining the reconciliation matrix used to adjust forecasts for accuracy. If the calibration and testing distributions are aligned, the reconciliation process is more likely to yield adjustments that are effective and representative of real-world demand patterns. This alignment strengthens the validity of the forecasting model, as the adjustments made during calibration will mirror the conditions encountered during testing, leading to improved forecast reliability. Thus, the observed distributional consistency across datasets is a promising indicator of both model robustness and the effectiveness of the reconciliation process.

3.2. Cluster structure

Before moving to electricity demand forecasting, we investigate the cluster structure in terms of spatio-temporal load patterns in the Alberta region. We adopt the hierarchical spatio-temporal clustering based on the mixed inertia introduced in Equation (2). Optimizing the parameter α using the procedure proposed in Chavent et al. (2018) and described in Section 2.1 gives some indication about the relevance of geographical information for the spatio-temporal clustering: the larger α is, the more importance is given to the spatial information. In the case of temporal clustering, we simply set $\alpha = 0$. In the case of Alberta electricity demand patterns, as shown in Figure 5, the temporal homogeneity and spatial homogeneity curves cross when α takes a value equal to 0.3. Therefore, we find that the spatial dimension is important to adjust the partition based on the temporal dimension only, which still remains the most relevant in defining the cluster structure of electricity load patterns. Using the automatic elbow algorithm, we detect $k = 4$ for the temporal clustering and $k = 2$ for the spatio-temporal approach. Figure 6 shows the comparison of the obtained cluster structures, considering both temporal and spatio-temporal approaches. Black points in the figure correspond to the weather stations used to extract information on temperature. Temperature data for areas without weather stations is taken from the closest station in the geographical space. We remark that the temporal cluster structure only reflects demand fluctuations over time to define similarities and discards possible similarities due to spatial proximity. The resulting clusters do not exhibit strong spatial cohesion, and two larger groups with highly synchronized temporal profiles are found. Then, the other two clusters include only two AESO areas each (AREA 13 and AREA 20 on one side and AREA 21 and AREA 52 on the other side), which could be outliers in terms of load patterns. On the other side, spatio-temporal results reflect a balance between temporal synchronization and geographical location. In general, we notice that some areas that were temporally similar but geographically distant in the previous clustering now fall into different clusters, while ge-

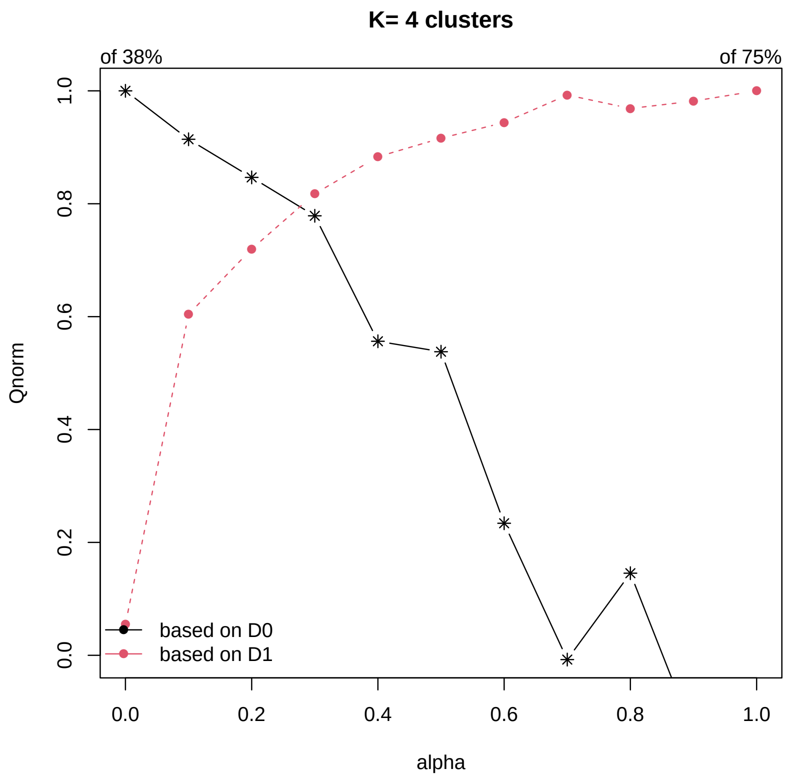

Figure 5: Optimal α, defining the relevance of spatial dimension, selected according to the procedure proposed in Chavent et al. (2018). The larger α, the more important is the spatial dimension.

ographically proximate areas are grouped despite minor differences in their temporal patterns. Comparing the results of the two clustering approaches more precisely, we observe that the spatio-temporal approach provides a more compacted partition, an isolated area in the fourth cluster is now matched with its neighbors. In general, the spatio-temporal approach suggests the existence of two large clusters in the Alberta region, that is, north and south, with distinct consumption patterns.

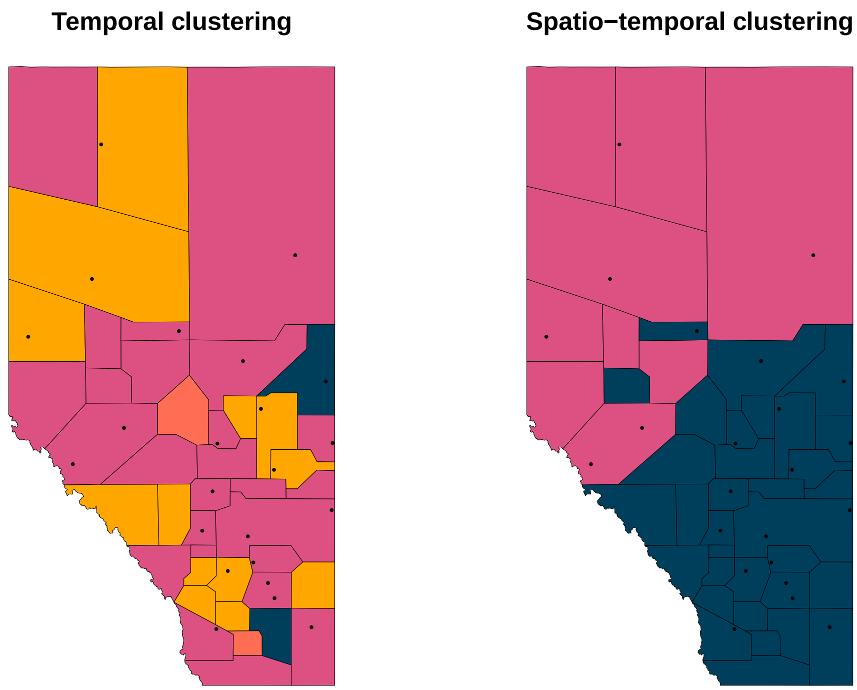

Figure 6: Clustering. Black points correspond the weather stations. If there is two or more stations in the it is because one is associated with the corresponding and the other(s) with close region(s).

3.3. Results

In what follows, we compare the out-of-sample accuracy of the unreconcilied forecasts with respect to different reconcilation approaches – bottom-up (BU), OLS and MinT – using either no cluster structure at all, where the matrix $\boldsymbol{A} = \mathbf{1}^{\mathrm{T}}$, temporal cluster structure and spatio-temporal as defined in Figure 6, and the "ALL" approach proposed in Mattera et al. (2024) where both cluster structures are combined. In the case of bottom-up, we consider three different approaches, that is, the standard bottom-up without the clustering structure and the bottom-up with clusters. In this last case, the forecast for the top-level series is obtained from the cluster aggregates (i.e. middle-level) rather the bottom series themselves. Table 2 presents a comparison of different reconciliation approaches for the top-level series, that is, the electricity demand for the whole Alberta region. The out-of-sample accuracy is evaluated under two distinct evaluations based on two metrics, the Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). The "Base" approach represents the baseline model, where no

Clustering	No RMSE · MAPE	Temporal RMSE · MAPE	Spatio-temporal RMSE · MAPE
Base	314 MW · 3.70 %	314 MW · 3.70 %	314 MW · 3.70 %
BU	484 MW · 5.41 %	320 MW · 3.80 %	318 MW · 3.77 %
OLS	311 MW · 3.66 %	312 MW · 3.67 %	311 MW · 3.66 %
MinT	**301 MW · 3.55%**	**302 MW · 3.56 %**	**291 MW · 3.42 %**

Table 2: RMSE and MAPE on testing data set for top-level series, that is, the entire Alberta region. The use of spatio-temporal clustering always improves the out-of-sample accuracy. Considering the alternative reconciliation approaches, the MinT based on spatio-temporal clustering provides the best results.

reconciliation is applied, resulting in a RMSE of 314 MW and a MAPE of 3.70%. This performance serves as a reference point for evaluating the effects of different reconciliation methods. The BU approach generally yields higher RMSE and MAPE values compared to this baseline. Specifically, in the case of no clustering, BU results in 484 MW for RMSE and 5.41% for MAPE, both of which are significantly worse than the baseline performance. The OLS reconciliation method delivers results that are close to, although a bit more accurate than, the baseline. Surprisingly, the results of the OLS reconciliation do not differ much considering alternative assumptions on the cluster structure. With no clustering, OLS achieves an RMSE of 311 MW and a MAPE of 3.66%, which is slightly better than the baseline. The performance remains relatively stable across different clustering structures. Indeed, for temporal clustering, OLS achieves an RMSE of 312 MW and a MAPE of 3.67%, while for spatio-temporal clustering, the results are almost identical to the no-clustering case with 311 MW and 3.66% for RMSE and MAPE, respectively. The MinT reconciliation, however, provides the best overall performance across all scenarios. Without any clustering, MinT achieves the lowest RMSE of 301 MW and the lowest MAPE of 3.55% compared other reconciliation approaches, thus outperforming both the baseline and the other reconciliation methods. When temporal clustering is applied, the MinT performance does not improve, with RMSE of 302 MW and MAPE of 3.56%. However, the best improvements are found with the spatio-temporal clustering with MAPE dropping to 3.42% and the lowest RMSE equal to 291 MW. This result is the best across all combinations of reconciliation methods and clustering structures, suggesting that spatio-temporal clustering is effective in achieving a more accurate prediction of electricity demand in Alberta.

We then test if the differences in Table 2 are statistically significant. Table 3 shows the results of pairwise Diebold and Mariano (2002) test for the entire Alberta region. Under the null hypothesis, the forecasting approaches in the rows of Table 3 have better accuracy than the model in the column. A negative value of the statistics suggest that the method in the column is more accurate than the one in rows. We compare the MinT reconciliation

Reconciliation approach	Clustering approach	Benchmark model	
		Base	MinT (Spatio-temporal)
Bottom-up	No	-44.46	-53.74
MinT	No	27.12***	-17.44
MinT	Temporal	15.44***	-23.44
MinT	Spatio-temporal	27.06***	-

Table 3: Results of pairwise Diebold and Mariano (2002) test for the entire Alberta region (i.e. top-level series). Under the null hypothesis, the forecasting approach in the rows has better accuracy than the model in the column. The statistics is reported, while *** indicates the significance of the test at 99% confidence.

approaches, that are those providing the best performances out-of-sample for all considered clustering structures, with the base unreconciled forecasts. Next, we compare all the approaches with the MinT reconciliation based on spatio-temporal clustering. The results indicate that all the reconciliation methods, with the only exception on the bottom-up approach, allow improving the base unreconciled forecasts. This result is in line with those of previous papers adopting forecast reconciliation to load forecasting (e.g. see Brégère and Huard, 2022, Mesquita Lopes Cabreira et al., 2024). However, our results indicate that reconciliation approaches based on clustering improve compared to that assuming no cluster structure. In particular, the MinT approach based on spatio-temporal clustering provides the most accurate forecasts, as suggested by the negative statistics in the last column of Table 3. The differences are statistically significant, given that we do not reject the null hypothesis of superior performance of the MinT with spatio-temporal clustering compared to the benchmarks. In sum, our findings suggest that the use of spatio-temporal clustering provides significant improvement in the out-of-sample accuracy of the top-level series, that is, the energy demand in the Alberta region. We then evaluate if significant improvements can be found also for the bottom-level series. Table 4 shows the average out-of-sample RMSE for the 42 AESO areas in Alberta. As in previous table, Base provides the average results for the unreconciled forecasts, while OLS and MinT reconciliation are compared, under no clustering assumption and both temporal and spatio-temporal cluster structure. For the bottom-level series, we find that the MinT reconciliation with temporal clustering structure shows the largest improvements on average, but the improvements obtained with the spatio-temporal clustering are comparable. The RMSE for MinT with spatio-temporal clustering is indeed 17.86, that is much lower than 21.93 obtained with the unreconciled approach. The results obtained with temporal clustering are similarly good, given that the average RMSE equals 17.68, while without clustering structure the average RMSE is about 18. Therefore, in the case of bottom-level series, it seems important the inclusion of the clustering structure. The reconciliation with OLS instead provides a too little improvement compared to the unreconciled case, and in this case the use of spatio-temporal clustering allows for improving the performance most compared to both temporal and no clustering assumptions.

Clustering	No	Temporal	Spatio-temporal
Base	21.93 MW	21.93 MW	21.93 MW
OLS	21.83 MW	21.98 MW	21.78 MW
MinT	17.98 MW	17.68 MW	17.86 MW

Table 4: Average RMSE on testing data set for the 42 bottom-level series, that is, the AESO areas.

The results shown in Table 4 provide average results. In Table 5, we evaluate for how many areas the use of reconciliation improves compared to the base approach. The OLS is confirmed to be not much effective in improving out-of-sample accuracy, as the RMSE is improved only for the 54% of the bottom level series, regardless the clustering approach adopted. Much larger improvements can be found with MinT reconciliation, specifically those using cluster structure. In particular, temporal clustering improves the forecast accuracy for more than 90% of the AESO areas and spatio-temporal clustering about 83%, while without the cluster structure the improvement is found for 80% of the bottom series.

Clustering	No	Temporal	Spatio-temporal
OLS	54.7 %	54.2 %	54.7 %
MinT	80.1 %	90.4 %	83.3%

Table 5: Percentage of AESO areas for which RMSE has been improved with reconciliation.

In sum, our results suggest that considerable improvements can be obtained using the MinT and spatio-temporal clustering, both for top-level and bottom-level series. The spatio-temporal clustering, indeed, seems to be better suited for forecasting electricity demand than temporal clustering.

4. Conclusions

To ensure the reliable operation of the electricity grid 24 hours a day, balancing supply and demand is crucial. Accurate predictions of electricity demand at various levels of aggregation are necessary for this balance. Predicting the overall electricity demand is vital for injecting the appropriate amount of electricity into the grid at any given time. Similarly, forecasting demand at a more localized level is essential to ensure proper distribution across the grid. In recent decades, the field of reconciliation forecasting has significantly advanced, introducing methods making forecasts coherent and more accurate for all the considered levels of aggregation.

When dealing with regional electricity data, one common aggregation approach is ensuring that the sum of individual regional profiles matches the total demand for that region. However, regions often exhibit clustering patterns in terms of the similarity of their demand behaviors. Recent studies on forecast reconciliation have suggested that it is possible to derive more accurate forecasts by accounting for such unknown clustered structures, as opposed to assuming that no structure exists. Clustering methods, however, are different and time series that fit multiple hierarchical structures are known as grouped time series. In line with the assumption that different grouping structures may be present, previous research (e.g. Yang and Shang, 2022) demonstrated that the choice of aggregation method plays a critical role in forecast accuracy. This principle holds true even when the exact group structures are unknown.

In this paper, we evaluate the performance of grouped structures based on the temporal similarity of consumption patterns against spatio-temporal approaches. Our case study focuses on electricity demand in the Alberta region, Canada. We show that spatio-temporal approaches outperform temporal-only methods in forecasting electricity demand in Alberta. In sum, our contribution is threefold. First, we provide new evidence supporting the effectiveness of forecast reconciliation in electricity demand forecasting. Second, we extend forecast reconciliation to account for unknown spatio-temporal cluster structures, demonstrating that incorporating such structures improves forecasting accuracy both at top and bottom levels. Therefore, we introduce a new MinT reconciliation procedure based on spatio-temporal clustering, highlighting its superiority over temporal-only clustering in forecasting electricity demand. Our results show that integrating spatial factors into the clustering process yields more accurate reconciled forecasts. Finally, we offer a simple method to apply the Chavent et al. (2018) ClustGeo algorithm to hourly time series data.

We stress that our procedure is greedy, in the sense that it relies on optimizing the summation matrix (through clustering), and only then the reconciliation matrix. The optimization of these two matrices in unison constitutes a highly complex problem, opening up a whole field of research.

References

Abdi, H., O'Toole, A. J., Valentin, D., and Edelman, B. (2005). Distatis: The analysis of multiple distance matrices. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, pages 42–42. IEEE.

Abdi, H., Williams, L. J., Valentin, D., and Bennani-Dosse, M. (2012). Statis and distatis: optimum multitable principal component analysis and three way metric multidimensional scaling. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(2):124–167.

Alonso, A. M., Nogales, F. J., and Ruiz, C. (2020). Hierarchical clustering for smart meter electricity loads based on quantile autocovariances. *IEEE Transactions on Smart Grid*, 11(5):4522–4530.

Ansari, M. Y., Ahmad, A., Khan, S. S., Bhushan, G., and Mainuddin (2020). Spatiotemporal clustering: a review. *Artificial Intelligence Review*, 53:2381–2423.

Athanasopoulos, G., Hyndman, R. J., Kourentzes, N., and Panagiotelis, A. (2023). Forecast reconciliation: A review. *International Journal of Forecasting*.

Brégère, M. and Huard, M. (2022). Online hierarchical forecasting for power consumption data. *International Journal of Forecasting*, 38(1):339–351.

Bucci, A., Ippoliti, L., Valentini, P., et al. (2023). Analysing spatiotemporal patterns of covid-19 confirmed deaths at the nuts-2 regional level. *Regional Statistics*, 13(2):214–239.

Chavent, M., Kuentz-Simonet, V., Labenne, A., and Saracco, J. (2018). Clustgeo: an r package for hierarchical clustering with spatial constraints. *Computational Statistics*, 33(4):1799–1822.

Chen, B.-J., Chang, M.-W., et al. (2004). Load forecasting using support vector machines: A study on eunite competition 2001. *IEEE transactions on power systems*, 19(4):1821–1830.

Chévez, P., Barbero, D., Martini, I., and Discoli, C. (2017). Application of the k-means clustering method for the detection and analysis of areas of homogeneous residential electricity consumption at the great la plata region, buenos aires, argentina. *Sustainable cities and society*, 32:115–129.

Chicco, G. (2012). Overview and performance assessment of the clustering methods for electrical load pattern grouping. *Energy*, 42(1):68–80.

Díaz, S. P. and Vilar, J. A. (2010). Comparing several parametric and non-parametric approaches to time series clustering: a simulation study. *Journal of classification*, 27:333–362.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144.

Dong, X., Deng, S., and Wang, D. (2022). A short-term power load forecasting method based on k-means and svm. *Journal of Ambient Intelligence and Humanized Computing*, 13(11):5253–5267.

Dudek, G. (2015). Short-term load forecasting using random forests. In *Intelligent Systems' 2014: Proceedings of the 7th IEEE International Conference Intelligent Systems IS'2014, September 24-26, 2014, Warsaw, Poland, Volume 2: Tools, Architectures, Systems, Applications*, pages 821–828. Springer.

Durante, F., Gatto, A., and Ravazzolo, F. (2023). Understanding relationships with the aggregate zonal imbalance using copulas. *Statistical Methods & Applications*, pages 1–42.

D'Urso, P., De Giovanni, L., Disegna, M., and Massari, R. (2019). Fuzzy clustering with spatial–temporal information. *Spatial Statistics*, 30:71–102.

Fan, S. and Hyndman, R. J. (2010). Forecast short-term electricity demand using semi-parametric additive model. In *2010 20th Australasian Universities Power Engineering Conference*, pages 1–6. IEEE.

Fard, A. K. and Akbari-Zadeh, M.-R. (2014). A hybrid method based on wavelet, ann and arima model for short-term load forecasting. *Journal of Experimental & Theoretical Artificial Intelligence*, 26(2):167–182.

Fouedjio, F. (2016). A hierarchical clustering method for multivariate geostatistical data. *Spatial Statistics*, 18:333–351.

Gaillard, P., Goude, Y., and Nedellec, R. (2016). Additive models and robust aggregation for gefcom2014 probabilistic electric load and electricity price forecasting. *International Journal of forecasting*, 32(3):1038–1050.

Goehry, B., Goude, Y., Massart, P., and Poggi, J.-M. (2019). Aggregation of multi-scale experts for bottom-up load forecasting. *IEEE Transactions on Smart Grid*, 11(3):1895–1904.

Goude, Y., Nedellec, R., and Kong, N. (2013). Local short and middle term electricity load forecasting with semi-parametric additive models. *IEEE transactions on smart grid*, 5(1):440–446.

Hyndman, R. J., Ahmed, R. A., Athanasopoulos, G., and Shang, H. L. (2011). Optimal combination forecasts for hierarchical time series. *Computational Statistics & Data Analysis*, 55(9):2579–2589.

Keisler, J., Claudel, S., Cabriel, G., and Brégère, M. (2024). Automated deep learning for load forecasting. *arXiv preprint arXiv:2405.08842*.

Kisilevich, S., Mansmann, F., Nanni, M., and Rinzivillo, S. (2010). *Spatio-temporal clustering*. Springer.

Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., and Zhang, Y. (2017). Short-term residential load forecasting based on lstm recurrent neural network. *IEEE transactions on smart grid*, 10(1):841–851.

Li, H., Li, H., Lu, Y., and Panagiotelis, A. (2019). A forecast reconciliation approach to cause-of-death mortality modeling. *Insurance: Mathematics and Economics*, 86:122–133.

Massaoudi, M., Abu-Rub, H., Refaat, S. S., Chihi, I., and Oueslati, F. S. (2021). Deep learning in smart grid technology: A review of recent advancements and future prospects. *IEEE Access*, 9:54558–54578.

Mattera, R. (2022). A weighted approach for spatio-temporal clustering of covid-19 spread in Italy. *Spatial and Spatio-temporal Epidemiology*, 41:100500.

Mattera, R., Athanasopoulos, G., and Hyndman, R. J. (2024). Improving out-of-sample forecasts of stock price indexes with forecast reconciliation and clustering. *Quantitative Finance*. doi: https://doi.org/10.1080/14697688.2024.2412687.

Mattera, R. and Franses, P. H. (2023). Are african business cycles synchronized? evidence from spatio-temporal modeling. *Economic Modelling*, 128:106485.

Mattera, R. and Franses, P. H. (2025). Forecasting house price growth rates with factor models and spatio-temporal clustering. *International Journal of Forecasting*, 41(1):398–417.

Mesquita Lopes Cabreira, M., Leite Coelho da Silva, F., da Silva Cordeiro, J., Ureta Tolentino, J. M., Carbo-Bustinza, N., Canas Rodrigues, P., and López-Gonzales, J. L. (2024). Comparison between hierarchical time series forecasting approaches for the electricity consumption in the brazilian industrial sector. *Applied Stochastic Models in Business and Industry*.

Morelli, C., Boccaletti, S., Maranzano, P., and Otto, P. (2024). Multidimensional spatiotemporal clustering–an application to environmental sustainability scores in europe. *arXiv preprint arXiv:2405.20191*.

Morgado, E., Martino, L., and San Millán-Castillo, R. (2023). Universal and automatic elbow detection for learning the effective number of components in model selection problems. *Digital Signal Processing*, 140:104103.

Obst, D., De Vilmarest, J., and Goude, Y. (2021). Adaptive methods for short-term electricity load forecasting during covid-19 lockdown in france. *IEEE transactions on power systems*, 36(5):4754–4763.

Pham, D. L. (2001). Spatial models for fuzzy clustering. *Computer vision and image understanding*, 84(2):285–297.

Pierrot, A. and Goude, Y. (2011). Short-term electricity load forecasting with generalized additive models. *Proceedings of ISAP power*, 2011.

Syed, D., Abu-Rub, H., Ghrayeb, A., Refaat, S. S., Houchati, M., Bouhali, O., and Bañales, S. (2021). Deep learning-based short-term load forecasting approach in smart grid with clustering and consumption pattern recognition. *IEEE access*, 9:54992–55008.

Taieb, S. B. and Hyndman, R. J. (2014). A gradient boosting approach to the kaggle load forecasting competition. *International journal of forecasting*, 30(2):382–394.

Voulis, N., Warnier, M., and Brazier, F. M. (2018). Understanding spatio-temporal electricity demand at different urban scales: A data-driven approach. *Applied Energy*, 230:1157–1171.

Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244.

Wickramasuriya, S. L., Athanasopoulos, G., and Hyndman, R. J. (2019). Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. *Journal of the American Statistical Association*, 114(526):804–819.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. chapman and hall/CRC.

Yang, Y. and Shang, H. L. (2022). Is the group structure important in grouped functional time series? *Journal of Data Science*, 20(3):303–324.

Zhang, B., Panagiotelis, A., and Li, H. (2024). Constructing hierarchical time series through clustering: Is there an optimal way for forecasting? *arXiv preprint arXiv:2404.06064*.

Zhang, Z., Cheshmehzangi, A., and Ardakani, S. P. (2021). A data-driven clustering analysis for the impact of covid-19 on the electricity consumption pattern of zhejiang province, china. *Energies*, 14(23):8187.