Simulating Low and High-Frequency Energy Demand Scenarios in a Unified Framework – Part II: High-Frequency Simulation

Guilherme Bodin, Raphael Saavedra, Érica Telles, Thuener Silva, André Milhorance, Cristiano Fernandes, Alexandre Street, Armando Leite
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marquês de São Vicente, 225 – Edifício Cardeal Leme, Sala L435
guilherme.b.moraes@gmail.com

RESUMO

O segundo trabalho desta série propõe uma nova metodologia para simular cenários de demanda de energia de longo prazo e alta frequência em cada barra de um sistema de distribuição elétrica. O modelo proposto gera cenários de alta frequência e alta dimensão em base horária para cada barra do sistema em função dos cenários de baixa frequência e baixa dimensão simulados no primeiro trabalho desta série. O método proposto é baseado em um processo de desagregação que é treinado com dados históricos e para gerar cenários de demanda de longo prazo. Um estudo de casos com dados reais de uma empresa brasileira de distribuição de energia é apresentado. Reiteramos que este método pode ser útil para várias aplicações em sistemas de potência.

PALAVRAS CHAVE. Demanda de energia. Modelos de desagregação. Regularização.

ABSTRACT

The second of this two-part paper proposes a novel methodology to simulate long-term, high-frequency scenarios for the power demand in each bus of a distribution system. The proposed model generates high-frequency and high-dimensional scenarios, on an hourly basis for each bus of the system, as a function of the low-frequency and low-dimensional scenarios simulated by the first part of this work. Hence, the proposed method relies on a disaggregation procedure that is trained within observed data and applied to long-run simulated scenarios. A case study with real data from the Brazilian power system is presented and relevant conclusions are drawn. We highlight that this method can be useful for a wide range of applications in power systems.

KEYWORDS. Energy demand. Disaggregation models. Regularization.

This work was partially supported by the Energisa Group through R&D project ANEEL PD-00405-1701/2017.

https://proceedings.science/p/85334?lang=pt-br

1. Introduction

In the first paper of this two-part series [Saavedra et al., 2018], we have motivated the need to correctly simulate the maximum power imported from the transmission system to feed the optimization framework that decided the optimal amount of usage of the transmission system. One possible way to obtain accurate scenarios of imported power from the transmission system is to model the power demand in all buses of the distribution system (load/renewable energy injection buses) in hourly basis[1]. This allows capturing the effect of network contingencies and topological changes by combining load with contingency scenarios in a joint high-dimensional and high-frequency simulation. However, the characterization of such a high-frequency and high-dimensional process is a challenging task, since many different seasonal cycles and stylized facts are found in the many different time scales (hourly, daily, weekly, monthly, and so on) for each bus.

To solve this problem we propose to split the problem in two steps. In the first step we model the aggregated (low-dimension) energy/demand consumption time series comprising components of the total system demand in a monthly basis (low-frequency). This part has been addressed in the first part [Saavedra et al., 2018]. In this paper we address the second step of the proposed methodology. More specifically, we propose a disaggregation process that receives as input the low-dimensional and low-frequency signal representing the total consumption series and outputs the the hourly power demand for each load/injection bus.

Many approaches have been proposed in the literature to load prediction. [Handschin and Dornemann, 1988] proposes the use of normalized load curves to model different types of buses on a distribution system and forecast load using daily time series such as hourly temperature and brightness. [Panapakidis and Papagiannis, 2014] analyses Greek electric distribution system using artificial neural networks, while [Franco et al., 2017] uses Knowledge Discovery in Databases techniques to very short-term forecasting. [Sinha and Mondal, 1999] presents dynamic state techniques and [Dordonnat et al., 2008] used a periodic state-space model to forecast the hourly one-day-ahead French electricity load. [Dordonnat et al., 2009] extended this model with the addition of dynamic factors and later the same authors presented in [Dordonnat et al., 2012] a dynamic multivariate periodic regression model with a similar goal. [Soares and Medeiros, 2008] proposed a two-level seasonal autoregressive model to forecast hourly electricity load, with a Brazilian case study. Aggregate methods such as clustering and multi-layer perceptron (MLP) in [Salgado et al., 2009] and forecast-aided state estimator/MLP [Amjady, 2007] were also developed. However, all previously reported works present methods to predict short-term bus load, forecasting only up to a few weeks ahead. Notwithstanding, none of the previously reported works address the problem of simulating high-frequency load scenarios for a long-term horizon.

The proposed methodology utilizes linear programming to perform a regularized quantile regression between the monthly signal and the hourly demand for each bus of the system. Note that the desegregation methodology proposed in this work applies to a wide range of power system applications. Other relevant examples of applications are: i) day-ahead unit commitment or economic dispatch applications, where aggregated power forecasts must be decomposed throughout the buses in the system, and ii) transmission planning problems, where typically economical studies determine a long-term aggregated load forecast for the system-wide demand and a disaggregation process is needed. Thus, this framework can be used to predict any high-dimensional and high-frequency phenomenon taking into consideration available low-dimensional and low-frequency forecasts. In our proposed framework, each bus, typical hour of the day and day of the week has its own regression model. Furthermore, additional explanatory variables such as climate indexes (temperature, rain, solar radiation, etc) can also be accounted for.

The remaining sections of this paper are organized as follows. Section 2 contains the problem description and the utilized notation. Section 3 presents the model employed in this paper.

[1] In this work we consider as load the net power extraction of each bus, i.e., demand minus renewable generation injection.

https://proceedings.science/p/85334?lang=pt-br

A case study with real data from the Brazilian system is presented in Section 4. Lastly, Section 5 presents the conclusion and raises some ideas being developed for future work.

2. Problem description

In this section we explain the problem context and the main notation used throughout this work.

Nomenclature

Functions

$d(t)$ Day of instant t

$h(t)$ Hour of instant t

$m(t)$ Month of instant t

Other Symbols

λ_w Regularization parameter

$w_{d(t),h(t),s,i}$ Regression coefficient of day $d(t)$, hour $h(t)$, low frequency serie s and bus i

Sets

$\mathcal{B}$ Set of low-level buses

$\mathcal{D}$ Set of days of the week $\mathcal{D} = \{Sun, Mon, ..., Sat\}$

$\mathcal{H}$ Set of hours of the day $\mathcal{H} = \{1, 2, ..., 24\}$

$\mathcal{M}$ Set of months

$\mathcal{S}$ Set of explanatory low-frequency series

$\mathcal{T}_{hist}$ Set of high-frequency instants of time (hours or 15 minutes intervals) of the in-sample power demand history

$\mathcal{T}_{sim}$ Set of high-frequency instants of time (hours or 15 minutes intervals) of the simulation horizon

Ω Set of scenarios

Time Series

$\hat{p}^{\alpha}_{i,t}(\omega)$ Power demand scenario at instant t on bus i for the α quantile with low-frequency uncertainty characterized

$E_{s,m(t)}$ Low frequency serie s at intant t

$p_{i,t}$ Power demand at instant t on bus i

$p_{i,t}(\omega)$ Power demand scenario at instant t on bus i with low and high-frequency uncertainty characterized

The problem is to estimate a multiple regression model that will allow one to build a forecast and an ensemble of scenarios for the power demand of each bus of the system, i.e., $\{\hat{p}_{i,t}\}_{t\in\mathcal{T}_{sim},i\in\mathcal{B}}$ and $\{p_{i,t}(\omega)\}_{t\in\mathcal{T}_{sim},i\in\mathcal{B},\omega\in\Omega}$, respectively. To do so, we dispose of historical data for each bus and hour within a time window, $\{p_{i,t}\}_{t\in\mathcal{T}_{hist},i\in\mathcal{B}}$, a set of low-frequency/dimension historical data, $\{E_{s,m(t)}\}_{t\in\mathcal{T}_{hist},s\in\mathcal{S}}$, and an ensemble of scenarios within a simulation time window (in the future), $\{E_{s,m(t)}(\omega)\}_{t\in\mathcal{T}_{sim},s\in\mathcal{S},\omega\in\Omega}$ (generated in [Saavedra et al., 2018]). The methodology to approach this problem is described in the next section.

https://proceedings.science/p/85334?lang=pt-br

3. Disaggregation Methodology

The proposed regression model receives as input variables the low-frequency/dimen-sion values and returns the power demand for each bus and hour within a month. The input values can be obtained through simulation processes as described in [Saavedra et al., 2018]. The model is an array of multiple quantile regressions for each typical day of the week and hour of the day. To explain the power demand history let's assume there is a set of explanatory series such that $\{E_{s,m}\}_{s\in\mathcal{S},m\in\mathcal{M}}$ at instant m, where m stands for a low-frequency time index such as months. To bridge low frequency and the high frequency datasets, we use a map function, $m(t)$, that associates hour t with its respective month m. Hence, for each bus, i, of the system, the power demand is estimated by the following convex optimization problem

$$\min_{\epsilon,w} \frac{1}{|\mathcal{T}_{hist}|} \sum_{t\in\mathcal{T}_{hist}} (\alpha(\epsilon_{i,t})^+ + (1-\alpha)(\epsilon_{i,t})^-) + \frac{\lambda_w}{|\mathcal{D}||\mathcal{H}||S|} \sum_{\substack{s\in\mathcal{S}\\ d\in\mathcal{D}\\ h\in\mathcal{H}}} |w_{d,h,s,i}|, \tag{1}$$

subject to

$$\epsilon_{i,t} = p_{i,t} - w_{d(t),h(t),0,i} + \sum_{s\in\mathcal{S}} E_{s,m(t)} w_{d(t),h(t),s,i} \quad \forall t \in T_{hist}. \tag{2}$$

Where, $(\cdot)^+$ and $(\cdot)^-$ stand for the positive and negative truncation functions, respectively.

Note that the objective function is composed of two summation terms. The first term stands for the penalized prediction error, whereas the second one accounts for a norm-one penalization to regularize the model following the LASSO scheme [Robert,]. To solve this problem in practice, a linear programming equivalent formulation can be used. To do that, the error variable, $\epsilon_{i,t}$, in (2) is decomposed in two new variables representing its positive and negative parts, i.e., $\epsilon_{i,t} = \epsilon^+_{i,t} - \epsilon^-_{i,t}$. In the objective function, the first summation term, which precisely accounts for those two parcels of the error are then updated as follows: $(\epsilon_{i,t})^+ = \epsilon^+_{i,t}$ and $(\epsilon_{i,t})^- = \epsilon^-_{i,t}$. The same standard linear programming transformation approach applies to the second summation term of the objective function related to the LASSO regularization penalty.

Finally, to adjust the regularization parameter λ_w, balancing between good out-sample results and a better in-sample fit, Bayesian principles [Huri and Feder, 2016] can be used. We perform a k-fold cross-validation method, where the optimal parameters are estimated for $\frac{k-1}{k}$ of the database, previously divided in k folds, and evaluated in the remaining one. The evaluation of a given λ_w is made by repeating this process k times, one for each fold (assumed as out-of-sample), and associating this λ_w a score. In this work, we use the sum of the R^2 for all k out-of-sample predictions as a evaluation score.

Once the problem is solved, we can use the optimum (estimated) vector of parameters, $\hat{\mathbf{w}}$, to generate both a forecast and an ensemble of scenarios for future power demand. Note that in order to evaluate the estimator for the entire system we need to solve this linear problem for each bus $i \in \mathcal{B}$.

To produce the high-frequency load/injection forecast for a given bus i, we take all the scenarios generated for the low-frequency/dimension explanatory variables, $\{E_{s,m(t)}(\omega)\}_{t\in\mathcal{T}_{sim},s\in\mathcal{S},\omega\in\Omega}$ as per [Saavedra et al., 2018] and evaluate the estimated high-frequency quantile model as follows

$$\hat{p}^{\alpha}_{i,t}(\omega) = \hat{w}_{d(t),h(t),0,i} + \sum_{s\in\mathcal{S}} E_{s,m(t)}(\omega)\hat{w}_{d(t),h(t),s,i} \quad \forall t \in T_{sim}, \omega \in \Omega, i \in \mathcal{B}. \tag{3}$$

The ensemble of scenarios generated by the above described process constitutes a α-quantile prediction for multiple low-frequency/dimension scenarios. Those scenarios account solely for the uncertainty in the long-term trend (demand growth, generation seasonality, and climate variation). Notwithstanding, to account for the stochasticity related to the high-frequency estimates,

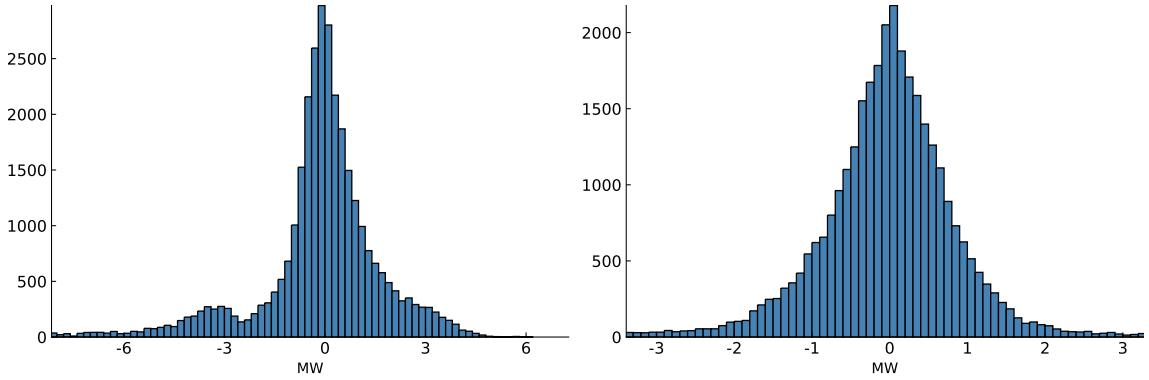
https://proceedings.science/p/85334?lang=pt-br

one should use the median ($\alpha = 0.5$) and store the estimated errors $\{\epsilon_{i,t}\}_{i\in\mathcal{B},t\in\mathcal{T}_{hist}}$, to be used in a sampling scheme. In this sampling procedure, the historical errors give rise to a sample set from which errors could be randomly sampled preserving the spacial dependency among buses. Under the hypothesis that estimated errors are not correlated to the explanatory variables, a given scenario ω of power demand can be obtained by paring the explanatory variable scenarios with a sampled error randomly, i.e.,

$$p_{i,t}(\omega) = \hat{w}_{d(t),h(t),0,i} + \sum_{s\in\mathcal{S}} E_{s,m(t)}(\omega)\hat{w}_{d(t),h(t),s,i} + \epsilon_{i,t}(\omega) \quad \forall t \in T_{sim}, \omega \in \Omega, i \in \mathcal{B}. \tag{4}$$

4. Case study

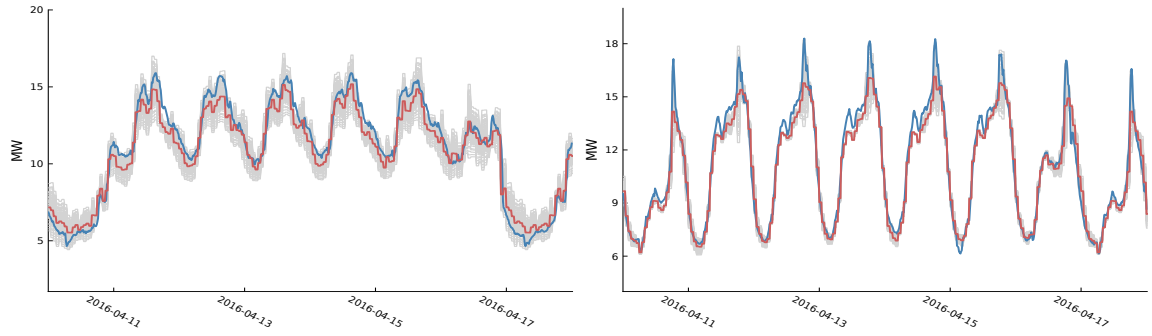
An implementation of the disaggregation model was done in Julia with data from a Brazilian energy distribution company (Energisa S.A.). We applied the disaggregation methodology using data from 2015 to train the model (in sample set) and the first nine months of 2016 (out-of-sample set) to evaluate the results. We will use only one explanatory serie, the monthly energy consumed by the system. The scenarios were generated in [Saavedra et al., 2018]. To illustrate results graphics in this section refer to three different low-level buses $i = 1, 2$.



(a) bus 1 (b) bus 2

Figure 1: Histogram of in-sample errors

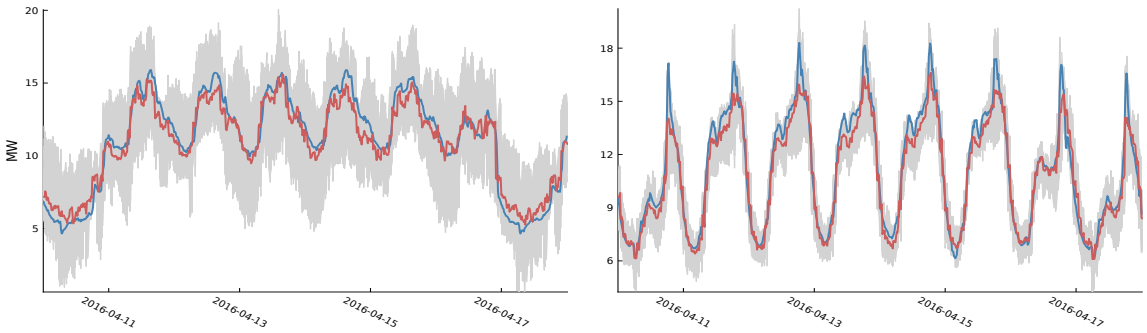
The next thing would be to simulate the power demand for the nine following months, taking the second week of April as reference to illustrate the results we can see the difference between accounting only for the low-frequency stochasticity described in Equation (3) and accounting the low and high-frequency stochasticity added in Equation (4) in Section 3.



(a) $p^{0.5}_{1,t}(\omega)$ (b) $p^{0.5}_{2,t}(\omega)$

Figure 2: Simulation of $p^{0.5}_{i,t}(\omega)$, vertical axis: (MW) – horizontal axis: time (hours)

https://proceedings.science/p/85334?lang=pt-br



(a) $p_{1,t}(\omega)$ (b) $p_{2,t}(\omega)$

Figure 3: Simulation of $p_{i,t}(\omega)$, vertical axis: (MW) – horizontal axis: time (hours)

The measurements of model specification of each bus are shown in Table 1

Table 1: Results

Bus	1		2	
	In-sample	Out-of-sample	In-sample	Out-of-sample
MAE (MW)	1.49	1.31	0.704	0.774
SMAPE	0.0442	0.0341	0.0161	0.0185
R^2	0.486	0.691	0.886	0.909

5. Conclusions and future work

This paper proposes a methodology to simulate future long-term and high-frequency bus load scenarios based on a novel framework. The out-of-sample goodness-of-fit results and diagnostics indicate that the model is well-specified. The disaggregation rise for the possibility to apply risk analysis methods for any application that requires bus load prediction due to the uncertainty characterization.

Future work consists in improving methodology by including lags as explanatory variables for a more robust time-series analysis. Ideas like the characterization of unusual bus load curves as in holydays can be explored.

https://proceedings.science/p/85334?lang=pt-br

References

Amjady, N. (2007). Short-term bus load forecasting of power systems by a new hybrid method. *IEEE Transactions on Power Systems*, 22(1):333–341.

Dordonnat, V., Koopman, S. J., and Ooms, M. (2009). Dynamic factors in state-space models for hourly electricity load signal decomposition and forecasting. In *Power & Energy Society General Meeting, 2009. PES'09. IEEE*, pages 1–8. IEEE.

Dordonnat, V., Koopman, S. J., and Ooms, M. (2012). Dynamic factors in periodic time-varying regressions with an application to hourly electricity load modelling. *Computational Statistics & Data Analysis*, 56(11):3134–3152.

Dordonnat, V., Koopman, S. J., Ooms, M., Dessertaine, A., and Collet, J. (2008). An hourly periodic state space model for modelling French national electricity load. *International Journal of Forecasting*, 24(4):566–587.

Franco, E. F., Ohishi, T., and Salgado, R. M. (2017). Very short-term bus reactive load forecasting models based on kdd approach. In *2017 IEEE 7th International Conference on Power and Energy Systems (ICPES)*, pages 34–41.

Handschin, E. and Dornemann, C. (1988). Bus load modelling and forecasting. *IEEE Transactions on Power Systems*, 3(2):627–633.

Huri, N. and Feder, M. (2016). Selecting the lasso regularization parameter via bayesian principles. In *2016 IEEE International Conference on the Science of Electrical Engineering (ICSEE)*, pages 1–5.

Panapakidis, I. P. and Papagiannis, G. K. (2014). Application of the load profiling methodology in short-term bus load forecasting. In *MedPower 2014*, pages 1–8.

Robert, T. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(3):273–282.

Saavedra, R., Bodin, G., Telles, E., Silva, T., Milhorance, A., Fernandes, C., Street, A., and Leite, A. (2018). Simulating low and high-frequency energy demand scenarios in a unified framework – part i: Low-frequency simulation. In *L Simpósio Brasileiro de Pesquisa Operacional*.

Salgado, R. M., Ballini, R., and Ohishi, T. (2009). An aggregate model applied to the short-term bus load forecasting problem. In *2009 IEEE/PES Power Systems Conference and Exposition*, pages 1–8.

Sinha, A. K. and Mondal, J. K. (1999). Dynamic state estimator using ann based bus load prediction. *IEEE Transactions on Power Systems*, 14(4):1219–1225.

Soares, L. J. and Medeiros, M. C. (2008). Modeling and forecasting short-term electricity load: A comparison of methods with an application to Brazilian data. *International Journal of Forecasting*, 24(4):630–644.

https://proceedings.science/p/85334?lang=pt-br