Applied Energy 282 (2021) 116180

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Very short term load forecasting of residential electricity consumption using the Markov-chain mixture distribution (MCM) model

Joakim Munkhammar *, Dennis van der Meer, Joakim Widén

Department of Engineering Sciences, Uppsala University, SE-751 21 Uppsala, Sweden

ARTICLE INFO

Keywords:
Electricity consumption
MCM model
Probabilistic forecasting

ABSTRACT

This study utilizes the Markov-chain mixture distribution model (MCM) for very short term load forecasting of residential electricity consumption. The model is used to forecast one step ahead half hour resolution residential electricity consumption data from Australia. The results are compared with Quantile Regression (QR) and Persistence Ensemble (PeEn) as advanced and simple benchmark models. The results were compared in terms of reliability, reliability mean absolute error (rMAE), prediction interval normalized average width (PINAW) and normalized continuous ranked probability score (nCRPS). For 10 steps conditioning for QR and PeEn, the MCM results were on par with QR, and superior to PeEn. As a sensitivity analysis, simulations were performed where the number of data points for conditioning QR and PeEn was varied and compared to the MCM output, which is based on only one data point for conditioning. It was shown that in terms of nCRPS and rMAE the QR results converged towards the MCM results for lower number of conditioning points included in QR. The nCRPS of PeEn never reached the superior MCM and QR results, but in rMAE, for number of conditioning points above 24, PeEn was the most reliable. Based on the sparse complexity design of MCM, high computational speed and competitive performance, it is suggested as a candidate for benchmark model in probabilistic forecasting of electricity consumption.

1. Introduction

Electricity consumption forecasts have been a fundamental business problem for over a century [1]. This business includes a wide span of energy system topics such as operations, energy trading and power systems planning [1]. Short-term electricity consumption forecasting, as a special case, is of importance for demand-side management [1,2], also for the policy making to reduce building energy use [3]. The majority of electricity consumption research has focused on point forecasting, with an increase of probabilistic load forecasting in the last few decades [1,4].

In general, probabilistic load forecasts aim to provide probabilistic prediction of the future state [4], which may be more valuable to stakeholders than single point predictions [1,5]. In general, a probabilistic forecast contains information regarding the uncertainty of the random variable instead of only the expected value. This gives decision-makers to run, e.g., worst-case optimizations (with an extreme quantile) or stochastic optimizations (with samples from the distribution), which can increase the revenue. Probabilistic models can produce median or mean values of the predictive distribution, which can be interpreted as point forecasts [4]. A large difference between point forecasts and probabilistic forecasts is also the use of metrics for quantifying the forecast, where for example root mean square estimation (RMSE) is popular for point forecasts, and Continuous Ranked Probability Score (CRPS) may be used in probabilistic forecasts [4]. However, it has been argued that probabilistic load forecasting, to a higher degree than point load forecasts, lacks well-established evaluation metrics [1].

Probabilistic forecasting of electricity consumption was extensively reviewed in [1,4], and a conclusion was that there is a wide variety of forecasting resolutions and model choices in the literature. There is a large amount of different models that have been used in the literature, such as stochastic time-series combined with Bayesian Inference (BI) [6], hybrid Kalman-Filters [7], Artificial Neural Networks [8,9], Quantile regression (QR) [10,11], Gaussian and lognormal processes [5,12] and Random Forests [13].

In terms of probabilistic very short term load forecasting, i.e. forecasting within a day, an early example is [7], where hybrid Kalman Filters were used to forecast load with a temporal resolution of 5 min. In [10] half-hour resolution forecasting of electricity consumption was performed with quantile regression (QR) combined with gradient

* Corresponding author.
E-mail addresses: joakim.munkhammar@angstrom.uu.se (J. Munkhammar), dennis.vandermeer@angstrom.uu.se (D. van der Meer), joakim.widen@angstrom.uu.se (J. Widén).

https://doi.org/10.1016/j.apenergy.2020.116180
Received 10 August 2020; Received in revised form 5 October 2020; Accepted 30 October 2020
Available online 10 November 2020

boosting. In [12] Gaussian processes, along with an ARIMA benchmark-model, were used for forecasting half-hour resolution residential load. In [5] half-hour resolution of residential load was forecasted using a lognormal process, and compared it with a Gaussian process as benchmark. In [13], hour resolution forecasts of electricity use were performed using Random Forests, compared with regression trees and support vector regression as benchmarks. Generally, based on the literature, a conclusion is that there is a lack of both standards for benchmark models and performance metrics in probabilistic load forecasting [4]. This is an effect of the relative immaturity of the field of PLF compared to deterministic load forecasting and is a matter still to be settled [1,4].

1.1. Research gaps and scientific contributions

Markov-chain models have been developed and used for monthly load forecasting [14], but there appears to be no example of short term probabilistic load forecasting using Markov chains. In other energy fields, such as wind power we have for example [15], where a first and second order Markov-chain model was used. We also have Markov-chain forecasting of solar irradiance [16,17], where the particular Markov-chain mixture distribution model (MCM) was shown to have high predictive accuracy, in particular for short horizons.

There are no preferred benchmark methods for probabilistic load forecasting [4]. There is also a wide variety of different settings for performance metrics for probabilistic forecasts, where proper probabilistic forecasting metrics are not always used [4]. The MCM model has not been investigated as forecasting method for very short term load forecasting previously.

This study utilizes the MCM model for probabilistic forecasting of residential electricity consumption of individual houses. The focus of this study is probabilistic forecasting in terms of model development, simulations and performance metrics. The focus is on very short term load forecasting, for which the MCM model was shown to be particularly useful in forecasting solar irradiance [16,17]. Results are compared with simulations from a Quantile Regression (QR) model, as an advanced form of benchmark, and a Persistence Ensemble (PeEn) model, as a simple form of benchmark. The standard probabilistic forecasting metrics of reliability diagrams, PINAW and normalized CRPS (nCRPS), are used. The paper aims to present a potential benchmark for probabilistic load forecasting, and proper use of performance metrics.

This paper is organized as follows. In Section 2 the models, data sets and performance metrics are presented, and in Section 3 the results are presented. Finally in Section 4 conclusions are drawn and the prospects for MCM in electricity consumption forecasting are presented.

2. Methodology

2.1. Markov-chain mixture (MCM) distribution model

The MCM model used in this study is based on the N-state Markov-chain mixture distribution model, which was developed in [18] for generating synthetic stationary time-series. The main concept is that a given time-series can be approximately represented by a Markov-chain mixture distribution based on weighted disjoint distributions, a conceptually straight-forward and computationally efficient form of a *Hidden Markov Model*, see [19, p. 312] for general information on Hidden Markov Models, and see [16,17] for the particular aspects of the MCM model.

The MCM model is based on a single observation data point for conditioning the model. In this study that data point is the electricity consumption in one time-step, and the model produces a probabilistic forecast for the electricity use in the next time-step.

The model is based on (1) binning a training data set and (2) setting up a transition matrix based on calculated transition probabilities between the bins. Once trained, one input test data value is used to estimate the non-parametric probability distribution for the next time-step. This is then repeated for the entire test data set. See Fig. 1 for a conceptual illustration of the MCM model forecasting procedure.

In formal terms, the MCM model is defined by the following procedure [16]:

1. Define training and test residential electricity use time-series bounded within limits $[a, b]$.
2. Choose a number of bins N for the model.
3. Bin the training time-series into N bins evenly divided on the range $[a, b]$.
4. Estimate an $N \times N$ transition matrix M from the transitions occurring between bins in the training data time-series.
5. Take an observation $x(t)$ for some time-step t of the test data and determine which bin $i \in [1, \dots, N]$ it belongs to.
6. The i:th row of the transition matrix M then corresponds to a piece-wise uniform distribution forecast for $x(t+1)$.
7. For forecasting the electricity consumption in the next time-step, alter the time-step $t \rightarrow t+1$, then use the new observation $x(t+1)$ and repeat from step 5. This makes the forecasting method recursive.

The estimation of the probabilities in the transition matrix M, in step 4, can implementation-wise be performed by counting the number of transitions from one state to another in the training data. That is, the transition matrix of probabilities M_{ij} of state transitions $i \rightarrow j$ for $i, j \in [1, N]$ can be estimated from the training data according to:

$$M_{ij} = \frac{n_{ij}}{\sum_{k=1}^{N} n_{ik}}, \tag{1}$$

where n_{ij} are the number of transitions from state i to state j from one time-step to the next (including transitions from the state to itself) throughout the entire training time-series. The formula in (1) estimates the number of transitions from state i to state j and, to keep the total probability of transitions to unity, normalizes it with the sum of transitions from state i. Note that the transition matrix dependents on resolution, and in this study half hour resolution is used. Note that the model needs to be re-trained if the resolution is changed.

In conclusion, the predictive probability distribution for $x(t)$ in step 6 is equal to the piece-wise uniform distribution corresponding to the i:th row of the transition matrix. In formal terms, the probability density function f as prediction for $x(t+1)$ can be expressed as:

$$f(x) = \sum_{i=1}^{N} M_{in} f_i(x) \tag{2}$$

where n is the bin of the observation data point, and each $f_i(x) = \mathcal{U}(x_i, x_{i+1})$ is a uniform probability density distribution defined for $x_i = \frac{i-1}{N}(b-a)+a$.

In this study $N = 100$ bins were used, which is an arbitrary number $N > 20$, as the model performance has been shown to saturate for high N [16]. Generally, for lower number of bins, non-uniform distributions such as Gaussian, log-normal or polynomial, could be useful to improve the forecast accuracy.

It should also be emphasized, in particular for high number of states (bins) N, that there may be states which have zero probability of transition from (even to stay in the state). This could occur if there were no transitions into (and out of) this particular bin in the training data set.

If a value in the test data set resides within such a bin, the model, according to the algorithm, is unable to estimate a probability density distribution for the next time-step. This never occurred during the simulations in this study, but if this indeed would occur, the authors recommend using a uniform distribution over the entire range $[a, b]$ as prediction for the transition to the next states.

An additional potential problem is that observation data points may reside outside the range of the training data, and thus outside the model

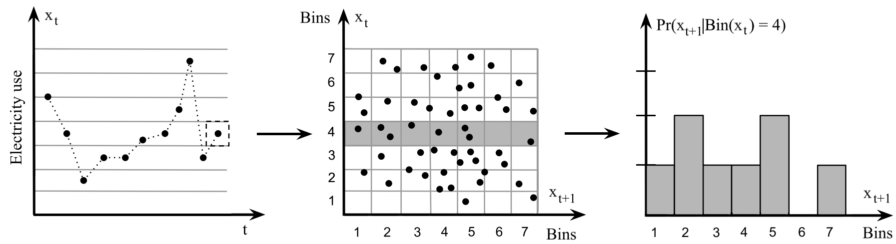

Fig. 1. Schematic illustration of the trained MCM model with a time-series of observations (left), followed by identification of bin for the latest observation (middle) and finally the non-parametric piece-wise uniform predictive distribution for next step (right).

range. This was not encountered in the simulations in this study either, but if it occurs, a reasonable solution to this would be to assume that the observed data point belongs to the bin which is closest to that point of observation. Typically, this would be a bin at the boundary of the range.

Although the MCM model is a probabilistic forecasting model, it should be emphasized that it can be used for point forecasting as well, and then the point forecast would be defined as either the expected value or the median of the predictive distribution. See [16] for examples of this.

The MCM model is available in Python and R via *SheperoMah/MCM-distribution-forecasting* at GitHub [20].

2.2. Benchmark models

2.2.1. Persistence ensemble

As a simple benchmark model, the *persistence ensemble* (PeEn) is utilized in this study. It is a common probabilistic forecasting benchmark model in which the predicted distribution for, in this case, $x(t+1)$ for one step ahead electricity use forecast is defined as the empirical distribution of h previous electricity consumption data points $x(t), x(t-1), \ldots, x(t-(h-1))$ [4,16,21]. Here h was chosen to be 10, which is an arbitrary choice. However, to mend this ambiguity, a sensitivity analysis was performed, and in that case h was varied, see Section 2.5. An intuitive interpretation of the persistence ensemble is that it generalizes the persistence forecast benchmark of deterministic forecasting to a probabilistic forecasting setting [4,21].

2.2.2. Quantile regression

As an alternative to the persistence ensemble, we use *quantile regression* (QR), as an advanced form of benchmark, to produce probabilistic forecasts of the electricity consumption in this study. We utilize the *quantreg* package in R [22,23] to implement the electricity use adapted QR model. A brief introduction to the general QR model follows, including the particular application to forecasting electricity use.

QR – in similarly with linear regression – proposes a linear relationship between the explanatory variables x and output y [24]:

$$y = x\boldsymbol{\beta} + \epsilon, \tag{3}$$

where $\boldsymbol{\beta}$ is a vector of parameters and ϵ is the random error. In order to learn the parameters $\hat{\boldsymbol{\beta}}$, the following minimization problem has to be solved:

$$\hat{\beta}_\tau = \arg_\beta \sum_{t=1}^{T} \rho_\tau \left(y(t) - x(t)\beta\right), \tag{4}$$

where τ represents the quantile probability ($0 < \tau < 1$) and ρ_τ is the pinball loss function, a typical loss function that is used in QR forecasting, defined as [24]

$$\rho_\tau(u) = \begin{cases} \tau u & \text{if } u \geq 0 \\ (\tau - 1)u & \text{if } u < 0. \end{cases} \tag{5}$$

In accordance with (4) each quantile is estimated independently. In that, there is a risk that the model could violate the monotonicity property [10]:

$$\hat{q}_{\tau_1} \leq \hat{q}_{\tau_2} \ \forall \ \tau_1, \ \tau_2 \text{ subject to } \tau_1 \leq \tau_2, \tag{6}$$

where $\hat{q}_\tau$ is the quantile forecast with nominal probability τ. A solution to this, which is used in this paper, is to sort quantiles in ascending order if this happens.

In this study, $y = y(\tau)$ is the forecasted quantile value of quantile $\tau \in [1, \ldots, Q]$ of the electricity consumption $x(t+1)$ for the forecasted time-step $t+1$. This is calculated from previous time-step values for $x(t)$ and the previously trained $\hat{\boldsymbol{\beta}}_\tau$ parameter. For this study, the number of quantiles Q was, as is common, set to 19, evenly spaced on $[0.05, \ldots, 0.95]$.

2.3. Performance metrics

In order to assess the performance of the forecasts, a series of probabilistic forecasting performance metrics were used: Reliability Mean Absolute Error (rMAE), Prediction Interval Normalized Width (PINAW), and normalized Continuous Ranked Probability Score (nCRPS). In addition to this, reliability diagrams are also presented in order to visualize reliability.

As a means of comparing models in terms of reliability, a mean absolute deviation from the diagonal and mean quantile value $q(\tau)$ is estimated, denoted reliability MAE (rMAE) and is expressed as:

$$\text{rMAE} = \frac{1}{Q} \sum_{\tau=1}^{Q} \left| q(\tau) - \frac{\tau}{Q} \right|, \tag{7}$$

where Q is the number of quantiles. The score was also used in [16]. In the literature, this score has also been referred to as the *discrepancy* or *reliability index* [25].

The prediction interval sharpness is measured by the PINAW score, which is defined as [4]:

$$\text{PINAW}(\eta) = \frac{1}{TR} \sum_{t=1}^{T} \left(q_{t,\tau=1-\frac{1-\eta}{2}} - q_{t,\tau=\frac{1-\eta}{2}} \right) \tag{8}$$

where T is the number of data points of the time-series, η is the nominal coverage ratio, $q_{t,\tau}$ is the quantile τ for observation t, and R is a normalization constant, commonly defined as the difference between the maximum and the minimum of the test data set. Here the normalization constant is defined as $R = |b - a|$, where $[a, b]$ is the range of the testing data set.

The sharpness and reliability is quantified by the nCRPS score, which for observation t is defined as:

$$\text{nCRPS}_\text{t} = \frac{1}{R} \int_{-\infty}^{\infty} (F_t(x) - \mathbb{1}_{y_t \leq x})^2 \text{dx} \tag{9}$$

where F_t is the cumulative distribution function of the predictive distribution, y_t is the observation and R is the range of the testing data set. The nCRPS is in this study averaged over each years worth of forecasts.

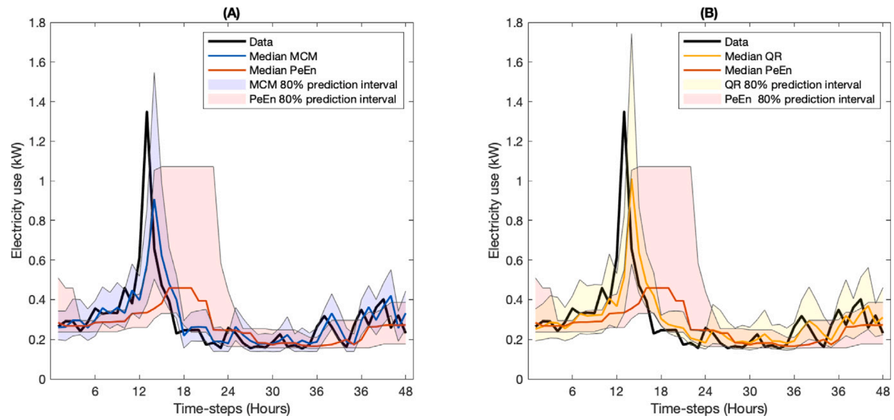

Fig. 2. Example electricity use data, median and 80 percent prediction interval forecasts from forecasting 48 h (starting first half hour after midnight) of house 69 year 1, with day 2 and 3 chosen in that data set, for (A) MCM and PeEn and (B) QR and PeEn.

Table 1
Mean value (kW)/standard deviation (kW) for each of the electricity use data sets used in this study.

Year	House 69	House 74	House 157	House 211	House 274
1	0.52/0.30	0.26/0.40	0.43/0.44	0.31/0.22	0.28/0.47
2	0.45/0.27	0.29/0.43	0.48/0.46	0.27/0.16	0.24/0.41
3	0.38/0.23	0.33/0.49	0.54/0.43	0.24/0.20	0.28/0.47

2.4. Data sets

The data used in this study comprises residential electricity consumption from the Sydney metropolitan area, Australia [26]. This data set is publicly available and contains 300 de-identified residential customers, of whom their consumption has been measured on half hourly basis from 1 July 2010 until 30 June 2013. After a careful data cleaning process, properly described in [27], data from 54 customers for three years remain, of which we have randomly selected customer numbers 69, 74, 157, 211 and 274 for forecasting in this study. For each customer there were in total three year's worth of data, each year having 17'536 data points. Statistical features of the electricity use data sets in terms of mean value and standard deviation is presented in Table 1. More information regarding the data set and the data cleaning process can be found in Ratnam et al. [27].

2.5. Simulation setup

The simulations were divided into scenarios, where in each scenario electricity consumption for one year for one customer was forecasted one step ahead, one half hour. This was based on electricity consumption training data from the two remaining years of the same customer. That way the models were used to produce forecasts for each house (69, 74, 157, 211 and 273) for each year (1, 2, 3), which in total generated 15 simulated scenarios. The training data sets contained 35'072 data points each, and the test data sets contained 17'536 data points each. For the scenarios, in order for 10 step conditioning to work for QR and PeEn, the first data point which was forecasted in each test data set was data point 11 out of 17'536.

MCM only uses one data point for producing each forecast, while QR and PeEn can use any lag of data points as conditioning for each forecast. In order to quantify the effect of this, a sensitivity analysis in terms of nCRPS and rMAE from each model was employed for number of lag data points (number of time-steps electricity use) included in the forecasting. In those cases, the first data points that were forecasted were lag h plus one. Note that neither MCM nor QR models are re-trained with the test data.

3. Results

Example electricity use data, model median and prediction interval forecasts from MCM, QR and PeEn is shown in Fig. 2. This gives an example of how the models react to ramps of electricity use as well as periods of low load. It is possible to note a similarity between MCM and QR, but with qualitative differences, where one particular example is the difference in prediction interval from the different models associated with the large peak at hour 13. Here both MCM and QR have a short high peak in prediction interval width, while PeEn has a lower but wider peak, the latter of which is because the high peak is included in the PeEn several steps after it has occurred.

Reliability diagrams associated with MCM, QR and PeEn simulations for the scenario-defined sets of customers and years are presented in Fig. 3. Generally, all models approximately align with the diagonal, with some variability that is specific to each data set, except for PeEn, which generally underestimates the variance of the data. In particular note MCM and QR variability for house 157 year 3 (C3) and house 211 year 3 (D3).

As a formal measure of the reliability, the rMAE was computed for the scenarios in Fig. 3 and the results are presented in Fig. 4. The results vary considerably from house to house and year to year with the general conclusion that PeEn has highest rMAE in all scenarios, which implies that MCM and QR are more reliable. Among the scenarios MCM had lowest rMAE for 6 scenarios, and QR had lowest rMAE for 9 scenarios.

In Fig. 5 PINAW plots for MCM, QR and PeEn simulations are presented. Do note that since PeEn forecasts are comprised of only 10 data points, it has fewer quantiles represented in the plot. In general there is similar sharpness for all models, with some variability between customers and years. Overall, MCM and QR are remarkably similar, with generally a slight favor for QR in terms of sharpness.

In Fig. 6 nCRPS results for MCM, QR and PeEn are presented. MCM and QR have superior mean nCRPS compared with PeEn for all tested cases, and in detail QR had lowest mean nCRPS in 11 scenarios, while MCM had lowest mean nCRPS in 4 scenarios. In terms of mean value the similarity is high with mean nCRPS of 2.50 for MCM and 2.45 for QR, respectively, a mere 2 percent lower nCRPS for QR. These results all regard the 10 steps lag QR, which is a special case chosen for this study mainly to investigate where MCM and QR perform similar.

In order to compare results with different included number of data points (lag), a sensitivity analysis was performed, and results in terms of nCRPS and rMAE are presented in Fig. 7. Since MCM is only dependent on one time-step lag, it is constant in these simulations. In terms of nCRPS, MCM and QR converge for lower lag, while PeEn has a more variable pattern, and is inferior, for all lags, to MCM and QR. As regards rMAE, MCM and QR converge for lower lags in a similar fashion as

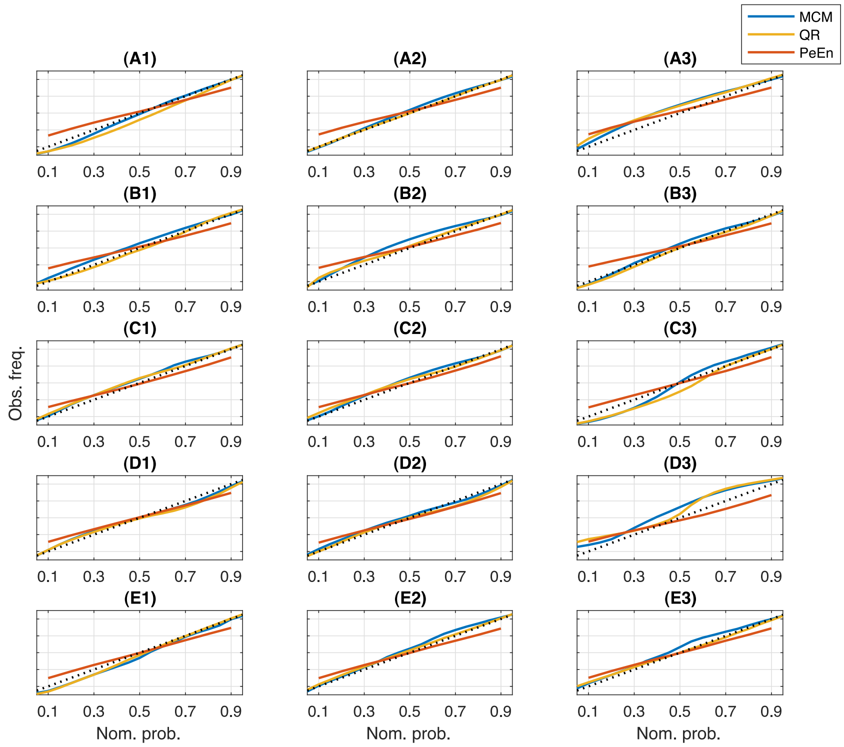

Fig. 3. Reliability diagrams for MCM, QR and PeEn for houses 69, 74, 157, 211 and 273, respectively in plot rows top to bottom, and for years 1, 2 and 3, respectively, in columns left to right (1–3). The dotted diagonal line is optimal reliability.

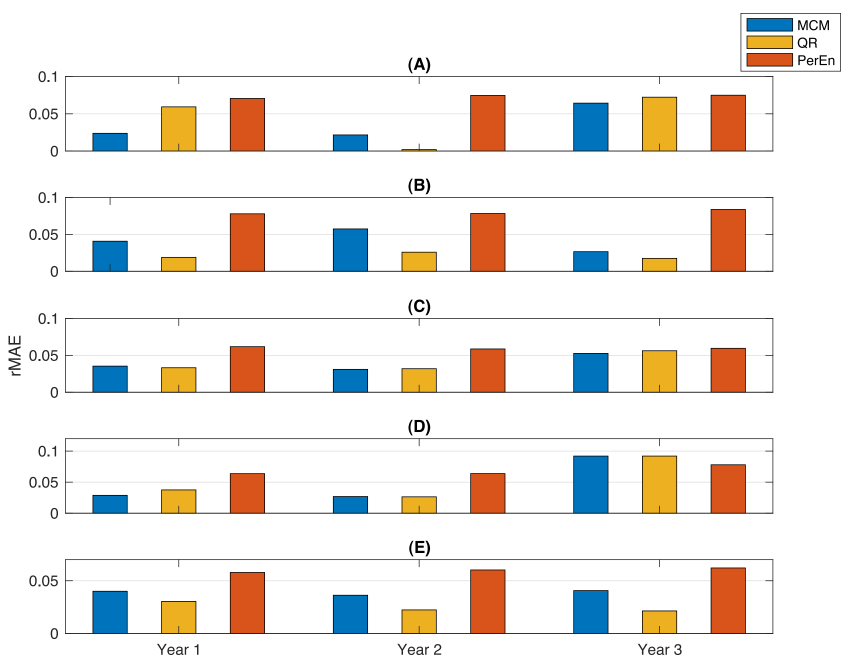

Fig. 4. rMAE plots for MCM, QR and PeEn for houses 69, 74, 157, 211 and 273, respectively in plot rows top to bottom (A)–(E), and for years 1, 2 and 3, respectively, in columns left to right. The dotted central line is optimal reliability.

with nCRPS, but in this case PeEn descends and approaches zero as the number of lags increase. This is because the distribution of PeEn converges to the distribution of the testing data set as the lag increases. For lower than approximately 24 lags, MCM and QR are superior in rMAE.

In order to analyze the sensitivity of the choice of N for the MCM model, estimates of nCRPS for $N = 20$ to $N = 200$ was performed, and the results are presented in Fig. 7. The results show a minimum at $N = 60$, and a maximum at $N = 20$, but a variance of $1.3 \cdot 10^{-6}$ indicates a comparative insensitivity to the choice of N.

4. Conclusion

The MCM model was used in very short term load forecasting of electricity consumption and was compared with Quantile Regression,

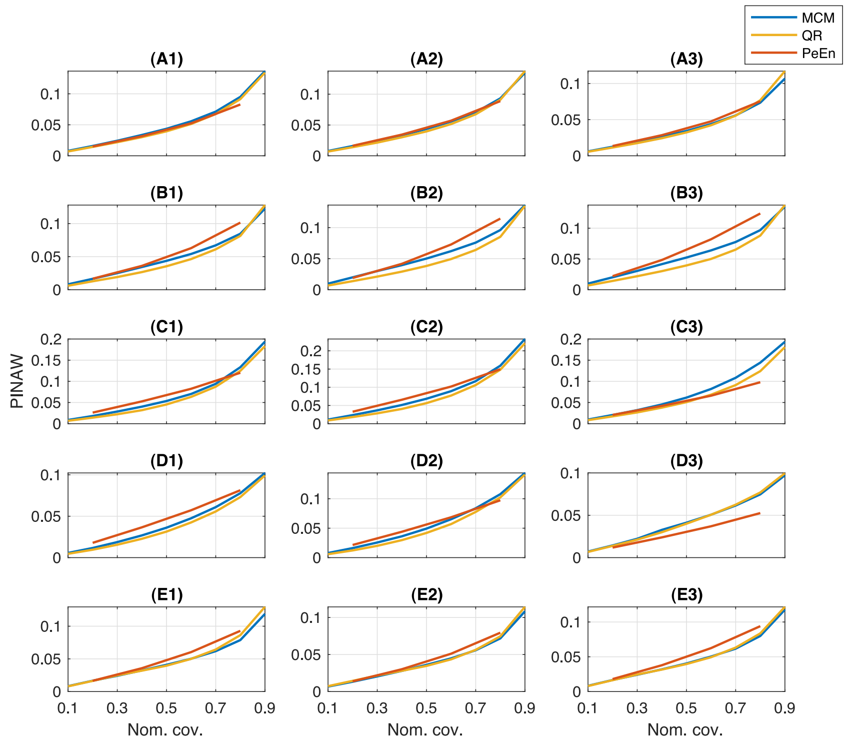

Fig. 5. PINAW plots for MCM, QR and PeEn for houses 69, 74, 157, 211 and 273, respectively in plot rows top to bottom, and for years 1, 2 and 3, respectively, in columns left to right (1–3).

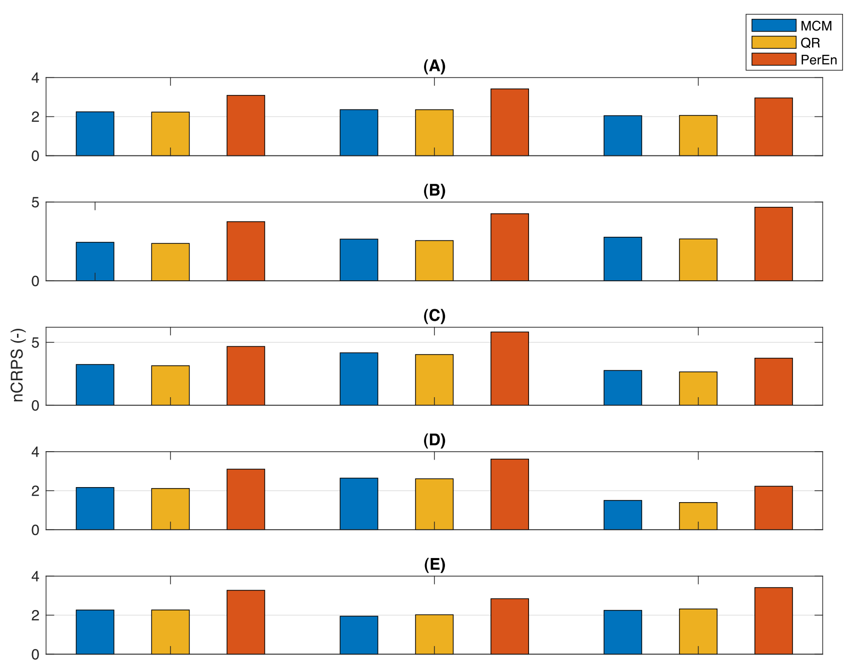

Fig. 6. nCRPS plots for MCM, QR and PeEn for houses 69, 74, 157, 211 and 273, respectively in plot rows top to bottom (A)–(E), and for years 1, 2 and 3, respectively, in columns left to right.

as an advanced model comparison, and also compared with the Persistence Ensemble as standard benchmark reference. Quantile Regression and the Persistence Ensemble used 10 steps conditioning, and for that setting the performance between MCM and Quantile Regression is similar in terms of reliability, prediction interval normalized average width and normalized continuous ranked probability score, while both MCM and Quantile Regression are generally superior when compared with the Persistence Ensemble. A sensitivity analysis of increasing the included lag of Quantile Regression and the Persistence Ensemble simulations was performed, which showed a monotonic improvement for Quantile Regression in normalized continuous ranked probability score, while the Persistence Ensemble had less of a discernible trend. In terms of reliability mean absolute error both Quantile Regression and the Persistence Ensemble improved with number of points lag included, with the Persistence Ensemble having highest reliability above approximately 24 time-steps lag.

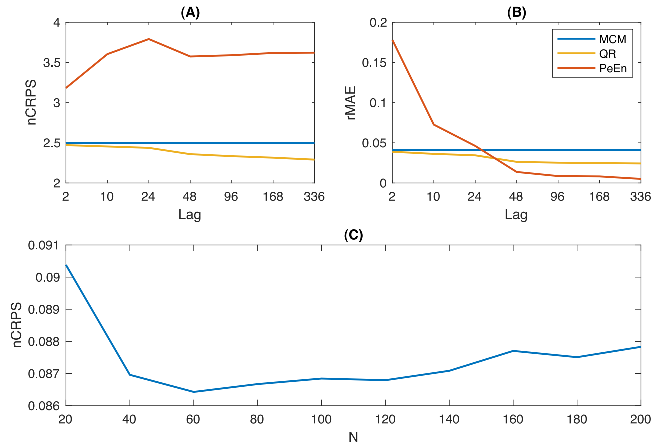

Fig. 7. Mean nCRPS (over all houses and years) in subplot (A) and rMAE in subplot (B) for a set of included time-steps back (lag) in the forecast. In subplot (C) the mean nCRPS for house 69 year 1 was computed for a series of N for the MCM model.

The MCM model is conceptually sparse, insensitive to parameter settings, computationally inexpensive and performs on par with QR for the 10 time-step lag scenario of one-step ahead residential electricity use forecasting. Based on this it is suggested as a benchmark for probabilistic forecasting of electricity consumption. As such, it could be compared with future methods in terms of both output accuracy in the suggested performance metrics and also in terms of computational complexity.

The use of the MCM model for forecasting other types of electricity use time-series would be interesting, as well as for utilizing different forecasting horizons. Also, the use of MCM on step-changes of electricity use could be fruitful to investigate, since such time-series are stationary, which is particularly fitting for the MCM model.

CRediT authorship contribution statement

Joakim Munkhammar: Co-developed the model, Executed most simulations, Wrote the majority of the paper. **Dennis van der Meer:** Co-developed the model, Executed QR simulations and metric, Assisted in writing of the paper. **Joakim Widén:** Co-developed the model, Assisted in writing of the paper.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This project was funded by the projects "Probabilistic Forecasting for Battery Management" and "Development and evaluation of forecasting models for solar power and electricity use over space and time", both funded by the Swedish Energy Agency. This work was also supported by the Swedish strategic research programme StandUp for Energy.

References

[1] Hong T, Fan S. Probabilistic electric load forecasting: A tutorial review. Int J Forecast 2016;32:914–38. http://dx.doi.org/10.1016/j.ijforecast.2015.11.011.

[2] Hu Y, Ye C, Ding Y, Xu C. Short-term load forecasting utilizing wavelet transform and time series considering accuracy feedback. Int Trans Electr Energy Syst 2020;30:e12455.

[3] Somu N, Raman GMR, Ramamritham K. A hybrid model for building energy consumption forecasting using long short term memory networks. Appl Energy 2020;261:114131.

[4] van der Meer D, Widén J, Munkhammar J. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. Renew Sustain Energy Rev 2018;81:1484–512. http://dx.doi.org/10.1016/j.rser.2017.05.212.

[5] Shepero M, van der Meer D, Munkhammar J, Widén J. Residential probabilistic load forecasting: A method using Gaussian process designed for electric load data. Appl Energy 2018;218:159–72.

[6] Bracale A, Carpinelli G, Falco PD. A Bayesian-based approach for the short-term forecasting of electric loads in smart grids. Part II: numerical applications. In: International symposium on power electronics, electrical drivers, and automated motion. 2016, p. 1760–71. http://dx.doi.org/10.1109/SPEEDAM.2016.752602.

[7] Guan C, Luh PB, Laurent DM, Chi Z. Hybrid Kalman filters for very short-term load forecasting and prediction interval estimation. IEEE Trans Power Syst 2013;28:3806–17.

[8] Quan H, Srinivasan D, Kosravi A. Short-term load and wind forecasting using neural network-based prediction intervals. IEEE Trans Neural Netw Learn Syst 2014;25:303–15. http://dx.doi.org/10.1109/TNNLS.2013.2276053.

[9] Quan H, Srinivasan D, Kosravi A. Uncertainty handling using neural network-based prediction intervals for electrical load forecasting. Energy 2014;73:916–25. http://dx.doi.org/10.1016/j.energy.2014.06.104.

[10] Taieb SB, Huser R, Hyndman RJ, Genton MG. Forecasting uncertainty in electricity Smart Meter Data by Boosting additive quantile regression. IEEE Trans Smart Grid 2016;5:2448–55. http://dx.doi.org/10.1109/TSG.2016.2527820.

[11] Zhang S, Wang Y, Zhang Y, Wang D, Zhang N. Load probability density forecasting by transforming and combining quantile forecasts. Appl Energy 2020;277:115600.

[12] van der Meer D, Shepero M, Svensson A, Widén J, Munkhammar J. Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using gaussian processes. Appl Energy 2018;213:195–207. http://dx.doi.org/10.1016/j.apenergy.2017.12.104.

[13] Wang Z, Wang Y, Zeng R, Srinivasan RS, Ahrentzen S. Random Forest based hourly building energy prediction. Energy Build 2018;171:11–25.

[14] Zhao W, Wang J, Lu H. Combining forecasts of electricity consumption in China with time-varying weights updated by a high-order Markov chain model. Omega 2014;45:80–91.

[15] Carpinone A, Giorgio M, Langella R, Testa A. Markov chain modeling for very-short-term wind power forecasting. Electr Power Syst Res 2015;122:152–8. http://dx.doi.org/10.1016/j.epsr.2014.12.025.

[16] Munkhammar J, Widén J, van der Meer D. Probabilistic forecasting of high-resolution clear-sky index time-series using a Markov-chain mixture distribution model. Sol Energy 2019;184:688–95. http://dx.doi.org/10.1016/j.solener.2019.04.014.

[17] Yang D, van der Meer D, Munkhammar J. Probabilistic solar forecasting benchmarks on a standardized dataset at Folsom, California. Sol Energy 2020;20:628–39.

[18] Munkhammar J, Widén J. An N-state Markov-chain mixture distribution model of the clear-sky index. Sol Energy 2018;173:487–95. http://dx.doi.org/10.1016/j.solener.2018.07.056.

[19] Murphy KP. Machine learning: A probabilistic perspective. Cambridge, USA: MIT press; 2012.

[20] Shepero M, Munkhammar J. R and python code for the Markov-chain mixture distribution (MCM) model. In: SheperoMah/MCM-distribution-forecasting. GitHub; 2019.

[21] Alessandrini S, Monache LD, Sperati S, Cervone G. An analog ensemble for short-term probabilistic solar power forecast. Appl Energy 2015;157:95–110. http://dx.doi.org/10.1016/j.apenergy.2015.08.011.

[22] Koenker R. quantreg: Quantile Regression in R. version 5.35, https://CRAN.R-project.org/package=quantreg.

[23] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2018, https://www.R-project.org/.

[24] Koenker R, Bassett G. Regression quantiles. Econometrica 1978;46:33–50.

[25] Gneiting T, Stanberry LI, Grimit EP, Held LL, Johnson NA. Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble predictions of surface winds. Test 2008;17:2011–235.

[26] Solar home electricity data - Ausgrid. 2020, https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Solar-home-electricity-data (2020-02-17).

[27] Ratnam EL, Weller SR, Kellett CM, Murray AT. Residential load and rooftop PV generation: an Australian distribution network dataset. Int J Sustain Energy 2017;36:787–806.