Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 70 (2023) 372–379

Transportation Research Procedia

www.elsevier.com/locate/procedia

8th International Electric Vehicle Conference (EVC 2023)

Modelling the Intensity of Electric Vehicle Arrivals at Charging Points

Yvenn Amara-Ouali[a,b,] *, Yannig Goude[a,b], Jean-Michel Poggi[b,c]

[a]*EDF R&D (OSIRIS), Palaiseau 91120, France*
[b]*University Paris-Saclay (LMO), Orsay 91400, France*
[c]*Universite´ de Paris, IUT STID, Paris 75016, France*

Abstract

With Electric Vehicles' (EV) market adoption surging in recent years, the smart grid paradigm requires accurate forecasts of EV arrivals at charging points. One efficient way to model these arrivals is to use Point Processes. This study introduces an additive model using both spline and wavelet effects for fitting the intensity of a non-homogeneous Poisson process applied to EV arrivals at charging points. The key contribution of this work is a novel estimation procedure inspired from backfitting which is illustrated by a case study on real-world EV arrivals at charging points. The results obtained show that this approach can help better capturing EV arrival peaks.

Keywords: Forecasting; Smart Grids; Peak Arrivals; Non-homogeneous Poisson Process; Generalised Additive Models; Splines; Wavelets

1. Introduction

Integrating Electric Vehicles (EVs) into the large-scale distribution network represent a major challenge. Poor management of this new market can have a negative impact on the load curve such as increased demand peaks. In addition, this new constraint on the network can lead to the overloading of certain system components linked to voltage and frequency imbalances Adderly et al. (2018). By moving to a smart charging paradigm, EVs can be used as a solution rather than a problem a to better balance the grid Aslam et al. (2020). This requires minimising the uncertainty of EV arrivals in order to better plan charging schedules. Furthermore, modelling EV arrivals can contribute to better infrastructure sizing Gjelaj et al. (2020). Finally, knowing ahead the arrival patterns of EVs raises the prospect of storing renewable energy in EVs as a viable strategy Ghotge et al. (2020). Modelling arrivals is a typical task for non-

* Corresponding author. Yvenn Amara-Ouali
E-mail address: yvenn.amara-ouali@universite-paris-saclay.fr

2352-1465
10.1016/j.trpro.2023.11.042

homogeneous Poisson Processes (NHPPs). NHPPs have been used to model a diverse range of phenomenons in the literature (e.g., seismology Aktas et al. 2009, imaging Streit 2010, extreme weather events Ngailo et al. 2016). NHPPs have also been applied to model EV arrivals in recent work (e.g., Zhang and Grijalva 2015 and Lahariya et al. 2020). The objective of the paper is to develop an NHPP model to predict EV arrivals at charging points, focusing on enhancing peak arrival forecasts. It presents an additive model using spline and wavelet effects to capture NHPP intensity for EV arrivals. The paper is organized as follows: Section 2 reviews related work, Section 3 introduces model specifics and the estimation procedure. Lastly, Section 4 presents a case study on real world data.

2. Related Work

NHPP intensity estimation. Historically, point process estimation has been thoroughly detailed in Daley and Vere-Jones (2003) including NHPP. The likelihood of a NHPP process with n observed arrival times $\{\widehat{t_i}\}_{i\in\{1\cdots n\}}$ can generally be written as follows:

$$L = \left(\prod_{i=1}^{n} \lambda(\widehat{t_i})\right) \exp\left(-\int_0^T \lambda(s)\, ds\right) \tag{1}$$

Estimators of the intensity function can thus be obtained by maximising L. In Brillinger (1997), a semi-parametric estimator of the conditional intensity of temporal point processes was introduced with a Tukey shrinkage procedure. The theoretical properties of wavelet coefficient estimators were later studied for the first-order intensity of multi-dimensional NHPPs (de Miranda and Morettin, 2011) with results applicable to Haar wavelets in practice. In Bigot et al. (2013), Meyer wavelets are preferred to Haar wavelets for estimating the intensity function for *n* independent realisations of a NHPP. A wider class of wavelets is studied in Reynaud-Bouret and Rivoirard (2010) with biorthogonal wavelets. More recently, Youngman and Economou (2017) have proposed fitting procedure for the intensity of NHPP with splines basis instead of wavelets.

Additive models. A general procedure for fitting semi-parametric regression models with additive functional effects was proposed in Hastie and Tibshirani (1986) with backfitting. In parallel, a General Cross Validation (GCV) criterion first defined in Craven and Wahba (1979) was used for estimating a single smoothing spline and extended in Gu and Wahba (1991) to multiple smoothing parameters. This result was used to propose a PIRLS fitting procedure for generalised additive models (GAMs) in Wood (2000). While splines have been the basis of choice for additive semi- parametric models, wavelet basis have gained more and more interest over time. Wand and Ormerod (2011) proposes to integrate wavelets into semi-parametric regression in a similar way as it is done for splines in generalised additive models (Wood, 2017). While splines estimated with PIRLS contain a ridge-like penalty, wavelets are given a lasso- type penalty. The close relationship of the lasso penalty with soft thresholding established in Antoniadis and Fan (2001) is what makes it our choice in our work.

Gaps. The literature review indicates a scarcity of papers combining NHPP intensity estimation with additive models, with most studies focusing on either splines or wavelets. A hybrid model proposed in Liu et al. (2017) did combine backfitted splines and kernel methods. However, very rarely splines and wavelets are combined in the same model specification. The only occurrence of this kind of work the authors have found is in Amato and Antoniadis (2017) where a hybrid approach for Gaussian regression is proposed with an application of this model to electricity demand in Amato et al. (2021). Therefore, the goal of the work presented in this paper is to extend these papers and study more precisely an additive model of the first-order intensity function of NHPPs combining both splines and wavelet basis as presented in equation (2).

3. Problem Formulation

Model. This study considers an additive model of the intensity function of a NHPP of the following form:

$$\log \lambda(t) = \beta_0 + \sum_{l=1}^{L_p} \beta_l\, x_l^p(t) + \sum_{l=1}^{L_s} s_l\left(x_l^s(t)\right) + \sum_{l=1}^{L_w} w_l\left(x_l^w(t)\right) \tag{2}$$

with, β_0 the intercept, $\beta_.$ the coefficients of the linear component, $s_.$ and $w_.$ respectively spline and wavelet basis expansions. In addition, $\boldsymbol{x(t)} = (\boldsymbol{x}^p(t), \boldsymbol{x}^s(t), \boldsymbol{x}^w(t))$ the vector of covariates evaluated at time $t\epsilon[0,T](T\epsilon\mathbb{R}^+$ being the final time at which the NHPP is observed). $x^p(t) = \left\{x_l^p(t)\right\}_{\{1,\cdots,L\}}$, $x^s(t) = \{x_l^s(t)\}_{\{1,\cdots,L\}}$ and $x^w(t) = \{x_l^w(t)\}_{\{1,\cdots,L\}}$ are the vectors of covariates respectively used for the linear, spline and wavelet effects. The idea behind this model is to decompose the signal into linear, smooth and irregular components. Also, the additive structure was chosen to have an interpretable model with the contribution of each component made clear for analysis.

Likelihood estimation. In practice, the likelihood presented in equation (1) is intractable unless approximations are made. In particular, the intensity function can be considered piecewise constant provided that a small enough timestep is taken relative to the phenomenon modelled. With that approximation, the integral term becomes a discrete sum over the number of timesteps. Therefore, equation (1) becomes:

$$L = \left(\prod_{i=1}^{n} \lambda(\hat{t_i})\right) \exp\left(-\sum_{t\epsilon[0,T]\cap\mathbb{N}} \lambda(t)\right) \tag{3}$$

Here, the intensity function is expressed in the unit of the timestep chosen so that there is no need to multiply each term of the sum by the timestep. This approximation is widely used as it is enough to make the likelihood tractable under reasonable assumptions on the intensity function. The intensity function of a Poisson process is often confused with the rate (also denoted by λ in most cases), even though they are conceptually different. These two quantities coincide exactly when the intensity function is assumed piecewise constant. Adopting a Poisson regression approach leads to another version of the likelihood presented in equation (3) which depends on time through different temporal signals $\boldsymbol{x}(t)$. In the context of our problem, a sample $\{(Y_i, \boldsymbol{x}_i), i \in \{1 \ldots n\}\}$ of size $n \epsilon \mathbb{N}$ is observed. Therefore the likelihood of the equivalent regression formulation of the NHPP likelihood from equation (3) can be rewritten as follows: $L(\theta) = \prod_{i=1}^{n} \frac{\exp(-\lambda_i(\theta))\lambda_i(\theta)^{y_i}}{y_i!}$ with θ the vector containing the parameters for all effects. Finally, the log-likelihood is as follows: $l(\theta) = \sum_{i=1}^{n} Y_i\, log(\lambda_i(\theta)) - \sum_{i=1}^{n} \lambda_i\,(\theta) - \sum_{i=1}^{n} log\,(Y_i!)$ with the last term of this log-likelihood independent of λ which makes it irrelevant for maximising the log-likelihood.

Background. First, the Generalized Linear Model (GLM) extends traditional linear regression by allowing a wider range of statistical distributions for the response variable and incorporating a non-linear relationship between the response and covariates using a link function. In our case, the GLM corresponds to the parametric part of the model, estimated using the iteratively (reweighted) least squares (IRLS) algorithm: $\log\lambda(t) = \beta_0 + \sum_{l=1}^{Lp} \beta_l x_l^p(t)$. IRLS updates the parameter estimates iteratively, taking into account the residuals of each observation. Next, the Generalized Additive Model (GAM) represents the non-parametric part of the model. It is formulated as a sum of smooth functions, with each function modeling a specific covariate $\log\lambda(t) = \sum_{l=1}^{Ls} s_l(x_l^s(t))$. The backfitting procedure is commonly used to fit GAMs (Hastie and Tibshirani, 1986), and it has been proven to converge to an optimal solution (Ansley and Kohn, 1994). A more recent approach, called penalized IRLS (PIRLS), introduced a penalty term on the second-order derivatives of the smooth functions to ensure smoother estimates while maintaining computational efficiency (Wood, 2000). Lastly, the Penalized Wavelet Additive Model is similar to GAM but uses wavelet basis functions instead of splines $\log\lambda(t) = \sum_{l=1}^{Lw} w_l(x_l^w(t))$. This model is suitable for capturing the piecewise constant nature of the function being modeled. To optimize the LASSO-penalized likelihood, a coordinate gradient descent (CGD) approach is chosen (Hastie et al., 2022). Cross-validation is used to select the penalty coefficient (gamma), and the soft-thresholding operator is employed to perform the coordinate-wise parameter updates.

Combining wavelets and splines basis. Wavelets are versatile tools for signal processing and representing functions of varying regularity (Mallat, 1999). They offer the ability to capture local behavior at different time scales. Orthonormal wavelet bases provide a comprehensive analysis of irregularities and allow for efficient decomposition algorithms with linear complexity. While splines are powerful for functional estimation, wavelets complement them by offering better temporal localization, particularly in representing peaks and discontinuities (Vidakovic, 2009). Wavelets excel in representing such functions with a sparse set of non-zero coefficients. When combining splines and wavelets in a model, it is recommended to use a sequential estimation approach. Splines are employed first to estimate low frequencies and regular parts, followed by wavelets to focus on breaks or peaks. The Haar wavelet is especially

suitable for this purpose due to its optimal localization properties. Increasing the number of zero moments (N) can provide additional regularity by using the Daubechies wavelet family, denoted as db_{2N}.

Proposed approach. Using the stepping stones introduced above, an algorithm can be proposed to successfully fit model (2). It is referred to as BAC, being a version of backfitting applied to this model. It consists in fitting sequentially the different effects of the model. Firstly, the linear part is fitted. Then, for each variable considered in the non-parametric part of the model, the splines components are first fitted and then the wavelet effects. The idea behind this algorithm is to move from the lowest frequency (linear part and splines with not too many degrees of freedom) to the highest frequency of the signal (wavelet basis of relative high-order). The BAC algorithm does not involve an a priori on the order in which the different effects should be fitted. In practice, effects are fitted in a random order. Each time an effect is fitted, the rest of the model fitted up until this iteration is subtracted from the target response. So only the residuals of the current model iteration are fitted at each step. Like backfitting, this procedure is repeated multiple times until convergence. Algorithm 1 formally presents the implementation of this approach with the notations introduced in the introduction. In addition, $\eta_{-l}(\theta_{-l})$ is the additive model without the l-th effect which can also be extended to components. For instance, $\eta_{-l}^{-s}(\theta_{-l}^{-s}) = \beta_0 + \sum_{l'=1}^{Lp} \beta_{l'}\, x_{l'}^{p}(t) + \sum_{l'=1;l'\neq l}^{Ls} s_{l'}\left(x_{l'}^{s}(t)\right) + \sum_{l'=1}^{Lw} w_{l'}\left(x_{l'}^{w}(t)\right)$, which is simply the same model as in equation (2) without the spline component. Similarly, $\eta_l(\theta_l)$ is the additive model only with the l-th effect composed of its spline $\eta_l^s(\theta_l^s)$ and wavelet $\eta_w(\theta_l^w)$ counterparts.

4. Problem Formulation

Data. The dataset in the scope of this application gathers charging session information in the United Kingdom (UK) during 2017 (Amara-Ouali et al., 2021). It concerns domestic chargers ranging from 3kW to 22kW. However, it is expected that the great majority of EV chargers in this dataset are 3kW or 7kW chargers. This data was collected by the UK department of transport. One key finding on arrival times (or plug-in times) of EVs was that domestic charging events were more frequent with different patterns on weekdays than on weekends. That is why this study focuses solely on weekdays. The circumstances of EV uptake in 2017 in the UK was strong. The total number of plug-in cars on UK roads passed 130,000 that year. The best BEV seller was the Nissan LEAF with more than 13,000 registrations. Across the whole country, the largest sales were made in London and Eastern England with Scotland and South West garnering the fastest growth (UK Department for Transport, 2018). While public EV charging infrastructure is increasing at a fast pace, domestic charging remains the first choice for a majority of EV users. In addition to this dataset, temperatures in the UK from 8 of the top 10 cities in terms of population from the Iowa Environment Mesonet website (Salmon, 2016) were gathered. That is, London (8.9 million), Birmingham (1.15 million), Glasgow (612 thousand), Liverpool (579 thousand), Bristol (572 thousand), Manchester (554 thousand), Leeds (503 thousand), Edinburgh (508 thousand). The temperature for each city is not particularly recorded at the same time nor at a regular timestep. Therefore, all these temperature were interpolated with cubic splines. In order to have a more compact model, a weighted average version of the temperature can be calculated as follows: $temp(t) = \frac{1}{\sum_{a=1}^{8} pop_s} \sum_{s=1}^{8} po\, p_s T_{s,t}$ where $T_{s,t}$ is the temperature recorded at time t by station s and temp(t) is the weighted mean temperature which will be used in the modelling experiments.

Experimental setting. The experimental protocol chosen to test the methods is close to the operational setting. In fact, it is a rolling forecast origin procedure, where the model is trained on all the data available up to a certain date in order to forecast the following week. The first training set runs from 1 January to 29 September 2017. As weekends are not included, it comprises 10 test weeks from 4 September to 8 December 2017. The algorithm considered in our experiments is `BAC` as defined in section 3. The algorithm was tested with three different variations. All of them include a linear part, which is simply the indicator of the day of the week. The first variation only takes into account the spline components for the hour of the day, the weighted temperatures and the time as an index (BAC_s). The second variation is the same with only the wavelet components (BAC_w). Finally, the third variation includes both splines and wavelet components only, with the time of day modelled by a wavelet component and the other effects assumed to be smooth enough to be captured by the splines component only (BAC_sw). Our assumption is that the linear and spline components will capture most of the variations in the intensity function, while the wavelets would improve performance during peak and/or irregular times. The dataset and code used for this experiment are available at Amara-Ouali (2023).

Algorithm 1: *BAC*

Target: $\boldsymbol{Y}$;
Number of non-linear effects to be fitted separately: L;
Covariates: $\boldsymbol{X} = (\boldsymbol{X}^p, \boldsymbol{X}^s, \boldsymbol{X}^w)$;
Parameters: $\boldsymbol{\theta} = (\boldsymbol{\theta}^p, \boldsymbol{\theta}^s, \boldsymbol{\theta}^w)$;
Objective: $\underset{\boldsymbol{\theta}}{\operatorname{argmax}}\, l(\boldsymbol{\theta})$;
$\epsilon \leftarrow 10^{-3}$;
$\boldsymbol{\theta}^p = \underset{\boldsymbol{\theta}^p}{\operatorname{argmax}}\, l(\boldsymbol{\theta}, \boldsymbol{Y}, \boldsymbol{X})$ (IRLS) ; /* Linear component */
$\boldsymbol{\theta}^{s,(0)} =\leftarrow \boldsymbol{0}$ and $\boldsymbol{\theta}^{w,(0)} \leftarrow \boldsymbol{0}$;
$t \leftarrow 1$;
do
  index = shuffle ($\{1 \dots L\}$) ;
  for *l in index* **do**
    $\tilde{\boldsymbol{Y}} = \boldsymbol{Y} - \exp\left(\eta^s_{-l}(\boldsymbol{\theta}_{-l})\right)$;
    $\boldsymbol{\theta}^{s,(t)}_l = \underset{\boldsymbol{\theta}^s_l}{\operatorname{argmax}}\, l(\boldsymbol{\theta}, \tilde{\boldsymbol{Y}}, \boldsymbol{X})$, (PIRLS) ; /* Spline component */
    $\tilde{\boldsymbol{Y}} = \boldsymbol{Y} - \exp\left(\eta^w_{-l}(\boldsymbol{\theta}_{-l})\right)$;
    $\boldsymbol{\theta}^{w,(t)}_l = \underset{\boldsymbol{\theta}^w_l}{\operatorname{argmax}}\, l(\boldsymbol{\theta}, \tilde{\boldsymbol{Y}}, \boldsymbol{X})$, (CGD) ; /* Wavelet component */
    $F^{(t)}_l = \eta^s_l(\boldsymbol{\theta}^{s,(t)}_l) + \eta^w_l(\boldsymbol{\theta}^w_l)$
  end
  $t \leftarrow t + 1$;
while $\forall l \in \{1, \dots, L\}$, $\|F^{(t)}_l - F^{(t-1)}_l\|_2 > \epsilon \|F^{(t-1)}_l\|_2$ *and* $t < 100$;
Result: $\boldsymbol{\theta}$

Metrics. To assess the efficiency of the various approaches, different metrics were retained: mean absolute error (MAE), root-mean squared error (RMSE), peak RMSE, deviance and dynamic time warping (DTW). The MAE and RMSE are the L^1 and L^2 norms divided by the number of observations. The peak RMSE is the RMSE restricted to daily peaks. So the value of the peak observed is compared to the prediction at the same time of day. Deviance is used as a goodness of fit test. In the context of Poisson regression, the deviance formula can be written as follows $D = 2\sum_{i=1}^{n}\left(Y_i \log\left(\frac{Y_i}{\hat{Y}_i}\right) - (Y_i - \hat{Y}_i)\right)$ with $\hat{Y}_i$ being the model prediction. This formula can be derived from the likelihood ratio test comparing the proposed model and the saturated model. The latter predicts exactly the observed value (like an oracle model). Finally, unlike the peak RMSE, DTW is used to measure the similarity between two time series that may be slightly out of sync. It calculates the distance between two time series by aligning them in a way that minimises the differences between corresponding points (Müller, 2007).

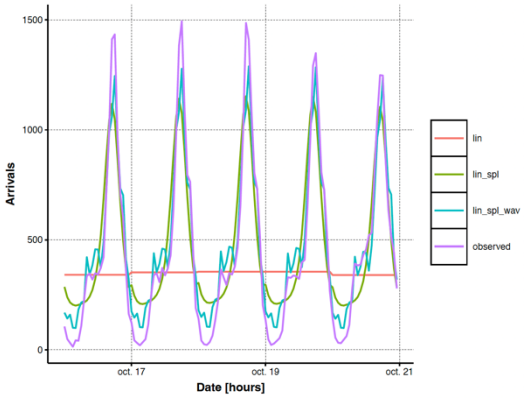

Fig. 1: Decomposition of the predicted signal approach on week 3 of test data (lin - linear component only; lin spl - linear and spline components; lin spl wav - all components).

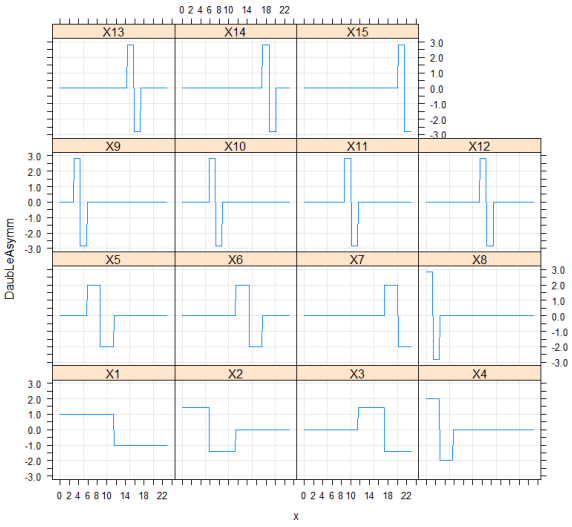

Fig. 2: Haar basis **before** fit on hour of day variable

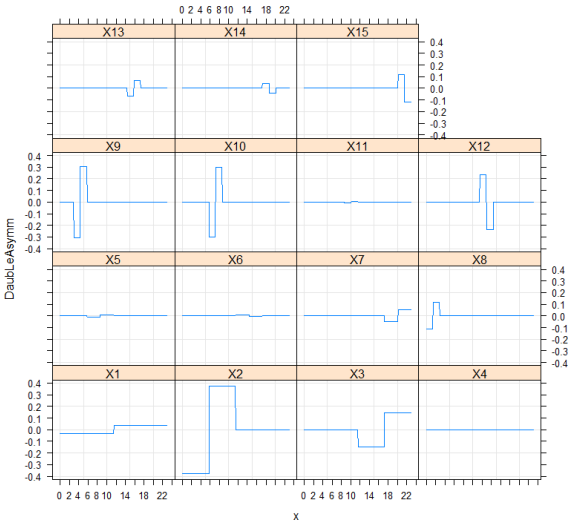

Fig. 3: Haar basis **after** fit on hour of day variable

Results. Fig. 1 shows a fitted BAC_sw approach for a random week of the test set. From this fit, it is interesting to note that the benefit of the wavelet component is clearly observed on all peak estimates. In addition, the benefit of the wavelets can also be seen on the ascending and descending parts of the curve where the arrivals are most irregular. This confirms our a priori assumption that wavelets could help to better capture peaks and irregularities in the data. In figure 2 the wavelet basis expansion of the time of day effect can be observed. This effect does not have too many levels and is easier to understand. The Haar basis is used here because of the assumption of piecewise constant intensity function and also because it is easy to implement. However, the proposed procedure is not restricted to any particular type of wavelet. It is interesting to compare what happens to this basis after fitting. In particular, the soft- thresholding procedure obtained by the LASSO fit is illustrated in Fig. 3. It shows the same wavelet functions as in Fig. 2, multiplied by the coefficients estimated in the BACswapproach. The sum of all these individual functions can be represented to show the the time (hour) of day - *tod* effect for the BACswapproach as well as the splines. The same thing can be done for the *toy* (time of year) and *temp* (temperature) effects used in the model. Fig. 4 achieves exactly this. The *toy* (time of year) and *temp* (temperature) evolve smoothly while *tod* requires a wavelet component to capture irregularities in the ascendant part of the intensity function (between 8am and 10am).

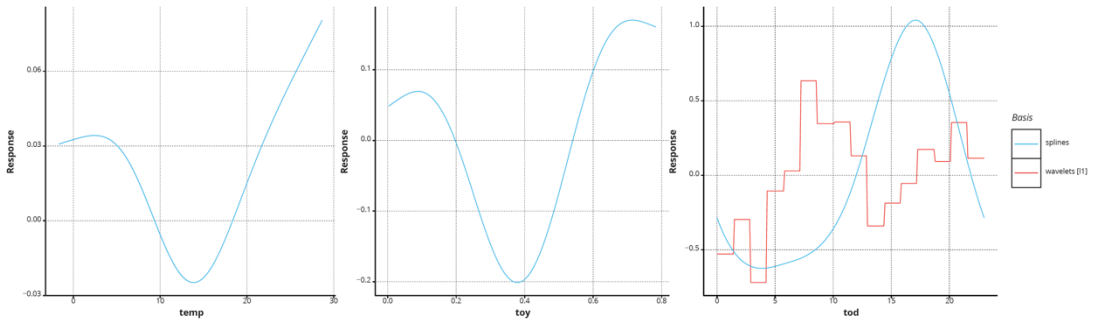

Fig. 4: All effects fitted with their splines and wavelet (only for tod - time of day [hours]) components

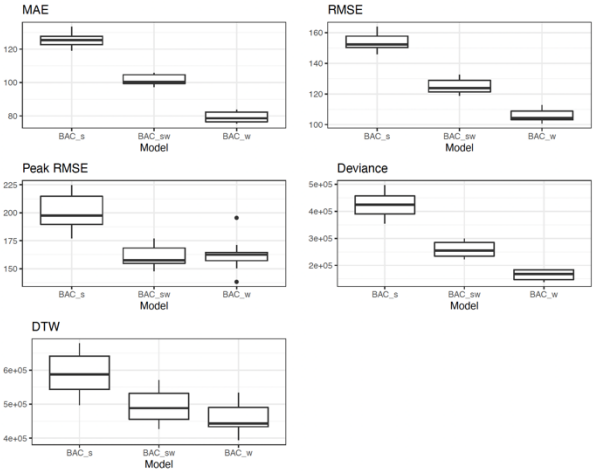

Fig. 5: Boxplots of performance metrics on the rolling train set

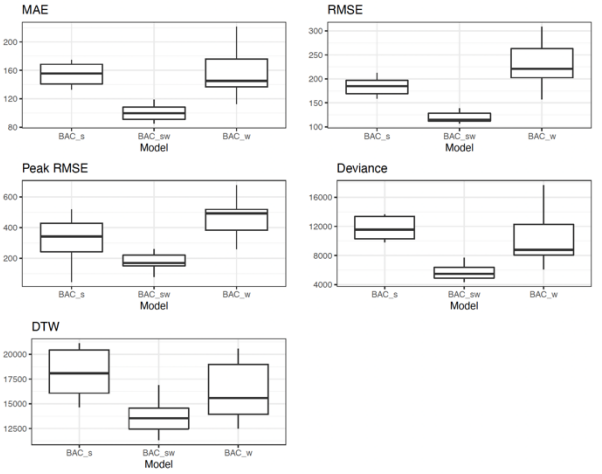

Fig. 6: Boxplots of performance metrics on the 10 weeks of test data

In terms of performance on the training sets, Fig. 5 shows the boxplots of the 10 rolling training sets for the 5 performance metrics retained in this analysis. The 5 metrics seem to agree that the BAC_w performs best on fitted data. However, Fig. 6 shows that BAC w actually performs worse when predicting arrivals on new data. This indicates overfitting by the BAC_w variant. The BAC_sw showcases more generalising capabilities as it is the best model across all metrics on the test set. Furthermore, the metrics for BAC_sw do not change significantly from training to test sets.

5. Conclusion

In this paper, an additive model with both wavelet and spline components for estimating the first-order intensity function of NHPP was studied. A novel algorithm inspired by backfitting has been proposed: BAC. This study shows that a model with both spline and wavelet components can help to better capture the peak arrivals of EVs at charging points, using wavelets to capture irregularities and sudden changes in the intensity function. The proposed methodology can be extended to any time step (as long as it is constant) as well as to other wavelet bases (e.g. Daubechies, Meyer). Better performance could be obtained by giving more degrees of freedom (knots) to the splines, but this can sometimes lead to overfitting and could defeat the purpose of the model.

References

Adderly, S.A., Manukian, D., Sullivan, T.D., Son, M., 2018. Electric vehicles and natural disaster policy implications. Energy Policy 112, 437–448.

Aktas, S., Konsuk, H., Yˆıg¯ˆıter, A., 2009. Estimation of change point and compound Poisson process parameters for the earthquake data in Turkey. Environmetrics 20, 416–427. doi:10.1002/env.937.

Amara-Ouali, Y., 2023. Sw-additive. https://github.com/yvenn-amara/SW-additive.

Amara-Ouali, Y., Goude, Y., Massart, P., Poggi, J.M., Yan, H., 2021. A Review of Electric Vehicle Load Open Data and Models. Energies 14, 2233. doi:10.3390/en14082233. number: 8 Publisher: Multidisciplinary Digital Publishing Institute.

Amato, U., Antoniadis, A., 2017. Estimation and group variable selection for additive partial linear models with wavelets and splines. South

African Statistical Journal 51, 37.

Amato, U., Antoniadis, A., De Feis, I., Goude, Y., Lagache, A., 2021. Forecasting high resolution electricity demand data with additive models including smooth and jagged components. International Journal of Forecasting 37, 171–185. doi:10.1016/j.ijforecast.2020.04.001. Ansley, C.F., Kohn, R., 1994. Convergence of the backfitting algorithm for additive models. Journal of the Australian Mathematical Society. Series A. Pure Mathematics and Statistics 57, 316–329. doi:10.1017/S1446788700037721.

Antoniadis, A., Fan, J., 2001. Regularization of Wavelet Approximations. Journal of the American Statistical Association 96, 939–967.

doi:10.1198/016214501753208942.

Aslam, S., Khalid, A., Javaid, N., 2020. Towards efficient energy management in smart grids considering microgrids with day-ahead energy forecasting. Electric Power Systems Research 182, 106232.

Bigot, J., Gadat, S., Klein, T., Marteau, C., 2013. Intensity estimation of non-homogeneous Poisson processes from shifted trajectories. Electronic Journal of Statistics 7. doi:10.1214/13-EJS794.

Brillinger, D., 1997. Some wavelet analyses of point process data, in: Conference Record of the Thirty-First Asilomar Conference on Signals,

Systems and Computers (Cat. No.97CB36136), IEEE Comput. Soc, Pacific Grove, CA, USA. pp. 1087–1091. doi:10.1109/ACSSC.1997. 679073.

Craven, P., Wahba, G., 1979. Smoothing noisy data with spline functions. Numerische Mathematik 31, 377–403.

Daley, D.J., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes. Probability and its Applications, Springer-Verlag, New York. doi:10.1007/b97277.

Ghotge, R., Snow, Y., Farahani, S., Lukszo, Z., van Wijk, A., 2020. Optimized scheduling of ev charging in solar parking lots for local peak

reduction under ev demand uncertainty. Energies 13, 1275.

Gjelaj, M., Hashemi, S., Andersen, P.B., Traeholt, C., 2020. Optimal infrastructure planning for ev fast-charging stations based on prediction of user behaviour. IET Electrical Systems in Transportation 10, 1–12.

Gu, C., Wahba, G., 1991. Minimizing GCV/GML Scores with Multiple Smoothing Parameters via the Newton Method. SIAM Journal on Scientific and Statistical Computing 12, 383–398. doi:10.1137/0912021.

Hastie, T., Qian, J., Tay, K., 2022. An Introduction to glmnet.

Hastie, T., Tibshirani, R., 1986. Generalized Additive Models. Statistical Science 1, 297–318.

Lahariya, M., Benoit, D., Develder, C., 2020. Defining a synthetic data generator for realistic electric vehicle charging sessions, in: Proceedings of the Eleventh ACM International Conference on Future Energy Systems, pp. 406–407.

Liu, R., Ha¨rdle, W.K., Zhang, G., 2017. Statistical inference for generalized additive partially linear models. Journal of Multivariate Analysis 162, 1–15. doi:10.1016/j.jmva.2017.07.011.

Mallat, S., 1999. A wavelet tour of signal processing. Elsevier.

de Miranda, J.C.S., Morettin, P.A., 2011. Estimation of the intensity of non-homogeneous point processes via wavelets. Annals of the Institute of Statistical Mathematics 63, 1221–1246. doi:10.1007/s10463-010-0283-8.

Mu¨ller, M., 2007. Dynamic time warping. Information retrieval for music and motion , 69–84.

Ngailo, T., Shaban, N., Reuder, J., Rutalebwa, E., Mugume, I., 2016. Non Homogeneous Poisson Process Modelling of Seasonal Extreme Rainfall Events in Tanzania. International Journal of Science and Research 5, 12.

Reynaud-Bouret, P., Rivoirard, V., 2010. Near optimal thresholding estimation of a Poisson intensity on the real line. Electronic Journal of Statistics 4. doi:10.1214/08-EJS319.

Salmon, M., 2016. riem: Accesses Weather Data from the Iowa Environment Mesonet.

Streit, R.L., 2010. Poisson Point Processes. Springer US, Boston, MA. doi:10.1007/978-1-4419-6923-1.

UK Department for Transport, 2018. Domestics Chargepoint Analysis 2017. URL: https://data.gov.uk/dataset/ 5438d88d-695b-4381-a5f2-6ea03bf3dcf0/electric-chargepoint-analysis-2017-domestics. library Catalog: data.gov.uk.

Vidakovic, B., 2009. Statistical modeling by wavelets. John Wiley & Sons.

Wand, M., Ormerod, J., 2011. Penalized wavelets: Embedding wavelets into semiparametric regression. Electronic Journal of Statistics 5. doi:10.1214/11-EJS652.

Wood, S.N., 2000. Modelling and smoothing parameter estimation with multiple quadratic penalties. Journal of the Royal Statistical Society:

Series B (Statistical Methodology) 62, 413–428. doi:10.1111/1467-9868.00240.

Wood, S.N., 2017. Generalized Additive Models: An Introduction with R. 2 ed., Chapman and Hall/CRC. doi:10.1201/9781315370279. Youngman, B.D., Economou, T., 2017. Generalised additive point process models for natural hazard occurrence: Generalised additive point process models. Environmetrics 28, e2444. doi:10.1002/env.2444.

Zhang, X., Grijalva, S., 2015. An advanced data driven model for residential electric vehicle charging demand, in: 2015 IEEE Power & Energy Society General Meeting, IEEE, Denver, CO, USA. pp. 1–5. doi:10.1109/PESGM.2015.7286396.