Energy Reports 11 (2024) 2535–2553

Contents lists available at ScienceDirect

Energy Reports

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

Research paper

Probabilistic net load forecasting framework for application in distributed integrated renewable energy systems

Jan-Simon Telle *, Ajay Upadhaya, Patrik Schönfeldt, Thomas Steens, Benedikt Hanke, Karsten von Maydell

DLR Institute of Networked Energy Systems - Department of Energy Systems Technology, Carl-von-Ossietzky-Straße 15, 26129, Oldenburg, Germany

ARTICLE INFO

Keywords:
Probabilistic net load
Sector integrated systems
Probabilistic forecasting
Quantile personalized standard load profile

ABSTRACT

Integrating various sectors enhances resilience in distributed sector-integrated energy systems. Forecasting is vital for unlocking full potential and enabling well-informed decisions in energy management. Given the inherent variability in generation and demand prediction, quantification of uncertainty is crucial. Therefore, probabilistic forecasting is becoming imperative compared to deterministic forecasting, as it ensures a more comprehensive depiction of uncertainty. This paper introduces probabilistic net load forecasting framework (PNLFF), a non-blackbox approach that is robust, non-parametric, computational and data inexpensive, and adaptable across sectors. It utilizes the personalized standard load profile for deterministic forecasts, and integrates quantile regression to generate probabilistic forecast. The cumulative distribution function is approximated from quantiles of probabilistic forecast using piecewise cubic hermite interpolating polynomial, and then it is derived to probability density function (PDF). Then the probabilistic net load was obtained by the convolution of PDFs for electricity demand, heat demand and PV generation. A case study demonstrates its application in operational optimization for a distributed energy system of the logistics facility. In the first stage of the PNLFF, the results of the personalized standard load profiles clearly show that they can be applied in all sectors and outperform their respective benchmarks. The second stage, the probabilistic expansion using quantile regression, also performs promisingly across all sectors, with the best results being achieved in particular with a small training data set of 30 days. With the extension of the quantiles and interpolation, it was demonstrated how a PDF can be approximated without prior knowledge of the distribution of the data. The result of the case study demonstrate that the PNL, as an aggregated PDF of the different sectors by convolution, can be used for decision making under uncertainty, e.g. for the planning of flexible loads.

1. Introduction

1.1. Background

Sector integration is already considered a cost-effective and efficient approach for decarbonizing the energy system and transforming it from a fossil and centralized basis, to a renewable and more decentralized one (Pavičević et al., 2020; This makes sector coupling an important building block towards a climate-neutral energy system in 2050, 2022). This makes sector integration an important component for achieving a climate-neutral energy system in 2050, which is a goal of the European Union (European Comission, 2018, 2019). Ever more decentralized generation and electrification in all sectors (i.e., "sector coupling", sector integration, or P2X) in local grids are creating a growing need for energy and load management at this level. This also applies to commercial and industrial properties that have not previously come into contact with the issue of energy beyond the paying of their energy bills (European Commission, 2020; Fridgen et al., 2020).

At the decentralized level, it is becoming increasingly important to operate across generation and consumption, as well as across sectors (i.g., electricity, heating and cooling, transportation, etc.). Thus, power demand and generation forecasts must also be considered together. The increasing number of electrified consumers, such as heat pumps (Baumann and Kepplinger, 2023) or electric cars, poses challenges for local energy and load management initiatives. However, opportunities also arise from optimized operation of the energy system. For example, own-consumption can be maximized, peak loads reduced, and CO_2 emissions minimized. This can reduce independence from energy prices and minimize the need to expand the public electricity grid. In order to optimize the operation of devices that are part of a functioning energy system, both forecasts and the quantification of their uncertainties are important.

* Corresponding author.
E-mail address: jan-simon.telle@dlr.de (J.-S. Telle).

https://doi.org/10.1016/j.egyr.2024.02.015
Received 23 October 2023; Received in revised form 12 January 2024; Accepted 10 February 2024
Available online 16 February 2024

Nomenclature

Abbreviations

BEV	Battery Electric Vehicle
CDF	Cumulative Density Function
DNL	Deterministic Net Load
el	Electricity
ht	Heat
IQR	Interquartile Range
MAE	Mean Absolute Error
$MAPE$	Mean Absolute Percentage Error
max, opt	Optimal Maximum Value
max	Maximum Value
$MIPW$	Mean Prediction Interval Width
ML	Machine Learning
NL	Net Load
$PCHIP$	Piecewise Cubic Hermite Interpolating Polynomial
PDF	Probability Density Function
PI	Prediction Interval
$PICP$	Prediction Interval Coverage Probability
PNL	Probabilistic Net Load
$PNLFF$	Probabilistic Net Load Forecasting Framework
$PSLP$	Personalized Standard Load Profile
$QPSLP$	Quantile Personalized Standard Load Profile
QR	Quantile Regression
QRA	Quantile Regression Averaging
$QRNN$	Quantile Regression Neural Network
$RMSE$	Root Mean Square Error
SLP	Standard Load Profile
ssa	Summer Saturday
ssu	Summer Sunday
std	standard
sw	Summer Working Day
$threshold$	Threshold Value
tsa	Transition Saturday
tsu	Transition Sunday
tw	Transition Working Day
wsa	winter Saturday
wsu	Winter Sunday
ww	Winter Working Day

Symbols

$*$	Convolution
α	Conficdence Interval
δ	Interval Width
$\frac{dP}{dx}$	PCHIP Derivative
$\hat{y}$	Predicted Value
$\mu, E(X)$	Expected Value
ρ	Quantile loss function
τ	Quantile Probability Level
C	Coverage Factor
f_{XY}	Joint PDF of Random Variables X and Y
f_X, f_Y	PDF of Random Variables X and Y
i, t	Time Point
n, N	Number of Samples
$P(x)$	Cubic Hermite Interpolation Polynomial
q	Predicted Quantile
U, L	Upper, Lower Interval
u	Error Term
X, Y	Random Variables
x_k, x_{k+1}	Sub-Intervals
y	Measured Value
z	Sum of Random Variables X and Y

Our view on this is that the renewable sectors integrated energy systems (P2X applications) and the expansion of decentralized generation will increase the need for forecast-based decentralized energy management solutions. Many households, companies, and other small decentralized energy systems will not be able to afford big data analysis or high costs for energy management services, but would nevertheless benefit from data driven energy management. For them, simple, highly automated, and pragmatic solutions are needed. Decentralized data acquisition (e.g., smart meters, solar photovoltaic (PV), and battery inverters) is the basis and makes it possible to operate more independently from third party providers and cloud services. If such data is collected and stored locally across sectors, it can be used to help plan and operate distributed integrated energy systems in a cheap and resilient manner. This paper offers solutions by leveraging such data to formulate cross-sector predictions and the quantification of its uncertainties.

1.2. Literature review

In order to be able to plan and operate within the temporal differences between generation and consumption, and optimize, e.g., grid operation, costs, or emissions, forecasts have been developed and used for a long period of time (Petropoulos et al., 2022; Ming et al., 2017; Dong et al., 2023). In recent years, deterministic artificial intelligence (AI) based machine learning (ML) models have gained interest among the researchers (Gaamouche et al., 2022). They are extensively used and further developed in energy domain, for both load demand and renewable energy forecasting (Wang et al., 2019b; Mawson and Hughes, 2020; Benti et al., 2023). A comprehensive reviews were conducted on machine learning and deep learning techniques for forecasting renewable energy generation (Gaamouche et al., 2022; Benti et al., 2023) and electricity demand (Eren and Küçükdemiral, 2024). Despite their high prediction accuracy and ability to capture non-linearity in volatile data, they might not be the perfect fit for application in small-scale decentralized energy management solutions. This is due to the fact that they are highly complex, requires big data handling, time and expertise demanding for data pre- and post-processing, computationally expensive, which ultimately results in high cost for energy management services. In particular, deterministic machine learning approaches have been further developed in recent years. Today, however, distribution network operators often still rely on so-called standard load profiles (SLPs) for planning and balancing consumption and the generation of small consumer loads, such as for residential buildings or small businesses (VDEW, 1999; Association et al., 1997). At the distribution grid level, SLPs are used to approximate consumption when no measurements are available and are applied at different temporal and spatial scales. In their study, “Are standard load profiles suitable for modern electricity grid models?”, Peters et al. show that a spatially higher resolution leads to better forecasting scores (Peters et al., 2020). With the “smart meter rollout” taking place globally is leading to an increased availability of data at the decentralized level, which can be effectively utilized in grid operation and energy management (Wang et al., 2019a) and to improve SLPs (Hinterstocker et al., 2014; Scholz and Müsgens, 2017). In recent years, deterministic prediction of electricity (Petropoulos et al., 2022; Nti et al., 2020; Klyuev et al., 2022) and heat load (Leiprecht et al., 2021; Bergsteinsson et al., 2021), as well as PV generation power (Antonanzas et al., 2016a; Das et al., 2018), has also become a common tool. Personalized standard load profiles (PSLPs) (Hinterstocker et al., 2014), on the other hand, are more of an evolutionary step beyond today's standard of grid operators (VDEW, 1999). We compared this approach using locally-collected data in conjunction with machine learning techniques and found that there can be advantages, especially when data availability or computational power are limited (Telle et al., 2020; Steens et al., 2021). In order to optimize the use of renewable energy, it is necessary to align the generation of renewable energy sources and consumption, minimizing the residual or net loads. The capacity to deterministically predict the net load has been further developed in recent years (Telle et al., 2020; Zhou et al., 2021; Beichter et al., 2022). For robust and safe optimization and scheduling of a power system, it is indispensable to be able to evaluate the uncertainties inherent to corresponding the forecasts (do Amaral Burghi et al., 2020). Probabilistic forecasts are therefore developed to assess these (Bjerregård et al., 2021). Literature on probabilistic forecasting is limited compared to point forecasting but research interest on this has grown significantly in past years (Hong et al., 2016). Many probabilistic approaches to generation and load prediction have merged in recent years (Klyuev et al., 2022; Hong and Fan, 2016; Zhang et al., 2014). Critical to the application of predictions and evaluation of uncertainty is the choice of evaluation metrics. For this purpose, detailed reviews of known techniques, application examples, and challenges have been published (Abdar et al., 2021; Bjerregård et al., 2021). In addition to the metrics, it is also necessary to determine whether the uncertainty is aleatoric or epistemic (Badings et al., 2023). There are a few examples of applications in which multiple sectors are taken into account (Fatema et al., 2023), e.g., the use transportation data to better predict electricity demand.

In this study, we combine PSLP with our approaches to build a low-effort generation forecasting model for small-scale (residential) PV power systems (Hanke et al., 2018; Maitanova et al., 2020) to create a probabilistic net load prediction. In order to do so, we employ quantile regression (QR) (Cleophas and Zwinderman, 2021; Koenker and Bassett, 1978) - a non-parametric approach, which has the advantage that no prior knowledge regarding the distribution of the data is required, but it may require a higher computational effort in comparison to a parametric approach (Zhang et al., 2014; Wang et al., 2018). QR has already been shown to be applicable in the field of load or energy prediction (González Ordiano et al., 2020; Liu et al., 2017). Point forecasts from several sister deterministic models are used to provide explanatory variables in QR with an approach called quantile regression averaging (QRA) (Liu et al., 2015; Mei et al., 2020). However, it requires big data, computationally expensive, and increase complexity in implementation due to the need of several deterministic models. Furthermore, high level of expertise for data pre-processing is required. To simplify the approach we use single statistical deterministic model whose output is used as an explanatory variable in QR model to generate quantile predictions. For the operational optimization of energy systems with a high share of renewable generation, the net load is a key parameter. In recent years, several approaches to predicting deterministic net load (DNL) have been unveiled (Falces et al., 2023; Zhang et al., 2023). However, the DNL do not provide uncertainty information and hence shift towards the probabilistic net load (PNL) becomes vital. The PNL predictions can be performed at various levels of aggregation (Beichter et al., 2022), (Browell and Fasiolo, 2021; Wang et al., 2018). Determining the cumulative density function (CDF) and probability density function (PDF) from a non-parametric prediction requires further calculation steps. In order to determine the CDF from predicted quantiles, an interpolation between the quantiles is required. Cubic spline interpolations can be used for this purpose. The piecewise cubic hermite interpolating polynomial (PCHIP) interpolation can be used to preserve the geometric shape of the predicted data (Fritsch and Butland, 1984; Ariffin and Karim, 2014; Barker and McDougall, 2020). When aggregating individual probabilistic predictions, the probability densities of two or more continuous random variables must be convolved as a joint probability density and, if applicable, including the dependencies to each other (Li et al., 2020; Beichter et al., 2022). This allows the uncertainties from various forecasts to be aggregated. The summary of related work is given in Table 1.

1.3. Contribution and paper outline

This paper makes the following contributions:

(a) In this study, the probabilistic net load forecasting framework (PNLFF) was developed to generate probabilistic net load (PNL) by convoluting forecast PDFs (electricity demand, heat demand, and PV generation) in distributed sector-integrated energy systems. The framework is characterized by its simplicity, robustness, low cost, non-parametric nature, computational efficiency, adaptability, and reliance on minimal locally collected measurement data. As a part of PNLFF, the following sub-contributions ((b), (c), (d), and (e)) were made.
(b) The PSLP was enhanced by introducing new modes ("fix" and "variable"), and possibilities to choose seasonality and type days as an option. Its application was extended to the heat and PV generation sector for deterministic forecasting.
(c) Simplistic approach to generate probabilistic forecast from deterministic forecast using QR was adopted in this study. The point forecast output from PSLP serves as an explanatory variable for the QR, wherein uncertainty within the PSLP is accounted for and transforms into the quantile personalized standard load profile (QPSLP).
(d) Approach how the CDF can be approximated and PDF derived from the quantiles of the QPSLP with the application of a PCHIP interpolation, to further determine the PNL (joint PDF) of all sectors via convolution.
(e) Lastly, a case study was demonstrated to present the application of the framework in operational optimization of the distributed energy system under uncertainty. This study provides the simplistic approach to determine the PNL and use them to quantify uncertainty and optimize such system technically, economically and ecologically.

This study outlines the following topics:

In Section 1 of the paper, a comprehensive framework is introduced with the aim of predicting distributed generation and load demand. Specifically, the framework focuses on the quantile personalized standard load profile (QPSLP) pertaining to the sectors of electricity, heat, and photovoltaic (PV) generation (see Fig. 1). This PNLFF also seeks to assess and quantify the inherent uncertainty associated with forecasts across these distinct sectors, along with the aggregated forecast of a distributed energy system. Section 2 includes a presentation and detailed description of the PNLFF development. For this purpose, an overview of all stages of PNLFF is first presented in order to introduce them step by step. The first step of the section introduces the case study and the data used. For this purpose, the decentralized energy system of a logistics facility is employed as an example to demonstrate the PNLFF. In the subsequent sections of the methodology, the deterministic (PSLP and Benchmark) and probabilistic prediction methods (QR and Benchmark) used are introduced, and the evaluation metrics and quantification of uncertainty are then explained. Furthermore, the aggregation methods for determining the decentralized residual QPSLP and the evaluation of its uncertainty are outlined. The results, reported in Section 3, present the outcome of the forecasting framework and the uncertainty quantification using case study 4 as an example. Using the example of an additional electrical load to be integrated, the PNL is used to demonstrate an optimized application The results are then evaluated in the discussion section reviewed and for their transferability; possible future work lines of research is also proposed.

Table 1
Summary of the related work.

Source	Method/Main Idea	Summary	Limitations	Relation to this work
Wang et al. (2019b), Gaamouche et al. (2022), Benti et al. (2023)	ML models for deterministic forecasting renewable generations (PV, and wind power)	Discussed and reviewed ML and deep learning techniques for renewable energy forecasting. It is found to be more accurate than statistical method.	Highly computational, complex, big data requirement and handling, expertise demanding for pre- and post-processing, expensive to manage for application in small-scale level	Gaps to be addressed in order to develop low-cost and computationally inexpensive forecasting approach for decentralized energy system
VDEW (1999)	SLP method for load forecasting	Distribution network operators often use standard load profiles of 1209 buildings which are grouped into residential, commercial, and agriculture for forecasting small consumer loads.	Lacks accuracy in forecasting individual load as they lack updates and the standard load profiles might have less correlation with specific building of interest	Limitations of SLPs in forecasting for small-scale and sector-integrated energy system were acknowledged.
Hinterstocker et al. (2014)	PSLP model	An extension of SLP, PSLP was introduced. This method uses locally collected measured data for creating load profiles that are divided into various curves depending on type of days like weekdays, Saturday, Sunday, Holidays, and by seasons like winter, summer, and period of transition. PSLP showed significant improvement in load forecasting for individual building of interest compared to SLP.	Limited to electricity load forecasting.	This method was further improved and developed for making it applicable for heat demand and PV generation forecast.
Telle et al. (2020), Steens et al. (2021)	PSLP model improvement	Further developed and implemented PSLP model (Hinterstocker et al., 2014) for different building profile to forecast electricity demand. The comparison with machine learning algorithms shows that the PSLP performs only slightly worse but with significantly less data and computational effort.	The authors only performed PSLP on electricity demand forecast.	In our work, PSLP was further developed and extended to perform deterministic forecasting for heat demand and PV generation.
Telle et al. (2020)	Deterministic net load (DNL) forecasting for building.	An approach to forecast the deterministic net load (the difference between forecast energy demand and renewable supply) for a commercial building to optimize integrated energy system	Deterministic residual load fails to provide stochastic optimization for uncertainty quantification.	The acknowledgment of the significance of probabilistic net load, as opposed to DNL, in energy optimization and control was noted. This recognition seeks to quantify uncertainty in forecasting, thereby improving decision-making capabilities in local energy management systems.
Beichter et al. (2022), Li et al. (2020)	Probabilistic net load forecasting (PNL)	Importance of net load forecasting in maintaining stability under high proportion of renewable generation is discussed. The paper compares three different net load forecasting approaches that exploits different levels of aggregation. Both deterministic and probabilistic net load forecasting were obtained using three different aggregation strategies.	In this case, the PNL was determined for a large energy system and a parametric approach was chosen for the probabilistic prediction. This requires assumptions to be made about the distribution of the data.	The aggregation approach for probabilistic net load forecasting proposed in Beichter et al. (2022) is adopted in our work, but in our case with an non-parametric approach.
Fritsch and Butland (1984), Ariffin and Karim (2014), Barker and McDougall (2020)	PCHIP	Interpolation of time series with cubic spline methods can lead to unrealistic behavior especially in the edge cases. Different variations (rotated PCHIPs [Barker and McDougall, 2020]) of PCHIP Interpolation shows the advantage of PCHIP interpolation compared to cubic spline interpolation. The advantages of a PCHIP interpolation compared to spline one is the lower effort and fewer oscillations if the data are not smooth and no overshoots occur.	PCHIP interpolation was applied in different areas, on different time series. So far, however, it has not been used to approximate the PDF from the quantiles of a non-parametric prediction.	In this work, the PCHIP interpolation was chosen in order to be able to interpolate the boundary calculations in particular without overshoots. Which would not be possible with cubic spline interpolation.
Hong et al. (2016)	"Global Energy Forecasting Competitions 2014 and beyond."	There has been a significant surge in research interest in the probabilistic forecasting in energy domain.	Detailed description of methods is missing.	This paper motivates us to shift from traditional point forecasting to probabilistic forecasting in order to quantify uncertainty with increase in decentralization, penetration of renewable sources, and increase in variable loads. Importance of connection between probabilistic forecasting and point forecasting was also realized.
Hong and Fan (2016), Zhang et al. (2023)	Probabilistic methods for load and generation forecasts	Reviews of various methods for load and generation forecasting. Comparisons between parametric and non-parametric methods.		Motivation for our paper to chose a non-parametric approach

(*continued on next page*)

Table 1 (*continued*).

Source	Method/Main Idea	Summary	Limitations	Relation to this work
Koenker and Bassett (1978)	QR	QR is an extension to ordinary least square regression model. It models the relationship between independent variables or explanatory variables (X), and the conditional quantiles of dependent variable (y).	Choosing explanatory variable that provides high correlation with the forecasted data is challenging.	This method is combined with PSLP, where the deterministic forecast is used as input variable to QR to generate quantile predictions.
Liu et al. (2015)	QRA for load forecasting	Point forecast outputs from several deterministic models are used as explanatory variable in QR model to generate probabilistic forecast at different quantile levels.	Highly complex and big data requirements due to need of several point forecast models.	The literature review on this methodology provides us an idea to further simplify the approach by using only point forecast output from single deterministic model as an explanatory variable in QR model.
Mei et al. (2020)	QRA for PV generation forecasting	Ensembles a group of independent long short-term memory (LSTM) deterministic forecasting models for obtaining the probabilistic forecasting of PV output.	Highly complex and big data requirements due to need of several point forecast models. Use of machine learning based deterministic models further increases complexity, and computational time.	This paper gives us an idea to use statistical deterministic model to generate point forecast as an input variable in QR instead of using machine learning based deterministic model in order to reduce complexity, computational time and expertise requirements.

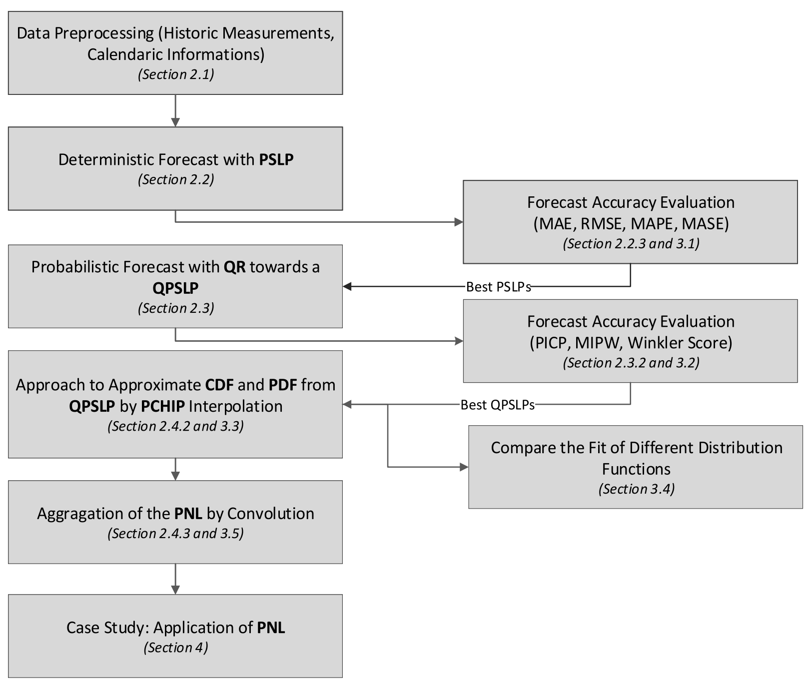

Fig. 1. Flow chart of the probabilistic net load forecasting framework.

2. Methodology

2.1. Probabilistic net load forecasting framework

2.1.1. Framework overview

The PNLFF, illustrated in Fig. 1, can be roughly explained in four steps. In the first, locally-collected data is processed and calendar features are created (e.g. public holidays). In the second step, the data are utilized as input for the creation of a deterministic forecast. As a deterministic model, the PSLP is used to forecast the next 24 h (day-ahead) of the electricity demand, heat demand, and PV generation. Persistence models are employed as a benchmark, or rather for naïve predictions. The PSLP prediction accuracy is evaluated using the four commonly used metrics of mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and mean absolute scaled error (MASE) and compared with the benchmark in order to select the best models for further steps (Antonanzas et al., 2016b; Hyndman and Athanasopoulos, 2018). The functionality and application of the PSLP are described in detail in Section 2.2. For the third step, we performed a QR based on the PSLP, as a probabilistic extension towards a quantile PSLP or QPSLP. In order to measure the performance and uncertainty of the QPSLP, the evaluation metrics including PICP, MIPW, and Winkler score are considered. In the last step, an approach is presented to approximate the CDF and PDF from the QPSLP and how convolution can best determine the joint density functions and PNL.

2.1.2. Data description

To demonstrate the PNLFF, the energy system of a logistics facility (Anon, 2019-2024) was used. A detailed description of the cross-sector energy system was described by the authors in Telle et al. (2022). In this paper, the focus is initially on the data. The most important parameters for the data used are shown in Table 2. The electricity and heat demand data were used from measurements collected as part of the ELogZ ("Energieversorgungskonzepte für Klimaneutrale Logistikzentren") project (Anon, 2019-2024). The heat is provided within the system by heat pumps. The electricity demand of the heat pumps is not included in the measured electricity data. The PV generation

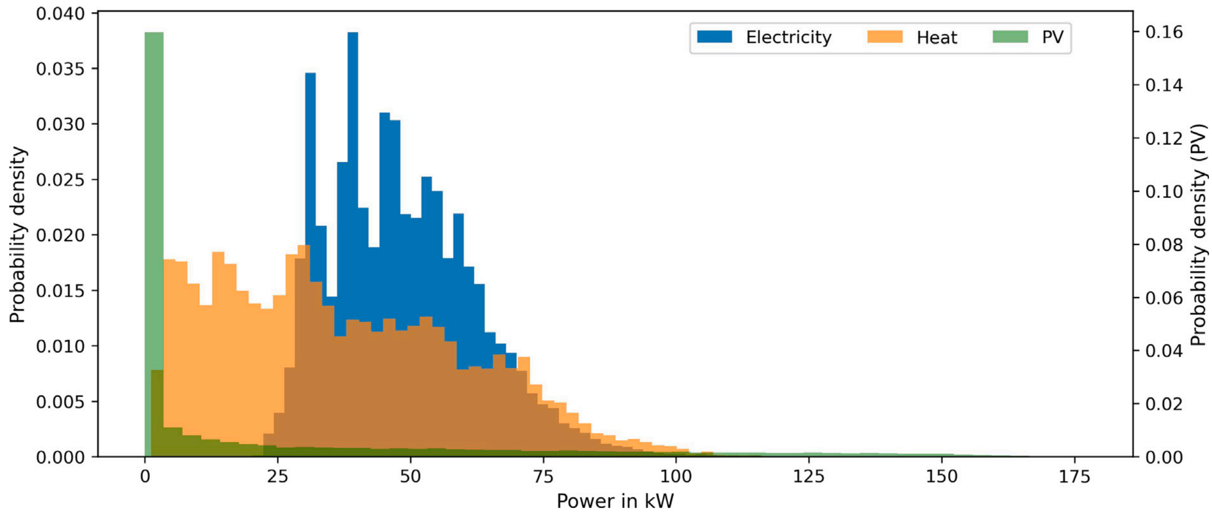

Fig. 2. Density distribution of electricity, heat load, and PV generation over the entire observation period.

Table 2
Data description.

Data	Resources	Details
Electricity Demand	measurements (Anon, 2019-2024)	15-min resolution, power in kW
Heat Demand	measurements (Anon, 2019-2024)	15-min resolution, power in kW, supplied by heat pumps
PV Generation	PV Lib simulation (Holmgren et al., 2018) with publicly available weather reports from Anon (2023)	East–west oriented, 200 kW PV system, 10° inclination angle, 15-min resolution, power in kW
Observation period		September 5, 2021 to September 30, 2022

power was generated using the Python library pvlib (Holmgren et al., 2018). For this purpose, a system that can realistically be installed on the demonstration premises with a nominal power of 200 kW was assumed. The modules were oriented half to the east and half to the west, with an inclination angle of 10°. Weather data, especially radiation, temperature and wind speed, which served as inputs for the PV simulation, were obtained from open DWD (Anon, 2023). All data was utilized in a 15-min resolution. Fig. 2 shows the distributions, in the form of a density distribution, of the data used. The PNLFF described in detail in the following sections is subject to the principles listed in Table 3. The observation period of the data is from September 5, 2021 to September 30, 2022.

2.1.3. Framework modeling and simulation fundamentals

The PNLFF described in detail in the following sections is subject to the basics listed in Table 3. The framework was developed in the programming language Python. The PSLP and QPSLP forecasts are run in a so-called rolling forecasts horizon, i.e. every day the training window (e.g. 30 days historical measurements or historical PSLPs) was moved forward by one day and the parameters of the models were retrained for the next day. Since the data is available in a 15-min resolution, 96 time steps are predicted in each forecasting window.

The interpolation of the CDF and PDF from the QPSLP is carried out on an intraday basis. This produces a instantaneous CDF and PDF for each point in time of the forecast horizon. The PNL is then determined as the product of the convolution of the PDFs of the individual sectors. Table 3 also lists the most important Python libraries that were used as a part for modeling the QR, PCHIP interpolation and convolution.

2.2. Personalized standard load profile - PSLP

2.2.1. PSLP fundamentals

In this study, the PSLP was applied and extended to include the sectors of heat load and PV generation. Different adaptation options were presented to address various load and generation behaviors. A persistence approach or naïve forecast is used as a benchmark for preparing deterministic forecasts. Three different approaches are considered, wherein the naive forecast corresponds to the values of the last day (d-1), the day before last (d-2), or the values from a week ago (d-7). The PSLP process and its extensions are described in detail in the following section.

The PSLP was first described by Hinterstocker et al. (2014) and further developed in other studies (Telle et al., 2020; Steens et al., 2021). It provides the ability to generate forecasts from historical measured data, which marks a significant improvement over the SLP (Telle et al., 2020; Steens et al., 2021). It starts from the basic idea that the use of standard load profiles is no longer sufficient (Peters et al., 2020). For instance, the SLP (VDEW, 1999) relies on annual energy demand, wherein the year is divided into three seasons and each day is assigned to a characteristic day (see Table 4) and public holidays are assigned to a Sunday. The following SLPs were proposed: seven SLPs were available for commercial/industrial enterprises, two for agricultural enterprises, and one for residential buildings. With the help of a dynamic modification factor, the profiles can be made even more variable. An annual consumption of 1000 kWh/a was defined as the standardization basis, i.e., the sum of all consumption values of one year results in 1000 kWh. The PSLP also builds up on these characterizations. For the PSLP, locally-collected measured data on the power demand are required. This results in the following characteristics: winter workday (ww); winter Saturday (wsa); winter Sunday (wsu); summer workday (sw); summer Saturday (ssa); summer Sunday (ssu); transition workday (tw); transition Saturday (tsa); and transition Sunday (tsu). The process of the PSLP can be traced in Fig. 3.

2.2.2. PSLP modes and options

In order to be able to also use the PSLP for production or consumption time series, which do not depend on weekdays, e.g., PV production, there is the option of switching off the characterizations. That is, it is

Table 3
Modeling and simulation parameters of the PNLFF.

Method	Forecast Horizon	Training Days	Training Frequency	Python Libraries
PSLP	day-ahead (96 time-steps)	variable, depends on PSLP mode (e.g. 30 days in standard mode)	daily	Own development
QPSLP (QR)	day-ahead (96 time-steps)	30, 90 or 120 days	daily	statsmodels (Seabold and Perktold, 2010), (smf.quantreg)
Method	**Execution Horizon**	**Aggregation Level**	**Execution Frequency**	
PCHIP	intraday (1 time step)	sector-specific (CDF and PDF interpolation)	each time step of the forecast horizon (96 times)	SciPy (Virtanen et al., 2020) (pchip_interpolate)
Convolution	intraday (1 time step)	aggregated (PNL)	each time step of the forecast horizon	SciPy (Virtanen et al., 2020) (convolve) (convolve)

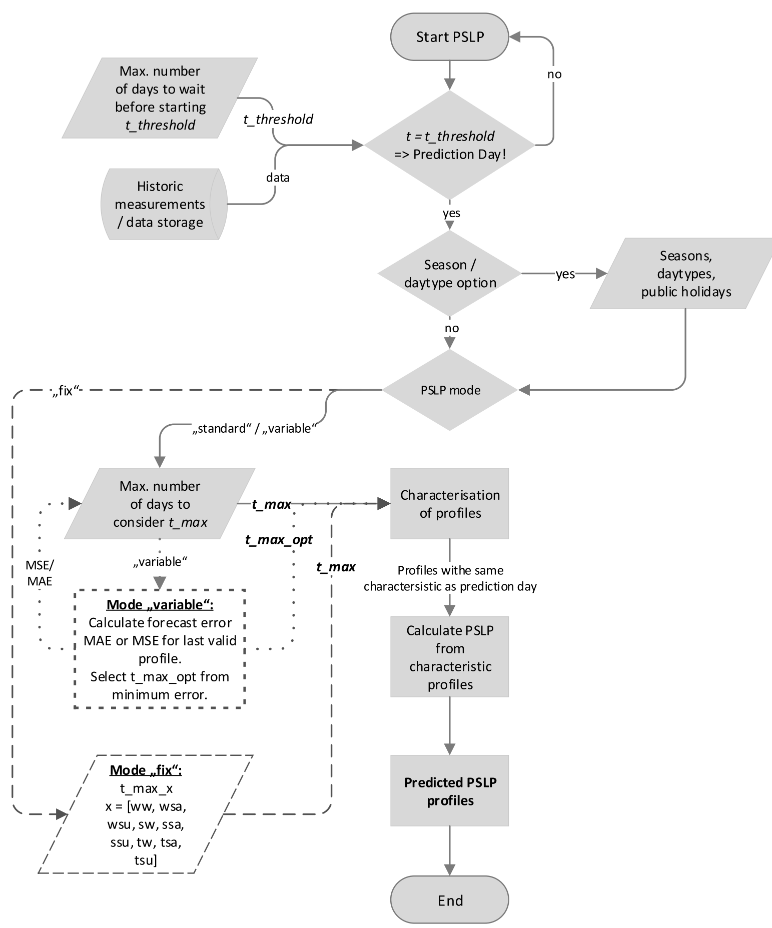

Fig. 3. Flow Chart of the PSLP algorithm. The three different modes "standard" (solid lines), "fixed" (dashed), and "variable" (dotted) are distinguished in the selection method of the maximum past days (t_{max} or $t_{max,opt}$).

resolved to make a division after weekdays (working days, Saturday, Sunday) and/or after seasons (winter, summer, transitions). In the first instance, there would only be division into winter, summer, and transitions. In the second, there would only be a distinction between working days, Saturdays and Sundays. If both options are switched off, there is no more differentiation and all days that were specified are used, with the maximum number of past days t_{max} to be considered for the PSLP forecast.

The process of a standard PSLP can be seen in Fig. 3. Before the PSLP prediction can be made, a minimum amount of historical data is needed for the data storage. This means, that for the respective day that is prognosticated, at least one day from the past is needed that has the same characteristics, namely: ww, wsa, wsu, sw, ssa, ssu, tw, tsa or tsu (see Table 4). However, one day, would only correspond to a naïve forecast. Assuming the threshold $t_{threshold}$ is set to 21 days, the first forecast can only be made for day 22. Within the PSLP, there

Table 4
PSLP characteristics.

Characteristic season	Characteristic dates	Characteristic day
Winter (w)	01.11.-20.03.	Workday (w), Saturday (sa), Sunday/holidays (su)
Summer (s)	15.05-14.0.	w, sa, su
Transition (t)	31.03-14.05. 15.09.-31.10.	w, sa, su

Table 5
PSLP variation possibilities.

PSLP type	PSLP modes
	standard, fix, variable
Aggregation mode	Mean, median
Season	True, False
Day type	True, False

is an option to determine the days from the past to be considered in the forecast. For this purpose, a value t_{max} must be specified. If this is, e.g., 21 days, the previous 21 days are taken into account for the creation of the forecast. The value t_{max} cannot be higher than $t_{threshold}$: $t_{max} <= t_{threshold}$. In the special case when the day to be predicted falls in a new season and there are not yet any corresponding days in the data storage, the last day that matches the characteristics attributes is used. The prediction profiles are then calculated using the data, incorporating the same characteristics as the day to be predicted. If more than one historic profile is available with the same characteristics, the mean value is calculated for each time point within the profiles.

The PSLP can be run in three different modes, which are differentiated in Fig. 3 as "standard" (solid lines), the newly introduced "fix" (dashed), and "variable" (dotted) modes Table 5. The "standard" mode corresponds to the previous description with a fixed value of t_{max}. In the "fix" mode, a separate t_{max} can be specified for each of the nine characteristics (ww, wsa, wsu, sw, ssa, ssu, tw, tsa or tsu). This can help better take into account seasonal or daytype differences. That is, in order to determine the forecast profiles, different numbers of historical daily profiles are used depending on the season and day characteristics if different values of t_{max} were chosen (see Table 5).

If the variable mode is selected, a maximum value t_{max} is stated. However, before each prediction, the optimal $t_{max,opt}$ is checked, as is shown in Fig. 3. This is the number of historical day profiles that should be used for the forecast. For this, the last available day with the same characteristics as the prediction day at timepoint t is selected as a reference point. Afterwards, the data storage is scanned from $t-1$ to $t-t_{max}$. In the first step, only one day with the same characteristics is looked at and the error against the reference point is determined. Either the MAE (1) or MSE (2) can be selected for the error calculation (see also Section 2.2.3). Thus, on how large t_{max} must optimally be in relation to the reference point is analyzed in an itarative form. Furthermore, a kind of "early stopping" is used, i.e., after a certain number of iteration steps in which the calculated error has not improved, it is stopped and the best result or t_{max_opt} with the smallest errors is used.

2.2.3. PSLP evaluation metrics

This study utilizes four commonly-employed evaluation metrics for the assessment of deterministic forecasts (cf. Antonanzas et al. (2016b), Hyndman and Athanasopoulos (2018), Hyndman and Koehler (2006)). These metrics include the MAE, MSE, RMSE, MAPE, and MASE as given by Eqs. (1), (2), (3), (4), and (5).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}| \tag{1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2} \tag{3}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}}{y_i}\right| \tag{4}$$

$$MASE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{\frac{1}{n-m}\sum_{j=m+1}^{n}\left|y_j - y_{j-m}\right|} \tag{5}$$

where y_i is the measured value and $\hat{y}$ the predicted value at time point i, n is the number of samples, and y_{j-m} a naive forecast with m as the number of previous days for a naive forecast. Compared to the MAE, MSE, and RMSE, the MASE and MAPE are scale-independent. MASE is obtained by comparing the MAE of the forecast with the MAE of a naive forecast. Thus, a forecasting method can be considered reliable if the MASE < 1. The MAPE is only used to evaluate deterministic forecasts in the electricity and heat sectors, as PV generation (y_i) often has zero values and is therefore only suitable for $y_i > 0$.

2.3. Quantile personalized standard load profile - QPSLP

The following section describes how the PSLP is extended to a probabilistic forecast, a quantile PSLP (QPSLP), using quantile regression (QR).

2.3.1. Quantile regression

QR is a statistic approach developed in the 1970s by Koenker and Bassett (1978) as an extension of linear models. It essentially models the relationship between independent variables or explanatory variables (X), and the conditional quantiles of dependent variable (y). As opposed to linear regression, QR provides a more comprehensive picture of the effect of the independent variable on the dependent one. QR is expressed in linear form as:

$$Q_y(q|x) = \beta_q X_t = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_p x_{i,p}, \tag{6}$$

where Q_y is the conditional q_{th} quantile of the load/generation distribution(y), q the quantile level, x the feature vector, β_q an estimated vector of parameters for quantile q with an unknown coefficient, and X_t the corresponding input feature vector at time t.

Quantiles are estimated by giving asymmetric weights to the error defined by the pinball loss function given in Eq. (7).

$$\rho_\tau(u) = \begin{cases} \tau u & \text{if } u \geq 0 \\ (\tau - 1)u & \text{if } u < 0 \end{cases} \tag{7}$$

where τ is the quantile probability level and ranges between 0 and 1 and u is the error term. The error term is given by $y_t - q_t$, where q_t is the quantile forecast and y_t the actual value.

Given the pinball loss function, the optimization problem of QR can be expressed as Eq. (8):

$$\begin{aligned}\hat{\beta}\tau &= \arg\min \beta \sum_{i=1}^{N}\rho_\tau(y_i - \beta x_i) \\ &= \arg\min \beta \sum_{i=1}^{N}\rho_\tau(y_i - \beta_0 + \beta_1 x_{i,1} \\ &\qquad + \cdots + \beta_p x_{i,p})\end{aligned} \tag{8}$$

Minimization of the quantile loss function ρ_τ is conducted separately for each τ. Following the minimization problem, regression coefficients are obtained for each quantile (τ), which could in turn be used to obtain the distribution of the forecast at different quantile levels. In this work, we use the point forecast from PSLP model as a single feature (explanatory variable) in QR. For this purpose, a historical period, training or calibration window, was determined (spanning 30, 90 and 120 days) from which the PSLPs forecast were taken into account. The QPSLP, as well as the PSLP, was run as a rolling forecast of a day-ahead prediction. Within the rolling scheme, i.e., every day the training window (30, 90 or 120 days) was moved forward by one day and the parameters of the models were recalculated for the next day or the next QPSLP. This was performed separately for all of the sectors considered.

2.3.2. QPSLP evaluation metrics

This paper evaluates probabilistic forecasting based on three key properties, namely: sharpness, reliability (calibration), and resolution (Dang et al., 2022). In order to fully assess these aspects, a set of well-established evaluation metrics are employed, namely prediction interval coverage probability (PICP), mean prediction interval width (MPIW), and Winkler Score (WS). The PICP quantifies the reliability of the forecasts, the MPIW measures their sharpness, and the WS holistically considers both of these factors. It is important to note that there is often a trade-off between PICP and MPIW, as a higher PICP and lower MPIW are both desired but can conflict with one another (Khosravi et al., 2010). The PICP, and MPIW are expressed by Eqs. (9), and (10), respectively. For a central $(1-\alpha)$ prediction interval (PI), WS is expressed as Eq. (11).

$$PICP = \frac{1}{N}\sum_{i=1}^{N} C_i \tag{9}$$

where C_i is defined as:

$$C_i = \begin{cases} 1, & \text{if } y_i \in [L_i, U_i], \\ 0, & \text{if } y_i \notin [L_i, U_i] \end{cases}$$

$$MPIW = \frac{1}{N}\sum_{i=1}^{N}(U_i - L_i) \tag{10}$$

$$WS_i = \begin{cases} \delta + \frac{2(L_i - y_i)}{\alpha} & \text{if } y_i < L_i \\ \delta & \text{if } L_i \le y_i \le U_i \\ \delta + \frac{2(y_i - U_i)}{\alpha} & \text{if } U_i < y_i \end{cases} \tag{11}$$

where L_i and U_i are lower and upper quantile values of the evaluated PI, N is the number of samples, C_i the coverage factor of the PI, y_i the measure value at time t, δ is the interval width given by $\delta = U_i - L_i$, y_i the true load demand at time step i.

2.4. Uncertainty evaluation and probabilistic distributed net load

2.4.1. Deterministic net load

In order to determine the uncertainty in the QPSLP and calculate the PNL, it is presented below how the predicted quantiles must be approximated to a CDF and PDF and how the PNL can be determined. To achieve this, we first defined the predicted deterministic net load (NL). The predicted NL is defined as the difference between the sum of all predicted electricity loads $y_{Load,i}$ and that of predicted distributed renewable generation $y_{Gen.,i}$, as expressed in Eq. (12). In our case the difference between the electricity load PSLP and PV generation PSLP.

$$NL = \sum_{i=1}^{N} y_{Load,i} - y_{Gen.,i} \tag{12}$$

Therefore, a positive NL indicates that additional power is drawn from the public grid. On the other hand, a negative NL implies there is a generation surplus i.e., grid feed-in. The NL is therefore essential for the optimization of flexible load operation and shows times at which the integration of flexible loads or storage is optimal in the case of e.g. own-consumption, or if it makes sense to schedule a flexible load or storage unit to at later time.

2.4.2. Empirical CDF and PDF from the QPSLP

In order to further apply the QPSLPs to determining the residual or net load or the total uncertainty within a distributed integrated energy system, the PDFs of each random variable are required. For this purpose, the empirical PDFs of the respective QPSLPs are first established on a time-intercept basis. Then, we determine the instantaneous empirical cumulative distribution function from over the predicted quantiles. For each time step t, the CDF is approximated by a PCHIP interpolation. This special case of spline interpolation allows to avoid the phenomenon of oscillating edges of an interval, which is also known as "Runge's phenomenon". Therefore, according to Fritsch and Butland (1984), a PCHIP using the Python based PCHIP interpolation function from SciPy (Virtanen et al., 2020) is used in this study. The empirical PDF is then determined from its first derivative, which is continuous without any jumps.

The condition for the interpolation is that the data are monotonically increasing, as the PCHIP interpolant $P(x)$ uses piecewise monotone cubic splines to compute new values of points and its associated derivative (Fritsch and Butland, 1984). For each sub-interval $x_k \le x \le x_{k+1}$, P(x) is a cubic hermite interpolation polynomial with a specific derivative at the interpolation points. It is given that the first derivative ($\frac{dP}{dx}$) is continuous. The advantages of a PCHIP interpolation compared to spline one is the lower effort and fewer oscillations if the data are not smooth and no overshoots occur. A spline interpolation can provide for a more accurate interpolation, and the second derivative is still continuous. A detailed description of the PCHIP algorithm and calculation of the one-sided three-point estimation of the slopes at the endpoints is described in Fritsch and Butland (1984), and Moler (2004).

In order to approximate a complete empirical CDF or PDF from QPSLP as exemplified in Fig. 4, with quantiles between 0.1 and 0.9 predicted as a basis, the quantile points 0.0 and 1.0 were estimated and added before the interpolation. The maximum and minimum occurring value of the final X days from the historic data store of the respective sectors are taken as an estimate. For example, if the CDF for the electrical load is to be determined for timepoint 12:00, it could present as follows. The measured electrical load profiles of the last 30 days for this time point are first selected. From this, the base load (minimum occurring power) is selected as a 0.0 (q = 0.0) quantile value and the peak load (maximum occurred value) is selected as a 1.0 (q = 1.0) quantile value for the considered time. The condition must be fulfilled i.e., the quantile values for q = 0.0 < q = 0.1 and q = 1.0 > q = 0.9. For highly variable or seasonally dependent load or generation profiles, care must be taken to ensure that the number of retrospective days is not too large.

2.4.3. Convolution

Assessing joint uncertainty within an integrated system or calculating the distributed net load necessitates the aggregation of random variables, such as electricity load and PV generation. For this the combination of the individual PDFs, a convolution (*) of the PDFs is necessary. To achieve this, the determination of joint PDFs becomes imperative. The PDFs $f_X, f_Y : R \to R+$ and their joint PDF f_{XY} of two independent continuous random variables X and Y can be determined as described in Eq. (13), as a convolution of f_X and f_Y. The joint PDF of X + Y, denotes as $f_{X+Y}(z)$, constitutes a continuous random variable in itself.

$$(f_X * f_Y)(z) = f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-x) f_Y(x) dx, z \in (-\infty, \infty) \tag{13}$$

For a non-parametric approach with discrete approximated PDF's, the discrete convolution is defined as Eq. (14):

$$(f_X * f_Y)(z) = \sum_{x=-\infty}^{\infty} f_X(z-x) f_Y(x), z \in (-\infty, \infty) \tag{14}$$

For the convolution of two approximated PDFs, discretized probability mass function (PMF) and the same symmetrical value range is used for the x-axis. The resulting PMFs can then be convoluted and subsequently multiplied by a discretization factor. The sum of the PMF's and the convolution must be equal to one. For the implementation of the convolution, the Python library SciPy (Virtanen et al., 2020) is used.

The expected value E(X) or μ of a discrete random variable X is given by Eq. (15):

$$\mu = E(X) = \sum f_X(x) x \tag{15}$$

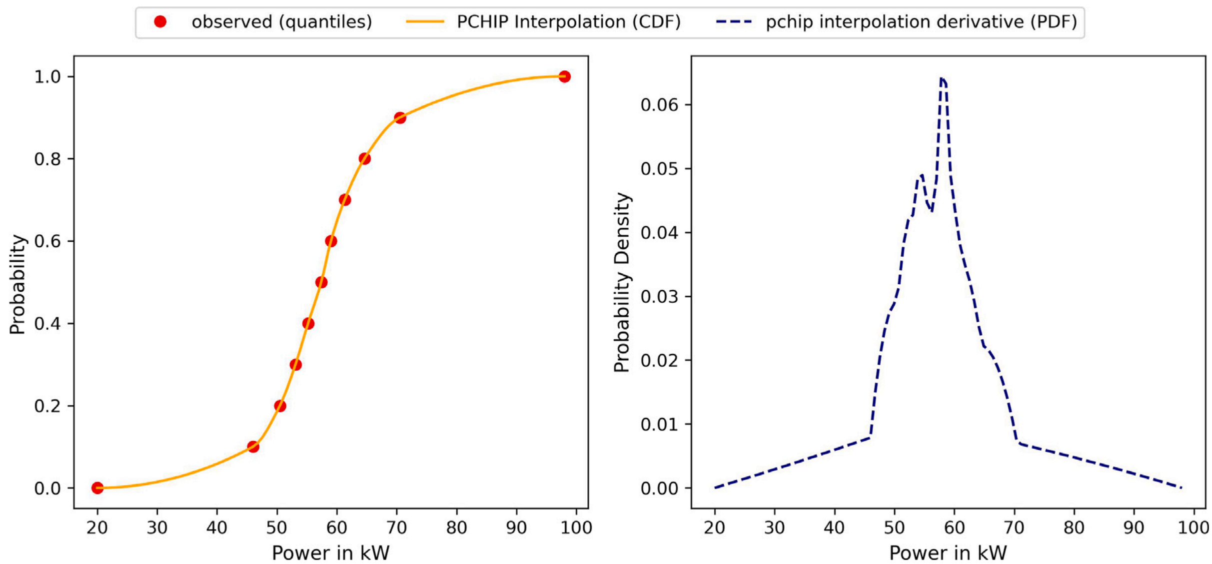

Fig. 4. Typical ECDF by PCHIP interpolation of the quantiles [0.1:0.9] for one time increment (left graph) and EPDF/derivative (right graph) for the electricity sector.

Table 6
Mean MAE and RMSE forecasting errors of the best PSLP.

Scores	Sector		
	Electricity	Heat	PV
MAE, kW	6.85	6.62	10.02
RMSE, kW	8.58	8.01	18.09

Table 7
Mean MAPE and MASE of the best PSLP forecasts.

Scores	Sector		
	Electricity	Heat	PV
MAPE, %	14.82	25.04	–
MASE	0.79	0.96	0.93

3. Forecasting framework results

3.1. PSLP forecasting evaluation

In the following section, the prediction results of the different PSLP modes and naive predictions are presented and compared. For this purpose, the forecast accuracy for the electricity, heat demand, and PV generation were evaluated using the metrics MAE, RMSE, MAPE, and MASE for the period from October 5_{th} 2021 to the end of September 2022. Before the first forecasts are made, a waiting period of 21 days was specified to serve as a basis for PSLP generation. Subsequently, the measured values of the previous day were taken into account in the forecast of the next day. In this manner, the data storage would constantly grow. This means that the forecasts were created in a rolling format for each day, using a training data set whose size depends on the mode of the PSLP (see also sub- Section 2.2.2). Thus, a real world scenario can be simulated with changing training data-sets and seasons.

The results of the PSLP predictions of each sector, mode, and daytype and/or season classification are shown in Fig. 5. The box marks the interquartile range (IQR) (25 % - 75 % quantile). The black line within the box marks the median (50 % quantile). The upper and lower whiskers extend from the box by a maximum of 1.5 times the IQR. From the boxplots in Fig. 5 and Table 6 it can be seen that all PSLP models outperform the benchmarks or the naive predictions. Looking at the PSLP prediction accuracy across different sectors (electricity, heat and PV) shows that the PSLPs performance depends on the different types of data profiles. The forecast error distributions in the boxplots of Fig. 5 show the differences between the individual modes (standard, fixed and variable) but also the influence of characterizing by day types and/or seasons. In the heat and PV forecasts, the larger IQRs are noticeable compared to the electricity sector. In addition, higher outliers occur and the smaller differences compared to the naïve forecasts become visible as pictured out in Fig. 5 and can be read from the MASE in Table 7. The performance of the different PSLP modes and the option to whether consider or not the day types (d) or seasons (s) within the different sectors is also remarkable. While the standard (std) PSLP performs best in the power sector, it is significantly worse in the heat and PV sectors. For these sectors, variable (var) mode of the PSLP demonstrates the best performance. Furthermore, the PSLP in the electricity sector is best when day type and season are taken into account, whereas PSLP is best for PV and heat sector when day or seasonal characteristics are not considered as in Fig. 5. Based on the MAPE of the best prediction results, the fundamental performance differences of the predictions between electricity and heat demand can be shown. Although the average MAPE in the electricity one is 14.56%, it is significantly higher for heat sector with 25.04%, as shown in Table 7.

For the probabilistic extension, the mean standard PSLP with day and season characteristics (std_el_d_s) was used for the electricity one, and the mean variable PSLP (var_ht, var_pv) was selected for the heat and PV sectors. The mean and median values of evaluation metrics of the selected PSLP for each sector are listed in comparison to the best performing naive predictions in Table 6. The selection can also be traced using the red markings in 5. Fig. 6 shows the PSLPs according to the selected PSLP modes of each sector and the associated DNL (Section 2.4.1) for an randomly selected day. Based on the prediction errors and example from Fig. 5, it is clear that the PSLP performs differently depending on the sector in which it is applied. Particularly for highly variable profiles like daily PV generation or heat demand, higher forecasting errors tend to occur than for the electricity load profile.

3.2. QPSLP forecasting evaluation

For the generation of the QPSLPs, as for the PSLP, a rolling forecast was used in which the training window was limited to 30, 90, or 120 days. All metrics were calculated for referral to the 80% PI (0.1–0.9 quantile). The probabilistic extension of the PSLP to a QPSLP is exemplified for the different sectors in Fig. 7. In Fig. 7, the gray solid

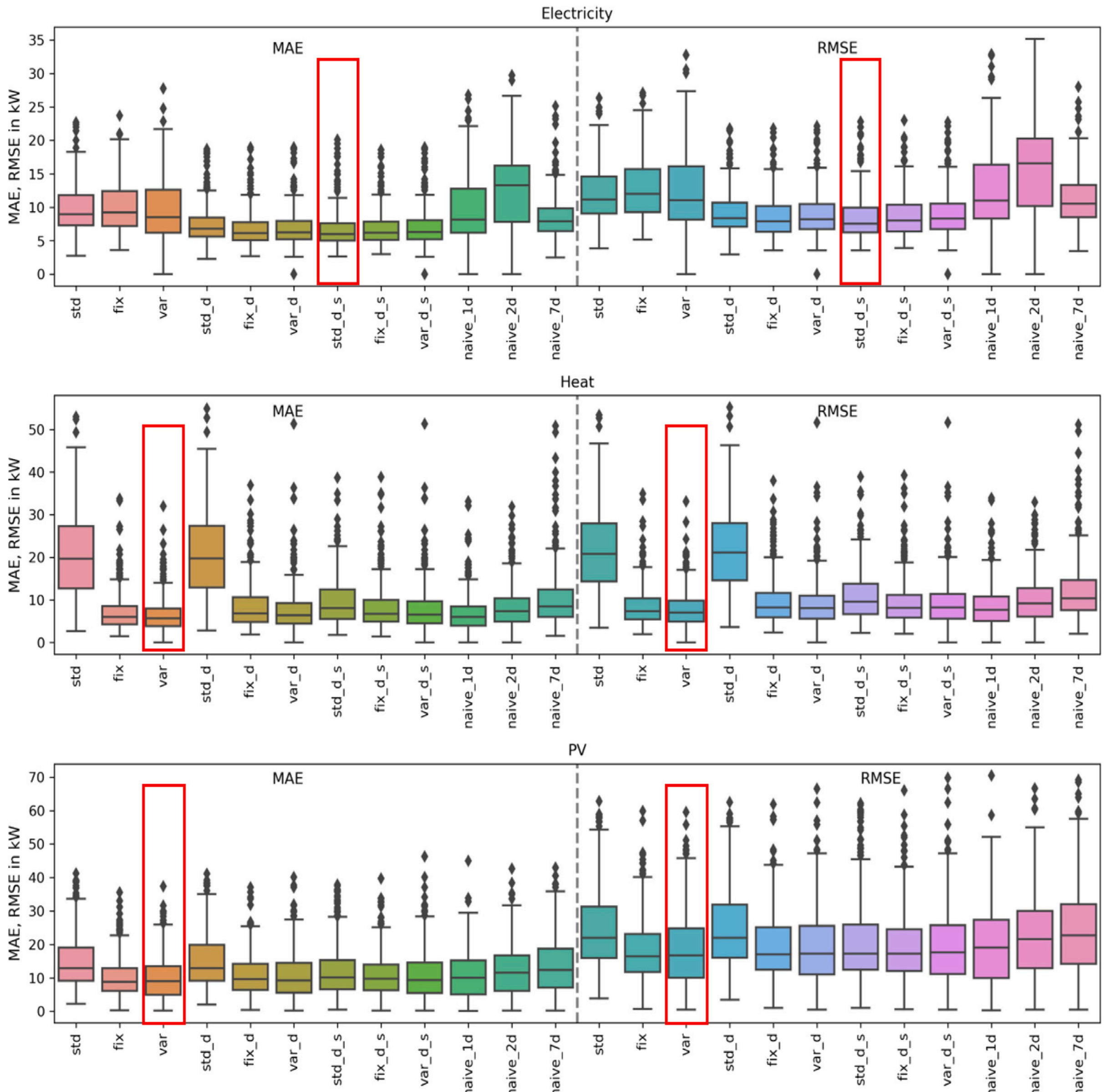

Fig. 5. Evaluation of the forecasting results by MAE and RMSE for the electricity load, comparing different PSLP modes and naive forecasts. The red marker shows the best average value of the respective metric overall PSLP modes.

line indicates quantiles 0.1 and 0.9, the dashed lines are the predicted quantiles (0.2, 0.3, 0.4, 0.6, 0.7, and 0.8), the red line indicates the median (0.5 quantile), the blue line is the PSLP point forecast, and the green dots depicts the actual measured values. The example day, as in Fig. 7, shows that the PI width can vary greatly, particularly during period of fluctuating PV generation, leading to large interval width.

The accuracy of the probabilistic forecasts were evaluated using the PICP, MPIW and Winkler score metrics, which are described in Section 2.3.2. The forecasting metrics of the QPSLP's for each sector are shown in Fig. 8 as boxplot and also in Table 8. The median (orange line) of the PICP, as well as the mean (see. Table 8) are around 80 %, which is to be expected with the PI of 80 %. However, there are very large differences in the range between the QPSLP in the electricity sector compared to the other sectors. This is especially reflected in the sharpness of the probabilistic prediction, as depicted by the MPIW metric. If we look at the mean values given in Table 8, the differences are also obvious and the predictions differ between the sectors. While the PICP, expected to be 80% for PI between 0.1 and 0.9 quantiles, remains at approximately the same level or higher, indicating the high reliability of QPSLP. There are significant differences in the MPIW and Winkler score across sectors. In particular, the highly variable PV generation is not predicted as efficiently compared to electricity and heat demand. The evaluation metrics over the entire observation period can be seen in the IQR (boxes) of the boxplots in Fig. 8. Again, the IQR of the MPIW and Winkler scores of the PV generation forecast is particularly striking, as it is significantly larger than in the other two cases. Next, to quantify the uncertainty, it is necessary to obtain a view on the CDF or PDF of the obtained predictions.

3.3. CDF and PDF approximation

The CDF and PDF from the QPSLPs quantiles were approximated using PCHIP interpolation and its derivatives, as described in Section 2.4.2. As an example, Fig. 9 shows the approximated instantaneous CDFs (yellow line) and PDFs (blue, dotted line) for electricity and heat demand and PV generation. As a plausibility check, the PDF's were discretized into PMFs and it was checked whether their sum was equal to one.

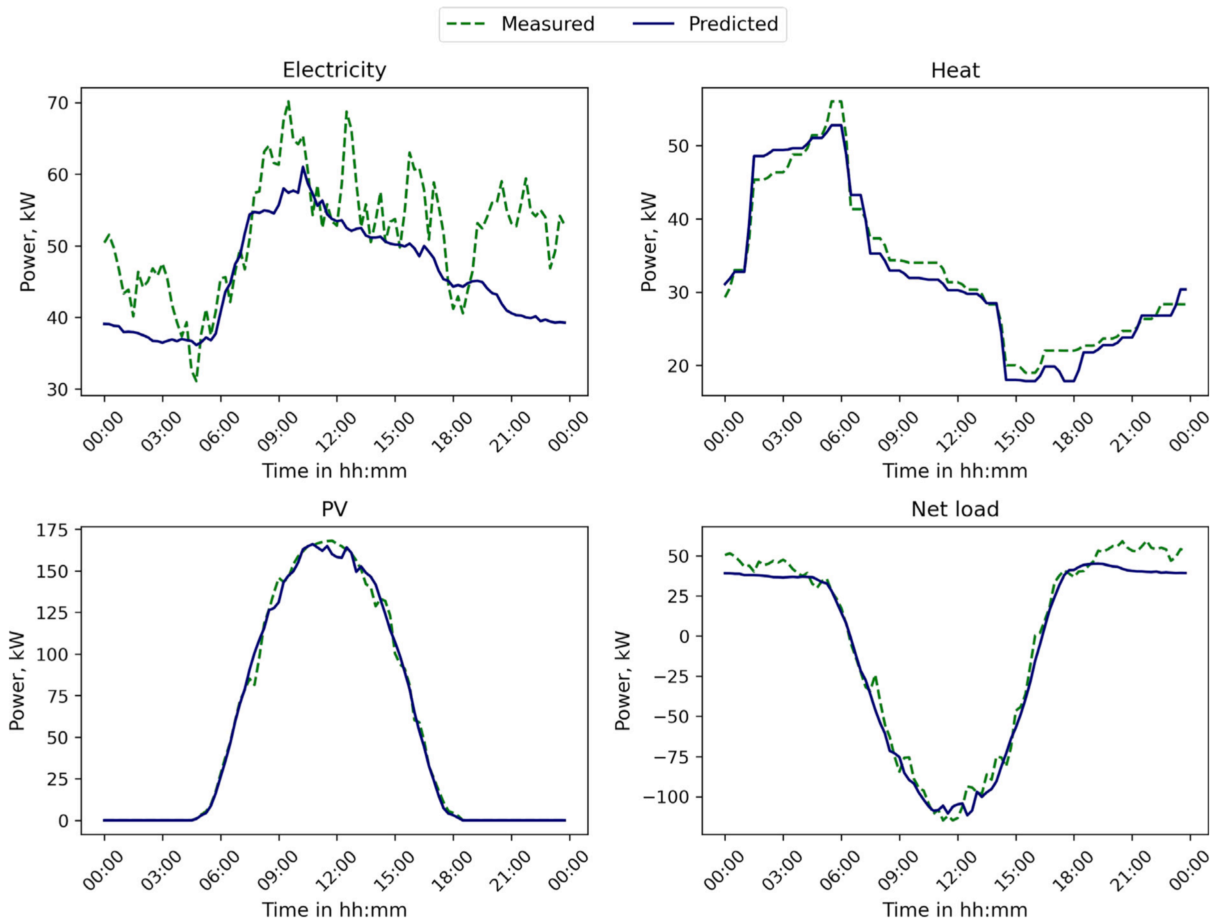

Fig. 6. Example day (April 20th 2022) of PSLP forecast for all sectors and the net load (electricity - PV).

Table 8
Mean QPSLP (QR) scores of the rolling forecast.

Sector/ Rolling days	Scores		
	PICP %	MPIW kW	Winkler Score
Electricity			
30 days	83.2	24.44	33.46
90 days	83.31	24.87	34.85
120 days	83.91	25.51	34.93
Heat			
30 days	83.91	25.51	34.93
90 days	84.54	26.84	36.25
120 days	84.62	27.01	36.5
PV			
30 days	80.57	41.23	59.3
90 days	86.34	43.81	59.98
120 days	88.17	44.55	60.39

3.4. Comparison of different distribution functions

Comparing the PCHIP-approximated CDFs with a fit of a Gaussian, beta or gamma distribution by the quantiles, the advantage of the proposed method becomes apparent, as is shown in Fig. 10. Quantiles fitting a CDF with Gaussian, beta, and gamma distributions were identified in 11 instances, which were also used for PCHIP interpolation in Section 3.3. In Fig. 10, the quantiles, the normal or Gaussian, the beta, the gamma distributions, and the pchip approximation for electricity, heat and PV sectors are shown in each subplot. It is apparent that these "standard" distributions sometimes work better and sometimes worse depending on the application (e.g., beta distribution works nearly good for PV generation, but fails for the other two applications). In comparison, the PCHIP approximation is clearly better adapted and has the advantage that no prior assumption about the distribution of the data must be known.

3.5. PNL forecasting evaluation

In order to add or subtract the probability densities of different predictions (random variables), they must be convoluted, as was described in Section 2.4.3. The results of such a convolution for a sample time point are displayed in Fig. 11. The top graph in Fig. 11 shows the approximated PDFs of the electricity and heat load forecasts, as well as PV generation (with negative powers). In the bottom graph, the convolution of the PDF's can be seen. The PNL (green-dashed line) results from the convolution between the power consumption minus PV generation. If one is interested in the joint PDF of several random variables (here, for example, electricity demand, heat demand, and PV generation), the convolution must be performed with all variables, as is shown in Fig. 11. The joint density function or PNL can then be used to determine the probability of occurrence for the different power ranges.

4. Case study

In application, the PNL can help making decisions in the operational optimization of distributed energy systems. It aids in decision-making regarding the optimal times for connecting an additional flexible load or energy storage (e.g., a BEV or electrolyzer), or when additional power can be provided by a flexible generation or storage unit. If the NL is positive, it indicates a power deficit, and additional power needs

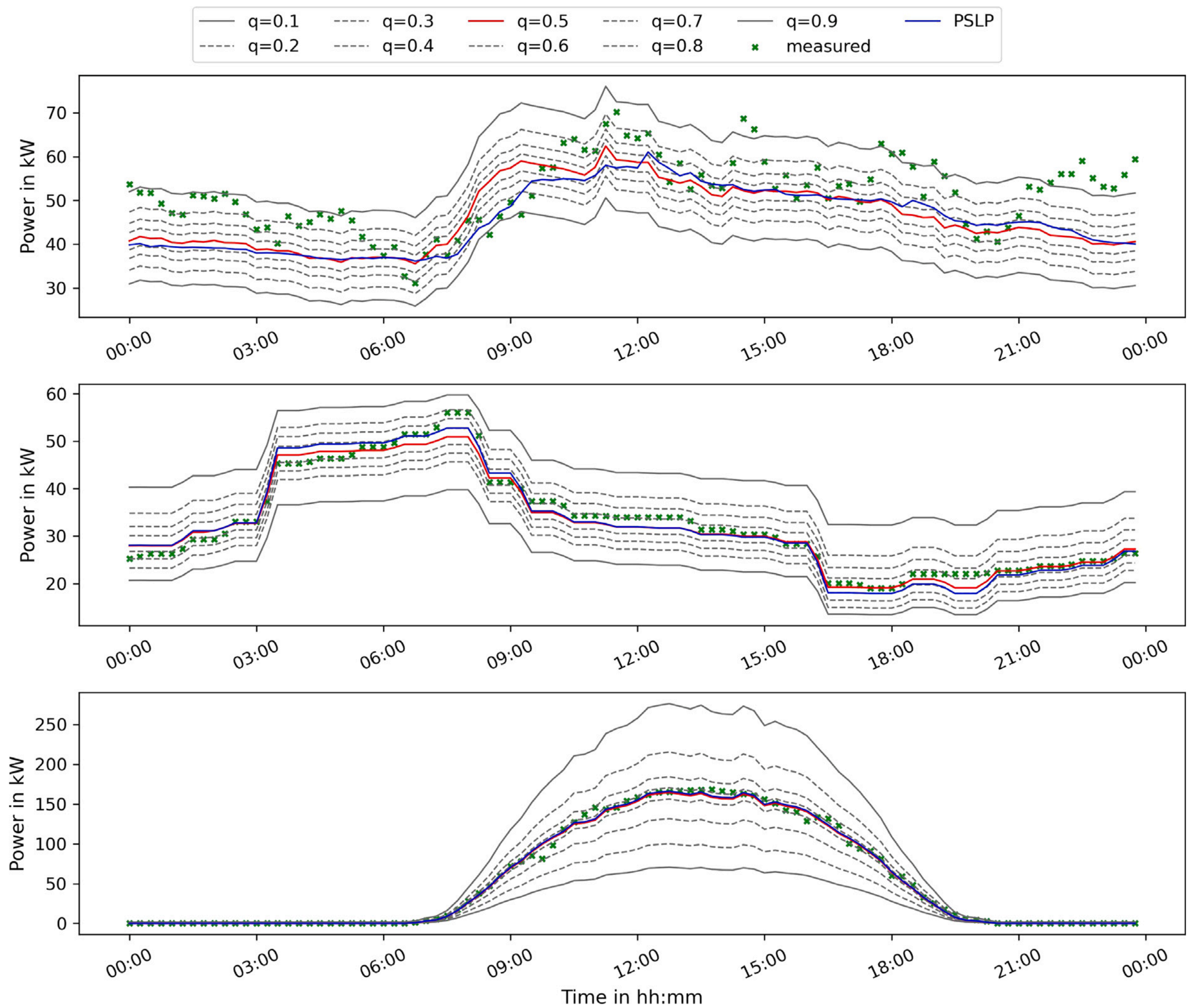

Fig. 7. Example day (April 20th 2022) for QPSLP of Electricity and heat load and PV-Generation using QR.

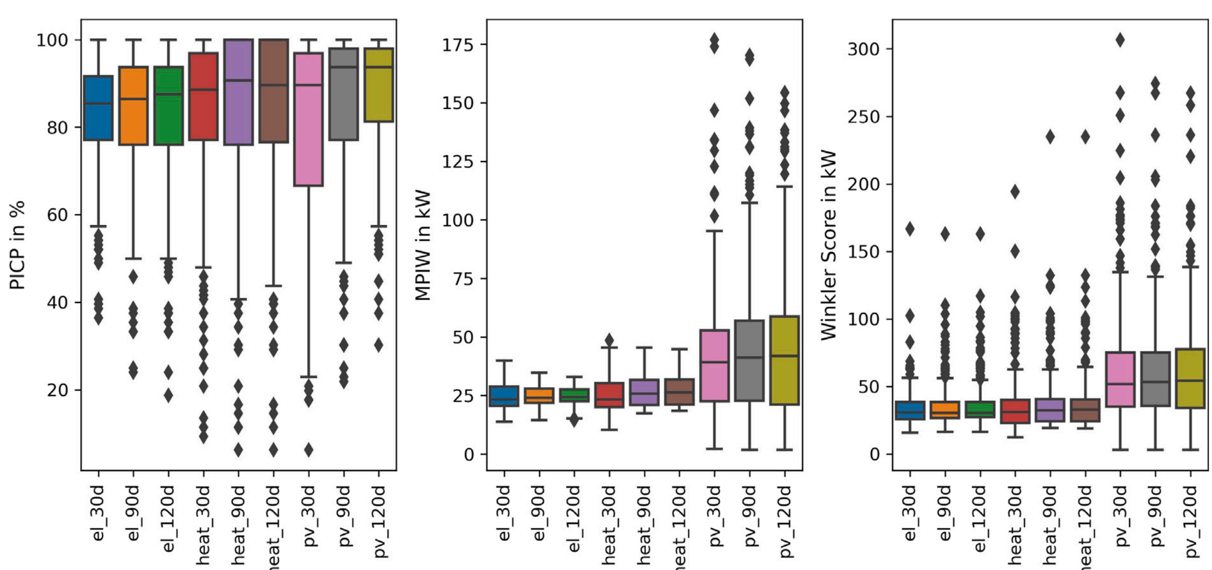

Fig. 8. Boxplots of probabilistic forecasting scores of the PICP, MPIW and Winkler score for electricity, heat and PV, with 30, 90 and 120 day training window, respectively.

to be purchased from the public electricity grid. On the other hand, if it is negative, it signifies a surplus of decentralized renewable power being generated. In that case, demand should be shifted or energy can be stored in energy storage technologies or feed into the public grid. The convolution of probability densities of stochastic supply sources and demands is necessary to aggregate uncertainties in different sectors

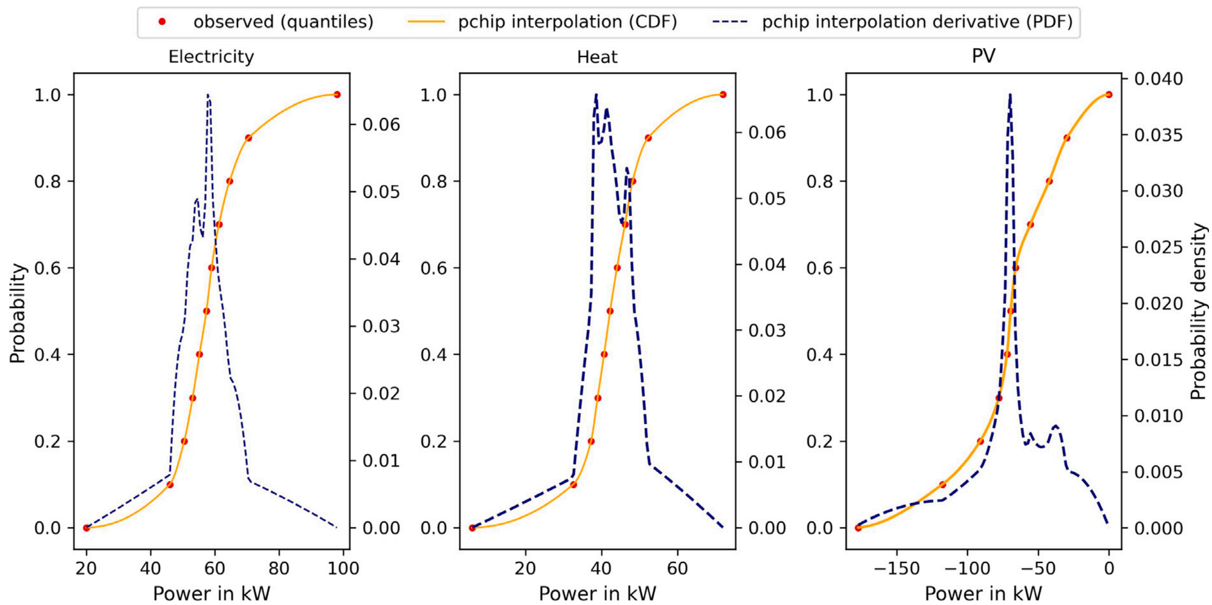

Fig. 9. PCHIP interpolation of the QPSLP quantiles to a CDF and PDF from its derivative for all sectors.

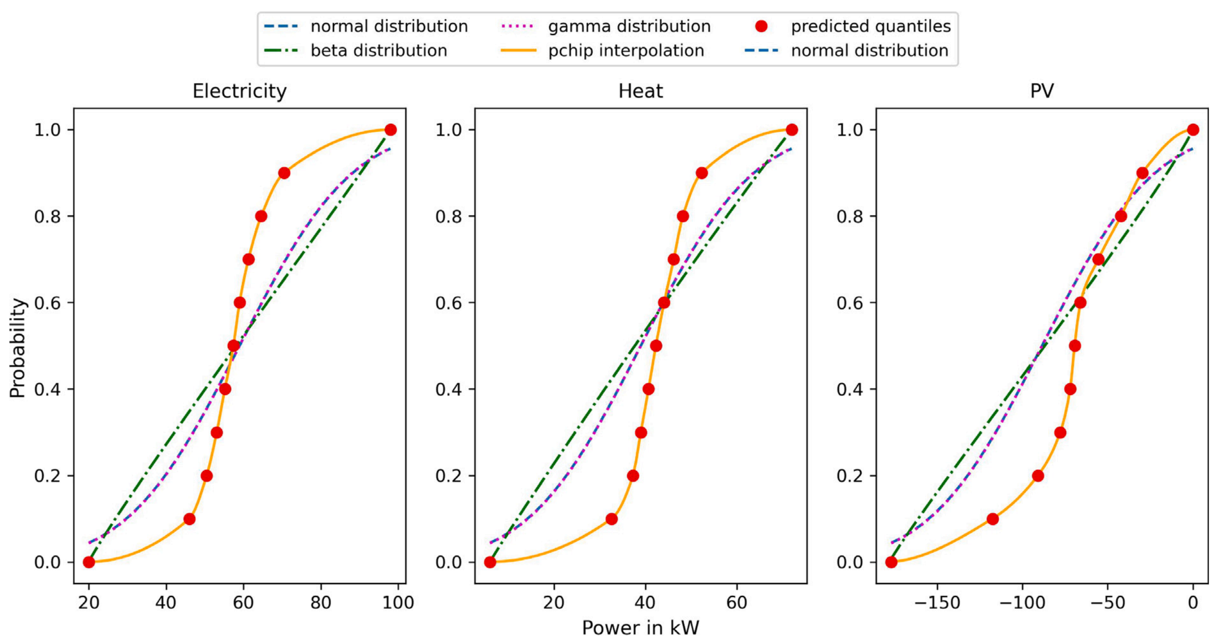

Fig. 10. Comparison of the PCHIP CDF approximation and the Gaussian, beta and gamma distribution fit to the quantiles.

and apply them in the form of a common PDF. In the following, a short case study is presented to demonstrate the advantages of the approach presented in this paper.

The case study consists of the distributed power system of a logistics property. In the first step, a probabilistic forecast and PNL of the power demand and PV generation for the subsequent 24 h is generated. The example of the logistics sector was chosen due to the facts that it is often time-critical, the transformation of the transport sector is proceeding rapidly, and solutions must be identified at a decentralized level. In this case, the flexible load is presented as an electrical preconditioning process, e.g. pre-cooling from a certain temperature to a temperature set-point of a refrigerated trailer (reefer). The refrigerated trailer stands at the logistics site for four hours and can only be electrically-preconditioned during this time. In the study (Telle et al., 2022), the authors described which cost advantage and reduced emissions the electrical process compared to the diesel preconditioning, especially when leveraging own-consumption with PV.

In the first example described in Fig. 13, a randomly selected day between 6:00 am and 10:00 am is considered. In the observation, the full hour is considered in each case. Start time = 06:00, t1 = 07:00, t2 = 08:00, t3 = 09:00, and the end time 10:00. Fig. 12 shows the respective PDFs of the NL for the respective times.

It is fixed as an assumption that own-consumption of PV-generated power incurs lower costs and CO_2-emissions compared to the supply from the public electricity grid. Therefore, the optimization goal is to maximize PV own-consumption whenever feasible, i.e., ≤ 0 kW (right vertical gray dashed line in Fig. 12). For the sake of simplicity, we assume the electrical pre-cooling takes one hour with an additional load of 10 kW. The left vertical gray dash-dotted line (≤ −10 kW) in Fig. 12 marks when it can be served solely from locally produced electricity.

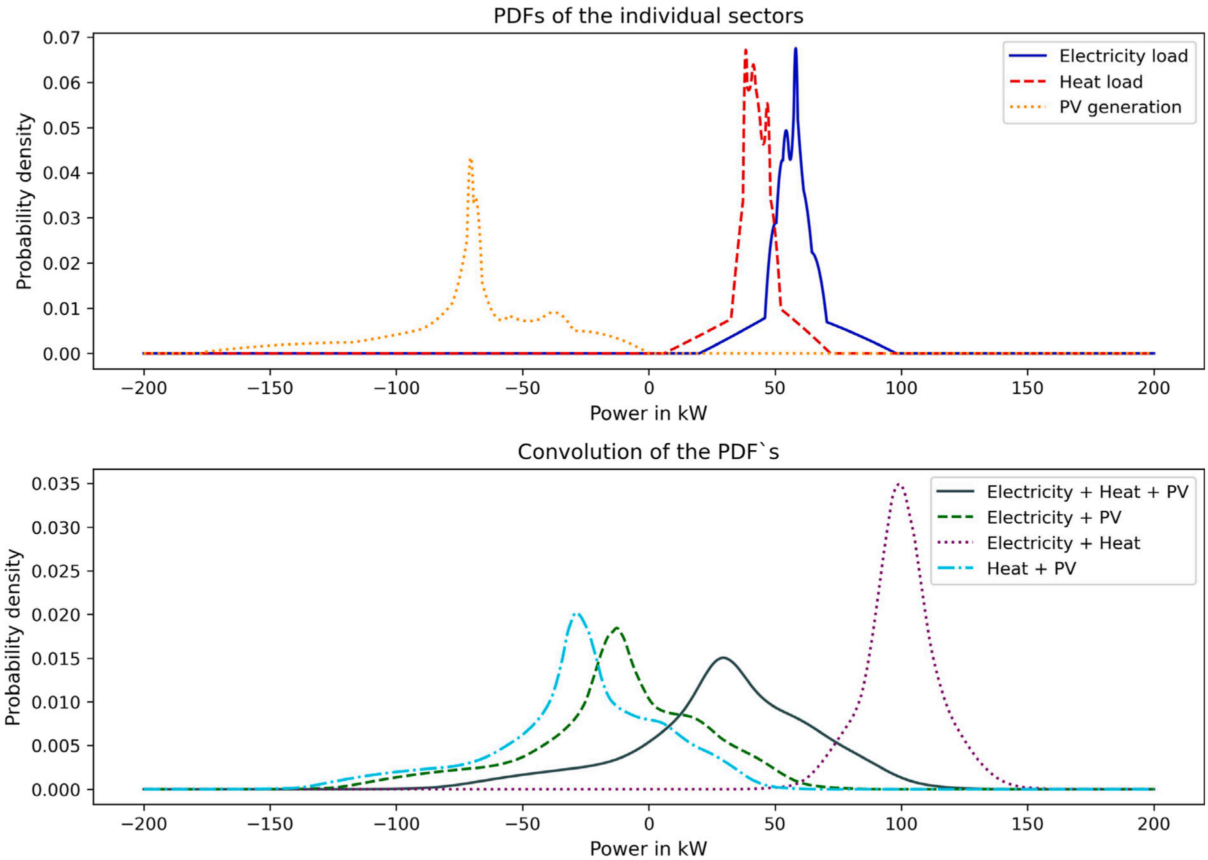

Fig. 11. Separate PDF of each sector (upper graph) and the joint PDF's after convolution (lower graph).

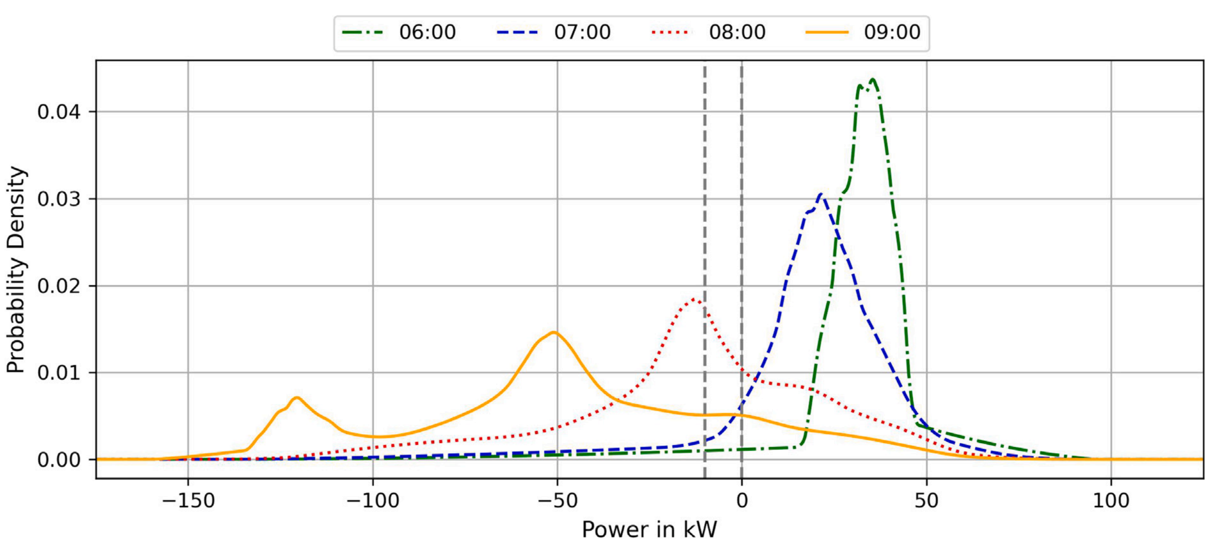

Fig. 12. PNL of different time points in the cast study (the left vertical line marks −10 kW, the right one 0 kW).

In order to make the decision under uncertainty, the probabilities to reaching the criteria are accumulated. Table 9 presents these values for both, the ≤ 0 kW (own-consumption) and ≤ −10 kW (+10 kW pre-cooling load) criteria. For comparison, it also displays the expected value μ for each time step.

Table 9
Accumulated probabilities and expected values of the case study criteria.

Timestep	p(P ≤ 0 kW)	p(P ≤ −10 kW)	μ in kW
06:00 – 07:00	5.69%	4.56%	30.81
07:00 – 08:00	11.86%	8.21%	17.92
08:00 – 09:00	65.39%	52.29%	−12.89
09:00 – 10:00	83.53%	83.05%	−47.67

In the first time step (starting 6:00), there is most likely no PV generation yet (see Fig. 12) and the PNL is therefore approximately 95 % ≤ 0 kW. By 10:00 am, PV power generation increases, as does probability of a negative NL,

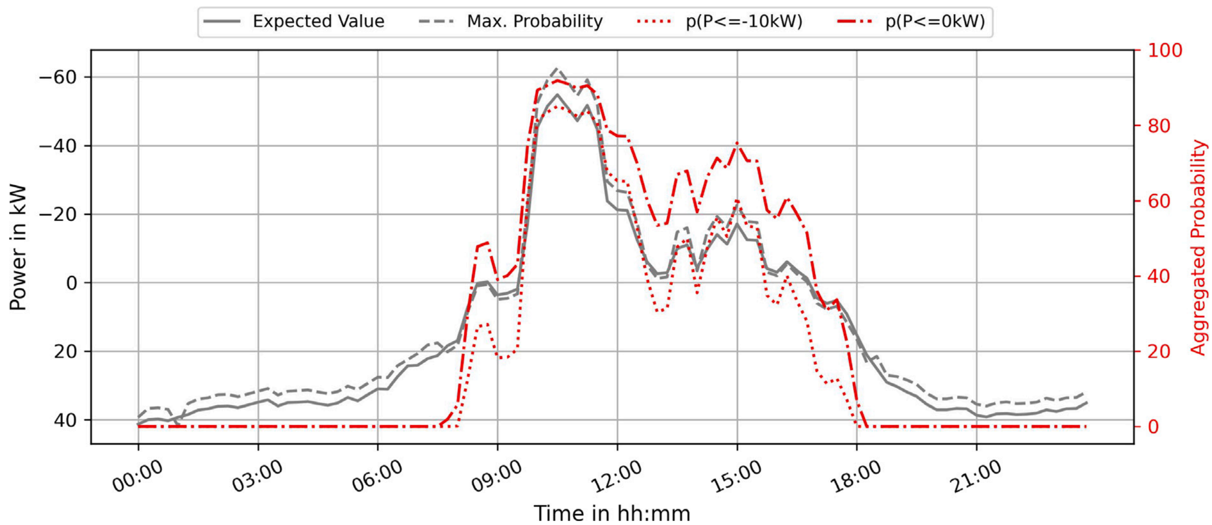

Fig. 13. Important characteristics values from the PNL, at another example day.

the expected values of the PNL shift towards negative values. At this point, the suggested method already helps to quantify the impact of a decision in the schedule. If the pre-cooling is carried out between 09:00 and 10:00, it will increase the probability to need electricity from the grid by 0.5%. However, in the presented period of time, judging based on the expected net load would yield the same decision, as the load also takes the lowest value in that hour.

Now, the example is extended to consider a full day. For this, the expected NL and the aggregated probabilities as used before are considered. In Fig. 13, the value with highest probability of occurrence is also added, as it might also serve as a criterion for decisions. It can be directly seen that a pre-cooling operation between 10:00 and 12:00 is beneficial to have a high probability of PV own-consumption, while having a strongly negative net load. Again, the aggregated probability (red dotted and dash-dotted lines in Fig. 13), the expected value (solid line), and the value of the maximum probability (dashed line) yield the same schedule. However, the situation becomes more interesting in cases where it is not possible to pick this optimal solution. For example, with a set departure time of 8:00, it would make no difference at which point in time the pre-cooling process is started. This fact is only visible as the probability of own-consumption stays at 0 %, while the load values already start shifting. At these times, it makes sense to consider another decision criterion.

5. Discussion

With the QPSLP and PNL, a method was presented that takes up the concept of the standard load profile and extends it to make it applicable to other sectors. It is designed to optimize sector-integrated distributed energy systems. This tool can be used to integrate renewable generation (e.g., PV) and new flexible loads (e.g., battery electric-vehicles) in distributed energy systems in a technically- and economically-optimized way. Therefore, this tool can help to decarbonize distributed energy systems in particular. The PNLFF, being a non-blackbox approach, is easy to implement and requires no more than locally collected measurements from various sectors for forecasting load demands, generations, and the net load. It can be used by a large number of users, particularly in small distributed energy systems, like households, neighborhoods, or SMEs, where high investment in a forecasting system may not be economically viable.

It was shown that the PSLP can also be applied to the heat and PV sectors. The PSLP performance in different sectors is very different, e.g., the average MAPE of electricity (14.82%) and heat (25.04%) load differs by more than 10%, as is shown in Table 6. A strong scattering of the errors can be observed, particularly in the PV prediction as shown in the box plots in Fig. 5. This can be interpreted in the higher stochastic within the data and large fluctuations between days. It can be noted that the PSLP performs well, particularly for the electricity sector compared to heat demand and PV generation. This is mainly due to the fact that the classification of electricity consumption by day of the week works much better. In contrast, PV generation, for example, is not based on days of the week. However, it was also shown that the addition of the “fix” and “variable” modes, as well as the option to turn off the characterizations by day type or season, brought about significant improvements in the PSLP (see Table 6 and Fig. 5) and outperformed the benchmarks. Individual hyper-parameters should be set for different energy systems and sectors, such as whether to classify by weekdays, workdays, or seasons, or which aggregation method is to be used. In the case of heat demand, the outdoor temperature, if temperature data are available, can be used to make the classification into different seasons, as described in Section 2.2.2, instead of a fixed date (day type characterization). However, this only carries an advantage if there is a correspondingly high correlation between outside temperature and heat demand. In the case of coupling between the heat and electricity sectors by means of heat pumps, for example, it could also become relevant for electricity demand. Furthermore, in the aggregation method, other options than the mean value, e.g., median or maximum value, could be tested. The maximum value could be successful, especially for PV, shown in Hanke et al. (2018). In addition, the transportation sector could be included as an outlook and the electric power demand of BEVs or charging stations (Boulakhbar et al., 2022) included in the PNLFF. In contrast to deterministic prediction, probabilistic prediction offers the advantage that uncertainty can be taken into account.

The probabilistic extension towards a QPSLP shows how to use PSLP as simply an input feature for QR in order to make a probabilistic forecast while achieving promising results. The QPSLP was evaluated for different training windows (30, 90, and 120 days). For the PICP, the expected value is 80 % with PI between 0.1 to 0.9 quantiles, as shown in Table 8. This is slightly overestimated in the mean value with 83.2 % (30 days) and 83.91 % (30 days) in the electricity and heat sectors, respectively, and is closest to the expected value, with 80.57 % in the PV sector. The PICP demonstrates that the PNLFF can be used to generate highly reliable probabilistic forecasts in all sectors with some outliers at the bottom (see Fig. 8). However, looking at the distribution of the other metrics (MPIW and Winkler score), it is clear that the magnitude and range of errors, as with the PSLP, are clearly the highest for PV (see Fig. 8). A short training window of 30 days performs best, but is less relevant in the electricity and heat sectors. The significant differences for PV are due to the changeable weather and the seasonal dependency of PV performance. This approach can be further optimized for PV in future work. However, there is an increased potential for optimization, especially in PV generation. For example, temperature time series or calendared features could be used to improve the QPSLP. It is also possible to see whether a parametric approach can be used, if the distribution of the load time series is known. The advantage of a non-parametric approach (QR) in our work is that no assumptions need to be made about the distribution of the data as shown in Section 3.4 and Fig. 10. We quantify this, and initially, it outweighs the disadvantages of higher computation times and the risk of quantile crossing, which was described by Das et al. (2019); a solution for QR without quantile crossing was presented in Anon (2022). However, the task of monitoring the results and taking counter measures against quantile-crossing would be essential in future work. In the study by van der Heijden et al. (2022), the approach involving quantile regression neural networks (QRNN) demonstrates a promising approach to address the issue of quantile crossing that could be investigated in future work. In order to be able to determine uncertainty within the QPSLPs, a method was presented to which the CDF and PDF from the quantiles of the QPSLPs (probabilistic forecasts) could be approximated. It was also shown how a joint PDF can then be determined from the convolution of two independent variables. For example, the PDF of the net load can be determined from convolution 2.4.3 of the electricity and PV PDFs. The aggregation to a common PDF is necessary to quantify the total uncertainty from different methods. If different forecasts are to be used in the application, for example to create operation schedules for an energy system or to make decisions under uncertainty, the joint PDF or PNL is required. The benefit of the PNLFF is that it evaluates the probabilistic predictions using conventional metrics, but also offers the possibility of approximating the PDF and creating a joint PDF or PNL for an energy system with several forecasts from different sectors. If the random variables are not independent of each other, then the correlation between their PDFs must be determined. According to Sklar's theorem (Van Vliet, 2023), Copulas can be used to consider the dependencies of two random variables in their joint distribution (Li et al., 2020). Convolution allows the joint probability density to be determined for different sectors at each timestamp, and to apply it in uncertainty quantification or risk optimization. Thus, the PNLFF can provide the basis for stochastic operational optimization as demonstrated by Kaisermayer et al. (2021). In contrast to the consideration of large-scale energy systems, rapid changes in load behavior or in the generation of decentralized energy generation are more critical for a distributed energy system. These rapid changes are almost impossible, costly or impossible to forecast. However, our approach offers the possibility of taking the uncertainty from the forecast errors into account in the operational optimization and scheduling of flexible units, and that with a low expenditure of data and computing time. Future work must show that this approach performs better in an stochastic optimization than deterministic optimization. Another advantage is that no assumption has to be made about the distribution of the data. It is important to highlight that depending on the time of day or season, there can be strong differences in the distribution of the data. In the last step of our work, a case study was used to show how the predictions or PNL can be used to optimize the integration or operation of additional electrical loads under uncertainty. The advantage of considering the probabilistic net load is obvious. If all electrical consumers and generators are taken into account, it can be used for decision-making under uncertainty. Thus, not only technical limitations (e.g. the house connection point) but also economic ecological factors can be optimized. As is shown in the example of the optimized integration of the electric refrigerated trailer, a decision can be made in the simplest way. Then, in a further step, operation schedules for decentralized sector-integrated energy systems can be generated in a stochastic optimization process for load, renewable generation, and storage. The limitations of the work and future work derived from it can be summarized as follows. The framework was tested on one dataset and should be tested on different datasets and different sectors in future work. The PSLP was used for the first time for heat load and PV generation and should be compared with other prediction methods in future work. A basic quantile regression model was used for probabilistic forecasting. An extension through the use of neural networks in quantile regression can help to reduce sources of error. The PCHIP interpolation of CDF and PDF was used for the first time for such an application. From a mathematical point of view, convolution is only possible if the variables to be convolved are independent of each other. In a real world scenario, a dependency must be assumed. In further work, as described above, copulas could be used to describe these dependencies. The transferability to other datasets and sectors must be verified in future work. The applicability of the approach was demonstrated over a limited period of time. Future work could show how scenarios could be sampled from the PNL CDF and PDF for e.g. a stochastic operational optimization of sector integrated energy systems.

6. Conclusions

In this study, we hypothesized that there is an increased need for forecast-based energy and load management in distributed energy systems to support integrated electrification in all sectors, as well as renewable generation. In the case of small-scaled energy systems, these should cost as little as possible, so as to be able to generate and be adapted on the basis of the system's own collected data (e.g., smart meters or PV inverters). Thereby, these can be usable for energy management by means of uncertainty assessment. For this purpose, a framework was developed that enables the creation of PSLPs for the electricity, heat, and PV sectors, and the generation of a non-parametric probabilistic forecast from them using QR, thus moving towards QPSLP. To this end, the question was answered on how PSLP and QPSLP can be transferred from the electricity sector to others (i.e., heat and PV generation). For further applications, an approach was then presented on how to derive the respective empirical CDF and PDF from the QPSLPs and thus by convolution of the PDF's of electricity, heat and PV generation forecasts to determine the probabilistic distributed net load. Based on the PNL, a case study for the use of the PNLFF in energy management was ultimately shown. The study can be understood as a proof-of-concept in which it was demonstrated that it is possible to predict highly fluctuating consumers and generators with less computational effort and to aggregate them into a probabilistic decentralized net load with just locally collected measured data. This in turn can be used in the application to make decisions in operational optimization under uncertainty, which was shown in an application example. As additional novelty, it was shown that the PSLP with the proposed adjustments can be transferred to the heating sector and PV generation. In addition, it was shown for the first time that the PCHIP interpolation can be used to interpolate the non-parametric probabilistic forecast in the form of quantiles to a CDF. The convolution made it possible to create a joint distribution that takes sector coupling into account and has not yet been presented in this form in distributed sector-integrated energy system. In the first section of the paper, we focused on the "related work" of previous developments, advantages and disadvantages in the field of standard load profiles, and personalized standard load profiles. In addition, probabilistic forecasts or extensions with a focus on a simple QR and the evaluation of uncertainty are highlighted. In the second section, the methodology that describes an overview of the developed PNLFF and the data basis used were presented. Based on this, deterministic forecasting methods were first presented. As a benchmark, a naïve forecast model built on the data from previous one, two, and seven days were used. For the PSLP, different options were presented that can be applied depending on the sector and data. As a novelty, the PSLP was transferred to the heat and PV generation sectors and options for switching on and off day type characterization and seasonal consideration were introduced. The PSLP was then used as an explanatory variable and input into a QR to generate a probabilistic forecast. In order to determine the net load of the QPSLPs and to be able to optimize own-consumption, a new approach for determining the empirical CDF and PDF by means of a so-called PCHIP interpolation was introduced. The PNL was then calculated via convolution from the individual PDF's. Finally, a case study was used to demonstrate the use of the PNL, which shows how the results of the framework could be used to optimize the operation of the distributed power system under uncertainty. The PNLFF considers energy systems as a holistic, cross-sectoral system, requiring just locally-collected measurement data. It includes the characterization of the load and generation data (PSLP), the generation of probabilistic forecasts (QPSLP), the approximation of empirical density distributions (CDF and PDF), and the calculation of net load (PNL) as a basis for optimization under uncertainty. Due to its simplicity, there will be proprietary methods and tools that can produce a more accurate prediction with increased resources, but it is holistic, widely applicable, and transferable, making it an important contribution to the development for cleaner energy systems.

CRediT authorship contribution statement

Jan-Simon Telle: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Ajay Upadhaya:** Data curation, Formal analysis, Software, Writing – review & editing. **Patrik Schönfeldt:** Conceptualization, Formal analysis, Methodology, Writing – review & editing. **Thomas Steens:** Data curation, Methodology, Software. **Benedikt Hanke:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **Karsten von Maydell:** Funding acquisition, Investigation, Supervision.

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.

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors would like to acknowledge the financial support of the "Federal Ministry for Economic Affairs and Climate Action" of the Federal Republic of Germany for the project "ELogZ: Energieversorgungskonzepte für klimaneutrale Logistikzentren" (03EN1015F). For more details, visit www.elogz.de The presented study was conducted as part of this project. A special thanks is due to Paneuropa Transport GmbH for the provision of data and know how.

References

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U.R., Makarenkov, V., Nahavandi, S., 2021. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Inf. Fusion 76, 243–297. http://dx.doi.org/10.1016/j.inffus.2021.05.008, URL https://www.sciencedirect.com/science/article/pii/S1566253521001081.

Anon, 2019-2024. Energieversorgungskonzepte für klimaneutrale Logistikzentren. URL https://elogz.de/.

Anon, 2022. Proceedings of the 25th international conference on artificial intelligence and statistics. In: Camps-Valls, G., Ruiz, F.J.R., Valera, I. (Eds.), In: Proceedings of Machine Learning Research, PMLR.

Anon, 2023. Deutscher Wetterdienst - Open Data. URL https://opendata.dwd.de/, Free provision of spatial data of the DWD via the DWD's Open Data Server.

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., de Pison, F.M., Antonanzas-Torres, F., 2016a. Review of photovoltaic power forecasting. Sol. Energy 136, 78–111. http://dx.doi.org/10.1016/j.solener.2016.06.069.

Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de Pison, F.J., Antonanzas-Torres, F., 2016b. Review of photovoltaic power forecasting. Sol. Energy 136, 78–111. http://dx.doi.org/10.1016/j.solener.2016.06.069.

Ariffin, S., Karim, A., 2014. Cubic spline interpolation for petroleum engineering data. Appl. Math. Sci. 8, http://dx.doi.org/10.12988/ams.2014.44284.

Association, E., et al., 1997. Load profiles and their use in electricity settlement. ukedc, UKEDC, URL https://data.ukedc.rl.ac.uk/browse/edc/efficiency/residential/LoadProfile/Load_Profiles.pdf.

Badings, T., Simão, T.D., Suilen, M., Jansen, N., 2023. Decision-making under uncertainty: Beyond probabilities. Int. J. Softw. Tools Technol. Transf. http://dx.doi.org/10.1007/s10009-023-00704-3.

Barker, P.M., McDougall, J., 2020. Two interpolation methods using multiply-rotated piecewise cubic Hermite interpolating polynomials. J. Atmos. Ocean. Technol. http://dx.doi.org/10.1175/JTECH-D-19-0211.1.

Baumann, C., Kepplinger, P., 2023. Application of a flexibility estimation method for domestic heat pumps with reduced system information and data. Clean. Energy Syst. 6, 100081. http://dx.doi.org/10.1016/j.cles.2023.100081, URL https://www.sciencedirect.com/science/article/pii/S2772783123000316.

Beichter, M., Phipps, K., Frysztacki, M.M., Mikut, R., Hagenmeyer, V., Ludwig, N., 2022. Net load forecasting using different aggregation levels. Energy Inform. 5 (1), 19. http://dx.doi.org/10.1186/s42162-022-00213-8.

Benti, N.E., Chaka, M.D., Semie, A.G., 2023. Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects. Sustainability 15 (9), 7087.

Bergsteinsson, H.G., Kloppenborg Mø ller, J., Nystrup, P., Pálsson, Ó.P., Guericke, D., Henrik, M., 2021. Heat load forecasting using adaptive temporal hierarchies. Appl. Energy 292, http://dx.doi.org/10.1016/j.apenergy.2021.116872.

Bjerregård, M.B., Møller, J.K., Madsen, H., 2021. An introduction to multivariate probabilistic forecast evaluation. Energy AI 4, 100058. http://dx.doi.org/10.1016/j.egyai.2021.100058, URL https://www.sciencedirect.com/science/article/pii/S2666546821000124.

Boulakhbar, M., Farag, M., Benabdelaziz, K., Kousksou, T., Zazi, M., 2022. A deep learning approach for prediction of electrical vehicle charging stations power demand in regulated electricity markets: The case of Morocco. Clean. Energy Syst. 3, 100039. http://dx.doi.org/10.1016/j.cles.2022.100039, URL https://www.sciencedirect.com/science/article/pii/S2772783122000371.

Browell, J., Fasiolo, M., 2021. Probabilistic forecasting of regional net-load with conditional extremes and gridded NWP. IEEE Trans. Smart Grid 12 (6), 5011–5019. http://dx.doi.org/10.1109/TSG.2021.3107159.

Cleophas, T.J., Zwinderman, A.H., 2021. Quantile regression. In: Regression Analysis in Medical Research: For Starters and 2nd Levelers. Springer International Publishing, Cham, pp. 453–467. http://dx.doi.org/10.1007/978-3-030-61394-5_27.

Dang, S., Peng, L., Zhao, J., Li, J., Kong, Z., 2022. A quantile regression random forest-based short-term load probabilistic forecasting method. Energies 15 (2), 663.

Das, K., Krzywinski, M., Altman, N., 2019. Quantile regression. Nature Methods 16 (6), 451–452. http://dx.doi.org/10.1038/s41592-019-0406-y.

Das, U.K., Tey, K.S., Seyedmahmoudian, M., Mekhilef, S., Idris, M.Y.I., Deventer, W.V., Horan, B., Stojcevski, A., 2018. Forecasting of photovoltaic power generation and model optimization: A review. Renew. Sustain. Energy Rev. 81, 912–928. http://dx.doi.org/10.1016/j.rser.2017.08.017.

do Amaral Burghi, A.C., Hirsch, T., Pitz-Paal, R., 2020. Artificial learning dispatch planning with probabilistic forecasts: Using uncertainties as an asset. Energies 13 (3), http://dx.doi.org/10.3390/en13030616, URL https://www.mdpi.com/1996-1073/13/3/616.

Dong, W., Sun, H., Mei, C., Li, Z., Zhang, J., Yang, H., 2023. Forecast-driven stochastic optimization scheduling of an energy management system for an isolated hydrogen microgrid. Energy Convers. Manage. 277, 116640. http://dx.doi.org/10.1016/j.enconman.2022.116640, URL https://www.sciencedirect.com/science/article/pii/S0196890422014182.

Eren, Y., Küçükdemiral, İ., 2024. A comprehensive review on deep learning approaches for short-term load forecasting. Renew. Sustain. Energy Rev. 189, 114031.

European Comission, 2018. 2050 Long-term strategy. URL https://climate.ec.europa.eu/eu-action/climate-strategies-targets/2050-long-term-strategy_en.

European Comission, 2019. The European Green Deal. Tech. Rep., European Comission.

European Commission, 2020. Powering a climate-neutral economy: An EU strategy for energy system integration. European Commission, European Commission Brussels, Belgium, URL https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2020:299:FIN.

Falces, A., Capellan-Villacian, C., Mendoza-Villena, M., Zorzano-Santamaria, P.J., Lara-Santillan, P.M., Garcia-Garrido, E., Fernandez-Jimenez, L.A., Zorzano-Alba, E., 2023. Short-term net load forecast in distribution networks with PV penetration behind the meter. Energy Rep. 9, 115–122. http://dx.doi.org/10.1016/j.egyr.2022.12.103, URL https://www.sciencedirect.com/science/article/pii/S2352484722027044.

Fatema, I., Lei, G., Kong, X., 2023. Probabilistic forecasting of electricity demand incorporating mobility data. Appl. Sci. 13 (11), http://dx.doi.org/10.3390/app13116520, URL https://www.mdpi.com/2076-3417/13/11/6520.

Fridgen, G., Keller, R., Körner, M.-F., Schöpf, M., 2020. A holistic view on sector coupling. Energy Policy 147, 111913. http://dx.doi.org/10.1016/j.enpol.2020.111913.

Fritsch, F.N., Butland, J., 1984. A method for constructing local monotone piecewise cubic interpolants. SIAM J. Sci. Stat. Comput. 5 (2), 300–304. http://dx.doi.org/10.1137/0905021.

Gaamouche, R., Chinnici, M., Lahby, M., Abakarim, Y., Hasnaoui, A.E., 2022. Machine learning techniques for renewable energy forecasting: A comprehensive review. In: Computational Intelligence Techniques for Green Smart Cities. Springer, pp. 3–39.

González Ordiano, J.Á., Gröll, L., Mikut, R., Hagenmeyer, V., 2020. Probabilistic energy forecasting using the nearest neighbors quantile filter and quantile regression. Int. J. Forecast. 36 (2), 310–323. http://dx.doi.org/10.1016/j.ijforecast.2019.06.003, URL https://www.sciencedirect.com/science/article/pii/S0169207019301621.

Hanke, B., Bottega, M., Peters, D., Maitanova, N., Telle, J.-S., Grottke, M., von Maydell, K., Agert, C., 2018. Fully automated photovoltaic system modelling for low cost energy management applications based on power measurement data. In: 35th European Photovoltaic Solar Energy Conference and Exhibition. pp. 1588–1593. http://dx.doi.org/10.4229/35thEUPVSEC20182018-6BO.8.2.

Hinterstocker, M., von Roon, S., Rau, M., 2014. Bewertung der aktuellen Standardlastprofile Österreichs und Analyse zukünftiger Anpassungsmöglichkeiten im Strommarkt. In: Symposium Energieinnovationen, Graz. pp. 1–7.

Holmgren, W.F., W. Hansen, C., A. Mikofski, M., 2018. pvlib Python: A Python package for modeling solar energy systems. J. Open Source Softw. 3 (29), 884. http://dx.doi.org/10.21105/joss.00884.

Hong, T., Fan, S., 2016. Probabilistic electric load forecasting: A tutorial review. Int. J. Forecast. 32 (3), 914–938. http://dx.doi.org/10.1016/j.ijforecast.2015.11.011.

Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., Hyndman, R.J., 2016. Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. Int. J. Forecast. 32 (3), 896–913.

Hyndman, R.J., Athanasopoulos, G., 2018. Forecasting: Principles and Practice. OTexts.

Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. Int. J. Forecast. 22 (4), 679–688.

Kaisermayer, V., Muschick, D., Horn, M., Gölles, M., 2021. Progressive hedging for stochastic energy management systems. Energy Syst. 12 (1), 1–29. http://dx.doi.org/10.1007/s12667-020-00401-z.

Khosravi, A., Nahavandi, S., Creighton, D., Atiya, A.F., 2010. Lower upper bound estimation method for construction of neural network-based prediction intervals. IEEE Transa. Neural Netw. 22 (3), 337–346.

Klyuev, R.V., Morgoev, I.D., Morgoeva, A.D., Gavrina, O.A., Martyushev, N.V., Efremenkov, E.A., Mengxu, Q., 2022. Methods of forecasting electric energy consumption: A literature review. Energies 15, http://dx.doi.org/10.3390/en15238919.

Koenker, R., Bassett, G., 1978. Regression quantiles. Econometrica 46 (1), 33. http://dx.doi.org/10.2307/1913643.

Leiprecht, S., Beherens, F., Faber, T., Finkenrath, M., 2021. A comprehensive thermal load forecasting analysis based on machine learning algorithms. Energy Reportsy 7, 319–326. http://dx.doi.org/10.1016/j.egyr.2021.08.140.

Li, B., Zhang, J., Hobbs, B.F., 2020. A copula enhanced convolution for uncertainty aggregation. In: 2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference. ISGT, pp. 1–5. http://dx.doi.org/10.1109/ISGT45199.2020.9087644.

Liu, B., Nowotarski, J., Hong, T., Weron, R., 2015. Probabilistic load forecasting via quantile regression averaging on sister forecasts. IEEE Trans. Smart Grid 8 (2), 730–737.

Liu, B., Nowotarski, J., Hong, T., Weron, R., 2017. Probabilistic load forecasting via quantile regression averaging on sister forecasts. IEEE Trans. Smart Grid 8 (2), 730–737. http://dx.doi.org/10.1109/TSG.2015.2437877.

Maitanova, N., Telle, J.-S., Hanke, B., Grottke, M., Schmidt, T., Maydell, K.v., Agert, C., 2020. A machine learning approach to low-cost photovoltaic power prediction based on publicly available weather reports. Energies 13 (3), 735.

Mawson, V.J., Hughes, B.R., 2020. Deep learning techniques for energy forecasting and condition monitoring in the manufacturing sector. Energy Build. 217, 109966.

Mei, F., Gu, J., Lu, J., Lu, J., Zhang, J., Jiang, Y., Shi, T., Zheng, J., 2020. Day-ahead nonparametric probabilistic forecasting of photovoltaic power generation based on the LSTM-QRA ensemble model. IEEE Access 8, 166138–166149.

Ming, M., Wang, R., Zha, Y., Zhang, T., 2017. Multi-objective optimization of hybrid renewable energy system with load forecasting. In: 2017 IEEE International Conference on Energy Internet. ICEI, IEEE, pp. 113–118.

Moler, C.B., 2004. Numerical Computing with Matlab. Society for Industrial and Applied Mathematics, http://dx.doi.org/10.1137/1.9780898717952, arXiv:https://epubs.siam.org/doi/pdf/10.1137/1.9780898717952, URL https://epubs.siam.org/doi/abs/10.1137/1.9780898717952.

Nti, I.K., Teimeh, M., Nyarko-Boateng, O., Adekoya, A.F., 2020. Electricity load forecasting: A systematic review. J. Electr. Syst. Inf. Technol. 13, 19. http://dx.doi.org/10.1186/s43067-020-00021-8.

Pavičević, M., Mangipinto, A., Nijs, W., Lombardi, F., Kavvadias, K., Jiménez Navarro, J.P., Colombo, E., Quoilin, S., 2020. The potential of sector coupling in future European energy systems: Soft linking between the Dispa-SET and JRC-EU-TIMES models. Appl. Energy 267, 115100. http://dx.doi.org/10.1016/j.apenergy.2020.115100, URL https://www.sciencedirect.com/science/article/pii/S0306261920306127.

Peters, D., Völker, R., Schuldt, F., von Maydell, K., 2020. Are standard load profiles suitable for modern electricity grid models? In: 2020 17th International Conference on the European Energy Market. EEM, pp. 1–6. http://dx.doi.org/10.1109/EEM49802.2020.9221967.

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M.Z., Barrow, D.K., Ben Taieb, S., Bergmeir, C., Bessa, R.J., Bijak, J., Boylan, J.E., Browell, J., Carnevale, C., Castle, J.L., Cirillo, P., Clements, M.P., Cordeiro, C., Cyrino Oliveira, F.L., De Baets, S., Dokumentov, A., Ellison, J., Fiszeder, P., Franses, P.H., Frazier, D.T., Gilliland, M., Gönül, M.S., Goodwin, P., Grossi, L., Grushka-Cockayne, Y., Guidolin, M., Guidolin, M., Gunter, U., Guo, X., Guseo, R., Harvey, N., Hendry, D.F., Hollyman, R., Januschowski, T., Jeon, J., Jose, V.R.R., Kang, Y., Koehler, A.B., Kolassa, S., Kourentzes, N., Leva, S., Li, F., Litsiou, K., Makridakis, S., Martin, G.M., Martinez, A.B., Meeran, S., Modis, T., Nikolopoulos, K., Önkal, D., Paccagnini, A., Panagiotelis, A., Panapakidis, I., Pavía, J.M., Pedio, M., Pedregal, D.J., Pinson, P., Ramos, P., Rapach, D.E., Reade, J.J., Rostami-Tabar, B., Rubaszek, M., Sermpinis, G., Shang, H.L., Spiliotis, E., Syntetos, A.A., Talagala, P.D., Talagala, T.S., Tashman, L., Thomakos, D., Thorarinsdottir, T., Todini, E., Trapero Arenas, J.R., Wang, X., Winkl, R.L., 2022. Forecasting: Theory and practice. Int. J. Forecast. 38 (3), 705–871. http://dx.doi.org/10.1016/j.ijforecast.2021.11.001, URL https://www.sciencedirect.com/science/article/pii/S0169207021001758.

Scholz, D., Müsgens, F., 2017. How to improve standard load profiles: Updating, regionalization and smart meter data. In: 2017 14th International Conference on the European Energy Market. EEM, pp. 1–6. http://dx.doi.org/10.1109/EEM.2017.7981939.

Seabold, S., Perktold, J., 2010. statsmodels: Econometric and statistical modeling with Python. In: 9th Python in Science Conference.

Steens, T., Telle, J.-S., Hanke, B., von Maydell, K., Agert, C., Di Modica, G.-L., Engel, B., Grottke, M., 2021. A forecast-based load management approach for commercial buildings demonstrated on an integration of BEV. Energies 14 (12), 3576.

Telle, J.-S., Maitanova, N., Steens, T., Hanke, B., von Maydell, K., Grottke, M., 2020. Combined PV power and load prediction for building-level energy management applications. In: 2020 Fifteenth International Conference on Ecological Vehicles and Renewable Energies. EVER, pp. 1–15. http://dx.doi.org/10.1109/EVER48776.2020.9243026.

Telle, J.-S., Schlüters, S., Schönfeldt, P., Hanke, B., von Maydell, K., Agert, C., 2022. The optimized integration of temperature-controlled transports into distributed sector-integrated energy systems. Energy Convers. Manage. 269, 116148. http://dx.doi.org/10.1016/j.enconman.2022.116148.

This makes sector coupling an important building block towards a climate-neutral energy system in 2050, 2022. Sector Coupling: A Key Concept for Accelerating the Energy Transformation. Tech. Rep., International Renewable Energy Agency, Abu Dhabi.

van der Heijden, T., Palensky, P., van de Giesen, N., Abraham, E., 2022. Day ahead market price scenario generation using a combined quantile regression deep neural network and a non-parametric Bayesian network. In: 2022 IEEE International Conference on Power Systems Technology. POWERCON, pp. 1–5. http://dx.doi.org/10.1109/POWERCON53406.2022.9929940.

Van Vliet, B., 2023. Abe Sklar's. http://dx.doi.org/10.2139/ssrn.4198458, The Original Document and an English Translation (March 3, 2023), URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4198458.

VDEW, B.C., 1999. Repräsentative VDEW- lastprofile. In: 2017 14th International Conference on the European Energy Market. EEM, p. 46, URL https://www.bdew.de/media/documents/1999_Repraesentative-VDEW-Lastprofile.pdf.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors, 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 17, 261–272. http://dx.doi.org/10.1038/s41592-019-0686-2.

Wang, Y., Chen, Q., Hong, T., Kang, C., 2019a. Review of smart meter data analytics: Applications, methodologies, and challenges. IEEE Trans. Smart Grid 10 (3), 3125–3148. http://dx.doi.org/10.1109/TSG.2018.2818167.

Wang, H., Lei, Z., Zhang, X., Zhou, B., Peng, J., 2019b. A review of deep learning for renewable energy forecasting. Energy Convers. Manage. 198, 111799.

Wang, Y., Zhang, N., Chen, Q., Kirschen, D.S., Li, P., Xia, Q., 2018. Data-driven probabilistic net load forecasting with high penetration of behind-the-meter PV. IEEE Trans. Power Syst. 33 (3), 3255–3264. http://dx.doi.org/10.1109/TPWRS.2017.2762599.

Zhang, Q., Chen, J., Xiao, G., He, S., Deng, K., 2023. TransformGraph: A novel short-term electricity net load forecasting model. Energy Rep. 9, 2705–2717. http://dx.doi.org/10.1016/j.egyr.2023.01.050, URL https://www.sciencedirect.com/science/article/pii/S2352484723000604.

Zhang, Y., Wang, J., Wang, X., 2014. Review on probabilistic forecasting of wind power generation. Renew. Sustain. Energy Rev. 32, 255–270. http://dx.doi.org/10.1016/j.rser.2014.01.033.

Zhou, B., Meng, Y., Huang, W., Wang, H., Deng, L., Huang, S., Wei, J., 2021. Multi-energy net load forecasting for integrated local energy systems with heterogeneous prosumers. Int. J. Electr. Power Energy Syst. 126, 106542. http://dx.doi.org/10.1016/j.ijepes.2020.106542.