

A Spatial Forecasting Model for Solar PV Generation and
Its Application in Household Electricity Management

Thanet Chitsuphaphan

A thesis submitted for the degree of
Doctor of Philosophy

Department of Mathematical Sciences

University of Essex

September, 2022

Abstract

A time-series data that depends on sunlight, such as solar photovoltaic generation output, principally consists of double sinusoidal components since solar irradiance reaches the Earth's surface while it spins and orbits. In addition, a part of solar energy is absorbed in the atmosphere by water vapour, dust, and ozone. In other words, geographical coordinates and regional weather conditions influence the data patterns. Therefore, this study focuses on developing a novel forecasting model to capture double seasonal patterns by adopting the spatial information of the data such as sunshine duration, cloudiness, and geographical coordinates. To explain an intra-day cyclical movement, a sine wave function with a predicted magnitude is considered to be integrated as a main part of the proposed model. Besides, an additive seasonal exponential smoothing model, which is a classical decomposition approach widely used for short-term forecasts, plays a role in adjusting the step-by-step error of forecasting over the daily pattern of the sine wave. Aggregated (intra-day to daily) data and regional daily weather-related variables are a real-world data set used in an empirical analysis by a regression model. The numerical results showed how the performance of the proposed model at different time horizons (e.g., one-step and one-period ahead forecasts) compared with existing models by the mean error (ME), mean absolute error (MAE), and root mean square error (RMSE).

Moreover, we present an application of the proposed forecasting model using a stochastic programming (SP) model to optimise electricity usage in a household with a photovoltaic system and an electric vehicle (EV). The several types of solar panel

electricity systems, including on-grid and hybrid, with EV battery incentive schemes will be developed by the two-stage SP model to investigate how to balance day-ahead electricity supply and demand and minimise daily electricity costs.

Acknowledgements

I am so grateful to Dr Xinan Yang and Dr Dai Hongsheng for their kind help and support over my PhD lifetime. Without your professional knowledge and experience, I cannot continue with my work, concentrate on such an interesting topic, and write this thesis. During the pandemic, I especially appreciate your patience and understanding. I will always remember all your kind lessons forever. Also, I wish to express my hearty appreciation for all the surrounding friendships I have obtained. I am thankful to my family and my wife for understanding my circumstance during the PhD, which makes me unable to take care of my parents and spend time together. Moreover, I want to thank all my Essex friends-P'Nicky, P'Nong, and P'Chom-for their support and re-boost energy for each other when we repeatedly feel failure and thank everyone I did not mention. Finally, I would like to appreciate the Royal Thai Government Scholarship Programme for supporting costs while studying here and the Thai Government Students' Office team for facilitating all educational processes.

Preface

I launched this topic with the problem of increasing the value of data by incorporating spatial information into a statistical-based optimisation approach. My problem initially aims to study how to set bid prices for renewables-based electricity for energy suppliers. After considering the issue for a while, energy storage development seems to be an important key to the energy revolution for a successful shift from fossil fuel energy generation to new renewable energy. Therefore, the objective has been reframed to concern increasing the potential of energy storage by focusing on a PV power statistical forecast and stochastic programming model for day-ahead planning of household electricity consumption. My raised issues hopefully make readers interested and decide to read the thesis. Also, I deeply expect these aims' results to be more valuable and worthwhile for their application.

Publications

As a long-distance PhD learning, the following publications have been produced as the result and methodology reflections during the research, which are closely related to each part of the thesis chapter.

1. T. Chitsuphaphan, X. Yang and H. Dai, “Stochastic Programming for Residential Energy Management with Electric Vehicle under Photovoltaic Power Generation Uncertainty”, 2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Liege, Belgium, 2020, pp. 1-6, doi: 10.1109/PMAPS47429.2020.9183393.
2. Forecasting high-frequency time series with double sinusoidal patterns using spatial exponential smoothing (Under revision)
3. X. Yang, T. Chitsuphaphan, H. Dai, and F. Meng, “EVB-Supportive Energy Management for Residential Systems with Renewable Energy Supply”, World Electric Vehicle Journal. 2022; 13(7):122. <https://doi.org/10.3390/wevj13070122>.

Contents

Abstract	i
Acknowledgements	iii
Preface	iv
Publications	v
List of Abbreviations	xii
List of Symbols	xiii
Nomenclature	xiv
1 Introduction	1
1.1 Background	1
1.2 Solar-based time series	4
2 Literature review	8
2.1 Exponential Smoothing	8
2.2 Solar PV generation forecasting	15
2.3 Theoretical background	16
2.3.1 Single seasonal exponential smoothing	16
2.3.2 Double seasonal exponential smoothing	18
2.3.3 Innovations state space models for exponential smoothing . .	19

2.3.4	Exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components (TBATS)	20
3	Spatial forecasting model	23
3.1	Basic principles in solar radiation	23
3.2	Spatial forecasting analysis	25
3.2.1	Sinusoidal model with explanatory variables	25
3.2.2	Spatial exponential smoothing model	26
3.3	Two-stage estimation	28
3.3.1	Stage 1: Estimation of regression coefficients	30
3.3.2	Stage 2: Determining the optimal values of smoothing parameters	31
3.4	Numerical study	32
3.4.1	Measuring forecast accuracy	32
3.4.2	The empirical data	33
3.4.3	Solar-PV Generation Data in Cambridge, East Midlands (United Kingdom)	34
3.4.4	Solar-PV Generation Data in Plymouth, Southwest England (United Kingdom)	36
4	Its Application via Stochastic Programming	40
4.1	Renewable energy for households	40
4.2	Stochastic programming	44
4.3	Default setting and system parameters	46
4.4	Day-ahead probabilistic PV generation forecast	47
4.5	Optimization model	50
4.5.1	Notations and parameters	50
4.5.2	System layouts and models	50
4.6	Numerical Results	59

4.6.1	Experiment settings and system parameters	59
4.6.2	Scenarios of solar PV generation	62
4.6.3	Results of SP models	63
5	Conclusions and Discussion	67
5.1	Conclusion	67
5.2	Discussion	68
5.2.1	Future Research	69
5.2.2	Limitations of the Study	70
A	Sunrise and sunset calculation	72

List of Figures

1.1	Model classification according to spatial and temporal resolutions [1].	3
1.2	The solar-based time series with high-frequency: (a) orbital and rotational characteristics of the Earth with 365.256-day sidereal period and 24-hour period, respectively, (b) the data that generate by the sunlight throughout four years, (c) daily peak of time series throughout four orbital Earth, and (d) the daily cycle of time series patterns.	5
2.1	Time series corresponding to the ETS models with different types of error, trend and seasonality [2]	13
3.1	The idea behinds sine wave function: (a) sunrise and sunset time in different places around the world, and (b) the Earth's orbit and rotation.	24
3.2	The spatial exponential smoothing model building process	30
3.3	Solar radiation map of the UK (source: MCS Design Guide), and average annual sunshine hours of study points.	35
3.4	Comparing the one-step ahead forecasts (48 iterations on 1 August 2013) and one-period (1 August 2013): actual data (solid), SES(dashed), ARIMA (dotdash), TBATS (longdash), Sp-Exp (twodash).	38
3.5	Five days of day-ahead forecasts: actual data (solid), SES(dashed), ARIMA (dotdash), TBATS (longdash), Sp-Exp (twodash).	39
4.1	The average household electricity consumption profile for a typical UK home (Watts) [3]	42

4.2	Graphical representation of home energy consumption and battery usage concepts	47
4.3	The 24-hour ahead solar PV output forecast when giving different states of cloud cover using real data for model training	49
4.4	Alternative layouts of home electric systems	52
4.5	Average 24-hour UK household electricity profile and PV generation curves on summer weekdays	61
4.6	Daily household consumption costs by solar PV and battery capacities	63

List of Tables

2.1	Timeline: A brief history of exponential-smoothing approaches	9
2.2	ETS Components and Models	11
2.3	The linear forms of the well-known exponential smoothing for ETS Components [4]	12
2.4	Notation for the well-known exponential smoothing models	14
3.1	The performance results of short-term forecasts models involved for solar-based time series.	36
3.2	The comparative assessment of all forecasting methods for one-step- ahead and one-day-ahead forecasts during 1st - 31st August 2013. . .	37
3.3	The estimation results of short-term forecasts models involved for solar- based time series	38
4.1	Abbreviations and system notations	51
4.2	Household power system setups	52
4.3	Forecasting estimation	62

List of Abbreviations

The following abbreviations are used throughout this thesis:

Abbreviation	Definition
SES	Simple Exponential Smoothing
ARIMA	Autoregressive Integrated Moving Average,
HW	Holt-Winters exponential smoothing
Sp-Exp	Spatial exponential smoothing
PI	Prediction Interval
ETS	Error-Trend-Seasonality models or Innovations state space models for exponential smoothing identified by three relatives: Error, Trend, and Seasonality
HWT	Holt-Winters-Taylor exponential smoothing
TBATS	Exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components
GPS	Global Positioning System
PV	Photovoltaic
HES	Home Energy Systems
SExp	Spatial exponential smoothing model
SP	Stochastic Programming
SDP	Stochastic Dynamic Programming
ESS	Energy Storage System
kW	Kilowatt
kWh	Kilowatt hour
SOC	State of Charge (0-100%)
DOD	Depth of Discharge (0-100%)
EV	Electric Vehicle
EVB	Electric Vehicle Battery (kWh)
HB	Home Battery (kWh)
PEV	Plug-in Electric Vehicle
PHEV	Plug-in Hybrid Electric Vehicle

List of Symbols

Symbols	Definition
S	residential solar photovoltaic system
G	regional electricity grid
E	electric vehicle battery (EVB)
B	home energy storage
V	plug-in hybrid electric vehicle (PHEV)
T	typical household appliances
D	small household appliances and electric devices

Nomenclature

Nomenclature	Definition
P_i^t	cost of electricity from resource i at time t , $i \in \{S, G\}, t \in \{1, 2, \dots, h\}$, [£/kW]
A_i^t	amount of electricity supply from source i at time t , $i \in \{S, G\}, t \in \{1, 2, \dots, h\}$, [kW]
U_j^t	amount of electricity demand from category j at time t , $j \in \{V, T, D\}, t \in \{1, 2, \dots, h\}$, [kW]
C_k	capacity of energy storage k , $k \in \{E, B\}$, [kW]
δ^t	indicator showing if EV is available at home during interval t , $t \in \{1, 2, \dots, h\}$
ρ	conversion rate of power transmission
γ	energy selling price as a proportion to the lowest ToU tariff (when selling of energy is allowed)
$x_{i,j}^t$	amount of electricity transmission from i to j at time t , $i \in \{S, G\}, j \in \{V, T, D\}, t \in \{1, 2, \dots, h\}$, [kW]
$y_{i,k}^t$	amount of electricity transmission from i to k at time t , $i \in \{S, G\}, k \in \{E, B\}, t \in \{1, 2, \dots, h\}$, [kW]
$w_{k,k'}^t$	amount of electricity transmission between energy storage at time t , $k, k' \in \{E, B\}, t \in \{1, 2, \dots, h\}$, [kW]
$z_{k,j}^t$	amount of electricity transmission from k to j at time t , $k \in \{E, B\}, j \in \{V, T, D\}, t \in \{1, 2, \dots, h\}$, [kW]
l_k^t	storage level of type k battery at time t , $k \in \{E, B\}, t \in \{1, 2, \dots, h\}$, [kW]
s^t	amount of surplus electricity generated from solar PV during interval t , $t \in \{1, 2, \dots, h\}$, [kW]

Chapter 1

Introduction

This chapter gives reasons why we are going to focus on developing a spatial forecasting model and method for a multiple seasonality time series, and also presents the main objectives of this study.

1.1 Background

Solar Photovoltaic (PV) forecasts

For decades, renewable energy and energy storage technologies have been developed to conserve natural resources and reduce global carbon emissions from fossil fuels [5, 6]. Numerous worldwide energy leaders are shifting their output away from fossil fuels and toward renewable energy sources such as solar, wind, hydro, and biomass, which are frequently referred to as clean energies. Renewable energy accounted for more than a quarter of worldwide power consumption by the end of 2018, and its capacity has increased by 8%, led by wind energy and solar PV[7]. Renewable electricity generation in the UK accounted for 35.8 per cent of total supply in 2019, led by wind, bioenergy, solar photovoltaic, and hydropower [8]. Renewable energy sources are classified in a given location by taking into account existing energy flows and geographical and temporal changes [9]. Solar panels and tiny wind turbines are two examples of fea-

sible domestic renewable energy deployments in the United Kingdom [10, 11]. Solar PV adoption has accelerated in recent years, as the International Energy Agency (IEA) declared solar photovoltaic (PV) power to be the new king of energy, as it is continuously less expensive than fossil-fuel power plants. However, major uncertainties exist in the generation of electricity from natural resources; for example, solar energy and wind turbine systems are highly dependent on weather conditions[12].

Therefore, to reduce the uncertainty caused by natural conditions, forecasting tools are one of the solutions for understanding variation sources. By examining the relationship between horizons, models, and activities in Fig 1.1, it was determined that statistical approaches seem to be appropriate for projecting intra-day to day-ahead energy consumption in residential (small-scale of spatial resolution) energy systems (HES). What is the most effective statistical response approach for its applications? The statistical models that are viable are dependent on the data available to forecasters, such as the photovoltaic system, historical power production statistics, and weather conditions [13, 14]. As may well be noticed, several of the smartphone applications related to renewable energy attempt to forecast PV output to help owners balance and plan their energy usage.

For instance, PV Forecast: Solar Power Generation Forecasts [15] is a mobile application that provides hour-by-hour forecasts of solar roof electricity generation for the next 48 hours and daily forecasts for the next 7 days for any location, making it the ideal tool for PV rooftop owners to connect the system without using Bluetooth or Wi-Fi. The information in the output, which is provided to the user for optimizing self-consumption (Household PVs), is calculated using parameter inputs such as PV system capacity, sun altitude and azimuth, sunrise and sunset times, daylight duration time, and solar noontime.

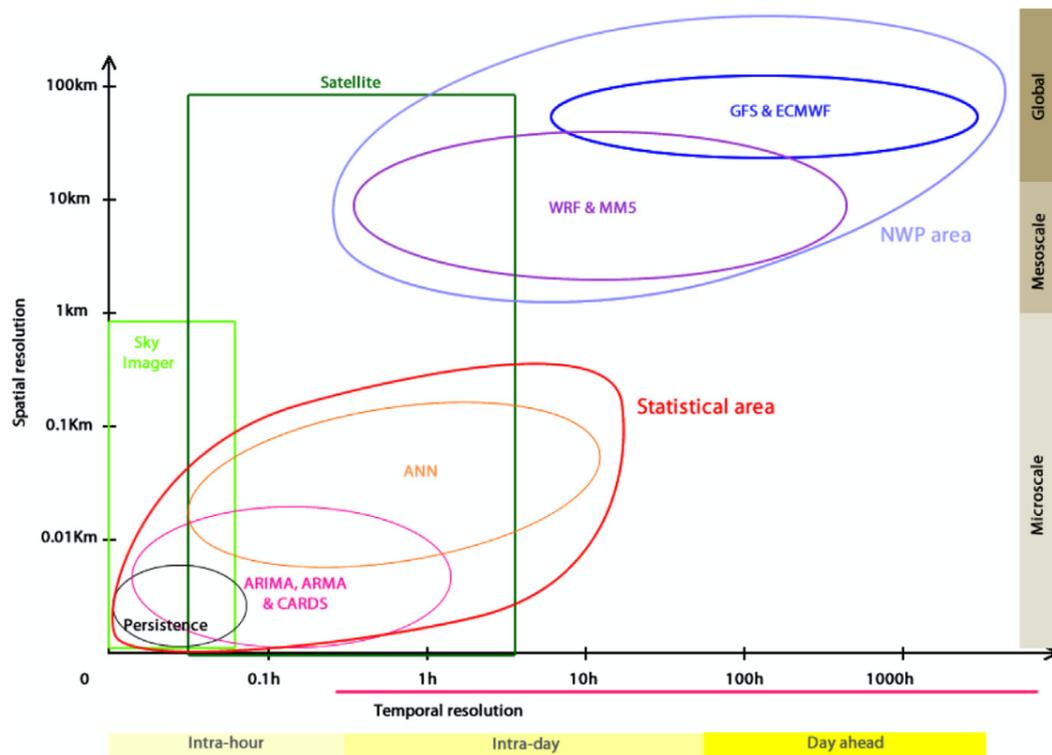


Figure 1.1: Model classification according to spatial and temporal resolutions [1].

The uncertainty of solar energy

In general, solar power production is uncertain due to meteorological variables such as clouds that can block sunlight from the sun. It brings about numerous problems relating to an inefficient allocations of resources. In the context of photovoltaic power potential, maximizing the self-consumption efficiency of electricity generation, planning energy usage, and storing electricity for later use all appear to be possible options without requiring end-user behaviour change.

Energy storage technologies are a key success factor to unlocking the full potential of sustainable renewables[16], allowing energy produced by clean, affordable sources, like wind and solar power, to be stored for later use. Moreover, batteries will be crucial factor in the transition from fossil fuels to renewable energy. According to the characteristics of storage technologies[17], battery solutions are an essential to balance demand and supply in household (as electricity end-user) during a day for improving renewable energy self-consumption.

1.2 Solar-based time series

In order to clarify the concept and terminology of multiple seasonality time series connected with the sunlight at every spot on earth - which is discussed throughout this study - we shall define the concept and terms involved as follows.

Solar radiation

Solar radiation, also known for short as sunlight, refers to the electromagnetic radiation released by the sun. Solar radiation is a renewable source of energy that may be captured and converted into usable forms of energy such as heat and electricity using a variety of technologies. The amount of solar radiation reaching any given area on the Earth's surface varies by the distance from Earth to the sun, and the Earth's rotation. In other words, the solar intensity varies according to geographic location, time of day, season, local landscape, and local weather.

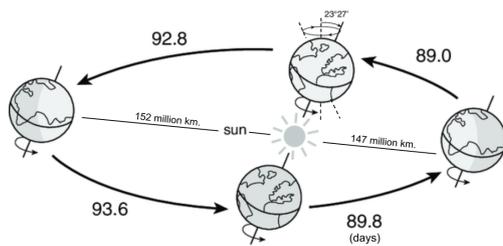
Components of the solar-based time series

Classical methods of time series analysis are concerned with decomposing a series into four fundamental components: trend, seasonal variations, cyclical variations, and irregular components. The solar-based time series is a univariate data affected by the sun, the daily pattern of time series that results directly from the earth's exposure to the sun (see Figure 1.2(b)) varies according to the amount of sunlight that reaches the earth's surface at indifferent rotation and orbit places (see the earth's repeated movements in Figure 1.2(a)). To have a better understanding of the earth's movement, the following three features are critical to consider.

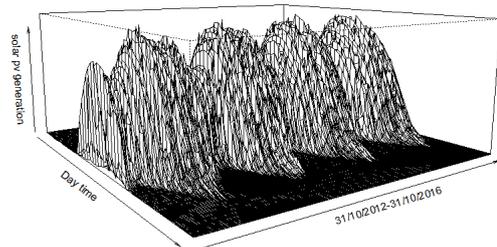
- The earth rotates on its axis, resulting in daily non-negative sine wave patterns as shown in Figure 1.2(d);
- Sine wave patterns at each location according to geographical coordinates have different sunrise and sunset times even though the location has the same local

time, i.e. there are different daytime lengths according to spatial data due to the earth's axis tilt;

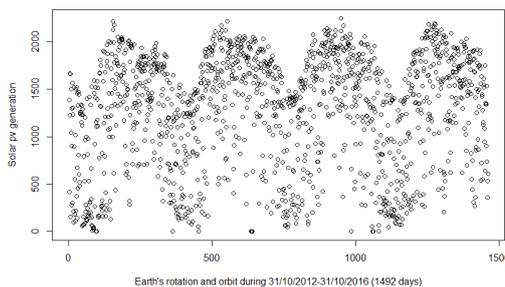
- The earth's orbit movement on the ecliptic plane results in seasons of the year. Figure 1.2(c) illustrates the annual maximum of solar irradiance (power per unit area received from the Sun) on the planet's surface as a function of the distance between the earth and the sun in an elliptical orbit, as depicted graphically in Figure 1.2(a). For example, when solar energy reaches the earth's surface in summer (near-side), as illustrated in Figure 1.2(a), it results in a higher solar intensity than when solar energy reaches the earth's surface in winter (far-side), as peak daily solar power generation exhibits upper and lower amplitude bounds for each period, as illustrated in Figure 1.2(c).



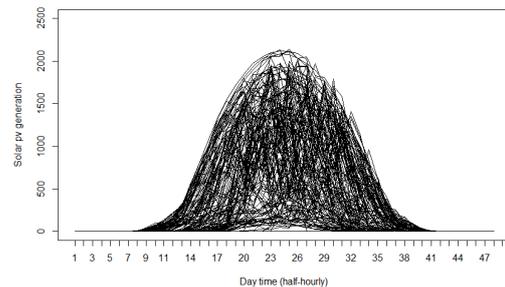
(a) Earth's rotation and orbit
(Source: Berger and Loutre, 1994b, as cited in Berger and Yin, 2012 [18])



(b) Solar-based time series characteristics



(c) Daily peak solar-based time series



(d) High-frequency time series each day

Figure 1.2: The solar-based time series with high-frequency: (a) orbital and rotational characteristics of the Earth with 365.256-day sidereal period and 24-hour period, respectively, (b) the data that generate by the sunlight throughout four years, (c) daily peak of time series throughout four orbital Earth, and (d) the daily cycle of time series patterns.

Therefore, in all of the above-mentioned, the components of data in relation to sunlight, contain the following:

- first-seasonal (intra-day) variations by the earth as it spins on its axis¹,
- second-seasonal (annual) variations by the earth as it orbits the sun (meteorological Seasons)²,
- irregular component caused by the uncertainty of weather and other random factors.

Recently, there has been no statistical literature that has looked at the components of solar-based time series. So, this study will focus on how to make a short-term spatial forecasting model. The main objectives of this study are as follows:

1. Clarifying the concept of a solar-based time series
2. Developing spatial forecasting for double-sinusoidal time series by combining a short-term forecasting model with a causal model included
 - Data aggregation of the time series and related weather conditions for regression analysis,
 - Boundary value analysis of non-negative forecast defined by spatial information such as latitude and longitude,
 - Additive seasonal exponential smoothing,
 - Spatial exponential smoothing.
 - Model
 - Method

¹Earth currently has an axial tilt of about 23 degrees approximately away from vertical, the North Pole stays in full sunlight all day in the summer, and falls in the twilight until early October during the Autumn

²The Earth's orbit is elliptical, with the Sun nearer one end than the other. This means Earth's distance from the Sun varies year-round.

3. Applying a two-stage stochastic programming model for increasing the potential of energy storage for HES based on 24-hourly PV outputs simulated by the proposed approach.

The remainder of this report is organised as follows. The second chapter examines the literature on statistical forecasting approaches for single and multiple seasonality time series, ranging from classical to state-of-the-art models. Furthermore, causal forecasts involved are thought to be the relationship between weather predictions and peak PV output. Chapter 3 describes how to develop a spatial model by combining a short-term forecasting technique with a causal forecast. Also, actual data from three different daytime places are used to analyse the robustness of the proposed model, in terms of spatial scale, in comparison to the statistical models involved, using the ME, MAD, and RMSE criteria. Chapter 4 presents its application by using a two-stage stochastic programming approach for planning electricity consumption associated with household energy systems (HES) when an electric vehicle (EV) and energy storage are used. Based on UK household electricity survey data, a two-stage programming approach was used to balance intra-day electricity use in PV-installed households and energy batteries. The final chapter discusses the proposed model's findings and its application to HES.

Chapter 2

Literature review

This chapter provides history and theoretical background on univariate time series with single and multiple seasonal patterns based on exponential smoothing approaches used to capture and explain characteristics of time series concerning what issues the proposed model deal with. In addition, we will explore solar PV forecasting tools and their application on electricity control in solar-panel households.

2.1 Exponential Smoothing

[19] This section gives a brief summary of how exponential smoothing forecasting has changed over time, as shown by the history timeline in Table 2.1 until 2014, when TBATS with regressors was created, as well as how ETS taxonomy was developed because of innovative exponential smoothing findings.

A simple exponential smoothing (SES) algorithm was developed during World War II by Robert G. Brown for fire-control information on the submarine position[20]. Brown improved the concept and created techniques for trend detection and seasonality detection, and introduced exponential smoothing into inventory management, planning, and control in 1956[21]. During the 1950s, Charles C. Holt worked independently of Brown to create a comparable approach for exponential smoothing additive trends and a completely new method for smoothing seasonal data, i.e. the data

Table 2.1: Timeline: A brief history of exponential-smoothing approaches

PY	Findings	Title
1956	SES	Exponential Smoothing (originated during World War II[20]) for Predicting Demand[21], Robert G. Brown (1923-2013): the father of exponential smoothing
1957	Holt	Forecasting seasonals and trends by exponentially weighted moving averages, reprinted 2004[22], i.e. no trend, additive and multiplicative trend, non-seasonal, and additive seasonal series with additive or multiplicative error structure are examined
1960	Holt-Winters	Forecasting Sales by Exponentially Weighted Moving Averages (+ multiplicative seasonal effect) [23]
1969	ETS(3x3)	Exponential Forecasting: Some New Variations, the ETS taxonomy originated [24]
1985	ETS(2x4x3)	Exponential smoothing: The state of the art (added additive damped trend into the ETS)[25]
1991	PI	Prediction intervals (PIs) for multiplicative Holt-Winters[26]
2001		Forecasting Models and Prediction Intervals for the Multiplicative Holt-Winters Method[27]
2002		A state space framework for automatic forecasting using exponential smoothing methods[28]
2003	HWT	Short-term electricity demand forecasting using double seasonal exponential smoothing [29]
2003	+ Damped T	Exponential smoothing with a damped multiplicative trend [30]
2006	ETS(2x5x3)	Exponential smoothing: The state of the art — Part II (added multiplicative damped trend into the ETS) [31]
2008		Forecasting time series with multiple seasonal patterns [32]
2008		An evaluation of methods for very short-term load forecasting using minute-by-minute British data (+ weather forecasts)[33]
2010	Triple S	Triple seasonal methods for short-term electricity demand forecasting[34]
2011	TBATS	Forecasting time series with complex seasonal patterns using exponential smoothing[35] [Multiple seasonality (non-integer seasonal length) exponential smoothing based method + ARIMA + Box-Cox transformation]
2012		Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing[36]
2014	+ Covariates	TBATS with regressors [37]

with no trend, additive and multiplicative trend, non-season and additive seasonality with additive or multiplicative error are examined by exponential smoothing. Holt's first research was reported in an ONR brief in 1957 but remained unpublished until 2004[22]. However, Holt's concept received widespread attention in 1960. Winters in 1960 established the Holt-Winters forecasting system by putting Holt's ideas to the test using actual data for capturing multiplicative seasonality in a publication[23].

Pegals (1969) initially proposed an extension and flexibility of exponential smoothing based on the concept of time series components. Each method is denoted by a pair of letters (T, S) denoting one of three types of trend and seasonality, namely none (N), additive (A), or multiplicative (M) [24], as illustrated in the model with * in Table 2.2. For example, (N, N) denotes a simple exponential smoothing (SES) method; (A, N) is an exponential smoothing method with no seasonal effect and an additive trend, also known as the Holt's linear method; (A, A) is an exponential smoothing method with an additive trend and seasonal effect, also known as the Holt-Winters' method; and so on.

Gardner (1985) later extended it and included it into the ETS taxonomy, based on the same concept. ETS is an acronym for Error-Trend-Seasonality and describes how the components interact. The ETS incorporates an additive damped trend with two distinct effects of error: additive and multiplicative[25]. Thus, the ETS allows for possible $2 \times 4 \times 3 = 24$ models.

Taylor (2003) developed double additive seasonality exponential smoothing, shortly Holt-Winters-Taylor (HWT), to accommodate the intraday and intraweek cycles in intraday data[29], also introduced the multiplicative damped trend component which has performed well in numerous empirical studies after that[30].

Gardner (2006) provides a summary of the evolution of Brown and Holt's original work since the 1950s, as well as some of its most significant concepts. In addition to a single seasonal pattern, extended smoothing may handle seasonal patterns in their whole, including their level and trend components. Due to its simplicity of usage and

ability to compete with more complex approaches such as a regression-based model or a seasonal ARIMA model, this method has been frequently used to generate point predictions with a monthly or quarterly (low-frequency) seasonal pattern. Additionally, the ETS is updated using a multiplicative damped trend by Taylor, which means that the ETS may be supported by $2 \times 5 \times 3 = 30$ exponential smoothing models[31] as displayed characteristics as Fig 2.1, ETS models with additive or multiplicative error in Table 2.2, and expressed well-known exponential smoothing equations in Table 2.3 with the model notations in Table 2.4.

Table 2.2: ETS Components and Models

Trend component	Seasonal component		
	N (None)	A (Additive)	M (Multiplicative)
N (None)	NN*	NA*	NM*
A (Additive)	AN*	AA*	AM*
Ad (Additive damped)	AdN	AdA	AdM
M (Multiplicative)	MN*	MA*	MM*
Md (Multiplicative damped)	MdN	MdA	MdM

Taylor (2007) presented a new result for forecasting very high-frequency electricity demand by employing an innovative state-space model. The method implicitly assumes an identical daily seasonal pattern for all days, while the state-space model allows the daily seasonal pattern to be different according to different types of days [32, 38]. Moreover, Taylor still found that combining the method based on weather forecasts with the Holt-Winters' adaptation resulted in forecasts that outperformed all other methods beyond about an hour ahead [33]. A new triple seasonal Holt-Winters exponential smoothing method, which incorporates the third seasonal component (*annual pattern*) into the existing double seasonal Holt-Winters method with daily and weekly patterns, which was published three years later [34].

While many researchers are familiar with time series forecasting, they have difficulty with certain types of time series data. For the first time, the length of a time series expressed as a decimal number, such as the length of the Lunar calendar, has

been considered when developing forecasting techniques. De Livera et al. (2011) proposed the TBATS model, which is capable of handling complex seasonalities (e.g., non-integer seasonality, non-nested seasonality, and large-period seasonality) without imposing seasonality constraints, enabling the generation of detailed, long-term forecasts [35]. The model was developed by combining an exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components, to capture and forecast time series. TBATS with regressors was presented in 2014 by Hyndman using “Automatic Time Series Forecasting” with R to include regression variables, i.e., covariates in a time series model[37, 39].

Table 2.3: The linear forms of the well-known exponential smoothing for ETS Components [4]

ETS	Model	Smoothing method	Remark
ANN	$y_t = l_{t-1} + e_t$ $l_t = l_{t-1} + \alpha e_t$	$\hat{y}_t = \hat{l}_{t-1}$ $\hat{l}_t = \hat{l}_{t-1} + \alpha(y_t - \hat{y}_t)$	SES (1956)
AAN	$y_t = l_{t-1} + b_{t-1} + e_t$ $l_t = l_{t-1} + b_{t-1} + \alpha e_t$ $b_t = b_{t-1} + \alpha v e_t$	$\hat{y}_t = \hat{l}_{t-1} + \hat{b}_{t-1}$ $\hat{l}_t = \hat{l}_{t-1} + \hat{b}_{t-1} + \alpha(y_t - \hat{y}_t)$ $\hat{b}_t = \hat{b}_{t-1} + \alpha v(y_t - \hat{y}_t)$	Holt (1957)
AAA	$y_t = l_{t-1} + b_{t-1} + s_{t-m} + e_t$ $l_t = l_{t-1} + b_{t-1} + \alpha e_t$ $b_t = b_{t-1} + \alpha v e_t$ $s_t = s_{t-m} + \gamma e_t$	$\hat{y}_t = \hat{l}_{t-1} + \hat{b}_{t-1} + \hat{s}_{t-m}$ $\hat{l}_t = \hat{l}_{t-1} + \hat{b}_{t-1} + \alpha(y_t - \hat{y}_t)$ $\hat{b}_t = \hat{b}_{t-1} + \alpha v(y_t - \hat{y}_t)$ $\hat{s}_t = \hat{s}_{t-m} + \gamma(y_t - \hat{y}_t)$	Holt-Winters (1960)
AAdN	$y_t = l_{t-1} + b_{t-1} + e_t$ $l_t = l_{t-1} + b_{t-1} + \alpha e_t$ $b_t = \phi b_{t-1} + \alpha v e_t$	$\hat{y}_t = \hat{l}_{t-1} + \hat{b}_{t-1}$ $\hat{l}_t = \hat{l}_{t-1} + \hat{b}_{t-1} + \alpha(y_t - \hat{y}_t)$ $\hat{b}_t = \phi \hat{b}_{t-1} + \alpha v(y_t - \hat{y}_t)$	Additive Damped trend (1985)
AAdA	$y_t = l_{t-1} + b_{t-1} + e_t$ $l_t = l_{t-1} + b_{t-1} + \alpha e_t$ $b_t = \phi b_{t-1} + \alpha v e_t$ $s_t = s_{t-m} + \gamma e_t$	$\hat{y}_t = \hat{l}_{t-1} + \hat{b}_{t-1}$ $\hat{l}_t = \hat{l}_{t-1} + \hat{b}_{t-1} + \alpha(y_t - \hat{y}_t)$ $\hat{b}_t = \phi \hat{b}_{t-1} + \alpha v(y_t - \hat{y}_t)$ $\hat{s}_t = \hat{s}_{t-m} + \gamma e_t$	Multiplicative Damped trend (2003)

As all the above-mentioned, the exponential smoothing can not accommodate more than one seasonal pattern in its traditional structure but is excellent for the short-term forecast because it only considers one seasonal pattern in its calculations. This limita-

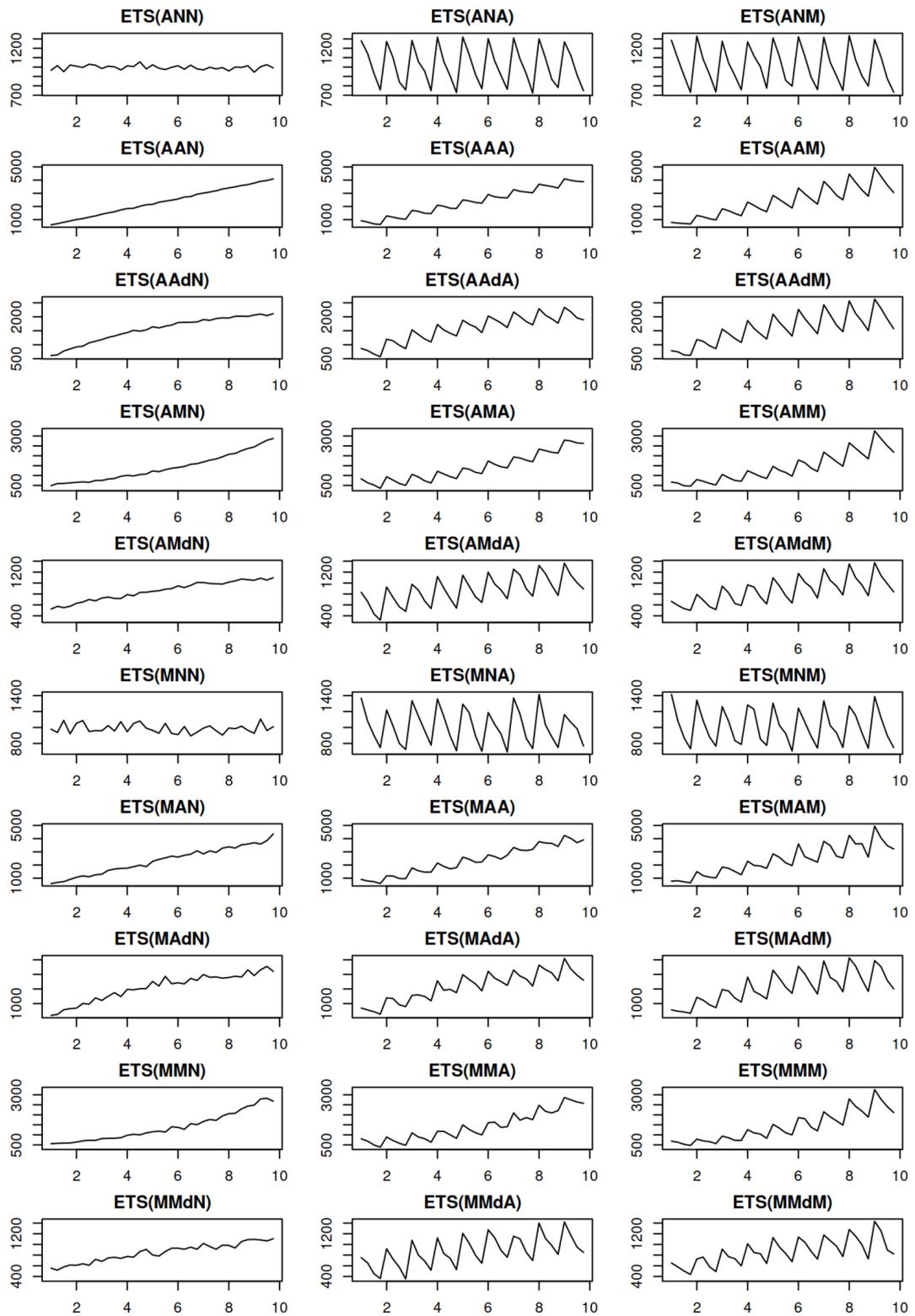


Figure 2.1: Time series corresponding to the ETS models with different types of error, trend and seasonality [2]

Table 2.4: Notation for the well-known exponential smoothing models

Symbol	Definition
l_t	level component of the series in period t
b_t	trend component of the series in period t
s_t	season component of the series in period t
α	Smoothing parameter for the level of the series, $0 \leq \alpha \leq 1$
v	Smoothing parameter for the trend, $0 \leq v \leq 1$
γ	Smoothing parameter for seasonal indices, $0 \leq \gamma \leq 1$
ϕ	Damping parameter, $0 \leq \phi \leq 1$
m	Number of periods in the seasonal cycle
y_t	Observed value of the time series in period t
\hat{y}_t	Forecast for 1 periods ahead from origin t , i.e., $\hat{y}_t = \hat{y}_{t-1}(1)$
e_t	One-step-ahead forecast error at time t , i.e., $y_t - \hat{y}_t$
\hat{l}_t	Estimated local level at time t
\hat{b}_t	Estimated trend at time t
\hat{s}_t	Estimated seasonal indices in period t

tion of the forecasting approach has not been recognized in high-frequency time series such as minutely-to-hourly time series, and so on until multiple seasonality methods are developed. However, these methods request a long horizon of time series and a high memory to recognise parameters for each seasonality for a time series observed on a very frequent basis (hourly or half-hourly). As a result, a solar-based time series, which is defined in Chapter I as double-seasonal high-frequency data, is unlikely to be suitable for multiple seasonal exponential smoothing methods such as ETS and TBATS. All of the strengths of these solutions can be used to capture each component of data and to reduce the amount of data required and memory used in a forecasting procedure.

Covariates are a feasible factor to apply to study a relation to the amplitude of sinusoidal annual characteristics (the second seasonal component of the solar-based time series as illustrated in Figure 1.2(c)) caused by the Earth's orbits in order to simplify and reduce the amount of memory required for parameter estimation. The magnitude of sine waves in each period of time (day) shown in Figure 1.2(d) directly depends on which the solar radiation have a relationship with sunshine duration [40–44] and cloud cover [40, 45–49] as coveriates, i.e. a regression model can be used to describe the so-

lar radiation and its relationship with geographical and weather conditions. Therefore, in this study, a multiple regression model as a causal model is a model used to forecast a daily peak PV generation which is a part of sinusoidal function.

2.2 Solar PV generation forecasting

Solar radiation requires reliable forecast information to make efficient use of the fluctuating energy output of PV systems. Solar irradiance, which refers to the amount of solar radiation received per unit area by a given surface (W/m^2), will be mentioned instead of radiation when discussing the electricity generation output from photovoltaic systems.

In the perspective of harvesting energy from the sunlight, stakeholders of a PV electricity market contain grid operators, electricity retailers, end-users, and government from four different parts such as generator, transmission, distribution, and customer. Electricity demands are frequently predicted from a resident-based (end-user) perspective for home energy systems (HES) with photovoltaics (PV)[50–52]. Another aspect is considered in this study: residential PV systems with varying capacity sizes are forecasted using the proposed spatial model for day-ahead electricity planning in households. Spatial data, such as coordinates and regional weather conditions (covariates), have a significant impact on the intra-day and daily patterns of solar-based time series, the following reviews involve conditions in solar PV generation output.

Diagne et al. (2013) look at solar irradiance forecasting methods in detail, including statistical approaches and techniques based on pictures of clouds. They also describe the current methods for predicting solar irradiance to help with method selection and to find future solar irradiance forecasting methods for small-scale grid management. Methods for predicting seasonality based on historical data of solar irradiance have been used successfully in time series forecasting for short time horizons up to one hour [1].

Barbieri et al. (2017) define temporal resolution relative to the spatial range of observation for various forms of input data and show how cloud elements influence the accuracy of a power forecast. In addition, time-series models can capture short-term features employing whole-sky cameras (clouds) for residential load balancing and transmission scheduling. Nonetheless, cloud types that correspond to cloud categories are depicted in a spectrum of blockage altitudes that varies [53].

A model that uses the Fourier series is known as the exponential smoothing state space model (ESSS, also known as the ETS). Its forecasting performance is compared with classic forecasting approaches like as ARIMA, ARMA, and so on, and it is discovered that the ESSS model produces more accurate predictions over a short-term forecasting horizon. It has also been observed that the accuracy of the model increases when the inputs have a high correlation factor with the outputs of the model [54–56].

2.3 Theoretical background

2.3.1 Single seasonal exponential smoothing

The solar-based time series data used in this study indicate double additive seasonality, so among the various forms of exponential smoothing methods, we start with the single additive seasonal exponential smoothing method (Holt-Winters, 1960) for forecasting time series. Assuming that seasonal time series can be decomposed into three unobserved components — level, trend, and season — the Holt-Winters method has been widely adopted as a standard forecasting technique.

The error correction form of the simple exponential smoothing model for forecasting an observed series (y_t) at time t consist of three updating equations for a level (l_t), a trend (b_t), and a seasonal pattern (s_t) for $t = 1, 2, \dots, n$ as expressed the AAA model

(Table 2.3) states as:

$$y_t = l_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2)$$

$$l_t = l_{t-1} + b_{t-1} + \alpha\varepsilon_t$$

$$b_t = b_{t-1} + \alpha v \varepsilon_t$$

$$s_t = s_{t-m} + \gamma\varepsilon_t,$$

where m is the length of the seasonal pattern. α, v and γ are smoothing parameters for the level, trend, and seasonal components, respectively. As a result of the additive Holt-Winters technique, we define the h prediction at the time t as

$$\hat{y}_{t+h|t} = l_t + hb_t + s_{t+h-m},$$

and one-step-ahead forecast error at the time t as

$$e_t = y_t - \hat{y}_{t|t-1} = y_t - (l_{t-1} + b_{t-1} + s_{t-m}).$$

Prediction intervals are a crucial component of the forecasting process, designed to reflect the uncertainty in point forecasts. Utilizing theoretical formulas based on the best-fitting model is the most frequent way for computing PIs. For the additive error model with a single source of error $\varepsilon_t \sim N(0, \sigma^2)$, a $100(1 - \theta)\%$ prediction interval (PI) for h -step forecast is

$$\left[\hat{y}_{t+h|t} - z_{\frac{\theta}{2}} \hat{\sigma}, \hat{y}_{t+h|t} + z_{\frac{\theta}{2}} \hat{\sigma} \right]$$

where θ is a significance level for PI, $z_{\frac{\theta}{2}}$ is the $\frac{\theta}{2}$ upper quantile of the standard normal, and $\hat{\sigma}$ is an estimate of the single source of error of the h -step forecast distribution.

This limitation of the Holt-Winters method is inapplicable to time series containing two or more seasonal patterns. If the standard Holt-Winters method must be applied to

multiple seasonalities, we must select only one of the possible seasonal patterns.

2.3.2 Double seasonal exponential smoothing

More than one seasonal pattern cannot be accommodated by the existing Holt-Winters method due to its limitation to a single season, as discussed in the previous subsection. Taylor (2003) proposed a novel variant of the exponential smoothing approach, called the double seasonal Holt-Winters exponential smoothing method with two seasonal patterns. In addition to the current seasonal index $s_{1,t}$, a new seasonal index $s_{2,t}$ was added to the new extra updating equation to expand the standard Holt-Winters approach.

The error-correction form of the additive double seasonal Holt-Winters exponential smoothing technique is given using four update equations for each component.

$$\begin{aligned} y_t &= l_{t-1} + b_{t-1} + s_{1,t-m_1} + s_{2,t-m_2} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2), \\ l_t &= l_{t-1} + b_{t-1} + \alpha\varepsilon_t, \\ b_t &= b_{t-1} + \alpha v\varepsilon_t, \\ s_{1,t} &= s_{1,t-m_1} + \gamma_1(1 - \alpha)\varepsilon_t, \\ s_{2,t} &= s_{2,t-m_2} + \gamma_2(1 - \alpha)\varepsilon_t, \end{aligned}$$

where $s_{1,t}$ and $s_{2,t}$ are the seasonal indices with seasonal lengths m_1 and m_2 , s.t. $m_1 \leq m_2$, respectively. α , v , γ_1 and γ_2 are smoothing parameters for the level, trend, first and second seasonal components, respectively. The h -step-ahead forecast at time t is

$$\hat{y}_{t+h|t} = l_t + hb_t + s_{1,t+h-m_1} + s_{2,t+h-m_2},$$

and one-step-ahead forecast error at the time t as

$$e_t = y_t - \hat{y}_{t|t-1} = y_t - (l_{t-1} + b_{t-1} + s_{1,t-m_1} + s_{2,t-m_2}).$$

As mentioned by Taylor et al. (2006), the double seasonal Holt-Winters method requires no model specification, whereas the double seasonal exponential smoothing method inherits the simplicity and robustness of traditional exponential smoothing methods. As with other exponential smoothing techniques, *the double seasonal Holt-Winters exponential smoothing technique may not provide accurate prediction intervals for point forecasts* [57].

2.3.3 Innovations state space models for exponential smoothing

Innovations state space models for exponential smoothing identified by three relatives: Error, Trend, and Seasonality (ETS). Based on the type of error, trend and seasonality, Pegels (1969) proposed a taxonomy, which was then developed further by Hyndman et al. (2002) and refined by Hyndman et al. (2008). According to this taxonomy, here defines a three-character string as shown in Table 2.2.

The ETS components have well-defined meanings. Additionally, ETS enables 30 models expressing a variety of error, trend, and seasonality characteristics. Table 2.2 and Figure 2.1 illustrates various time series with deterministic levels, trends, seasonality, and time series with an additive error term. The combination of ETS models obtained linear forms from the above components as expressed in Table 2.3.

The ETS model with a single source of error [58, Appendix A.4]

Gould et al. (2008) have made a contribution by extending the single seasonal innovations state-space model to accommodate innovations state-space models with either single or double seasonality that accommodate more than one seasonal pattern [32]. The state-space model for the double additive seasonal Holt-Winters exponential

smoothing method is derived and written by

$$\begin{aligned}
 y_t &= l_{t-1} + b_{t-1} + s_{1,t-m_1} + s_{2,t-m_2} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2) \\
 l_t &= \alpha_1(y_t - s_{1,t-m_1} - s_{2,t-m_2}) + (1 - \alpha_1)(l_{t-1} + b_{t-1}) \\
 b_t &= \frac{\alpha_2}{\alpha_1}(l_t - l_{t-1}) + (1 - \frac{\alpha_2}{\alpha_1})b_{t-1} \\
 s_{1,t} &= \frac{\alpha_3}{1 - \alpha_1}(y_t - l_t - s_{2,t-m_2}) + (1 - \frac{\alpha_3}{1 - \alpha_1})s_{1,t-m_1} \\
 s_{2,t} &= \frac{\alpha_4}{1 - \alpha_1}(y_t - l_t - s_{1,t-m_1}) + (1 - \frac{\alpha_4}{1 - \alpha_1})s_{2,t-m_2},
 \end{aligned}$$

where the ranges of the parameters of the structural model are

$$0 < \alpha_1, \frac{\alpha_2}{\alpha_1}, \frac{\alpha_3}{1 - \alpha_1}, \frac{\alpha_4}{1 - \alpha_1} < 1.$$

Therefore, the innovations state-space technique for the double seasonal Holt-Winters method is essentially identical to Taylor (2003)'s original concept. Because the smoothing parameters $(\alpha, \nu, \gamma_1, \gamma_2)$ of Taylor's double seasonal Holt-Winters technique should range from 0 to 1,

2.3.4 Exponential smoothing state space model with Box-Cox transformation, ARMA errors, Trend and Seasonal components (TBATS)

A TBATS model is a forecasting technique for a time series with complex seasonal patterns using exponential smoothing originally proposed by De Livera et al. (2011) that combines many of the components of statistical models into one single automated framework. The details of the model are as follows:

- **T**rigonometric terms for seasonality
- **B**ox-Cox transformation for heterogeneity
- **A**RMA errors for short-term dynamics

- Trend (possibly damped)
- Seasonal (including multiple and non-integer periods)

The TBATS model can be expressed mathematically by the following equations:

Box-Cox transformation:

$$y_t^{(\omega)} = \begin{cases} \frac{1}{\omega}(y^\omega - 1), \omega \neq 0 \\ \log y_t, \omega = 0 \end{cases},$$

Seasonal periods:

$$y_t^{(\omega)} = l_{t-1} + \phi b_{t-1} + \sum_{i=1}^T S_{t-m_i}^{(i)} + d_t,$$

Global and local trend:

$$l_t = l_{t-1} + \phi b_{t-1} + \alpha d_t,$$

$$b_t = (1 - \phi)b + b_{t-1} + \nu d_t,$$

ARMA error:

$$d_t = \sum_{i=1}^p \phi_i d_{t-i} + \sum_{j=1}^q \varepsilon_{t-j} + \varepsilon_t,$$

Fourier seasonal terms:

$$S_t^{(i)} = \sum_{j=1}^{k_i} S_{j,t}^{(i)},$$

$$S_{j,t}^{(i)} = S_{j,t-1}^{(i)} \cos \lambda_j^{(i)} + S_{j,t-1}^{*(i)} \sin \lambda_j^{(i)} + \gamma_1^{(i)} d_t,$$

$$S_{j,t}^{*(i)} = -S_{j,t-1}^{(i)} \sin \lambda_j^{(i)} + S_{j,t-1}^{*(i)} \cos \lambda_j^{(i)} + \gamma_2^{(i)} d_t.$$

Where y_t is a time series at time t , $y_t^{(\omega)}$ represents Box-Cox transformation of y_t with the parameter ω , l_t is the local level in period t , b_t is the short-term trend in period t , ϕ is the trend damping parameter, $S_t^{(i)}$ denotes the i th seasonal component at time t , m_1, m_2, \dots, m_T refer to the length of seasonal periods, d_t represents an ARMA(p, q) process and ε_t is a Gaussian white noise process with zero mean and constant variance σ^2 . The smoothing parameters are given by $\alpha, \nu, \gamma_1^{(i)}, \gamma_2^{(i)}$ for $i = 1, 2, \dots, T$, and $\lambda_j^{(i)} = \frac{2\pi j}{m_i}$ for $j = 1, 2, \dots, k_i$. $S_{j,t}^{(i)}$ represents the stochastic level of i th seasonal component and $S_{j,t}^{*(i)}$ represents the change in the stochastic level of i th seasonal component over time. Finally, $k_i = \frac{m_i}{2}$ denotes the required number of harmonics for the i th seasonal component. In general form, the TBATS model is designated as TBATS($\omega, \phi, p, q, \{m_1, k_1\}, \{m_2, k_2\}, \dots, \{m_T, k_T\}$).

In conclusion, the TBATS model is extremely versatile and can handle a wide variety of time series, making it especially suitable for data with large seasonal periods and multiple seasonal periods. Despite the fact that the model's point forecasts may appear accurate, the prediction intervals it generates are frequently overly large. In addition, the automation makes things pretty slow, particularly for very long time series, because it must evaluate a multitude of potential ways of assembling the model.

The motivation behind the TBATS, which adds a Fourier series (a periodic function) into the exponential smoothing state-space model for capturing various non-nested and non-integer seasonal components of high-frequency time series, is now acknowledged. In particular, the example provides an illustration of how to use a periodic function to capture the non-integer seasonal pattern that is a partial characteristic of solar-based time series formed by the influence of the Earth's orbit around the sun.

Chapter 3

Spatial forecasting model

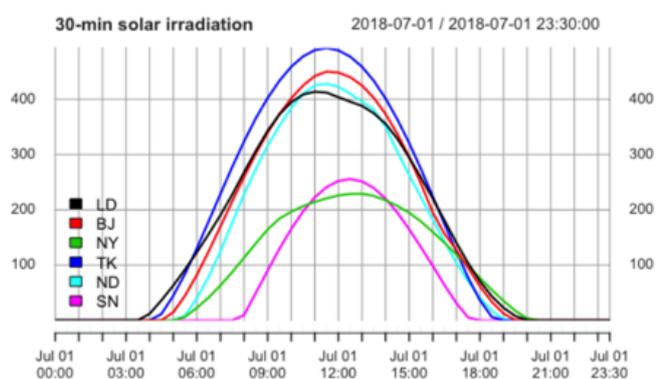
The PV generation output is a solar-based time series that indicates a very distinct length of double seasonality (intra-day and annual patterns). According to Taylor (2008) and Gould, et.al.(2008), double seasonal exponential smoothing methods for high-frequency time series under different patterns of days by weather forecasts [32, 33] a good performance in numerous empirical studies. In addition, in order to forecast, the TBATS model needs a significant amount of data storage for historical data with high frequency and seasonality. As a consequence of this, we really need to understand how to simplify the method so that it can be applied in the real world by making use of the fewest possible model parameters and reducing the amount of work that needs to be done computationally. In order to simplify the process of parameter estimation and cut down on the amount of memory needed, we are going to develop a spatial model. Estimation was broken down into two stages by the proposed model: (1) regression analysis on aggregating time series data to daily, and (2) smoothing and initial determinations on a time-series frequency.

3.1 Basic principles in solar radiation

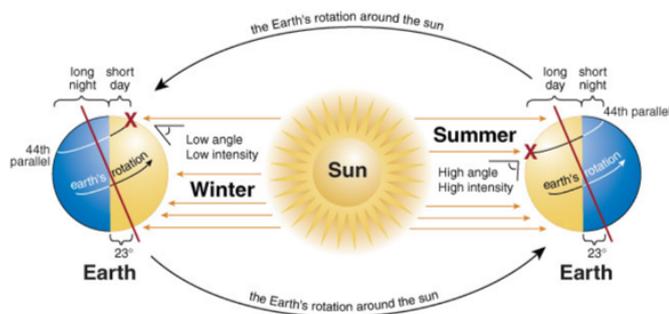
Every location on Earth's surface receives sunlight for at least a portion of the year. As shown in Figure 1.2, the amount of solar radiation that reaches any given spot

on the Earth's surface varies based on geographic location, time of day, season, local landscape, and local climate. Because the Earth is spherical, the rotation of the Earth around its tilted axis creates a daily cycle of the Sun's power that varies in duration based on geographic coordinates. Additionally, the Earth has an elliptical orbit around the sun and is closer to the sun during a portion of the year, as shown in Figure 1.2(a), which divides the year (orbital period) according to the distance between the Earth and the sun, also known as a season.

Figure 3.1 shows that geographic location information, including London, Beijing, New York, Tokyo, New Delhi, and Sydney, have impacted the variation in sunrise and sunset times, and sunshine duration, which affects directly both solar electricity generation time and solar radiation pattern.



(a) Solar radiation cycle on 1st July 2018 in different places



(b) Seasonal and weather conditions affected by the Earth's orbit and rotation

Figure 3.1: The idea behinds sine wave function: (a) sunrise and sunset time in different places around the world, and (b) the Earth's orbit and rotation.

The following section is going to give the details of how incorporating spatial information into statistical-based forecasting approaches over the basic physical principles of photovoltaics (PV).

3.2 Spatial forecasting analysis

The intensity of solar radiation, as energy input from the sun to photovoltaic cells, on the Earth's surface on a daily rotation has a sine wave or sinusoid pattern. Therefore, a day-ahead forecasting model of the solar intensity mainly runs on a sinusoidal model based on the season of the year by applying local-weather conditions.

3.2.1 Sinusoidal model with explanatory variables

Let $\{y_{\vec{s},t}\}, t = 1, 2, \dots, n$ represents a high-frequency time series with double seasonal cycles reflected by the sunlight, namely a solar-based time series as defined in Section 1.2. The data is generated and measured on Earth's surface at position $\vec{s} = (x_1, x_2)$ with reference to global positioning system (GPS) coordinates, s.t. x_1, x_2 represent the latitude and longitude of a location in decimal, respectively. For example, GPS coordinates of University of Essex, United Kingdom (Latitude: 51.8763 Longitude: 0.9449), i.e., $\vec{s} = (51.88, 0.94)$. The solar-based time series has general characteristics as shown in Figure 1.2(b). Solar irradiance on any point of the earth's surface has principally a sine wave pattern on daytime with different frequency, day length, and magnitude, depending on spatial information and weather conditions.

The general formulation of the sine wave is expressed by

$$\psi_t = A \sin(2\pi ft + \varphi) + V, t = 1, 2, \dots, T$$

where ψ_t represents observations at time t , A is an amplitude, f is a frequency of period, φ is a horizontal shift, and V is a vertical shift.

We highlight the sine function used to detect both the daily and annual seasonality of the data, which is the horizontal shift (φ) and the vertical shift (V) adjusted by spatial and weather-related data. In other words, geographical information where the measured data will be used to determine the duration between sunrise and sunset (the range of positive values of the data each day) and daily frequency, whereas the meteorological information, e.g., sunshine duration and cloud cover [59], used to analyse in a regression model to forecast the daily magnitude as annual seasonal characteristics. Therefore, the adjusted sine function will be an asymmetric non-negative curve corresponding to the diurnal variation of solar radiation.

According to the regression model, we assume that the daily seasonal factor is an observed dependent variable aggregated by daily peak or average of the observed time series depending on the application of the model. Also, meteorological variables, considered as explanatory variables [60, 61] measure sunshine duration and cloud cover which are provided by an atmospheric model from a meteorological authority, such as Met Office, National Oceanic and Atmospheric Administration (NOAA), or weather forecasts mobile applications. Therefore, the non-negative sine function used to describe daily seasonal shapes of solar-based time series for a daylight cycle at different (latitude) locations given weather-related covariates can be expressed as the following equations:

$$\begin{aligned} \psi_{\vec{s},t} &= \frac{A_{\vec{s},j}}{2} \sin(2\pi f_j t_j - \frac{\pi}{2}) + \frac{A_{\vec{s},j}}{2}, t \in [u_{\vec{s},t}, v_{\vec{s},t}] \\ \tilde{A}_{\vec{s},j} &= \beta_0 + \beta_1 x_{\vec{s},1j} + \beta_2 x_{\vec{s},2j} + \dots + \beta_k x_{\vec{s},kj} + \eta_j, \eta_j \sim N(0, \sigma_a^2) \end{aligned} \quad (3.1)$$

where

3.2.2 Spatial exponential smoothing model

The seasonal exponential smoothing model, also known as Holt-Winters (HW), is the generalised approach to deal with the seasonal variation which originally rely on three basic models, simple exponential smoothing (local level), trend corrected exponential

- $\psi_{\vec{s},t}$ represents the seasonal component value at the \vec{s} -position and time t ,
 $\frac{A_{\vec{s},j}}{2}$ denotes the \vec{s} -position true amplitude at seasonal period j , $j = \lceil t/m \rceil$, when m is a frequency of the time series,
 $\tilde{A}_{\vec{s},j}$ denotes the \vec{s} -position observed amplitude value with observation errors at seasonal period j , i.e. $A_{\vec{s},j} = \mathbb{E}(\tilde{A}_{\vec{s},j})$,
 $u_{\vec{s},t}, v_{\vec{s},t}$ denote sunrise and sunset time series indices (starting and ending point of sine function) in period j , respectively, s.t. $m*(j-1)+1 \leq u_{\vec{s},t} < v_{\vec{s},t} \leq m*j$ calculated by Almanac for Computers (1990), i.e. $u_{\vec{s},t} = m*(j-1) + \lceil \text{sunrise time} \rceil$, $v_{\vec{s},t} = m*(j-1) + \lceil \text{sunset time} \rceil$
 f_j is a frequency of the function at period j s.t. $f_j = \frac{1}{n_j}$, $n_j = v_{\vec{s},j} - u_{\vec{s},j} + 1$,
 t_j represents the new time index in period j over the daylight time s.t. $t_j = t - u_{\vec{s},j} + 1$,
 $\beta_0, \beta_1, \dots, \beta_k$ denote regression coefficients,
 $x_{\vec{s},1j}, \dots, x_{\vec{s},kj}$ represent regression variables as meteorological information at position \vec{s} and period j , which can be predicted by the atmospheric model, and
 η_j denotes a Gaussian random disturbance in period j with zero mean and variance σ_a^2 .

smoothing (trend), and additive/multiplicative seasonal variation (season). Speaking of which, the model is suitable to explain one seasonal time series characteristic for short-term forecasting.

In the case of a high-frequency time series that has double seasonal cycles and shows a huge difference between two seasonal frequencies (e.g. if a half-hourly time series consists of daily and annual seasonal patterns), the frequency for the first and the second seasonalities will be 48 and $48 \times 365.25636 = 17,532.30528$ (365.25636 days is the orbital period of the earth around the Sun), respectively. The double seasonal exponential smoothing methods expanded by [29] and the TBATS approach developed by [35] which can accommodate two seasonalities, need a high data-storage of the previous seasonal data for forecasting. Therefore, it is important to know how to simplify the method for practical use by using the least possible number of model parameters and reducing the computational burden.

The sun's total solar irradiance changes slowly over time [62]. In other words, it is unlikely to make a difference to PV generation, i.e. there is no significant change in the

long-term trend. Therefore, the high-frequency time series of sunlight as a climate time series basically consists of a slowly varying trend, multiple and nested seasonality.

The current study attempts to replace the sine wave model as periodic phenomena Equation 3.1 with the seasonal component of additive seasonal exponential smoothing model. The adjusted sine function plays an important role in a-day-ahead forecast while exponential smoothing provides one-step-ahead procedures for detecting and adjusting to changes in the previous state. Therefore, the spatial exponential smoothing model with covariates for solar-based time series can be expressed in the error-correction form by

$$\begin{aligned}
y_{\bar{s},t} &= l_{t-1} + b_{t-1} + s_t + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2), \\
l_t &= l_{t-1} + b_{t-1} + \alpha\varepsilon_t, \\
b_t &= b_{t-1} + v\varepsilon_t, \\
s_t &= \psi_{\bar{s},t} + \gamma\varepsilon_{t-1}, \\
\psi_{\bar{s},t} &= A_{\bar{s},j} \sin(2\pi f_j t_j - \pi/2) + A_{\bar{s},j}, t \in [u_{\bar{s},t}, v_{\bar{s},t}], \\
\tilde{A}_{\bar{s},j} &= \beta_0 + \beta_1 x_{\bar{s},1j} + \beta_2 x_{\bar{s},2j} + \dots + \beta_k x_{\bar{s},kj} + \eta_j, \eta_j \sim N(0, \sigma_a^2),
\end{aligned} \tag{3.2}$$

where l_t, b_t, s_t denote the local level, trend, and spatial seasonal components of time series at time t , respectively, and α, v, γ are the smoothing values as in the traditional approaches. The model has manifested that the local and trend states depended on previous states whilst seasonal state depended mainly on spatial seasonal factor and previous states of error.

3.3 Two-stage estimation

Suppose that a solar-based time series, $\{y_{\bar{s},t}\}, t = 1, 2, \dots, n$, consists of daily and annual seasonal cycles with frequency m and $365.256 \times m$, respectively, and $\{x_{\bar{s},1}, x_{\bar{s},2}, \dots, x_{\bar{s},[n/m]}\}$ is an explanatory variable (daily data) which has a linear relationship with daily peak of the time series $\{A_{\bar{s},1}, A_{\bar{s},2}, \dots, A_{\bar{s},[n/m]}\}$. For example, a half-hourly solar PV gen-

eration time series has daily and annual frequency equal to 48 and 17532.288, respectively.

The traditional procedure of exponential smoothing requires estimation of the smoothing parameters and the initial states. However, the spatial exponential smoothing model is simply the merger between exponential smoothing and spatial/covariate model with the different time intervals, i.e., the exponential smoothing fitted by high-frequency time series in small time units such as every minute, half-hourly, hourly, whereas the regression model fitted by daily observations.

The likelihood function of parameters including the smoothing parameters, initial values, and regression coefficients given the observations and covariates is defined by

$$\begin{aligned}
 L(\boldsymbol{\theta}, \mathbf{z}_0 | \mathbf{y}_{\vec{s}}, \mathbf{x}_{\vec{s}}) &= \prod_{t=1}^n f(y_{\vec{s},t} | x_{\vec{s},\lceil t/m \rceil}, \boldsymbol{\theta}, \mathbf{z}_0) \\
 &= \prod_{j=1}^{\lceil n/m \rceil} \prod_{i=1}^m f(y_{\vec{s},m(j-1)+i} | x_{\vec{s},j}, A_{\vec{s},j}, \boldsymbol{\theta}, \mathbf{z}_0), \text{ if } m(j-1) + i \leq n \\
 &= \prod_{j=1}^{\lceil n/m \rceil} \prod_{i=1}^m f(y_{\vec{s},m(j-1)+i} | A_{\vec{s},j}, \boldsymbol{\theta} \setminus \boldsymbol{\theta}_1, \mathbf{z}_0) f(A_{\vec{s},j} | x_{\vec{s},j}, \boldsymbol{\theta}_1)
 \end{aligned}$$

where

$$\mathbf{z}_0 = \{l_0, b_0\}, \boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \sigma^2\}, \boldsymbol{\theta}_1 = \{\beta_0, \beta_1, \dots, \beta_k, \sigma_a^2\}, \boldsymbol{\theta}_2 = \{\alpha, v, \gamma\}.$$

Observe that the full likelihood, which is a product of two conditional functions with different time intervals, is too computationally expensive to estimate using direct existing methods. Therefore, we propose to approximate the model parameters under two-stage approaches as shown in the flowchart in Figure 3.2. Also, the historical time series data set used in the proposed model is divided into two segments: training data for model fitting and test data set for validation.

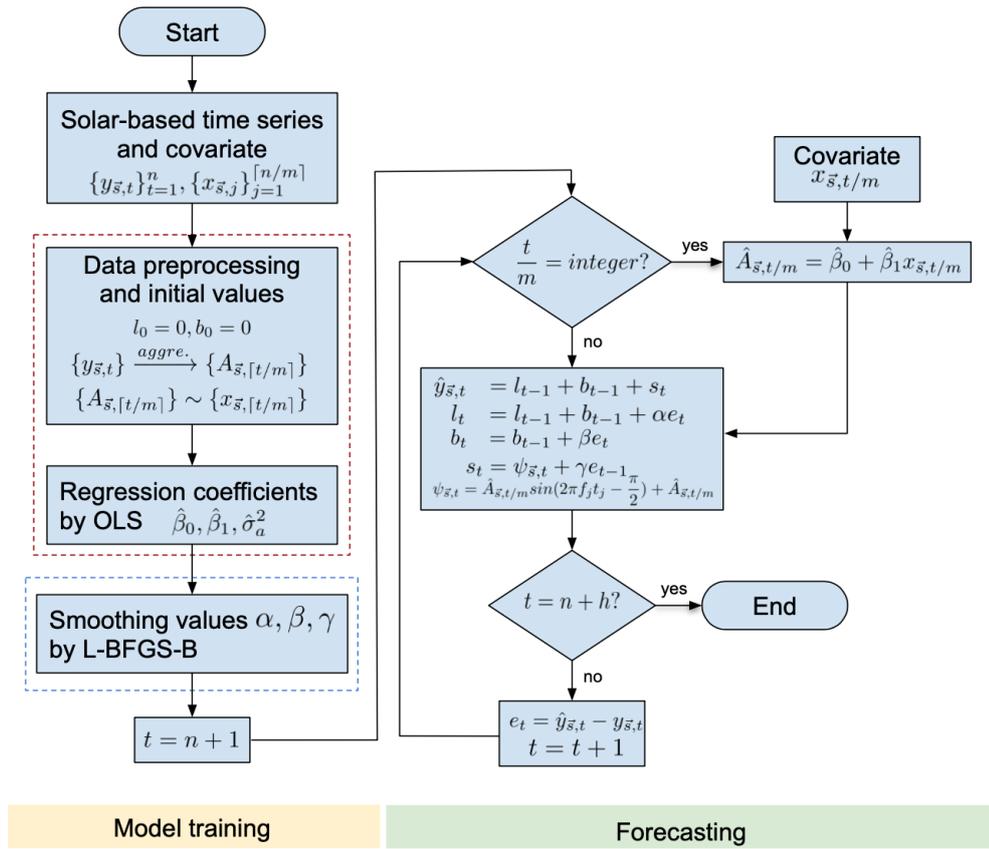


Figure 3.2: The spatial exponential smoothing model building process

3.3.1 Stage 1: Estimation of regression coefficients

In the first stage, the multiple linear regression model, $\tilde{A}_{\vec{s},[t/m]}$ defined in Equation 3.1 can be rewrite in matrix notation as follows

$$\tilde{\mathbf{A}}_{\vec{s}} = \mathbf{X}_{\vec{s}} \boldsymbol{\beta} + \boldsymbol{\eta}, \boldsymbol{\eta} \sim N(0, \sigma_a^2 I)$$

where

$$\tilde{\mathbf{A}}_{\vec{s}} = \begin{bmatrix} \tilde{A}_{\vec{s},1} \\ \tilde{A}_{\vec{s},2} \\ \vdots \\ \tilde{A}_{\vec{s},[t/m]} \end{bmatrix}, \mathbf{X}_{\vec{s}} = \begin{bmatrix} 1 & x_{\vec{s},11} & \cdots & x_{\vec{s},1k} \\ 1 & x_{\vec{s},21} & \cdots & x_{\vec{s},2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{\vec{s},[t/m]1} & \cdots & x_{\vec{s},[t/m]k} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \boldsymbol{\eta} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_{[t/m]} \end{bmatrix}$$

We wish to approximate the regression coefficients, $\boldsymbol{\beta}$, which minimise the resid-

uals sum of the squares, $S(\boldsymbol{\beta}) = \sum_{j=1}^{\lceil t/m \rceil} \eta_j^2 = \boldsymbol{\eta}'\boldsymbol{\eta} = (\tilde{\mathbf{A}}_{\bar{s}} - \mathbf{X}_{\bar{s}}\boldsymbol{\beta})'(\tilde{\mathbf{A}}_{\bar{s}} - \mathbf{X}_{\bar{s}}\boldsymbol{\beta})$, i.e. minimising the amount of the covariates variance that is not explained by the regression model, namely using an ordinary least squares (OLS) method.

Thus, the least squares estimators of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'_{\bar{s}}\mathbf{X}_{\bar{s}})^{-1}(\mathbf{X}'_{\bar{s}}\tilde{\mathbf{A}}_{\bar{s}}), \quad (3.3)$$

such that $\mathbb{E}[\hat{\boldsymbol{\beta}}|\mathbf{X}_{\bar{s}}] = \boldsymbol{\beta}$, $\mathbb{V}[\hat{\boldsymbol{\beta}}|\mathbf{X}_{\bar{s}}] = \sigma_a^2(\mathbf{X}'_{\bar{s}}\mathbf{X}_{\bar{s}})^{-1}$, and an estimator of the population variance of the covariates, σ_a^2 , is estimated by substituting $\hat{\boldsymbol{\beta}}$ into the residual sum of square, then

$$SS_{Res} = (\tilde{\mathbf{A}}_{\bar{s}} - \mathbf{X}_{\bar{s}}\hat{\boldsymbol{\beta}})'(\tilde{\mathbf{A}}_{\bar{s}} - \mathbf{X}_{\bar{s}}\hat{\boldsymbol{\beta}}) = \tilde{\mathbf{A}}'_{\bar{s}}\tilde{\mathbf{A}}_{\bar{s}} - \hat{\boldsymbol{\beta}}'\mathbf{X}'_{\bar{s}}\tilde{\mathbf{A}}_{\bar{s}}.$$

Therefore, the least squares estimators of σ_a^2 is

$$\hat{\sigma}_a^2 = SS_{Res}/(\lceil t/m \rceil - k + 1). \quad (3.4)$$

This stage of the application would predict the expected value of daily peak solar-based time series pattern via sine wave model corresponding to real weather throughout the year via daily expected weather conditions.

3.3.2 Stage 2: Determining the optimal values of smoothing parameters

Regarding the sine wave cycle as daily pattern, there is no significant change in local level and trend components, and also solar-based time series at the beginning of the period (midnight) should be zero except for the places that are able to harvest the midnight sunlight in the summer. Therefore, the initial state values $\mathbf{z}_0 = \{l_0, b_0\}$ are assumed to be fixed with zero values. Substituting (3.3)-(3.4) and initial values \mathbf{z}_0 into

the likelihood function, we get the reduced form of the likelihood as

$$\begin{aligned}
L^*(\alpha, \nu, \gamma, \sigma^2) &= \prod_{j=1}^{\lceil n/m \rceil} \prod_{i=1}^m f(y_{\bar{s}, m(j-1)+i} | A_{\bar{s}, j}, \alpha, \nu, \gamma, \sigma^2) f(A_{\bar{s}, j} | x_{\bar{s}, j}, \hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k, \hat{\sigma}_a^2) \\
&= \prod_{t=1}^n f(y_{\bar{s}, t} | A_{\bar{s}, \lceil t/m \rceil}, \alpha, \nu, \gamma, \sigma^2) \\
&= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(\frac{-1}{2\sigma^2} \sum_{t=1}^n \varepsilon_t^2\right)
\end{aligned}$$

Focusing on the population variance σ^2 by maximising the likelihood function, thus

$$\hat{\sigma}^2 = n^{-1} \sum_{t=1}^n \varepsilon_t^2.$$

Therefore, smoothing parameters, α, ν, γ , will be subsequently estimated as arbitrary parameters using gradient projection methods for solving (0,1)-bound constrained nonlinear optimisation problem. Refer to the Broyden-Fletcher-Goldfarb-Shanno algorithm (BFGS) is a method that tries to solve a general nonlinear optimization problem without any constraints. Thus, a limited memory BFGS method, also known as L-BFGS-B is a version of BFGS that lets “box” constraints, which are restrictions of the form $0 \leq \alpha, \nu, \gamma \leq 1$, be suitable for estimating all (0,1)-bound smoothing parameters. [63].

3.4 Numerical study

3.4.1 Measuring forecast accuracy

As mentioned above, we need to know how the performance of the proposed approach and relevant methods when applied in real-life situations for forecasting pv generation one-step ahead up to one-day ahead. The diagnostic checks in this study look at the residuals of both training and test data sets. The one-step-ahead forecast errors for

training data are given by

$$e_t = y_t - \hat{y}_{t-1}(1) = y_t - \hat{y}_t.$$

To evaluate various the aspects of the performance such as the overall forecast bias, the error characteristics and values, we compared forecast errors using the mean error (abbreviated ME), mean absolute error (MAE), and root mean square error (RMSE), defined as

$$\begin{aligned} \text{ME} &= \frac{1}{n} \sum_{t=1}^n e_t \\ \text{MAE} &= \frac{1}{n} \sum_{t=1}^n |e_t|, \\ \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}. \end{aligned}$$

3.4.2 The empirical data

To explore the possibility of applying the proposed approach to the solar-based time series, a description of the characteristics of the effect of the sun on earth and how the spatial component can be utilised is provided in this section.

According to the regression model in equation 3.2, the accuracy of the proposed model mainly relies on the seasonal component as the direction of daily patterns of the data. Thus, this study focuses on what daily measured meteorological conditions values should be considered to predict the characteristic of photovoltaics output. Based on statistical approaches, there is a strong relationship between the pattern of solar PV output and solar irradiance in a clear-sky day [61]. However, it is not easy to measure solar power in many areas due to the cost and maintenance of the measuring equipment. At this point, the relationship between global solar radiation and sunshine duration plays an important role in estimating the global solar power [43, 44]. Several expressions have been used to estimate global solar radiation from sunshine

hours [59, 64–66]. Moreover, atmospheric conditions such as cloud cover variations are also the most important condition to reduce incoming radiation by up to 90 percent by reflecting the radiation back into space which mainly affects the daily profiles of pv generation [67–69]. Therefore, cloud cover variations and sunshine duration will be considered as explanatory variables in regression analysis. In this study, these regressors are assumed to be weather-related information obtained from other practical methods (see e.g. in [70, 71]).

In order to examine the performance of the proposed model in comparison with existing models, a data set of historical high-frequency pv time series and relevant meteorological information were derived from the PVOutput project¹ and some solar pv farm in the UK, and also sunshine hours provided by nearby weather station . Three different geographical coordinates as shown in Figure 3.3 were adopted in this numerical study.

3.4.3 Solar-PV Generation Data in Cambridge, East Midlands (United Kingdom)

From June 1 to August 31, 2013, there are 4,416 half-hourly PV generation observations from the Cambridge solar farm. The first two months are used for model training, and the remaining for evaluation. Also, daily sunshine duration as a weather-related covariate during the given period measured at the Digital Technology Group (DTG), the University of Cambridge, is used to reflect real weather as a second seasonal component of data.

The relationship between daily peak time series, which is the maximum amount of time series each day, and sunshine hours as modeled by the regression model in Equation 3.2, was found in the training estimation stage to be significant with a Pearson’s correlation coefficient equal to 0.4215.

¹the PV output data is available at <https://pvoutput.org>

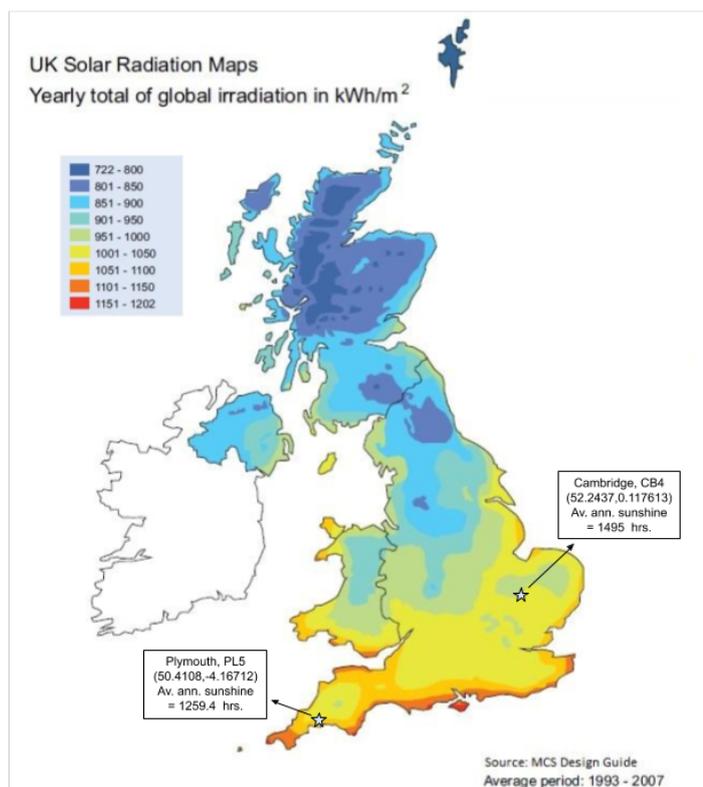


Figure 3.3: Solar radiation map of the UK (source: MCS Design Guide), and average annual sunshine hours of study points.

Table 3.1 displays the results of the performance evaluation of training for all models, including the simple seasonal exponential smoothing (SES) or Holt-Winters model, the autoregressive integrated moving average (ARIMA), the spatial exponential smoothing (Sp-Exp), and the TBATS model for one seasonal pattern. The number of parameters, the error measurement, and the amount of time spent computing are included in these results (unit: second). To evaluate the performance of the proposed model to the previously popular extremely short-term forecast models for forecasting high-frequency time series one day in advance using the test data set, the results are illustrated separately into two parts based on forecasts horizon: one-step ahead and 24-hour(one-period)-ahead forecasts shown in Table 3.2 and the graph puzzle (24-hour ahead of a particular day) of the test sample forecasts in Figure 3.4.

The limitation of the high-frequency time series is the amount of the data produced. It is not enough to detect second seasonality by ARIMA and TBATS. These methods

require at least two years of data, i.e. the methods need $24 \times 365.25 \times 2 = 17,532$ observations at least. This numerical example, therefore, compares the results in terms of short-term forecasts up to one-period-ahead for high-frequency data. It is noticeable that SES achieved the results by giving a small number of parameters and mean errors (ME) and reducing the computational time for model fitting. TBATS is likely to be the best training model but holds the biggest number of parameters, excluding seed values. Moreover, it suits the data in the given period.

Table 3.1: The performance results of short-term forecasts models involved for solar-based time series.

Model	Parameters	Error measurement			Comp. time (sec.)
		ME	MAE	RMSE	
SES	4	-2.865	88.252	189.772	0.021
ARIMA	6	8.487	81.312	163.315	80.333
TBATS	9	11.761	68.179	132.631	30.385
Sp-Exp ^a	7	0.034	72.765	138.703	184.688

^a Smoothing parameters searched by L-BFGS-B algorithm, and coefficient of determination (R^2) for regression model equals 0.5146

The problem with point of Sp-Exp is its high computational burden under the two-stage estimation approach spent which is smoothing search operations. Therefore, Sp-Exp has the highest estimation time, but it is better than double-seasonal methods like ARIMA and TBATS for one-step-ahead forecasting. Moreover, the performance of Sp-Exp is highly dependent on the correlation of the weather-related input and daily magnitude. Thus, the forecasting accuracy of the proposed model can be improved by entering more covariates [61], which will be shown in the next application.

3.4.4 Solar-PV Generation Data in Plymouth, Southwest England (United Kingdom)

This application examines the model with the time series, which has higher frequency than the previous example, on significantly different location, and adds more weather-related conditions, such as cloud cover. Cloud is a mass of particle in the sky that

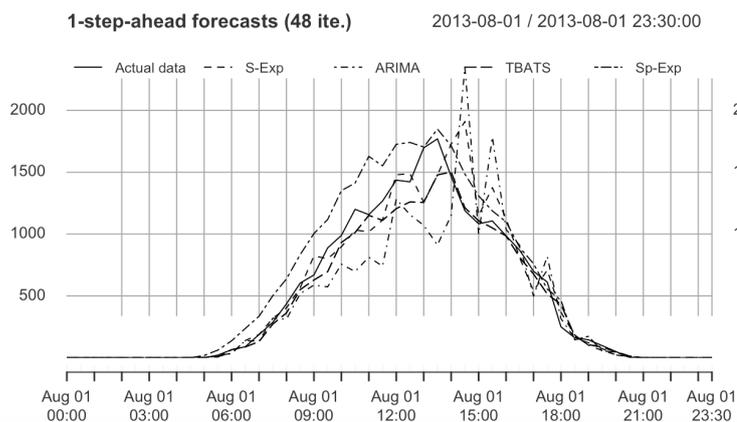
Table 3.2: The comparative assessment of all forecasting methods for one-step-ahead and one-day-ahead forecasts during 1st - 31st August 2013.

Model	1-step-ahead (1488 iterations)			1-day(48s)-ahead (31 iterations)			Comp.time/day (sec.)
	ME	MAE	RMSE	ME	MAE	RMSE	
SES	9.824	100.882	190.805	140.380	190.006	333.918	0.043
ARIMA	11.984	105.543	202.938	36.692	159.749	280.547	163.399
TBATS	-2.179	191.154	379.944	91.980	167.132	292.645	6.980
Sp-Exp	0.279	93.918	167.820	-138.513	205.882	326.492	0.751

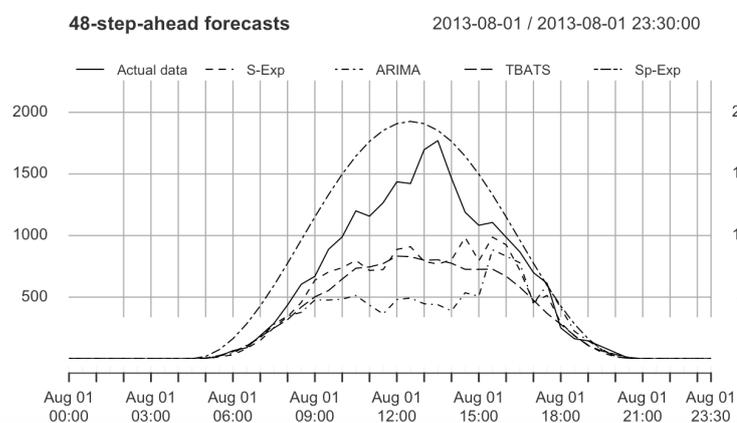
directly absorb solar power before reaching the earth's surface cloud cover was considered in the regression model.

The historical 15-minute PV output time series of 5.800 kilowatts (kW) solar PV farm with cloud conditions in Plymouth, PL5 (coordinates in decimal: 52.2437, 0.117613) during 1st April - 30th June 2019 is provided by the PV-Output project. The 5-km gridded sunshine duration data from Plymouth Live Weather Station covers approximately a solar farm. Therefore, sunshine hours collected at the station can be used [59] to forecast daily peak PV generation. In addition, cloud conditions during the periods are measured at four levels of cover (clear sky, partly cloudy, mostly cloudy, and cloudy). As can be seen in Table 3.3, the covariate model used for magnitude prediction accounts for 78.09% of the overall variance (R-squared values 0.7809) based on sunshine duration and cloud cover variables.

It is noticeable that the proposed model still keeps less processing time than ARIMA and TBATS while error measurements have no significant difference for one-period-ahead forecasts as shown in Table 3.3. Besides, a five-days example in Figure 3.5 displays the role of the sine wave model used to predict the expected daily PV generation.



(a) One-step ahead forecasts



(b) One-day (48-step) ahead forecasts

Figure 3.4: Comparing the one-step ahead forecasts (48 iterations on 1 August 2013) and one-period (1 August 2013): actual data (solid), SES(dashed), ARIMA (dot-dash), TBATS (longdash), Sp-Exp (twodash).

Table 3.3: The estimation results of short-term forecasts models involved for solar-based time series

Model	Model training			One-day-ahead (30 iterations)			Comp. time / day (sec.)
	ME	MAE	RMSE	ME	MAE	RMSE	
SES	-6.522	209.375	407.141	-29.252	503.629	819.350	0.437
ARIMA	3.012	273.808	551.466	12.454	414.712	720.056	33.797
TBATS	51.846	194.049	365.302	160.761	350.065	595.442	9.731
Sp-Exp ^a	5.247	187.024	353.712	-252.847	469.656	735.845	2.495

^a The coefficient of determination (R^2) for regression model is equal to 0.7809.

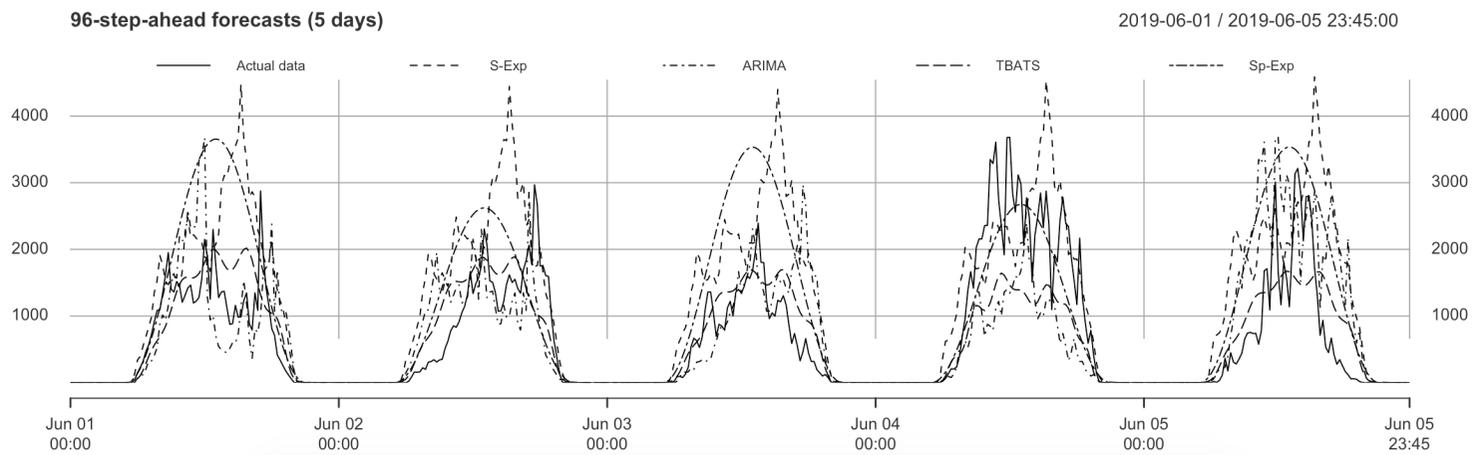


Figure 3.5: Five days of day-ahead forecasts: actual data (solid), SES(dashed), ARIMA (dotdash), TBATS (longdash), Sp-Exp (twodash).

Chapter 4

Its Application via Stochastic Programming

This chapter focuses on an application of the developed forecasting model to optimize the energy storage utilization in a home with a solar photovoltaic (PV) system and an electric vehicle (EV). Consideration is given to developing a two-stage linear stochastic programming (SP) model to propose and investigate the different residential energy connections of electric vehicle batteries (EVB) to the residential solar photovoltaic (PV) system and to home appliances in order to determine the efficacy of deploying EVB in future households. The novel forecasting model (as demonstrated by the equation 3.1) plays a crucial role in simulating potential 24-hour-ahead solar-panel outputs as one of the system parameters in order to achieve an optimal solution of the system under solar PV generation uncertainty. In addition, the CPLEX optimization studio is utilized to address the large-scale mathematical programming problem and conduct stochastic numerical investigations.

4.1 Renewable energy for households

For decades, renewable energy and energy storage technologies have been developed to conserve natural resources and minimize global carbon emissions from fossil fuels

[5–7]. Numerous global energy leaders tend to shift output away from fossil fuels and toward renewables such as solar, wind, and biomass, which are frequently referred to as clean energies. Solar power and wind turbine systems, for example, are highly reliant on the weather conditions and so are subject to great uncertainty. The fact that electricity generation from natural resources is largely unregulated by people and may not fit the demand pattern is one of the greatest obstacles. According to numerous sources, energy storage increases the fraction of energy consumption at the site of generation, i.e. energy self-consumption [72–75]. The key to success is in the approach to forecasting and the utilization of energy storage to carry energy from the generation periods to the consumption periods.

More than a quarter of the world's electricity consumption was met by renewable sources by the end of 2018, and their capacity grew by 8 percent, led by wind and solar [7]. Wind power, bio-energy, solar photovoltaic (PV), and hydropower account for the majority of the renewable electricity produced in the United Kingdom in 2019 (35.8 percent of the total). Renewable energy sources are characterized at a specific location by taking into account the existing energy flows and the spatial & temporal variations [76]. In the United Kingdom, solar panels and small wind turbines are examples of suitable residential renewable energy deployments [10, 77]. Particularly, solar PV adoption has increased rapidly over the past few years as the International Energy Agency (IEA) declared that solar PV power is the new king of electricity because it is consistently cheaper than fossil-fuel power plants [78].

It is common knowledge that significant uncertainties exist in renewable energy generation, such as solar power and wind turbine systems, which are highly dependent on weather conditions. In addition, this type of renewable energy generation does not typically correspond to the daily consumption patterns of households. Various home appliances, such as kitchen and cooking appliances, washing machines and tumble dryers, refrigerators and freezers, microwaves, ovens, toasters, humidifiers, coffee makers, televisions, DVD players, video games, telephones, laptops, etc., are avail-

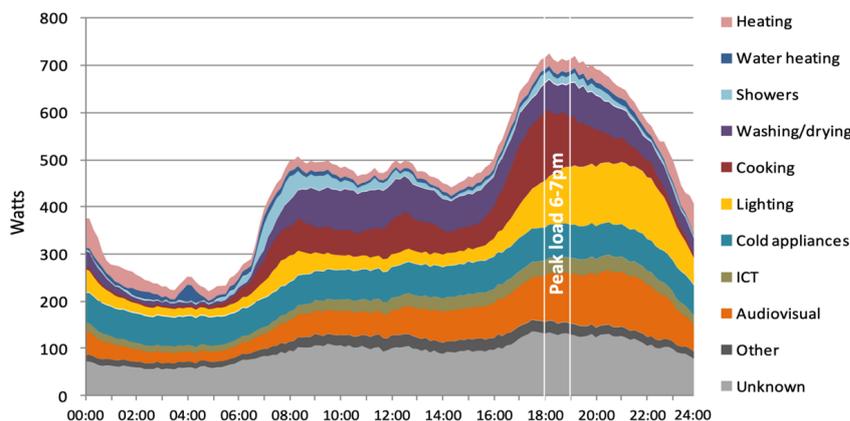


Figure 4.1: The average household electricity consumption profile for a typical UK home (Watts) [3]

able for household electric end-uses. The electricity consumption of these end-uses is contingent upon lighting and dwelling size, local temperature, utility cost, consumer lifestyles, etc. Figure 4.1 illustrates the usage profile of home appliances and peak periods throughout the day based on the survey results of 250 British households in 2010-2011 [3, 52]. It can be observed that daily electricity profiles and peak demand occur between 6 and 9 p.m., which corresponds to the time for cooking, lighting, and audiovisual demands after work hours, which also corresponds to the time of high (peak load) prices. One way to reduce electricity costs is to purchase electricity during off-peak hours and transport it to meet peak demand, while another is to maximize self-consumption when a home has its own renewable energy generators such as solar panels. By installing electricity storage, both objectives can be attained (e.g. batteries).

According to numerous publications [79–83], self-consumption, i.e., the use of electricity on the site of its generation, is one of the most efficient forms of energy consumption, and its level can be enhanced by the deployment of energy storage. As a result of technological and material advancements, solar PV panels and energy storage have become increasingly affordable [79–81, 84]. It can be inferred that the quality of weather forecasting and the optimal use of energy storage are the primary factors influencing the deployment of renewable energy at the residential level.

Alternatively, electric vehicles (EVs) play a crucial role in achieving the environ-

mental objectives of lowering local air pollution and combating climate change. In many nations, efforts are made to promote the development and use of electric vehicles (EVs) in order to lower pollution levels by establishing EV policy support measures in various forms of government policies, such as financial incentives, tax breaks, etc. [85, 86]. The global market for electric vehicles (EVs) developed significantly over the past decade, fueled by favorable policies and technology developments at all levels of government [87]. In addition to these policies, the fact that an EV is powered by a battery and the driver takes it home with them leads to discussions about using the EV battery to power home electronics, also known as EV-to-house power supply. In 2021, for instance, Mitsubishi Motors Corporation and the Electricity Generating Authority of Thailand have agreed to build a system that would allow electric vehicles to power homes for other purposes [88]. In addition, Tesla's battery supplier, a leader in energy storage and EV technology, has announced plans to introduce a cheaper battery with a million-mile lifespan [89].

Based on the analysis above, it is critical to have an optimal control system to coordinate the operations of renewable energy resources and demand, given the uncertain distributed renewable generation (e.g., micro wind turbines, on-roof solar panels for local communities and households) and various consumption needs. Home energy management energy management system (HEMS), which could serve in this capacity, can be viewed as a demand response tool conceptualized on the basis of solar energy management system [90]. Existing literature on HEMS has presented a number of optimization models and techniques, such as stochastic dynamic programming (SDP) [91], convex programming [92], mixed integer linear programming (MILP) [93] and stochastic programming (SP) [94, 95] to optimize energy usage in smart homes with renewable energy, energy storage system (ESS), and Plug-in electric (PEV).

Schram et al. (2018), for instance, proposed a simulation model for optimizing the residential battery with various household load profiles and PV systems. In addition, stochastic management of energy storage is one of the crucial areas to investigate in

order to reveal the potential of renewable energy sources [96]. Mathematical optimization, model predictive control, and heuristic control are used to improve the energy efficiency of household appliances [97]. The operating policies, including vehicle-to-home supply and peak-load shifting [93], have been proposed and implemented in an effective manner.

Most existing works design and optimize a system based on the communication between devices and smart meters for real-time monitoring of electricity consumption in households with real-time pricing (RTP). In this study, we assume that the hourly charging rate of batteries can be pre-set in smart meters based on a forecast of weather conditions, electricity prices, and demand for the following 24 hours.

The supportive policies of energy storage systems, such as electric vehicle batteries (EVBs) and residential energy storage, are viewed as a means of managing day-ahead electricity consumption. In addition to the difficulty of managing HEMS as a demand response tool to improve the energy load consumption based on the price of electricity, solar PV forecasting is crucial for coping with parameter system uncertainty.

4.2 Stochastic programming

Stochastic programming is a way to model optimization problems where there are a lot of things we do not know. In deterministic optimization problems, all of the parameters are known, but in real-world problems, almost all of the parameters are unknown. Robust optimization is a way to solve problems when only a small number of parameters are known. The goal here is to find a solution that works with all of these pieces of information and is better in some way. In a similar way, stochastic programming models take advantage of the fact that the data's probability distributions can be known or estimated. Here, the goal is to find a policy that works for all (or almost all) possible data instances and makes the expectation of a function of decisions and random variables as good as possible. In a broader sense, these models are created, numerically

solved, and analysed in such a way that a decision-maker can learn something useful in a variety of applications ranging from finance to transportation to energy optimization.

Two-stage stochastic programming [98]

Two-stage stochastic programming involves making decisions for *two different time periods* based on a set of random parameters determined by either previous experience or the results of some kind of survey. The objective function for formulating two-stage stochastic programming with fixed recourse consists of two parts: the first-stage forecast and the second-stage fixed decisions based on the results of the experiment. Both parts are included in the objective function. Instead of making a few adjustments to the requirements and technological assets, the constraints are more like standard optimization techniques.

The fundamental concept behind two-stage stochastic programming is that the most appropriate choices should be formulated on the basis of the information that is readily available at the time that the choices are being formulated, rather than on the basis of future observations. In the field of stochastic programming, the two-stage formulation is very widespread. The following is an example of the general formulation of a stochastic programming problem with two stages:

$$\min_{x \in X} \{g(x) = f(x) + E[Q(x, \xi)]\},$$

where $Q(x, \xi)$ is the optimal value of the second-stage problem,

$$\min_y \{q(x, \xi) | T(\xi) + W(\xi)y = h(\xi)\}.$$

The classical two-stage linear stochastic programming (SP) models can be performed as

$$\min_{x \in R^n} g(x) = c^T x + E[Q(x, \xi)],$$

subject to

$$\begin{aligned} Ax &= b \\ x &\geq 0, \end{aligned}$$

where $x \in R^n$ is the first-stage decision variable vector, and $Q(x, \xi)$ is the optimal value of the second-stage problem, s.t.

$$\min_{y \in R^m} q(\xi)^T y,$$

subject to

$$\begin{aligned} T(\xi)x + W(\xi)y &= h(\xi) \\ y &\geq 0, \end{aligned}$$

where $y \in R^m$ is the second-stage decision variable vector, and $\xi(q, T, W, h)$ is a random vector of the data of the second-stage problem. In this formulation, the first step requires us to make a decision about x in the here-and-now, before the realisation of the uncertainty data ξ is known. In the second stage, once we have a realisation of ξ and are able to do so, we optimise our behaviour by finding a solution to an appropriate optimization problem.

4.3 Default setting and system parameters

Figure 4.2 represents the electricity flow inside a household that has a solar PV system, a home battery, and an EV. Demands are classified into three categories, i.e., EV, typical home appliances, and small appliances. It shows the default system where all devices are connected to the grid. A home battery is installed and connected to the solar panel to increase self-consumption. Electricity generated by solar panels or pulled from the

grid can directly supply home appliances and/or charge the home battery, while the home battery can supply the appliances as well. We assume that solar panels can directly supply home appliances without going through the home battery. A stochastic programming model will be developed in this work to explore the optimal usage of batteries under this connection setting under different solar PV generation levels and home battery capacities. The EV is driven by its own battery, which can be charged from the grid. Due to the voltage and existing restrictions, direct charge from the solar panel to the EV battery is prohibited by this default layout. In later discussions, however, we will lift this restriction to explore the potential usage of EV batteries in balancing demands and supply, and look into the possibility of replacing home batteries with EV batteries in a household with renewable energy supply.

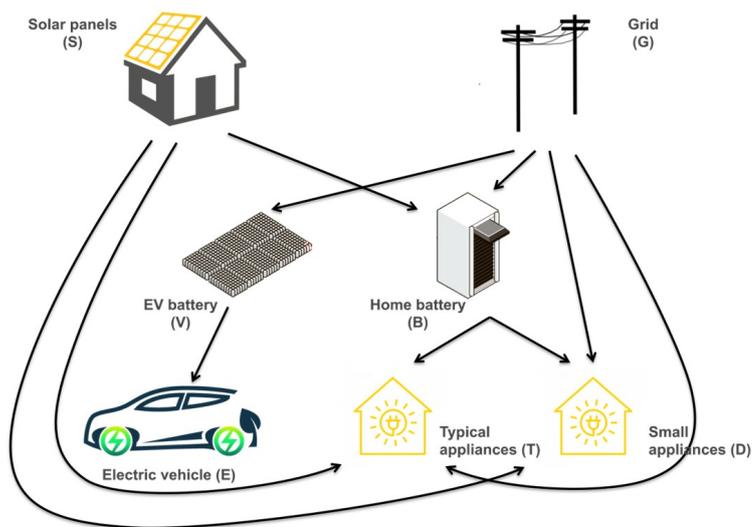


Figure 4.2: Graphical representation of home energy consumption and battery usage concepts

4.4 Day-ahead probabilistic PV generation forecast

The proposed forecasting model with exogenous variables at the day ahead horizon generates scenarios for the hourly output of peak PV generation over the next twenty-four hours.

The time-series methodology is one of the three major categories for micro-scale and intra-day resolutions in solar PV power forecasting, according to [99]. Particularly, exponential-based and ARIMA-based approaches are suitable for short-term solar PV planning [100–104]. The majority of these methods utilize meteorological data as exogenous variables, such as solar radiation, sunshine hour, weather conditions, cloud cover, etc. The intensity of solar radiation or sunlight, which is a direct input of PV through solar panels, reaches the earth's surface and has a strong correlation with satellite (images) data [45, 105], historical solar radiation data [95, 106], cloud conditions [40, 45–49], and sunshine duration [40–44]. It also depends on coordinates, season, time of day, cloud cover, and altitude, among other variables.

A model for forecasting proposed by Chitsuphaphan et al. (2020) [95, Section. 2(E)] incorporates local information, such as coordinate, and weather-related conditions, into the exponential-based method that can simulate solar PV output for the following day. To forecast the day-ahead solar PV output in micro-scale resolution for residential energy consumption as a stochastic variable, the cloud cover, which includes five levels such as sunny, partly cloudy, mostly cloudy, cloudy, and rainy, is considered as an independent variable (as an indicator variable that represents categorical data). A five-level categorical variable represents cloud cover. In the regression analysis (as a causal model), cloud conditions are transformed to be a 4-dimensional dummy variable $x_{\vec{s},1j}, x_{\vec{s},2j}, x_{\vec{s},3j}, x_{\vec{s},4j}$, with $(0, 0, 0, 0)$ standing for sunny (fine sky), $(1, 0, 0, 0)$ standing for partly cloudy, $(0, 1, 0, 0)$ standing for mostly cloudy, $(0, 0, 1, 0)$ standing for cloudy and $(0, 0, 0, 1)$ standing for rainy, where \vec{s} indicates the PV's coordinates (latitude and longitude) and j denotes the j th day.

Let $(P_{\vec{s},j}, \mathbf{x}_{\vec{s},j})$ represents a historical data set of the peak-PV generation and a cloud forecast vector, i.e. $\mathbf{x}_{\vec{s},j} = (x_{\vec{s},1j}, x_{\vec{s},2j}, \dots, x_{\vec{s},4j})$, at \vec{s} position on day j , and $A_{S|\mathbf{x}_{\vec{s},j}}^t$ denotes an amount of hourly- t solar electricity generation given \mathbf{x}_j . Therefore, the forecasting model by using five levels of cloudiness used to simulate the scenarios of

hourly electricity supply for second-stage problem of the SP model is

$$A_{S|\mathbf{x}_{\vec{s},j}}^t = \frac{\hat{P}_{\vec{s},j}}{2} \sin\left(2\pi f_j t_j - \frac{\pi}{2}\right) + \frac{\hat{P}_{\vec{s},j}}{2} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2), t \in [u_{\vec{s},j}, v_{\vec{s},j}],$$

$$= 0, \text{ elsewhere,}$$

$$P_{\vec{s},j} = \beta_0 + \beta_1 x_{\vec{s},1j} + \beta_2 x_{\vec{s},2j} + \beta_3 x_{\vec{s},3j} + \beta_4 x_{\vec{s},4j} + \eta_j, \eta_j \sim N(0, \sigma_a^2),$$

where $\hat{P}_{\vec{s},j}$ represents the magnitude forecast at \vec{s} -position on day j when $\mathbf{x}_{\vec{s},j}$ given, $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4$ denote regression coefficients, η_j denotes a normally distributed disturbance term, s.t. $\mathbb{E}[\eta_j] = 0, \mathbb{V}[\eta_j] = \sigma_a^2$, $u_{\vec{s},j}, v_{\vec{s},j}$ denote sunrise and sunset time index, respectively, f_j denotes a frequency of sine wave function s.t. $f_j = \frac{1}{n_j}, n_j = v_{\vec{s},j} - u_{\vec{s},j} + 1, t_j$ denotes a new time index on day j , s.t. $t_j = t - u_{\vec{s},j} + 1$.

Figure 4.3 displays forecasts for the next four days of the 24-hour outputs of PV generation based on the model that was applied given different cloudiness conditions on July 1, 3, 6, and 10, 2019 with partly cloudy, fine, cloudy, and costly cloudy, respectively. As can be seen, the hourly forecast for each day relies heavily on the sinusoidal function that is derived from the daily cloudiness forecast.

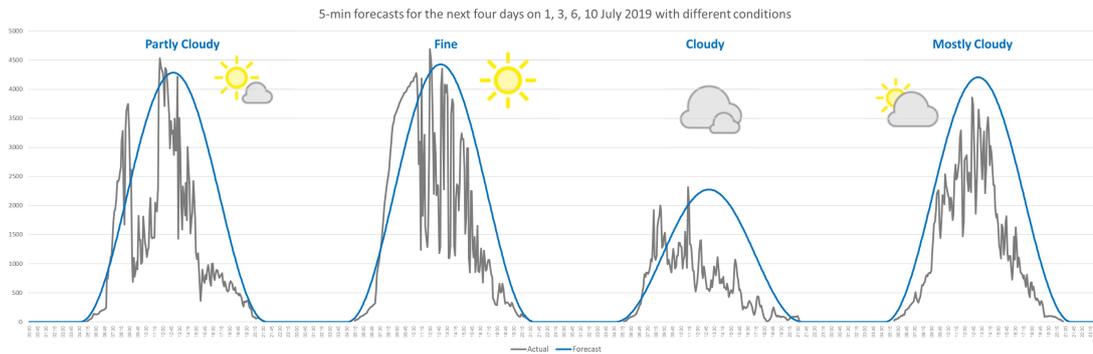


Figure 4.3: The 24-hour ahead solar PV output forecast when giving different states of cloud cover using real data for model training

The historical time series and cloud cover conditions (five levels of cloudiness) are serviced to download by the PVOutput platform (see <https://pvoutput.org/>). The data set is used to estimate the parameters of the forecasting model and simulate the scenarios afterward. In the UK, a 3 kW solar panel system, which requires at least 21

square metres of surface area to install, can sufficiently generate electricity for daily consumption of a small-to-medium household size (1 - 3 residents). Therefore, we will generate three sets of solar electricity scenarios ($A_S^\omega, \omega \in \Omega, s.t. |\Omega| = 20$) based on 1, 2, and 3 kW peak PV system.

4.5 Optimization model

4.5.1 Notations and parameters

To develop the two-stage SP models for household power systems with PV generation system and EV. Model parameters, decision variables and abbreviations are firstly declared to represent components of system programming as shown in Table 4.1.

4.5.2 System layouts and models

To understand the different roles played by each component of the household energy system, we look into different systems with respect to the availability of batteries and/or connections between batteries and devices. Table 4.2 summarizes the detailed systems we are examining with the connection links demonstrated in Figure 4.4. These scenarios are going to be compared with the default setting as displayed in Figure 4.2.

Note that these proposed systems take into account both the current applicability and the potential future development of solar PV and energy storage devices. The on-grid system is the simplest system for a home with solar PV and an EV, as it is connected to the grid and all consumption is drawn directly from the grid when needed. To investigate how batteries assist in transporting energy from the generation phase (cheap hours) to the demand phase (peak hours) so as to maximise self-consumption, the default system (Figure 4.2) is proposed, which reflects the standard configuration of a household with solar PV, an electric vehicle (EV), and home energy storage (battery) using state-of-the-art technology. The hybrid system, on the other hand, improves the default system by allowing power transmission from home battery to EV battery so

Table 4.1: Abbreviations and system notations

Indicators	Definition
S	solar PV generation system
G	regional electricity grid
E	electric vehicle battery (EVB)
B	home battery
V	electric vehicle
T	typical household appliances or major appliances used for routine housekeeping tasks such as cooking, washing laundry, or food preservation
D	small household appliances including portable/semi-portable machines such as microwave ovens, toasters, humidifiers, coffeemakers, and other electronic devices
Indices	
i	electricity supply sources, $i \in \{S, G\}$
j	power demand categories, $j \in \{V, T, D\}$
k	energy storage, $k \in \{E, B\}$
t	time intervals, $t \in \{1, 2, \dots, h\}$, where $h = 24$
Parameters	
P_i^t	[\\$/kW], cost of electricity from resource i at time t , $i \in \{S, G\}, t \in \{1, 2, \dots, h\}$
A_i^t	[kW], amount of electricity supply from source i at time t , $i \in \{S, G\}, t \in \{1, 2, \dots, h\}$
U_j^t	[kW], amount of electricity demand from category j at time t , $j \in \{V, T, D\}, t \in \{1, 2, \dots, h\}$
C_k	[kW], capacity of energy storage k , $k \in \{E, B\}$
δ^t	indicator showing if EV is available at home during interval t , $t \in \{1, 2, \dots, h\}$
ρ	conversion rate of power transmission
γ	energy selling price as a proportion to the lowest ToU tariff (when selling of energy is allowed)
Variables	
$x_{i,j}^t$	[kW], amount of electricity transmission from i to j at time t , $i \in \{S, G\}, j \in \{V, T, D\}, t \in \{1, 2, \dots, h\}$
$y_{i,k}^t$	[kW], amount of electricity transmission from i to k at time t , $i \in \{S, G\}, k \in \{E, B\}, t \in \{1, 2, \dots, h\}$
$w_{k,k'}^t$	[kW], amount of electricity transmission between energy storage at time t , $k, k' \in \{E, B\}, t \in \{1, 2, \dots, h\}$
$z_{k,j}^t$	[kW], amount of electricity transmission from k to j at time t , $k \in \{E, B\}, j \in \{V, T, D\}, t \in \{1, 2, \dots, h\}$
l_k^t	[kW], storage level of type k battery at time t , $k \in \{E, B\}, t \in \{1, 2, \dots, h\}$
s^t	[kW], amount of surplus electricity generated from solar PV during interval t , $t \in \{1, 2, \dots, h\}$

that excess solar PV generation will not be wasted if it exceeds standard household consumption.

On top of the structure of the hybrid system and the on-grid system, additional

Table 4.2: Household power system setups

Systems	Residential power systems with solar PV and EV			
	including B	allowing charge of E from B	allowing T& D supplied from E	allowing selling energy to the grid
Default System	Yes	No	No	No
(a) Hybrid System	Yes	Yes	No	No
(b) Hybrid System with EVB Supply	Yes	Yes	Yes	No
(c) On-grid System	No	No	No	Yes
(d) On-grid System with EVB Supply	No	No	Yes	Yes

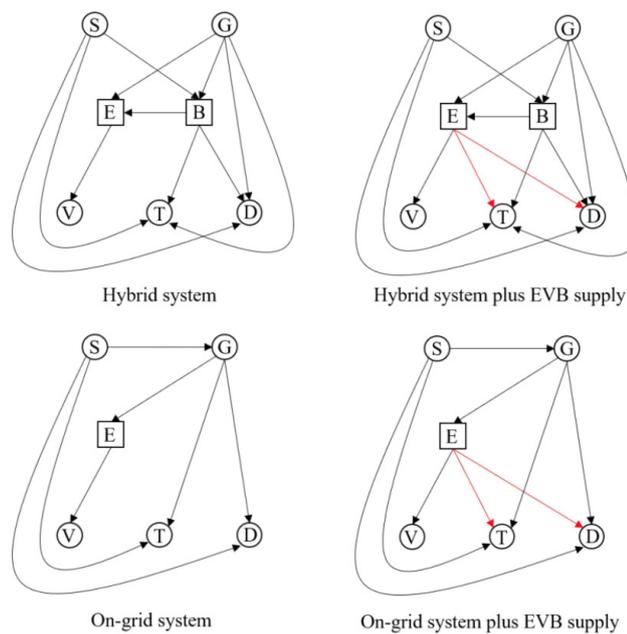


Figure 4.4: Alternative layouts of home electric systems

transmission options from the EV battery to home appliances are proposed for the other two systems. Incorporating/replacing the home battery with an EV battery is enabled by doing so. Given that EV is a global trend for the next generation of vehicles, these systems provide information on the optimal use of EV batteries in a broader household energy context. Note that all links to EV batteries are only valid when the vehicle is parked at home.

SP model for the default setting

Default system, as shown in Figure 4.2, is a modern PV system that combines solar panels and battery storage in one place. The electricity generated from solar panels can be consumed right away or be stored in home battery for later usage. This can improve self-consumption and ease the load and pressure on the grid. In addition, the home battery can also be used as a route to reduce energy bills by carrying electricity from off-peak hours to peak hours. We assume that the EV battery can only be charged at home from the grid, so essentially the two storage systems are not interacting with each other and the EV system is not influenced by the power supply from solar panels.

To simplify the problem structure, a two-stage SP is designed, with *the first-stage deciding proactively the best amount to pull from the grid into the energy storage (EV and home battery)*, and *the second stage reacting to the actual supply and demand to minimize the overall cost*.

Let $y_{G,k}^t, k \in \{E, B\}$ and l_E^t be the first-stage variables and $(x_{i,j}^t)^\omega, (z_{B,j}^t)^\omega, (y_{S,B}^t)^\omega, (l_B^t)^\omega, \forall i \in \{S, G\}, \forall j \in \{T, D\}, \forall k, \forall t, \forall \omega$ be second-stage recourse variables, where ω denotes the outcome of the random weather condition. The objective function of the two-stage SP model to minimise total daily electricity cost under weather conditions is expressed by

$$\min \sum_t \sum_k P_G^t y_{G,k}^t + \mathbb{E}_{\omega \in \Omega} [\varphi(y_{G,k}^t, \omega)],$$

s.t.

EV battery flows (BEV-to-EV linkage shows energy transfer during travel time t):

$$\rho \delta^t y_{G,E}^t - U_E^t = l_E^t - l_E^{t-1}, \forall t,$$

EV battery capacity (BEV level at node B in period t lies between 0 and capacity):

$$0 \leq l_E^t \leq C_E, \forall t,$$

EV battery daily power cycle (Start the day with the same energy at node E):

$$l_E^{24} = l_E^0$$

Non-negativity (Non-negative grid-to-battery electricity transmission is prohibited):

$$y_{G,k}^t \geq 0, \forall k, \forall t,$$

where $\varphi(y_{G,k}^t, \omega)$ represents the optimal objective of the second-stage problem under scenario ω . The second-stage problem can thus be written as

$$\varphi(y_{G,k}^t, \omega) = \min \sum_t \sum_{j \in \{T,D\}} P_G^t(x_{G,j}^t)^\omega,$$

s.t.

Solar PV supply (The power used to run appliances and stored in a home battery doesn't exceed a solar PV system's output in period t under the scenario):

$$\sum_{j \in \{T,D\}} (x_{S,j}^t)^\omega + (y_{S,B}^t)^\omega \leq (A_S^t)^\omega, \forall t, \forall \omega$$

Home battery flow (Total pull and push electricity at node B in period t under the scenario):

$$\rho [y_{G,B}^t + (y_{S,B}^t)^\omega] - \sum_{j \in \{T,D\}} (z_{B,j}^t)^\omega = (l_B^t)^\omega - (l_B^{t-1})^\omega, \forall t, \forall \omega$$

Electricity demands (In period t , total electricity from energy suppliers and home batteries meets home appliance demand at nodes T and D under the scenario):

$$\sum_{i \in \{S, G\}} (x_{i,j}^t)^\omega + (z_{B,j}^t)^\omega = U_j^t, \forall j \in \{T, D\}, \forall t, \forall \omega$$

Home battery supply (Under the scenario, a home battery's energy supply cannot exceed battery level in period t):

$$\sum_{j \in \{T, D\}} (z_{B,j}^t)^\omega \leq (l_B^t)^\omega, \forall \omega$$

Home battery daily power cycle (start the day with the same energy at node B under the scenario):

$$(l_B^{24})^\omega = (l_B^0)^\omega, \forall \omega$$

Home battery capacity (Home battery level at node B in period t under the scenario lies between 0 and capacity):

$$0 \leq (l_B^t)^\omega \leq C_B, \forall t, \forall \omega$$

Non-negativity (The transmission of electricity between nodes is greater than or equal to zero in period t under the scenario):

$$(y_{S,B}^t)^\omega \geq 0, \forall t, \forall \omega$$

$$(x_{i,j}^t)^\omega \geq 0, \forall i, \forall j \in \{T, D\}, \forall t, \forall \omega$$

$$(z_{B,j}^t)^\omega \geq 0, \forall j \in \{T, D\}, \forall t, \forall \omega$$

where $(A_S^t)^\omega$ represents the 24-hour PV output that are generated by the forecasting model in Section 4.4, which only become available in second stage.

SP model for (b) Hybrid System with EVB Supply

(b) Hybrid System with EVB Supply is an extension of the hybrid system by allowing home appliances be supplied from the electric vehicle battery (EVB). This system is motivated by the fact that normally the house occupiers come back home together with the EV, so that the EVB connection is available during the peak demand hours. By pushing remaining energy from EVB to supply home appliances when parking at home, one can further reduce the electricity purchase from the grid during peak rate hours. This would enable the maximum usage of the EV battery as an energy storage, which is available to be charged during the off-peak night hours and to be discharged to supply home appliances during the peak hours. This scenario should work well when the home battery is relatively small. Major differences from system (a) lie in three constraints:

EV battery flow:

$$\rho \delta^t [y_{G,E}^t + (w_{B,E}^t)^\omega] - \sum_{j \in \{T,D\}} \delta^t (z_{E,j}^t)^\omega - U_E^t = (l_E^t)^\omega - (l_E^{t-1})^\omega, \forall t, \forall \omega$$

EV battery supply:

$$\sum_{j \in \{T,D\}} \delta^t (z_{E,j}^t)^\omega + U_E^t \leq (l_E^t)^\omega, \forall t, \forall \omega$$

Electricity demands:

$$\sum_{i \in \{S,G\}} (x_{i,j}^t)^\omega + (z_{B,j}^t)^\omega + \delta^t (z_{E,j}^t)^\omega = U_j^t, \forall j \in \{T, D\}, \forall t, \forall \omega$$

where $(z_{E,j}^t)^\omega \geq 0, j \in \{T, D\}$ denotes the electricity flow from EV battery to home appliances in hour t .

SP model for (c) On-grid System

(c) On-grid System is a benchmark system which has no home battery installed. The excess solar electricity after supplying home appliances at time t , denoted by s^t , can be uploaded to the grid to get power credits. This is shown by the link between the solar panel and the grids in Figure 4.4. Provided that the surplus can be sold back to the grid at a discounted price with discount rate γ , the objective function for second stage problem has to be updated to include the income.

$$\min \sum_t P_G^t y_{G,E}^t + \mathbb{E}_{\omega \in \Omega} [\varphi(y_{G,E}^t, \omega)],$$

s.t.

EV battery flow:

$$\rho \delta^t y_{G,E}^t - U_E^t = l_E^t - l_E^{t-1}, \forall t,$$

EV battery daily power cycle:

$$l_E^{24} = l_E^0,$$

EV battery capacity:

$$0 \leq y_{G,E}^t \leq C_E, \forall t,$$

with the second-stage problem written as

$$\varphi(y_{G,E}^t, \omega) = \min \sum_t \left\{ P_G^t \sum_{j \in \{T,D\}} (x_{G,j}^t)^\omega - \gamma P_G^t (s^t)^\omega \right\},$$

s.t.

Solar PV supply:

$$\sum_{j \in \{T, D\}} (x_{S,j}^t)^\omega + (s^t)^\omega \leq (A_S^t)^\omega, \forall t, \forall \omega$$

Electricity demands:

$$\sum_{i \in \{S, G\}} (x_{i,j}^t)^\omega = U_j^t, \forall j \in \{T, D\}, \forall t, \forall \omega$$

Non-negativity:

$$(x_{i,j}^t)^\omega \geq 0, \forall i, \forall j \in \{T, D\}, \forall t, \forall \omega$$

$$(s^t)^\omega \geq 0, \forall t, \forall \omega$$

where $(s^t)^\omega \geq 0$ denotes the surplus electricity generated from solar panels in hour t . Note that as there is no storage available, this surplus has to be uploaded to the grid, disregard whether this uploading is paid or not.

SP model for (d) On-grid System with EVB Supply

(d) On-grid System with EVB Supply is an extension of the on-grid system. It assumes no home battery but allows home appliances to be driven by EV battery. Essentially we assume the EV battery takes the role of home battery. While a major difference lies in the fact that the EV battery is not available throughout the day and it can only be charged/used when the EV is parked at home, so it will miss the peak hours of solar generation and therefore has limited ability to store surplus solar power generations. The EV batter is more likely to serve only as a tool to bring cheaper

electricity to peak hours. The SP model can be modified based on the one for on-grid system.

EV battery flow (moved to the second stage as it now depends on scenario ω):

$$\rho \delta^t y_{G,E}^t - \sum_{j \in \{T,D\}} \delta^t (z_{E,j}^t)^\omega - U_E^t = (l_E^t)^\omega - (l_E^{t-1})^\omega, \forall t, \forall \omega$$

EV battery daily power cycle (moved to the second stage as it now depends on scenario ω):

$$(l_E^{24})^\omega = (l_E^0)^\omega, \forall \omega$$

Electricity demands:

$$\sum_{i \in \{S,G\}} (x_{i,j}^t)^\omega + \delta^t (z_{E,j}^t)^\omega = U_j^t, \forall j \in \{T,D\}, \forall t, \forall \omega,$$

4.6 Numerical Results

4.6.1 Experiment settings and system parameters

To draw valuable insights from solving the proposed SP model, we set the problem parameters to reflect practical situation according to reliable sources.

Electricity supply: $A_i^t, i \in \{S, G\}$ In practice, the electricity grid can supply as much energy as what is needed by a household, so this problem has no constraints on the amount of electricity from the grid (A_G^t). On the other hand, how much electricity we can generate from the PV system installed at home is under significant uncertainty, which has to be predicted to feed into the model.

The electricity output from solar panels depends on several factors such as size, ca-

capacity, location, and weather conditions. In the UK, most residential solar PV systems installed typically have a capacity between 1 kW and 4 kW by taking around 8 - 28 square metres (m²) of space. In this study, we will examine into three different peak PV generation capacities at 1 kW, 2 kW and 3 kW in clear sky conditions (sunny day) and study how electricity storage would act in these three settings. Re-scaled according to the chosen capacities, the spatial exponential smoothing model as presented in section 4.4 is used to generate 20 random scenarios of the 24-hour electricity generation, which are used as the $\{(A_S^t)^\omega, t = 1, \dots, 24, \omega = 1, \dots, 20\}$ values in the model.

Electricity demand: $U_j^t, j \in \{V, T, D\}$ The households consumption data (without EV), i.e., U_T^t and U_D^t , are extracted from the household electricity survey [52]. In this study we deploy the average household consumption data over all household types, so as to explore the optimal household electricity system settings under the typical UK consumption pattern. Devices (excluding EV) in households are categorised into two groups, i.e., major appliances (white goods) and small appliances & electronic devices (brown goods). Figure 4.5 summarises the hourly consumption level of these two categories (green and orange lines), together with other key information of the system like average daily PV generation and electricity price patterns. It can be seen that the peak demand occurs between 6-9 p.m., which is the typical cooking and family entertaining time after working hours. High peak-load price applies to this period's consumption, which forms the major part of high electricity bills.

For the EV consumption (U_V^t), we consider urban commuter driving profiles by assuming specifications as follows: (a) the household has an EV with 30 kWh battery, with a typical daily consumption of 10 kWh, (b) the EV is only charged at home and its available time ranges from 7 pm to 7 am the next morning. Let δ^t be binary indicators stating the available charging time at home for the EV which is set to one between 7 pm - 7 am.

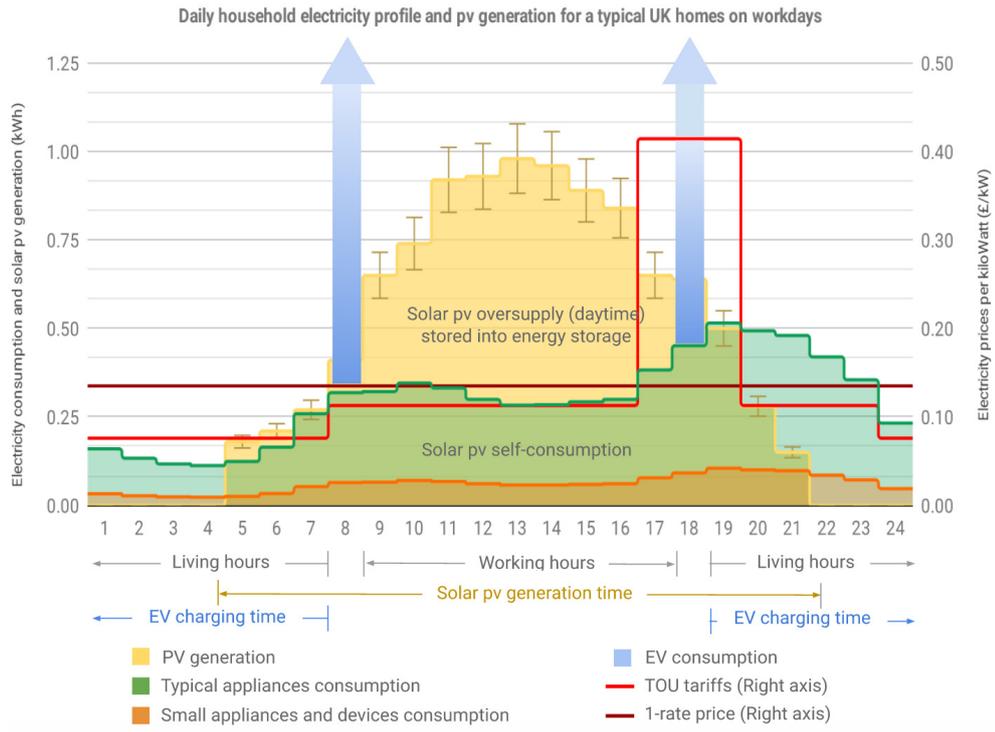


Figure 4.5: Average 24-hour UK household electricity profile and PV generation curves on summer weekdays

Electricity prices: $P_i^t, i \in \{S, G\}$ For electricity price we take a typical time-of-use (TOU) electricity tariff that is provided by the eastern region household electricity surveys [52]. A day is divided into four time intervals and each has its own electricity prices (P_G^t) as shown by the red line in Figure 4.5. On the other hand, we assume that the electricity generated by solar panels (P_S^t) is free of charge by ignoring the installation and maintenance costs. When customer upload energy back to the grid they receive credits for doing so and this is called Feed-in-tariff. In this study we test through different levels of Feed-in-tariff, γ , between 30 - 60% of the lowest price of the TOU tariffs.

Energy storage capacity: $C_k, k \in \{E, B\}$ Being consistent with industrial production, here we assume that a 30 kW (C_E) lithium-ion electric-vehicle battery is installed which provides up to 160 km range per charge. As for the home battery C_B , we are aiming to find the influence of it so we examine into the options ranging from 0-30

kWh. The battery round-trip efficiency is the fraction of energy put into the storage that can be retrieved afterwards. Here we set it to 80% according to [107].

4.6.2 Scenarios of solar PV generation

In this subsection we fit the forecasting model of Section 4.4 to a real dataset recorded at a solar farm in Cambridge during 1st - 30th June 2019. This dataset consists of 1,440 observations of half-hourly PV outputs with cloud cover conditions (five levels of cloudiness). The fitted model is then rescaled according to home PV capacity so as to generate simulation scenarios for our stochastic programming models.

In regression analysis (first stage of the proposed model), the aggregated-daily PV output as a dependent variable is modelled with weather conditions using enter (variable selection) method.

Table 4.3: Forecasting estimation

Stage	Parameter	Estimate			Remark (Cloud level)
		1 kW	2 kW	3 kW	
1st stage	β_0	0.805	1.610	2.414	Fine
	β_1	-0.025	-0.051	-0.076	Partly cloudy
	β_2	-0.041	-0.081	-0.122	Mostly cloudy
	β_3	-0.391	-0.782	-1.173	Cloudy
	β_4	-0.660	-1.321	-1.981	Showers
	σ_a^2	0.009	0.037	0.083	
2nd stage	σ^2	0.014	0.055	0.137	

The estimation results from the forecasting model as presented in Table 4.3 shown estimates which is used to simulate possible PV supplies. Therefore, the 1 kW PV generator for 24-hour ahead (on 1st July 2019, $j = 31$) PV outputs at time $t = 720$ when giving weather conditions and spatial information as follows: (a) a clear-sky day (fine), i.e. $\mathbf{x}_{\vec{s},31} = (0, 0, 0, 0)$, (b) the sunrise and sunset times on that day in Cambridge is 04.42 [sunrise time falling in 04.00-04.59 or the 5th hour of the day ($u_{\vec{s},31} = 24 * (31 - 1) + \lceil 4.42 \rceil = 725$), so the first hour for solar PV generation output of the day at time 725] and 21.24 ($v_{\vec{s},31} = 720 + \lceil 21.24 \rceil = 742$), respectively, i.e., a frequency of sine function on $j = 31$ equals $f_j = \frac{1}{v_{\vec{s},j} - u_{\vec{s},j} + 1} = \frac{1}{18}$, is written by

$$A_{S|\bar{s},\mathbf{x}_{\bar{s},31}}^t = \frac{\hat{P}_{\bar{s},31}}{2} \sin \left[\frac{1}{9} \pi (t - 724) - \frac{\pi}{2} \right] + \frac{\hat{P}_{\bar{s},31}}{2} + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2), t \in [725, 742],$$

$$= 0, \text{ elsewhere,}$$

$$P_{\bar{s},31} = \hat{\beta}_0 + \eta_j, \eta_j \sim N(0, \hat{\sigma}_a^2),$$

where $(\hat{\beta}_0, \hat{\sigma}_a^2, \hat{\sigma})$ for 1 kW to 3 kW peak PV system capacity are (0.805,0.009,0.014), (1.610,0.037,0.055) and (2.414,0.083,0.137), respectively.

To feed in the stochastic programming model we generate 20 random scenarios of hourly solar PV output for a specific weather condition, e.g., sunny, which form the inputs of the second-stage optimization model.

4.6.3 Results of SP models

To evaluate the energy efficiency of the proposed systems, we test them with the parameters listed above. Experiments are conducted using ILOG CPLEX V12.9.0 on an Intel Core i9-7940X 3.1GHz. Figure 4.6 shows a comparison of the energy bills under different system settings as presented in Section 4.5.2.

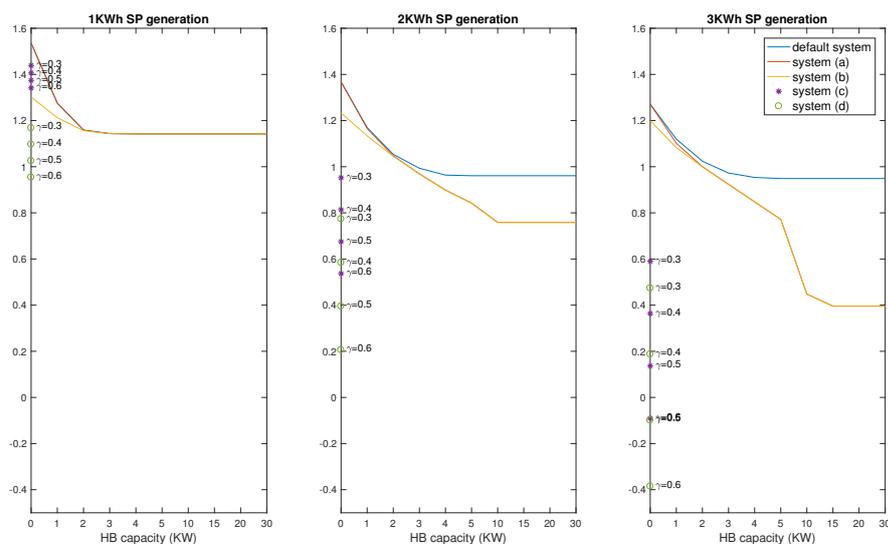


Figure 4.6: Daily household consumption costs by solar PV and battery capacities

Figure 4.6 shows the daily energy bills payable by the household under different system settings with different level of PV generation and Home Battery (HB) capacities. Note that some results are the same for different systems; to display them clearly some lines are thickened in Figure 4.6 and having another line drawn on top of it.

Not surprisingly, the daily usage cost decreases with increased size of solar PV capacity. The cost also decreases with an increased size of home battery, under the same solar PV generation. Upper limits on the home battery capacity can be identified for each system setting from the results. The larger the solar PV generation is, the higher HB should be installed to facilitate full usage of the power generated. According to Figure 4.6, suitable capacity of HB for 1 kWh, 2 kWh and 3 kWh PV under the default systems are, respectively, 3 kW, 4 kW and 5 kW. Higher home battery capacity won't decrease the electricity bill further. While the best capacity of HB under systems (a) and (b) are, 3 kW, 10 kW and 15 kW for 1 kWh, 2 kWh and 3 kWh PV generation.

System (b) (allowing EVB supply to home appliances) saves electricity bill by 5.6 – 15% when no HB is installed in the household, and by 3.3 – 5% when HB capacity is very low (HB=1). This is achieved by using the EVB to carry low-cost electricity from off-peak hours to peak hours. The saving is getting smaller as the solar PV generation is getting higher. If we take a closer look we can see that the electricity bills are roughly the same for system (b) when HB=0, regardless the size of the solar PV. This is because the EVB operates independently with the solar system if there's no home battery connecting them. On the other hand, the cost of the default system when HB=0 decreases with the increased generation capacity of solar PV, so the gap between system (b) and the default system becomes less significant.

System (a) (allowing transmission from HB to EVB) converge to default system when home battery capacity is low and converge to system (b) when home battery capacity is high. Indeed, when HB capacity is low, it carries less energy than the household demands so there won't be surplus to supply the EVB. When HB capacity is high, however, allowing EVB supply to home appliances won't add further value

to the system. In addition, savings achieved by system (a) become more significant when solar system gets larger; it saves 21% (58%) the daily electricity bill by allowing transmission from HB to EVB on top of the default system with 2 kWh (3 kWh) solar generation. Comparing the 2 kWh results with that of 3 kWh, it is not hard to see that both systems are restricted by the battery size until the HB capacity reaches 5 kW (they have same decreasing pattern until HB = 5 kW). Further savings can only be achieved with higher solar generation.

For system (c) and (d), their results are shown in the figures by the scatters aligned on the left of each graph, because they don't have home batteries (HB = 0). The scatters are labelled with the selling price rate considered, e.g., $\gamma = 0.3$ one can sell the surplus electricity at $0.3 * \min\{P_G^t, t = 1, \dots, h\}$ pounds per kW. When selling price is zero system (c) is identical to the default system (and system (a)) with HB = 0, while system (d) is identical to system (c) with HB = 0. We can see from the results that when selling is allowed, the daily bill decreases approximately linearly with the increase of the selling price. System (d) saves around 20% on system (c) by allowing EVB supply to home appliances when $\gamma = 0.3$, and the savings increase with the size of solar panels. When $\gamma = 0.3$, system (d) leads to similar daily bill to system (a) and (b) with the highest possible home battery installed, and System (d) is certainly better than system (a) and (b) when $\gamma = 0.4$ or higher.

Managerial insights:

- when HB=0 (is low), allowing EVB supply to home appliances saves 9 – 20% (5 – 8%) the daily bill. So assume EV is everywhere we can reduce the HB size and use EVB as alternative storage and supply.
- when home battery is large (over 10 kW), allowing transmission from HB to EVB saves 39 – 86% the daily bill with minimum 2 kWh solar generation.
- When solar generation is higher than 2 kWh, allowing HB supply to EVB (Sys-

tem (b)) become a good idea as this would increase the self-consumption of the electricity generated by the solar panel.

- Overall results: EV battery cannot fully replace HB because a) it cannot be charged by solar panel due to low voltage, b) it's not at home when solar generates electricity.

Chapter 5

Conclusions and Discussion

This study aims to: (a) define the solar-based time series used to explain the data that consists of two seasonal patterns with significant different lengths like the amount of hourly electricity generated from the PV system; (b) develop the forecasting model and method incorporating with spatial information, such as GPS coordinates, weather-related conditions (e.g. sunshine duration, cloud, etc.), to simulate a 24-hour (or many consecutive days) solar PV generation output in advance for balancing household electricity demand and supply; and (3) propose and investigate the different residential energy connections of electric vehicle batteries (EVB) to the residential solar photovoltaic (PV) system and to home appliances in order to determine the efficacy of deploying EVB in future households using a two-stage stochastic programming approach.

5.1 Conclusion

Weather-related information, such as sunshine duration and cloud conditions from weather forecasts, can reflect the daily peak of solar irradiation (the magnitude of the adjusted sine wave function) by the regression model before it is transformed into PV energy via the solar panel. The sinusoidal model with the magnitude obtained in the first stage of the spatial forecasting model, which is used to replace a traditional sea-

sonal component of Holt-Winters's exponential smoothing to predict the daily characteristics of the solar-based time series, plays an important role in extending the range of forecasts from one step up to a day ahead. While remaining components of the spatial forecasting model are applied to detect and adjust the short-term pattern for a step-by-step forecast in the second stage.

To estimate and run the forecasting model, the model needs a small sample size of the data. Relevant models, on the other hand, need at least two sidereal periods of the time series in order to train the model. The incorporation of a covariate into the regression model is another method that can be utilised to mitigate forecasting errors.

The optimal solutions for the default system and four different connections between batteries, devices, and selling electricity policy under parameter setting without considering the PV installation and maintenance costs of the PV system and home battery (HB) are totally explained in Chapter 4. In light of our findings, we would like to provide additional details regarding the potential for EVB to power home appliances. The average UK household consumes approximately 10 kWh of electricity per day¹, depending on its size, the appliances used, and the season. This means that a typical battery of an electric vehicle (EVB) with a capacity of 30 kWh, as configured in our systems, can store enough energy to meet the energy demands of a typical home for a couple of days. In other words, EVB supply to power home energy systems (HES) as set forth in systems (b) and (d) are effective solutions if the load profiles in households, which are reflected by power consumption behaviours, are not allowed to change.

5.2 Discussion

From the beginning of this study to its conclusion, it was challenging to propose the best model and method for a specific time series. In addition, we also design various residential energy systems for battery and device connections in an effort to discover

¹Celia Topping, "Average electricity usage in the UK: how many kWh does your home use?", 18 March 2021, URL: <https://www.ovoenergy.com/guides/energy-guides/how-much-electricity-does-a-home-use>

new findings in EVB-supportive energy management. We have used the spatial model to investigate the potentials of residential energy management based on the current parameter settings, which include things like the amount of energy that is lost during transmission and the average efficiency of solar panels, among other things. It's possible that some of these studies will be revised in the future to take into account developments in energy storage and photovoltaic solar technology.

5.2.1 Future Research

In the future, if the following problems in both the statistical and optimization sections have optimal solutions, the proposed methodology will yield superior

- In the event that the weather conditions have changed from a sunny day, which was used as the basis for the solar PV output simulation in Section 4.4, to a day that is either partly cloudy, mostly cloudy, cloudy, or rainy as shown an example of the next-day forecast in Figure 4.3, the results of the two-stage SP model will probably be affected. This is because the simulation was performed using a sunny day as the basis.
- To reduce the amount of energy drawn from the grid and to smooth the home energy system over consecutive days with varying weather conditions, it is necessary to investigate the correlation between energy storage levels at the beginning of each day for a more effective solution.
- Weather forecasting using digital image processing, such as a weather camera, can be applied to the proposed model for optimising the management of household energy consumption in real time.
- Estimating prediction intervals for the proposed forecasting by using a state-space model, which is a generalisation of the exponential smoothing model. For

estimating the PIs, the bootstrap method and the Monte Carlo Markov Chain (MCMC) methods can also be considered.

- In order to simulate the outputs from PV household systems for the purpose of long-term electricity planning, the weather forecasts for multiple days in a row that are provided by meteorological authorities can be applied to the model that was developed.
- In order to simulate the outputs from PV household systems for the purpose of long-term electricity planning, the weather forecasts for multiple days in a row that are provided by meteorological authorities can be applied to the model that was developed.
- There are reliable methods from existing pieces of literature for forecasting electricity demand that is better suited for specific types of buildings or households than survey (empirical) data.
- Under the influence of solar variation and climate change, the developed method is possibly applied to forecast solar-based time series related, such as a dam's evaporation rate, for sustainable water management in agriculture, etc.

5.2.2 Limitations of the Study

The findings of this study have to be seen in light of some limitations as following.

Lack of previous research studies: The currently available research on forecasting models was developed for a diverse range of data applications. It is not suitable for a particular time series, such as a high-frequency time series with double seasonality, a non-negative spatial-and-temporal time series, or another similar type of time series. Thus, there is a lack of prior research on the use of spatial information in statistical modelling for double seasonality data characterised by huge different seasonal lengths.

Lack of spatial information for the study: The proposed spatial forecasting model needs to know how the weather is in the area where the solar PV system is located. Because of this, information about a particular area may not always be available that is why only two data set are analytical study in Section 3.4.

Appendix A

Sunrise and sunset calculation

According to the seasonal equation as 3.2, this section will demonstrate how the sunrise and sunset time calculate by spatial information. Considering the Earth's axis and a rotation of the Earth that the Earth spins around its axis and rotates around the sun at the same time, that is why we have different daytime, nighttime, and season depend on geography. Now, we are interested in what time the sunrise and sunset are each position and how to determine a new index of the time series for the proposed model. The algorithm to calculate the sunrise and sunset time following [108] published by Nautical Almanac Office consists of the following steps:

Inputs:

day, month, year:	date of sunrise/sunset
latitude, longitude:	location for sunrise/sunset
zenith ¹	Sun's zenith for sunrise/sunset
	official = 90° 50'
	civil = 96°
	nautical = 102°
	astronomical = 108°

Notice:

1. Longitude is positive for East and negative for West.
2. The trigonometry functions in the algorithm calculated in degree.

Calculation:

¹the point on the celestial sphere directly above an observer on the Earth.

1. The day of the year, N .

$$N = \left\lfloor \frac{275 * month}{9} \right\rfloor - \left\lfloor \frac{month + 9}{12} \right\rfloor * \left\lfloor 1 + \frac{year - 4 * \left\lfloor \frac{year}{4} \right\rfloor + 2}{3} \right\rfloor + day - 30$$

2. Convert the longitude to hour value, $lngHour$, and calculate an approximate time, t .

$$lngHour = \frac{longitude}{15}$$

$$t = \begin{cases} N + \left(\frac{6 - lngHour}{24} \right), & \text{rising time} \\ N + \left(\frac{18 - lngHour}{24} \right), & \text{setting time} \end{cases}$$

3. The Sun's mean anomaly, M .

$$M = (0.9856 * t) - 3.289$$

4. The Sun's true longitude, L .

$$L = M + (1.916 * \sin(M)) + (0.020 * \sin(2 * M)) + 282.634$$

Notice: L potentially needs to be adjusted into the range $[0,360)$ by adding/subtracting 360.

5. The Sun's right ascension, RA .

$$RA = \frac{1}{15} \left[atan(0.91764 * \tan(L)) + 90 * \left\{ \left\lfloor \frac{L}{90} \right\rfloor - \left\lfloor \frac{atan(0.91764 * \tan(L))}{90} \right\rfloor \right\} \right]$$

6. The Sun's declination, $sinDec$.

$$sinDec = 0.39782 * \sin(L)$$

$$cosDec = \cos(asin(sinDec))$$

7. The Sun's local hour angle, H .

$$\cos H = \frac{\cos(\text{zenith}) - \sin Dec * \sin(\text{latitude})}{\cos Dec * \cos(\text{latitude})}$$

$$H = \begin{cases} \frac{1}{15} * [360 - \text{acos}(\cos H)], & \text{rising time} \\ \frac{1}{15} * [\text{acos}(\cos H)], & \text{setting time} \end{cases}$$

Notice: if $\cos H > 1$ or $\cos H < -1$, the sun never rises and sets on this location (on the specified date).

8. Local mean time of rising/setting

$$T = H + RA - (0.06571 * t) - 6.622$$

9. Adjust back to UTC, UT .

$$UT = T - \text{lngHour}$$

Note: UT potentially needs to be adjusted into the range [0,24) by adding/subtracting 24.

10. Convert UT value to local time zone of latitude/longitude, $localT$.

$$localT = UT + \text{localOffset}$$

Sunrise and sunset time index:

Sunrise and sunset time index of time series at time t over the position $\vec{s} = (x_1, x_2)$ are

respectively as follows:

$$u_{\vec{s},t} = m * (j - 1) + [2 * f(x_1, x_2, zenith, date, sun.mode = 1, UTC)]$$

$$v_{\vec{s},t} = m * (j - 1) + [2 * f(x_1, x_2, zenith, date, sun.mode = 2, UTC)]$$

where $j, j = \lceil t/m \rceil$ denotes a seasonal period.

To start with defining a time series in time order and inputs spatial data such as latitude and longitude, (x_1, x_2) , of a data collecting location, sun's zenith angle (approximately 90°), time zone (including daylight saving time in position), and subsequently calculate the above sunrise-sunset seasonal time index as the following formulas. For example, calculating sunrise and sunset time on 22nd September 2018 in Colchester, (Coordinate: $51^\circ 53' 35'' N, 0^\circ 54' 16'' E$). The results from above formulas for sunrise is 6.6812502 (6:40 am), and sunset time is 18.9322 (6:55 pm). Therefore the seasonal time index for sunrise and sunset assuming the given day is a first day of a half-hourly time series ($m = 48$) are 14 and 38, respectively.

Bibliography

- [1] Maimouna Diagne, Mathieu David, Philippe Lauret, John Boland, and Nicolas Schmutz. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renewable and Sustainable Energy Reviews*, 27:65–76, 2013.
- [2] Ivan Svetunkov. Forecasting and analytics with adam. OpenForecast, 2022. (version: [current date]).
- [3] Steven Firth Tom Kane Daniel Godoy-Shimizu Peter Pope Jason Palmer, Nicola Terry. Energy use at home: models, labels and unusual appliances. *Further Analysis of the Household Electricity Survey, Reference 475/09/2012*, February 2014.
- [4] Ralph Snyder, Anne Koehler, Rob Hyndman, and Keith Ord. Exponential smoothing for inventory control: Means and variances of lead-time demand. 03 2002.
- [5] RENA. *Renewable capacity statistics 2020 International Renewable Energy Agency (IRENA)*. Abu Dhabi, 2020.
- [6] Phebe Asantewaa Owusu and Samuel Asumadu-Sarkodie. A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Engineering*, 3(1):1167990, 2016.
- [7] REN21. Renewables 2019 global status report. Technical report, 2019.

- [8] Department for Business, Energy & Industrial Strategy. Uk energy statistics, q1 2019, June 2019.
- [9] Leidy Contreras, Santiago Lain, Mohamad Issa, and Adrian Ilinca. *4 - Renewable energy systems*, pages 103–177. 12 2020.
- [10] Mahdi Sharifzadeh, Helena Lubiano-Walochik, and Nilay Shah. Integrated renewable electricity generation considering uncertainties: The uk roadmap to 50energies. *Renewable and Sustainable Energy Reviews*, 72:385 – 398, 2017.
- [11] Department for Business, Energy & Industrial Strategy. Uk renewable electricity capacity and generation, December 2020.
- [12] Liam Stoker. Iea: Solar the 'new king' of power, will break records for decades to come, Oct 2020.
- [13] Sophie Pelland, Jan Remund, Jan Kleissl, Takashi Oozeki, and Karel De Brabandere. *Photovoltaic and Solar Forecasting: State of the Art*. 10 2013.
- [14] Bella Espinar, José-Luis Aznarte, Robin Girard, Alfred Moussa, and George Kariniotakis. Photovoltaic forecasting: A state of the art. *5th European PV-Hybrid and Mini-Grid Conference-Tarragona, Spain*, 04 2010.
- [15] NRG Labs. Pv forecast: Solar power generation forecasts. 2022.[Mobile app].[Accessed 15 March 2022].
- [16] Ahmed Zayed AL Shaqsi, Kamaruzzaman Sopian, and Amer Al-Hinai. Review of energy storage services, applications, limitations, and benefits. *Energy Reports*, 6:288–306, 2020. SI:Energy Storage - driving towards a clean energy future.
- [17] Eric Confais and Ward van den Berg. Business models in energy storage. *Roland Berger*, May, 2017.

- [18] André Berger and Qiuzhen Yin. Astronomical theory and orbital forcing. *The SAGE Handbook of Environmental Change: Volume 1*, 1:405–425, 01 2012.
- [19] Robin John Hyndman and George Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, Australia, 2nd edition, 2018.
- [20] Saul I. Gass and Carl M. Harris. *Encyclopedia of operations research and management science*. Springer, 1996.
- [21] Robert G. Brown. *Exponential Smoothing for Predicting Demand*. Cambridge, Massachusetts: Arthur D. Little Inc., 1956.
- [22] Charles C. Holt. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1):5 – 10, 2004.
- [23] Peter R Winters. Forecasting sales by exponentially weighted moving averages. *Management science*, 6(3):324–342, 1960.
- [24] C. Carl Pegels. Exponential forecasting: Some new variations. *Management Science*, 15(5):311–315, 1969.
- [25] Everette S. Gardner Jr. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1):1–28, 1985.
- [26] Chris Chatfield and Mohammed Yar. Prediction intervals for multiplicative holt-winters. *International Journal of Forecasting*, 7(1):31–37, 1991.
- [27] Keith Ord, Ralph Snyder, and Anne Koehler. Forecasting models and prediction intervals for the multiplicative holt-winters method. *International Journal of Forecasting*, 17:269–286, 02 2001.
- [28] Rob J Hyndman, Anne B Koehler, Ralph D Snyder, and Simone Grose. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of forecasting*, 18(3):439–454, 2002.

- [29] James W Taylor. Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society*, 54(8):799–805, 2003.
- [30] James W. Taylor. Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19(4):715 – 725, 2003.
- [31] Everette S. Gardner. Exponential smoothing: The state of the art—part ii. *International Journal of Forecasting*, 22(4):637 – 666, 2006.
- [32] Phillip G. Gould, Anne B. Koehler, J. Keith Ord, Ralph D. Snyder, Rob J. Hyndman, and Farshid Vahid-Araghi. Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research*, 191(1):207–222, 2008.
- [33] James Taylor. An evaluation of methods for very short-term load forecasting using minute-by-minute british data. *International Journal of Forecasting*, 24:645–658, 10 2008.
- [34] James W. Taylor. Triple seasonal methods for short-term electricity demand forecasting. *European Journal of Operational Research*, 204(1):139 – 152, 2010.
- [35] Alysha M De Livera, Rob J Hyndman, and Ralph D Snyder. Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496):1513–1527, 2011.
- [36] James W. Taylor and Ralph D. Snyder. Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing. *Omega*, 40(6):748–757, 2012. Special Issue on Forecasting in Management Science.
- [37] Rob J. Hyndman. TbatS with regressors, Oct 2014.

- [38] James W Taylor. Exponential smoothing for time series with multiple seasonal cycles. *The 27th International Symposium on Forecasting*, pages 1–15, June 2007.
- [39] Rob J. Hyndman and Yeasmin Khandakar. Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, 27(3):1–22, 2008.
- [40] Ji-Long Chen, Guo-Sheng Li, and Sheng-Jun Wu. Assessing the potential of support vector machine for estimating daily solar radiation using sunshine duration. *Energy Conversion and Management*, 75:311 – 318, 2013.
- [41] M. Rivington, G. Bellocchi, K.B. Matthews, and K. Buchan. Evaluation of three model estimations of solar radiation at 24 uk stations. *Agricultural and Forest Meteorology*, 132(3):228 – 243, 2005.
- [42] D.G. Miller, M. Rivington, K.B. Matthews, K. Buchan, and G. Bellocchi. Testing the spatial applicability of the johnson - woodward method for estimating solar radiation from sunshine duration data. *Agricultural and Forest Meteorology*, 148(3):466 – 480, 2008.
- [43] A. A. Teyabeen and A. E. Jwaid. Sunshine duration-based models for predicting global solar radiation. In *2017 UKSim-AMSS 19th International Conference on Computer Modelling Simulation (UKSim)*, pages 168–172, April 2017.
- [44] Harry Suehrcke, Ross S. Bowden, and K.G.T. Hollands. Relationship between sunshine duration and solar radiation. *Solar Energy*, 92:160 – 171, 2013.
- [45] Chi Wai Chow, Bryan Urquhart, Matthew Lave, Anthony Dominguez, Jan Kleissl, Janet Shields, and Byron Washom. Intra-hour forecasting with a total sky imager at the uc san diego solar energy testbed. *Solar Energy*, 85(11):2881 – 2893, 2011.

- [46] Dorota Matuszko. Influence of the extent and genera of cloud cover on solar radiation intensity. *International Journal of Climatology*, 32(15):2403–2414, 2012.
- [47] Md. Nazmul Islam Sarkar. Estimation of solar radiation from cloud cover data of bangladesh. *Renewables: Wind, Water, and Solar*, 3, 03 2016.
- [48] S. Rangarajan, M.S. Swaminathan, and A. Mani. Computation of solar radiation from observations of cloud cover. *Solar Energy*, 32(4):553–556, 1984.
- [49] Gordon Reikard. Predicting solar radiation at high resolutions: A comparison of time series forecasts. *Solar Energy*, 83(3):342–349, 2009.
- [50] Jacopo Torriti. A review of time use models of residential electricity demand. *Renewable and Sustainable Energy Reviews*, 37:265 – 272, 2014.
- [51] Department of Energy & Climate Change. Household electricity survey, Jan 2014.
- [52] Department for Communities and Local Government. *English Housing Survey HOUSEHOLDS 2010-11: Annual report on England’s households, 2010–11*. JULY 2012.
- [53] Florian Barbieri, Sumedha Rajakaruna, and Arindam Ghosh. Very short-term photovoltaic power forecasting with cloud modeling: A review. *Renewable and Sustainable Energy Reviews*, 75:242–263, 2017.
- [54] R. Ahmed, V. Sreeram, Y. Mishra, and M.D. Arif. A review and evaluation of the state-of-the-art in pv solar power forecasting: Techniques and optimization. *Renewable and Sustainable Energy Reviews*, 124:109792, 2020.
- [55] Zibo Dong, Dazhi Yang, Thomas Reindl, and Wilfred M. Walsh. Short-term solar irradiance forecasting using exponential smoothing state space model. *Energy*, 55:1104–1113, 2013.

- [56] Harendra Kumar Yadav, Yash Pal, and M.M. Tripathi. Photovoltaic power forecasting methods in smart power grid. In *2015 Annual IEEE India Conference (INDICON)*, pages 1–6, 2015.
- [57] James W Taylor, Lilian M De Menezes, and Patrick E McSharry. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting*, 22(1):1–16, 2006.
- [58] Moonyoung Baek. *A study on forecasting high -frequency time series with multiple seasonal patterns*. PhD thesis, 2008. Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2019-10-17.
- [59] Dougal Burnett, Edward Barbour, and Gareth P. Harrison. The uk solar energy resource and the impact of climate change. *Renewable Energy*, 71:333 – 343, 2014.
- [60] J.A. Sabbagh, A.A.M. Sayigh, and E.M.A. El-Salam. Estimation of the total solar radiation from meteorological data. *Sol. Energy; (United States)*.
- [61] Utpal Kumar Das, Kok Soon Tey, Mehdi Seyedmahmoudian, Saad Mekhilef, Moh Yamani Idna Idris, Willem Van Deventer, Bend Horan, and Alex Stojcevski. Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*, 81:912 – 928, 2018.
- [62] Claus Fröhlich and Judith Lean. The sun’s total irradiance: Cycles, trends and related climate change uncertainties since 1976. *Geophysical Research Letters*, 25(23):4377–4380, 1998.
- [63] Richard H. Byrd, Peihuang Lu, Jorge Nocedal, and Ciyu Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal of Scientific Computing*, 16:1190–1208, 9 1995.

- [64] J. N. Black, C. W. Bonython, and J. A. Prescott. Solar radiation and the duration of sunshine. *Quarterly Journal of the Royal Meteorological Society*, 80(344):231–235, 1954.
- [65] David B. Ampratwum and Atsu S.S. Dorvlo. Estimation of solar radiation from the number of sunshine hours. *Applied Energy*, 63(3):161 – 167, 1999.
- [66] A. Q. Jakhrani, A. K. Othman, A. R. H. Rigit, and S. R. Samo. A simple method for the estimation of global solar radiation from sunshine hours and other meteorological parameters. In *2010 IEEE International Conference on Sustainable Energy Technologies (ICSET)*, pages 1–6, Dec 2010.
- [67] M. Suri, T. Cebecauer, T. Skoczek, R. Marais, C. Mushwana, and J. Reinecke. Cloud cover impact on photovoltaic power production in south africa. *Proceedings: 2nd South African Solar Energy Conference.*, 2014.
- [68] Chrobak, Pavel, Skovajsa, Jan, and Zalesak, Martin. Effect of cloudiness on the production of electricity by photovoltaic panels. *MATEC Web Conf.*, 76:02010, 2016.
- [69] A. Woyte, V. Van Thong, R. Belmans, and J. Nijs. Voltage fluctuations on distribution level introduced by photovoltaic systems. *IEEE Transactions on Energy Conversion*, 21(1):202–209, March 2006.
- [70] Cédric Bertrand, Colienne Demain, and Michel Journée. Estimating daily sunshine duration over belgium by combination of station and satellite data. *Remote Sensing Letters*, 4(8):735–744, 2013.
- [71] Steffen Kothe, Uwe Pfeifroth, Roswitha Cremer, Jörg Trentmann, and Rainer Hollmann. A satellite-based sunshine duration climate data record for europe and africa. *Remote Sensing*, 9(5), 2017.
- [72] Sylvain Quoilin, Konstantinos Kavvadias, Arnaud Mercier, Irene Pappone, and Andreas Zucker. Quantifying self-consumption linked to solar home battery

- systems: Statistical analysis and economic assessment. *Applied Energy*, 182:58 – 67, 2016.
- [73] G.B.M.A. Litjens, B.B. Kausika, E. Worrell, and W.G.J.H.M. van Sark. A spatio-temporal city-scale assessment of residential photovoltaic power integration scenarios. *Solar Energy*, 174:1185 – 1197, 2018.
- [74] Rasmus Luthander, Joakim Widén, Daniel Nilsson, and Jenny Palm. Photovoltaic self-consumption in buildings: A review. *Applied Energy*, 142:80 – 94, 2015.
- [75] Wessam El-Baz, Peter Tzscheutschler, and Ulrich Wagner. Day-ahead probabilistic pv generation forecast for buildings energy management systems. *Solar Energy*, 171:478 – 490, 2018.
- [76] Leidy Tatiana Contreras Montoya, Santiago Lain, Mohamad Issa, and Adrian Il-inca. 4 - renewable energy systems. In Ersan Kabalci, editor, *Hybrid Renewable Energy Systems and Microgrids*, pages 103 – 177. Academic Press, 2021.
- [77] Department for Business, Energy & Industrial Strategy. *UK Renewable electricity capacity and generation, April to June 2020*. December 2020.
- [78] Liam Stoker. Iea: Solar the ‘new king’ of power, will break records for decades to come. <https://www.pv-tech.org>, 2020. Accessed: 2020-12-12.
- [79] Sylvain Quoilin, Konstantinos Kavvadias, Arnaud Mercier, Irene Pappone, and Andreas Zucker. Quantifying self-consumption linked to solar home battery systems: Statistical analysis and economic assessment. *Applied Energy*, 182:58 – 67, 2016.
- [80] G.B.M.A. Litjens, B.B. Kausika, E. Worrell, and W.G.J.H.M. van Sark. A spatio-temporal city-scale assessment of residential photovoltaic power integration scenarios. *Solar Energy*, 174:1185 – 1197, 2018.

- [81] Rasmus Luthander, Joakim Widén, Daniel Nilsson, and Jenny Palm. Photovoltaic self-consumption in buildings: A review. *Applied Energy*, 142:80 – 94, 2015.
- [82] Solène Goy and Ana Sancho-Tomás. 4 - load management in buildings. In Ursula Eicker, editor, *Urban Energy Systems for Low-Carbon Cities*, pages 137 – 179. Academic Press, 2019.
- [83] Michael Schmela, Thomas Döring, Andrés Pinto-Bello Gómez, and Alexandre Roesch. 3 - solar power in europe: Status and outlook. In Trevor M. Letcher and Vasilis M. Fthenakis, editors, *A Comprehensive Guide to Solar Energy Systems*, pages 37 – 52. Academic Press, 2018.
- [84] Wessam El-Baz, Peter Tzscheutschler, and Ulrich Wagner. Day-ahead probabilistic pv generation forecast for buildings energy management systems. *Solar Energy*, 171:478 – 490, 2018.
- [85] Xian Zhang, Ke Wang, Yu Hao, Jing-Li Fan, and Yi-Ming Wei. The impact of government policy on preference for nevs: The evidence from china. *Energy Policy*, 61:382 – 393, 2013.
- [86] Automotive Dialogue(AD). The impact of government policy on promoting new energy vehicles (nevs) - the evidence in apec economies. report APEC 217-CT-01.3, Beijing Capital Energy Technology Co. Ltd, March 2017.
- [87] IEA. Global ev outlook 2020. Technical report, IEA, Paris, 2020. <https://www.iea.org/reports/global-ev-outlook-2020>.
- [88] Outlander phev debuts in thailand to offer a new environmentally-friendly option. <https://www.mitsubishi-motors.com/en/newsrelease/2020/detail1303>, 2020. Accessed: 2020-12-01.

- [89] Tim Mullaney. Tesla and the science behind the next-generation, lower-cost, ‘million-mile’ electric-car battery. <https://www.cnbc.com/2020/06/30/>, 2020. Accessed: 2020-12-01.
- [90] Bandana Mahapatra and Anand Nayyar. Home energy management system (hems): concept, architecture, infrastructure, challenges and energy management schemes. *Energy Systems*, 11 2019.
- [91] Xiaohua Wu, Xiaosong Hu, Scott Moura, Xiaofeng Yin, and Volker Pickert. Stochastic control of smart home energy management with plug-in electric vehicle battery energy storage and photovoltaic array. *Journal of Power Sources*, 333:203 – 212, 2016.
- [92] Xiaohua Wu, Xiaosong Hu, Yanqiong Teng, Shide Qian, and Rui Cheng. Optimal integration of a hybrid solar-battery power source into smart home nanogrid with plug-in electric vehicle. *Journal of Power Sources*, 363:277 – 283, 2017.
- [93] X. Hou, J. Wang, P. Wang, and T. Wang. Smart home energy management optimization method considering ess and pev. In *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pages 101–106, 2019.
- [94] Anton J. Kleywegt and Alexander Shapiro. *Stochastic Optimization*, chapter 102, pages 2625–2649. John Wiley & Sons, Ltd, 2001.
- [95] T. Chitsuphaphan, X. Yang, and H. Dai. Stochastic programming for residential energy management with electric vehicle under photovoltaic power generation uncertainty. In *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, pages 1–6, 2020.
- [96] Wouter L. Schram, Ioannis Lampropoulos, and Wilfried G.J.H.M. van Sark. Photovoltaic systems coupled with batteries that are optimally sized for household self-consumption: Assessment of peak shaving potential. *Applied Energy*, 223:69 – 81, 2018.

- [97] Marc Beaudin and Hamidreza Zareipour. Home energy management systems: A review of modelling and complexity. *Renewable and Sustainable Energy Reviews*, 45:318 – 335, 2015.
- [98] John R Birge. *Introduction to stochastic programming*. Springer New York, New York, NY, c2011.
- [99] Sobrina Sobri, Sam Koohi-Kamali, and Nasrudin Abd Rahim. Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management*, 156:459–497, 2018.
- [100] E. G. Kardakos, M. C. Alexiadis, S. I. Vagropoulos, C. K. Simoglou, P. N. Biskas, and A. G. Bakirtzis. Application of time series and artificial neural network models in short-term forecasting of pv power generation. In *2013 48th International Universities' Power Engineering Conference (UPEC)*, pages 1–6, 2013.
- [101] M. G. De Giorgi, P. M. Congedo, and M. Malvoni. Photovoltaic power forecasting using statistical methods: impact of weather data. *IET Science, Measurement Technology*, 8(3):90–97, 2014.
- [102] C. Yang, A. A. Thatte, and L. Xie. Multitime-scale data-driven spatio-temporal forecast of photovoltaic generation. *IEEE Transactions on Sustainable Energy*, 6(1):104–112, 2015.
- [103] M. Bouzerdoum, A. Mellit, and A. Massi Pavan. A hybrid model (sarima–svm) for short-term power forecasting of a small-scale grid-connected photovoltaic plant. *Solar Energy*, 98:226 – 235, 2013.
- [104] Hasimah Abdul Rahman Yuan-Kang Wu, Chao-Rong Chen. A novel hybrid model for short-term forecasting in pv power generation. *International Journal of Photoenergy*, 2014(9), 2014.

- [105] A. Hammer, D. Heinemann, E. Lorenz, and B. Lücke. Short-term forecasting of solar radiation: a statistical approach using satellite data. *Solar Energy*, 67(1):139 – 150, 1999.
- [106] S.X. Chen, H.B. Gooi, and M.Q. Wang. Solar radiation forecast based on fuzzy logic and neural networks. *Renewable Energy*, 60:195 – 201, 2013.
- [107] U.S. Energy Information Administration. Utility-scale batteries and pumped storage return about 80% of the electricity they store. FEBRUARY 12, 2021.
- [108] *Almanac for Computers*. U.S. Government Printing Office, 1990.