IEC TR 63043

Edition 1.0 2020-11

TECHNICAL REPORT

colour inside

Renewable energy power forecasting technology

IEC TR 63043:2020-11(en)

IEC Central Office
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
info@iec.ch
www.iec.ch

About the IEC
The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

About IEC publications
The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigendum or an amendment might have been published.

IEC publications search - webstore.iec.ch/advsearchform
The advanced search enables to find IEC publications by a variety of criteria (reference number, text, technical committee,…). It also gives information on projects, replaced and withdrawn publications.

IEC Just Published - webstore.iec.ch/justpublished
Stay up to date on all new IEC publications. Just Published details all new publications released. Available online and once a month by email.

IEC Customer Service Centre - webstore.iec.ch/csc
If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: sales@iec.ch.

Electropedia - www.electropedia.org
The world's leading online dictionary on electrotechnology, containing more than 22 000 terminological entries in English and French, with equivalent terms in 16 additional languages. Also known as the International Electrotechnical Vocabulary (IEV) online.

IEC Glossary - std.iec.ch/glossary
67 000 electrotechnical terminology entries in English and French extracted from the Terms and Definitions clause of IEC publications issued since 2002. Some entries have been collected from earlier publications of IEC TC 37, 77, 86 and CISPR.

IEC TR 63043

Edition 1.0 2020-11

TECHNICAL REPORT

colour inside

Renewable energy power forecasting technology

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

ICS 29.020 ISBN 978-2-8322-9079-8

CONTENTS

INTERNATIONAL ELECTROTECHNICAL COMMISSION

RENEWABLE ENERGY POWER FORECASTING TECHNOLOGY

FOREWORD

1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
6) All users should ensure that they have the latest edition of this publication.
7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
9) Attention is drawn to the possibility that some of the elements of this IEC Publication may be the subject of patent rights. IEC shall not be held responsible for identifying any or all such patent rights.

IEC 63043 has been prepared by subcommittee 8A: Grid Integration of Renewable Energy Generation, of IEC technical committee 8: Systems aspects of electrical energy supply. It is a Technical Report.

The text of this Technical Report is based on the following documents:

Draft	Report on voting
8A/71/DTR	8A/73/RVDTR

Full information on the voting for its approval can be found in the report on voting indicated in the above table.

The language used for the development of this Technical Report is English.

This document was drafted in accordance with ISO/IEC Directives, Part 2, and developed in accordance with ISO/IEC Directives, Part 1 and ISO/IEC Directives, IEC Supplement, available at www.iec.ch/members_experts/refdocs. The main document types developed by IEC are described in greater detail at www.iec.ch/standardsdev/publications.

The committee has decided that the contents of this document will remain unchanged until the stability date indicated on the IEC website under "http://webstore.iec.ch" in the data related to the specific document. At this date, the document will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

IMPORTANT – The 'colour inside' logo on the cover page of this publication indicates that it contains colours which are considered to be useful for the correct understanding of its contents. Users should therefore print this document using a colour printer.

INTRODUCTION

The purpose of this IEC Technical Report (TR) is to describe common practices and the state of the art for renewable energy power forecasting, which includes general data requirements, methods for renewable energy power forecasting and forecast error evaluation.

Various stakeholders, including transmission system operators, transmission system owners, utilities, renewable energy generation plant developers, academic units, research institutions, certifying bodies and standardization groups, require a common understanding of renewable energy power forecasting methods, data and evaluation techniques so they can incorporate them in their operations.

Renewable energy power forecasting finds a broad application in many areas of electrical engineering related to design, analysis, market trading, and optimisation of the power system. Among others, forecasting could be as an input to the operation and management of the renewable energy generation plants and can improve the economic efficiency and reliability of the power system.

Renewable energy power forecasting is increasingly important in multi-stakeholder systems where renewable plant manufacturers, renewable energy generation plant developers and operators, as well as the power system operators, need to have a common understanding about the capabilities and methods associated with renewable energy power forecasting.

RENEWABLE ENERGY POWER FORECASTING TECHNOLOGY

1 Scope

This Technical Report, which is informative in its nature, describes common practices and state of the art for renewable energy power forecasting technology, including general data demands, renewable energy power forecasting methods and forecasting error evaluation. For the purposes of this document, renewable energy refers to variable renewable energy, which mainly comprises wind power and photovoltaic (PV) power – these are the focus of the document. Other variable renewable energies, like concentrating solar power, wave power and tidal power, etc., are not presented in this document, since their capacity is small, while hydro power forecasting is a significantly different field, and so not covered here.

The objects of renewable energy power forecasting can be wind turbines, or a wind farm, or a region with lots of wind farms (respectively PV systems, PV power stations and regions with high PV penetration). This document focuses on providing technical guidance concerning forecasting technologies of multiple spatial and temporal scales, probabilistic forecasting, and ramp event forecasting for wind power and PV power.

This document outlines the basic aspects of renewable energy power forecasting technology. This is the first IEC document related to renewable energy power forecasting. The contents of this document will find an application in the following potential areas:

- support the development and future research for renewable energy power forecasting technology, by showing current state of the art;
- evaluation of the forecasting performance during the design and operation of renewable energy power forecasting system;
- provide information for benchmarking renewable forecasting technologies, including methods used, data required and evaluation techniques.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC 61400-12-2, *Wind turbines – Part 12-2: Power performance of electricity-producing wind turbines based on nacelle anemometry*

3 Terms, definitions and abbreviated terms

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- IEC Electropedia: available at http://www.electropedia.org/
- ISO Online browsing platform: available at http://www.iso.org/obp

3.1 Terms and definitions

3.1.1
accumulation method
wind power forecasting method for wind farm clusters that directly accumulates the forecasting results of each wind farm output

3.1.2
combination approaches, pl.
approaches usually used to describe the forecasting models which combine physical approaches with statistical approaches

3.1.3
data assimilation
assimilation combining recent observation data with the background (prior) forecast in real time, to adjust forecast model trajectories and update the background field

Note 1 to entry: In numerical weather prediction (NWP), this helps the weather forecast become more accurate.

3.1.4
day ahead power forecasting
timescale used in forecasting, which is from the current time to the next 24 hours to 72 hours

3.1.5
days to week ahead power forecasting
timescale used in forecasting, which is from the day ahead to the week ahead

3.1.6
deterministic forecasting
kind of forecasting method that can output deterministic results

3.1.7
deterministic forecasting evaluation
evaluation of the specific forecasting value at a certain moment in the future

3.1.8
distributed PV power forecasting
a method to forecast the total power output of distributed PVs in a certain region

3.1.9
ensemble forecasting
NWP method which produces a set of forecasts instead of making a single forecast

Note 1 to entry: This set of forecasts aims to give an indication of the range of possible future states of the atmosphere.

3.1.10
event forecasting evaluation
evaluation of the forecasting results of specific events

Note 1 to entry: For example wind or solar ramp events.

3.1.11
horizon
length of the forecast look-ahead time

3.1.12
minutes and hours ahead power forecasting
type of forecasting which is in minute scale (up to 15 min) or ultra-short-term (15 min to 6 h or 8 h)

3.1.13
nonparametric modelling
distribution-free without any assumptions of the distribution type

Note 1 to entry: It directly calculates the quantile or distribution function of the unknown random variable by means of data analysis methods.

3.1.14
numerical weather prediction
NWP
method to predict weather which numerically solves the basic equations of atmospheric motion based on the most recent observations that best represent the current atmospheric conditions

3.1.15
parameterization schemes
methods to capture the quantitative physical characteristics of radiative, convective and diffusive processes in the atmosphere and at the interface between the atmosphere and the surface.

Note 1 to entry: These processes are often determined by relatively small spatial scales, and are used in NWP models.

3.1.16
parametric modelling
model to use a predetermined distribution type describing the probability density function (PDF) of the unknown random variable

3.1.17
persistence forecasting
a method to use the measured power value at the current moment as the forecasted power at the future time

3.1.18
physical approaches
mathematical and physical models which are used to describe the physical factors

3.1.19
power forecasting of renewable plant clusters
the forecasting of the overall output of wind or solar PV clusters

3.1.20
probabilistic forecasting
a kind of forecasting methods that focuses on the uncertainty of power output

Note 1 to entry: Including wind power probabilistic forecasting and PV power probabilistic forecasting. The forecasting results could be the PDF, cumulative distribution function (CDF) of the random variable of power or the prediction intervals at certain probability levels.

3.1.21
probabilistic forecasting evaluation
evaluation of the forecasting results of the uncertainty of power output

3.1.22
ramp events, pl.
significant changes of power output in a short period

Note 1 to entry: These may refer specifically to those events not caused by the expected change due to the expected change in output of solar PV. Such events are prone to cause frequency fluctuation and power quality deterioration, potentially impacting the reliable operation of the power grid.

3.1.23
ramp magnitude
variation of power output in the observation period

3.1.24
ramp rate
variation rate of wind power in the observation period

3.1.25
resolution
spatial or temporal scales at which forecasts are made, measured in kilometers (spatial) or minutes/hours

3.1.26
statistical approach
mathematical model which is used to describe the relationship between historical NWP data, weather data and historical power output of a wind farm or a PV power station

3.1.27
statistical upscaling method
establishment of an upscaling model with part of a set of points to estimate the total

Note 1 to entry: It is the total regional power output estimated from a subset of wind farms or PV stations.

3.1.28
stochastic process for forecast development
stochastic process/model to incorporate random variation

Note 1 to entry: Usually based on fluctuations observed in historical data for a selected period using standard time-series techniques. The use of solely historical data is the main difference compared to probabilistic processes/models, which apply variations generated from some type of perturbation of the prediction process.

3.1.29
wake effect
phenomenon of wind speed decreasing after wind turbines extract power from the wind

3.2 Abbreviated terms

AE	analog ensemble
ANN	artificial neural network
AR	autoregressive
ARMA	autoregressive moving average
BPNN	back propagation neural network
BS	Brier score
BSS	Brier skill score
CART	classification and regression tree
CDF	cumulative distribution function
CFD	computational fluid dynamics

CRPS	continuous rank probability score
CSI	critical success index
DC	direct current
EC	efficiency coefficient
ECM	ensemble composite models
ECMWF	European Centre for Medium-range Weather Forecasts
EMD	empirical mode decomposition
EPS	ensemble prediction system
ETS	equitable threat score
FAR	false alarm ratio
FBI	frequency bias index
FLA	fuzzy logic algorithm
GBM	gradient boosted machine
GEM	global environmental multiscale model
GFS	Global Forecasting System
GRAPES	global/regional assimilation and prediction system
GRNN	general regression neural network
GSM	global spectrum model
GSS	Gilbert skill score
HRRR	high resolution rapid refresh
HSS	Heidke skill score
IEA	international energy agency
ISET	Institute of Solar Energy Technology
ISO	independent system operator
KDE	kernel density estimation
KNN	k-nearest neighbours
LS	logarithmic score
MA	moving average
MAE	mean absolute error
MAPE	mean absolute percentage error
MBE	mean bias error
MBPE	mean bias percentage error
MLP	multi-layer perceptron
MLR	multiple linear regression
MOS	model output statistics
MPE	maximum prediction error
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NWP	numerical weather prediction
PCA	principal component analysis
PDF	probability density function
POD	probability of detection
POE	probability of exceedance

PR	pass rate
PV	photovoltaic
PVPF	photovoltaic power forecasting
QDR	quantile deviation rate
QNDR	quantile negative deviation rate
QPDR	quantile positive deviation rate
QR	quantile regression
QS	quadratic score
RASS	radio acoustic sounding system
RBF	radial basis function
RBFNN	radial basis function neural network
RelS	reliability score
RF	random forests
RMSE	root mean square error
RPF	renewable energy power forecasting
RPS	ranked probability score
RPSS	ranked probability skill score
RWP	radar wind profiler
SS	skill score
SVM	support vector machine
SVR	support vector regression
TDM	temporal distortion mix
TS	threat score
TSO	transmission system operator
TWRA	Tehachapi wind resource area
WPF	wind power forecasting
WPPT	wind power prediction tool
WRF	weather research and forecasting

4 General introduction to renewable energy power forecasting

4.1 History of RPF

4.1.1 General

Renewable energy has increased substantially in recent times, and is expected to continue to do so, due to a mix of policies encouraging uptake of renewables, cost reductions and societal preferences. Renewable energy includes wind, solar, tidal, geothermal, and biomass energy, but the development and utilization of renewable energy has mostly focused on wind and solar (particularly solar photovoltaic (PV)) energy in recent times. As the sources of wind and solar PV energy are both variable and uncertain, their output is also variable and uncertain. Large-scale integration to the grid has the potential to make it more difficult to balance power generation and consumption, which might significantly increase challenges related to power system operation. Uncertainty related to output can lead to increased operating reserve requirements, for example, or can result in curtailment of renewable output.

To help meet this challenge, many industry stakeholders have carried out research on wind and PV power forecasting technologies. Various advanced power forecasting systems have been developed to reduce the impact of uncertainty, improve ability of the system to consume renewable energy (by reducing curtailment) and ensure the efficient and reliable operation of the power system. Clause 4 introduces the development, usage and technologies used in renewable energy power forecasting.

4.1.2 Development of wind power forecasting

The development process of WPF can be divided into three stages. Before 1990 there was the initial stage; from 1990 to 2000 there was the stage of rapid development; and since 2000, various technologies have emerged rapidly, with a large number of commercial providers now offering their services in this area.

There are several historical overviews of state of the art wind and solar power forecasting technologies. The most extensive ones for wind power were by Giebel et al [1][1] and Monteiro et al [2]. For solar power, Diagne et al [3], Inman et al [4] and the IEA [5] provided overviews, and Antonanzas et al [6] reviewed the subject and proposed common metrics. Kariniotakis' book [7] and a report by the Weather Intelligence for Renewable Energy COST Action [8] both offered an overview of wind and solar power together.

Brown, Katz and Murphy wrote the first paper to consider WPF in 1984 [9]. There, wind speeds were transformed to a Gaussian distribution, forecasted using an AR process, upscaled to hub height with a power law (whilst discussing the potential benefit of using a log law), and then the power was predicted using a measured power curve. Additionally, they discussed the removal of seasonal and diurnal swings in the AR components alongside forecasting intervals and probabilistic forecasts.

For pure time series models, in 1985 Bossanyi [10] used a Kalman filter with the last six values as input, obtaining a 10 % improvement in RMSE over persistence for one-minute averaged data, for forecasting the next time step. Many other authors reported the use of ANNs, FLA approaches, Box–Jenkins models and other AR efforts; a 10 % improvement in RMSE over persistence can be easily attained by a number of methods.

McCarthy [11] developed one of the earliest WPF models using offsite meteorological input for the Central California Wind Resource Area. It was operated in the summers of 1985–87 on an HP 41CX programmable calculator, using meteorological observations and local upper-air observations. The program was built around a climatological study of the site and had a forecasting horizon of 24 h. It forecasted daily average wind speeds with better skill than either persistence or climatology.

For any forecasting horizon beyond several hours, the use of NWP is necessary. In 1990, Landberg [12] and Troen [13] developed a short-term forecasting model, known as Prediktor, based on physical reasoning which is similar to the methodology developed for the European Wind Atlas [14]. The idea was to use NWP wind speed and direction, and adjust the winds according to the local characteristics, then apply a power curve, and finally modify it with the park efficiency. This model was first operationally used in a control room at Elkraft System, TSO for eastern Denmark in 1993. The Institute for Informatics and Mathematical Modelling of the Technical University of Denmark developed WPPT. WPPT has been running operationally in the western part of Denmark since 1994 and in the eastern part since 1999. The modelling system combines traditional linear models with conditional parametric models. For online applications, it is advantageous to allow the function estimates to be modified with available data. Furthermore, because the system may change slowly over time, observations should be down-weighted as they become older. The time adaptivity of the estimation is an important property because the total system consists of a wind farm or area and its surroundings, and the NWP model is subject to changes over time. This is caused by effects such as the aging of wind

[1] Numbers in square brackets refer to the Bibliography.

turbines, changes in the surrounding vegetation, and perhaps more significantly, upgrades in NWP models as well as changes in the population of wind turbines at a wind farm site or region.

In the late 1990s, the University of Oldenburg [15],[16], ARMINES [17], TrueWind [18] and ISET in Germany (now Fraunhofer IEE) also developed WPF systems. From the 2000s, many academic institutions started to develop WPF models, and companies were set up to commercialize these. Many global players are now jointly working together in the IEA Wind Task 36 on Forecasting [19].

4.1.3 Development of PV power forecasting

Compared with WPF, PV power forecasting started relatively late and the research methods and technical routes are also different due to the following aspects: (1) the solar movement law (which can precisely state the solar angle from astronomical principles) and atmospheric state fluctuation together lead to complex characteristics of irradiance fluctuation, and it is difficult to grasp the hourly variation in a day; (2) rapid and drastic changes of surface irradiance caused by the formation, dissipation, and motion of cloud clusters and variation in PV output may occur every minute in cloudy weather; this is difficult to capture in NWP models. In view of the above difficulties, scholars have carried out research on PV power forecasting using a variety of technical means and models and will continue to do so. Many of the same universities and companies which provide wind power forecasts also provide solar power forecasts.

Various institutes, including those in Europe, Asia and North America, have carried out research on solar PV forecasting since at least the early 2000s [20] to [22]. These can predict PV power output and perform resource assessment, as well as identify suitable locations for PV panels. Some are self-learning and self-calibrating systems based on physical models and advanced machine learning. These combine the historical data of output power, short-term NWP information, geographical information, calendar date and other factors, using an adaptive statistical model to predict the short-term output of a PV power generation system. Such systems have been providing power forecasting and optimization services for renewable power in Europe, North America, Australia, and other countries for more than 10 years.

4.2 Use of RPF

4.2.1 General

There are many studies on the benefits of using renewable energy power forecasting [23] to [25]. Different user groups may have different interests in electricity market and grid operation environments. For instance, system operators are interested in minimizing total system operational costs while maintaining a high level of reliability. In contrast, wind farm/PV plant operators are mainly interested in maximizing their income in the electricity market. The value of renewable energy power forecasting for TSOs and wind farms/PV plants is described separately below. It should be noted that there are now many forecasting companies (at least several dozen) across the world providing some or all of these services using a variety of combinations of the methods described later. All system operators that are integrating even relatively low shares of renewable resources have some type of forecasting system, typically procured from commercial providers though sometimes developed in-house. Such technologies are common and readily available, although, as discussed in Clause 10, there is a need to evaluate the forecasts provided by commercial providers.

4.2.2 RPF for system operations

For transmission system operators (TSOs) or independent system operators (ISOs), renewable energy power forecasting is one of the most cost effective measures for mitigating the results of variability and uncertainty. For TSOs/ISOs, renewable energy power forecasting systems are used for several purposes. For system operators, renewable energy power forecasting systems are used for two purposes that support their purpose of maintaining reliability. The first purpose is to facilitate the efficient scheduling of the system, which in market regions implies the buying (selling) of the amounts of renewable energy generation [26]. In many regions, such as Europe, this is primarily focused on the real-time market, which relies on renewable energy power forecasting updated every few minutes (though participants use the forecasts in day ahead also, TSOs tend to focus on real time operations). In the US ISOs, forecasts are also used in the day ahead processes, including day ahead markets and the day ahead reliability unit commitment. In the non-market regions day ahead (sometimes week ahead) and real time operations also require accurate forecasts.

The second purpose is to determine the amount of reserve capacity required to manage the uncertainty associated with renewables; in many regions this is set based on forecasted conditions, again at different look-ahead times from days ahead to real time. The assessment of the needs and participation in reserve markets are thus influenced by the forecast accuracy.

In recent years, many renewable power large-scale integration studies have been performed, with TSOs typically providing data backing these. A number of reports and facts from Denmark, the Netherlands, Greece, the United States, the United Kingdom, Spain, Ireland, etc., have shown that an accurate renewable energy power forecasting system plays a significant role in renewable energy integration. Therefore, sizeable efforts have been made to improve forecasting [27].

4.2.3 RPF for power trading

For wind farms/PV plants, renewable energy power forecasting is the basis for participating in the electricity market, where one exists. Short-term forecasting and ultra-short-term rolling forecasting comprise the most important information for generating schedules, bidding into the market, balancing any deviations the power plant is responsible for, and so on. For example, in Europe, the main functions are as follows: (1) In the day ahead markets, the wind farms/PV plants (often aggregated together as balancing responsible parties) participate in the market using day ahead forecasts [28]. Forecasts form the basis (together with any adjustments traders may want to make based on risk exposure) of the quantity of power that is traded in the day ahead market. If the shares of wind and solar are high, they may impact prices significantly, and can potentially result in negative prices, or prices closer to zero. If the forecast is poor, there may be a significant imbalance in the market that needs to be covered in real time. (2) In the intra-day market, the wind farms/PV plants adjust the hourly generation plan based on the rapidly updated ultra-short-term forecasting. Higher forecasting accuracy results in less balancing needed at the system level, and can result in reduced balancing payments and greater confidence in participation in the day ahead markets, where there is typically more liquidity. The narrower the gap between the generation plan and the real output is, the less TSO adjustment is required to regulate the market. As a result, the wind farm/PV plants' payments are reduced in some markets based on forecast error. In other regions, such costs are mostly socialized by the TSO.

In general, the accuracy and the application of the renewable energy power forecasting system helps inform the economic benefit of wind farms/PV plants. It is often the most important technical tool for the wind farms/PV plants to participate in the market profitably. Therefore, many professional trading teams include meteorologists, and buy their forecasts from several companies for security of supply, but also to improve the individual forecasts. Traders also use such forecasts to set their positions in the market, either with or without renewables in their portfolio.

In non-market regions, forecasts are used to determine fuel procurement, unit commitment and dispatch of conventional generation, and to set reserve requirements. They may also be used for trading with neighbouring regions.

4.2.4 RPF for operations and maintenance

Shorter horizons can also be considered for maintenance to ensure that crews can safely return from the offshore turbines in the evening. For example, in 2009 during a North Sea storm, a maintenance crew of the German offshore wind farm Alpha Ventus was trapped for 2 days due to the failure of early warning of weather prediction. The distribution system operator in German is integrating wind power forecasting into transformer maintenance routines to assess the line loading of rerouted electricity flows [29]. Additionally, even longer time scales would be interesting for the maintenance planning of large power plant components, wind turbines, or transmission lines. However, the accuracy of weather prediction decreases strongly 5 days to 7 days in advance.

4.3 Methods for forecasting renewable power

4.3.1 General

Wind power and PV power forecasting all need to consider the time scale of the forecasting plant, the spatial range, the type of input data and the forecasting model. Therefore, the division of WPF technology and PV power forecasting technology follows similar principles.

4.3.2 Classification of forecasting methods

Commonly used RPF technology classifications include those based on time scale, spatial range, modelling input data, forecasting model and forecasting form. The classification, characteristics, applicable scope, and typical methods of commonly used RPF techniques are summarized in Table 1.

Table 1 – Classification of RPF methods

Classification standard	Type	Characteristics and scope of application	Main model methods msed
Classification based on time scale	Minutes and hours ahead, intra-day (ultra-short-term)	Forecasts in this time scale predict the output in the next several minutes or hours. It can be used for commitment and dispatch, reserves optimization, situational awareness and trading.	Statistical extrapolation, assimilation methods (e.g. ensemble Kalman filter), and persistence method. For PV also use of sky-imagers.
	Day ahead (short-term)	Forecasts in this time scale can be used to make generation plans, optimize cold and hot standby, and rationally dispatch power grid resources, etc. For example, a method might predict the active power of the wind farm for three days with a time resolution of 15 min.	Physical method, dynamic statistics method.
	Days to week ahead (medium-term)	Forecasts in this time scale are mainly used to commit long start resources, schedule maintenance of power generation, transmission and distribution equipment, and schedule long duration storage resources.	Physical method based on NWP.

Classification standard	Type	Characteristics and scope of application	Main model methods msed
Classification based on spatial range	Single turbine power forecasting	This method predicts the output power of a single wind turbine.	Physical method and statistical method.
	Single wind farm or PV plant power forecasting	This method forecasts the output power of a single wind farm or PV plant.	Physical method, time series method, and artificial intelligence method.
	Wind power cluster power forecasting	This method predicts the overall output of a cluster of power plants.	Accumulation, statistical method, and spatial resource matching method.
	Distributed PV power forecasting	This method predicts the overall output of distributed, often behind the meter, PV across a region.	Statistical, physical methods.
Classification based on modelling input data	Input data without NWP data	This method uses historical power data as the input to establish the forecasting model and deduce and the power forecasting results.	Statistical and machine learning methods.
	Input data using NWP data	With topographic and geomorphological information near the power plant as input, this method deduces physical information such as wind speed and direction or irradiation and temperatures through the micro-meteorological physical model, and it further obtains the forecasting power.	Physical method and artificial intelligence method.
Classification based on the forecasting model	Persistence method	Using the measured power value at the current moment as the predicted power at the future time, this method is simple and suitable for predicting short time scales.	Persistence method
	Physical method	Considering the physical information of the wind farm, this model uses NWP data instead of accumulating a large amount of historical data. It adopts the method based on wind speed, which is suitable for short-term RPF.	Physical method
	Statistical and learning methods	This method establishes the mapping relationship between NWP data and measured power, which is suitable for all horizons. This can use simple statistical methods or more advanced machine learning or other artificial intelligence methods	ARMA and Kalman filtering method, neural network method, SVM, deep learning, and gradient boosting method.
	Multivariate combination method	This is a kind of weighted average WPF method that combines the characteristics of the wind power data and meteorological data. It adopts the weighted average RPF method.	Forecasting based on time series method and neural network.
Classification based on forecasting form	Deterministic forecasting	This method provides the future point-by-point expected value of power. Its forecasting accuracy is typically high, but the forecasting cannot quantitatively reflect the uncertainty of wind power.	Physical method and statistical method.
	Probabilistic forecasting	The forecasting of prediction intervals or the density function of wind/solar power at a future time can be divided into the parameterized method, non-parameterized statistical methods and methods based on NWP ensemble forecasts.	Statistical methods such as QR, KDE, physical methods based on NWP ensemble.
	Event forecasting	The forecasting of the wind/solar power ramp event at a future time, focusing on the significant changes of power in a time period, which can be divided into the direct method and indirect method.	SVM, ASD and BPNN, NWP ensemble based forecasts, ensemble Kalman filter.

4.3.3 Classification based on time scale

The division of the forecasting horizon is affected by various factors such as forecasting feasibility, forecasting accuracy and usage demand. Although different countries and organizations do not have a unified standard for classifying the RPF horizon, research institutions and operators have divided RPF into ultra-short-term, short-term and medium-term. Forecasting of PV power at different spatial and temporal scales is shown in Figure 1. Introduced data for PV power forecasting at different spatial and temporal scales is shown in Figure 2. Similar methods can be used for wind power forecasting.

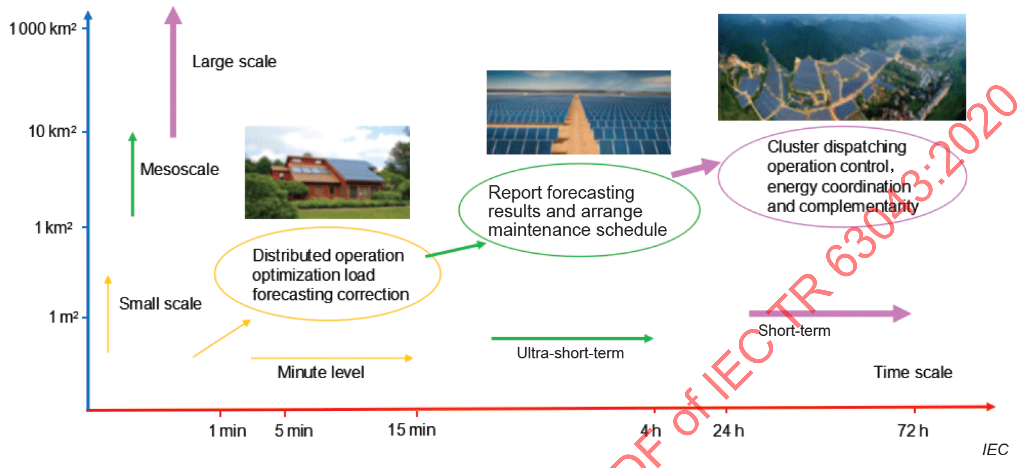


Figure 1 – Forecasting of PV power at different spatial and temporal scales

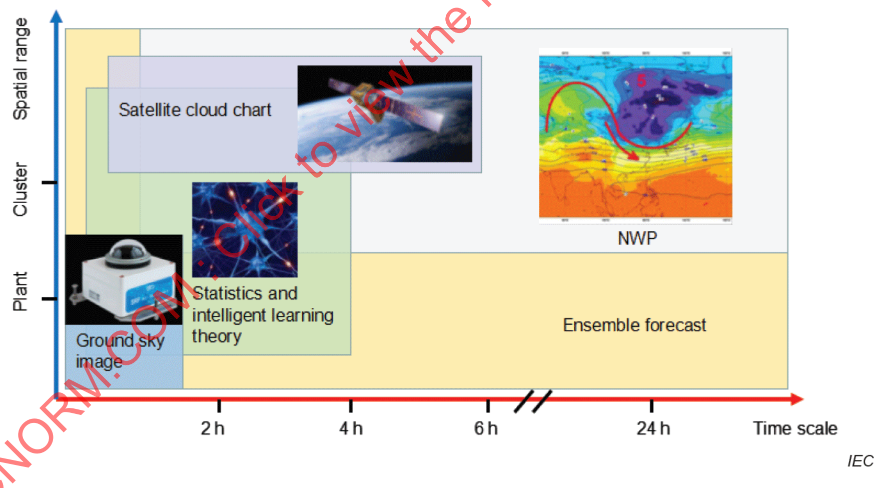


Figure 2 – Introduced data for PV power forecasting at different spatial and temporal scales

a) Minute and hours ahead

The forecast horizon of a few minutes to several hours (such as 4 h or 6 h) is sometimes called ultra-short-term forecasting, very-short-term or in more meteorological terms, nowcasting. Ultra-short-term forecasting is to predict wind/solar power 0 h to 8 h into the future. The use of forecasting in this time range is mainly for situational awareness, reserve procurement, short-term trading, balancing or economic dispatch, and commitment of quick start resources. Common methods for predicting renewable power in this time range are statistical extrapolation, two- and three-dimensional data assimilation methods (e.g. inverted ensemble Kalman filter [30]) and persistence methods using measurements of power output or meteorological measurements to assimilate or adopt long-term forecasts to the measurement state.

The time horizon of ultra-short-term power forecasting ranges from a few minutes to several hours, usually dependent on the validity of measurements to be used in the forecasting process.

For PV, surface irradiance is affected by the generation, dissipation, and motion of cloud clusters during cloudy weather, and its changes are variable, uncertain, rapid and severe, which seriously affects the accuracy of traditional forecasting algorithms. At the same time, a large number of field data shows that PV output fluctuates from 10 s to a minute scale in cloudy and fast-changing weather. Therefore, the predicted time resolution should correspond to the output fluctuation.

b) Day ahead

The time horizon of short-term renewable energy power forecasting ranges from the ultra-short-term limit up to 2 days or 3 days. In most European and many US markets, the main power trading is done before 12 noon for the coming day, i.e. based on forecasts 12 h to 36 h ahead. As an example, for State Grid Corporation of China, the requirement for short-term forecasting is to predict the wind power output of 72 h from 12 a.m. the next day with a time resolution of 15 min. China's short-term forecasting is mainly used to formulate a power generation plan. 48 h short-term forecasting in the United States is mainly used in the operation of the day ahead market (or unit commitment in the non-ISO regions), while 72 h short-term forecasting is mainly used to adjust the commitment of longer start resources, and to procure fuel and plan short-term maintenance. Similar time frames are used in Europe, Australia and elsewhere. Common methods for predicting short-term wind power are the continuous, physical, and dynamic statistics methods or physical power curve methods.

c) Week ahead and seasonal

Medium and long-term forecasting is generally used to predict the amount of electricity generated over seven days or even several weeks or months. Medium-term forecasting is mainly used to commit resources with long start-up, procure fuel, adjust maintenance schedules and adjust the operation of long duration energy storage. Week ahead and seasonal forecasting is mainly used to schedule maintenance in this time frame, and adjust very long duration storage (e.g. large hydro reservoirs).

4.3.4 Classification based on spatial range

This RPF method can be divided into single turbine or PV panel or roof installation power forecasting, single wind farm or PV power plant power forecasting and regional power forecasting according to the spatial scale of the research object. It should be noted that spatial scale and temporal scale are related – for example, second scale forecast for an entire wind farm is not going to give much realism, as the natural time scale of intra-wind farm smoothing is in the order of minutes.

Single turbine power forecasting refers to predicting the output power of a single wind turbine. The corresponding activity in the solar world is forecasting for a single PV panel or roof installation. Single wind farm power forecasting refers to predicting the output power of a single wind farm. At present, WPF research mostly focuses on single wind farms. Forecasting of an entire PV power plant is the equivalent.

Wind power cluster power forecasting refers to predicting the overall output of wind power clusters composed of multiple wind farms in a large space. Common wind power cluster power forecasting methods include the accumulation method, statistical upscaling method, and space resource matching method. For PV, similar methods are used for distributed PV forecasting, with a particular focus on those distributed PV regions where a large amount of the PV is behind the meter and therefore may not be directly observable to the system operator or those making forecasts. Spatial forecasts are described in more detail in the relevant wind and solar specific clauses later (Clause 8 and Clause 9).

4.3.5 Classification based on the forecasting model

According to different power forecasting models, RPF can be divided into persistence methods, physical methods, statistical learning methods, and multi-component combination methods. In addition, PV power forecasting has unique methods based on sky imagers and satellite cloud images.

a) Persistence method

The persistence method uses the measured power value at the current time as the predicted power at a future time. The model is simple, and its forecasting accuracy decreases rapidly as the forecasting time increases. It is usually applied to forecasting with shorter time scales. Smart persistence, which uses known information about the change in solar output based on the position of the sun, is often used for persistence PV forecasting.

b) Physical methods

The principle behind physical methods is to establish a fluid dynamics model that conforms to the meteorological characteristics of the wind farm based on the topographic and geomorphological information and physical information around it, and then use the model to predict wind speed and wind direction at the height of the wind turbine hub, leading to further predict the power of the wind farm. These methods have several important factors to consider. First, NWP data is needed to consider the physical information of the wind farm, rather than a large amount of historical data. Second, it does not correct for forecast deviations, for example in wind speed or irradiance. These make them more suitable for short-term RPF.

NWP-based power forecasting is currently the main method used in short-term forecasting with a time horizon of more than 4 h. NWP is a method for forecasting the atmospheric motion state and weather phenomena in a certain future time by numerically solving the fluid mechanics and thermodynamic equations that describe the weather evolution process. Due to the complexity of the evolution of weather conditions, it is generally necessary to rely on NWP to obtain meteorological data such as irradiance or wind speed to predict short-term and longer-term PV or wind power. The mapping relationship between predicted meteorological parameters and the output of renewable power plants can be constructed by physical or statistical methods to realize renewable energy power forecasting. NWP technology has been used in renewable energy power forecasting for many years and has performed well on the time scale of 6 h to one week.

c) Statistical methods

Statistical methods, used by themselves, do not consider the physical processes that impact wind and solar output. Based on an analysis of historical statistical and, if available, NWP data, they establish the mapping relationship between historical and/or NWP data and renewable output power. The methods directly use the historical and/or NWP data to predict the output power. They are applicable to ultra-short-term, short-term, medium-term, and long-term forecasting. Ultra-short-term RPF only requires historical data, while historical and NWP data is used together for short-term, medium-term and long-term forecasting. Compared with the physical method, the statistical method is simple and uses a single data source, but the processing ability of the mutation information is poor. Common statistical methods include ARMA, Kalman filtering and so on.

Traditional statistical methods have recently been enhanced by use of artificial intelligence methods, such as regression and machine learning; this has resulted in improvements in the subhourly time frame using just historical data, or multiple hour time scale using NWP data. These new methods use analytical equations to describe the relationship between input and output, while the former uses historical data, NWP data, or local time series extrapolated result data as input information to establish a nonlinear mapping relationship between output and multiple variables. Artificial intelligence methods require a large amount of historical observation data to build a model, but have the benefit of convenient modification and high precision. They are applicable to forecasting of all time scales. Ultra-short-term RPF only requires historical data, while historical and NWP data are used together for short-term, medium-term, and long-term forecasting. Commonly used artificial intelligence methods include various regression and machine learning methods, such as artificial neural networks, support vector machine (SVM), deep learning and gradient boosting methods.

d) Multi-component combination methods

Multi-combination methods combine the characteristics of renewable power data and meteorological data, i.e. the various methods above. They adopt the weighted average RPF method by applying different weights to different forecasting methods such as physical methods, statistical methods, and artificial intelligence methods to maximize the performance of each model and improve their forecasting accuracy. They include various machine learning and regression learning methods, and potentially other AI-based methods such as deep learning, etc. Most commercial forecasting systems use some form of combination approaches.

e) PV power forecasting method based on sky imagers and satellite cloud images

The main principle of this method, which is unique to solar, is to predict the distribution of future cloud clusters by tracking the cloud image in the sky image or satellite cloud image, then calculate the corresponding irradiance as well as solar power. The difference between the two kinds of images is mainly about the spatial resolution and update frequency of the images. Sky imagers are local and updated frequently, and as such mainly used for minute-level forecasting of a single solar power plant or local group of distributed PV, while satellite based forecasts cover a larger space, though are updated less frequently (minutes rather than seconds). They can thus be used to predict output from a few minutes to 6 h over a larger, but lower spatial resolution, area.

4.3.6 Classification based on the forecasting form

According to different forecasting forms, RPF can be divided into deterministic forecasting, probabilistic forecasting and event forecasting.

a) Deterministic forecasting

The forecasting content of deterministic forecasting is the expected value of renewable power in the future. The forecasting cannot quantitatively reflect the uncertainty of renewable power, which will lead to inevitable forecasting errors. Physical methods, statistical methods, artificial intelligence methods, and multivariate combination methods can all be applied to the deterministic forecasting of renewable power.

b) Probabilistic forecasting

Probabilistic forecasting uses NWP forecasting data as inputs to predict renewable power uncertainty information. It is an extension of single-point forecasting and can be used to assess the uncertainty of the power prediction for example for trading purposes, unit commitment, cut-off risks of wind farms, congestion or other operational risk factors in the power system. The forecasting content of probabilistic forecasting is the prediction interval or density function of power in the future, which can be divided into the statistical parameterized, non-parametric methods and NWP ensemble based methods. Typical non-parametric methods include quantile regression (QR) and kernel density estimation (KDE), while typical parameterized methods include vector autoregressive (AR) and generalized error distribution models. Typical NWP ensemble methods are the physical multi-model, multi-scheme methods or methods based on perturbation such as Monte Carlo, singular vector, breeding or ensemble Kalman filter methods (see also 5.5.2).

c) Event forecasting

Event forecasting is usually done in a statistical or a NWP ensemble based framework. The statistical methods use the random sequence of output power and the discrete renewable plant output power probability distribution to replace the power uncertainty and predict the output power probability distribution. These methods can describe the time correlation and spatial relevance of power in the future from historical date. Common scene generation methods include the Monte Carlo method, approximate Monte Carlo method, and optimal scene generation method.

The other methodology is based on NWP ensemble forecasts that generate a physical uncertainty of the current and future weather parameters of interest. The major difference of this physical methodology is that it is capable of finding extremes in the forecasts, which are of special value when extremes did not happen before or have extreme long return periods and are not present in past historical training data.

4.4 Summary

Clause 4 introduced the historical development of wind and PV power forecasting technology across the world, and summarized the technical characteristics of the current common renewable energy power forecasting system. It also expounded on the role of power forecasting for different use cases, primarily system operations, power trading and plant operations and maintenance. And finally, it classified forecasts, based on time scale, spatial range, forecasting method and forecasting form. This helps identify how improvements and assessment of forecasting can be performed for different applications, based on the specific needs of the forecast user, for example for time horizon or spatial scale.

5 NWP technology

5.1 General

Renewable energy power forecasting, especially for the hours and day(s) ahead time horizon, requires the input of meteorological forecasting data, which is mainly fed from numerical weather prediction (NWP) model outputs. The predicted output power of wind farms and PV power plants is very sensitive to the accuracy of the NWP input data. NWP technology has several complicated processes that are introduced here. To improve forecasts, the factors that affect the prediction accuracy need to be analysed in detail, and targeted methods should be used to reduce the forecasting error.

5.2 Concept and characteristics of NWP

NWP numerically solves the basic equations of atmospheric motion given initial and boundary conditions and predicts the future state of the atmosphere from a best-known initial moment [31].

NWP is divided into two models based on spatial range: the global and regional models. The global model covers the entire world, and it solves synoptic-scale atmospheric conditions. There are currently fourteen National Meteorological Services in the world running global NWPs, including ECMWF's Integrated Forecasting System (IFS), the United States' Global Forecast System (GFS) of NOAA, UK Met Office's Unified Model (global configuration), the German Weather Service DWD's Icosahedral Nonhydrostatic (ICON) model, Canada's Global Environmental Multiscale Model (GEM), Meteo France's ARPEGE model, China's Global/Regional Assimilation and Prediction System (GRAPES), etc.

Global NWP models provide the necessary boundary conditions for regional NWP models. The horizontal spatial resolution of global models is generally on the order of tens of kilometers. Due to their lower resolution, global models struggle to reflect fine changes of the wind speed, clouds, and irradiance affected by local orography and sub-grid weather phenomena. Therefore, the NWP for RPF often requires a more refined regional model. The horizontal spatial resolution of regional models is generally several kilometers, which allows them to simulate the micro-terrain and micro-processes more accurately. In addition to higher horizontal resolution, regional models also benefit from more data assimilation cycles than global models, with new observation data being assimilated more frequently. This means regional model prediction results can be more accurate than those of global models. A few of the regional NWP models include the Weather Research and Forecasting (WRF), the UK Met Office's Euro4 and UKV models, DWD's COSMO-EU and COSMO-DE, Meteo France's AROME, and GRAPES-MESO model. The overall operational flow of a regional model is shown in Figure 3. Not all models will run exactly the same components, for example machine learning may not always be used in post processing.

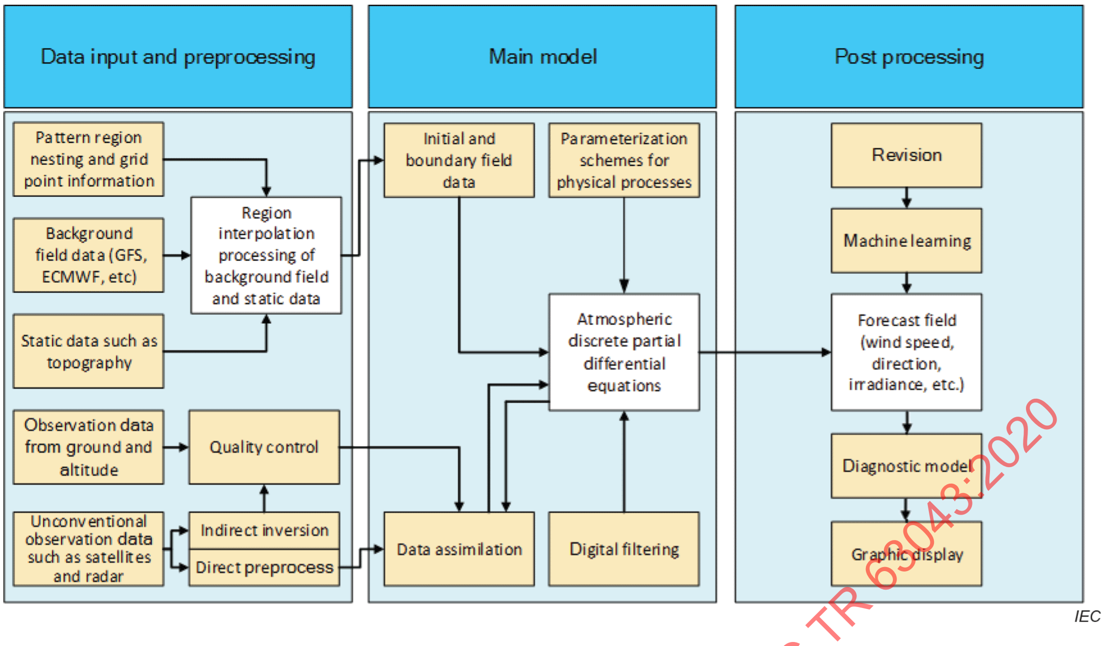


Figure 3 – Typical process for running a regional model

For RPF purposes, the most relevant meteorological fields are those closely related to renewable power generation (e.g. wind speed). Spatial and temporal resolutions are also important to help resolve small scale features. The paragraphs below highlight other specific requirements:

a) Focus on wind and solar

The magnitude of wind speed and solar irradiance directly determines the magnitude of wind power and PV output, so they are the most critical meteorological elements in power forecasting. However, to improve WPF accuracy, it is often necessary to introduce more weather fields, such as wind direction, temperature, pressure and humidity, etc. For PVPF, it is often necessary to introduce wind speed, wind direction and temperature, etc., in addition to solar irradiance.

b) Time resolution needs to be consistent with use case requirements

The purpose of RPF is to predict the output of wind and PV power to support the various use cases described in 4.2, including system operations, trading and operations and maintenance. For scheduling applications, wind and solar forecasts usually adopt a time resolution that is the same as the dispatch resolution of the system of interest, for example every 15 min in the hour ahead time frame, or hourly resolution in the day ahead time frame.

c) Quantitative forecasting

Deterministic NWP provides a single value weather forecast for a particular site and time. For example, at 10:30 on July 22, 2018, the wind speed of a certain point on a wind farm is 10,2 m/s, the wind direction is 93°, and so on. In addition, ensemble prediction systems (EPS) enable power quantitative probabilistic forecasting. This is given by a variety of probabilistic forecasting products, for example in the form of quantiles derived from the ensemble PDF, or individual outcomes from each ensemble member.

d) Forecasting look-ahead is consistent with needs of the use case

The look-ahead should be as long as that used for the processes in which they are integrated. For example, for unit commitment models used in power system operations, the look-ahead might need to be at least 72 h, and as much as a week ahead. This varies depending on the system of interest and should be based on characteristics of the power system.

5.3 Influence on RPF accuracy

5.3.1 Sensitivity analysis

NWPs of wind speed and solar irradiance are the most critical inputs for the RPF model, and the key factors affecting power forecast error. Taking WPF as an example, at the steep part of the power curve, the wind power is proportional to the cubic of the wind speed, meaning the power is very sensitive to any variations in wind speed. The power curve of typical wind turbines is shown in Figure 4. As illustrated in the power curve shown in Figure 4, when the wind speed is 8 m/s, if the predicted wind speed deviates by only 1 m/s, the predicted power error will reach 20,5 %. Therefore, power forecasting is very sensitive to the NWP error. Much of the time, wind turbines operate in the steep part of the power curve, which leads to NWP bias being the main source of wind power prediction error.

NWP skill, namely that for wind speed and irradiance, varies across different regions, phenomenal scales and historical distribution of weather events. Generally speaking, the forecasting error of complex terrain and land cover is larger than that of flat terrain and single land cover. Compared to large-scale weather phenomena (e.g. frontal winds), the forecasting error of a convective-scale process (e.g. turbulence) can be larger. For example, in China, the forecasting error in the northwest region is generally higher than that in the eastern coastal areas, mainly due to the complex terrain in the northwest, changeable dominant weather system, and scarcity of observation data. The limitation of historical data available may also compromise the analysis of extreme weather events, leading to biased results [32]. The use of extreme-value theory may hence be applied to avoid a biased statistical analysis of severe and rare weather events.

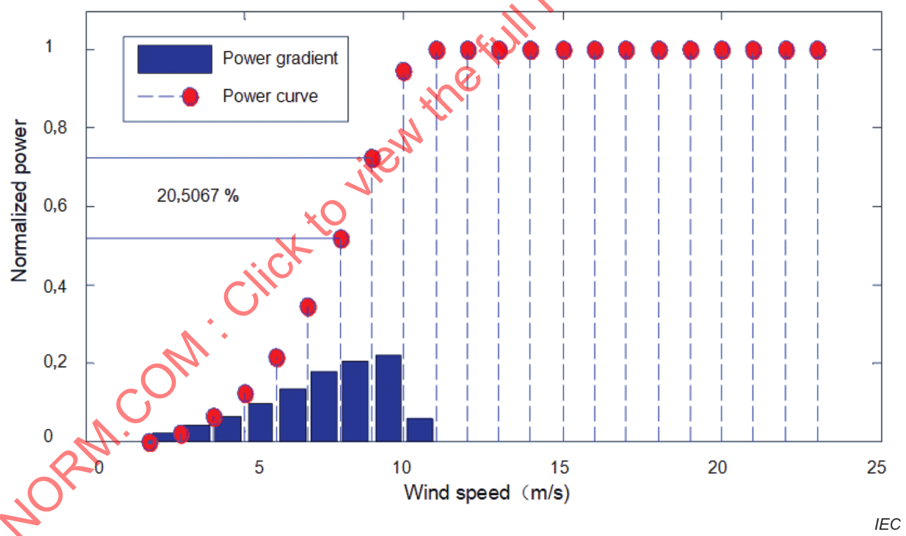


Figure 4 – Power curve of typical wind turbines

Figure 5 shows an example of time series of predicted (blue) wind speed and solar irradiance compared to what was observed (red) at a specific location, highlighting three types of deviations. The first is amplitude deviation—that is, where the trend is accurately predicted, but the maximum and minimum values of the fluctuation appear to differ significantly. The second is phase deviation—that is, where the amplitude of fluctuation is accurate, but the time of the fluctuation deviates from the actual. The third is other deviations—that is, where the deviation characteristics cannot be attributed to the first two, for example, a missing forecast or complete anti-phase of a fluctuation process often occurs when the simulation process has a large error. Wind power ramp events may be caused by sudden and/or large variations of the wind, and changes in cloud cover may lead to PV power fluctuations. Variations in wind and solar resources, leading to fluctuations in power forecast and an increase in the prediction error, can be revealed by the above three kinds of errors.

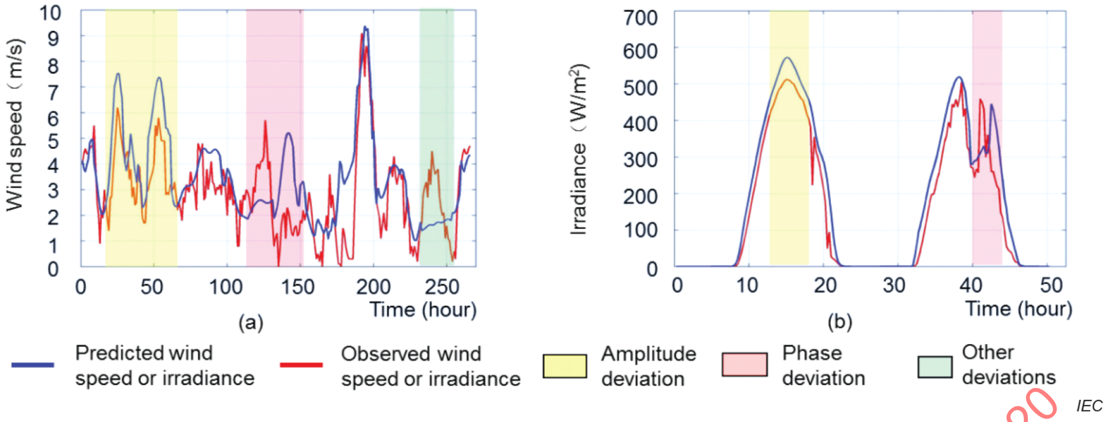


Figure 5 – Characteristics of three kinds of forecasting errors

5.3.2 Error source analysis

There are no perfect models; by definition, they are incomplete representations of the systems they aim to simulate. Given the complex physical phenomena and dependencies being represented in NWP models, the sources of error and uncertainty are manifold. First, the NWP model is a discrete computing system of continuous time and space, which inevitably leads to incomplete representations of the orography and meteorological fields. The computing machines used to perform those calculations are also limited by the number of significant digits they are able to compute. Second, the observation data is limited, and the instruments taking measurements also have limited sensitivity. These contribute to errors in the NWP background field, as the observed data is assimilated into the model at that stage. In addition, parameterization schemes describing the sub-grid and micro-scale phenomena, such as turbulence, radiation, clouds and aerosols are not perfect, because their internal mechanisms and interaction with other atmospheric components are complex, not yet fully understood and formulated. Finally, the atmosphere is an extremely chaotic and nonlinear system. As an initial value problem model, this means that the equations describing the dynamics of the atmosphere and its thermal processes are highly sensitive to the initial conditions of the NWP. Therefore, the uncertainty in NWPs increases with the forecast lead time.

There are different ways to account for and reduce NWP errors, many of which are undergone as an integral part of a model suite, in various post-processing stages, for example, by using in-situ observations to reduce systematic biases. In addition, uncertainty in the weather forecast is addressed by running ensemble prediction systems. Ensembles are generated by making controlled changes to the initial conditions, in order to emulate the chaotic behaviour of the atmosphere. The different output scenarios are then analysed to quantify the uncertainty in the forecast. Thus, a combination of methods is required to reduce NWP errors. The paragraphs below address the three types of NWP error described in 5.3.1.

a) Amplitude deviation

Despite the increase in horizontal spatial resolution of mesoscale regional models, the acceleration or attenuation of wind speed caused by sub-grid orography cannot be accurately described. Where physical corrections accounting for orographic and height adjustments, or other statistical corrections, have not been applied on a post-processing stage, this may result in a systematic amplitude deviation.

Sub-grid processes leading to cloud formation and movement, aerosol and dust dispersion, among others, affect the amount of solar irradiance that reaches the ground or plane of array. These are parameterised in the NWP model, which may contribute to a systematic amplitude deviation.

b) Phase deviation

Phase deviation means that, although the corresponding value has been correctly predicted, it has been so at the wrong time.

For irradiance, the main contributor to spatial deviation is the deviation of the predicted positions of cloud clusters [33], [34]. Due to the error of the relevant parameterization scheme and the low resolution of the model itself, it is difficult to accurately simulate the location of cloud clusters.

c) Other deviation

NWP sometimes does not predict the actual fluctuation process or predict complete anti-phase of the fluctuation, indicating that there is a large forecasting error which may be due to background field, parameterization scheme, micro-topography, observation data, or calculation accuracy, etc. It may also be caused by the sensitivity of initial errors of the NWP model in the forecasted region and period.

5.4 Technology progress for improving NWP

5.4.1 General

NWP global configurations provide boundary data to drive nested higher resolution regional models. Subclauses 5.4.2 and 5.4.3 describe the main characteristics of global and regional models.

5.4.2 Global model

Global models are complex, with details for the entire world. In addition to the atmospheric motion calculation contained in the regional model described next, the global model also contains calculation of the motion of the ocean, ice layer, etc. The major weather centres' computing ability keeps growing, but the complexity of the global models is also limited by the delivery times of the assimilated data and the necessary output, for example, if an operational global model runs 4 times a day, then the cycle of data intake, assimilation, model run and output preparation shall reliably be over within 6 h, which limits the resolution and amount of ensemble members a weather centre can deliver. There are currently a dozen national meteorological departments in the world that are able to produce and operate global models, with a smaller number of private companies also running these. Two relevant examples include:

- The European Centre for Medium-Range Weather Forecasts (ECMWF) is an independent intergovernmental organization supported by 34 states. As a research institute and an operational service, it invests in scientific and technical research to improve global NWP, through international collaboration. Figure 6 shows ECMWF's forecast skill for the 3-, 5-, 7-, and 10-day forecasts, over the Northern and Southern hemispheres. The skill over the forecast range has improved by about one day per decade, for the past 40 years. This means that the 7-day ahead forecast today is about as skilful as the 5-day ahead forecast was back in the 1990s. The convergence of the Northern Hemisphere and Southern Hemisphere curves is a result of developments in the satellite data assimilation scheme, whereby 3-D Var was upgraded to 4-D Var.
- The National Centers for Environmental Prediction, which is part of the United Stated National Oceanic and Atmospheric Administration, runs the Global Forecast System (GFS). The GFS background field is widely used in the world.

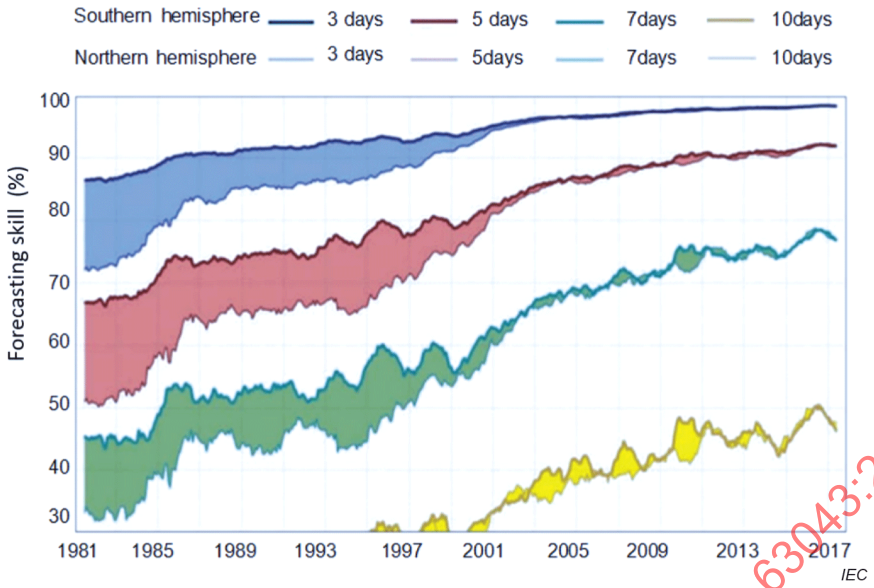


Figure 6 – Evolution of ECMWF's forecasting skills for the 500 hPa potential height [35], [54]

The growth of computing power available for meteorological research and forecasting has allowed for a continuous increase in global NWP model resolutions. Table 2 shows the main forecasting features of global NWP models around the world. The highest horizontal resolution reaches 0,1° × 0,1°. However, refinement creates a problem where increased data take longer to download and calculate. Therefore, many organizations today provide background field data of 0,5° × 0,5° or even 1° × 1° resolution.

The resolution of the global model is likely to continue to become more refined in the future. The coupling effect among the atmosphere, land, and ocean will also be improved. The data assimilation method tends to use the hybrid assimilation of four-dimensional variation, Kalman filtering, and ensemble methods, etc., to absorb more sophisticated observational data. These actions will improve the forecasting quality of wind speed, ground irradiance, clouds and vapour, etc., and are described as areas of future research in the conclusions clause (Clause 10).

Table 2 – Features of global NWP models

Model	Producer	Horizontal resolution	Vertical layers No.	Time resolution	Forecasting length	Space discretization
GFS	America NCEP	0,25°× 0,25°	64	1 h	16 days	Quadrilateral grids
FV3	America NCEP	0,125°× 0,125°	64	1 h	16 days	Hexagonal grids
ECMWF	Europe ECMWF	0,1°× 0,1°	137	1 h	10 days	Spectrum
GEM	Canada CMC	0,14°× 0,14°	84	3 h	10 days	Quadrilateral grids
GRAPES	China CMA	0,15°× 0,15°	90	3 h	10 days	Quadrilateral grids
GSM	Japan JMA	0,25°× 0,25°	100	3 h	11 days	Spectrum
UM-Global	UK Met Office	0,1°× 0,1°	70	1 h	6 days	Quadrilateral grids
ICON	Germany DWD	0,125°× 0,125°	90	1 h	7 days	Icosahedral grids

Model	Producer	Horizontal resolution	Vertical layers No.	Time resolution	Forecasting length	Space discretization
ARPEGE	Météo France	0,1°× 0,1° for Europe, 0,5°× 0,5° for other regions	105	3 h	5 days	Stretching Spectrum
NAVGEM	America FNMOC/NRL	0,28°× 0,28°	50	3 h	7,5 days	Spectrum
GDAPS	Republic of Korea KMA	0,1°× 0,1°	70	1 h	12 days	Quadrilateral grids
ACCESS-G	Australia BoM	0,25°× 0,25°	70	1 h	10 days	Quadrilateral grids
NCUM	India NCMRWF	0,125°× 0,125°	90	1 h	10 days	Quadrilateral grids
BESM	Brazil CPTEC/INPE	0,1°× 0,1°	64	1 h	10 days	Quadrilateral grids

5.4.3 Regional model

A regional model benefits from higher resolutions, enabling more accurate microphysical and convection-permitting schemes, more detailed surface orography modelling, etc. Regional models are more widely used in power forecasting than global models. They are driven by the coarse resolution data from global models, and generally have a resolution from 1 km to 10 km. A large number of regional observation data can be absorbed into regional models by using many assimilation methods. Several relevant examples include:

- The WRF model developed by NCAR is the most widely used regional model. After more than 20 years of development, WRF has many numerical methods and physical process parameterization schemes with multiple grid nesting capabilities. WRF also has many data assimilation interfaces, for example WRF-DA, GSI, DART, etc. As for irradiance prediction, NCAR developed the WRF-Solar model based on WRF and improved the physical parameterization scheme of aerosol, cloud microphysics, and radiation interaction, which can assimilate satellite and other irradiance data.
- The regional model UKV of the Met Office operates in a fine grids resolution up to a 1,5 km inner domain and a 4 km outer domain; its initialisation of the atmosphere is by 4D-VAR data assimilation with hourly cycling.
- The DWD COSMO-DE regional model in operation has a 2,2 km grids resolution and a 3 h cycling, and encompasses many local details of the landscape and related flow phenomena that are not contained in DWD global model ICON.
- The regional model of Météo France in operation is AROME-HD, it has a 1,5 km resolution and 6 h cycling.
- The regional model of the China Meteorology Administration in operation is GRAPES-MESO. It has a resolution of 3 km × 3 km, and has a 3 h cycling.

The above list is non-exhaustive and intended to illustrate different models, which are under constant development and improvement. Subclause 5.5 introduces key technologies to improve the prediction accuracy of the regional model—especially for wind speed and solar irradiance.

5.5 Key techniques for improving the forecast accuracy of regional models

5.5.1 Improve the accuracy of the initial conditions

Operating the regional model requires initial and boundary conditions. For short-term forecasting, the initial conditions are more important than the boundary conditions (for long-term climate prediction, the boundary conditions are more important than the initial conditions), and largely determine the accuracy of the short-term forecast. The improvement of the accuracy of the initial conditions is mainly reflected in the following aspects:

a) Meteorological satellites and radar observation data assimilation

Data assimilation makes the best use of observations data and background forecasts to adjust the initial field and move it closer to reality. Often, large-scale PV and wind power plants are installed in remote and sparsely populated areas. However, meteorological observation stations in these areas are often relatively sparse, and there is not enough observation data for assimilation. Meteorological satellites and radars may have a wide range of observations in potential sites for deploying wind and/or solar farms. This contributes to improving the accuracy of wind and solar forecasts [36], [37]. Currently, most satellite and radar assimilation methods are based on the variational method. The assimilation provides the benefits of initial field improvement, especially for surface water vapour, aerosol, dust, clouds, and irradiance.

b) Wind and PV power plants observational data assimilation

Wind and PV power plants generally have station observation devices. The standard observations of wind farms include the wind speed and wind direction at 10 m, 30 m, 50 m, 70 m, and at hub height. The standard observations of PV power plants include the global horizontal irradiance, direct normal irradiance, diffuse irradiance, air temperature, air pressure, relative humidity, total cloud cover, etc. Introducing local observations may help improve the accuracy of the forecast. If only the observation data of a single station is assimilated in the model, the assimilation can only work in the first few hours of the forecast, because the correction of the initial field by a single observation point will disappear with the movement of the atmosphere. But if the observation data of a cluster of plants is used, the assimilation effect will be prolonged.

There are also some challenges when aiming to incorporate these observations into assimilation: observation quality control is necessary according to existing standards, for example WMO’s standards, the observation instruments need to be calibrated regularly, the observation data shall be sent on time for data assimilation, etc. These local wind/solar farm observations can also be used in the post-processing stages to reduce forecasting errors.

For the situation where there is no observational data of wind farms and PV power plants, as much data as possible from adjacent observation sites should be collected.

c) Rapid update cycle for regional models

In the regional model running process, if the initial field is from the global background field, this is called a cold start. If the initial field is from the last forecasting result, this is called a warm start. The cold start has no cloud and water vapour information at the beginning; cloud and water vapour are gradually generated by the parameterization scheme. The warm start may retain the cloud and vapour information, but after several warm start forecasting, the forecast field may deviate greatly from the real. To rationally reduce cold start times and increase warm start times may bring benefits: computing time can be reduced as the cold start times are reduced, and the accuracy of the meteorology field can be promoted by this combination of cold start and warm start. This process is also called rapid update cycle (RUC). With the aim of not deviating from the big background field, RUC is in favour of assimilating satellite and radar data more frequently.

5.5.2 Ensemble prediction systems

5.5.2.1 General

The ensemble forecasting method has become a widely used approach to deal with the initial error sensitivity of NWP [38] and general uncertainty of the weather forecast from the deterministic models [39]. An ensemble forecasting sketch is shown in Figure 7.

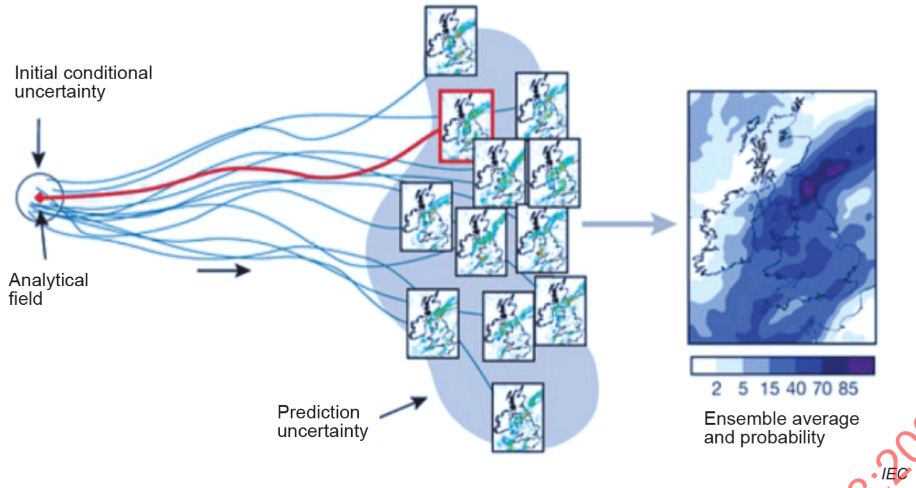


Figure 7 – Ensemble forecasting sketch [54]

Ensemble forecasting started in the 1990s (see e.g. [39] to [46]) as a means to provide information on risk management to society, where the first operational ensembles forecasting system was started in 1993.

Cheung's review of ensemble forecasting techniques (2001) [47] provides a comprehensive summary of the available methodologies and developments since the 1990s. The most important aspect to understand regarding ensemble forecasting is that an ensemble consists of a collection of forecasts that together estimate the inherent uncertainties in a numerical forecast. A properly designed ensemble should be a finite approximation of the probability density function of the atmospheric state in phase space. In other words, each ensemble member should ideally be an equally probable state of the real atmosphere.

The method of producing perturbations depends on the particular system under consideration and its associated spatial scale. The simplest way is to add random noise to the original analysis (termed as Monte Carlo forecast [48], see 5.5.2.2), but this is not an optimal method because the error characteristics in an analysis are often organised or correlated in some way. The ensemble approach requires the distribution of perturbations to be close to that of the initial state errors. This distribution in general is unknown and explains why the perturbation methodology is central to any discussion on ensemble forecasting [49].

Here the focus is on the aspects concerning the power industry and the challenges of applying various techniques for operational forecasting of wind and solar power at system operator level, at utility scale and for balance responsible parties. A more detailed description of the methodologies can be found in [47].

5.5.2.2 Monte Carlo approach

The principle of Monte Carlo forecasts is to produce random samples of an expected outcome, build a probability density function (PDF) and analyse the statistical behaviour of the outcome. Harrison [48] described the overall intention of Monte Carlo simulation as simulations that use random sampling and statistical modelling to estimate mathematical functions and mimic the operations of complex systems.

Monte Carlo simulations have a long history in statistics and any type of description of uncertainty. In the context of weather forecasting, they have also been the forerunner of today's ensemble forecasting systems. The difficulty of the approach comes with the size of the sample. Zhang and Pu [49] estimate that for a common real-model with 107 degrees of freedom, a 107 × 107 dimension calculation for estimating the PDF will be involved. This deficiency of the Monte Carlo simulations meant that since the early 1990s, other methodologies have been developed and brought into operation.

5.5.2.3 Initial conditions perturbation approach

The ensemble approach with the next longest history of operational use is those approaches where the initial conditions are perturbed. A large number of ensemble prediction systems exist today (see e.g. the TIGGE project [50]). The most common methods are:

- Breeding vector method

  This method is based on the argument that fast-growing perturbations develop naturally in a data assimilation cycle and will continue to grow as short- and medium-range forecast errors.

- Singular sector method

  In the singular vector (SV) based method, the directions of the fastest perturbation growth are identified (see e.g. [51], [42]). Singular vectors maximize growth over a finite time interval and are consequently expected to dominate forecast errors at the end of that interval and possibly beyond. Instead of using a selective sampling procedure, initial conditions may be assimilated with randomly perturbed observations [52], using different model versions in a number of independent data assimilation cycles. This Monte Carlo like procedure is referred to here as the perturbed observation (PO) approach [53].

One of the main characteristics of these approaches is that they work with so-called target times or areas. This means that the uncertainty generated by the perturbations is only valid at that target time or area. This is worth noting, because for most applications in the power industry, especially wind energy and PV forecasting, there exists a discrepancy between such meteorological lead times of these ensemble techniques and the requirements of uncertainty in the power industry. While meteorologists focus on specific times such as 24 h ahead or 72 h ahead, the power industry has a need to know the uncertainty in a continuum, i.e. in every time step of the forecast.

The major calibration methods to circumvent such issues will be given in 5.5.2.5.

5.5.2.4 Physical approaches

The only ensemble techniques that have inherent physical computations of uncertainty in every hour of the forecast lead time are the multi-model technique and the multi-scheme technique, which will be described briefly hereafter. A more detailed description of these methods and examples can be found in [47].

a) The multi-model approach

The multi-model approach uses the output of many different deterministic forecast models to create an ensemble. The drawback of this approach is that a multi-model ensemble is likely to be rather under-dispersive, because deterministic models usually suppress extremes and thereby generate too little spread. There are other disadvantages with the multi-model approach such as the difficulty to maintain many different NWP models or even collect output data from many different NWP models. However, the most critical issue with multi-model ensembles in operation is certainly the fact that deterministic models are tuned to provide best average forecasts, which often means that extremes are suppressed and therefore missed. Without extremes, the forecast spread however will not resemble a realistic uncertainty and the probability density function will be skewed. To overcome such skewness, post-processing methodologies are necessary.

b) The multi-scheme approach

The multi-scheme approach was developed due to the needs for limited areas and short-range probabilistic weather prediction, driven by studies for flooding risks [55].

The predictability horizon for precipitation as well as for other surface or near-surface variables such as wind, radiation temperature, is shorter than for more conservative parameters like mean sea level pressure, geopotential height or even upper level temperature. It is well known that those processes that cannot be explicitly resolved in the numerical models have to be approximated through different parameterization schemes. And it is exactly these schemes that are major sources of model error [54].

The success in the form of ensemble spread, correctness and consistency of the probability distribution of the multi-scheme ensemble members is however closely dependent on the choice of processes that are computed differently within the numerical prediction model. If the chosen processes that have varied formulations are not directly connected to the variables that are under focus, the spread and hence uncertainty estimate and the skill of the ensemble to predict the uncertainty of specific variables lose their value [47].

For the application in the power industry, this also means that a multi-scheme approach needs to be designed or calibrated to the target variables in order to generate the desired uncertainty of the variables responsible for the wind and/or solar generation. Here, the fast meteorological processes of (1) dynamical advection, (2) vertical mixing and (3) condensation are most relevant for the simulation of fronts and the friction between the atmosphere and the earth's surface relevant for wind and solar forecasting.

The multi-scheme approach has been demonstrated for other surface or near-surface variables to be the only approach that combines the advantages of the physical uncertainty computation with a reasonable computational effort. This is so, because the kernel of the NWP model is the same for all ensemble members. Specific physical and dynamical processes are then computed with different, but physically equivalent approaches [56] to [58]. Meng and Zhang [58] found in their experiments of a severe storm that a combination of different parameterisation schemes has the potential to provide better background error covariance estimation and smaller ensemble bias.

c) Ensemble Kalman filter approach

The ensemble Kalman filter (EnKF) method was first described by Evensen in 1994 and later by Houtekamer and Mitchell in 1998 [59] to [61]. EnKF became feasible in the context of operational atmospheric data assimilation around 2005. EnKF has its base in the computation of the background error covariance. The difficulty in the approach lies in that the true state of the atmosphere is unknown. This makes the estimation of the background error covariance difficult and expensive [49].

An example of an EnKF ensemble is the EnKF EPS at the Meteorological Service of Canada that has been used operationally since 2005 as a medium-range ensemble prediction system (EPS) and since 2007 as a short-range EPS, [61] to [63].

In the Canadian EnKF algorithm, a strategy currently with four sub-ensembles [60], [64] is used to preserve a representative ensemble. Consequently, in the absence of any differences between the model and the atmosphere or between the true and assumed observation-error statistics, the EnKF should maintain ensemble statistics that are representative of the actual error in the ensemble mean. It is thus possible to predict the analysis quality from the ensemble statistics for a hypothetical environment without model error. The negative impact of the different sources of model error on forecast quality can subsequently be quantified from the increase in ensemble spread as these components are added to the EnKF environment [63].

5.5.2.5 Ensemble calibration methodologies

It is only recently that the energy meteorology research community has developed ensemble calibration methodologies for the output from ensemble prediction systems using stochastic perturbation methods such as singular vectors or breeding [41], [42]. The issue with these approaches is that they are designed for specific lead times, which are usually greater than 48 h. It means that the spread between the ensemble members is too small in the beginning of the forecast and often too large at the end of the forecast. In the scientific community, these problems are referred to as under-dispersiveness and over-dispersiveness, respectively. The first shows so-called u-shaped histograms, while the latter shows bell-shaped histograms when the spread is evaluated objectively [65], [66].

To get around this issue, calibration methods have been developed, such as ensemble model output statistics (EMOS) methodologies [65], [66] and derivatives thereof, such as:

- BIAS corrections on each ensemble member
- bivariate EMOS [67]
- variance deficit calibration [68]
- ensemble kernel dressing [68]
- Bayesian model averaging [69]
- bivariate ensemble copula coupling [70]
- analogue ensemble calibration [71]

Univariate or bivariate Bayesian model averaging techniques are theoretically also useful, but their high computational costs and similar performance compared to EMOS approaches mean that they are not considered for practical applications at present [72].

5.5.2.6 Importance of the correct choice of ensemble forecasts

It is well-known among meteorologists that forecast errors in real-world applications arise not only because of initial errors, but also because of the use of imperfect models. Representing forecast uncertainty related to the use of imperfect models is thought to be of an even greater challenge than simulating initial-value-related errors [54]. The deficiencies and benefits of the various approaches are hence an imperative aspect to consider when choosing probabilistic forecasts for real-world applications such as wind power or solar power forecasting. In fact, the success and the value of using probabilistic forecasts is dependent on the correct choice of methodology and requirements to fulfil a specific task [47].

Because of the inherent risk in forming an ensemble with a number of deterministic models, the spread that determines the model error may not be consistent and rather small, because deterministic models often tend to suppress extremes, which are required to form a realistic description of model errors in the atmosphere.

Using an EPS formed with stochastic perturbations, a multi-scheme approach or EnKF for wind power or solar power predictions is therefore fundamentally different from using one consisting of a few deterministic weather prediction models, because severe weather and critical wind power events are different. The severity level increases with the wind speed in a weather context, while wind power has two different ranges of winds that cause strong ramping, one in the middle range and a narrow one just around the storm level (the cutoff level). Wind power forecasting models therefore have to be adapted to the use of the ensemble data.

In general, a wind power prediction model or module that is directly implemented into an EPS is different from traditional power prediction tools, because the ensemble approach is designed to provide an objective uncertainty of the power forecasts due to the weather uncertainty and requires adaptation to make use of the additional information provided by an ensemble [47].

Table 3 provides an overview and comparison of different ensemble forecasting technologies and their applicability in the power industry as aid in understanding and choosing the correct method for the application at hand.

Table 3 – Comparison of different ensemble prediction methodologies and their attributes [46], [73]

Approach	Monte Carlo simulations	Ensemble Kalman filter	Initial condition perturba-tion	Multi-model	Multi-scheme
Method	statistical	statistical	statistical	physical	physical
Member differences	random statistical perturbations of the initial conditions (“analysis”)	statistical perturbations from linearised equations – forecast error covariance	non-linearised lyapunov (bred) vectors or singular Vectors	every member is an Individual NWP model	computation of different processes inside one NWP model kernel
Application of the perturbations	random simulations to form probability density function (PDF)	differences are generated from perturbation of the initial conditions (analyse)	differences are a result of covariance error matrices of initial conditions	models as well as initial conditions are different	perturbations of the initial conditions and physical and dynamic processes inside the NWP model
No. of members	limited	limited	computation ally limited	unlimited	unlimited
Definitions and recognition of the differences	all differences are random and not unique	can be defined statistically but are not unique	can be defined statistically but are not unique	no differences can have a technical or physical reason	yes, well-defined
Expense (technically)	large, because of random simulation of uncertainty	reasonable because only one NWP model is required	low for small ensembles, large for larger ensembles	huge maintenance of many models is required	large, but manageable due to the maintenance, is limited to parts of the NWP model
Uncertainty validity	dependent on uncertainty simulations	predefined for example (>)3 days	predefined in error matrix for example (>)3 days	any forecast hour	the maintenance is limited to scheme performance
Applicability in intra-day	requires uncertainty computation to match time horizon	requires calibration	requires calibration	may miss extremes	parts of the NWP model
Applicability in day ahead	requires uncertainty computation to match time horizon	requires calibration	requires calibration	may miss extremes	yes
Applicability for futures ((>)2 days)	yes	yes	yes	may miss extremes	yes
Deficiencies	large computational effort to create valid uncertainty	requires calibration for many applications	computation ally expensive to avoid inbreeding	members deterministicall y tuned, extremes are suppressed, maintenance very expensive	computationally expensive

5.5.2.7 Best practices on the use of ensemble forecasts

It has been scientifically demonstrated since the early 2000s that ensemble forecasting is the state-of-the-art method to find the instabilities in the atmosphere that cause most forecast uncertainty and error. Thus, it is not clear whether it is actually best practise to search for the best RMSE (or similar deterministic metric) optimised forecast, because it has become apparent that optimisation on MAE/RMSE is often against, or at least not directly in line with, the interest of the system operator and/or consumers.

Over time, it has also been found that model changes that lead to better average deterministic forecasts lack internal perturbations that seem to be required to forecast instabilities. Such instabilities are difficult to forecast and therefore it is often better not to predict them at all, if only a single value can be used and if low RMSE is the target. The likelihood of a forecast extreme coming out with an incorrect phase and thereby being hit by a double punishment of errors is high. Thus, model systems that may develop a small insignificant low may not be punished as much as the system that tries to develop a bigger and stronger low, which may be more realistic.

In the light of the specific application in the power industry, where risks due to uncertainty in the weather development and the infeed of variable generation units have to be defined, it has to be noted that only ensemble forecasting systems have the capability to predict extremes that have not been observed in the data that is available to a statistical approach.

5.5.3 Establish regional customized forecasting model

The terrain, geomorphology, and weather types of different regions vary widely. An illustration of a parameterization schemes for sub-grid physical processes is shown in Figure 8. It is necessary to consider a combination of regionally targeted physical process parameterization schemes instead of using the same schemes for all regions (such as the simulation of physical process parameterization schemes Figure 8 shows). For wind farms located in complex topography and landforms, the grids in the horizontal and vertical directions should be encrypted as much as possible in consideration of calculation resources, and the parameterization schemes affecting the wind speed in the near-surface or boundary layer should be fine-tuned [74]. At the same time, cloud analysis should be used to assimilate satellite, radar, ground-based cloud maps, and other observational data to improve the irradiance prediction accuracy.

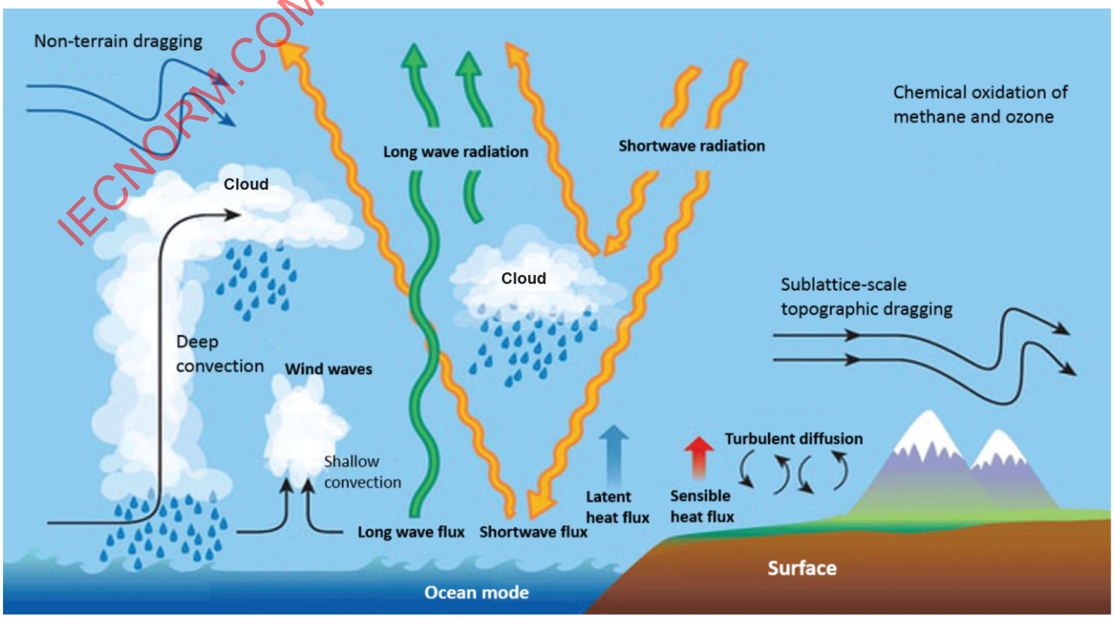


Figure 8 – Illustration of parameterization schemes for sub-grid physical processes [54]

5.5.4 NWP post-processing

For the operational NWP model, the forecasting results of the same or similar weather types are often very close, and the systematic deviations under the same weather type occur repeatedly. Based on the amount of measured meteorological observation data, a statistical post-processing correction module for the prediction model can be established—that is, the wind speed and irradiance forecasting results can be statistically revised based on the error analysis and the diagnosis of the current weather situation combined with statistics methods such as MOS and machine learning, etc. The MOS method generally achieves an effective bias correction after several years of data training, but its disadvantage is that the correction effect on short-term weather changes is not ideal. Kalman filtering is a machine learning method. It is a dynamic adaptive regression optimization process that requires only a small number of data samples and a short training period to quickly adapt the weather processes, seasonal changes and the upgrade of the NWP model. It is especially suitable for the boundary layer area where the new energy site is located. However, the disadvantage of Kalman filtering is that the bias correction of extreme events and rapid error changes is not ideal.

Machine learning methods are developing rapidly for statistics correction. IBM has used machine learning methods to make cloud cover forecasting, and the results accuracy has been superior to any model [75]. Caltech scholars consider a machine learning method based on global observation system to improve the earth system model accuracy [76]. China Meteorological Administration has applied a random forest based machine learning method to make bias correction and optimize a group of forecasts (including EC, NCEP, China T639 and Japan RJTD model), and the precipitation forecast error is globally reduced compared to the single model [77]. In 2017, ECMWF met for the topic of greyzone in modelling and forecasting [78]. This greyzone is referring to the area where traditional deterministic dynamics fails, including long/short wave radiation, convection, clouds and micro-physics, atmospheric composition, boundary layer, turbulence, etc. This greyzone may be lit up by machine learning methods in the future.

In addition, dynamic diagnostic tools such as CALMET, WAsP, and WindSim can be used to correct the wind speed change caused by micro-topography through the acceleration factor. The terrain resolution of these diagnostic modes is generally tens to hundreds of meters. The magnitude, using linear or nonlinear hydrodynamic wind field calculation methods, can effectively reduce the wind speed error caused by micro-topography.

5.6 Summary

Clause 5 focused on NWP technology applicable to RPF. The basic concept of NWP was introduced. The main characteristics of NWP that are suitable for RPF, including spatial resolution, time resolution and forecast duration, etc., were then described. Clause 5 also analysed the causes of NWP error and its impact on RPF error. The progress of NWP technology (for both global and regional model) was summarized before describing the methods to improve regional models and forecasting accuracy, including assimilation and RUC method for initial condition improvement, ensemble forecasting method for reducing the uncertainty, customized model method for localization, statistical and dynamic post-processing method, etc. As these models are important components of the overall system, it is important to understand the current state and areas where improvements may occur.

6 Statistical methods

6.1 General

Statistical forecasting techniques are applied to a very wide range of forecasting problems, and the forecasting of wind and solar power production are among them. Many textbooks have been published that supply an overview of statistical forecasting techniques [79], the details of specific methods [80] and the application of statistical methods to atmospheric forecasting problems [81]. A number of methods have been applied for short-term RPF. Clause 6 provides a high-level overview of the most widely applied methods in wind and solar power forecasting and how they are typically applied.

6.2 Methods

There are two fundamental types of statistical tools applied in forecasting applications: (1) data pre-processing methods and (2) deterministic or probabilistic forecasting model generation (training) and application techniques.

In forecasting applications, data pre-processing tools serve several functions. These include the identification and elimination of erroneous or inconsistent data (i.e. quality control), handling of missing data elements, normalization of the data and the possible grouping of data into categories or classes to facilitate the construction of better forecasting models. The grouping methods include approaches such as principal components analysis (PCA) and clustering techniques. Some of the most advanced forecasting model generation techniques implicitly incorporate data grouping concepts into the forecasting model training process and therefore the value of performing explicit grouping as a pre-processing step is somewhat dependent on the method subsequently applied to develop a forecasting model.

An enormous number of forecasting model generation techniques have been formulated, tested and widely applied in operational forecasting applications. Many of these have been and are being applied in renewable generation forecasting activities. The performance results from a wide range of wind forecasting applications indicate that no single method consistently exhibits the best performance although a small set of advanced methods have produced models that are consistently among the best performers.

One of the oldest and most basic statistical forecasting methods is linear regression. Typical applications employ multiple predictor variables. The use of linear regression with two or more predictor variables is widely referred to as multiple linear regression (MLR). MLR is extensively documented in most textbooks on basic statistical methods [81]. Many variations of MLR have been formulated. Although MLR generally does not perform as well as many of the recently formulated advanced forecasting methods when large high-quality training datasets are available, it is still a valuable tool in many situations when only limited amounts of data are available. However, the coupling of MLR with the application of sophisticated data pre-processing tools can sometimes achieve forecasting performance levels that are similar to the most advanced forecasting methods.

Another set of basic empirical forecasting methods that have been employed for many years is based on the analog concept. The analog approach has implicitly been widely applied by human forecasters throughout history by noting that a particular situation is "similar" to a set of previous scenarios and therefore the outcome of the previous scenarios can be applied as the basis for a forecast for the current situation [71]. Prior to the economic availability of high levels of computing power, this approach was generally subjectively employed with graphical displays to facilitate subjective pattern matching. However, the recent availability of high-performance computing has enabled this forecasting concept to be used in a more rigorous and quantitative manner. For example, it has been applied to the WPF under the name of "analog ensemble" (AE) by Monache et al [82]. This approach basically performs a search through the available historical records to find the cases that most closely match the current forecast situation based on a set of matching variables. The historical cases that most closely match the current situation based on a specified threshold value for a closeness parameter are identified and the outcomes of those cases are used to construct an ensemble of outcomes (i.e. the "analog ensemble"). A deterministic forecast can be created by constructing a composite of this ensemble and a probabilistic forecast can be constructed by generating a probability density function from the distribution of the ensemble members. The fundamental strength of the AE approach is that it essentially creates a custom grouping of historical cases based on each forecast situation. One of the key issues with this approach is that it does not have a basis for the forecasting of cases for which there is no satisfactory match in the historical sample. Other issues are the specification of the matching variables and the definition of closeness.

A relatively recent advancement in the regression concept that has been applied for WPF applications is support vector regression (SVR). The SVR method [83] is rooted in the support vector machines (SVM) concept [84], which originated as a tool for classification problems. SVMs are supervised learning models with associated learning algorithms that analyse data and recognize patterns. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

SVR was proposed by Drucker et al [85]. SVR is the SVM tool that is most commonly applied for WPF applications. The key attribute of SVR is that its model training and optimization process utilizes a forecasting error margin (a band) to reduce the importance of data points with small forecasting errors in the training process. In other words, it places more weight on the minimization of the errors that are beyond the margin. An example of the use of SVR in a time-series-based wind power forecasting scheme is provided in [86].

One of the advanced forecasting methods that has a long history of use in forecasting applications is artificial neural networks (ANNs) [80], [87]. ANNs are a family of statistical learning models inspired by biological neural networks (e.g. the central nervous systems of animals). ANNs are generally presented as systems of interconnected "neurons" which send messages to each other. Mathematically, neurons accept inputs and have an activation function that determines their output. The connections have numeric weights that can be tuned based on experience, making ANNs adaptive to inputs and capable of learning [77]. A common formulation of ANNs that is employed in WPF applications is the multi-layer perceptron (MLP) [80]. The MLP is a class of feed-forward ANNs that consist of at least three layers of nodes. ANNs can be very powerful tools for building complex forecasting models. However, their exceptional ability to model complex nonlinear functions also makes them prone to overfitting relationships in training datasets and thus producing forecasting models that do not generalize well. This is especially an issue in noisy and small datasets (relative to the number of predictor variables and neurons in the ANN). An example of the use of ANNs in wind power forecasting is given in [88].

In recent years a considerable amount of research and development effort has been focused on the application of the decision-tree concept [88], for the training of forecasting models. Tree models in which the target variable can take a discrete set of values are called classification trees. In these tree structures, leaves represent class labels and branches represent conjunctions of features (predictors) that lead to those class labels. Decision trees in which the target variable can take continuous values are called regression trees. The term classification and regression tree (CART) is an umbrella term first used by Breiman et al [89], to refer to both of the above procedures. Trees used for regression and trees used for classification have some similarities – but also some differences.

Many of the best performing tree-based techniques construct more than one decision tree and are often referred to as ensemble methods. One class of tree-based ensemble techniques is called “boosted trees” and these methods iteratively build an ensemble by training each new instance to emphasize the training instances previously mis-modelled. A second class of tree-based ensemble techniques is referred to as “bootstrap aggregated" (or bagged) trees. This approach builds multiple decision trees by repeatedly resampling (with replacement) the training data and then constructing a composite of the results of the ensemble of trees to produce a forecasting.

One method based on the decision-tree-ensemble approach that has seen considerable use in WPF is called "random forests" (RF). RF falls into the previously noted class of methods called bootstrap aggregated trees. It operates by constructing a multitude of decision trees at training time, and for the regression application the forecasting is the mean of the forecasting of the individual trees. The algorithm for inducing a random forest [90] was developed by Leo Breiman and Adele Cutler and they created the name "random forests". The subsampled ensemble approach used in RF corrects for the tendency of decision trees to overfit their training set. An example of the application of RF to WPF is provided in [91].

Another concept that been incorporated into recently developed advanced statistical forecasting techniques is “gradient boosting” [92]. In general, the boosting concept combines weak forecasting (“learning”) models into a single strong forecasting model in an iterative fashion. The concept can be applied in a number of ways to a wide range of forecasting models. The term “gradient boosting” refers to use of gradient descent algorithms to iteratively develop the set of weak forecasting models.

The use of decision-tree constructs as the basis for the forecasting models within the gradient boosting approach has yielded some of the best performing forecasting model generation algorithms currently available. One of the most popular implementations is known as Gradient boosted machine (GBM). Although GBM is based on the same forecasting model structure (decision trees) as RF, the two approaches have significant differences when considering which to employ for a specific application. In the RF approach, an ensemble of decision trees is fit to subsamples of the training data set. The final result is the mean of the result of each tree. However, GBM employs an iterative approach that produces the final results in a sequentially additive manner (rather than the conceptually parallel approach of random subsamples in RF). The GBM algorithm tries to fit to the residual (i.e. the error) of the previous set of trees. The result is based on the combination of a sequential set of trees with each tree in the sequence attempting to model the residuals (errors) from the previous set of trees. GBM typically produces shorter trees, which tend to reduce the overfitting of the training sample.

The GBM approach has been enhanced by a number of developers. The most widely used enhanced version is a python-based implementation called extreme gradient boosting, which is commonly referred to as XGBoost. In contrast to GBM, XGBoost uses regularized gradient boosting formulation to control over-fitting, which gives it better performance. In addition, XGBoost is formulated to achieve much higher computational performance through parallel processing.

A number of recent machine-learning-oriented forecasting competitions have been won by competitors using models generated by XGBoost [93]. It has also produced models that have consistently been among the best performers in a range of WPF applications although models based on this method have not always been the best performer. As a result, XGBoost is considered by many renewable energy forecasters to be the preferred statistical forecasting method for the statistical components of state-of-the-art wind power forecast systems when large high-quality datasets are available for training. Some examples of its performance are presented in 6.3.

6.3 Applications

6.3.1 General

Statistical techniques are employed in several ways in the forecasting of renewable power generation. There are four prominent types of applications: (1) time series forecasting using local area data, (2) refining the forecast output from NWP models (such as reducing the systematic errors for specific target variables or making forecasts of variables not explicitly predicted by NWP systems) via a procedure widely known as model output statistics (MOS), (3) construction of a composite of an ensemble of forecasts and (4) power output models that convert forecasting of meteorological variables to forecasts of power output. An overview of each of these applications is presented in 6.3.2 to 6.3.5.

6.3.2 Time series models

The objective of time series forecasting techniques is to use the predictive information contained in the values, trends or patterns of the recent history of the forecast variable or related variables. This is typically the best performing approach for wind forecasts with look-ahead times from a few minutes ahead to approximately 2 h to 3 h ahead because it has a number of advantages over NWP-based approaches for this look-ahead time frame [9].

There are a wide variety of statistical time series models that have been applied for short-term WPF. They are differentiated by (1) the amount and types (sources) of input data applied to train the models and generate forecasting, (2) the statistical methods applied to construct the forecasting model and (3) the specific configuration (e.g. values of the method's internal parameters) of the statistical methods or the algorithm applied to switch among methods or method configurations (e.g. selecting different methods or predictors by scenario such as weather regime or time of day).

The results from a recent experiment [94], conducted in the Tehachapi Wind Resource Area (TWRA) of California illustrate the variations in the skill level of 0 h to 3 h ahead forecasting of wind power production using only time series data (i.e. without NWP inputs) associated with changes in input data and statistical methods. The fundamental objective of this experiment was to assess the impact on forecast performance of data from a targeted atmospheric sensor network based on an observation targeting study [95] deployed in the vicinity of the TWRA based on a forecast sensitivity analysis. The regional generation capacity considered in this project was 2 319 MW although the total generation capacity for the TWRA at the time of the experiment was over 3 000 MW. However, some of the facilities in the region were excluded from the experimental dataset because of their limited availability of data or poor data quality. The "targeted sensors" deployed in this project consisted of three microwave radiometers, two SODARs (600 m vertical range), one mini-SODAR (200 4m vertical range), a radar wind profiler (RWP) and a radio acoustic sounding system (RASS).

The TWRA experiments analysed the impact of several factors on forecast performance including (1) the type (source) of predictor data, (2) the definition of the prediction variable and (3) the type of statistical method applied to build the forecasting model. All of the experiments were based on a 25-month dataset. This was the full duration of the project's targeted sensors deployment. The forecast evaluation was performed on a 12-month subset that extended from October 2015 to September 2016. In order to maximize the size of the training sample applied to build the statistical models, a rolling 24-month training procedure was employed. This procedure excluded the month for which forecasts were to be produced and applied data from the other 24 months to train the statistical models. Thus, 12 different statistical models were trained for the one-year forecast evaluation period.

The impact of different sources (types) of predictor data on the performance of 0 h to 3 h ahead time series forecasts for the regional aggregate 15 min average TWRA wind power production is illustrated in Figure 9 and Figure 10. All of the predictor source experiments employed the XGBoost method to train the forecasting models. The chart in Figure 9 depicts the mean absolute error (MAE) (% of capacity) versus look-ahead time for 0 h to 3 h forecasts of the 15 min average wind power production from the TWRA aggregate over the one-year period from October 2015 to September 2016 for each of 5 source-dependent sets of predictors employed in the predictor source category experiment. Figure 10 illustrates the percentage MAE reduction over persistence by look-ahead time achieved by each source-dependent set of predictors for 0 h to 3 h forecasts of the 15 min average TWRA aggregate (capacity of 2 319 MW) power production over the one-year period from October 2015 to September 2016.

The five sets of predictors are cumulative. That is, each set in the sequence employs all of the previous subsets of predictors plus a new subset. The first predictor subset is labelled "Persistence" and uses the measured 15 min average power production for the period ending at forecast time zero as the forecast for all periods in the forecast look-ahead time window (0 h to 3 h). This is a forecast of no change in production.

The second subset is labelled "Add time" and is based on a predictor pool of the time of the day and day of the year in addition to the persistence value. The addition of the date and time information to the predictor pool results in a slight reduction in MAE, especially for the longer look-ahead periods. These predictors essentially represent a type of production change climatology. Thus, it can be viewed as persistence plus an approximation of a climatological change.

The third predictor subset adds predictors from the recent time series of TWRA power production to the persistence plus climatology model. This is a type of autoregressive model and accounts for recent trends in production but uses no information from external sources (i.e. outside of the wind generation facilities). This model produces a substantial reduction in MAE relative to the persistence plus climatology model for all look-ahead times. However, the impact (i.e. the MAE reduction) is greatest for the 15 min look-ahead period for which a reduction of 23,2 % is achieved. The impact decreases with increasing look-ahead time to a minimum of 16,5 % for a 90 min ahead forecast. The MAE reduction is slightly higher for 2 h and 3 h ahead forecasts. These results indicate that there is considerable predictive information in the recent trends in power production, but the greatest value of this information is for very short look-ahead period (0 min to 30 min). This is well known, and this type of information is widely used for operational very-short-term forecasts.

The fourth subset is labelled "Add existing external data" and utilizes predictors derived from the external time series data from existing (i.e. not deployed by the project) meteorological monitoring stations in the TWRA area. These included airports and so-called "MesoNet" sites. All of these provided only in situ near-surface measurements (i.e. no remotely sensed data) such as the 10 m wind speed and direction, 2 m temperature and surface pressure. The addition of this data to the predictor pool produced an average MAE reduction relative to the "Add wind facility data" pool of 2,1 % for the 0 h to 3 h period. It ranged from a minimum reduction of 0,5 % for a 15 min forecast to a maximum of a 2,9 % for a 180 min forecast.

The fifth subset is denoted as "Add targeted sensors" and incorporates a set of predictors derived from the array of remote sensing devices. The addition of this data produces a substantial reduction in MAE and increase in skill over the persistence forecast relative to predictor sets #3 and #4 for look-ahead times of 60 min and longer. The average MAE reduction relative to the use of only the wind facility data (set #3) is 9,3 %. It ranges from approximately 2,5 % for a 15 min forecast to about 12,4 % for a 180 min forecast. There is also a considerable reduction in the MAE relative to the use of predictor set #4. This demonstrates the potential forecast benefit of data from remote sensing devices deployed at targeted locations in the vicinity of wind generation facilities.

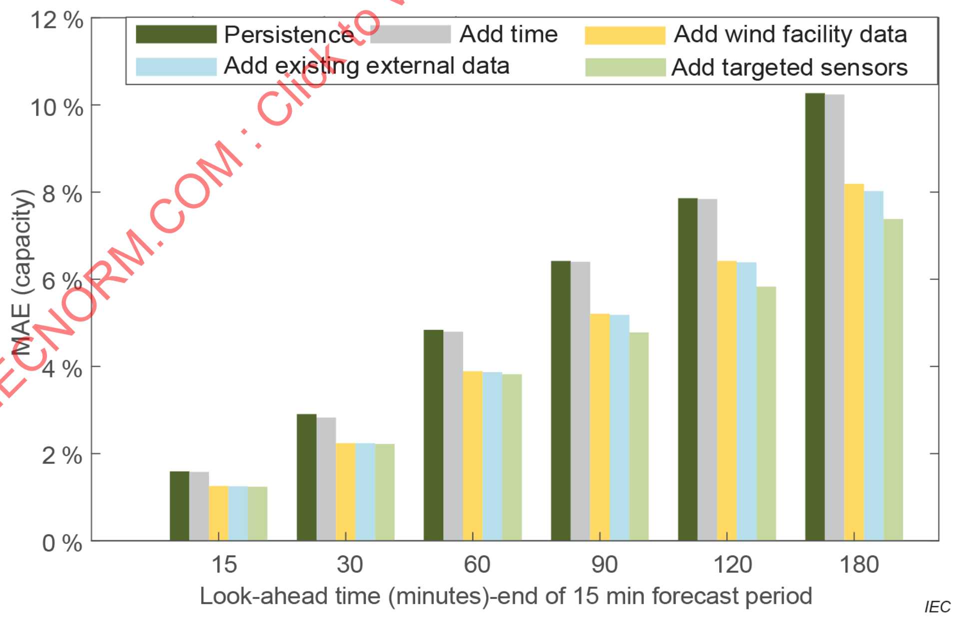


Figure 9 – MAE (% of capacity) versus look-ahead time for 0 h to 3 h forecasts of the 15 min average wind power production from the TWRA aggregate over the one-year period from October 2015 to September 2016 for each of 5 source-dependent sets of predictors employed in the predictor source category experiment [96]

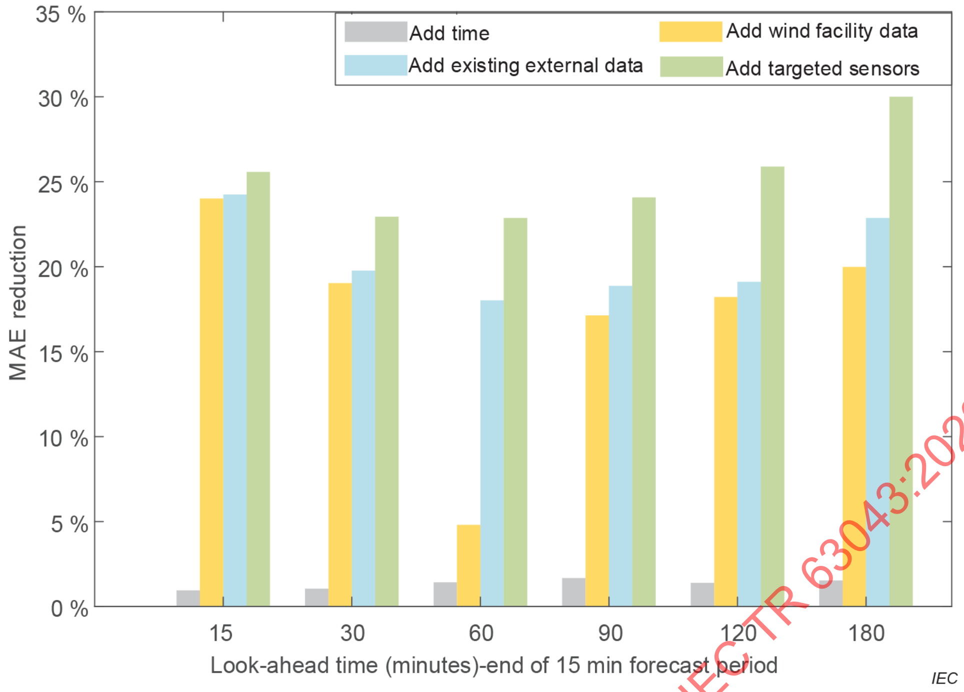


Figure 10 – Percentage MAE reduction over persistence by look-ahead time achieved by each source-dependent set of predictors for 0 h to 3 h forecasts of the 15 min average TWRA aggregate (capacity of 2 319 MW) power production over the one-year period from October 2015 to September 2016 [96]

As noted previously, the forecasting performance results shown in Figure 9 and Figure 10 were produced with an advanced machine learning method called XGBoost. The benefit obtained from using a state-of-the-art machine learning method was investigated by training a traditional MLR model with the same predictors and training sample as the XGBoost model for predictor sets #4 ("Add existing external data") and #5 ("Add targeted sensors"). The percentage MAE reduction by look-ahead time achieved by building forecasting models with the XGBoost method versus MLR for the "Add existing external data" (set #4) and "Add targeted sensors" (set #5) predictor sets for 0 h to 3 h forecasts of the 15 min average TWRA aggregate (capacity of 2 319 MW) power production over the one year period from October 2015 to September 2016 is shown in Figure 11. XGBoost provides a substantial MAE reduction relative to MLR for both predictors sets and for all look-ahead times. However, it is interesting to note that the benefit of XGBoost versus MLR is significantly greater for predictor set #5. This suggests that advanced machine learning algorithms provide more benefit over simpler traditional statistical forecasting methods when larger and more complex datasets are available.

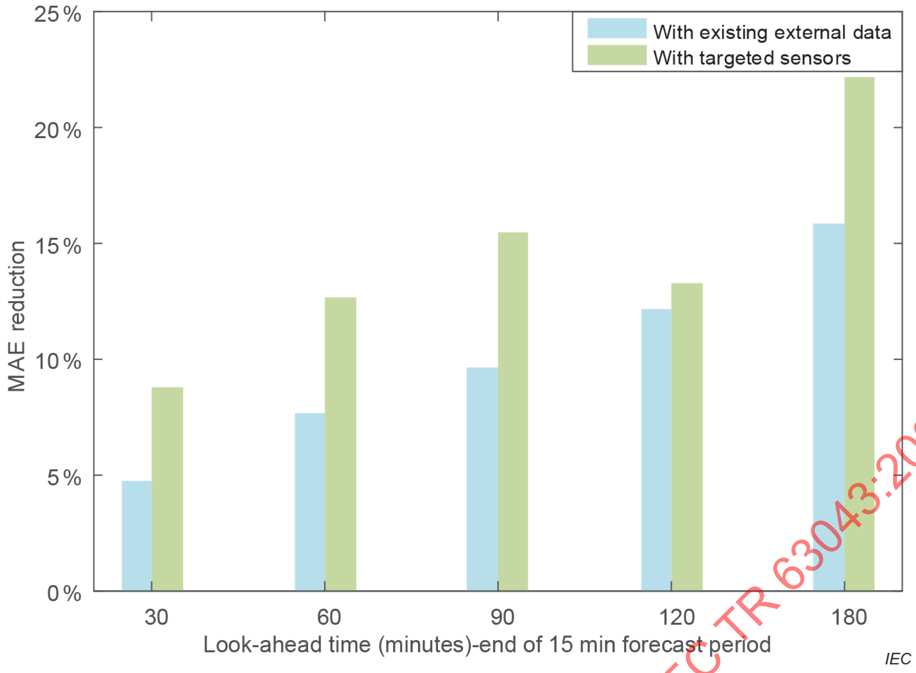


Figure 11 – Percentage MAE reduction by look-ahead time achieved by building forecasting models with the XGBoost method versus MLR for the "Add existing external data" (set #4) and "Add targeted sensors" (set #5) predictor sets for 0 h to 3 h forecasts of the 15 min average TWRA aggregate (capacity of 2 319 MW) power production over the one year period from October 2015 to September 2016 [96]

A third issue that was investigated in the TWRA project was the impact of the formulation of the predictand on forecast performance. While there is no doubt many complex predictand formulations could be concocted, two straightforward approaches are (1) direct statistical forecasting of the power production for the future time interval and (2) statistical forecasting of the change in production from time zero to the future interval and then the calculation of the future power production by adding the predicted change to the measured value at time zero. An experiment was conducted to assess the relative performance of these two approaches. The same 25-month dataset and rolling sample training procedure was used to train two XGBoost models. The first model employed the power production as the predictand and the second employed the change in power production as the predictand. A comparison of the MAE for these two sets of forecasts indicated that use of the change in production as the predictand yielded better performance. The percentage MAE reduction by look-ahead time achieved by using the "rate of change" (indirect forecasting) versus "the 15-min average power generation" (direct forecasting) as the target predictand for the XGBoost model for 0 h to 3 h forecasts of the 15 min average TWRA aggregate (capacity of 2 319 MW) power production over the one year period from October 2015 to September 2016 is shown in Figure 12. The impact is most significant for the very short look-ahead periods (15 min and 30 min) for which the MAE reduction is over 8 %. The impact decreases to about 3 % for a 60 min forecast and continue to monotonically decrease with increasing look-ahead time.

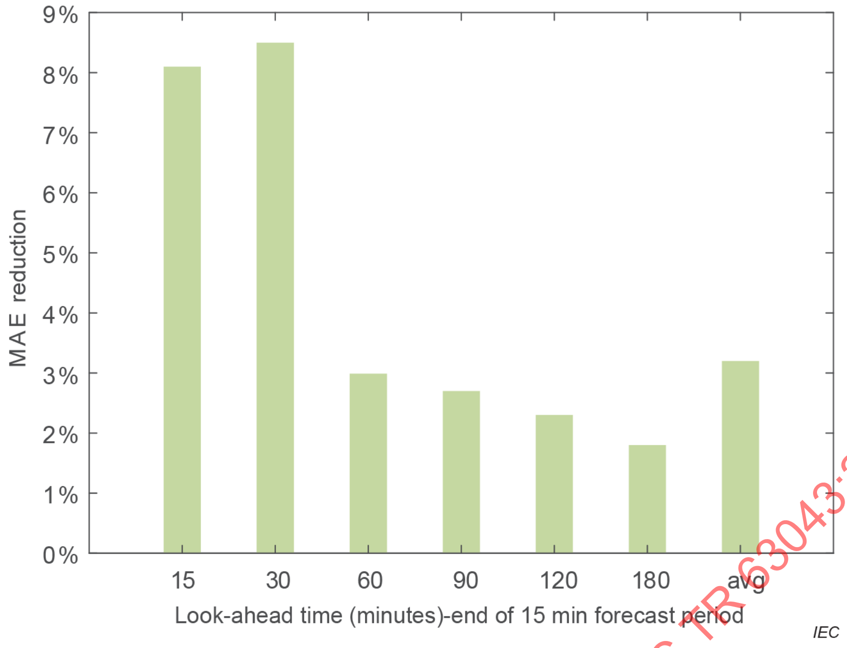


Figure 12 – Percentage MAE reduction by look-ahead time achieved by using the "rate of change" (indirect forecasting) versus "the 15 min average power generation" (direct forecasting) as the target predictand for the XGBoost model for 0 h to 3 h forecasts of the 15 min average TWRA aggregate (capacity of 2 319 MW) power production over the one year period from October 2015 to September 2016 [96]

6.3.3 Model output statistics (MOS)

A second major application of statistical techniques in WPF is model output statistics (MOS). The fundamental objective of model output statistics is to reduce the magnitude of systematic errors ("biases") in the forecasts from an underlying (often a physics-based such as an NWP system) forecasting model and/or generate forecasting of variables that are not explicitly produced by the underlying model. The term "MOS" originated in a paper by Glahn and Lowry [97] in a description of their method to predict a set of meteorological variables at airport measurement sites using output variables from early NWP models as predictors in a screening MLR model.

It should be noted that the original formulation of MOS and the use of the current use of the term cast it as a tool for the generation of improved deterministic forecasts of the variables of interest from data supplied by a deterministic NWP forecast. However, ensembles of NWP forecasts (from data or formulation perturbations of an NWP system from a single source or from multiple NWP systems from different sources) are used to produce probabilistic forecasts. There is a conceptually analogous process to MOS that is applied to the output of data from NWP ensembles in order to reduce systematic errors (i.e. biases) in probabilistic forecasts. This process is typically referred to as "calibration" of the ensemble. However, it performs a similar function to MOS: remove systematic errors. Subclause 6.3.3 will focus on the application of MOS in a deterministic forecast process. There are a number of papers in the literature (e.g. [98]) that address the ensemble calibration issue for generation of probabilistic renewable power forecasts and other applications.

As a result of its origins, the term MOS is often associated with the application of MLR methods to the output of NWP models. This approach is still widely used in general meteorological forecasting and in the forecasting of wind power production. However, the concept of MOS should be viewed in a much broader perspective. First, a broad set of advanced statistical techniques, such as those described in 5.5, can be employed for MOS applications. Second, the approach can be applied to many types of underlying non-statistical forecasting models, not only NWP systems.

There are a number of factors that should be considered in the application of MOS in the forecasting of renewable power production. One basic issue is whether to use the MOS approach to directly predict the power production or to employ MOS to improve upon the NWP forecasting of the relevant meteorological variables (such as the wind speed and direction) and subsequently employ an explicit power output model to create the power forecasts.

A second fundamental issue is the selection of the type of statistical forecasting method that is to be employed and the configuration of that method. MLR is the most basic method that is typically employed. However, in recent years, many of the advanced machines learning methods such as SVM, RF and XGBoost have been used in MOS-type components in renewable power forecast systems. It may seem logical to simply use one of the most advanced methods. However, they do not always yield better performance than basic traditional methods such as MLR and typically are much more computationally intensive. Experience has indicated that the benefits of the advanced methods are more frequently realized when large samples of high-quality training data are available. MAE in m/s of two 0 h to 18 h NWP-MOS forecasts of the maximum wind gust in a 15 min period for 33 sites over a 32-case sample of high wind events as a function of training sample size is shown in Figure 13. The results depicted by the blue line are from a linear regression model and the red line shows the performance of an XGBoost model. This chart presents the results of two different 0 h to 18 h NWP-MOS forecasts of the maximum wind gust in each of the 72 15 min intervals in the 18 h period. The forecasts were produced for 33 sites in a target region and evaluated with a 32-case sample. Thus, each case contributed 2,376 instances (72 time intervals × 33 sites) to the training sample and therefore a 10-case sample consisted of 23,760 training data points. The underlying NWP model was a custom configured system with a 1 km grid and the identical predictors were used for both MOS models. The trainings samples always excluded the case for which the forecasts were evaluated. The results indicate that the mean absolute error is reduced for both MOS models. However, the rate of MAE reduction as the sample size increases is substantially greater for the XGBoost MOS model. Thus, in this case, there is not much difference between XGBoost and linear regression for a small training sample size (10 cases) but the performance difference becomes substantial for the largest training sample size (31 cases).

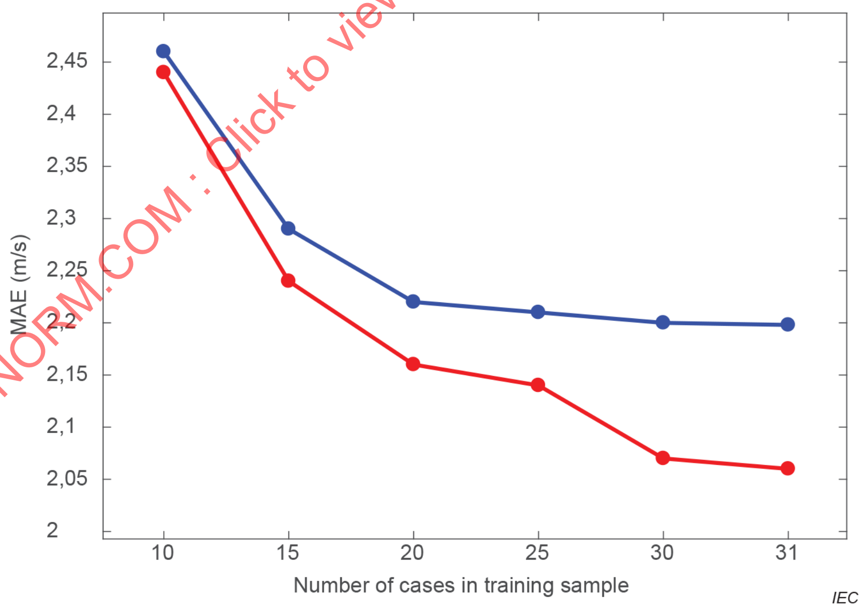


Figure 13 – Mean absolute error (MAE) in m/s of two 0 h to 18 h NWP-MOS forecasts of the maximum wind gust in a 15 min period for 33 sites over a 32-case sample of high wind events as a function of training sample size

Percentage reduction in the mean absolute error of NWP-based 0 h to 15 h wind power forecasts for the Tehachapi Wind Resource Area (TWRA) over a one-year period resulting from the application of 26 statistical forecasting methods to the output from the United States National Weather Service's High Resolution Rapid Refresh (HRRR) model is illustrated in Figure 14. The chart depicts the reduction in mean absolute error (MAE) for MOS forecasts from 26 different statistical methods applied to the output of an NWP model relative to an MLR technique applied to the same model. The model is the National Weather Service's HRRR model. The performance results are for 0 h to15 h wind power forecasts for the aggregate of all generation resources in the TWRA of California. The 26 methods are based on statistical forecasting modules available in the Python-based Scikit-learn package [99]. All of the MOS forecasts are based on the use of the same set of 9 predictor variables from each model and an identical training sample with a size of 11 months. The predictand (the target variable) was the 15-minute average aggregated TWRA (capacity of 2 319 MW) wind power production. Forecasts were evaluated over the one-year period extending from October 2015 to September 2016.

The evaluation results indicate that most of the statistical forecasting model generation methods available in the Scikit package do not perform significantly better than multiple linear regression (i.e. MLR) for either model. The methods that substantially outperform linear regression are highlighted with green and blue shading. The red shading denotes methods based on the decision tree concept and the blue shading denotes an artificial neural network model (i.e. MLP). A number of the methods perform substantially worse than linear regressions. In general, the improvement over linear regression is about 7 % to10 % in this example.

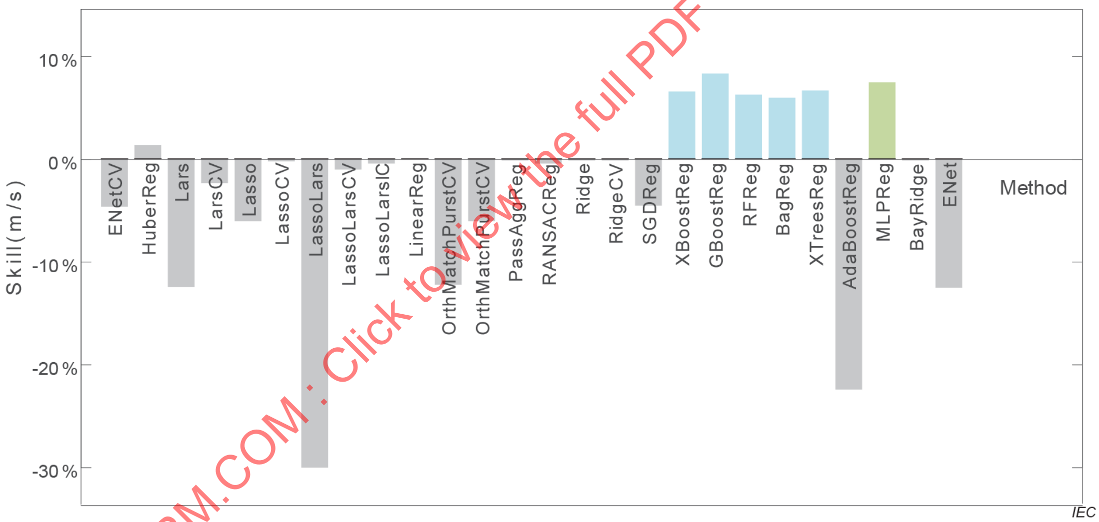


Figure 14 – Percentage reduction in the mean absolute error of NWP-based 0 h to 15 h wind power forecasts for the Tehachapi Wind Resource Area (TWRA) over a one-year period resulting from the application of 26 statistical forecasting methods to the output from the United States National Weather Service's High Resolution Rapid Refresh (HRRR) model [96]

A third issue is the representativeness of the data in the training sample. A common occurrence of this issue is when the attributes of the underlying model (e.g. the NWP model) that supplies the predictors to the MOS procedure change during the period for which data for the forecast target variable (e.g. measured meteorological data or actual power production data) is available for use in a training sample. This often is the result of modelling system upgrades such as an increase in NWP model resolution or changes to the formulation of the NWP model physics. These types of changes can significantly alter the patterns of systematic errors in the model forecasts. Therefore, the use of data from before the change in the underlying model may not represent the error patterns that occur after the change. If one chooses to use only data from after the model change, the training sample size will be more limited. Since upgrades to government-run NWP systems routinely occur, this drastically reduces the number of opportunities to use a very long training sample for a MOS application.

There are two approaches that are widely used to mitigate this limitation. First, many of the changes to the underlying model are minor and have little impact on the systematic error patterns of particular variables. Therefore, if the errors patterns in the variable of interest (e.g. wind speed) for the target locations did not exhibit any significant difference before and after the model change, a combined data sample (i.e. data from before and after the NWP model modification) can be used with little or no negative impact on MOS performance. However, a considerable amount of caution shall be used when applying this approach. The impact of a model change typically varies substantially among model variables and even between geographic areas for a specific variable. A second approach is to use predictors from an in-house model that is controlled by the user. This enables the user to generate a historical training dataset that is produced by a model with an unchanging configuration although even with this approach the characteristics of the input data, which are typically not under the control of the user, may have changed.

A fourth issue is the training sample strategy. There is an almost infinite number of possible training sample configurations. The training sample strategies are frequently classified into three broad categories: (1) static, (2) dynamic and (3) regime-based.

In the static approach, a single training sample is used to generate a fixed set of MOS equations that are infrequently updated. The training sample typically covers a long period of a year or more. The same set of equations is used for all subsequent forecast cycles. This was the approach used by Glahn and Lowry [97] in their first-ever MOS application for NWP systems. In this approach, the MOS models are typically retrained at infrequent intervals as more data is accumulated and the underlying NWP models are modified. This is a less computationally intensive approach, which was certainly an issue in the early days of NWP (1960s and 1970s).

In contrast, the dynamic approach typically uses a much smaller training sample that is frequently updated. A typical dynamic MOS configuration is a rolling 60-day sample that ends on the forecast cycle before the current forecast cycle. In this approach, the oldest data elements are deleted on each cycle and the recent ones are added. In this configuration, the composition of the training sample changes for each forecast cycle and a new set of MOS forecasting equations are used for each cycle. There are several advantages to this approach. First, the training sample is typically more representative of the current error patterns in the model forecasts since it is drawn from the current season and often from the current weather regime. Second, the use of a short-training sample avoids many of the issues associated with the impact of changes in the underlying model except immediately around the time of the model change. A disadvantage of this approach is that the MOS equations are less stable because they are derived from a smaller training sample.

The regime-based MOS approach is based on the concept of defining groups of cases for which the underlying model has similar error patterns. The use of the word “regime” in this context is often misinterpreted to mean “weather regime” but the regime definitions should be based on “NWP model error patterns” to have the most value. Certainly, in many cases the “model error regimes” are correlated with “weather regimes” and this information may be useful in designing a regime-based MOS approach, but ultimately it is the error regimes that are critical to identify. There are also many possible implementation strategies for the regime-based concept itself. A commonly employed strategy is to define a fixed set of error regimes either subjectively or via an objective approach such as a clustering algorithm or PCA (PCA). The training sample is then divided into “N” regimes (clusters) by one of these approaches and a separate set of MOS equations is derived for each regime by employing one or more statistical forecasting method(s). The production of a forecast is then accomplished by assigning the current forecast scenario to the most appropriate regime and using the MOS equations for that regime to generate the forecasting. The primary benefit of this approach is that more case-specific error regimes can be identified, which can yield more effective correction of errors in the underlying NWP (or other) model. However, a substantial disadvantage is that the division of the training sample into several sub groups results in the training sample being smaller for each regime. Since the regimes may not have equal size, some of the regimes may not have a statistically meaningful sample size. Another approach is to formulate a dynamic regime-based strategy in which there is no pre-defined set of regimes. Instead, a custom-regime is created for each forecast scenario by selecting historical cases that are the best matches for the current case. This is essentially the approach employed by the Analog Ensemble method [71]. In general, it is difficult to employ many of the advanced machine learning methods with the regime-based approach because of the small sample sizes that typically occur in the subdivided samples. However, the regime-based approach is essentially explicitly doing what many of the advanced techniques are implicitly doing (such as creating decision-trees which may implicitly represent error regimes).

The overall impact of MOS on the performance of unadjusted NWP forecasts can be quite large and highly dependent on the characteristics of the underlying NWP forecasts. Percentage reduction in the mean absolute error (MAE) of wind power forecasts relative to a baseline of a raw NWP forecast for three NWP models when a MOS procedure is applied to the NWP output (larger percentages are better) is shown in Figure 15. The data is for 0 h to 15 h wind power forecasts for the aggregated wind generation (capacity of 2 319 MW) for the TWRA of California over a one-year period. This chart depicts the percentage reduction in MAE for MOS 0 h to 15 h forecasts for the aggregated wind power production from the TWRA from each of two statistical methods (MLR and the best machine learning method for each model) relative to a baseline raw NWP forecast for each of three NWP models. The baseline raw NWP forecast was generated by interpolating the hub-height wind speed and direction from the three-dimensional NWP grid point dataset. The wind data is then used as input into a facility-scale power output model statistically derived from measured wind speed and power generation data. The three models are two models from the United States National Weather Service (i.e. the NAM and HRRR) and a custom configuration of the WRF model run with a grid cell size of 1 km. The percentage MAE reduction is 25 % to 30 % for the NAM model, 20 % to 30 % for the custom WRF model and about 60 % for the HRRR model. As might be expected, NWP models with large biases will obtain more benefit from the application of MOS. In the period of time covered by this experiment, the HRRR model had a large positive hub-height wind speed bias of about 3 m/s for the TWRA. The bias for the other two models was less than 1 m/s. Linear regression is quite effective at removing much of the bias. The benefit of the machine learning methods is largest for the custom WRF model with the 1 km grid.

NOTE The data is for 0 h to 15 h wind power forecasts for the aggregated wind generation (capacity of 2 319 MW) for the TWRA of California over a one-year period [96].

Figure 15 – Percentage reduction in the mean absolute error (MAE) of wind power forecasts relative to a baseline of a raw NWP forecast for three NWP models when a MOS procedure is applied to the NWP output (larger percentages are better)

6.3.4 Ensemble composite models (ECM)

The objective of the ECM in the context of 6.3.4 is to construct the best-performing deterministic composite forecast from an ensemble of individual forecasts. As noted previously, the statistical construction of the highest quality probabilistic forecast from a set of NWP (and possibly other) ensemble members is an analogous problem (i.e. removing systematic errors in the NWP ensemble data). However, only the deterministic version of the problem is addressed in 6.3.4 [76].

Although the term “ensemble” is frequently used to describe a set of NWP forecasts, an ensemble for a typical state of the art wind forecast system is often composed of forecasts produced by several types of methods including raw NWP forecasts [41], MOS-adjusted NWP forecasts, time series models, and feature detection and tracking models. The ECM is actually a form of MOS with the predictors coming from many forecasting models rather than a single model [42] to [45]. Therefore, most of the same techniques that are employed for the single model MOS are applicable to the ECM. But there are some issues that distinguish the ECM application from the typical single-model MOS application [39].

One issue in the ECM application is which members of a forecast ensemble are distinguishable and which are indistinguishable [40]. The members are indistinguishable if the members are generated by using the identical forecast model and randomly perturbing some aspect of the modelling system. This is typically applied to the initialization datasets for an NWP run in order to model initialization uncertainty [46]. In this case, there is no basis to distinguish one member of the ensemble from another and it is not useful to attempt to assign differential weighting to each of these members. However, the ensemble mean of an indistinguishable ensemble will typically produce a lower forecast error than an individual member. In contrast, the distinguishable members are produced through the use of characteristically different forecasting methods (e.g. different models or different configurations of the same model) or input datasets (e.g. systematically omitting or adding specific input datasets) [47]. In this case, there is a basis for differential weighting of the members since it is possible some methods or input datasets may yield better forecasts overall under specific circumstances (e.g. regimes). Therefore, a typical approach is to create ensemble mean variables and apply a MOS procedure to the ensemble-mean variables. This MOS-adjusted ensemble mean can then be used as a distinguishable input into a multi-method ensemble, some of whose members are ensembles themselves. This is sometimes referred to as a “super ensemble” [49].

A second issue in the formulation of an ECM is whether to use all of the raw predictors from each member method as input into the ECM (a “super-MOS” approach) or to first create a forecast for the ultimate target variable (e.g. hub height wind speed) from each method and then use only the target variable forecasts as input into the ECM. In general, it is better to apply MOS to each individual model and then combine the resulting forecasts since each model has its own unique error patterns, which may be difficult to distinguish in a training sample with forecasts from many other models.

A third consideration is the selection or filtering of member inputs into the ECM. It is tempting to use as many methods as could be available to the ECM. However, an indiscriminate application of that philosophy can be detrimental. The issue is based on two factors: (1) Intercorrelation of errors and (2) historical sample size available to the ECM. The first factor reflects the intuitive fact that the construction of a composite will have no benefit if the errors of all the methods are the same for each forecast period. In that case any composite of the methods will of course yield the same error as any individual method since all the errors are the same and there is no basis for distinguishing the performance of the methods. A less extreme and more typical occurrence is that the errors of individual methods are highly correlated. In this case, there will be minimal benefit in the construction of an ensemble. The point is that ensemble members with high error correlations to other members do not provide much value in the construction of the composite. However, the result can be worse than no impact. Indiscriminate use of highly correlated members in an ensemble composite can have detrimental effects. For example, in the case where some members have highly correlated errors and some don’t, a simple equally weighted ensemble average will have the beneficial impact of less intercorrelated members diluted by the highly intercorrelated members. In effect, one is implicitly placing heavier weight on the forecast represented by the highly correlated members since it essentially represents a multiple occurrence of the same forecast in the ensemble. On the other hand, poor performing uncorrelated forecasts will not be beneficial either.

The use of an appropriate statistical technique will serve to minimize the weighting of the highly correlated members and achieve an optimal blending of the high correlated and less correlated methods. This is where the training sample size can become an issue. The use of a large number of input methods along with an advanced statistical technique (with many adjustable parameters) can result in an overfitting issue if only a small training sample is available. The availability of a long representative training sample can minimize this issue, but this can often be difficult to assemble because many of the input models will change periodically.

6.3.5 Power output models

Once a high quality forecasting of the key meteorological variables is generated via a composite of the methods discussed in the previous subclauses, the meteorological data shall be transformed to a forecasting of wind or solar power production [55]. This is accomplished with the use of a power output model. This type of model can be constructed on different scales (e.g. single turbines, clusters of turbines. a generation facility, regional aggregate, etc.). However, the most common approach is to construct a facility-scale output model. The facility-scale power output model represents the relationship between the meteorological variables and the facility-scale power generation but also implicitly or explicitly accounts for non-meteorological effects.

Although the basic concepts and objectives are the same, a number of different modelling strategies can be employed and there are notable differences between them. The most fundamental option is whether to use an explicit or implicit power output model. The implicit approach predicts the power output at the MOS or ECM step by training the statistical model to go directly from forecasting of meteorological variables to power output in the MOS or ECM process. In this case the equivalent of a power output model is implicit within the MOS and/or ECM equations. This simplifies the forecasting process and also removes some of the need for high quality meteorological data from the generation facility. The disadvantage of this approach is that it makes it more difficult to separate the components of the forecast error that are associated with the meteorological forecasting from those that are associated with the power production model. This makes it more difficult to analyse the performance and refine the forecasting system. Results indicate that the explicit approach generally yields better forecast performance for facilities that supply high quality meteorological data while the implicit approach may be as good or better when high quality meteorological data is not available from the facility.

If the explicit approach is pursued, then the modelling method and the granularity of the model shall be selected. There are two fundamental types of power output models: (1) process (physics)-based and (2) bulk statistical. The process-based models attempt to simulate the behaviour of the facility based on the physical layout and hardware specifications of the facility. These models use meteorological and operating data and the hardware and layout specifications as input into quasi-physics-based model equations to calculate the power output. These models have considerable detail and the engineering processes are generally modelled quite well. However, the detail of input data required for these models to perform well is generally more than is typically available in typical operational applications. This generally limits their performance in an operational setting.

In most operational applications, statistical power output models are employed because they provide better performance. These models are statistical relationships between measured meteorological data and actual power output. Any of the statistical methods previously described can be employed for this purpose. However, the data may be noisy, and, in many cases, simple models will perform as well or better than more sophisticated methods. The facility-scale models can be constructed at different levels of granularity. For example, in the case of a wind generating facility, statistical relationships could be constructed for the output of each turbine or for the aggregated output of the facility. The aggregated approach is more typically employed because the data is often not available at higher granularity and even in cases where such data is available, the impact of modelling with additional granularity on forecast performance is often minimal, but the level effort is somewhat greater.

7 Wind power forecasting (WPF) technology

7.1 General

Clause 7 presents five main aspects of WPF, including short-term WPF, ultra-short-term WPF, probabilistic WPF, WPF for ramp events, and WPF for wind farm clusters. It also briefly introduces medium-term and long-term WPF and WPF for offshore wind farms.

7.2 Short-term WPF

Short-term WPF is usually considered to forecast wind power output from the ultra-short-term WPF limit up to 48 h, 72 h, or 168 h, which is mainly applied to the commitment of power generation and the clearing of markets. There are three short-term WPF approaches: physical approaches, statistical approaches, and a combination of the two.

Input and output parameters of the three-days-ahead WPF are shown in Figure 16. Historical NWP data, current NWP data, and the measured value of wind power output are applied as input parameters of the forecasting. The time resolution for the short-term WPF is usually 15 min or 30 min, though could be as high as 5 min. To improve accuracy, the WPF results—which are the equivalent measurement values of wind power output for a future moment—are usually applied as the input parameters of the following WPF [100].

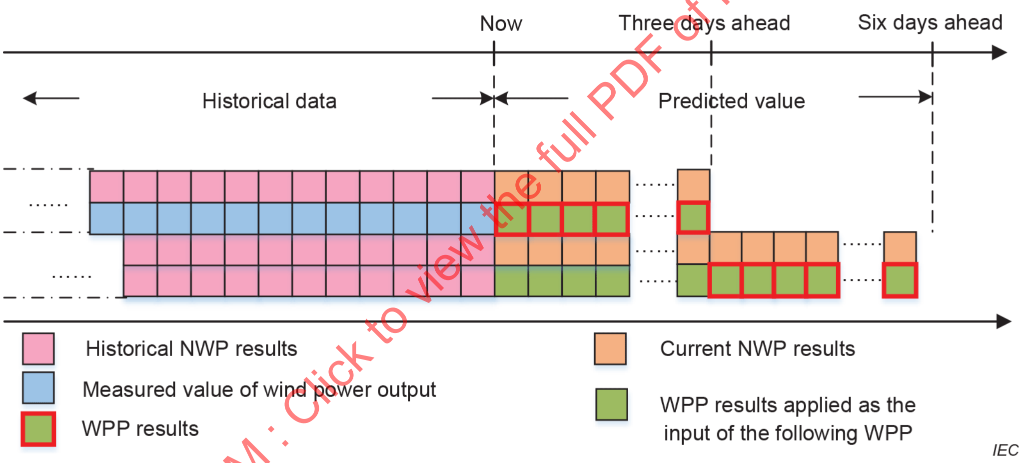


Figure 16 – Input and output parameters of the three-days-ahead WPF

7.2.1 Relationship between wind power output and meteorological elements

7.2.1.1 General

Wind power output is related to several meteorological elements including wind speed, wind direction, temperature, air pressure, and humidity.

7.2.1.2 Relationship between wind power output and wind speed

Wind speed is the most important meteorological element for wind power generation. The wind power output of a wind turbine is proportional to the cubic of the wind speed, which can be described according to Formula (1).

$$P = \frac{1}{2} C_{\mathrm{P}} A \rho v^3 \qquad (1)$$

where

P is the output power of a wind turbine (unit: kW),

C_P is the energy coefficient of the turbine,

A is the rotor swept area of the turbine (unit: m^2),

ρ is the air density (unit: kg/m^3), and

v is the wind speed (unit: m/s) [101], [102].

Figure 17 shows the wind power output for a typical 2 MW wind turbine at different wind speeds under the air density of 1,225 kg/m^3.

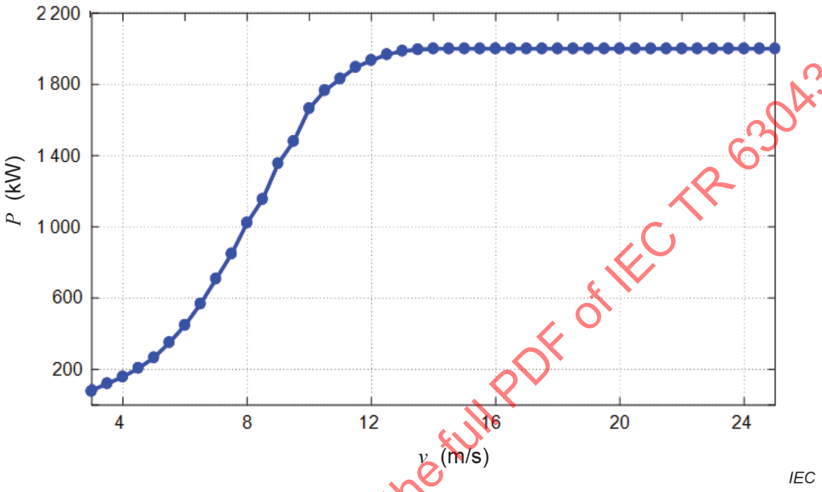


Figure 17 – Wind power output at different wind speeds under air density of 1,225 kg/m³ (a typical 2 MW wind turbine)

In Figure 17, the cut-in wind speed of the wind turbine is 3 m/s; the cut-out wind speed is 25 m/s; the effective wind speed is in the range of 3 m/s to 25 m/s. When wind speed is 14 m/s, the wind power output reaches the maximum value of the wind turbine. When the wind speed is in the range of 3 m/s and 12 m/s, the small variations of wind speed will lead to great changes in wind power output. Furthermore, the actual relationship between the forecasted wind speed and generation output frequently differs from the power curve obtained from a manufacturer due to the wake effect of the wind turbines and the ground roughness of the wind farm. Without considering the influence of the wake effect and the ground roughness, the power output of the wind farm is approximately proportional to the cubic of the wind speed when the speed is below the rated one.

7.2.1.3 Relationship between wind power output and wind direction

A wind farm contains many wind turbines geographically arranged according to certain rules. When the wind passes through them, part of its energy is converted into electricity and its speed decreases, leading to the reduction of energy obtained by downwind turbines. This is called the wake effect [103], [104].

To describe the influence of the wind direction on the wind power output, the EC of the wind farm is defined according to Formula (2).

$$\eta = P_m / P_f \qquad (2)$$

where

P_m is the measured wind power output of the wind farm at a certain wind speed and wind direction (unit: MW),

P_f is the ideal wind power output of the wind farm without the wake effect (unit: MW).

The EC distribution of a wind farm at different wind speeds and directions is shown in Figure 18.

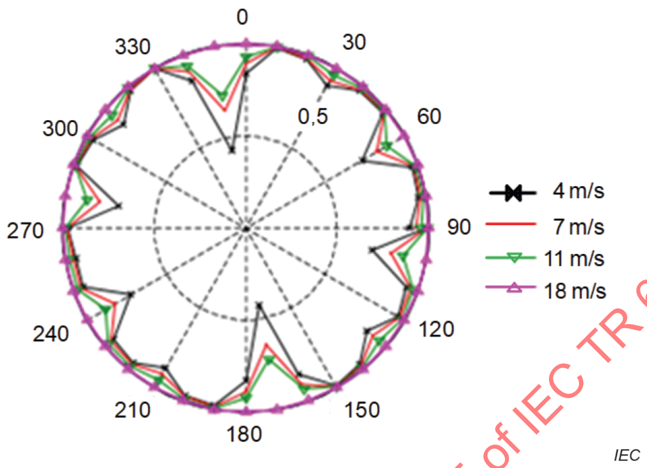


Figure 18 – EC distribution of a wind farm at different wind speeds and directions

Figure 18 shows that under low wind speeds, the EC of the wind farm in some directions is relatively low due to the influence of its ground roughness and the wake effect of the wind turbines. Further, the higher the wind speed, the higher the EC of the wind farm. When the wind speed exceeds a certain value, the wind direction no longer affects the output power of the wind farm.

7.2.1.4 Relationship between wind power output and air density

In Formula (1), one of the parameters is air density ρ, which is related to the amount of wind energy captured by wind turbines. Air density significantly affects the wind power output in high-altitude areas. Figure 19 shows one example of the wind power output of a wind turbine under different wind speeds and air densities. At a certain wind speed—from 4 m/s to 12 m/s—the output power of wind turbines increases with the air density.

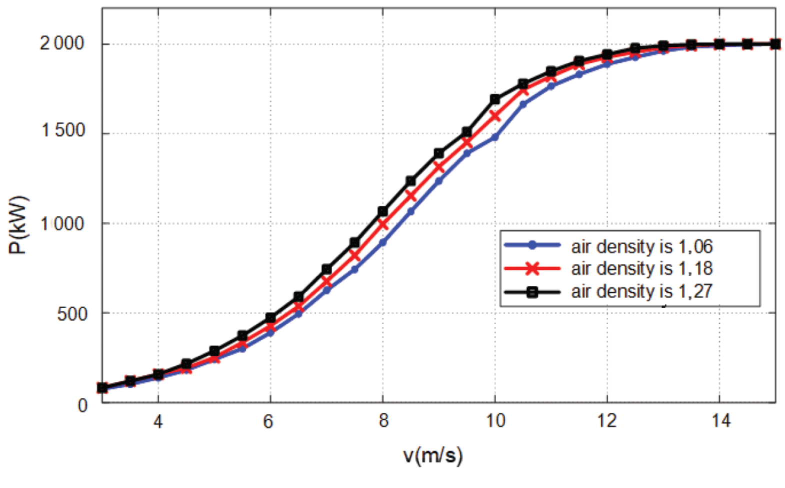


Figure 19 – Wind speed and wind power curves of wind turbines at different air densities

Air density can be described according to air pressure, temperature and humidity, the relationship of which is given in Formula (3).

$$\rho = \frac{1{,}276}{1+0{,}003\,66t} \times \frac{p-0{,}378\,p_\mathrm{w}}{1000} \tag{3}$$

where

p is the air pressure (unit: kPa),

t is the temperature (unit: ℃), and

p_w is the air pressure produced by water vapour, which is strongly related to air humidity (unit: hPa).

Formula (3) shows that air pressure, temperature, and air humidity are key factors affecting air density, which should constitute the input parameters for WPF besides wind speed and wind direction.

7.2.2 Framework of short-term WPF

Figure 20 shows a typical framework of short-term WPF, which is divided into three parts: input and output parameters of WPF, and the methods for WPF. Recently, improvements in combining geographically distributed time series and using a grid of NWPs have improved the state of the art shown in Figure 20.

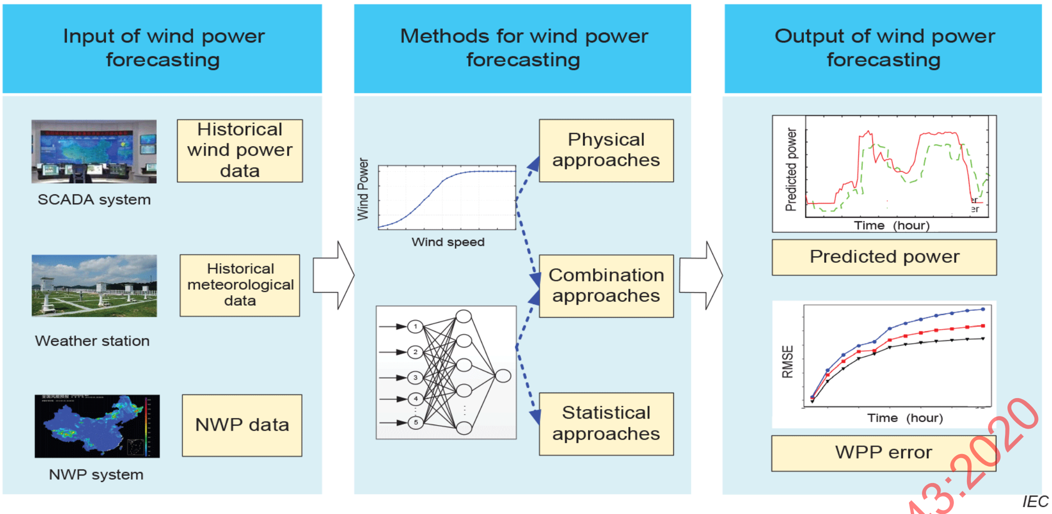


Figure 20 – Typical framework of short-term WPF

The input parameters of short-term WPF include historical wind power data, historical meteorological data, and NWP data. NWP data should be applied to the short-term WPF as the forecasting time scale exceeds 24 h. The physical parameters, such as wind speed, wind direction, temperature, air pressure and humidity have varying importance to wind power output, so it is essential to choose specific variables in the NWP data to improve the training speed of the forecasting models.

The output of short-term WPF is divided into two parts, including the forecasted wind power value of a specific period and the WPF error.

7.2.3 Short-term WPF methods

7.2.3.1 Physical approaches

Physical approaches are the earliest adopted approaches in the WPF. Figure 21 shows the principle of the short-term WPF based on physical approaches.

Three kinds of data—the roughness data, contour data and NWP data of the wind farm—are applied as input parameters to three kinds of models, which are the turbulence model, roughness model and terrain change model, respectively. Based on the roughness data, the roughness change model is applied to calculate the changes of wind passing through different underlying surfaces of the wind farm. Based on the contour data, the terrain change model is applied to measure the wind changes with the topographic changes of the wind farm. Based on the wind speed and wind direction in the NWP data, the turbulence model can be applied to estimate form drag on topography. Finally, the forecasted wind speed and wind direction at the hub height of the wind turbines are obtained.

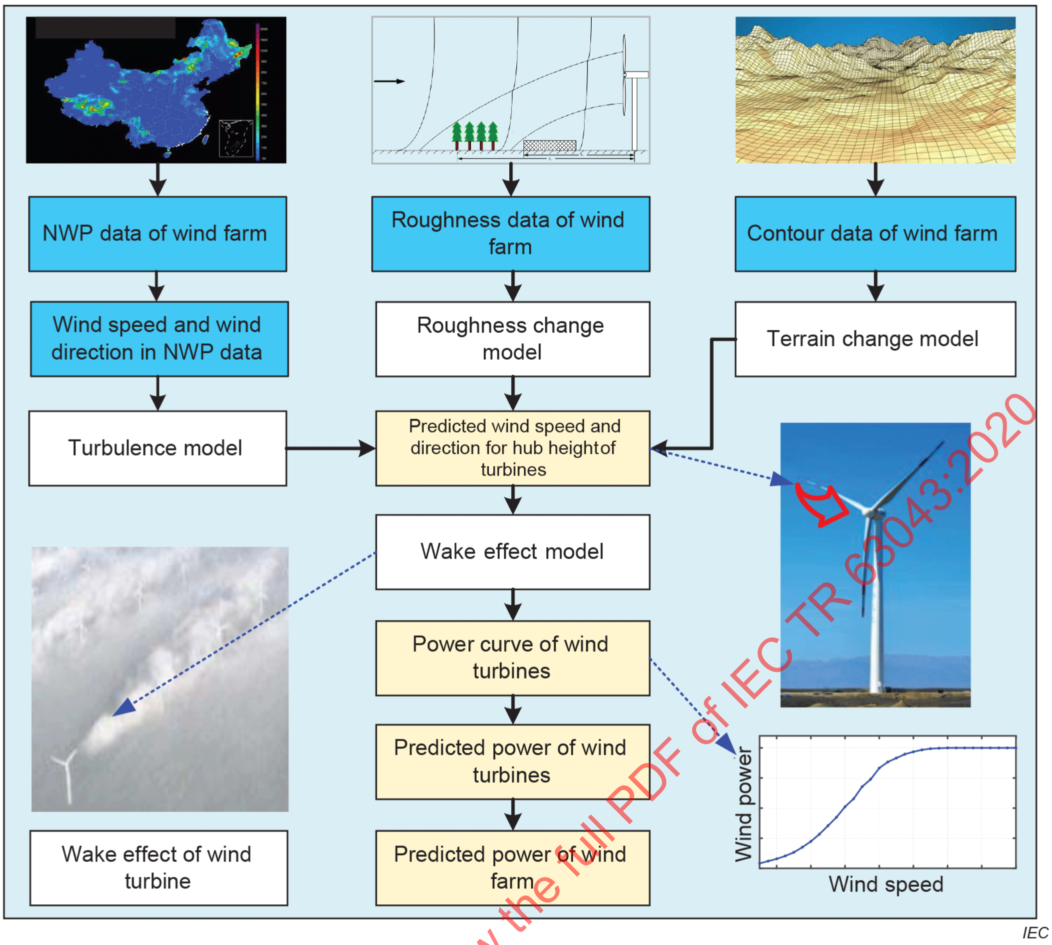


Figure 21 – Principle of short-term WPF based on physical approaches

Further, the wake effect model of wind turbines is applied to adjust the wind speed. Wind turbines extract energy from the wind. Thus, the downstream wind speed is slower than the upstream wind speed. The wake effect is the aggregated influence of wind speed reduction caused by the interaction of wakes between wind turbines. Based on the wake effect model, the predicted wind speed of each wind turbine at the hub height is obtained. According to the power curve of wind turbines, the predicted wind speed and wind direction can be converted into wind power output. Finally, the WPF results of the wind farm can be acquired by accumulating the WPF results of all the wind turbines in it [104].

The physical forecasting model, which is not based on historical data, requires only the location information, topographic information and wind turbine information. It can be applied to wind farms with few restrictions (e.g., new wind farms). However, the physical approaches are based on the physical process of wind power generation, which cannot be adjusted by forecasting errors because there is no feedback module within the model. For wind farms with significant NWP errors, the error of final WPF results will be further amplified.

7.2.3.2 Statistical approaches

Based on one or more statistical approaches, the relationship between historical NWP data and historical wind power output of a wind farm is established, which can be applied to the WPF of the wind farm in the following several days based on the current NWP data. Figure 22 shows a flowchart of short-term WPF based on statistical approaches. The time series method, Kalman filter method, and artificial intelligence methods (such as ANNs, SVMs, KNNs, and FLAs) are widely adopted statistical approaches in the short-term WPF [105] to [112].

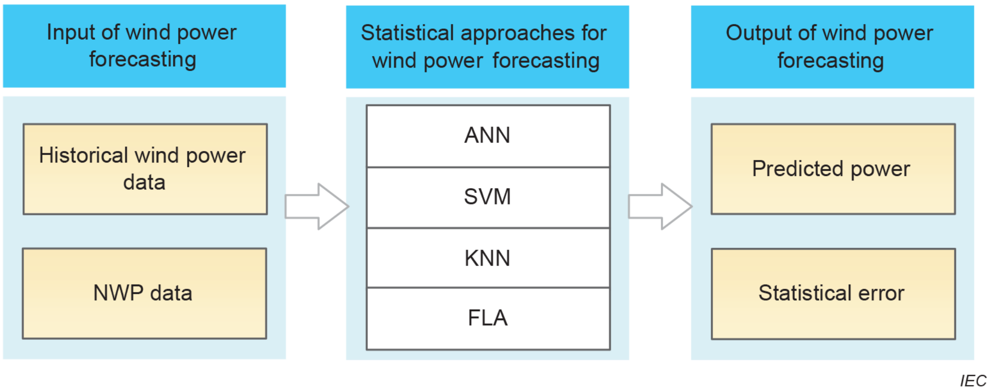


Figure 22 – Flowchart of short-term WPF based on statistical approaches

a) Time series method

The time series of wind or wind power contains the historical behaviour of a system that produces the sequence. For the time series method based WPF, current and previous limited length observation data is inputted as the training data, based on which the corresponding forecasting model can be established. There are three types of time series methods: AR model, MA model, and ARMA model. A large amount of data is usually required for training the models [113], [114].

b) Kalman filter method

For Kalman filter based WPF, the state space model—the state variable of which is usually wind speed—is established by the Kalman filter algorithm, which assumes that the statistical characteristics of the forecasting system are known. The challenge of the Kalman filter based WPF is to estimate the statistical characteristics of noise [115].

c) ANN

Due to its robustness, fault tolerance and strong generalizability, ANN has become the most widely applied statistical method in WPF. Figure 23 shows a short-term WPF model based on ANN. Wind speed, wind direction, air temperature and air pressure can be applied as the input parameters. The outputs of the forecasting are the forecasted power and forecasting error.

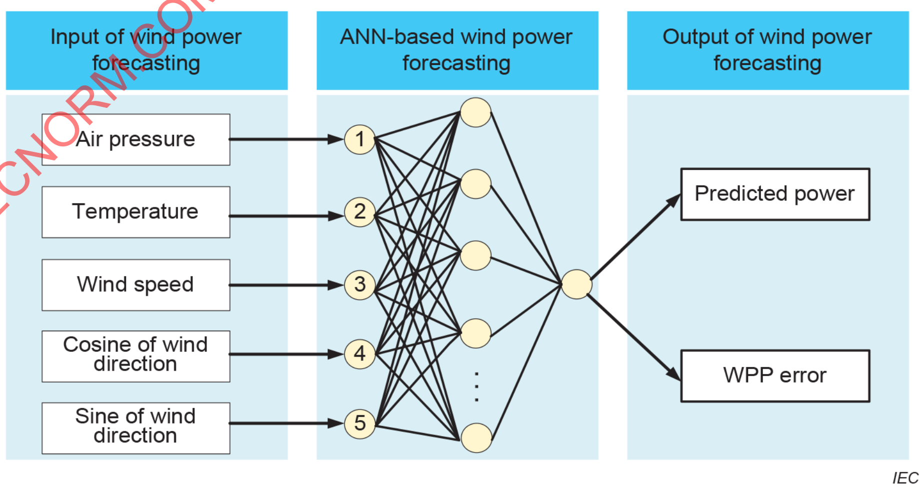


Figure 23 – Short-term WPF model based on ANN

Among the different types of ANN models, BPNN, RBFNN, GRNN and Elman network are popular methods reported to be successfully applied in WPF. BPNN has been widely applied to WPF due to its excellent characteristics of nonlinear mapping ability, generalizability and fault-tolerant ability. But the disadvantages of the BPNN-based WPF are also remarkable, including requiring too much time for the training process, difficulty in identifying the numbers of hidden layers and ease of falling into the local minimum value without getting the global optimal value. RBF also has the excellent characteristic of generalizability, and it is not easy to fall into the local minimum value during the training process, so it is widely applied to WPF. But the base function of RBF is not easy to select, and the convergence speed of RBF training is also slow. GRNN has a simple network structure and fast training speed, which is suitable for a small number of training samples. The Elman network has excellent performance for online WPF due to its dynamic feature mapping capability. However, it is easy to fall into the local minimum value during training, and the forecasting accuracy of the Elman network decreases significantly in obvious weather fluctuation processes [105] to [107].

d) SVM

Based on the Vapnik-Chervonenkis dimensional theory and the structural risk minimization principle of statistical learning theory, SVM has the excellent characteristics of robustness and generalizability, which can be applied to nonlinear systems, systems with small number of samples and systems with high-dimensional input parameters.

Based on the nonlinear sample data, the SVM-based WPF model can be mapped into a high-dimensional feature space by a certain nonlinear function, after which the regression estimation function can be created. Parameter selection is critical to the successful application of SVM-based WPF, which should be optimized. Commonly adopted optimization methods include the genetic algorithm, the particle swarm optimization, the ant colony algorithm, and the drosophila algorithm [108] to [110].

e) KNN

The KNN algorithm has been widely applied to various fields, including data classification, data forecasting and modelling based on historical data. KNN has a fast training speed and high accuracy, and it is insensitive to distorted data. However, the accuracy of KNN depends on the amount of available historical data. The forecasting time of the algorithm increases linearly with the number of factors affecting the forecasting model [111].

f) FLA

The basic principle of FLA is fuzzy mathematics theory (i.e., expressing experience and fuzzy information in the form of rules). In the fuzzy forecasting model, limited fuzzy rules are applied to describe the functional relationship between its input and output parameters.

Based on the fuzzy forecasting model, fuzzy factors (e.g., the influence of weather conditions) can be controlled. But the fuzzy forecasting model requires a larger amount of historical data than other statistical models do, and it has no learning ability, which is usually combined with the neural network for WPF. Neural networks are based on learning and connection structures, and they show ideal performance in low-level calculations using raw data. Conversely, FLA is based on human-like reasoning, and it shows ideal performance in advanced computing. Therefore, combining the two methods provides an effective new method for WPF [112].

7.2.3.3 Combination approaches

Different models have different advantages for WPF. The physical methods do not require a large amount of historical wind power data and other data, so they are applicable to newly built wind farms. The statistical methods have high accuracy but require a lot of historical data to train their forecasting models [116]. Combining the two types of methods will contribute to the improvement of the forecasting results.

The use of combination approaches is standard today. Essentially, all the approaches except for the minute scale use a combination of physical and statistical approaches. Nearly all models use both physical modelling and numerical modelling. Especially for wind farms with complex terrain, physical methods can be applied to obtain local meteorological information through high-resolution calculations and detailed physical descriptions, while statistical methods can be applied to acquire the pattern of the wind speed and output power of each wind turbine [117]. Combining physical methods with statistical methods considers the advantages of the wind resources at the location of each wind turbine and the changes in the wind power curve over time and the environment, which is helpful for improving WPF accuracy.

7.3 Ultra-short-term WPF

The time resolution of ultra-short-term WPF is usually not less than 15 min. Generally, the ultra-short-term WPF model is used to forecast the wind power of the next 15 min to 4 h, which is mainly appropriate for formulating intra-day power generation plans. Figure 24 shows the input and output parameters of the 4 h ultra-short-term WPF.

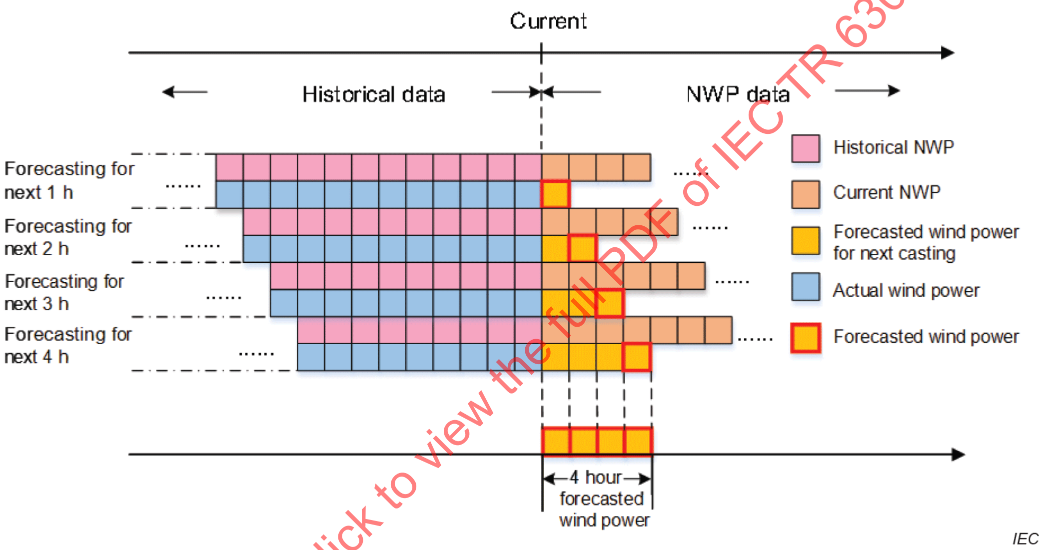


Figure 24 – Input and output parameters of the 4 h ultra-short-term WPF

Ultra-short-term WPF is mainly based on historical wind power data recorded by the SCADA system of wind farms, meteorological data from wind towers and NWP data. Wind speed and air density are the two meteorological impact factors of wind power generation. However, in most cases, the wind power change caused by air density is much smaller than wind speed is. Accordingly, wind speed is the major meteorological impact factor for the ultra-short-term WPF. The flowchart of ultra-short-term WPF is shown in Figure 25. The input parameters of the forecasting are the historical wind power data and NWP data. The output of the forecasting is the forecasted wind power, based on which the error and error band can be calculated.

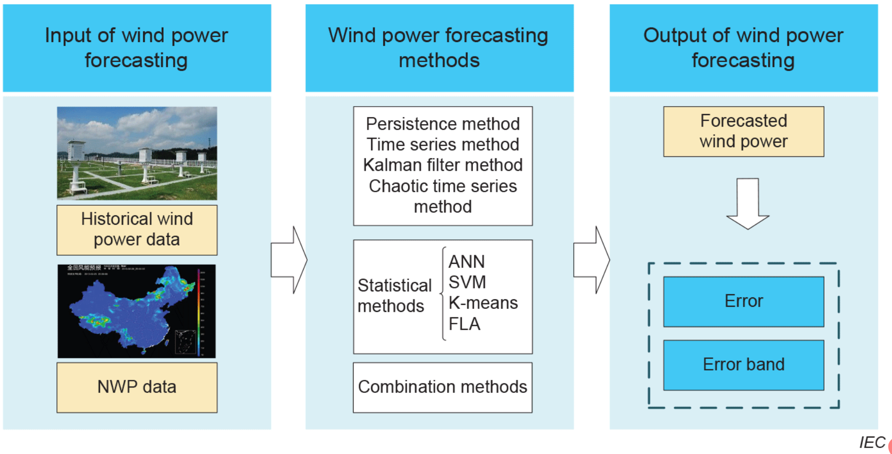


Figure 25 – Flowchart of ultra-short-term WPF

Statistical methods of ultra-short-term WPF are similar to those of short-term WPF. The difference is that the output power of ultra-short-term WPF is from 15 min to 4 h. The time series of wind power have chaotic characteristics, and the future trend in a short time can be forecasted according to the historical data, so the input of statistical methods applied to ultra-short-term WPF can comprise only historical wind power data. For WPF of more than 4 h, NWP data is also recommended as an input parameter of the forecasting methods. Nearly all methods of short-term WPF addressed in 8.2.4 could be applied in ultra-short-term WPF, such as the time series method and Kalman filter method; some other predication methods are also addressed as follows:

a) Persistence method

The persistence method is a simple and straightforward WPF method, based on which the actual wind power output value before the forecasting time—or the average value of the actual wind power output of several time points before the forecasting time—comprises the WPF output. The persistence method can only be applied to WPF in a very short period and has poor WPF accuracy, which is usually considered to be the benchmark method of evaluating other WPF techniques.

b) Chaotic time series method

The time series of wind power indicate typical chaotic characteristics. According to the theory of nonlinear dynamics, chaotic time series can be predicted in a short period, which means that chaotic theory can be applied to ultra-short-term WPF—the parameters of which can be optimized by phase space reconstruction. The development and change of the system can be reflected by reconstructing time series in high-dimensional phase space, and the evolution trajectory of the phase points is exactly the same as the topological structure of attractors. Therefore, ultra-short-term WPF can be transformed into the estimation of future wind power from the evolution process of reconstructed phase points [118].

c) Combination methods

Combination forecasting methods are based on maximum information utilization. They optimally combine the information contained in a variety of single models and consider the advantages of different models, which can significantly improve the accuracy of forecasting. Many factors affect the wind power of large-scale wind farms, which a single WPF model cannot easily describe [119]. Especially for extreme weather conditions, a single WPF model cannot be fully trained, which may lead to remarkable forecasting errors. Combination forecasting methods integrate different kinds of single models and include different factors that may play a role in the future.

The weights of at least two independent single WPF models directly determine the quality of combination WPF models. According to the different combinations of weight coefficients, combination forecast methods can be divided into fixed weight coefficient combination methods and variable weight coefficient combination methods. Fixed weight coefficient combination methods construct an objective function and determine the weight coefficients by minimizing the objective function under the constraints, such as minimizing the sum of the squares or the percentage errors of forecasting errors. Because the weight coefficients are fixed with time in the whole forecasting process, fixed weight coefficient combination methods can also be called linear combination methods. The weight coefficients of the variable weight coefficient combination methods change with time, reflecting a model with the highest forecasting accuracy at every moment. Therefore, variable weight coefficient combination methods are effective at improving the accuracy of the forecasting model and enhancing forecasting reliability. Because the weight coefficients of this combination model are difficult to determine, they are usually obtained using the nonlinear combination function, so variable weight coefficient combination methods can also be called nonlinear combination methods.

The combined WPF methods discussed above require at least two single WPF models. In addition, there are also some generalized combination forecasting methods for WPF—which are essentially single models—but the input, output or parameters of the WPF model will be optimized and adjusted, the principle of which is shown in Figure 26.

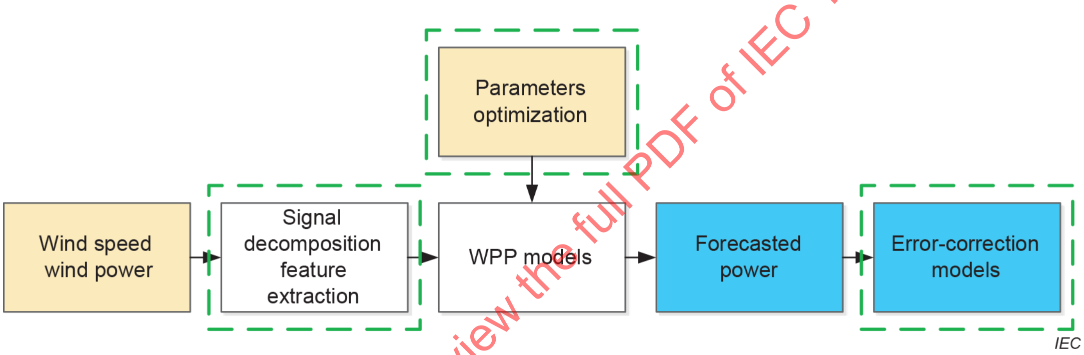


Figure 26 – Generalized combination methods of ultra-short-term WPF

In Figure 26, the generalized combination methods of ultra-short-term WPF can be achieved with the following four aspects:

1) The original sequence is decomposed into sub-sequences by wavelet decomposition or EMD, and each subsequence is modelled separately.
2) Data mining methods (e.g., rough set and PCA) are used to select features for the input of the WPF model.
3) Genetic algorithm, particle swarm optimization, fish swarm optimization and other optimization methods are used to optimize the parameters of the WPF model.
4) An error correction model is established for the output of the forecasting model. Error correction is a general method to improve forecasting accuracy, which is not limited to the specific forecasting process.

7.4 Probabilistic WPF

7.4.1 General

In past decades, a lot of research work has focused on deterministic WPF [120]. Nevertheless, considering the uncertainty, variability and volatility of wind and solar resources, it is not appropriate to ignore these terms in the forecasting process. Moreover, the detailed and accurate descriptions of the impacts of various weather-related factors that lead to forecast errors cannot be fully captured in a deterministic model. Consequently, the deterministic forecasting results inevitably contain uncertainties. Furthermore, the power grid planning, operation as well as safety and stability analysis require relatively accurate estimations of the fluctuation ranges of wind power, and only providing the single-valued expectations is not enough [121]. Therefore, it is necessary to develop wind power probabilistic forecasting to reveal the probability distribution of the forecasting error and to be able to quantify the inherent uncertainty of the (deterministic) forecast result. The wind power probabilistic forecasting can quantitatively describe the uncertainty in the process of forecasting and provide more comprehensive information for the decision-making in power system operation and regulation.

7.4.2 Basic concepts and model framework definition

Wind power probabilistic forecasting is a kind of WPF method that focuses on the uncertainty of wind power output. Based on the historical measurements, forecasts of meteorological variables and output power, the wind power probabilistic forecasting model can provide the fluctuation ranges or distribution functions of the wind power at different forecast time horizons.

The fundamental difference between the production of deterministic and probabilistic forecasting is the modelling of forecast uncertainty. The methods used for probabilistic forecasting are shown in Figure 27. According to whether the model is based on the deterministic forecasting results, the probabilistic forecasting approaches can be divided into four main types. Statistical approaches and physical approaches are based on so-called ensemble weather information. The latter, when converted into power output, can directly obtain the probability distribution of wind power. The former, where the error of wind power deterministic forecasting is the basis of the method, can obtain the forecasting intervals or distribution functions of the deterministic forecasting errors based on the deterministic forecasting results.

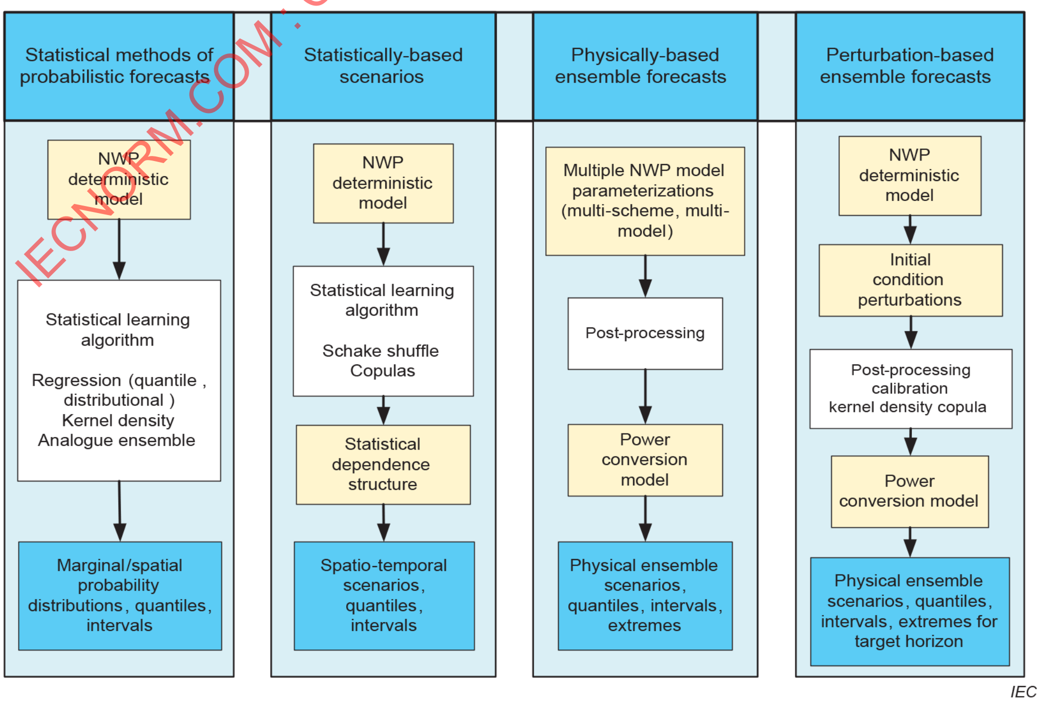


Figure 27 – Methods used for probabilistic forecasting

7.4.3 Uncertainty modeling approaches

7.4.3.1 General

The modelling of wind power probabilistic forecasting focuses on the uncertainty of wind power output. According to whether the model is based on the deterministic forecasting results, the modelling approaches can be divided into two types: wind power output and error of wind power deterministic forecasting. The overview of probabilistic wind power forecasting is shown in Figure 28. The former approach can directly obtain the probability distribution of wind power. The latter approach obtains the confidence intervals or distribution functions of the deterministic forecasting errors, and then attaches them to the deterministic forecasting results to form the probabilistic results.

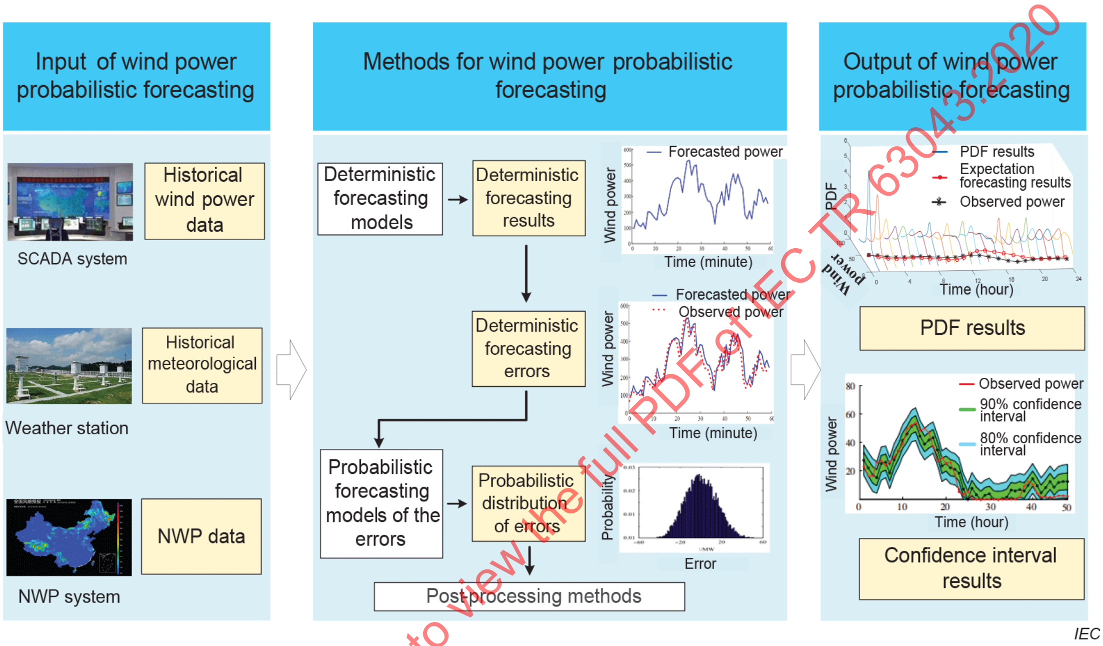


Figure 28 – Overview of probabilistic wind power forecasting

7.4.3.2 Inputs of the model

The input information of the wind power probabilistic forecasting model [122] mainly includes: a) historical measurements of wind power output; b) historical measurements of explanatory variables, for example relevant meteorological variables; c) forecasts of explanatory variables, for example NWP; d) wind power deterministic forecasting results.

7.4.3.3 Outputs of the model

The core output from a probabilistic forecasting model is a probability distribution, which can be expressed in many forms, for example, CDF (cumulative distribution function) or PDF (probability distribution function), confidence interval (often called forecasting interval) and quantile and POEs [123], [124]. The 24 h wind power probability distribution forecasting results are shown in Figure 29. As can be seen from Figure 29, in addition to providing the expectations of wind power output, the probabilistic forecasting model can also estimate the probability distributions of the error of the forecast expectation. The characteristics of the three kinds of output types are listed in Table 4.

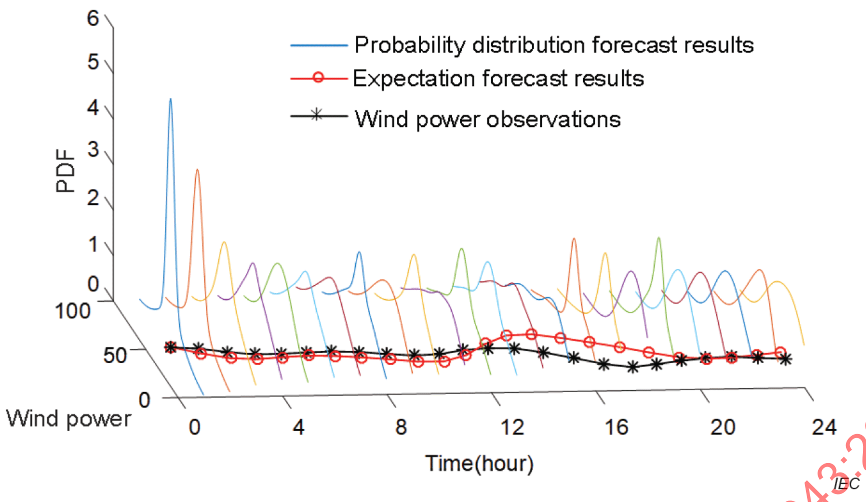


Figure 29 – Wind power probability distribution forecasting results

Table 4 – Output modes of probabilistic forecasting

Output type	Characteristics
PDF or CDF	Continuous distribution; intuitive and comprehensive; can be further converted to other output modes
Confidence (prediction) interval	Provide the fluctuation range of wind power at a certain confidence level; the most visualized mode
Quantile/ percentiles	The unilateral form of confidence interval forecasting result; discrete distribution

7.4.4 Probabilistic WPF model

7.4.4.1 General

Probabilistic WPF has several modelling methods. According to whether the modelling object is assumed to obey a predetermined distribution type, the probabilistic WPF models can be divided into parametric models and nonparametric models. According to the types of input NWP information, the wind power probabilistic forecasting models can be divided into deterministic NWP-based models and ensemble NWP-based models. According to the adopted forecasting theory, the probabilistic WPF models can be roughly divided into three categories: physical models, statistical models, and machine learning models.

7.4.4.2 Parametric modelling and nonparametric modelling

Parametric modelling uses a predetermined distribution type (such as Gaussian distribution [125], exponential distribution [126] or their segmented combinations [127]) to describe PDF. This method transforms the wind power probabilistic forecasting problem into a parameter estimation problem with low computational complexity. However, the predetermined distribution type is mostly based on the simplified empirical assumptions. Therefore, the inherent modelling errors in parametric modelling may affect the accuracy of the forecast probability distribution.

In contrast, without any assumptions of distribution type, the nonparametric modelling is distribution-free. It directly calculates the quantile or distribution function by means of data analysis methods such as QR, KDE, etc., which effectively avoids the error introduced by inappropriate modelling hypotheses. Since there are more parameters to be estimated, nonparametric modelling also has the drawbacks of complex calculation and requires a large amount of observation samples.

7.4.4.3 Modelling based on deterministic NWP and ensemble NWP

Wind power generation is significantly affected by meteorological factors, so the NWP information is usually applied in WPF on an hourly or daily scale. The input NWP information of wind power probabilistic forecasting includes two modes: deterministic NWP and ensemble NWP.

During the analysis of atmospheric motion, a deterministic NWP model solves a series of hydrodynamic and thermodynamic equations describing the weather evolution under certain boundary conditions and gives a single valued forecasting result for a future weather situation.

In contrast, the ensemble NWP consists of a series of separate deterministic NWPs that are derived from different initial conditions, some kind of statistical or physical perturbations, or different parameterization schemes. Compared with a deterministic NWP model, ensemble NWP members can provide more abundant weather prediction information (see details in 5.4.2_Ensemble_Prediction_Systems). When adopting ensemble NWP, the modelling approaches can be roughly divided into the three categories according to the specific processing methods for the ensemble NWP, which will be explained hereafter.

a) The filtering approach

The filtering approach with ensemble NWP as input is shown in Figure 30. As can be seen from Figure 30, in the first step, each member of the ensemble NWP is converted to power deterministic forecasting, which constitutes a set of power deterministic forecasting. Then, the post-processing methods are used to organize the uncalibrated power ensembles into a probabilistic forecasting result. During the conversion, the model is selected according to the different characteristics of each deterministic NWP member. Post-processing methods include analyzing the deterministic forecasting error to generate the probability distribution of wind power, or combining multiple deterministic forecasting results to form probability distribution, etc. Ensemble forecasting techniques can improve the accuracy compared to a single forecast. Ensemble forecasting techniques typically use static or dynamic weights or create statistical adaptation functions [120], [122] to [124], [128] to [130].

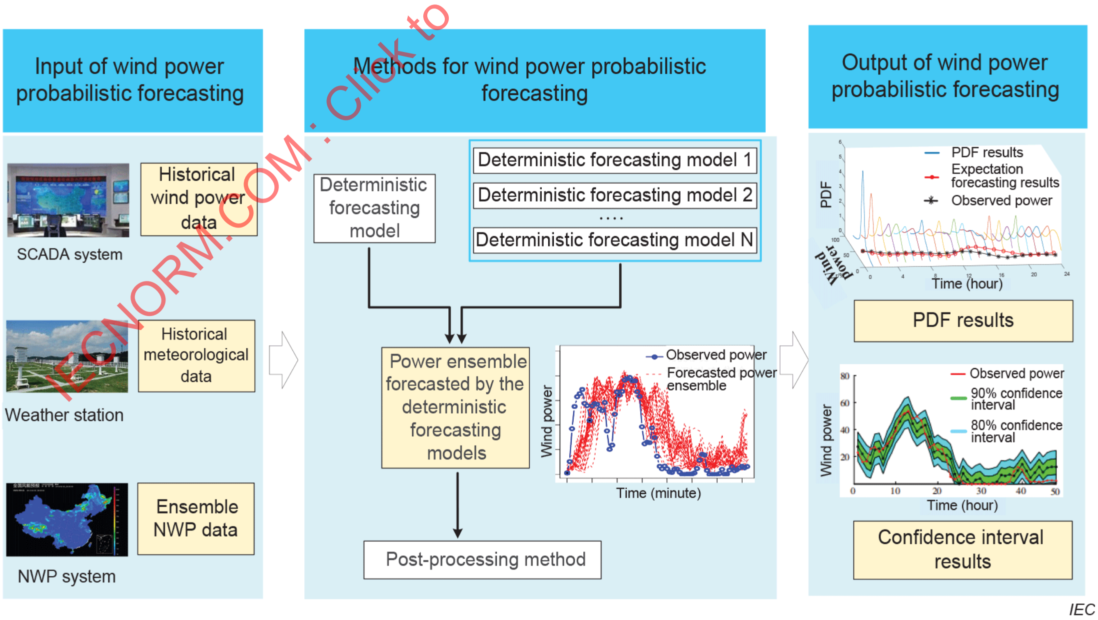


Figure 30 – Filtering approach with ensemble NWP as input

b) The dimension reduction approach

The dimension reduction approach with ensemble NWP as input is shown in Figure 31. As can be seen from Figure 31, in a first step, the dimension reduction approach extracts the central tendency information (e.g. the mean) and the spread information (e.g. the variance) of the ensemble NWP, reducing the input dimensionality. Then, these reduced inputs will be fed to a probabilistic forecasting model. This approach can significantly reduce the calculation costs, but the dimension reduction process should be carried out with caution to avoid deteriorating the accuracy of the forecast probability distribution.

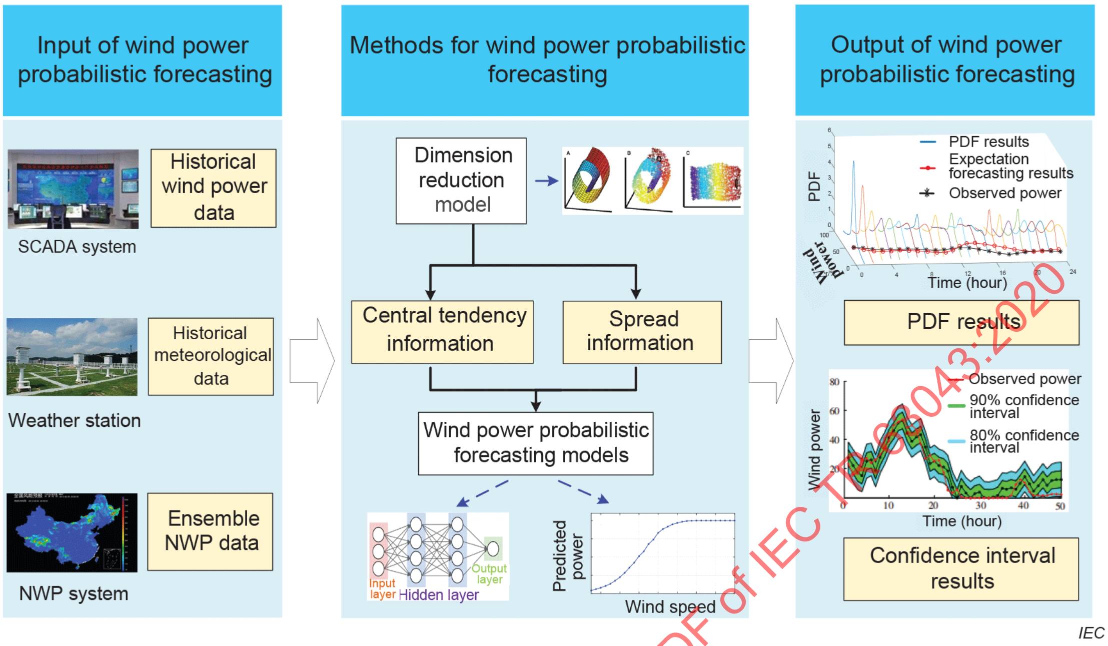


Figure 31 – Dimension reduction approach with ensemble NWP as input

c) The direct approach

The direct approach with ensemble NWP as input is shown in Figure 32. As can be seen from Figure 32, the direct approach inputs all members of the ensemble NWP into the probabilistic forecasting model. This approach has a higher dimension of the input information, a drastic increase in sample complexity and requires greater computational efficiency and more capacity [131] to [135].

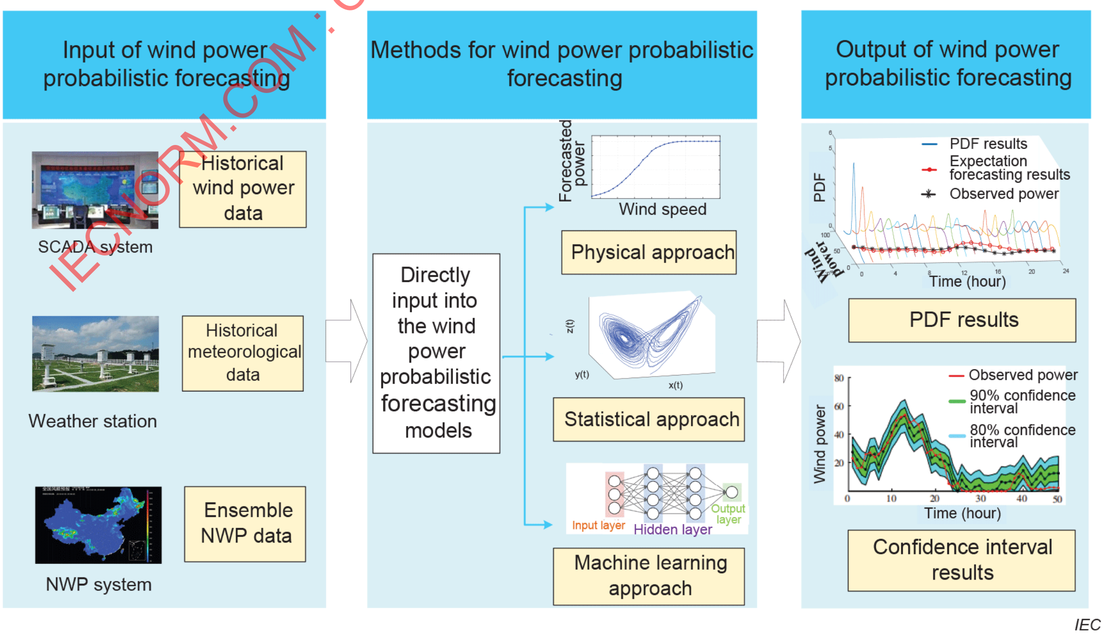


Figure 32 – Direct approach with ensemble NWP as input

7.4.4.4 Physical, statistical and machine learning models

a) The physical probabilistic forecasting model

Physical models generally adopt the filtering approach with ensemble NWP as input. They take topography, roughness, elevation, and other factors into consideration to establish the aerodynamic model to describe the flow field and put every NWP member as a boundary condition to calculate the wind speed and direction at the height of the wind turbine hub. Then, according to the wind turbine power curve, the input NWP ensembles can be converted into power ensembles [136] to [139]. After that, the wind power probability distribution can be obtained through a post-processing process.

b) The statistical probabilistic forecasting model

Statistical models are based on statistical algorithms. Under the premise of obtaining NWP information and other required data, the wind power probability distribution at the forecast target time can be inferred by mining the dependence relationship between the modelling object and the explanatory variables from historical samples. This method determines the structure and parameters of the model by analyzing historical samples, and hardly includes the simulation and consideration of the physical process.

KDE is a typical non-parametric statistical forecasting model [140]. Without any assumption of distribution type, the KDE studies the distribution characteristics of the modelling object from the observed samples. Besides, KDE is a typical representative of the forecasting technique requiring large samples, so it is not applicable to new wind farms. QR trains the regression relationship between explanatory variables and every quantile independently. It constitutes the regression model between the modelling object and explanatory variables at different probabilistic levels. As long as the probabilistic levels are set properly, the calculated quantiles can perfectly describe the probability distribution of the power fluctuation interval, and thus effectively grasp the uncertainty of wind power output [141].

In addition, other parametric/non-parametric statistical models, such as generalized error distribution and multivariate random variable analysis, are also widely used in wind power probabilistic forecasting, where the modelling object is chosen as the deterministic forecasting error. These statistical models have the advantages of fitting the parameters flexibly to reflect the ever-changing spatial-temporal correlations; in this way satisfactory results can be obtained, if the uncertainty has a correlation with past events.

c) The machine learning probabilistic forecasting model

The machine learning models employ artificial intelligence techniques to mine the highly complex nonlinear relationship from input to output. They have a stronger generalization ability and can also be used in wind power probabilistic forecasting to quantify the uncertainty.

Extreme learning machine is a kind of excellent-performance feed-forward neural network methods with a single layer or multiple layers of hidden nodes. Compared with the traditional neural network methods, Extreme learning machine greatly improves the generalization ability, nonlinear fitting ability as well as the learning speed. The practices in probabilistic forecasting reflect its application prospect of online dynamic updating [142]. Sparse Bayesian theory has also been adopted to wind power probabilistic forecasting. It considers the model parameters as random variables. After considering the prior distributions of these parameters, the method can effectively improve the accuracy of the forecasting results. The predefinition of the Gaussian distribution on the parameters endows this kind of model with excellent sparse characteristics.

In addition, machine learning algorithms such as wavelet neural network and evidence theory have also been successfully applied to wind power probabilistic forecasting.

7.5 Wind power ramp event forecasting

7.5.1 General

The significant changes of wind power in a short period are often referred to as wind power ramp events. It is shown that extreme weather, such as strong convection, low level jet stream and thunderstorms, can easily cause a sudden increase of wind speed and trigger ramp up events. Conversely, the precipitous drop in pressure gradient or the wind turbine protection shutdown caused by the wind speed that is continuously higher than the cut-off wind speed, will lead to a sudden decrease in power output and trigger ramp down events. It is worth noting that many ramp events occur with the subtle variations in wind speed and direction over an area as well, which is not a classical extreme weather event.

The wind power integrated into a power grid has the characteristics of large-scale and high concentration. Therefore, wind power ramp events will lead to unbalance between power generation and consumption, which is liable to cause frequency fluctuation and power quality deterioration, threatening the safe and stable operation of the power grid. Furthermore, severe wind power ramp events can even result in accidents such as load shedding and blackout over large areas, which may cause significant economic losses. Therefore, the quantitative analysis and accurate early warning of wind power ramp events can assist to optimize the unit combination, allocate the reserve rationally, mitigate the impact on the power grid, as well as enhance the safety and stability of the system operation [143].

7.5.2 Quantitative description of wind power ramp events

Wind power ramp events are usually described by five characteristic quantities:

1) Ramp magnitude ΔP_r: the variation of wind power from the start to the end of a ramp event.
2) Ramp direction: (the sign is either positive, for an upward ramp, or negative, for a downward ramp) the power generation at the end of the ramp period minus the generation at the beginning.
3) Ramp duration Δt: period of time from the start to the end of the ramp event.
4) Ramp rate $\Delta P_r/\Delta t$: the ramp magnitude divided by the ramp duration.
5) Ramp timing t_0, which can be defined either as the starting moment or the middle moment of the ramp.

Two ramp events of a wind farm located in Spain within 5 days are shown in Figure 33, in which the horizontal axis is the time and the vertical axis is the wind power observation normalized by the installed capacity of the target wind farm, i.e., P_r. The aforementioned five characteristic quantities describe these two ramp events as follows. The wind farm experienced a 4 h wind power ramp up event at 20:00 on January 24, during which the wind power output increased by 91 % P_r, and the ramp rate was 22,75 % P_r/h. On January 27, the target wind farm experienced a 3 h ramp down event since 0:00, during which the wind power output decreased by 74 % P_r, and the ramp rate is –24,7 % P_r/h.

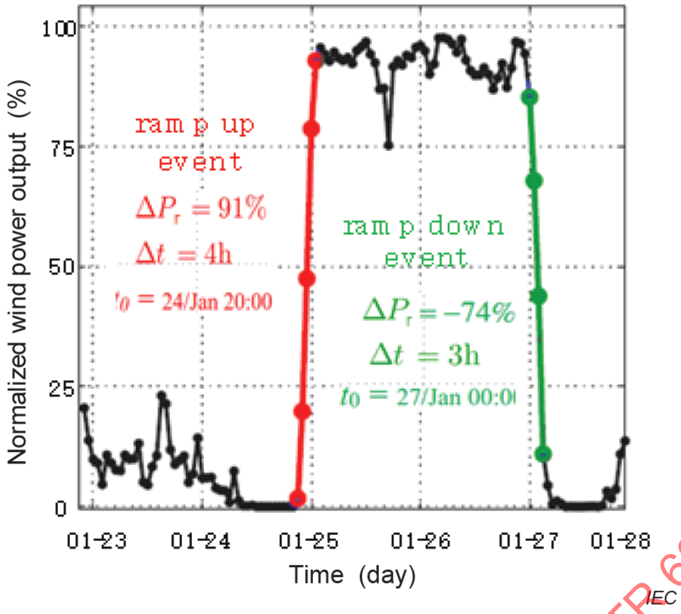


Figure 33 – Two ramp events of a wind farm

Based on the characteristic quantities of the ramp, the definitions of wind power ramp events are usually expressed in the following four ways [144]:

Definition 1: If the absolute variation of power observed at the starting and ending moments of period [t, $t+\Delta t$] is greater than the predetermined threshold, it can be judged that the wind power ramp event occurs during this observation period. The criterion is calculated by Formula (4).

$$\left|P(t+\Delta t)-P(t)\right|>P_{\varepsilon} \tag{4}$$

where

$P(t)$ represents the observed power at time t (unit: MW),

$P(t+\Delta t)$ represents the observed power at time $t+\Delta t$ (unit: MW) and

P_{ε} is the predetermined threshold of power variation (unit: MW).

Definition 2: If the difference between the maximum and the minimum of wind power observations during the period [t, $t+\Delta t$] is greater than the predetermined threshold, then a wind power ramp event is calculated by Formula (5).

$$\max\left(P[t,t+\Delta t]\right)-\min\left(P[t,t+\Delta t]\right)>P_{\varepsilon} \tag{5}$$

where

max (P [t, $t+\Delta t$]) represents the maximum observation (unit: MW), and

min (P [t, $t+\Delta t$]) is the minimum observation (unit: MW).

Definition 3: If the ratio of the absolute variation, i.e., |$P(t+\Delta t)$- $P(t)$|, and the length of the observation period, i.e., Δt, exceed the threshold, then a wind power ramp event is calculated by Formula (6).

$$\frac{\left|P(t+\Delta t)-P(t)\right|}{\Delta t}>P_{\varepsilon} \tag{6}$$

The above three kinds of wind power ramp events definitions are directly based on the observed wind power series and identify the occurrence of the ramp event by comparing the power variation in the observation period with the predetermined thresholds. In addition, an alternative criterion that firstly carries out the low-pass filter process for wind power series using the idea of difference can be adopted. This improved criterion avoids the disturbance of the second-level fast fluctuation of wind power or the inherent noise in the observation data. Using P_t to represent the observed wind power series, the filtered signal can be calculated by Formula (7).

$$P_t^f = \text{mean}\left(P_{t+h} - P_{t+h-n}\right), h = 1, 2, \cdots, n \quad (7)$$

where

h is the average difference estimator, i.e., the window width of the filtered signal,

n is the maximum value of the time window, and

mean(·) represents the average value within the target time window.

Definition 4: If the absolute value of the filtered power signal P_t^f exceeds the predetermined threshold, then a wind power ramp event is calculated by Formula (8).

$$\left|P_t^f\right| > P_\varepsilon \quad (8)$$

In the aforementioned four kinds of ramp event criteria, the threshold P_ε can be defined either as a specific megawatt or a percentage of the installed capacity. The length of the observation period generally ranges from 10 min to 4 h according to the application scenarios [145]. Current researches have not reached a consensus on the setting of P_ε and Δt. In most scenarios, when predetermining these values, the installed capacity of wind farm and the operating environment of the power grid should be taken into account.

There are significant differences in the four kinds of definitions of ramp events, and their advantages and disadvantages are shown in Table 5.

Table 5 – Advantages and disadvantages of ramp events definitions

Definition type	Advantages	Disadvantages
Definition 1	Simple identification; distinguishes the ramp direction	Ignores the power variation process; leads to inaccurate identification
Definition 2	Considers the power variation process; low failed reporting rate	Difficult to distinguish ramp direction
Definition 3	Reflects the ramp rate; distinguishes the ramp direction	Ignores the power variation process; leads to inaccurate identification
Definition 4	Removes the noise effect; low failed reporting rate	Difficult to distinguish ramp direction

7.5.3 Forecasting methods of wind power ramp events

7.5.3.1 Wind power ramp event deterministic forecasting

The modelling objects of wind power ramp event forecasting models can be divided into two types [146]. One is a binary variable indicating whether a ramp event occurs, i.e., the value of 1 represents the occurrence of a ramp event and the value of 0 indicates no ramp event occurs. The other is a continuous variable reflecting the characteristics most relevant to system ramp issues, such as ramp rate over a previous period. The forecast that takes the former as the modelling object is a binary classification problem, and algorithms such as SVM can be used to realize the early warning of the approaching ramp state. The forecast of the characteristic quantity of wind power ramp events is similar to the forecast of wind power. Current researches indicate that the statistical regression forecasting models and the intelligent learning forecasting models such as neural network [147] can also be applied to forecast the characteristic quantities of wind power ramp events. Wind power ramp event forecasting approaches can be roughly divided into two categories: indirect approach and direct approach.

The indirect approach is a mainstream approach of ramp event forecasting at present. This method forecasts the wind power time series by means of WPF technology, and then employs the predetermined definition to detect the occurrence of a ramp event. The corresponding characteristic quantities of ramp can be subsequently extracted based on the power forecasting results. However, the classical WPF technology often intentionally ignores the samples observed with extreme weather, so as to obtain a relatively smooth forecasting curve to minimize the overall forecast error. Therefore, the indirect approaches that rely on WPF results may underestimate the severity of ramp events, resulting in the omission of ramp warning.

The direct approach uses the historical sample data to extract the dependence relationship between the characteristic quantity of ramp and the regional meteorological information. Then, the mapping from the meteorological information to attributes is established without forecasting the wind power output in advance [148]. These kinds of approaches are more intuitive and have higher forecast accuracy, but the training of the forecast models depends on a large number of historical observation samples. Therefore, the completeness of samples and the accuracy of data will directly affect the performance of these approaches. In addition, a wind power ramp event is a typical low-probability event, and the inherent scarcity of its observation samples seriously restricts the application and development of these kinds of approaches.

7.5.3.2 Wind power ramp event probabilistic forecasting

The scarcity of ramp events from extreme weather conditions leads to the inevitable large errors in ramp event deterministic forecasting results. Under this background, probabilistic forecasting of ramp events, which can provide more detailed information for system operators, dispatchers and market participants, has therefore attracted widespread attention.

Statistical forecasting models such as multiple auto-regressive statistics and artificial intelligence forecasting models such as data-driven machine learning have been successfully applied to the probabilistic forecasting of the occurrence/non-occurrence of a ramp event [149], [150]. And the results obtained are more reliable than that of the deterministic forecasting. In terms of the probabilistic forecasting of characteristic quantities, the flexible application of ensemble NWP enables forecasting results to reliably quantify the uncertainty in the forecasts also in cases which have not been present in historic data and which can provide more comprehensive data support for the dispatching and control of the power grid in extreme events [140], [151] to [153].

7.5.3.3 Relationship of probabilistic ramp forecasting, extreme events and climate change

Extreme events in meteorology are often associated with return periods of 20 years or 50 years. The same is applicable for wind (and solar) power forecasts, but there are far more conditions leading to extreme events of wind power than extreme weather conditions, if we consider the power system. Today, it is not unrealistic to list 200 different weather conditions with 20-year return periods. That means approximately 10 events per year, where some kind of weather condition occurs, which causes an unusual pattern in the wind (and/or solar) power generation, which would not be expected to be seen more than once within 20 years. The definition of a 20-year event may also relate to climate change, because climate is changing on such horizons. A typical example of a meteorological extreme event can be at a weekend day with low demand and a strong temperature drop and wind speeds reaching more than 20 m/s.

Such extreme events for the power system can be, but are not necessarily defined as, meteorological extreme events. If we assume an event has a meteorological future return period of 20 years, a 20-year future period may be different than a 20-year past period, due to climate change and the fact that the next time the event is expected to occur, the wind power capacity may be significantly higher. Due to increasing penetration levels, the same meteorological event hence may be categorized as significantly more extreme from a power system perspective. The power system and the climate develop asynchronously, so the concept of a return period does not always help shed light on the issues that need to be managed from either a forecasting or power system operations perspective.

In the power system an extreme event of a long return period could be caused by a single small low pressure system, which is aligned at a particular time relative to the available generation capacity in such a way that a peak ramp rate occurs. The uncertainty factors can interact and result in sudden events with return periods of maybe 25 years or 50 years.

The relationship between ramp rates and hence ramp forecasting associated with extreme events can therefore not be neglected in areas where the penetration increases over approximately 30 % of the electricity demand. In Möhrlen et al. [158], this relationship is described in the example of the Irish grid operator. Here, ramp forecasts are used to dynamically allocate reserves. Due to the uncertainty associated with the ramp of the variable generation (wind and solar power), probabilistic forecasts are a necessity. In the study, the computation of the required reserve allocation is where the relationship between reserve and ramp, as well as the uncertainty of the forecasts, comes into play. Here, some of the reserves held will be a function of the uncertainty associated with the variable generation. If the non-scheduled variable generation is sufficiently high, a risk-based probabilistic forecasting is necessary in order to allocate reserves in advance. The optimal allocation of a reserve product can be determined from a cost function or from a profile of the uncertainty.

7.6 WPF for wind farm clusters

7.6.1 General

WPF for wind farm clusters is the forecasting of the overall output power, which consists of multiple wind farms. With the continuous increase of global wind power installed capacity, the distribution scale of wind farms has changed from a decentralized and small-scale distribution in the early stage of development to a clustered, large-scale distribution. Therefore, the integration of massive wind farm clusters to the power grid has a major impact on the safe operation and economic dispatch of the power system. WPF techniques are mostly concentrated on the power forecasting of a single wind farm. However, the power grid dispatching department is more concerned with the overall output of the power system, and it needs to effectively control the exchange power of the grid tie lines. Therefore, WPF for wind farm clusters is very important [154].

Subclauses 7.6.2 to 7.6.4 introduce the basic knowledge of WPF for wind farm clusters based on the following three aspects: the basic concepts of WPF for wind farm clusters, the overall framework of WPF for wind farm clusters and the physical levels of WPF for wind farm clusters.

7.6.2 Basic concepts of WPF for wind farm clusters

a) Wind farm clusters

Wind farm clusters represent wind farms with related factors such as geographical location or grid topology. However, there is no uniform standard for the number of wind farms in the cluster, and the number of wind farms is selected according to different partitioning requirements.

b) Sub-area division for wind farm clusters

Sub-area division refers to dividing a large-scale wind power cluster into different areas to obtain the wind power cluster WPF results according to the sub-area output [155].

c) Reference wind farm

Reference wind farms need to be designated for the WPF model of wind farm clusters based on the statistical upscaling method. Generally, wind farms with ideal data integrity and forecasting accuracy comprise reference wind farms [156].

d) Accumulation method

The accumulation method obtains the WPF results for wind farm clusters by directly accumulating the forecasting results of each wind farm output.

e) Statistical upscaling method

Statistical upscaling refers to the establishment of an upscaling model with sub-regional wind farm output as input and overall wind farm cluster output as output to obtain statistical methods for WPF results for wind farm clusters [157], such as the ensemble Kalman filter techniques [158].

f) Spatial resources matching method

The spatial resources matching method is the WPF method for wind farm clusters that utilizes spatial similarity to match the predicted wind speed with the historical wind speed to obtain a corresponding WPF result [157].

g) Input data for WPF of wind farm clusters

The input data for WPF of wind farm clusters can be divided into three categories: (1) NWP data; (2) historical meteorological data and historical wind power data; (3) real-time meteorological data from the anemometer towers and real-time wind power data from the SCADA system.

h) Output data of WPF for wind farm clusters

The output data of the WPF for wind farm clusters is the predicted results of the overall wind power output of wind farmers in a whole region. When the accumulative method is applied to the WPF of wind farm clusters, two phases of forecasting results can be obtained—the forecasting results of the single wind farm and the regional wind farms.

7.6.3 Overall framework of the WPF for wind farm clusters

Figure 34 shows the overall framework of the WPF for wind farm clusters, including data sources, sub-region divisions of wind farm clusters, WPF steps, and data flow.

The data source is the basic WPF for wind farm clusters. Data from the NWP system, data from the SCADA system and online data from the anemometer towers should be collected to generate the WPF database of wind farm clusters. The effective sub-region division of wind farm clusters is the premise of the forecasting, which can contribute to improving WPF efficiency. WPF for wind farm clusters contains five steps, including data preparation and pre-processing, sub-region division of wind farm clusters, WPF for individual wind farms, WPF for sub-regions and wind power summation of the sub-regions. Data flow of WPF for wind farm clusters contains power flow, information flow, meteorological flow, training process and forecasting process.

a) Data sources

The data sources for WPF include geographic data, NWP data, historical data and real-time data, the details of which are shown in Table 6. The geographic data includes the topography, layout and location information of the wind farms. The NWP data includes wind speed, wind direction, air pressure, temperature and humidity. The historical data includes historical meteorological data and historical wind power data. Finally, the real-time data includes real-time meteorological data from the anemometer towers and real-time wind power data from the SCADA system.

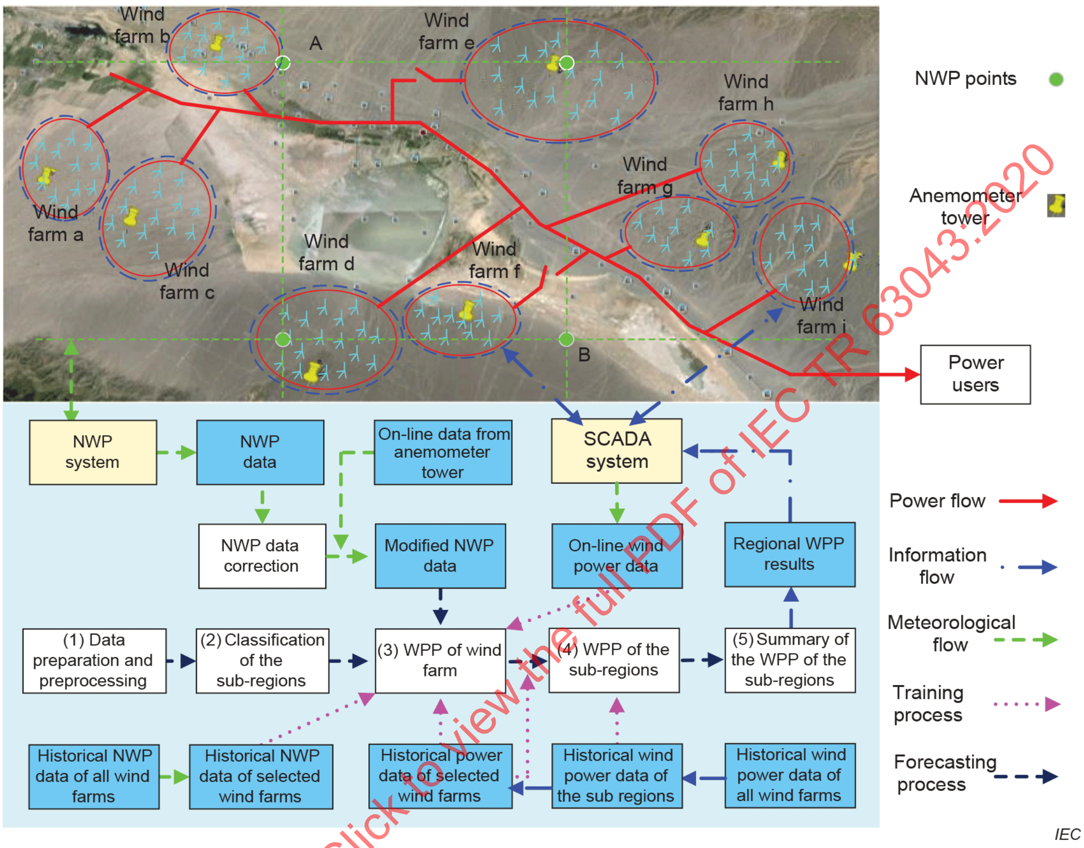


Figure 34 – Overall framework of the WPF system for wind farm clusters

Table 6 – Data sources of WPF for wind farm clusters

Data sources	Detailed information
Geographical data	Topography, layout and location information of the wind farms
NWP data	Wind speed, wind direction, air pressure, temperature and humidity
Historical data	Historical NWP data and historical wind power data
Real-time data	Real-time meteorological data of the anemometer towers and real-time wind power data from the SCADA system

b) Sub-region division of wind farm clusters

Sub-region division is an important step in WPF for wind farm clusters because it affects forecasting accuracy. By creating a reasonable sub-region division standard, the wind farm clusters are divided into different sub-areas, and their overall wind power output is obtained according to the summation of wind power output of the sub-regions. The sub-region division criteria can be established according to factors including geographical information, power grid topology, and wind resource correlation of wind farms.

c) WPF steps

WPF for wind farm clusters can be divided into five steps: data preparation and pre-processing, sub-region division of wind farm clusters, WPF for individual wind farms, WPF for sub-regions and summation of the forecasting results of sub-region wind farms.

d) Data flow

The data flow of the WPF for wind farm clusters is represented by different types of arrows in Figure 34, including power flow, information flow, meteorological flow, training process and forecasting process. Information flow represents the information exchange between the WPF and SCADA systems. Meteorological flow represents the information exchange between the WPF and NWP systems. The training process describes the training of the single wind farm forecasting model and regional wind farms forecasting model. The forecasting process is based on the five forecasting steps in Figure 34 and uses a single wind farm and the result of the sub-region wind farms to achieve the WPF outcomes of the wind farm clusters.

7.6.4 Physical hierarchy of WPF for wind farm clusters

The physical WPF hierarchy of the wind farm clusters can be divided into six layers, including the data layer, mapping layer, feature layer, model layer, feedback layer and output layer, as shown in Figure 35.

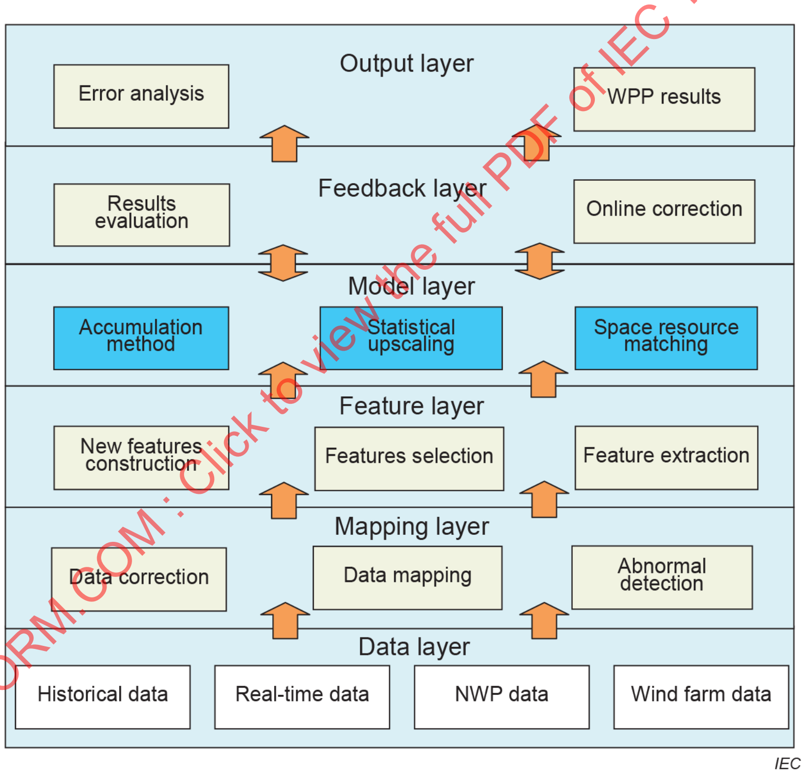


Figure 35 – Physical levels of WPF for wind farm clusters

The data layer is the data source of WPF for wind farm clusters, including the historical data, real-time data, NWP data and wind farm data (e.g., geographical data and distribution of wind farms). The mapping layer is the further processing of the original data, including abnormal data detection, data correction and data mapping. The feature layer contains three parts: feature extraction, new feature construction and feature selection. The model layer is the process for establishing a WPF model for wind farm clusters, and it contains the accumulation method, statistical upscaling method and spatial resource matching method, etc. The feedback layer is applied to evaluate and correct WPF outcomes of the wind farm clusters, including the results evaluation and online correction. The output layer contains the error analysis and the WPF result visualization.

7.6.5 WPF methods of wind farm clusters

7.6.5.1 General

WPF methods of wind farm clusters contain the accumulation method, statistical upscaling method, and space resource matching method.

7.6.5.2 Accumulation method

When the accumulation method is implemented, the output power of each wind farm is predicted separately, and the results are accumulatively merged to obtain the WPF results of wind farm clusters. The flowchart of the accumulation method is shown in Figure 36. The method is applicable to sparsely distributed and small-scale wind farm clusters, and it has high requirements for the entirety of the historical and NWP data of all the wind farms.

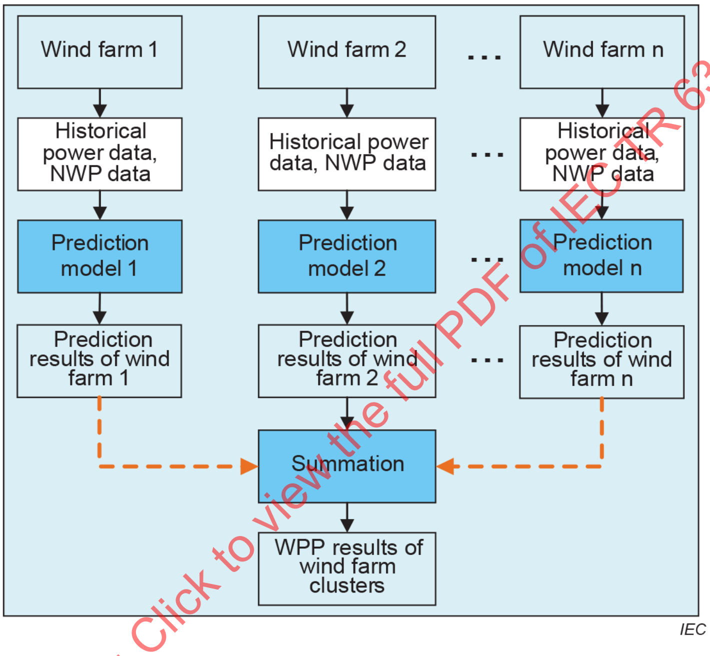


Figure 36 – Flow chart of the accumulation method

7.6.5.3 Statistical upscaling method

The statistical upscaling method is the most popular method of WPF for wind farm clusters, the flow chart of which is shown in Figure 37.

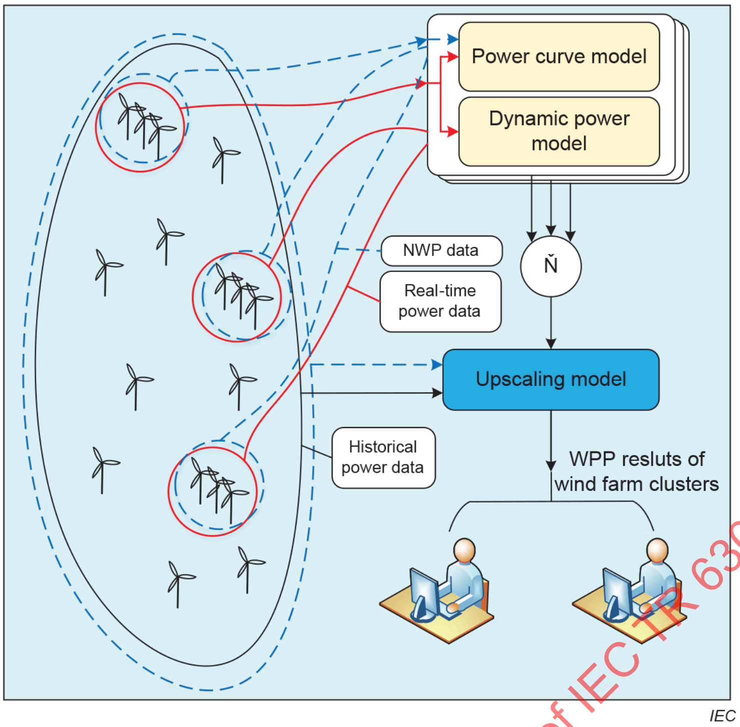


Figure 37 – Flow chart of the statistical upscaling method

First, the wind farms with ideal data integrity and forecasting accuracy are selected as the reference wind farms, and their wind power is forecasted. A statistical upscaling model is then established with the wind power output of the reference wind farms as the input and the overall wind power output of the wind farm clusters as the output to obtain the WPF results of the wind farm clusters. The statistical upscaling method has high forecasting accuracy and low requirements for the entirety of the historical and NWP data of all the wind farms. However, proper reference wind farms should be selected, which is highly important to the accuracy of the method [157].

7.6.5.4 Spatial resource matching method

The spatial resource matching method-based WPF for wind farm clusters requires fewer computing resources and has high forecasting accuracy, the flow chart of which is shown in Figure 38. First, the spatial resources in terms of wind speed matrices of the wind farm clusters are created, including the historical moment wind speed matrix (matrix A) and the wind speed matrix (matrix B) at the time of wind power to be predicted. Then the Euclidean distance between matrix A and matrix B is calculated. Further, a certain distance threshold is set to extract n sample sets, which are those with the highest spatial resource similarity to the time of wind power to be predicted. The weight coefficients of the n sample sets are calculated according to the Euclidean distance. Finally, the historical power of the n sample sets is weighted according to the weight coefficient and then summed to obtain the final WPF result.

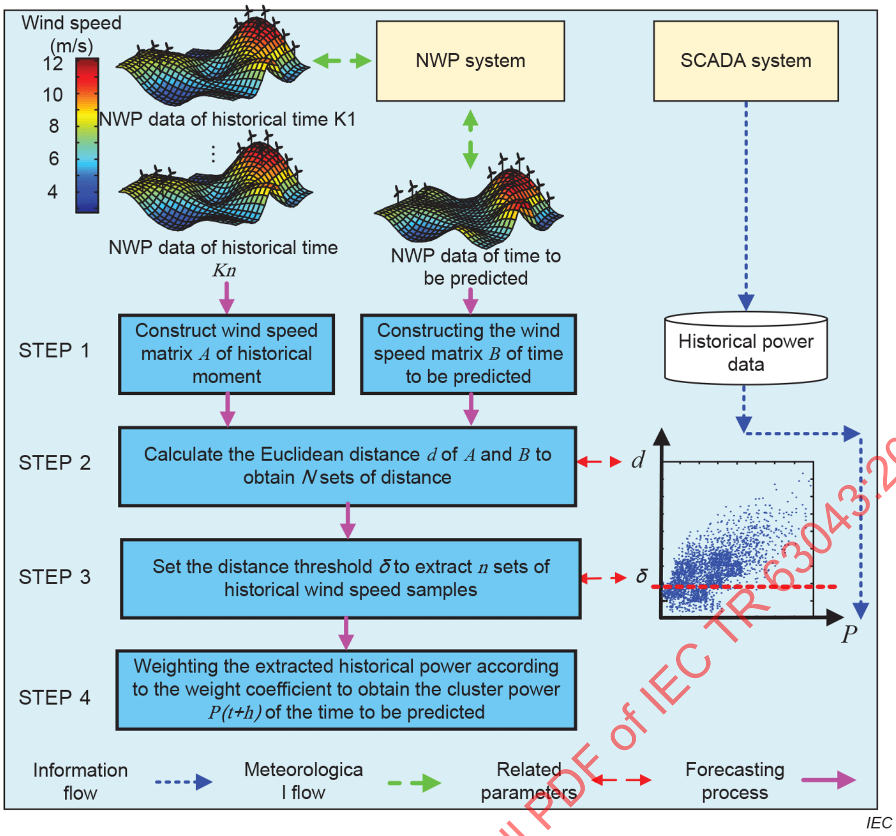


Figure 38 – Flow chart of the space resource matching method

7.6.5.5 Comparison of the WPF methods for wind farm clusters

Comparison of the WPF methods of wind farm clusters—the accumulation method, statistical upscaling method and space resource matching method—is shown in Table 7.

Table 7 – Comparison of WPF methods for wind farm clusters.

	Accumulation method	Statistical upscaling method	Spatial resource matching method
Principle of the methods	The power of each wind farm is separately predicted and then accumulatively merged to obtain the final results.	A statistical upscaling model is established with the wind power output of the reference wind farms and overall wind power output of the wind farm clusters.	Data mining is carried out to extract the samples with the highest spatial resource similarity to the time of wind power to be predicted.
Advantages of the methods	The model is easy to establish.	The method has high forecasting accuracy and low requirements for the entirety of the historical and NWP data of all the wind farms.	The method has high forecasting accuracy and requires very few computational resources.
Disadvantages of the methods	The method has high requirements for the entirety of the historical and NWP data of all the wind farms.	Selecting the reference wind farms is challenging.	Parameter optimization is challenging.
Scope of application	The method is suitable for sparsely distributed and small-scale wind farm clusters.	The method can be widely applied to WPF for wind farm clusters.	The method is suitable for large-scale wind farm clusters.

The accumulation method is easy to establish, but it has high requirements for the entirety of the historical and NWP data of all the wind farms and is suitable for sparsely distributed and small-scale wind farm clusters. The statistical upscaling method has high forecasting accuracy and low requirements for the entirety of the historical data and NWP data of all the wind farms but selecting the reference wind farms is challenging. The space resource matching method has high forecasting accuracy and requires very few computational resources, making it suitable for large-scale wind farm clusters.

7.7 Other WPF techniques

7.7.1 Medium-term and long-term WPF

Medium-term WPF is generally considered to be WPF for more than three days. Different countries have different views on the time boundaries of medium-term WPF. Argonne National Laboratory stipulates that it is limited to seven days, which is widely recognized. Medium-term WPF is mainly applied to adjusting the energy storage strategy of the wind farm, optimizing the standby power generation capacity and optimizing the maintenance scheduling of the wind turbines, etc.

The methods of the medium-term and short-term WPF are similar. However, due to the long forecasting duration, the accuracy of medium-term WPF is lower than that of short-term WPF. Two common techniques are applied to improve its accuracy: the hybrid NWP and hybrid WPF models. Various NWP systems such as the GFS, WRF, HIRLAM have been developed, but one NWP model usually cannot provide sufficient downscaling for a particular wind park in a particular location [159]. In a hybrid NWP model, meteorological data from more than one NWP system is assimilated, which can effectively improve the accuracy of NWP and WPF—especially for medium-term power forecasting. The principle of hybrid WPF models is similar to that of hybrid NWP models. Hybrid WPF models incorporate the individually superior features of various forecasting models, based on which an advanced forecasting method with high forecast accuracy and a wide forecast horizon can be obtained. Usually, hybrid WPF models employ a linear model to predict the linear component and a nonlinear model to predict the nonlinear component in wind power time series. In recent years, the hybrid models have drawn more and more attention.

Long-term WPF is generally considered to be the WPF for several weeks, several months, or even more than one year. There is no authoritative standard of the time boundaries of long-term WPF. Long-term forecasting is mainly applied to arranging the maintenance schedule of power equipment, evaluating wind power resources, and planning wind farm construction. Long-term WPF will contribute to planning the connection or disconnection of wind turbines or conventional generators [160], thus achieving low standby power generation capacity and optimal operating costs for the power system.

Long-term WPF methods include trend extrapolation methods, time series methods and regression methods. Compared with short-term and medium-term WPF, long-term WPF depends more on the autocorrelation and periodicity of wind power output. The key to long-term WPF is to collect a large amount of historical data and conduct a large number of experimental studies based on that data to summarize experiences.

7.7.2 WPF for offshore wind farms

Compared with onshore wind farms, offshore wind farms have special characteristics. First, the sea surface wind speed is more stable than it is on land, and the autocorrelation function attenuation coefficient of offshore wind speed is less than that of onshore wind speed, so the volatility of offshore wind power output is lower than that of onshore wind power [161].Second, the environment of offshore wind turbines is very harsh, so their failure rate is much greater than that of onshore wind turbines [162].

Due to the characteristics of offshore wind farms, their WPF methods are different from those of onshore wind farms.

Because of the lack of meteorological observation data, there is less input data for the offshore WPF model than that for the onshore WPF model. Therefore, more NWP data from different forecasting systems should be assimilated to improve the WPF accuracy of offshore wind farms [163].

Due to the harsh environment of offshore wind farms and high failure rate of offshore wind turbines, the real-time installed capacity at the time of forecasting may be different from the installed capacity at the historical time [164]. Therefore, a more complete forecasting model should be established from the aspects of data selection and correction factor construction of the wind turbine failure rate.

7.8 Summary

Clause 7 introduced WPF technologies based on the principles of time scale, forecasting objects, and application scenarios, including short-term forecasting, ultra-short-term forecasting, probabilistic forecasting, ramp event forecasting, and cluster forecasting. It emphatically expounded on the basic concepts and technical routes of various technologies and provided common forecasting methods. In addition, it analysed the advantages, disadvantages, and applicable conditions of various methods.

8 PV power forecasting technology

8.1 General

There is a certain similarity between PVPF and WPF in terms of technology and how the general principles described in the introduction and the clauses on NWP and statistical modelling are applied. However, due to the different influence factors of PV and wind power, there are also multiple differences between them, which explains the need for Clause 8 and Clause 7, both of which delve deeper into specific issues around wind (Clause 8) and PV (Clause 9). The regularity of the Earth's periodic revolution and rotation leads to solar irradiance variation. Theoretically, the extra-terrestrial irradiance outside the atmosphere anytime and anywhere on the Earth can be accurately calculated. The main factors of changes in solar radiation are variable aerosols and clouds. This increases the uncertainty of PVPF [165]-[167].

Clause 8 provides an overview of the main types of PVPF, namely, deterministic, probabilistic and distributed PVPF methods for forecast horizons ranging from minutes to days ahead. Note that Clause 8 adopts a similar nomenclature to WPF for consistency within the document, with callouts provided, where appropriate, to highlight terminology more commonly used in the PVPF literature, for example, intra-hour, intra-day and day ahead rather than minute-time-scale, ultra-short-term and short-term.

8.2 Short-term PVPF

8.2.1 General

Short-term forecasting of PV power generation needs to provide power output data of PV station in the next 1 day to 5 days. In the solar forecasting literature, this is commonly referred to as day ahead forecasting, but there is variation in the exact definition. Regardless, in general, data processing, analysis, and forecasting need to be performed at least daily, and the daily PV power curve will exhibit different fluctuation patterns due to the weather conditions around the PV station. Therefore, improving the accuracy of weather forecasts can directly improve the PVPF accuracy [168], [169].

8.2.2 Meteorological influence factors of PV power generation

8.2.2.1 General

The PV power generation is mainly determined by the local weather environment and can be influenced by multiple meteorological influence factors, including irradiance, PV module temperature, ambient temperature, wind speed, etc. The relationship between meteorological influence factors and PV power output is introduced below [170], [171].

8.2.2.2 Irradiance

Solar irradiance is the most important meteorological factor affecting PV power generation. In a PV power generation system, the output power P_S of a PV array is calculated by Formula (9) (valid for a silicon-based cell with a typical cell efficiency dependency of -0,005 / K).

$$P_S = \eta SI\left[1 - C_{temp}\left(t_0 + 25\right)\right] \tag{9}$$

where

η is the conversion efficiency of the PV module array (unit: %),

S is the array area (unit: m^2),

I is solar irradiance (unit: MW/m^2),

C_{temp} is the temperature coefficient (unit: %)

t_0 is the ambient temperature (unit: °C).

Using a simplified model, it is assumed that the conversion efficiency of the same type of PV module array is fixed, and the array area S is the unit area, regardless of the incident angle of sunlight and the panel tilt angle. Thus, the output power PS is a function of two variables: ambient temperature t0 and solar irradiance I. Accurate forecasting of solar irradiance is key for accurate prediction of PV power generation in this approach.

Figure 39 shows the volt-ampere characteristic curve of PV modules corresponding to different irradiance at a specific temperature. It can be seen that with the increase of solar irradiance, the open circuit voltage and short-circuit current of the PV module increase. The characteristic curve is gradually shifted towards the outside and the output power is increased. Generally, the stronger the irradiance, the greater the PV power. The irradiance is proportional to the current flowing through the PV module. When the irradiance varies from 100 W/m2 to 1 000 W/m2, the component current always increases linearly with the increase of the irradiance, while the irradiance has little effect on the component voltage. Under the condition of fixed temperature, the open circuit voltage of the PV module remains basically constant when the irradiance varies from 400 W/m to 1 000 W/m. Therefore, the output power of the PV module is substantially proportional to the irradiance.

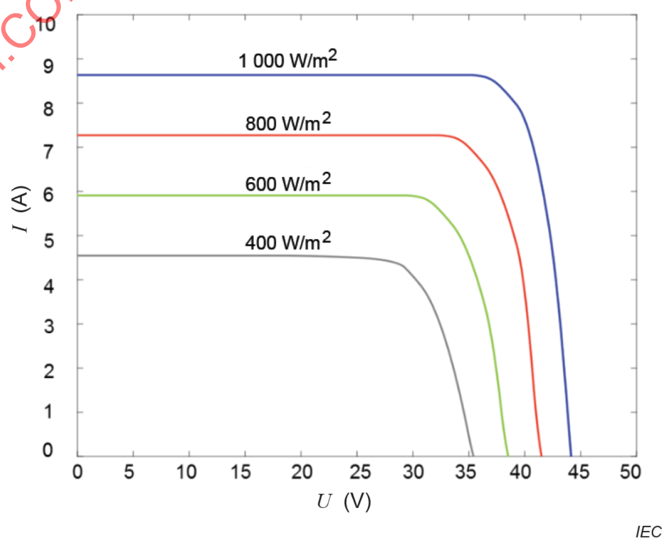


Figure 39 – Volt-ampere characteristic curve of PV modules corresponding to different irradiance

8.2.2.3 PV module temperature

Figure 40 shows the volt-ampere characteristics of PV modules at different temperatures. PV modules typically have an efficiency of 20 %, with most of the remaining radiation lost as heat, thereby increasing the temperature of the module. Therefore, the operating temperature of a solar PV module is usually higher than the ambient temperature.

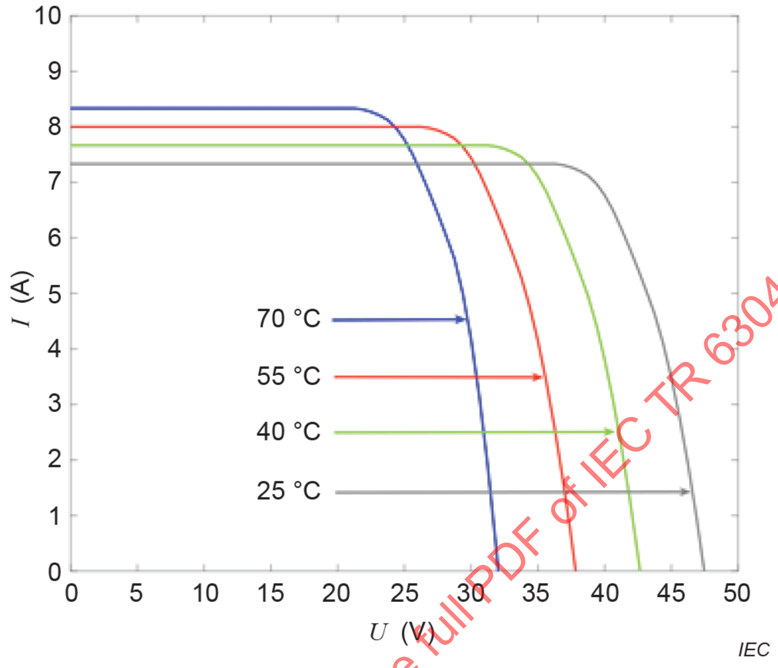


Figure 40 – Volt-ampere characteristics of PV modules at different temperatures

The influence of module temperature on the power generation of PV power stations is mainly reflected in the change of PV module performance. As the temperature of PV module increases, the open circuit voltage decreases, and the working efficiency decreases. Under normal circumstances, in the range of 20 °C to 100 °C, the voltage of the PV module will be reduced by 2 mV for every 1 °C increase; the current increases slightly with temperature; the current of each module increases by 0,1 % or 0,003 mA/(°C·cm^2) for every 1 °C increase. In general, the temperature rises, the power of the solar module decreases, and the typical temperature coefficient is 0,35 %/ °C, that is, the power is reduced by 0,35 % for every 1 °C increase in the temperature of the PV module – note this is valid for silicon cell type modules (which are most common). Therefore, PV module temperature has an impact on PV power generation.

8.2.2.4 Ambient temperature

The ambient temperature is a physical quantity indicating the degree of heat (or cold) of the air, and also reflects the average kinetic energy of the air molecules. On clear-sky days, the ambient temperature is directly affected by the solar irradiance, with increasing solar irradiance resulting in increased ambient temperature. However, on partly cloudy or overcast days, the relationship between solar irradiance and temperature is more complicated. Nonetheless, ambient temperature can be used as a proxy for approximating the irradiance over a site, with higher ambient temperatures generally correlating with higher irradiance and therefore power output. At the same time, higher ambient temperatures can result in higher PV module temperatures, thereby decreasing the module efficiency and the power output.

8.2.2.5 Wind speed

The increase of wind speed promotes the air flow on the surface of the PV module to a certain extent, thereby lowering the surface temperature of the component, which is beneficial to improve the photoelectric conversion efficiency and increase the output power (in the case of silicone-based modules). At the same time, however, the wind increases the probability of dust adhesion on the surface of the PV module, and the presence of dust will reduce the photoelectric conversion efficiency of the PV module. In general, high wind speeds allow more dust to adhere to the surface of the PV module, however, the light transmittance of dust adhesion layer generated in high wind speed is higher than that in low wind speed.

8.2.2.6 Relative humidity

Relative humidity provides a measure of the water content in the air, with larger relative humidity values corresponding to higher water vapor content. Humidity can impact PV generation in two ways. First, higher humidity can decrease convective heat transfer of the PV modules, thereby increasing the module temperature compared to low humidity conditions. Second, higher humidity, and therefore water vapor content, can increase the attenuation of solar radiation in the atmosphere, thereby decreasing the solar irradiance that reaches the PV modules and decreasing the power output. Additionally, relative humidity is a function of both the ambient temperature and pressure. With the pressure fixed, increasing ambient temperature can increase the humidity, thereby decreasing the power output.

8.2.2.7 Cloud cover

Cloud cover refers to the extent to which the cloud obscures the view of the sky. Generally, clouds are given in octas. One octa corresponds to 1/8 of the sky blocked by clouds. If there is no cloud in the sky, the amount of cloud is 0, if half of the sky is covered by clouds, the amount of clouds is 4. Solar radiation travels in the atmosphere and is decreased by the scattering and absorption of air molecules (mainly CO_2, O_3, O_2 and other gases), water vapor, and aerosol particles. In clear and cloudless weather, the cloud cover is small, the transparency of the atmosphere is high, and most of the solar radiation can reach the surface. Therefore the output of PV power stations will be large. In cloudy or rainy days, the cloud cover is large, and the transparency of the atmosphere is low, and only a small fraction of solar radiation can reach the surface. Therefore the output of the PV power station is small. In addition to scattering and absorbing solar radiation, the cloud also reflects a portion of the solar radiation back into space. The ability of the cloud to reflect depends mainly on the type, thickness, and cloud cover. The impact of clouds cover on PV station processing is reflected indirectly.

Note that, in addition to cloud cover, the type of cloud and its optical properties can have a significant influence on PV power generation. For example, a sky with 25 % cloud cover composed of a few optically thick clouds, for example cumulonimbus, may attenuate more irradiance over a PV plant compared to a sky with 100 % cloud cover composed of optically thin clouds, for example cirrus. In practice, cloud cover is more commonly integrated in forecasting methods than cloud optical properties, but there is ongoing research to accurately model and incorporate cloud optical properties into operational forecasts.

8.2.2.8 Air pressure

Air pressure is a comprehensive reflection of the weight and movement of air molecules. The change of air pressure will cause the change of the air flow on the surface of the PV module, which will affect its heat dissipation and the output of the PV power station.

8.2.3 Basic concepts for short-term PVPF

8.2.3.1 Input data

The input data can be divided into endogenous and exogenous datasets. Endogenous datasets are based on the output power of the PV power station, for example lagged values or the variability of the power at the time of forecast generation. Exogenous datasets are based on external data from sources relevant to PV power generation, for example weather stations, irradiance sensor networks, ground-based cloud imagery, satellite imagery and NWP forecasts. In general, the use of exogenous data enables improved forecast performance over forecasts based on endogenous-only data. For forecast horizons more than 24 h ahead, using NWP forecasts as input generally results in the best forecasting performance, but the specifics can vary between PV stations, climates, etc.

8.2.3.2 Output data

The short-term forecasting of PV power generation should be able to provide the future output power of PV stations in a manner that lines up with system scheduling practices. For example, the time is from midnight in the next day to the next 72 h to 120 h (although the details here vary around the world), with a time resolution of at least 5 minutes, but often 15 min or 60 min is sufficient depending on the use case. Short-term forecasting is performed at least twice a day, often hourly, and the time for one-time calculation should be less than 5 min to allow for input into the dispatch.

8.2.4 Short-term PVPF model

8.2.4.1 General

The short-term forecasting models of PV power generation are currently divided into two major categories: statistical forecast models that learn from data and physics-based forecast models that use physical modelling of the power based on irradiance, temperature, etc. [172] to [174].

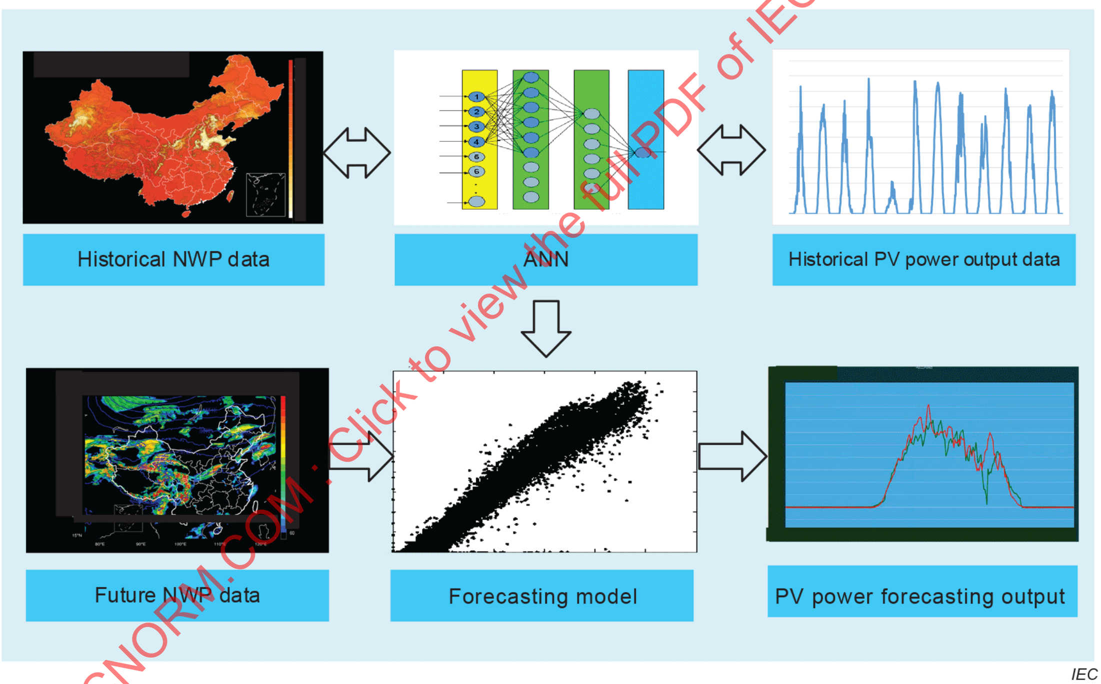


Figure 41 – Short-term forecasting models of PV power generation

8.2.4.2 Statistical forecasting model

Statistical forecast models learn to predict the power output from data, as described in Clause 7. In the training stage, a statistical model is defined, and its weights are tuned to create a mapping from input data, for example historical NWP forecasts of irradiance, wind speed and temperature, to output data, i.e., the power output of the PV station. After training, the model is "frozen", i.e., the learned weights are fixed, and then the model generates forecasts using new values of the inputs used in the training stage. Figure 41 illustrates these two stages. Note that some statistical forecast models use a training process to learn model parameters and then dynamically update the model parameters over time as new data is provided.

The key advantage of the statistical model approach is that the forecast learns directly from the data, with the forecast user simply guiding the process, and can therefore be applied to any PV station with historical data available for training. In practice, such methods can produce state-of-the-art performance and can result in good forecast performance even with endogenous-only features. However, such models can be difficult to train, suffer from overfitting, and require large amounts of training data to achieve robust performance. In general, the effectiveness of statistical methods depends to a large extent on whether there is a large, complete, and evenly distributed historical data of PV power stations to support the training of forecasting models. A PV clear sky model can be used to detrend PV output and improve forecasts, such as in [175], [176]. The redundant information introduced by the correlation coupling between multiple inputs will also directly affect model learning and training effects. In addition, changes in mapping relationships under different time scales, working conditions, and weather conditions may result in performance degradation or even failure of the model. These problems bring certain difficulties to the use of statistical forecasting methods in both theory and practice.

8.2.4.3 Physical forecasting model

A physical forecasting model approach is based on resource-to-power modelling of a PV station, rather than training a model to directly learn the mapping from inputs to power output from historical data. At minimum, the model estimates the power output as a function of the irradiance, but it can include a range of sub-models and inputs related to the PV modules, solar geometry, tracking system, inverters, etc. Once the model is developed, forecasts of the inputs are fed into the model to generate a power forecast. Therefore, the forecast accuracy is directly tied to the accuracy of 1) the model inputs and 2) the physical model itself.

a) Major meteorological influence factors forecasting. The main meteorological influence factors of PV power generation include irradiance, ambient temperature, relative humidity, and wind speed. The forecasting values of the above influence factors can be obtained using the NWP technique.

b) PV power system physical model. The power generation output characteristic model of the PV power station is one of the keys to achieving accurate forecasting of power generation after obtaining the forecasted value of the main meteorological influence factor of PV power generation through NWP. The above-mentioned meteorological parameter forecasting result is input into the PV power generation system output characteristic model to obtain the corresponding PV power generation forecasted value. The overall technical route is shown in Figure 42.

The physical approach focuses on the digital models of energy conversion devices (PV modules, inverters, transformers, etc.) and operational control systems from solar to electrical energy conversion processes. The effectiveness of this method depends on the degree of understanding of the internal composition of the research object, the laws it follows and the accuracy of the model parameters, such as PV module models, inverter models and component combination models. The physical model involves many links, complicated processes, and difficult parameter solving. The influence of some factors is especially difficult to simulate by mathematical methods, such as dust cover, rain wash, and will result in the change of the surface cleanliness of PV modules and the physical characteristics degradation of PV modules.

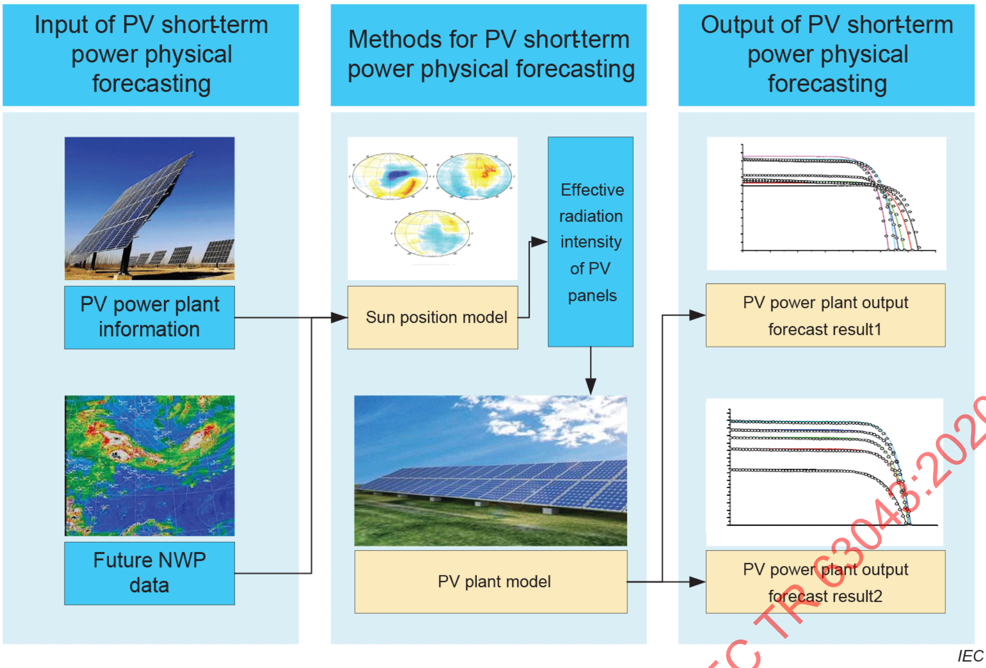


Figure 42 – PV short-term power physical forecasting method technical route

8.2.5 Trends in PVPF development and key technical issues

8.2.5.1 High precision NWP technology

As the input of the short-term PVPF model, NWP has a close relationship with PVPF in terms of accuracy, resolution and precision, which has a significant influence on PVPF results. Generally, error in the non-clear sky conditions is larger than that in the clear sky conditions. Therefore, in order to enhance the accuracy of PV power short-term forecasting, it is necessary to further enhance the precision of NWP [177], [178].

8.2.5.2 Hybrid forecast models

The relationship between the inputs and outputs of a forecast model can vary widely under different weather conditions, for example clear-sky versus partly cloudy versus overcast. In general, training a single model to generate forecasts in all conditions will result in lower accuracy than training a hybrid model that consists of multiple models, each trained to handle one category of weather conditions. However, a key challenge of achieving good performance with a hybrid model is determining which model to switch to for an upcoming period, when the weather conditions for the said period are not yet known. One approach is using a classification model that predicts the upcoming weather conditions, for example clear-sky, partly cloudy or overcast, and then using the predicted classification to switch to the appropriate model.

8.3 Ultra-short-term PVPF

8.3.1 General

Ultra-short-term forecasting of PV power generation focuses on predicting PV power data in the order of hours ahead. In the forecasting literature, such forecasts are generally referred to as intra-day forecasts, with forecast horizons of 1 h to 6 h being typical. Compared with short-term forecasting, the forecasted period of ultra-short-term PVPF is shorter and closer to the forecasting time, which will be used as the basis for real-time scheduling of the power grid.

At present, commonly used models and algorithms in the ultra-short-term power forecasting of PV power generation include machine learning methods, for example ANN and SVR, and time-series forecasting algorithms, for example the auto-regressive integrated moving average

(ARIMA) model. However, as with short-term PVPF, each ultra-short-term PVPF method has its own limitations and there is no single model guaranteed to provide accurate forecasts in all weather conditions [179].

8.3.2 Basic concepts for ultra-short-term PVPF

8.3.2.1 Input data

As with short-term PVPF, ultra-short-term PVPF can incorporate either endogenous or exogenous data sources. In general, exogenous features derived from remote sensing, for example visible channel images from geostationary satellites, or NWP data provide the most benefit to ultra-short-term PVPF.

8.3.2.2 Output data

Ultra-short-term forecasts provide predictions of PV power generation for horizons of 1 h to 6 h ahead, at temporal resolutions of at least 1 h, although 15 min or 5 min resolution forecasts are increasingly common. In contrast to short-term (day ahead) forecasts, ultra-short-term forecasts tend to be generated approximately every hour, although some applications require the forecasts to be generated on a fast cycle, for example every 15 min.

8.3.3 Ultra-short-term PVPF models

8.3.3.1 General

Ultra-short-term PVPF models tend to be based on machine learning methods, for example ANN or time-series prediction algorithms, for example ARIMA [180] to [182]. Current state-of-the-art models tend to use features derived from either remote sensing or NWP sources, with a large percentage of published studies relying on cloud motion vectors derived from sequential satellite images over a target area. However, similar to short-term PVPF, some ultra-short-term PVPF models are based on physical modelling. Figure 43 provides an overview of a typical approach to ultra-short-term PVPF.

8.3.3.2 Persistence forecasting methods

Persistence is one of the simplest forecasting methods and is generally used as a baseline forecast for comparing performance. Usually, it only takes the latest measurement, or the sliding average values of recent historical measurements, as the next forecasting value and therefore its performance can degrade rapidly as the forecast horizon increases. However, persistence is straightforward to implement for all locations, climates, etc., and therefore serves as a reliable lower bound on forecasting performance. In addition, the baseline persistence method can be improved using a correction based on the clear-sky irradiance (or power) model (often referred to as smart persistence in the literature).

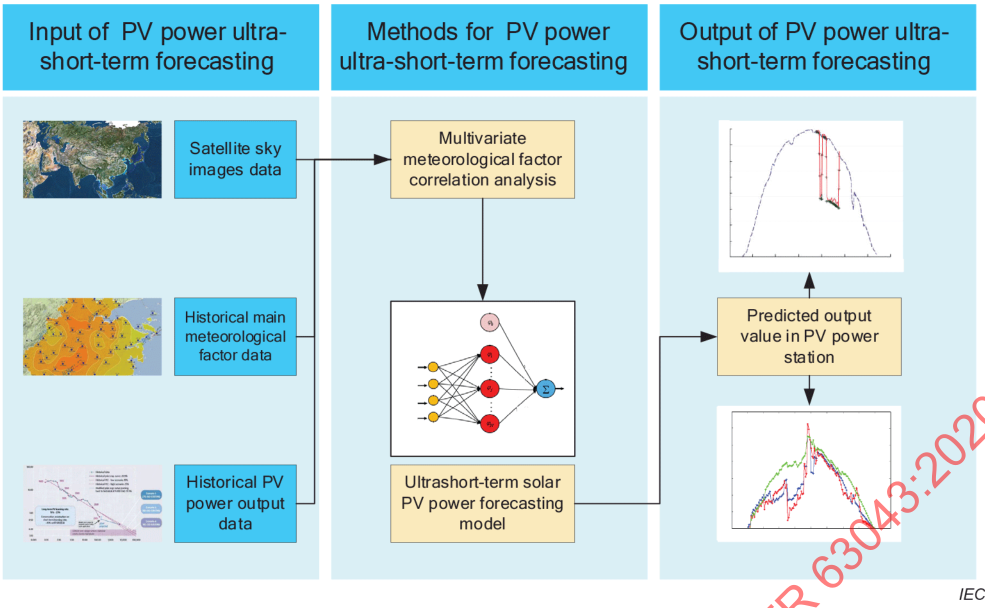


Figure 43 – Basic technology roadmap for pv power ultra-short-term forecasting

8.3.3.3 Machine learning model

Machine learning models can learn mappings from a set of inputs (features) to outputs (targets) by training on historical data. During training, the forecast model parameters are learned directly from the data in order to minimize the error between the model and the historical measurements. Once trained, the forecast models can be used to generate forecasts by providing new values for the required inputs, as shown in Figure 44. At present, the commonly used algorithms for building ultra-short-term power forecasting models of PV power generation based on machine learning include ANN, SVR, and random forest, among others.

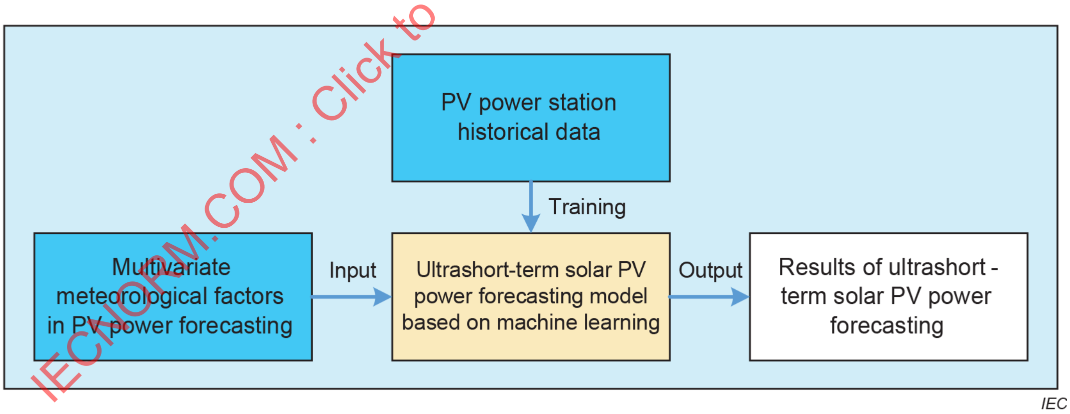


Figure 44 – Ultra-short-term PVPF based on machine learning model

8.3.3.4 Time series algorithm

Time-series prediction algorithms, for example AR, ARMA and ARIMA, are based on fitting to lagged time-series values to future values. Similar to machine learning methods, applying a time-series model for forecasting involves deciding on a model, estimating parameters based on historical data and validating the model. Generally, such methods are straightforward to implement, although model performance can vary.

Time-series methods can be applied to both stationary and non-stationary time-series. Typically, ARMA and related models are better suited for smooth, low-frequency sequences, while ANN and other machine learning methods are better suited for high-frequency sequences. In addition, the Kalman filter and wavelet decomposition methods are often used in the forecasting of time-series to modify the coefficient of forecasting equation in time to improve the forecasting accuracy at the next moment. In the Kalman filtering algorithm, the mathematical model describing the system is the equation of state and the measurement equation. Wavelet analysis and decomposition are commonly used in signal processing, but there is a growing number of studies using such techniques for forecasting.

8.3.3.5 Hybrid models

Hybrid forecasting models are based on maximizing the information provided by multiple models, each tuned to best capture different conditions. They combine the information contained in a variety of single models optimally and take the advantages of different models into consideration, thereby improving the forecast accuracy, and have become a standard approach in the forecasting literature. The generation capacity of large PV power stations is affected by many factors, and a single forecasting model cannot fully include all kinds of factors. Especially in extreme weather conditions, a single model without sufficient learning may lead to a large forecasting error. Select a certain combination mode to integrate multiple models for forecasting and include all the factors that will play a role in the future into the combination model as much as possible.

8.3.4 Trends in development and key technical issues

8.3.4.1 Deep learning methods

The quality of input characteristics determines the upper limit of the predicted results. The traditional ANN method requires manual feature extraction and researchers to have a deep understanding of the physical properties of PV forecasting problems. However, some historical data implies features that cannot be extracted manually. On the other hand, the generalization of the traditional shallow structure learning method of manually extracting features is limited. The influencing factors for the forecasting of PV output power are not completely the same in different regions or different time periods in the same region, and the forecasting of the same model results in poor robustness of the forecasting model. Deep learning methods, which enable end-to-end training and automatic feature extraction from a range of structured data, for example images, have seen a surge in use and success in recent years. While the use of deep learning for PVPF is still in its early stages, there is growing interest among the solar forecasting community.

8.3.4.2 Optimal ensemble methods

Ultra-short-term PVPF is strongly dependent on weather conditions, making it difficult to train single models to perform well in all conditions. One alternative approach is to train a range of forecast models and then combine them in some optimal way to produce an ensemble forecasting model. This approach is similar to the idea of hybrid models, but rather than switching between models, with only one model used for each time period, ensemble methods incorporate forecasts from multiple models for each time period to then generate a single forecast.

8.4 Minute-time-scale PVPF

Minute-time-scale PVPF focuses on forecasting power generation in the order of minutes ahead, when the generation, dissipation and movement of local cloud fields dominate the power output of a PV station. In the solar forecasting literature, such methods are typically referred to as intra-hour, with horizon ranges from 1 min to 30 min or 60 min and temporal resolutions of 1 min to 5 min or 10 min [183], [184].

8.4.1 Basic concepts for minute-time-scale solar power forecasting

8.4.1.1 Input data

As with short-term and ultra-short-term PVPF, minute-time-scale PVPF can benefit from exogenous inputs. In the case of minute-time-scale, the exogenous inputs are typically derived from sky cameras, that is, skycams, which are ground-based imagers that provide visible channel images of the sky. Current state-of-the-art ultra-short-term PVPF forecasts tend to use cloud motion vectors or similar features derived from consecutive images along with endogenous features derived from measured power output, for example lagged power output.

8.4.1.2 Output data

Minute-time-scale PVPF focus on predicting the output power over the next 1 min to 30 min or 60 min, typically at temporal resolutions of 1 min to 5 min.

8.4.2 Technique routine of minute-time-scale solar power forecasting

Surface solar irradiance is the main determinant factor of PV power output. The combined effects of sun and cloud motion determine their relative position and influence the surface solar irradiance directly. At the time scale of minute level, the motion of the sun is almost negligible. Then the motion of clouds becomes the key element to affect the irradiance value. Thus, it is the generation, dissipation, and movement of clouds that mainly cause surface irradiance changes at minute level time scales, and further lead to rapid fluctuations in PV power output.

In sky images recorded at different times, the differences of cloud distribution around a PV power station can be classified into two categories: a) the shape and position of the cloud are different, or b) the shape is basically the same, but the position of the cloud is different. Which category the cloud distribution belongs to depends on the time interval between sky images and the atmospheric environment. Longer time intervals and more variable atmospheric conditions lead to a higher probability of the former condition. However, the generation, dissipation, and deformation of clouds are complex atmospheric physics processes. Influenced by inertia, the changes in cloud shape need a certain amount of time to accumulate and then be reflected in the sky images. Usually, for two sky images with a time interval with a range between 1 min and up to 2 min, the atmospheric changes are not yet able to impact the shape of clouds significantly, except for a few unusual severe weather conditions. This time interval can be longer in the case of a stable atmospheric environment. In the above cases, it can be considered that the shape of the clouds remains the same; only the locations are different. A large number of actually measured sky images in different regions have shown this empirically. Therefore, it is theoretically feasible for a sky image based on linear extrapolation algorithm to predict the cloud distribution in a future sky image based on the cloud motion displacement vectors calculation using historical image sequences.

According to the sky images captured by a ground-based device such as sky camera, the solar irradiance forecasts can be transferred into the forecasting of cloud distribution in future sky images and the model about 'sky image – ground irradiance' respectively based on the research about cloud motion process and solar irradiance attenuation principle. The advantage of modelling by the above method is that the physical process and research routine are quite distinct and it is easier to forecast solar power output compared with other methods.

The technique routine of minute-time-scale solar power forecasting based on sky images is showed in Figure 45. Based on the assumption that the cloud motion velocity and deformation are consistent in the same sequence sky images which are captured by a ground-based device mounted on solar power stations, the image-based features of clouds are extracted by identification and matching so as to obtain the cloud motion displacement vector field and velocity. Then the cloud distribution in a future sky image is predicted based on linear extrapolation. Finally, the solar PV output power can be calculated according to the "sky image-ground solar irradiance-PV power output" model. The technique routine explained above mainly consists of three parts: cloud distinguish at the pixel level, cloud motion displacement calculation method and solar irradiance model based on sky images [185] to [188].

The latest developments related to sky imager-based forecasts include deep learning methods, and the stitching together of groups of imagers into a network to extend the horizon. The uncertainty of most imager-based forecasts is high and outperforms satellite-based forecasts only by being up to 15 min ahead, though networks of imagers may push this to 30 min or more in certain conditions.

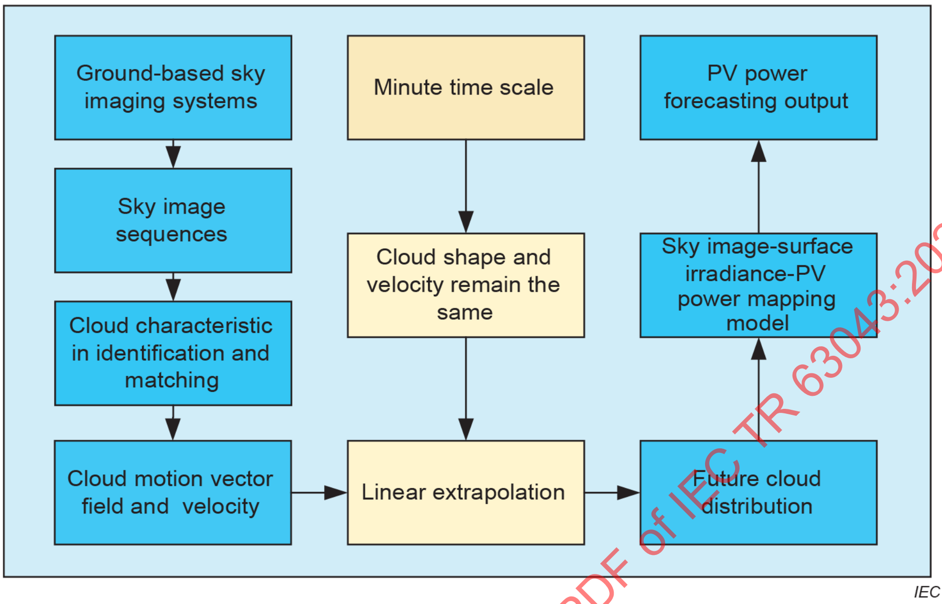


Figure 45 – Minute-time-scale solar power forecasting technique process

8.4.3 Trends in development and key technical issues

In the case when they are shaded by clouds, PV power generation systems without some form of buffer such as storage have a sharply volatile generation in a short time. Large-scale grid-connected solar power may impact on the reliable operations of grids in such cases, though geographic smoothing also has to be accounted for that can reduce the variability observed. One of the effective approaches to handle sharp volatility stemming from PV power during a short period is to employ minute-time-scale solar power forecasting. With the growing requirements of safe operations of grids, future technical demand for PV minute power forecasting will be further highlighted and existing technologies need to be developed further. With the improvement of cloud observation technology and artificial intelligence technology, future progress can be made in the following areas:

a) Cloud extraction technology based on multivariate observation information

With the launch of new and recent meteorological observation satellites, the observation means that can provide cloud information will be more abundant and the observation frequency will be more intensive in the future, which can be better applied to the PV minute power forecasting. With the input of information from polar orbit satellites, geostationary satellites, meteorological radar and all-sky imagers, etc, cloud extraction based on multivariate information and state recognition technology would be the future development of minute-level PVPF technology.

b) Cloud trajectory forecasting technology based on numerical model

Cloud trajectory forecasting is the main problem currently faced, where the method of mathematical derivation is mainly used. However, cloud movement is mainly driven by airflow, which is a non-linear process, leading to the low accuracy of current forecasting methods. Based on the numerical model, it is one of the research directions to improve the accuracy of cloud movement forecasting in the future by introducing the information of, for instance, current cloud location and state into the numerical model, starting with the basic principles of cloud movement, and developing the research of cloud movement forecasting technology.

c) Irradiance intensity cloud attenuation technology based on deep learning

Irradiance intensity attenuation is the basis of solar power forecasting. At present, the ANN method is mainly used to predict irradiance intensity attenuation. The mapping relationship between cloud grey level and irradiance intensity attenuation is constructed, and the effect is good. In the future, based on the premise of abundant sample data, using the powerful learning ability of deep learning theory, the relationship between cloud and irradiance intensity attenuation could be more accurately constructed, and the accuracy of irradiance intensity attenuation forecasting can be improved.

8.5 Probabilistic PVPF

At present, most of the existing PVPF methods focus on single point or regional forecasting (i.e. deterministic forecasting), which can only provide a single forecasted value of PV output power or the assumed output in a region. However, the information contained in the single value forecast is limited, thus it is difficult to describe the uncertainty of forecasting [189]. For regions with high penetration of PV, the forecasting accuracy of the existing single point PVPF methods is difficult to meet the requirements of the optimal operation of the power system. In this context, more and more attention has been paid to the PV power probabilistic forecasting, which can not only produce the PV output power forecasted values, but also provide the corresponding uncertainty information, thus providing better support for the operation of the power system [190] to [192]. Probabilistic forecasts can be divided into the same timescales as deterministic ones.

8.5.1 Basic concepts of PV power probabilistic forecasting

8.5.1.1 Input data

Input data can be divided into two categories according to the different sources of data: one is composed of current and lagged forecast data of a PV power station (so-called “poorman’s ensemble”), the other is based on ensemble NWP data.

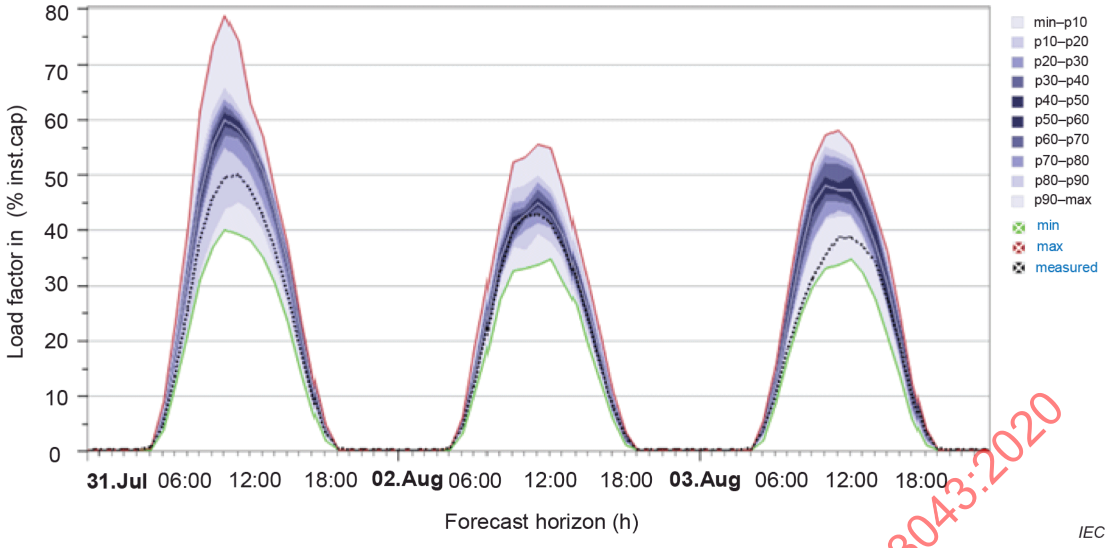
Forecast data (e.g. temperature, relative humidity, solar irradiance, cloud volume, wind speed, wind direction and air pressure, etc.) can be supplemented by local measurement data, ground-based cloud data, satellite image data for training of the forecast models or forecast adaptation in short-term forecasts.

In addition, the output power of PV power stations in a given region is likely to be temporally and spatially correlated. Therefore, incorporating the historical output power data, meteorological data and NWP data of near field stations in suitable areas as input to the forecasting model and mining the spatial-temporal dependence between the data can help to improve the accuracy of the PV power probabilistic forecasting.

More details on probabilistic NWP forecasting can be found in 5.5.2 and in 7.4.

8.5.1.2 Output form

There are three main output forms of PV power probabilistic forecasting results: PDF, CDF [193] and prediction (or confidence) interval [194], [195] and quantile [196]. Figure 46 shows an example of a probabilistic PV model. It can be seen from the figure that not only can probability distribution show mean value forecasting results, but it is also able to make a preliminary estimation of the probability distribution of error.



SOURCE: Courtesy WEPROG.

Figure 46 – Example of probabilistic PV model

8.5.2 Probabilistic PVPF model

8.5.2.1 Physical probabilistic PVPF

The physical probabilistic PVPF model is a short-term probabilistic forecasting model based on the solar radiation and PV power generation system model, where the probabilistic information stems from an ensemble NWP model. Figure 47 shows the basic idea of physical probabilistic PVPF model: firstly, simulating the probability distribution of irradiance (such as Normal [197], Weibull and Beta distribution [198], etc.) by selecting a parameter distribution according to empirical information; then, establishing the probabilistic forecasting model of PV power according to the functional relationship between PV output power and irradiance.

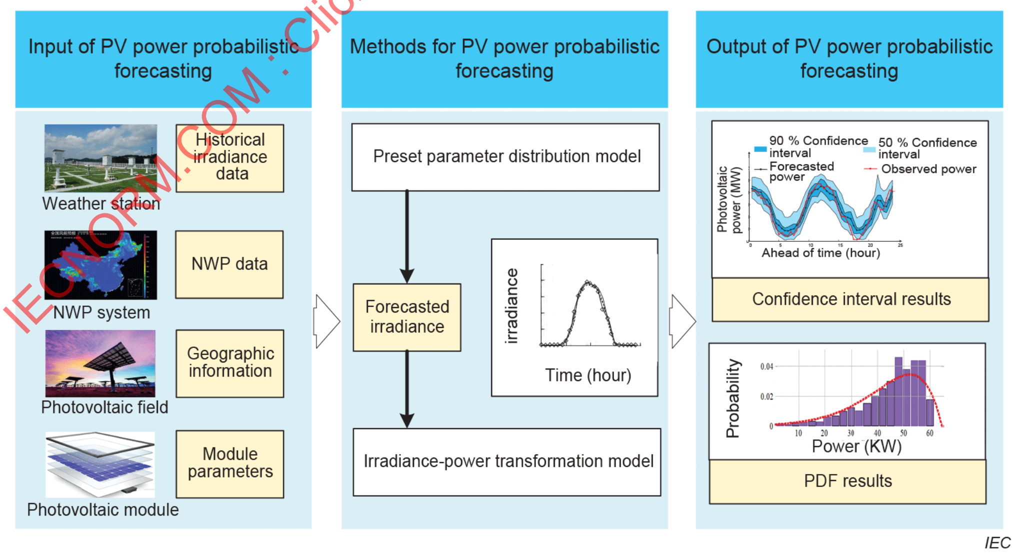


Figure 47 – Forecasting process of physical PV power probabilistic forecasting model

The physical probabilistic PVPF model is especially suitable for the newly built PV power station and at areas where extreme weather can and should be expected (see also 5.3.1 and 8.3.3). Once the technical specifications of the power station are acquired and the local NWP data is obtained, the probability distribution of the output of the PV power station can be obtained. However, detailed geographic information and component parameters of the PV power station are required for the construction of this model.

8.5.2.2 Statistical probabilistic PVPF model

The internal information of the PV power station system does not have to be obtained when modelling the statistical probabilistic PVPF model. It is a data-driven model that employs NWP data and PV power historical data to forecast future power output. Therefore, the data used for model training has a significant impact on the forecasting accuracy. Since the construction of statistical probabilistic PVPF model usually depends on a large number of historical data, this model is suitable for PV power stations that have been put into operation at least one year. The commonly used statistical probabilistic PVPF models are KDE, QR, etc. The forecasting process of the statistical probabilistic PVPF model is shown in Figure 48.

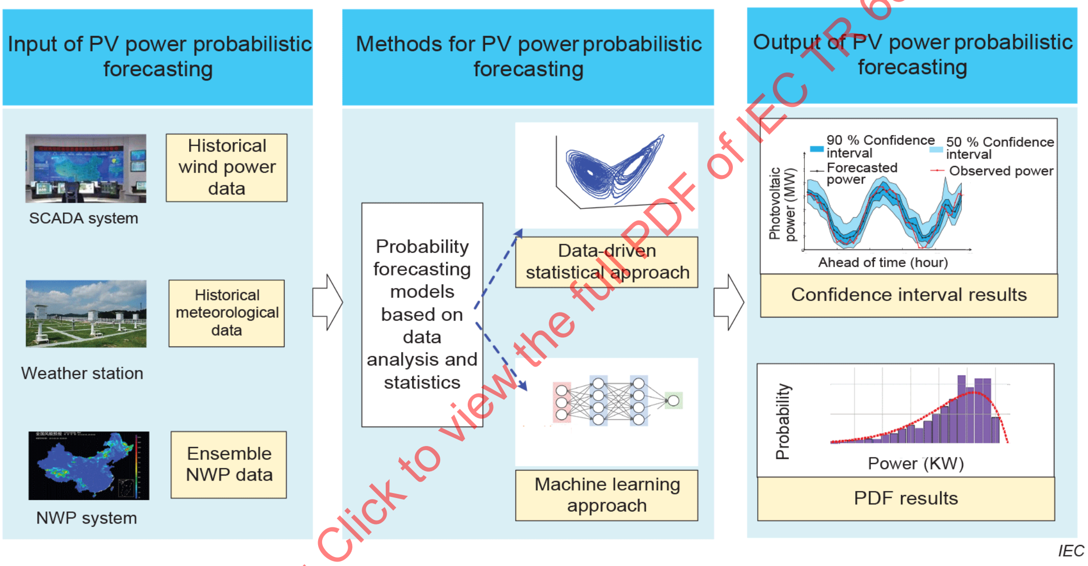


Figure 48 – Forecasting process of statistical probabilistic PVPF model

The KDE model is a historical data-based non-parametric statistical forecasting model which does not need to assume the distribution of PV power in advance. The model forecasting performance is not sensitive to the selection of kernel functions, mainly depending on the setting of bandwidth parameters.

The QR model fits the mapping relationship between explanatory variables and response variables directly, and it is also a non-parametric statistical probabilistic forecasting model [199]. The biggest advantage of the QR model is that the conversion relationship between the different explanatory variables and the corresponding variables can be obtained through the independent training process, thus forecasting the uncertainty of the PV output effectively.

8.5.2.3 Machine learning-based PV power probabilistic forecasting model

The machine learning-based probabilistic forecasting model can fully exploit the complex nonlinear relationship between input and output data. It can easily be applied to a forecasting case and has a great generalization ability. Machine learning methods commonly used in PV power probabilistic forecasting are Bayesian learning, neural networks [200], etc.

Bayesian learning is a conditional probabilistic forecasting model. This model considers the model parameters as random variables. After considering the prior distribution of the parameters, the posterior distribution of the parameters is estimated based on the historical sampling data. Bayesian Learning theory uses probability to represent the uncertainty of the preset distribution form in the parametric model and implement the learning and inference process of the distributed parameters through Bayesian rule. It can effectively improve the accuracy of the parameter estimation in the parametric forecasting model, thus obtaining the superior probability distribution of forecasting results.

Neural networks have a good capacity of generalization and fault tolerance as they can map complex nonlinear relationships. The nonlinear QR PV power probabilistic forecasting model which combines the neural network model with QR theory can provide the probability distribution function of PV power and has an excellent forecasting performance [201] to [204].

8.5.3 Trends in development and key technical issues

8.5.3.1 General

In general, the research on the PV power probabilistic forecasting model is still in its infancy all over the world. It still needs to be further studied to improve the capability of high-dimensional data processing and to extract the key information in the complex cloud image effectively.

8.5.3.2 Development of centralized PV power probabilistic forecasting model based on big data technology

The centralized distribution of solar energy resources results in the characteristics of large-scale and highly concentrated grid-connected PV power. Therefore, research on the probabilistic forecasting of centralized PV power seems to show a more important practical value. However, centralized PV power probabilistic forecasting exhibits many problems because of the large amount of data and the difficulty in modelling. Therefore, by applying big data processing technology to centralized PV power probabilistic forecasting, the data storage and processing capabilities of the model can be improved, and the correlation between the data can be excavated deeply. Combining big data technology with the existing centralized PV power probabilistic forecasting technology provides a new idea for improving the PVPF accuracy.

8.5.3.3 Deep learning based cloud image information mining technology

Images such as satellite cloud maps and ground-based cloud maps have attracted much attention in the research of PV power probabilistic forecasting model due to their intuitiveness, accuracy and diversity in the information display. However, current image processing and recognition technology fail to make full use of the characteristics of the image. As a result, the effect of cloud classification and cloud computing technology is unsatisfactory. On the other hand, deep learning theory has made major breakthroughs in the past decade, especially the convolutional neural network [205] which has been widely applied in the field of image recognition. Therefore, the PV power probabilistic forecasting model that adopts the satellite cloud image, ground cloud image and other images as the input data can efficiently extract the cloud information in the image and improve the accuracy of the PV probabilistic forecasting results by using the deep learning theory.

8.6 Distributed PVPF

8.6.1 General

In recent years, the installation of distributed PV has increased rapidly, due to a variety of factors including reduction in energy poverty, increased consumer choice and certain policy preferences, and its penetration rate in the power grid is continuously increasing. This could have some potential impacts on the safety and stability of the power system, and so it is necessary to predict the output power of distributed PV and incorporate distributed PV forecasting results into system operations.

The principle of power generation of distributed PV is the same as centralized PV, so the methods used for centralized PVPF also can be applied to the distributed PVPF. However, there are still some differences, which can be summarized as follows:

a) Compared to centralized PV, the distribution area of distributed PV is wider with the same installed capacity. Different from centralized PV, distributed PV panels are mainly installed on rooftops or near customer premises. Distributed PV installations have to adapt to the installation environment (for example non-optimal tilt), while the centralized PV is usually installed in a dedicated region and optimized for PV production. In some regions, measurement data may be challenging to procure to train models and update forecasts, depending on the regulatory schemes used to connect distributed PV. For certain regions, much of this data is available in real time, in other regions, it may be available after a time period has passed, and in other regions, it may not be available. This will have an impact on how many of the centralized PV power forecasting methods can be directly adopted.
b) Distributed PV is often produced and consumed at the same location, (i.e. self-production). After meeting the electricity demand of the customer, the rest of the electricity will be injected back to the power grid through distribution networks. Therefore, the distributed PV power fed into the grid is the residual power with the available PV power generation. Compared to centralized PVPF, distributed PVPF also needs an understanding of the load on the customer side to understand injection. Operators may be more concerned with the net injection of PV rather than overall PV production.

Therefore, centralized PVPF methods can be applied to distributed PVPF; however, these methods should be improved to adapt to the characteristics of distributed PV [206], [207].

8.6.2 Basic concepts for distributed PVPF

8.6.2.1 Basic requirements

According to the requirement of actual dispatch, short-term power forecasting needs to be carried out for distributed PV. The forecasting time scale should be based on the scheduling processes of the region of interest, and is typically 0 h to 24 h, 48 h, 72 h or 168 h, with a time resolution of 5 min, 15 min or 60 min, updated at least twice a day; this will be similar to other wind and solar PV forecasts.

8.6.2.2 Input data

Future weather changes and the output power characteristics should be considered when performing Distributed PVPF. Different forecasting methods require to input different data in the modelling stage. The clustering statistical method needs to forecast the NWP data with a time scale of 24 h to 168 h, the network power data and the installed capacity; the grid forecasting method needs to forecast the NWP data with the relevant time scale (24 h to 168 h), the parameters of the PV panels, and the parameters of the inverter (often lacking). After modelling, the input NWP data should be applied to power forecasting. If the real-time net power data is available and the quality of the data is high, the real-time net power data can be input to optimize the forecasting result.

8.6.2.3 Output data

Different methods may produce different types of results. The clustering statistical method predicts the net power (considering consumer load) directly, and the output data is the forecasting result of the net power, while the grid forecasting method does not consider the impacts of self-use electricity.

8.6.3 Distributed PVPF methods

8.6.3.1 General

Currently, there are two main kinds of distributed PVPF methods: the clustering statistical method and grid forecasting method [208] to [210]; however more may also be possible, such as various neural network-based models. Compared to single point forecasts, the lack of information about the single PV installation results in higher uncertainty. However, the geographical smoothing effect for solar radiation has a strong dampening effect. The local errors are smoothed out looking at regions.

8.6.3.2 Clustering statistical model

When distributed PV net power data is available, a clustering statistical method can be used to forecast distributed PVPF based on the net power output, while considering load. The net power data can be obtained by the linear expansion method. The overall framework is shown in Figure 49.

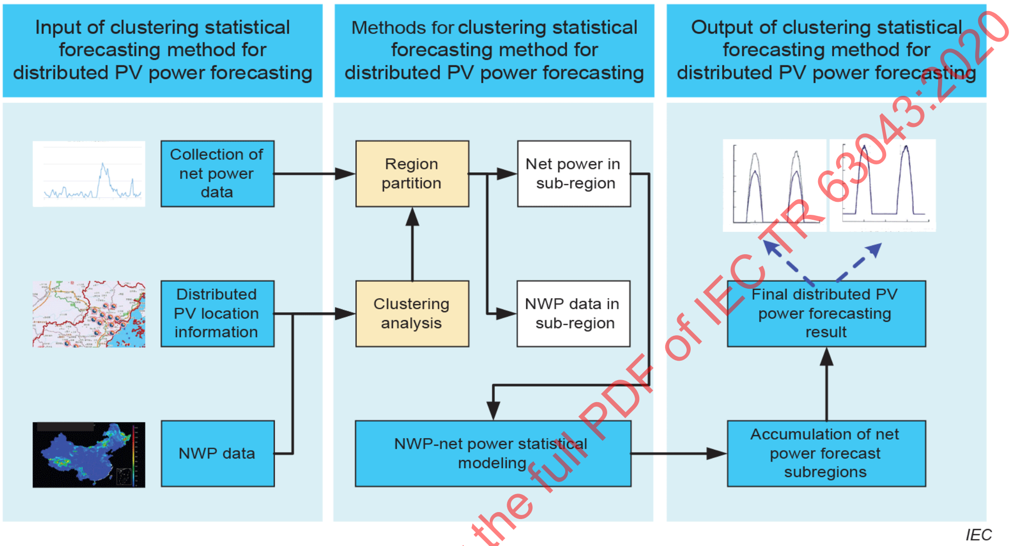


Figure 49 – Framework of clustering statistical forecasting method for distributed PVPF

Considering the wide distribution of distributed PV systems, the clustering analysis can be used to separate distributed PV systems into different sub-regions according to the similarity of output power characteristics such that the distributed PV systems in the same region share similar output power characteristics while those in different sub-regions exhibit different output power characteristics. Statistical methods such as ANN can be used to build the mapping relationship between the meteorological parameters (e.g. NWP irradiance, temperature, wind speed, wind direction, etc.) and the net power in each region. The distributed PVPF results in each sub-region can be obtained by inputting the NWP data into the built mapping model in the corresponding sub-region. Finally, distributed PVPF results in all sub-regions are accumulated to obtain the distributed PVPF results in the whole region.

8.6.3.3 Grid forecasting model

When the net power data is not available, the grid forecasting method can be used to realize distributed PVPF. The framework is shown in Figure 50.

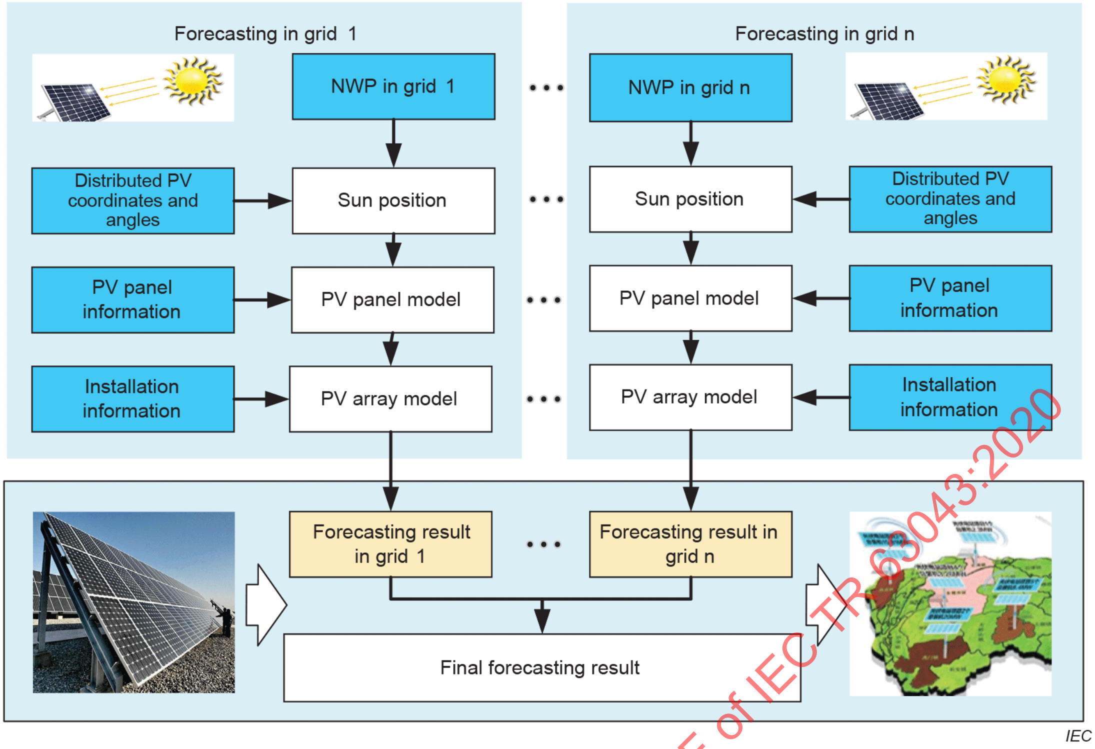


Figure 50 – Framework of grid forecasting method for distributed PVPF

The grid forecasting method is essentially a kind of physical forecasting method. The irradiance information contained in the NWP can be transformed to the irradiance actually received by PV panels according to the grid coordinates and installation angle of the distributed PV system and the sun position. The mathematic model of the PV panel can be built based on the information of the PV panel type, maximum voltage, maximum, current open-circuit voltage, short-circuit current, etc. Thus, the irradiance received by the PV panel can be transformed into the output power. All the PV panel output power in the same grid can be aggregated to obtain the PVPF results in each grid. The final forecasting result can be obtained by the summation of the forecasted distributed PV output power in all grids.

The comparison between the forecasting results obtained by the clustering statistical method and the grid forecast method is shown in Figure 51. It should be noticed that the clustering statistical method is able to directly forecast the net power of the distributed PV. Moreover, it shows high accuracy and is able to meet the requirements of practical power system dispatch. However, the modelling is challenging, since the accumulation of basic data for a certain period of time is required. Compared with the clustering statistical method, the grid forecast method shows better engineering applicability, but it does not consider the problem of self-use electricity, which has to be considered in separate load forecasting models and then grid-integrated.

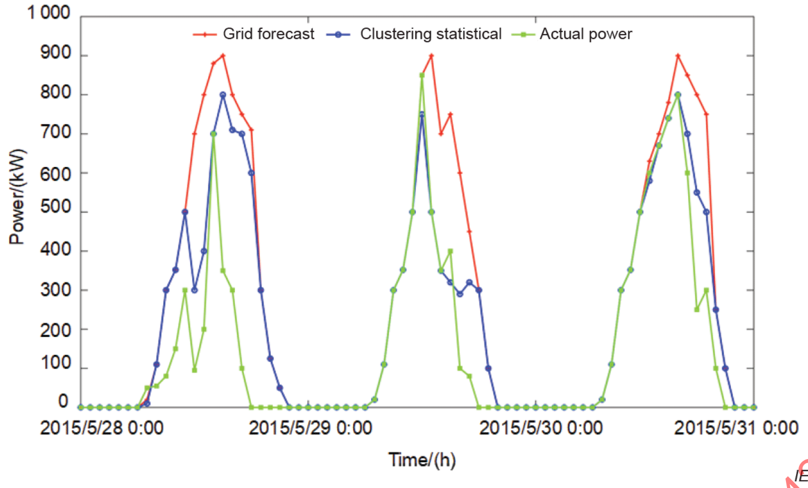


Figure 51 – Comparison between the forecasting results of the clustering statistical method and the grid forecast method

8.6.4 Trends in development and key technical issues

8.6.4.1 General

Data plays a pivotal role in distributed PVPF. In some regions, the collection and accumulation of the distributed PV output power data are deficient. With the further development of big data technology and the accumulation of the distributed PV data, the efficiency and accuracy of the distributed power forecasting will be effectively improved.

8.6.4.2 Distributed PVPF technology based on a big data platform

The modularity of the generation process of the PV power generation system is important in this key technical issue. By utilizing the key parameters, such as the position, tilt angle, types of the PV panels and so on, different power generation systems can be identified, and a distributed PV power generation physical model can be built. Inputting the gridding NWP results into the formulated model and utilizing the powerful data processing ability via a big data platform enable the realization of distributed PVPF at the national level.

8.6.4.3 Forecasting technology based on cloud information

The principle of power generation of distributed PV is exactly the same as the centralized PV. Thus, the cloud information based centralized PVPF methods can be applied to the distributed PVPF. However, distributed PV has a wider distribution area than centralized PV with the same installed capacity, which indicates that the cloud coverage changes gradually. In the future, it is necessary to consider the process of gradual changes in cloud coverage so that the adaptability of the cloud motion information-based forecasting method for distributed PVPF can be improved.

8.7 Summary

Clause 8 introduced the basic concepts and technical route of PVPF, including deterministic centralized PVPF, probabilistic PVPF, and distributed PVPF. Different time scales were covered, including short-term PVPF, ultra-short-term PVPF, minute-time-scale PVPF. The advantages, disadvantages and applicable conditions of the main methods were discussed. Finally, trends in development and key technologies in the future were introduced.

9 Renewable energy power forecasting (RPF) evaluation

9.1 General

In general, the term “renewable power forecast evaluation” refers to all aspects of the assessment of the performance of renewable power predictions. This includes the evaluation of the “quality” and “value” of the operational forecasts employed by end users to meet the needs of their applications as well as the performance of the components of these forecasts, which include the quality and availability of the input data and the error characteristics of the predictions by individual forecast models that contribute to the ultimate forecast employed by end users. The valuation of the individual components of a forecast is primarily of interest to forecast providers and researchers. The interest of most operational forecast users is on the quality and value of the forecasts that is relevant to their applications. Therefore, the focus of Clause 9 is on the concepts and methods used to assess quality and value from an application perspective. There are several key concepts that should be kept in mind in the process of formulating an approach to forecast evaluation.

The first key concept is the distinction between forecast “quality” and “value”. “Quality” is the more general term and refers to the degree of agreement between the forecasts and the actual outcomes. The degree of agreement can be measured through the use of a large variety of metrics and each of these is a valid measure of the quality of the forecasts. The distinction among them is that they measure different aspects of the degree of agreement. It may be tempting to interpret “forecast quality” as “forecast accuracy” but a simple concept of accuracy is better aligned with a deterministic forecast whereas the accuracy concept is more complex in the case of a probabilistic forecast. In contrast, the term “forecast value” has a more limited definition. In general, value is the benefit (economical or others) gained from the use of the forecasts in an application (e.g. in operational grid management decision-making). It may be expressed in a monetary unit or it may be expressed in the units of the forecast target variables. In either case, it attempts to measure the attribute(s) of the forecast error distribution (e.g. error magnitudes during ramp events, at times of peak load, etc.) to which the user’s application has a significant sensitivity. Thus, value metrics are a customized subset of the very large set of forecast quality metrics. The important point is that a generic assessment of forecast quality does not necessarily provide a good assessment of value. This is only the case if the quality metrics provide a good representation of the way in which a user’s application is sensitive to forecast error.

The second key concept is the relationship between quality metrics and the forecast type. From a broad perspective there are two key attributes of a forecast: (1) deterministic versus probabilistic and (2) continuous versus categorical. Deterministic forecasts only provide a prediction of the future value of the forecast target variable. A probabilistic forecast is based on a probability distribution of the forecast target variable that can be viewed as a prediction of the future value (e.g. the most likely value) and an estimate of the uncertainty of that prediction. Thus, a probabilistic forecast provides much more information than a deterministic forecast and a useful forecast evaluation process should evaluate the quality of all the additional information in the probabilistic forecast. A forecast for a continuous target variable attempts to predict the value of that variable or the probability distribution of the values of that variable. In contrast, a categorical forecast approach attempts to predict the occurrence or non-occurrence of a category (which may be conceptually associated with an event such as a ramp). If the event is predicted to occur, the forecast may also attempt to predict the continuous attributes of the event (e.g. amplitude, duration, timing of a ramp event).

These two key attributes result in four broad classes of forecasts: a) deterministic forecasts of continuous variables; b) probabilistic forecasts of continuous variables; c) deterministic forecasts of categorical (event) variables; d) probabilistic forecasts of categorical (event) variables.

From a forecast evaluation perspective, the key point is that there is a different set of metrics that are appropriate for the evaluation of each class of forecast. Subclauses 9.2 to 9.5 provide an overview of the methods and metrics for the evaluation of each of these classes.

9.2 Deterministic forecasts of continuous variables

9.2.1 General

Deterministic time series forecasts of a continuous variable are the most common form of short-term variable renewable energy predictions. These are typically predictions of the wind or solar power generation in MW for discrete forecast time intervals (e.g. every 15 min or hour). The deterministic forecast provides a value of the forecast target variable for each interval. The evaluation of these forecasts is focused on assessing the properties (average, distribution, extremes, etc.) of the deviations (i.e. which can be termed forecast errors if the observed values are reliable) of the forecasted values from the corresponding observed values over an evaluation sample.

9.2.2 Metrics

There are many metrics available to assess the error characteristics of this type of forecast. The most widely used are the mean bias error (MBE) which is also known as the “bias”, the mean absolute error (MAE) and the Root Mean Square Error (RMSE). These three metrics provide an indication of the “typical” error of the forecasts over the evaluation period. In a number of applications, these three basic error metrics are used as the basis for the calculation of a “skill score”, which is the percentage improvement in a specified metric (e.g. MAE or RMSE) produced by a target forecast relative a reference metric. A fourth metric that is widely used but less popular than the top three is the correlation coefficient. This measures the degree of agreement between the temporal shape of the sequence of forecasts and outcomes. It is not sensitive to the forecast bias. Another attribute that is often of interest to forecast users is the magnitude and/or frequency of large errors. Four metrics frequently used for this purpose are the maximum prediction error (MPE), the maximum error rate (MER), the 95 % quantile deviation rate and the pass rate. The first two of these metrics provide a measure of the largest magnitude error. The latter two provide a broader indication of the characteristics of the tails of the error distribution. Subclauses 9.2.8 to 9.2.10 provides additional details about each of these metrics.

9.2.3 Mean bias error

The mean bias error, which is typically abbreviated as MBE, is the average forecast error of a sample of forecasts. Since positive and negative errors will cancel in the summation, it provides a measure of net forecast error, which is also known as the “bias”, and which is sometimes referred to as the systematic forecast error. This is the tendency for the forecasts to be systematically “too high” or “too low”. The MBE is calculated by Formula (10).

$$e_{\text{MBE}} = \frac{1}{N}\sum_{n=1}^{N}\left(P_{\text{P},n} - P_{\text{M},n}\right) \qquad (10)$$

where

e_{MBE} is the mean bias error (unit: MW),

$P_{\text{M},n}$ is the measured power (unit: MW) production for the forecast period n,

$P_{\text{P},n}$ is the predicted power (MW) production for the same period,

N is the total number of forecast periods in the evaluation sample.

In many applications, it is desirable to calculate the MBE as a relative quantity. The most common approach is to express it as a percentage of the generating capacity of the forecast target facility. This is often called the mean bias percentage error (MBPE) and can be calculated from Formula (11).

$$e_{\text{MBPE}} = \frac{1}{N}\sum_{n=1}^{N}\frac{\left(P_{\text{P},n} - P_{\text{M},n}\right)}{C_n} \tag{11}$$

where

e_{MPBE} is the mean bias percentage error,

C_n is the generation capacity (unit: MW) of the forecast target entity for the forecast interval n.

The actual generation for forecast period n is sometimes used as the denominator of Formula (11). However, this approach can lead to misleading results during forecast intervals in which the actual generation is near zero since the MBPE can become very large in these cases even though the non-relative (i.e. in MW) MBPE is quite small (i.e. the bias becomes a very large percentage of a very low level of generation). Thus, if it is desired to express the MBE as a percentage of the actual generation, it is best to utilize the average actual generation over the entire forecast sample as the denominator (that is, move C_n outside the summation operation and use an average value over all N forecast periods).

9.2.4 Mean absolute error

The mean absolute error, which is typically referred to as MAE, provides a measure of the average magnitude of the forecast error of a sample of forecasts. It can be calculated by Formula (12). The use of the absolute value of the forecast error results in an average of the error magnitudes rather than the net error as in the case of the MBE.

$$e_{\text{MAE}} = \frac{1}{N}\sum_{n=1}^{N}\left|P_{\text{P},n} - P_{\text{M},n}\right| \tag{12}$$

The symbols in the Formula (12) are as defined in Formula (10). As in the case of the MBE, the MAPE is calculated by Formula 13).

$$e_{\text{MAPE}} = \frac{1}{N}\sum_{n=1}^{N}\frac{\left|P_{\text{P},n} - P_{\text{M},n}\right|}{C_n} \qquad 13)$$

As with the MBPE, C_n is typically specified as the generation capacity of the forecast target entity for interval n but in some forecasts, evaluators prefer to use the actual generation as the base of the percentage calculation. If this is the case, it is also advisable to move C_n outside of the summation operation and employ an average value over all N forecast intervals in the evaluation sample to avoid the problem of meaninglessly high (or infinite) percentage values when the actual generation is near or equal to zero.

9.2.5 Root mean square error

The RMSE is also a measure of the typical magnitude of the error of a forecast sample. However, it differs from the MAE due to the use of the squared error in the error summation. This weights larger errors more heavily in the averaging process. This is often desirable because larger errors have a disproportionate impact on many (but of course not all) applications. The RMSE is calculated by Formula (14).

$$e_{\text{RMSE}} = \frac{1}{N}\sqrt{\sum_{n=1}^{N}\left(P_{\text{P},n} - P_{\text{M},n}\right)^2} \tag{14}$$

The symbols in the formulae are as defined in Formula (10). The RMSE is also often expressed as a percentage of the generation capacity in a manner similar to that done in the MAPE calculation (i.e. by dividing each difference).

In addition to its heavier weighting on larger forecast errors, the RMSE is also widely used because of three additional attributes. First, the default optimization objective for many operational forecast production systems is the minimization of the squared error (i.e. least squares optimization). Therefore, the RMSE is the metric that is consistent with the optimization target of most systems. Second, the RMSE is the standard deviation of the forecast errors and therefore is a measure of the dispersion (spread) of the forecast errors. If the error distribution can be approximated by a Gaussian frequency distribution, then +/- RMSE (one standard deviation) from the MBE value will represent about 68 % of the forecast errors. A third useful attribute of the RMSE is that it can be decomposed into "bias" and "variance" components, which can provide insight into the nature of the forecast errors.

9.2.6 Skill score

In the context of the assessment of forecasts, the skill score (SS) is defined as the relative improvement in a forecast metric with respect to a reference forecast. Additional information about the skill score metric is presented in [211]. The SS is calculated by Formula (15).

$$e_{\mathrm{SS}} = \frac{\left(M_{\mathrm{R}} - M_{\mathrm{T}}\right)}{M_{\mathrm{R}}} \tag{15}$$

where

e_{SS} is the skill score,

M_{T} is the value of the "base" forecast performance metric for the target forecast,

M_{R} is the value of the same metric for the reference forecast.

This formulation of the SS assumes that lower values of the base metric represent better performance (e.g. smaller error). In this case, a value of zero for M_{T} will result in a SS of 1,0. If M_{T} = M_{R}, the SS will have a value of zero (i.e. no improvement over the reference forecast). If M_{T} is larger than M_{R} (e.g. higher error), then the SS will be negative, which indicates worse performance than the reference forecast. A "naive forecast" is often selected as the reference forecast. Typical specifications for a "naive forecast" are persistence for short-term power generation forecasts and climatology (long term average production) for day ahead and longer forecast periods. However, in principle, the reference forecast could be any other forecast of interest (such as an alternative forecast solution) to the user or evaluator.

9.2.7 Correlation coefficient

The correlation coefficient (r) can reflect the correlation between the predicted power and the actual power fluctuation. The r can be calculated with Formula (16).

$$r = \frac{\sum_{i=1}^{n}\left[\left(P_{\mathrm{M},i} - \overline{P}_{\mathrm{M}}\right)\left(P_{\mathrm{P},i} - \overline{P}_{\mathrm{P}}\right)\right]}{\sqrt{\sum_{i=1}^{n}\left(P_{\mathrm{M},i} - \overline{P}_{\mathrm{M}}\right)^{2}\sum_{i=1}^{n}\left(P_{\mathrm{P},i} - \overline{P}_{\mathrm{P}}\right)^{2}}} \tag{16}$$

where

$\overline{P_{\mathrm{M}}}$ is the average of all measured power data samples (unit: MW),

$\overline{P_{\text{P}}}$ is the average for all predicted power data samples (unit: MW).

Other symbols in the formula are as defined in Formula (10).

9.2.8 Maximum prediction error

In some applications, it is valuable to assess the extreme forecast errors rather than the typical error. One of the simplest ways to do this to use the maximum prediction error (MPE) metric. The MPE is defined as the maximum deviation of single-point power forecasting, and it can be calculated by Formula (17).

$$e_{\text{MPE}} = \max\left[\,|\,P_{\text{P},n} - P_{\text{M},n}\,|\,\right] \tag{17}$$

The symbols in the formula are as defined in Formula (10). The MPE is sometimes calculated based on the relative forecast error. In this case, it is often referred to as the maximum error rate (MER). The MER is calculated by Formula 18).

$$e_{\text{MER}} = \max\left[\frac{(P_{\text{P},n} - P_{\text{M},n})}{C_n}\right] \qquad 18)$$

The symbols in Formulae (17) and (18) are the same as those defined in Formula (10). As in the previously described metrics, the C_n in Formula (18) is typically defined as the installed generation capacity of the facility for forecast interval n. However, it could be defined as the actual generation for the forecast interval. As noted previously, this can be problematic because the actual generation for an interval can be near or equal to zero.

9.2.9 Pass rate

Another useful attribute of forecast performance is the fraction of forecasts whose absolute error is below a threshold. This is often relevant for operational activities that are primarily sensitive to errors above a threshold magnitude. The pass rate (Q_{R}) is defined as the percentage of deterministic forecasts that achieve an absolute error below a specified threshold T. For many electric grid operational applications, a T of 25 % of the installed capacity is used. The Q_{R} is a useful reference indicator for the value of forecasting results in power scheduling applications. It can be computed by the relationships in Formulae (19) and (20).

$$Q_{\text{R}} = \frac{1}{n}\sum_{i=1}^{n} B_i \times 100\% \tag{19}$$

$$B_i = \begin{cases} 1, & \dfrac{\left|P_{\text{P},i} - P_{\text{M},i}\right|}{C_i} < T \\[2ex] 0, & \dfrac{\left|P_{\text{P},i} - P_{\text{M},i}\right|}{C_i} \geq T \end{cases} \tag{20}$$

where

Q_{R} is the pass rate,

B_{i} is a binary variable that is 1 if the absolute error is within the specified threshold and zero otherwise.

The other symbols are as defined in Formula (10).

9.2.10 95 % QDR

95 % QDR includes the 95 % QPDR and the 95 % QNDR. The data set for positive deviation of single-point forecasting is sorted from the smallest to the largest during the evaluation, and the single-point positive forecasting positive deviation rate at the 95 % position is selected as the 95 % QPDR. The expression is shown in Formula (21).

$$\begin{cases} E_i = \dfrac{P_{\mathrm{P},i} - P_{\mathrm{M},i}}{C_i} \geq 0, i = 1,2\ldots n \\ E_j = \mathrm{sortp}(E_i), j = 1,2,\ldots m \\ P_{\mathrm{er}_{95}} = E_j, j = \mathrm{INT}(0{,}95 \cdot m) \end{cases} \tag{21}$$

The data set for negative deviation of single-point forecasting is sorted from the largest to the smallest during the evaluation, and the single-point forecasting negative deviation rate at the 95 % position is selected as the 95 % QNDR. The expression is shown in Formula (22).

$$\begin{cases} E_i = \dfrac{P_{\mathrm{P},i} - P_{\mathrm{M},i}}{C_i} \leq 0, i = 1,2\ldots n' \\ E_j = \mathrm{sortn}(E_i), j = 1,2,\ldots m' \\ P_{\mathrm{er}_{95}} = E_j, j = \mathrm{INT}(0{,}95 \cdot m') \end{cases} \tag{22}$$

where

$P_{\mathrm{er}_{95}}$ is the 95 % QDR, and the step is determined according to the specific situation,

E_i represents the predicted deviation rate at time i,

E_j is the single-point forecasting bias rate after ranking,

sortp(.) is a sort function from small to large,

sortn(.) is a sort function from large to small,

INT(.) is the rounding function, and

n and n' stand for the number of positive deviation samples and negative deviation samples in the evaluation period, respectively.

The evaluation period should cover at least one year of the period spanned by the data sample.

9.2.11 Customized metrics

A number of widely-used forecast performance metrics for deterministic forecasts of continuous variables have been described in the preceding subclauses. A few of them might be classified as "standard" methods. These include the MBE, MAE and RMSE. These are the metrics that are typically used by default in the process of evaluating deterministic forecast performance. They offer the advantage of an apparently easy comparison with the performance experienced by other users. There is some value in that perspective. However, there are many factors that cause variations in these metrics apart from the skill of a specific forecast solution. Most of these factors are related to the typical magnitudes and time scales of the variability of the power generation (i.e. the forecast target variable) experienced the forecast target entities. For example, facilities (i.e. locations) that have typically had a higher generation variability will generally have larger forecast errors. Thus, it is misleading to compare MBE, MAE or RMSE values among facilities. A better approach is to use a SS based on the MAE or RMSE and a suitable reference forecast. The SS will provide a more meaningful comparison of forecast performance among a set of locations or even the same location in different seasons that are characterized by different variability characteristics.

However, a more critical issue is the measurement of the attributes of forecast performance that are most important for a user's application. In some cases, standard or widely-used metrics will serve this purpose fairly well. However, in many cases, a customized metric should be used to obtain the most relevant assessment for a specific user. Ideally, the customized metric should be a "loss function" that models the way in which a user's application is sensitive to forecast error. This of course requires that the user has an adequate quantitative understanding of that sensitivity. However, the application-sensitivity to forecast error can often be quite complex and it may be a difficult project or model if this has not already been done. If this is not feasible, it is best to employ a matrix of forecast performance metrics to assess a range of performance attributes rather than relying on one or two standard metrics.

9.3 Deterministic forecasts of categorical (event) variables

9.3.1 General

A deterministic "yes or no" prediction of the occurrence or non-occurrence of a pre-defined set of event criteria within a forecast window is the most frequent application of this concept for renewable energy forecasting applications. The most common type of event that is the target for this application is a large change in power generation over a short period of time, which is typically referred to as a "ramp". Many different definitions of ramp events with different time scale, amplitude and initial and end time criteria have been employed. However, ramp events are not the only type of events that may be of interest in applications of wind and solar power forecasting.

The following overview of deterministic evaluation metrics focuses on single category events. In a single-category event there is only one type of positive outcome. A multi-category event includes more than one type of positive outcome. For example, there could be a "large ramp", a "small ramp", or "no ramp". Thus, there are two types of positive outcomes. All of the metrics can be extended to the multi-category case. However, while in some cases the extensions will be noted in the following metric descriptions, a comprehensive presentation of the multi-category cases is not included.

9.3.2 Occurrence/non-occurrence metrics

Deterministic event forecast metrics can in general be linked to a contingency table of forecasts and the corresponding outcomes. A conceptual representation of a single category forecast-outcome contingency table is shown in Table 8. The table has four boxes or cells that represent the four possible combinations of forecasts and outcomes in a deterministic single category event forecast. Each cell contains an "N" with a subscript. The N represents the number of forecast-outcome pairs in an evaluation sample that fall into that cell. The subscript is the cell label. The upper left cell is labelled "TP" for "true positive" and represents "Yes" forecasts and "Yes" outcomes (i.e. an event was forecasted and it occurred). The upper right cell is labelled "FP" for "false positive" and represents "Yes" forecasts and "No" outcomes. These are forecasted events that did not occur. The lower left cell is labelled "FN" for "false negative". These are events that occurred but were not forecasted. The bottom right cell is labelled "TN" for "true negative" and these are the cases in which the forecast correctly predicted that an event would not occur. The total number of forecast samples in this example is N, which is the sum of all four cells in the contingency table.

Table 8 – Contingency table for forecasts of the occurrence/non-occurrence of an event

Deterministic event forecast-outcome contingency table		**Outcomes**	
		Yes (occurrence)	**No (non occurrence)**
Forecasts	Yes (occurrence)	N_{TP}	N_{FP}
	No (non occurrence)	N_{FN}	N_{TN}

Subclauses 9.3.3 to 9.3.8 describe the conversion of the data in this conceptual contingency table into metrics that quantify a range of attributes of the performance of deterministic event forecasts.

9.3.3 Frequency bias

The frequency bias index (FBI) quantifies the relationship between the frequency of forecasting and observed events. The FBI is calculated by Formula (23).

$$e_{\text{FBI}} = \left(N_{\text{TP}} + N_{\text{FP}}\right) / \left(N_{\text{TP}} + N_{\text{FN}}\right) \tag{23}$$

An FBI of 1,0 indicates that there is no frequency bias. Values above 1,0 indicate that there are too many forecasted events (i.e. "overforecasting") while values below 1,0 indicate that there are too few forecasted events (i.e. "underforecasting").

9.3.4 Probability of detection

The probability of detection (POD) is the fraction of observed events that were correctly forecasted. It is also known as the hit rate (HR) and the POD is calculated by Formula (24).

$$P_{\text{OD}} = P_{\text{HR}} = \left(N_{\text{TP}}\right) / \left(N_{\text{TP}} + N_{\text{FN}}\right) \tag{24}$$

The POD ranges from 0 to 1 and a perfect score is 1,0 (i.e. every event occurrence was correctly forecasted). The POD is only sensitive to correct and missed forecasts of the events. That is, FPs and TNs do not affect this metric. Therefore, the POD can be improved by overforecasting events since it does not penalize "false alarms".

9.3.5 False alarm ratio

The false alarm ratio (FAR) is the ratio of the false positive outcomes to the total number of forecasts. The FAR is calculated by Formula (25).

$$R_{\text{FAR}} = \left(N_{\text{FP}}\right) / \left(N_{\text{TP}} + N_{\text{FP}}\right) \tag{25}$$

The FAR also ranges from 0 to 1 and a perfect score is 0,0, which indicates that there were no false positive outcomes. However, the FAR is only sensitive to false alarms (FNs) and hits (TPs) and therefore it can be improved by underforecasting.

9.3.6 Critical success index

The CSI is the ratio of the number of correct forecasts to the total number of forecasted and observed events. In the meteorological community, it is also known as the threat score (TS). The CSI is calculated by Formula (26).

$$I_{\text{CSI}} = S_{\text{TS}} = \left(N_{\text{TP}}\right) / \left(N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}}\right) \tag{26}$$

The CSI integrates information contained in both the POD and FAR and provides a more comprehensive assessment of deterministic event forecasts than either of those metrics. It does not consider true negative outcomes. However, CSI scores are sensitive to the climatological (long-term observed) frequency of the target events.

9.3.7 Equitable threat score

The ETS is an adjusted TS (also known as the CSI) that accounts for the frequency of the target event. It is sometimes referred to as the Gilbert skill score (GSS). The ETS is calculated by Formula (27).

$$S_{\text{ETS}} = \left(N_{\text{TP}} - R_{\text{TP}}\right) / \left(N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}} - R_{\text{TP}}\right) \tag{27}$$

where

R_{TP} is the number of true positives (i.e. “hits”) that would be expected from a random forecast based on the observed frequency of the events.

The R_{TP} is calculated by Formula (28).

$$R_{\text{TP}} = \left(N_{\text{TP}} + N_{\text{FP}}\right)\left(N_{\text{TP}} + N_{\text{FN}}\right) / N \tag{28}$$

This can also be estimated from long-term historical (climatological) data and this may be more appropriate if the evaluation sample is small. The use of the ETS provides a better comparison on the performance of forecasts for events with different observed frequencies.

9.3.8 Heidke skill score

The HSS is an attempt to provide a more robust reference point for the performance of deterministic event forecasts. It measures the fractional improvement over a random forecast generated with knowledge of the frequency of the observed event. The HSS is calculated by Formula (29).