Supplementary material of:

Towards a future-proof climate database for European energy system studies

Authors
Laurent Dubus, *RTE, Paris, France*
David J. Brayshaw, *Department of Meteorology, University of Reading, UK*
Daniel Huertas-Hernando, *Elia Transmission Belgium, Brussels, Belgium*
David Radu, *ENTSO-E, Brussels, Belgium*
Justin Sharp, *Sharply Focused LLC, Portland, OR, USA*
William Zappa, *TenneT TSO B.V., Arnhem, The Netherlands*
Laurens P. Stoop, *Information and Computing Science, and Copernicus Institute of Sustainable Development, Utrecht University, The Netherlands, Royal Netherlands Meteorological Institute, De Bilt, The Netherlands, TenneT TSO B.V., Arnhem, The Netherlands*

Corresponding author: Laurent.dubus@rte-france.com mailto:

1. Representing the Climate in Datasets

Climate if often thought of as the 'average weather'. More rigorously, it is the statistical description of the weather in terms of the mean and variance of relevant quantities over a period of time for a certain region. The period is usually 30 years as defined by the World Meteorological Organization (WMO 2017). When considering only a region, it really describes the average pattern of weather for that region.

While 30 years is still recommended as a standard averaging period for the calculation of quantile boundaries in climatological standard normal, the stability of most extreme statistics derived from that period is likely to be low for some variables (K. van der Wiel et al. 2019; Wohland, Eddine Omrani, et al. 2019). One could consider three approaches to solve that problem: first, by fitting a statistical distribution, such as a gamma distribution, to the observed data within a standard 30-year period (WMO 2007). Second, to use a period of data substantially longer than 30 years (Wohland et al. 2017; Wohland, Omrani, Keenlyside, et al. 2019). Or third, by using multiple model runs to sample a set period of 30 years (K. van der Wiel et al. 2019).

Each method of characterizing the climate while considering extremes has advantages and disadvantages. The use of a statistical distribution allows for very specific sampling for certain risks, but it relies on the input data and the sampled extremes therein. The use of a longer period is very effective in sampling multi-decadal modes in the climate, but this method obscures the effects of climate change on the occurrence of extremes and thereby might ignore changes in risks. The use of multiple model runs allows for representative sampling of risks, while allowing for climate change signals to be visible, but it requires extensive analysis of risks and biases in the model runs.

Types of Climatological datasets

When discussing climatological or meteorological information, there are different types of datasets that can be used (see Fig. 1). These can be classified into three categories: (1) observational data, (2) reanalysis datasets, and (3) climate projection-based datasets. Each of these dataset types has its strengths and weaknesses.

Observational datasets are extremely good in characterizing historical climate variables at a specific location. The variables under consideration are usually measured with a very high temporal resolution

(Cornes et al. 2018). Their local information is of interest when studying for instance the suitability of a specific wind turbine or solar panel at a specific location (Akpinar and Akpinar 2005b; Chenni et al. 2007). However, these types of datasets have one inherent downside, they are a collection of local measurements. The climatological information is thus only known for an extremely limited number of locations that do not cover all regions. Additionally, due to strict requirements for the way meteorological variables are measured, they generally only sample near-surface variables at open grass fields well outside of built-up areas with very similar topography(World Meteorological Organization 2018).

Reanalysis datasets represent gridded historical climate variables over a large region, based on assimilated observations of those variables and a climate model. The gridded nature of these datasets means that the variables are available for each grid box, but they also represent the whole grid box. The quality of a reanalysis dataset depends heavily on the quality of the assimilated observations, therefore most reanalysis cover the satellite era period as satellites provide consistent observations globally. Commonly used global datasets are ERA5 (Hersbach et al. 2020), MERRA-2 (Gelaro et al. 2017) and JRA55 (Harada et al. 2016; Kobayashi et al. 2015). There are also centennial reanalysis that focus on multi-decadal modes in the climate system (ERA-20C (Poli et al. 2016), CERA-20C(Laloyaux et al. 2018), ERA-20CM(Hersbach et al. 2015), and the 20th century reanalysis project (Compo et al. 2011)).

Climate projection datasets represent simulated gridded climate variables obtained by running global circulation models (GCMs). These models numerically simulate the dynamics of the climate system to derive long-term climate and weather projections. Climate projection datasets generally have lower resolution than reanalysis so they can model multiple iterations for long periods to create an ensemble dataset. When higher spatiotemporal resolution is needed, Regional Climate Models (RCMs) can be used. Transient GCMs typically model the period 1850-2100 with varying concentrations of greenhouse gases; they represent a realistic state of the climate under these emission scenarios. Equilibrium GCMs simulate the response of the climate to a constant concentration of greenhouse gases; they represent the stable state of the climate for a given concentration of greenhouse gases.

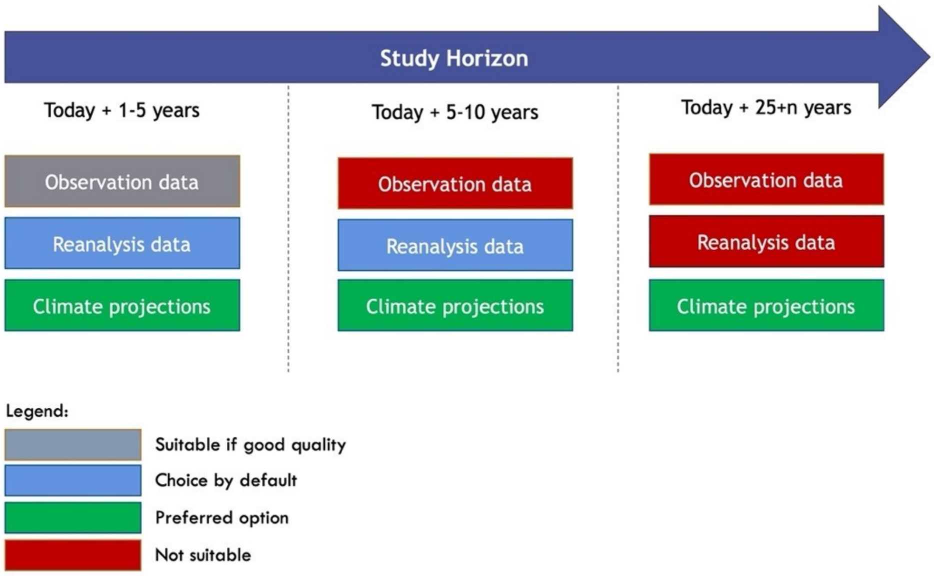

Fig. 1 General applicability of climate dataset type for different study horizons in Resource Adequacy Assessments. Depending on the type of study, different choices can be made.

2. Climate Change and Climate dependent Energy modelling

Usage of historically observed or reconstructed climate conditions

The most frequent approach so far has been to use in-situ historical records of climate information to represent the future (i.e., direct site-based meteorological observations). This climatological approach has the advantage that it provides an accurate view of the observed events. However, observations are not available everywhere and for long periods, and this becomes a real issue when considering RES generation from multiple locations.

The best solution to overcome this gap is to consider climate reanalysis that provide comprehensive datasets both in space and time, over long periods. State-of-the-art reanalysis like the ECMWF ERA5 have a spatial (ca 30 km) and temporal (1 hourly) resolution that makes them good candidates for such applications (Craig et al. 2022). As model-based and gridded products they do, however, come with the potential for biases and deficiencies that hinder their ability to estimate the actual observations which would have been taken at specific point locations. While ERA-5 has sufficient resolution to properly evaluate solar, and wind resources properly, at 30km resolution convection is still implicitly parameterized and the mesoscale phenomena driving many regional wind resources is not resolved at all(Kalverla et al. 2019; Sharp and Mass 2002). Still, they provide information everywhere and at all time steps, thus overcoming station observation data gaps. Reanalysis are frequently updated and extended, with improvements in the different components: the climate model itself, its spatial and temporal resolution, and the type, number, and quality of the assimilated observations.

However, actual observations and reanalysis represent only the past climate. Such historical information is valid only if the climate is stationary and provided that the observational sample is sufficiently large to estimate the relevant modes of climate variability appropriately. Under a changing climate, past data can no longer be regarded as representative of the climate in the future. In addition, observations and reanalysis have a limited length. For instance, ERA5 spans the period from 1950 to the present, but the quality of the reconstruction before the 1980s comes with higher uncertainty due to fewer observations assimilated into the model. Power systems studies need to consider extreme events that might not be sufficiently well sampled with 40 to 70 years of data.

Usage of detrended historical climate conditions

One solution to consider the climate evolution observed over the past decades is to evaluate the trends and extrapolate these to future years. Such an approach offers the advantage of getting data more representative of the recently observed changes (Harang, Heymann, and Stoop 2020). However, it assumes that the trends will remain the same in the future, which might not be true, depending on the extrapolation method and period considered (Liebmann et al. 2010; Rojas et al. 2013; Karin van der Wiel et al. 2019). The associated methods are also generally not dealing with multivariate trends calculation and do not easily allow to extrapolate the different variables in a physically sound way.

Trend extrapolation methods then sound acceptable for medium-term projections (for the next 5 to 10 years) but are certainly not optimal for longer-term extrapolation (Craig et al. 2022; Harang, Heymann, and Stoop 2020; Wohland, Omrani, Witthaut, et al. 2019), especially when considering multi-decadal variability (Wohland, Omrani, Keenlyside, et al. 2019). In addition, the changes in circulation patterns (Coumou, Lehmann, and Beckmann 2015) and climate change induced shifts in weather regimes over certain regions are not considered (Rojas et al. 2013). This severely limits the consideration of changes in risk due to the changing climate.

Usage of Climate projections

Climate projections, based on physical models, offer the option to have long-term information, where the physics between the different variables is explicitly modelled. In particular, high-resolution regional climate models offer a higher spatial resolution at the cost of. Multi-model exercises like the EURO-CORDEX experiment (Jacob et al. 2014), provide different models so that models' uncertainties can be estimated to some extent.

These models are run over past decades, which allows to validate them against reanalysis or observations and possibly adjust for their biases. They also provide continuous projections into the future, generally until the end of the 21st century, for different greenhouse gas emissions scenarios. While climate model output is perhaps not critical for medium-term studies (5-10 years ahead), it becomes essential for longer-term projections where the effects of forced anthropogenic climate change become more significant. With the new targets of the EU towards decarbonization in 2050, this kind of studies are becoming more and more frequent and vital, and relevant climate datasets are necessary for these.

Conversion models

Energy systems models do not use climate variables directly, but rather their transformation into energy-relevant variables: power consumption and generation, mainly for wind, solar and hydropower, through energy conversion models.

Two modelling methodologies can be considered when converting climate to energy variables. Statistical models (Akpinar and Akpinar 2005a; Dubus et al. 2021a; de Felice 2020; de Felice, Alessandri, and Ruti 2013; Ho et al. 2020; Pierro et al. 2017) use historical climate and generation data to make assumptions and build the relationship between climate and energy variables. Fully statistical models effectively capture historical relationships but are very dependent on the input data used (de Felice 2020). The assumption that past statistical relationships will still be valid in the future makes these models not very flexible (Harang, Heymann, and Stoop 2020). Physical and empirical models make assumptions about the technological relationship between climate and energy variables based on physical and empirical laws (Dubus et al. 2021; Y.-M. Saint-Drenan et al. 2018; Y. M. Saint-Drenan et al. 2020). Different technological assumptions can thus be made by adapting the parameters in these models accordingly.

REFERENCES

Akpinar, E. Kavak, and S. Akpinar. 2005a. "A Statistical Analysis of Wind Speed Data Used in Installation of Wind Energy Conversion Systems." *Energy Conversion and Management* 46(4): 515–32.

———. 2005b. "An Assessment on Seasonal Analysis of Wind Energy Characteristics and Wind Turbine Characteristics." *Energy Conversion and Management* 46(11–12): 1848–67.

Chenni, R., M. Makhlouf, T. Kerbache, and A. Bouzid. 2007. "A Detailed Modeling Method for Photovoltaic Cells." *Energy* 32(9): 1724–30.

Compo, G. P. et al. 2011. "The Twentieth Century Reanalysis Project." *Quarterly Journal of the Royal Meteorological Society* 137(654): 1–28.

Cornes, Richard C., Gerard van der Schrier, Else J.M. van den Besselaar, and Philip D. Jones. 2018. "An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets." *Journal of Geophysical Research: Atmospheres* 123(17): 9391–9409.

Coumou, Dim, Jascha Lehmann, and Johanna Beckmann. 2015. "The Weakening Summer Circulation in the Northern Hemisphere Mid-Latitudes." *Science* 348(6232): 324–27.

Craig, Michael T. et al. 2022. "Overcoming the Disconnect between Energy System and Climate Modeling." *Joule*: 1–13. http://arxiv.org/abs/2205.15636%0Ahttp://dx.doi.org/10.1016/j.joule.2022.05.010.

Dubus, Laurent et al. 2021a. "C3S Energy: An Operational Service to Deliver Power Demand and Supply for Different Electricity Sources, Time and Spatial Scales over Europe." http://doi.org/10.31223/X5MM06.

———. 2021b. "C3S Energy: An Operational Service to Deliver Power Demand and Supply for Different Electricity Sources, Time and Spatial Scales over Europe."

de Felice, Matteo. 2020. "Hydropower Information for Power System Modelling : The JRC-EFAS-Hydropower Dataset."

de Felice, Matteo, Andrea Alessandri, and Paolo M. Ruti. 2013. "Electricity Demand Forecasting over Italy: Potential Benefits Using Numerical Weather Prediction Models." *Electric Power Systems Research* 104: 71–79. http://dx.doi.org/10.1016/j.epsr.2013.06.004.

Gelaro, Ronald et al. 2017. "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)." *Journal of Climate* 30(14): 5419–54.

Harada, Yayoi et al. 2016. "The JRA-55 Reanalysis: Representation of Atmospheric Circulation and Climate Variability." *Journal of the Meteorological Society of Japan. Ser. II* 94(3): 269–302.

Harang, Inès, Fabian Heymann, and Laurens P. Stoop. 2020. "Incorporating Climate Change Effects into the European Power System Adequacy Assessment Using a Post-Processing Method." *Sustainable Energy, Grids and Networks* 24.

Hersbach, Hans et al. 2015. "ERA-20CM: A Twentieth-Century Atmospheric Model Ensemble." *Quarterly Journal of the Royal Meteorological Society* 141(691): 2350–75.

———. 2020. "The ERA5 Global Reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146(730).

Ho, Linh T.T., Laurent Dubus, Matteo de Felice, and Alberto Troccoli. 2020. "Reconstruction of Multidecadal Country-Aggregated Hydro Power Generation in Europe Based on a Random Forest Model." *Energies*.

Jacob, Daniela et al. 2014. "EURO-CORDEX: New High-Resolution Climate Change Projections for European Impact Research." *Regional Environmental Change* 14(2): 563–78.

Kalverla, Peter C., James B. Duncan, Gert Jan Steeneveld, and Albert A.M. Holtslag. 2019. "Low-Level Jets over the North Sea Based on ERA5 and Observations: Together They Do Better." *Wind Energy Science* 4(2): 193–209.

Kobayashi, Shinya et al. 2015. "The JRA-55 Reanalysis: General Specifications and Basic Characteristics." *Journal of the Meteorological Society of Japan. Ser. II* 93(1): 5–48.

Laloyaux, Patrick et al. 2018. "CERA-20C: A Coupled Reanalysis of the Twentieth Century." *Journal of Advances in Modeling Earth Systems* 10(5): 1172–95.

Liebmann, Brant et al. 2010. "Influence of Choice of Time Period on Global Surface Temperature Trend Estimates." *Bulletin of the American Meteorological Society* 91(11): 1485–91.

Pierro, Marco et al. 2017. "Data-Driven Upscaling Methods for Regional Photovoltaic Power Estimation and Forecast Using Satellite and Numerical Weather Prediction Data." *Solar Energy* 158(May): 1026–38. https://doi.org/10.1016/j.solener.2017.09.068.

Poli, Paul et al. 2016. "ERA-20C: An Atmospheric Reanalysis of the Twentieth Century." *Journal of Climate* 29(11): 4083–97.

Rojas, M. et al. 2013. "Winter Weather Regimes over the Mediterranean Region: Their Role for the Regional Climate and Projected Changes in the Twenty-First Century." *Climate Dynamics* 41(3–4): 551–71.

Saint-Drenan, Yves Marie et al. 2020. "A Parametric Model for Wind Turbine Power Curves Incorporating Environmental Conditions." *Renewable Energy* 157(May): 754–68.

Saint-Drenan, Yves-Marie et al. 2018. "An Approach for the Estimation of the Aggregated Photovoltaic Power Generated in Several European Countries from Meteorological Data." *Advances in Science and Research*.

Sharp, Justin, and Cliff Mass. 2002. "American Meteorological Society COLUMBIA GORGE GAP FLOW: Insights from Observational Analysis and Ultra-High-Resolution Simulation." *Source: Bulletin of the American Meteorological Society* 83(12): 1757–62.

van der Wiel, K. et al. 2019. "Meteorological Conditions Leading to Extreme Low Variable Renewable Energy Production and Extreme High Energy Shortfall." *Renewable and Sustainable Energy Reviews* 111: 261–75.

van der Wiel, Karin et al. 2019. "The Influence of Weather Regimes on European Renewable Energy Production and Demand." *Environmental Research Letters*.

WMO. 2007. "The Role of Climatological Normals in a Changing Climate." *World Climate Data and Monitoring Programme* WCDMP-No.(WMO-TD No. 1377).

———. 2017. *WMO Guidelines on the Calculation of Climate Normals, 2017 Edition*. Geneva.

Wohland, Jan, Nour Eddine Omrani, Noel Keenlyside, and Dirk Witthaut. 2019. "Significant Multidecadal Variability in German Wind Energy Generation." *Wind Energy Science* 4(3): 515–26.

Wohland, Jan, Nour Eddine Omrani, Noel Keenlyside, and Dirk Witthaut. 2019. "Significant Multidecadal Variability in German Wind Energy Generation." *Wind Energy Science* 4(3): 515–26.

Wohland, Jan, Nour Eddine Omrani, Dirk Witthaut, and Noel S. Keenlyside. 2019. "Inconsistent Wind Speed Trends in Current Twentieth Century Reanalyses." *Journal of Geophysical Research: Atmospheres* 124(4): 1931–40. https://onlinelibrary.wiley.com/doi/full/10.1029/2018JD030083 (April 22, 2022).

Wohland, Jan, Mark Reyers, Juliane Weber, and Dirk Witthaut. 2017. "More Homogeneous Wind Conditions under Strong Climate Change Decrease the Potential for Inter-State Balancing of Electricity in Europe." *Earth System Dynamics* 8(4): 1047–60.

World Meteorological Organization. 2018. I Guide to instruments and methods of observation *Volume I – Measurement of Meteorological Variables*. https://library.wmo.int/index.php?lvl=notice_display&id=12407%0Ahttps://library.wmo.int/index.php?lvl=notice_display&id=12407#.YkdSz3XMLio.