Article

Improving Energy Transition Analysis Tool through Hydropower Statistical Modelling

Cristobal Gallego-Castillo [1,*] **and Marta Victoria** [2,3]

1   Aircraft and Space Vehicles Department, Universidad Politécnica de Madrid, Plaza Cardenal Cisneros 3, 28040 Madrid, Spain
2   Department of Engineering, Aarhus University, Inge Lehmanns Gade 10, 8000 Aarhus, Denmark; mvp@eng.au.dk
3   iCLIMATE Interdisciplinary Centre for Climate Change, Aarhus University, 8000 Aarhus, Denmark
*   Correspondence: cristobaljose.gallego@upm.es

Abstract: Reservoir and pumped hydro storage facilities represent one of the best options for providing flexibility at low marginal cost and very low life cycle carbon emissions. However, hydropower generation is subject to physical, environmental and regulatory constraints, which introduce complexity in the modelling of hydropower in the context of transition energy analysis. In this article, a probabilistic model for hydropower generation is developed in order to improve an hourly-resolved tool for transition path analysis presented in previous research. The model is based on time series analysis, which exploits the fact that the different constraints affecting hydropower generation were met in the past. The upgraded version of the transition path analysis tool shows a decrease in the hydropower flexibility as compared with previous published results, providing a better picture of the benefits and drawbacks associated with a specific transition path under analysis, for example in terms of assessing the probability of unserved energy. The upgraded version of the tool was employed to analyse the Spanish National Energy and Climate Plan (NECP), finding consistence between proposals associated with the power system and related CO_2 reduction and share of renewable electricity targets.

Keywords: energy modelling; hydropower; renewable grid integration; time series models; energy transition; Spain

1. Introduction

The Paris Agreement [1] implies that carbon neutrality can be achieved by 2050. Since energy conversion is mainly responsible for anthropogenic greenhouse gases (GHG) emissions, plans for energy transition are being developed worldwide. The focus in the short-term (2020–2030) is placed mainly on the electricity sector, as several renewable technologies (PV and wind power) are more competitive than conventional technologies, and massive deployment is expected in the following years. In the framework of the European Union (EU) energy policy, the so-called Clean Energy for All Europeans package, Spain has recently submitted to the European Commission its National Energy and Climate Plan (NECP) [2]. According to this plan, renewables must cover 74% of electricity by 2030 (this percentage was around 40% in the last years [3]).

However, the integration of large amounts of non-dispatchable renewable generation represents a challenge, as power systems are constrained by the fact that supply and demand must be balanced at every moment. While flexibility provided by conventional generation may contribute to wind and PV integration during initial stages of the transition, additional sources of flexibility will be required as the non-dispatchable penetration increases and lower carbon emissions are required.

Reservoir and pumped hydro storage facilities represent one of the best options for providing both flexibility and storage at low marginal cost and very low life cycle carbon

emissions. In addition, according to a recent study in the EU, among the different storage technologies, only hydropower reservoirs and, to a small extent, batteries will be relevant by 2030 [4]. Indeed, hydropower plants are able to change rapidly the power output at will, since no thermal processes are involved. However, the constraints associated with hydropower generation exist, and they are of very different nature. For example, the fact that several hydropower facilities are placed on the same river or its tributaries means that the generation of the facility upstream places conditions on the rest of the hydropower generators downstream due to the distance between facilities. When modelling the aggregated hydropower generation in a power system, such constraints could be included by incorporating physical information of the water systems [5,6], but this information is often scarce or incomplete. Other limitations might be even more difficult to account for, like environmental and regulatory constraints [7]. Environmental constraints aim at reducing the impacts of hydropower facilities in the local environment. These impacts are mainly due to the presence of dams, as they affect water quality, aquatic species and the natural water flow. Regulatory constraints arise from the rights and needs of different participants of the water systems downstream of hydropower facilities, including farms, cities and industries. They also reflect other goals of the water use different from power generation, like recreational use, conservation of protected areas and flood control.

As a consequence of technical, environmental and regulatory constraints, hydropower facilities may be subject to limitations on the amount of water that can be released, the reservoir level and flow rate variations, affecting eventually to hydropower generation. According to Reference [7], the difficulty of modelling these constraints usually leads to inaccurate estimates of the ability of hydropower facilities to provide storage and flexibility. An example of including environmental and regulatory constraints in an hourly-resolved optimal provision of hydropower was described in Reference [8], but the small scale of the considered case study (a single facility of 150 MW) makes this approach rather limited for modelling the aggregated hydropower generation within a whole power system.

Assessing the viability of a power system in the context of the energy transition requires detailed simulations in order to account for the variability in the different timescales of solar radiation, wind, rain patterns and electricity demand, among others, together with considering the constraints of the available flexible generation and storage systems. In Reference [9], an hourly-resolved tool for transition path analysis was introduced, and employed for analysing energy transition paths in Spain up to 2030. The modelling of the hydropower technology was grounded on maximum capacity together with maximum and minimum reservoir levels, disregarding other constraints affecting hydropower time series dynamics. This led to a limited, and probably optimistic, representation of the hydropower flexibility.

In this work, a model of the hydropower generation fleet in peninsular Spain through time series probabilistic forecasting techniques is presented. The advantage of time series based models is that the dynamics of the considered variable are captured from historical data, meaning that prior knowledge on the underlying causes is not required. This represents an appropriate framework for incorporating technical, environmental and regulatory hydropower constraints, under the assumption that historical data are representative of hydropower dynamics subject to such constraints.

The proposed hydropower statistical model is coupled with the dispatch algorithm for power systems with high renewable energy source (RES) penetration presented in Reference [9], providing an upgraded version of the transition path analysis tool. The results obtained are compared with those obtained in the previous work [9] with two main goals. The first one is the analysis of the influence of the proposed model on the hydropower flexibility, with special emphasis on the observed hydropower hourly gradients. The second goal is to analyse the extent to which more realistic hydropower dynamics impact the transition path results, in terms of emissions, demand supplied with RES, security of supply, and so forth.

A third additional objective is included in this research, which consists in the analysis of the Spanish power system roadmap defined in the NECP [2] based on the upgraded version of the tool. By doing this, the potential bridge between NECP proposals associated with the power system and related CO_2 reduction and share of renewable electricity targets is explored.

The rest of the article is structured as follows. Section 2 provides a brief description of the tool introduced in Reference [9], and summarises the main hypotheses of the model. The proposed methodology for modelling of the hydropower generation in peninsular Spain is introduced in Section 3, and the procedure to obtain feasible hydropower values for an hour coherent with the dynamics observed for this variable in the past is described. In Section 4, the introduced hydropower model is employed to obtain a new set of results for the main transition path described in Reference [9]. Finally, the upgraded version of the tool (equipped with the hydropower model) is employed in Section 5 in order to analyse the transition path of the Spanish power system described in the Spanish NECP for period 2020–2030. The main conclusions of the work are described in Section 6.

2. Review of the Model Basis

This section summarises the main aspects of the tool for transition path analysis introduced in Reference [9]. The model uses an hourly dispatch algorithm based on heuristic rules that prioritise the use of renewable technologies. The capacities of the different technologies are fixed exogenously and follow different evolution determined by the transition path under analysis. The dispatch in every hour is determined in subsequent steps without assuming any knowledge on future hours. The methodology is divided into the following four steps:

- **Transition path definition.** This comprises the assumption of capacities of the different technologies for every year of the considered period. Of special relevance for a transition path are the assumptions of wind and PV installation rate and the decommissioning of coal and nuclear plants. The transition path definition also includes the assumption on the electricity demand for every year of the period.

- **Definition of the hourly capacity factors.** A total of 900 combinations of hourly capacity factors for a number of variables were built from historical data. The hourly capacity factor is defined as the ratio between the available generation and installed capacity in every hour. These variables are assumed to be imposed by weather and technological factors, and include non-dispatchable generation together with electricity demand and hydro inflow. Non-dispatchable generation comprises wind power, PV, small hydropower, cogeneration, nuclear, coal and Concentrated Solar Power (CSP) generation. Some of these technologies have a certain capacity for regulating its generation level. However, they are considered non-dispatchable in this work for convenience. Flexible generation, that is, hydro, biomass and gas, are not included, as the hourly generation levels of these technologies are determined by the dispatch algorithm, see below, in order to balance generation and demand. The reason behind this high number of hourly capacity factor combinations is to consider a broad range of situations during hourly dispatch simulation. By doing this, results are less affected by the use of historical time series, and a statistical characterisation of the performance indicators is obtained.

- **Hourly dispatch algorithm.** Every year within a transition path is solved for the 900 combinations of hourly capacity factors. The dispatch algorithm determines how much electricity demand is covered by non-dispatchable technologies in order to assess the needs for flexible generation. Flexible technologies are dispatched in the following order: hydropower generation, biomass and gas power plants. Situations with renewable energy excess lead to store electricity through pumped hydro storage and, eventually, to curtailments. Maximum generation (pumping) power for hydropower is constrained by the installed generation (pumping) capacity and by the historical maximum and minimum filling level of the hydropower reservoir. Accord-

ing to this algorithm, the aggregated hydropower generation can be decided by the algorithm every hour with a limited number of constraints, given by the hydropower installed capacity and the existing water reservoirs at the moment, with the objective of covering the demand before using polluting technologies (biomass and gas). This approach might result in hydropower generation time series that overestimate the hydropower flexibility. For this reason, additional constraints that reflect those mentioned in Section 1 are required. This is the goal of the methodology introduced in Section 3.

- **Performance indicators.** The outputs of the dispatch algorithm for the 900 combinations of time series and for every year of the considered period are employed to generate relevant outputs and performance indicators (CO_2 emissions, share of renewable electricity, etc.). One of the main indicators is the hourly coverage ratio, defined as the ratio between the total energy that could potentially be generated at a a certain hour and the aggregate electricity demand at the same hour. Critical h are defined as those for which an hourly coverage ratio between 1 and 1.2 is observed (this means that the total energy that could potentially be generated that hour is between 0% and 20% higher than the electricity demand at this hour). Values of this ratio below 1 are indicative of a non-zero probability of unserved energy.

Three transition paths were analysed in Reference [9]. The main one, referred to as transition path *A*, considered an ambitious but realistic renewable capacity deployment (annual installation rates of 4 GW/year for PV and 2 GW/year for wind power), a reasonable level of efficiency (leading to a +0.5% increase of the electricity demand every year) and the full decommissioning of nuclear and coal before 2025. Results showed that this transition path was feasible (zero hours with unserved energy), and would led to noticeable reductions of CO_2 emissions (more than 50% of reduction by 2030) and high renewable penetration levels in electricity generation (~80% by 2030). Two additional transition paths, referred to as *A2* and *B*, analysed the impacts of considering a slower nuclear phase-out and a higher installed capacity for dispatchable RES (biomass), respectively.

3. Statistical Model for Hydropower Generation

Time series models aim at explaining the dynamics of a certain variable of interest in the absence of knowledge about the underlying governing laws of the considered system. This approach is useful when these governing laws do not allow for a theoretical formulation, as in sociology, econometrics or any other field involving human behaviour, or they are too complicated to handle, for example, complex physical processes as in some problems related to biology or atmospheric dynamics, among others.

In a nutshell, time series modelling is grounded on the idea that the evolution in time of a certain variable y is due to two different contributions, one is deterministic and can be computed from available data, and the other is random. Mathematically,

$$y_{t+1} = f(\mathbf{x}_t|\mathbf{\Theta}) + \varepsilon_{t+1}, \tag{1}$$

where function $f(\cdot)$ represents the deterministic component and ε_t is the random component. $\mathbf{x}_t$ represents a set of available data at time t, including observations of the considered variable $[y_t, y_{t-1}, ...]$, but also information from any other relevant variable. $\mathbf{\Theta}$ is a vector containing the model parameters. The key of time series modelling is the implementation of appropriate deterministic functions $f(\cdot)$, including appropriate choices for input variables and the determination of the model parameters. This can be done through several criteria, such as maximum likelihood or minimum prediction error.

Forecasts provided by time series models can be in the form of point forecasts (the prediction consists of the expected value) or, more interestingly, probabilistic forecasts (predictive density functions or quantile forecasts). One of the most common approaches in probabilistic forecasting consists in dressing a point forecast of the variable of interest with predictive densities of the forecast error, obtained in similar conditions [10]. Indeed, even

though ε_t is usually assumed to be a sequence of independent Gaussian random variables, in most cases the gap between reality and the use of manageable and parsimonious models in $f(\cdot)$ lead to forecast error distributions showing structure and dependencies. This information can be captured and exploited within a probabilistic framework in order to better describe the dynamics of the considered system.

In this work, the hydropower generation in Spain is modelled in a probabilistic framework in order to describe the range of flexibility provided by this source when accommodating non-dispatchable sources in a context of energy transition. To this end, hourly time series for period 2010–2014 for the following variables were employed:

- h_t [GW], the aggregated hydropower generation in peninsular Spain.
- θ_t [GW], the thermal gap, defined as the demand that needs to be covered by dispatchable sources. In particular, in this work the thermal gap is defined as the electricity demand at time t minus nuclear generation, wind generation, solar generation and electricity supplied by cogeneration and waste incineration.)
- σ_t [TWh], the filling level of the aggregated capacity of hydropower reservoirs in peninsular Spain. According to the Spanish TSO, the total storage capacity for all the reservoirs in Spain is 18.5 TWh (https://www.ree.es/es/datos/publicaciones/series-estadisticas-nacionales).

In order to model the feasible values of the hydropower generation for one hour ahead, a two-staged approach is followed. First, a point-forecast model is implemented in order to provide the expected value of the hydropower generation for one hour ahead. This model accounts for the autocorrelation of the hydropower generation, resulting from the diverse constraints operating in the hydropower system. Second, the prediction errors obtained with the point-forecasting model are employed to characterise the flexibility of the hydropower generation observed in the past. This is done by analysing the prediction error distributions, and their dependencies with relevant variables. The underlying assumption is that observed forecast errors represent feasible departs of the hydropower generation with respect to the expected value provided by the inertia of the system.

As a result, the probabilistic model is able to provide a predictive density function of the hydropower generation at every time step, reflecting the range of the feasible hydropower generation values. This range will be introduced later in the simulation algorithm of energy transition paths in order to limit the hydropower generation that is determined by the dispatch algorithm every hour.

3.1. Modelling the Expected Hydropower Generation through a Point-Forecast Model

Two families of models are considered, linear Autoregressive (AR) models and Varying Coefficient Models (VCMs). Both models require the original time series to be stationary, which is the case, since hydro generation is a two-sides bounded variable (the minimum and the maximum values being zero and the aggregated installed capacity, respectively). AR models assume linear relationships between consecutive values in a time series. Despite their simplicity, AR models are widely employed due to their solid theoretical foundations, see for example, References [11,12], among others. Mathematically, an AR model of order p, AR(p), is expressed as follows:

$$\text{AR}: \quad h_{t+1} = \sum_{i=1}^{p} \varphi_i\, h_{t-i+1} + \varepsilon_{t+1}, \tag{2}$$

where φ_i are the model parameters, to be determined from historical data. The following version of the AR model including exogenous variables (such as the thermal gap and the filling level of the reservoirs), usually referred to as ARX models, will be also considered:

$$\text{ARX}: \quad h_{t+1} = \sum_{i=1}^{p} \varphi_i\, h_{t-i+1} + \varphi_\theta\, \theta_{t+1} + \varphi_\sigma\, \sigma_{t+1} + \varepsilon_{t+1}. \tag{3}$$

The model given in (3) assumes that the hydropower generation is a function of its recent past, the current thermal gap, θ_{t+1}, and the hydropower reservoir filling level, σ_{t+1}.

Linear models can be upgraded to account for non-linear dynamics in different ways, usually at the expense of the model simplicity. A good trade-off between non-linear capabilities and model simplicity is provided by Varying Coefficient Models [13], which generalise AR models by implementing smooth dependencies of the model coefficients with other variables. By doing so, model coefficients φ_i become smooth functions, $\varphi(\cdot)$. These dependencies usually provide insights and interpretability on the captured dynamics, which is a rare characteristic in non-linear models. Several methods can be followed to obtain the model coefficients $\varphi(\cdot)$, all of them being computationally intensive: polynomial spline [14], smoothing spline [15] and kernel-local polynomial smoothing ([16] and references therein). The main drawback of the polynomial spline approach is that the regression performed to obtain the coefficient functions are likely to be quite sensitive to the position of the knots between splines, specially when a reduced number of knots is considered. Concerning the smoothing spline approach, the splines are fitted by minimising a penalized least squares criterion. In this case, the choice for the smoothing parameters involved in the process is somehow arbitrary, while it usually has a noticeable impact on the obtained coefficient functions. For these reasons, the kernel-local polynomial smoothing approach will be employed in this work. For further details on these estimation methods, the reader is referred to Reference [17].

Two different dependencies in the framework of VCMs will be tested:

$$\text{VCM}(\theta): \quad h_{t+1} = \sum_{i=1}^{p} \varphi_i(\theta_{t+1}) \, h_{t-i+1} + \varepsilon_{t+1}, \tag{4}$$

and

$$\text{VCM}(\sigma): \quad h_{t+1} = \sum_{i=1}^{p} \varphi_i(\sigma_{t+1}) \, h_{t-i+1} + \varepsilon_{t+1}. \tag{5}$$

The model parameters were obtained from a training period comprising years 2010–2012, and evaluated for the test period (years 2013–2014). The evaluation is based on the Root Mean Square Error (RMSE) between observations, h_t, and forecasts, $\hat{h}_t$:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} (h_t - \hat{h}_t)^2} = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \varepsilon_t^2}, \tag{6}$$

where the forecast, $\hat{h}_t$, is the expectation of the considered variable, $\hat{h}_t = \mathbb{E}(h_t)$. Since ε_t is assumed to be zero-mean gaussian noise, forecasts are computed directly from the deterministic component in Equation (1).

Figure 1 shows the root mean square error (RMSE) for the different AR models considered, for model orders from $p = 1$ to $p = 28$ to cover the time scale of one day (left), and for model orders up to $p = 200$ to cover the time scale of one week (right).

Results in the daily scale (left) show negligible differences as compared with the important error reductions experienced at $p = 3$ and $p = 26$ achieved by all the models. This reveals two important regression time scales when accounting for hydropower generation dynamics. For this reason, VCMs were implemented for these regression model orders (results are also shown in Figure 1). However, VCMs only outperform all the AR models in the case of $p = 26$ and for the case of conditioning model coefficients to reservoir filling level. In any case, differences between models are negligible compared to the importance of including relevant time scales within the regression order.
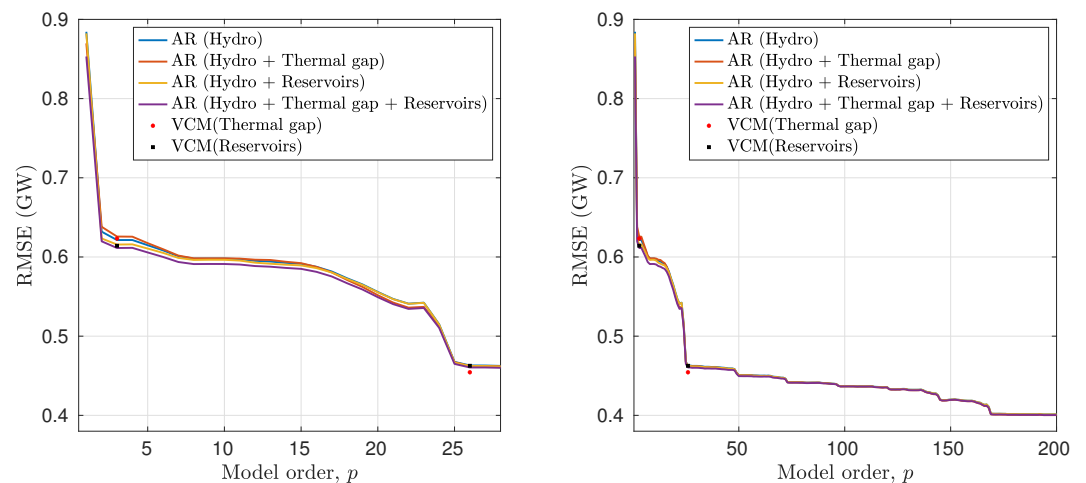
Figure 1. Root Mean Squared Error (RMSE) of the different point-forecasting models, as a function of the autoregression order, p.

Based on this, the analysis was extended to a week time scale. That was done only for the AR models, since the computational requirements for modelling hundred of coefficient-functions required for VCMs are very high, and probably not justified given the improvements obtained in the daily scale with respect to AR models. Figure 1 (right) shows that there exists a negligible difference in performance when considering exogenous variables. However, what becomes important is to consider a model order of $p = 169$ (one week contains 168 h). From this order on, no further noticeable improvement is obtained.

The performed analysis suggests that the aggregated hydropower generation in peninsular Spain in the considered period was mainly driven by the dynamics observed during the last week, while thermal gap and hydropower filling level contributed little to the level of hydropower generation. In view of these results, the selected point-forecast model is an AR(169) model, without exogenous variables.

3.2. Modelling the Prediction Error through a Probabilistic Model

An analysis of the predictive error obtained with the point-forecasting model described above reveals a certain structure in ε_{t+1}, as well as dependencies with exogenous variables. Figure 2 (top) shows the statistical distribution of ε_{t+1} conditioned to three levels of the predicted hydropower generation, $\hat{h}_{t+1}$. It can be seen that in situations where the point-forecast model predicts a negative aggregated hydropower generation (indicating that hydropower pumping overcomes hydropower generation) of -1 GW, the distribution of the related predictive errors shows less uncertainty (i.e., less variance) than that observed for situations in which the predicted hydropower generation is of 5 and 10 GW. Also, a different bias can be appraise in each situation.

A similar conclusion is obtained from Figure 2 (middle), in which the error predictive densities were conditioned to three different levels of thermal gap, θ_{t+1}. In this case, the difference in bias is even more prominent, indicating a positive bias for large thermal gaps.

Figure 2 (bottom) shows the statistical distribution of ε_{t+1} conditioned to three different values for the filling level, σ_{t+1}, reflecting different levels of uncertainty depending on this variable.
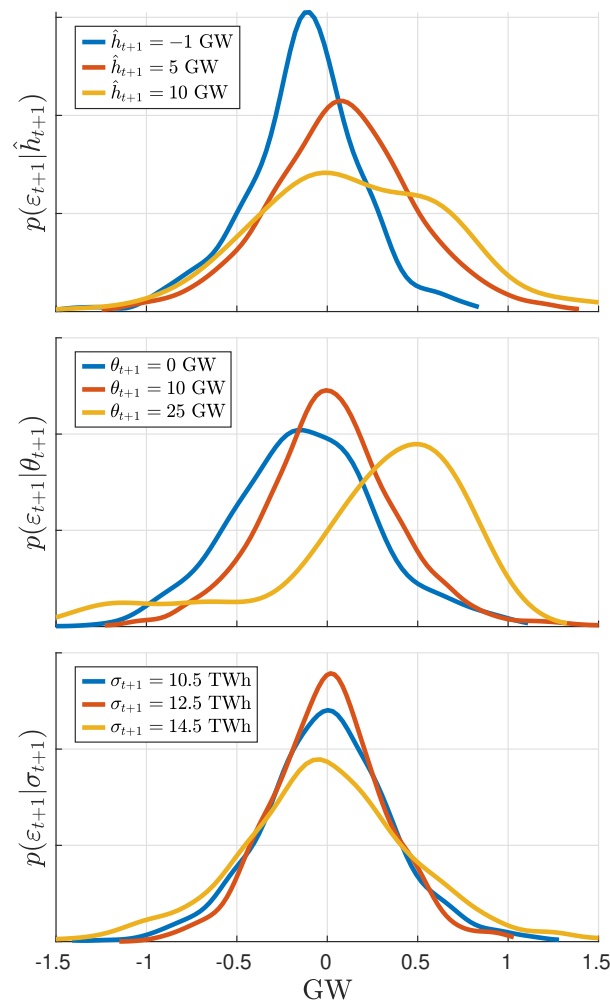
Figure 2. Statistical distribution of ε_{t+1} conditioned to three levels of: (**top**) predicted hydropower generation, $\hat{h}_{t+1}$; (**middle**) thermal gap, θ_{t+1}; (**bottom**) reservoir filling level, σ_{t+1}.

In view of the previous analysis, a probabilistic model based on the generation of conditioned predictive densities of the forecast error was built.

3.3. A Probabilistic Model for the Hydropower Generation

By combining the point-forecasting model and the error probabilistic model described in the previous sections, the probabilistic model for the hydropower generation is obtained. Mathematically, the predictive density of the hydropower generation estimated at time t for time $t + 1$ is given by:

$$p(h_{t+1}) = \hat{h}_{t+1} + p(\varepsilon_{t+1}|\hat{h}_{t+1}, \theta_{t+1}, \sigma_{t+1}). \tag{7}$$

Figure 3 shows an example of the hydropower generation up to time t, and the obtained predictive density for time $t + 1$, for a situation with $\sigma_{t+1} = 13.1$ TWh and a thermal gap of $\theta_{t+1} = 17.8$ GW. This figure shows that, according to the implemented model, the hydropower generation for the next time step is expected to be 3.5 GW. In addition, the conditional predictive density of the error forecast is informative about the uncertainty associated with the level of generation, in such a way that the probability for the hydropower generation is concentrated within the range from 2.5 to 5 GW.
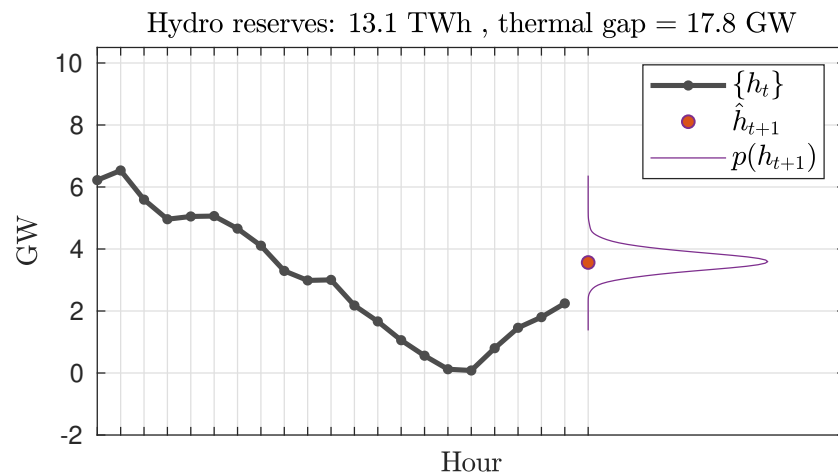
Figure 3. Example of the predictive density of the hydropower generation provided by the statistical model.

4. Revisiting Previous Results with the Upgraded Version of the Transition Path Analysis Tool

In this section, the results obtained for transition path *A* in Reference [9] are revisited and compared to those obtained with the upgraded version of the tool. The upgrade consists in the coupling of the probabilistic model described in Section 3 within the dispatch algorithm of the tool. The probabilistic model provides the dispatch algorithm with a range of feasible values for the hydropower generation at every time step. By *feasible* we mean that, from a statistical point of view, they are likely to happen based on past observations. This range is determined from the predictive density $p(h_{t+1})$ described in (7). In particular, percentiles P1 and P99 of the distribution are employed to define the minimum and maximum levels of hydropower generation for the related time instant. The set of results obtained for transition path *A* with the upgraded version of the model will be referred to as transition path *A+*.

Due to the nature of the implemented hydropower probabilistic model, it is expected that the new hydropower generation time series will show hourly gradients of lower magnitude. This follows as a consequence of providing a constrained range of feasible hydropower values to the dispatch algorithm. To illustrate this idea, Figure 4 shows 24 h of the mix generation for transition paths *A* and *A+*. The considered day belongs to year 2030, when large PV capacities have been installed, resulting in a critical ramp-down throughout the sunset. The flexibility required in such situations is one of the challenges for power systems with high PV penetration [18]. In both cases the figure shows that hydropower generation is storing water during the day, while there is an excess of PV generation. On the left, this situation is followed by a sharp increase in hydropower generation from 16 h to 19 h. During this period, gas generation is not required, conveying the idea that hydropower provides the required flexibility. This gradient in the hydropower generation is moderated in the plot on the right due to the constraints imposed by the probabilistic model, leading to the need for some gas generation during the PV ramp-down.

To delve more into this question, Figure 5 shows the histogram of the hourly hydropower gradient, normalised as a probability density function, for transition paths *A* and *A+* for years 2020, 2025 and 2030. To facilitate interpretability, a similar plot with historical data for 2015 is included.
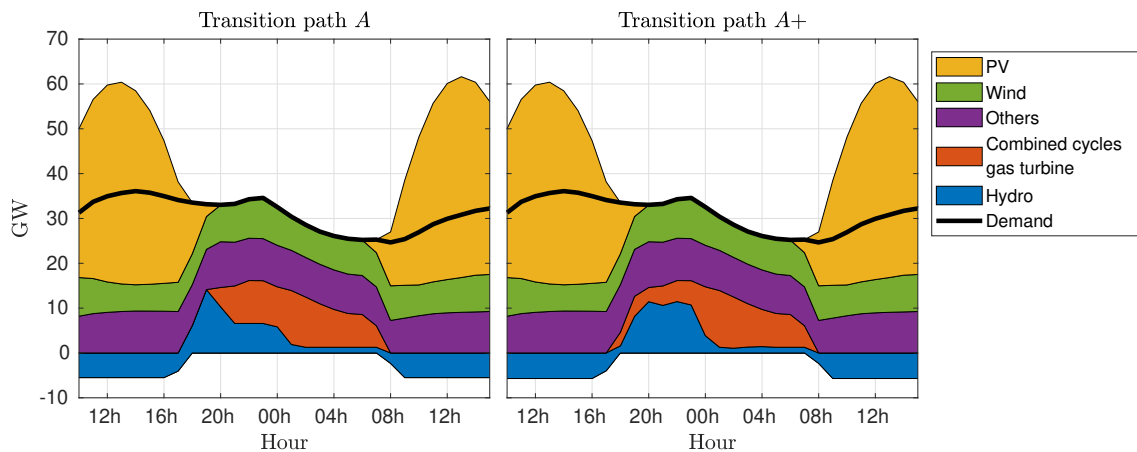
Figure 4. Power generation mix during a 24-h period in 2030 provided by the transition path analysis tool. **Left**: without considering the probabilistic model for hydropower generation (transition path *A*). **Right**: including the probabilistic model (transition path *A+*).
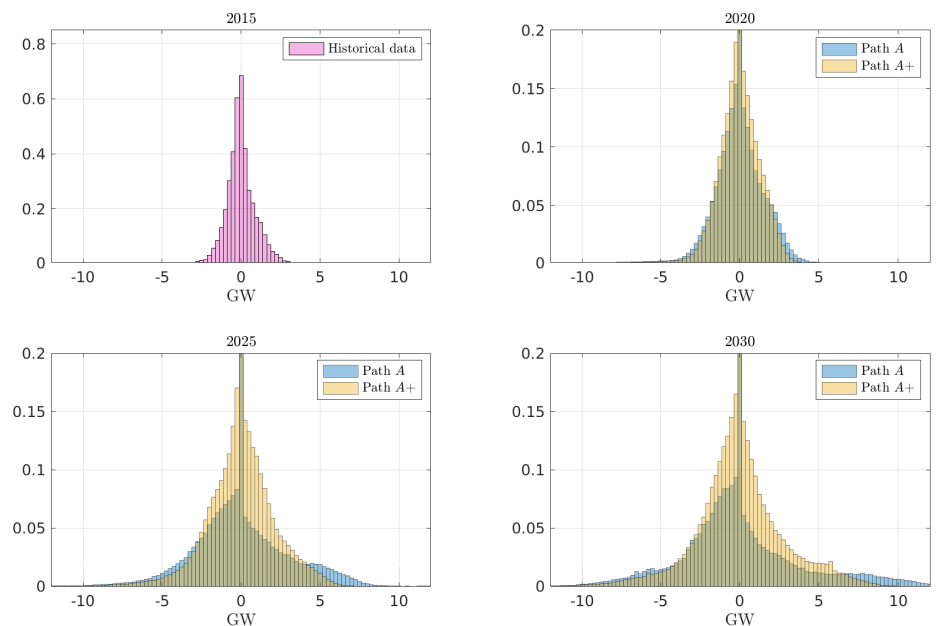

Figure 5. Histograms of hourly hydropower gradients normalised to probability density function. Top left: historical data (2015); rest of figures: results for transition paths *A* and *A+*.

In can be seen that hydropower gradients in the past were typically between ± 3 GW. The obtained simulations for 2020–2030 show the need for increasing the maximum hydropower gradient, up to $[-10, 12]$ GW by 2030. The reason for this is two-fold: (i) The increase in installed non-dispatchable RES technologies (wind and PV); (ii) the priority of the hydropower with respect to combined cycle gas turbines in the dispatch algorithm; this makes that maximum capacity of hydropower (provided by maximum installed power, water reservoirs availability and, in the upgraded version of the model, the range of feasible hydropower values provided by the statistical model) is employed to cover the demand before using gas.

Since transition path *A+* incorporates the impact of constraints on the hydropower dynamics, the modelled flexibility in the power system during the transition path is reduced. Consequently, it is expected that the performance of the transition path will

decrease. Figures 6 and 7 show the main performance indicators for transition paths *A* and *A+*.
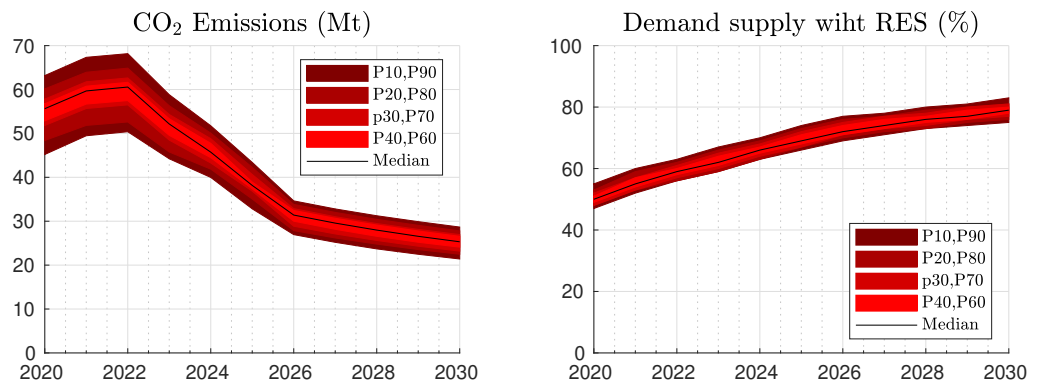

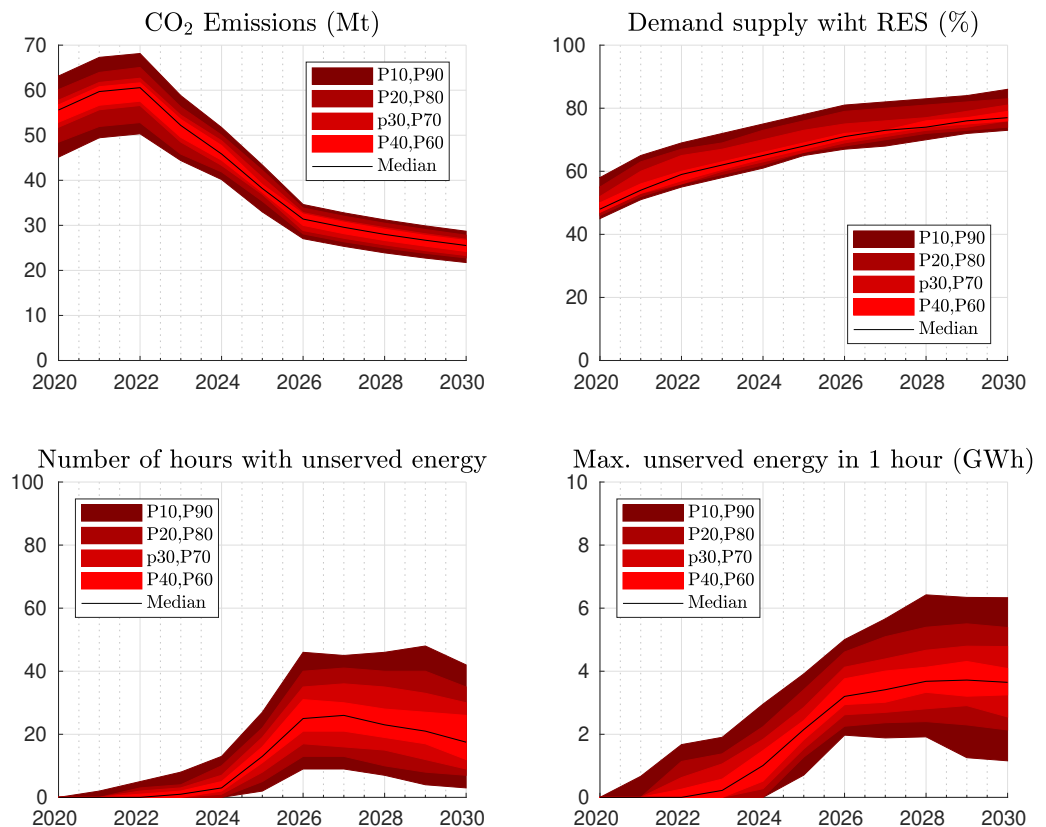**Figure 6.** Performance indicators for transition path *A*.

Figure 7. Performance indicators for transition path *A+*.

Figures 6 and 7 (top) illustrate the statistical distribution of the CO_2 emissions and the percentage of demand supplied with RES for every year. Both variables show similar levels in both transition paths, indicating that the hydropower modelling did not impact these figures noticeably. The main difference observed between transition paths *A* and *A+* are observed in terms of unserved energy. While there is no unserved energy in transition path *A*, Figure 7 (bottom left) reflects an increasing probability of having a number of hours every year with unserved energy in transition path *A+*. The number of hours with non-zero probability remains in all the period below 50 h per year, which represents less than 0.6% of the time. However, it is worth noting that the constraints in the hydropower dynamics introduced by the model led to such situations, and they deserve attention when analysing the proposed transition path. With this purpose in mind, the plot on the bottom right in Figure 7 was performed. This plot illustrates the probability distribution of the maximum

unserved energy in one hour observed for every year. This variable is important because it helps sizing the required additional resources for addressing situations with unserved energy. According to the plot, additional resources providing a maximum of ~6.5 GW would be required by 2028 to eliminate hours with non-zero probability of unserved energy. This power could be supplied in different forms: additional flexible generation capacity (i.e., combined cycles), storage (batteries), demand-side management, interconnections or a combination of them. Additional flexible generation should be considered with care, as it could involve additional emissions, and investments that could not be recovered. It is important to remark that batteries, demand-side management and interconnections are not considered in the current version of the tool, which represents current limitations to be addressed in future research. In particular, concerning the absence of interconnections in the tool implies that peninsular Spain is assumed to be electrically isolated. However, currently import capacities with neighbour countries are of 3 GW with France, 3.4 GW with Portugal and 0.6 GW with Morocco. Moreover, new interconnections with France will add 2 GW more by 2026–2027. Also, REE (the Spanish TSO), has implemented an interruptibility service, which affects mainly large-scale industry, and provides the possibility for rapid demand reductions in the order of 2–3 GW. With all this in mind, the conclusion is that the potential situations with unserved energy revealed by the upgraded version of the model in transition path *A+* could probably be solved with currently available mechanisms, and/or with little investments in additional flexibility mechanisms (batteries and demand-side management), which are likely to be available by the second half of the decade.

5. The Spanish Roadmap for 2030

The upgraded version of the model was employed to analyse the Spanish roadmap for 2030, gathered in the submitted National Energy and Climate Plan (NECP) [2]. A transition path referred to as *NECP* has been defined according to that document, and other relevant information. The main assumptions are:

- The installation rate for PV and wind power for peninsular Spain is of 2.7 GW/year and 2.1 GW/year, respectively, reaching 38.4 of PV and 48.5 GW of wind power by 2030 (data for peninsular Spain, see Reference [2], Annex D). Installed power capacity pumped hydro storage (PHS) increase linearly between 2019 (3.3 GW) and values detailed in Reference [2] for 2025 (4.2 GW) and 2030 (6.8 GW). Storage energy capacity increases accordingly, with 73 GWh in 2019, 92 GWh in 2025 and 150 GWh in 2030. No new installed capacity or storage capacity is considered for reservoir and mixed-storage hydroelectric power plants. Nuclear capacity follows the nuclear decommissioning plan recently defined by the Government in agreement with the owners of the nuclear facilities. This plan establishes the decommissioning of four reactors out of seven in the period 2020–2030, in particular, one reactor per year during 2027–2030. Concerning coal, a sharp decrease of installed capacity is expected in the first half of the period, with installed capacity in 2025 representing only 23% of that installed in 2019 (from 9215 MW to 2085 MW), and no coal plants operative by 2030 [2]. A linear decrease has been assumed in both periods 2019–2025 and 2025–2030.
- Yearly capacity factors for some renewable energy sources are expected to increase throughout 2020–2030, according to Reference [2]. In particular, full load hours of wind power scale up from 2200 to 2400 during 2025–2030 due to wind farms repowering and new offshore wind power.
- Electricity demand increases by 0.49% every year, reaching 263 TWh in 2030 [2].
- Hourly hydro inflow combines historical data employed in Reference [9] with corrections accounting for the expected impact of climate change. The corrections apply to both the annual hydro inflow and its seasonal profile. According to Reference [19], a reduction in the annual hydro inflow of 8% could be expected by 2030 in Spain. In Reference [20], hydro annual inflow reduction between −22% and −28% is predicted at the end of the century compared to todays values, for the RCP8.5 scenario. Slightly more severe are the reductions predicted in Reference [21], with an annual hydro

inflow change in Spain between −20% and −40% by 2050. With this in mind, in this work a linear reduction of the annual hydro inflow is assumed for the period 2020–2030, from −5% of reduction for 2020 to −10% for 2030 with respect to historical data gathered from 1991. Concerning the seasonal profile of hydropower inflow decrease, Reference [20] estimated that main decreases would take place during spring and autumn. Based on that work, a seasonal pattern of hourly hydro inflow decrease was built. Figure 8 (top) reflects this pattern for a reduction of 10% of the annual hydro inflow observed in Spain in 2010. Figure 8 (bottom) shows the original hourly hydro inflow in Spain for 2010 and the modified one.

The roadmap for the Spanish power system described in Reference [2] includes reinforcements of interconnections and 2.5 GW of batteries with Maximum Charge/Discharge time of 2 h by 2030. As mentioned in Section 4, both interconnections and batteries are not considered in the current version of the tool, what makes the obtained results a conservative picture of the real performance of the analysed transition path.
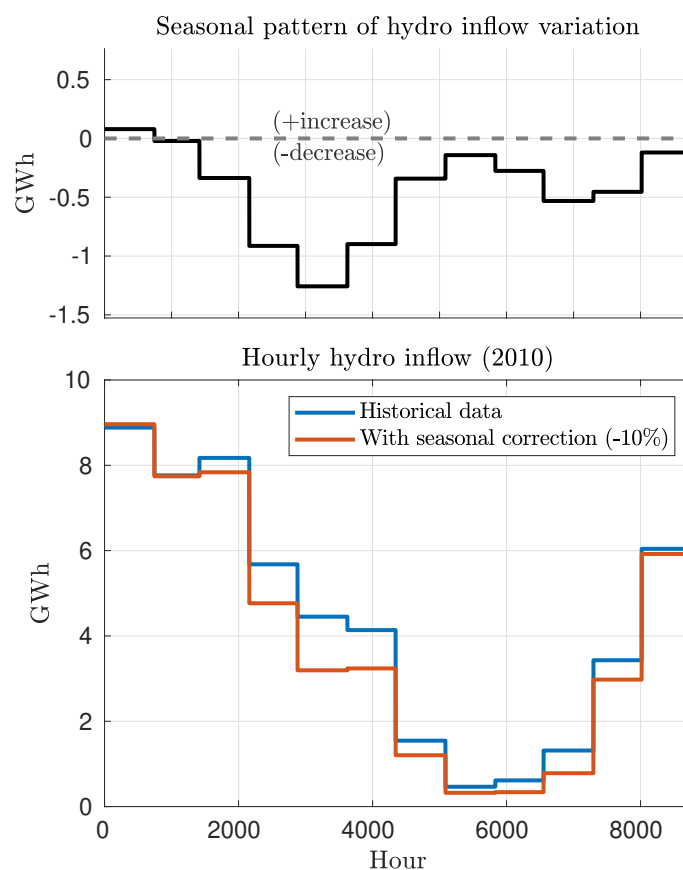

Figure 8. Top: Seasonal variation of hydro inflow, adapted from Reference [20], and computed for a reduction of 10% of the annual hydro inflow observed in Spain in 2010. **Bottom**: Hourly hydro inflow in Spain for 2010, and modified assuming a variation of −10% in the annual hydro inflow and the seasonal variation (see text for details).

Results for transition path *NECP* are shown in Figure 9. The two plots on the top show the distribution of the CO_2 emissions and the percentage of demand covered with RES, respectively, together with the targets established by Spanish authorities (NECP targets, in Reference [2]). Both sets of data fit notably well, which serves as a validation of the Spanish plans based on our tool for transition path analyses. It is important to remark that, while results obtained in our simulations correspond to peninsular Spain, NECP goals refer to Spain, thus, they include insular territories. This could explain why the obtained emissions remain slightly below the NECP target.
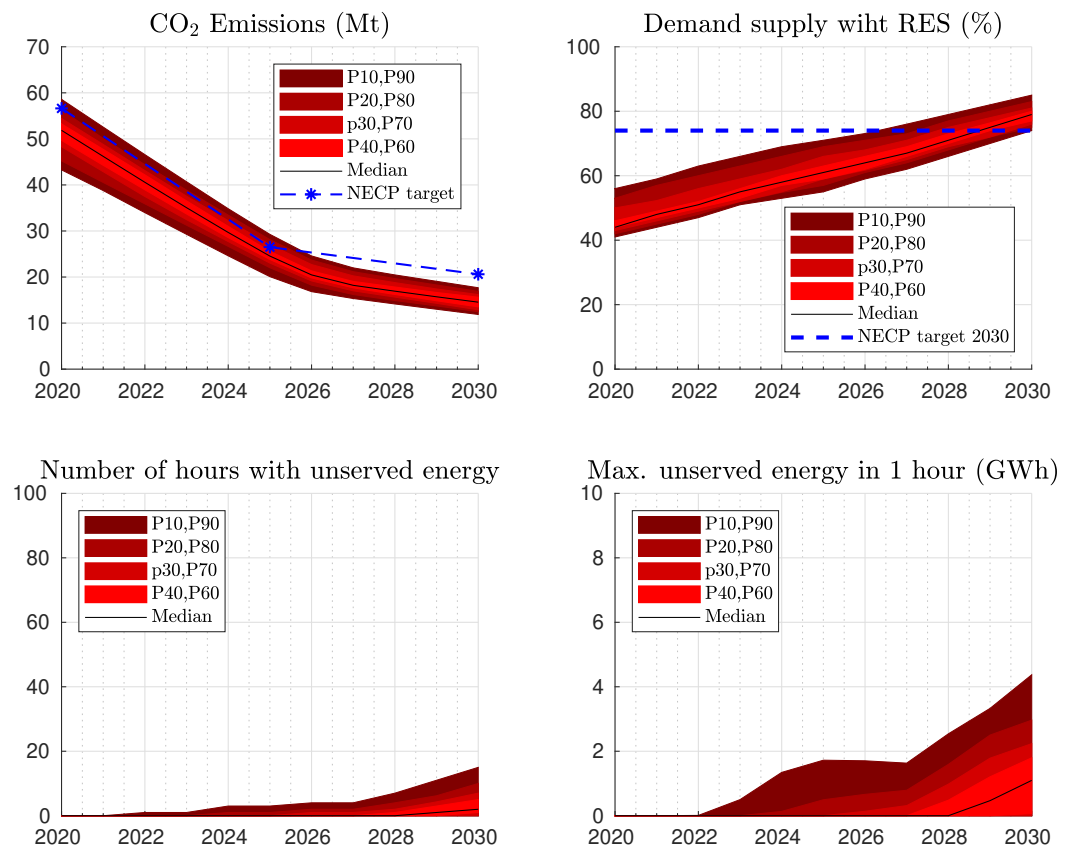
Figure 9. Performance indicators for transition path *NECP*.

Our results also suggest a certain probability of hours with unserved energy. The two plots on the bottom in Figure 9 show that this could be the case for around 15 h by 2030, the maximum unserved energy in one hour being lower that 4.5 GWh. Based on the limitations of the model mentioned at the end of Section 4, it is very likely that such situations could be safely managed with non-modelled elements, like current interconnection levels and interruptibility service, and/or the expected installed batteries and demand-side management to be developed in the near future.

*A More Realistic Picture for Coal in Spain*

As mentioned above, the Spanish power system roadmap described in Reference [2] considers a complete decommission of coal plants by 2030, with a noticeably decrease by 2025. However, during the elaboration of this research, coal decommissioning has unexpectedly accelerated, as several coal plants closed in 2019–2020. Additionally, several announcements make probably that all existing coal power plants in Spain are decommissioned by 2023, see Table 1. The main reason behind is the increasing cost of coal generation, mainly driven by increasing carbon prices in EU Emissions Trading Systems (ETS), in combination with the new RES capacity that is being installed in Spain. Actually, several decommission announcements are for carbon facilities that had already invested in adapting their installations to new European legislation on quality air, and which in principle could keep operating. Table 1 shows the remaining coal facilities in peninsular Spain and the installed power over the following year.

Thus, a relevant question is whether accelerated coal decommissioning may be incompatible with NECP roadmap, in the sense that the risk of unserved energy could increase noticeable, affecting the quality of electricity supply, specially during the first half of the period 2020–2030. To address this issue, a new transition path, referred to as *NECP-no coal*, has been defined by considering the total coal capacity detailed in Table 1. The rest of parameters are those included in *NECP* transition path.

Table 1. Expected installed coal capacity [MW] in peninsular Spain, according to different decommission announcements.

Coal Power Plant	2020	2021	2022	2023
As Pontes	1468	1468	0	0
Litoral	1159	1159	0	0
Los Barrios	589	589	0	0
Lada	358	358	0	0
Aboño	916	916	916	0
Soto de Ribera	350	350	350	0
TOTAL	4840	4840	1266	0

Results for this new transition path are shown in Figure 10. Two main differences can be appraise between *NECP* and *NECP-no coal* performance. First, CO_2 emissions are notably reduced in the *NECP-no coal* transition path, specially during the first half of the period. In particular, the accumulated emissions during the decade were reduced by 18.5% (from 313.7 Mt to 255.6 Mt). This reveals the extent to which power system emissions are driven by coal, thus, the importance of displacing this energy source in the context of climate policies. The counterpart is an increase in the maximum unserved energy in one hour, showing a need for additional flexible generation earlier in the period (2020–2023). However, the amount of required capacity during these first years remains low, below 2 GW.
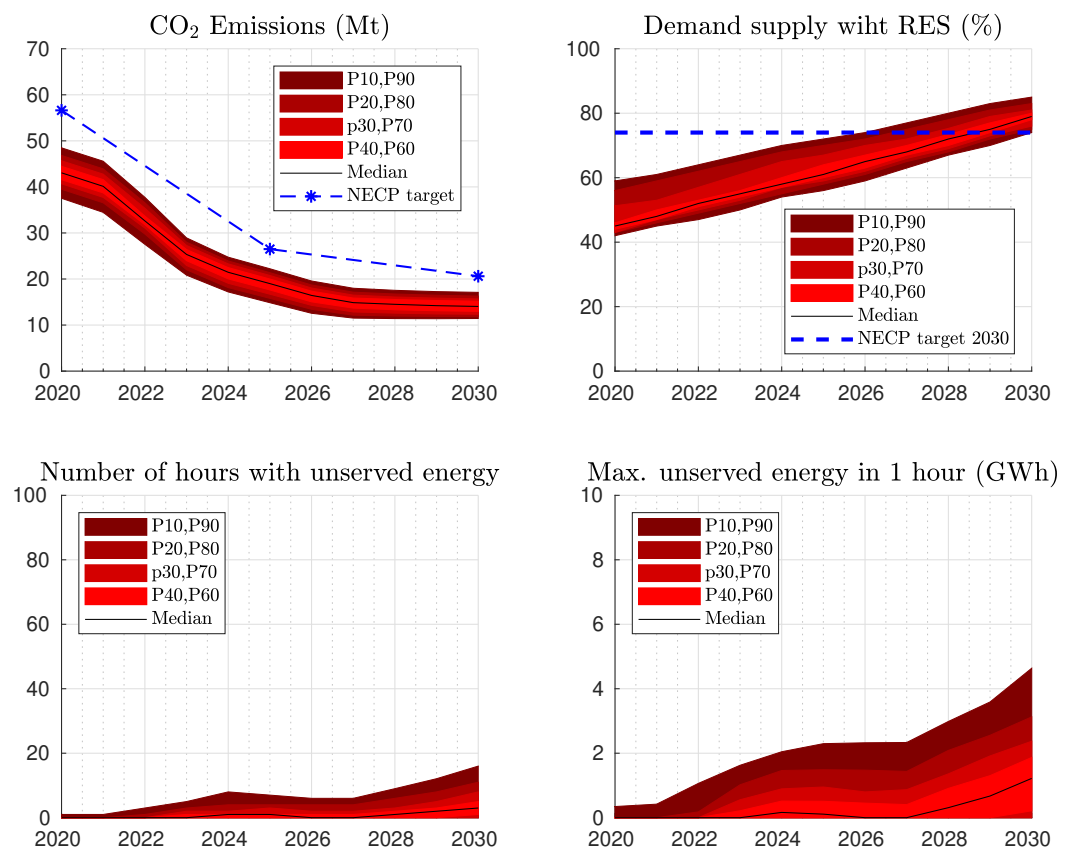
Figure 10. Performance indicators for transition path *NECP-no coal*.

6. Conclusions

In this work, a probabilistic model for hydropower generation was developed in order to improve an hourly-resolved tool for transition path analysis presented in previous research. The introduced model aimed to account for physical, environmental and regulatory constraints affecting hydropower dynamics. The model is grounded on time series analysis

of historical data, which exploits the fact that these constraints were met in the past. Results showed that the dynamics of the aggregated hydropower generation in peninsular Spain were properly accounted for by linear regression on the last week records. Dependencies with exogenous variables, such as thermal gap and reservoir level, were employed to determine predictive densities of the hydropower generation for the next hour.

The upgraded version of the transition path analysis tool showed a decrease in the hydropower flexibility, as compared with results previously published. In particular, hourly hydropower gradients were reduced, providing a better picture of the benefits and drawbacks associated with a specific transition path under analysis, for example, in terms of assessing the probability of unserved energy.

The upgraded version of the tool was employed to analyse the Spanish National Energy and Climate Plan 2020–2030 (NECP), and an updated version of this plan including the faster coal decommissioning already taking place in Spain. In both cases, a small probability of unserved energy was observed during the considered period. The quantification of the amount of maximum unserved energy in one hour showed that these situations could most probably be solved with currently available mechanisms not modelled by the tool (interconnections and interruptibility service), or with little investments in additional flexibility (batteries, demand-side management, etc.). The modelling of these flexibility sources as well as the coupling of the power system with other sectors such as heating/cooling and transport represent some lines for future research. Other lines of research should consider the impact of climate change on solar and wind resources, as well as on electricity demand. Finally, the obtained results indicate consistence between NECP proposals associated with the power system and related CO_2 reduction and share of renewable electricity targets. In particular, the expected faster coal decommissioning could led to -18.5% of the accumulated emissions during the decade, as compared to the original path for coal considered in the NECP.

References

1. UNFCCC. The Paris Agreement. In Proceedings of the 21st Conference of the Parties: Paris, France, 30 November–11 December 2015; United Nations/Framework Convention on Climate Change, Adoption of the Paris Agreement.
2. Ministerio para la Transición Ecológica y el Reto Demográfico. Integrated National Energy and Climate Plan 2021–2030. Technical Report. 2020. Available online: https://ec.europa.eu/energy/sites/ener/files/documents/es_final_necp_main_en.pdf (accessed on 1 September 2020).
3. Red Eléctrica de España. El Sistema Eléctrico Español 2019. Technical Report. Available online: https://www.ree.es (accessed on 10 September 2020).
4. European Commission. Study on Energy Storage. Contribution to the Security of the Electricity Supply in Europe. Technical Report. 2020. Available online: https://op.europa.eu/en/publication-detail/-/publication/a6eba083-932e-11ea-aac4-01aa75ed71a1/language-en (accessed on 15 September 2020).
5. Ibanez, E.; Magee, T.; Clement, M.; Brinkman, G.; Milligan, M.; Zagona, E. Enhancing hydropower modeling in variable generation integration studies. *Energy* **2014**, *74*, 518–528. [CrossRef]
6. Liu, H.; Andresen, G.B.; Brown, T.; Greiner, M. A high-resolution hydro power time-series model for energy systems analysis: Validated with Chinese hydro reservoirs. *MethodsX* **2019**, *6*, 1370–1378. [CrossRef] [PubMed]
7. Stoll, B.; Andrade, J.; Cohen, S.; Brinkman, G.; Brancucci Martinez-Anido, C. *Hydropower Modeling Challenges*; Technical Report; NREL: Denver, CO, USA, 2017. [CrossRef]

8. Edwards, B.; Flaim, S.J.; Howitt, R. Optimal Provision of Hydroelectric Power under Environmental and Regulatory Constraints. *Land Econ.* **1999**, *75*, 267–283. [CrossRef]
9. Victoria, M.; Gallego-Castillo, C. Hourly-resolution analysis of electricity decarbonization in Spain (2017–2030). *Appl. Energy* **2019**, *233–234*, 674–690. [CrossRef]
10. Pinson, P. Wind Energy: Forecasting Challenges for Its Operational Management. *Stat. Sci.* **2013**, *28*, 564–585. [CrossRef]
11. Tsay, R. *Analysis of Financial Time Series*; John Wiley & Sons: Hoboken, NJ, USA, 2002.
12. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*, 5th ed.; Wiley: Hoboken, NJ, USA, 2016.
13. Cleveland, W.; Grosse, E.; Shyu, W.; *Statistical Models in S*; Chapter Local Regression Models; CRC Press, Inc.: Boca Raton, FL, USA, 1991; pp. 309–376.
14. Huang, J.; Shen, H. Functional Coefficient Regression Models for Non-Linear Time Series: A Polynomial Spline Approach. *Scand. J. Stat.* **2004**, *31*, 515–534. [CrossRef]
15. Hastie, T.; Tibshirani, R. Varying-coefficient models. *J. R. Stat. Soc. Ser. B (Methodol.)* **1993**, *55*, 757–796. [CrossRef]
16. Fan, J.; Zhang, W. Statistical methods with Varying-Coefficient Models. *Stat. Interface* **2008**, *1*, 179–195. [CrossRef] [PubMed]
17. Gallego-Castillo, C. Statistical Models for Short-Term Wind Power Ramp Forecasting. Ph.D. Thesis, Escuela Técnica Superior de Ingenieros Aeronáuticos, Universidad Politécnica de Madrid, Madrid, Spain, 2013. Available online: http://oa.upm.es/21912 (accessed on 27 July 2020).
18. Victoria, M.; Zhu, K.; Brown, T.; Andresen, G.B.; Greiner, M. The role of storage technologies throughout the decarbonisation of the sector-coupled European energy system. *Energy Convers. Manag.* **2019**, *201*, 111977. [CrossRef]
19. Ministerio de Agricultura, Alimentación y Medio Ambiente. *Sexta Comunicación Nacional de España. Convención Marco de las Naciones Unidas sobre el Cambio Climático*; Technical Report; Ministerio de Agricultura, Alimentación y Medio Ambiente: Madrid, Spain, 2013.
20. Schlott, M.; Kies, A.; Brown, T.; Schramm, S.; Greiner, M. The impact of climate change on a cost-optimal highly renewable European electricity network. *Appl. Energy* **2018**, *230*, 1645–1659. [CrossRef]
21. Van Vliet, M.; Wiberg, D.; Leduc, S.; Riahi, K. Power-generation system vulnerability and adaptation to changes in climate and water resources. *Nat. Clim. Chang.* **2016**, *6*, 375–380. [CrossRef]