engineering proceedings

MDPI

Proceeding Paper

Optimizing Energy Planning in Countries with Extreme Climates: A Methodological Proposal for the Colombian Case †

César Dubbier Castro *, Lina Montuori * and Manuel Alcázar-Ortega

Institute for Energy Engineering, Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain; malcazar@iie.upv.es

* Correspondence: ccastro2@doctor.upv.es (C.D.C.); lmontuori@iie.upv.es (L.M.)

† Presented at the 1st International Online Conference on Inventions—Energy Security and Sustainable Development, 25–26 June 2026; Available online: https://sciforum.net/event/IOCIV2026.

Abstract

This research proposes an optimization model for the scheduling of power generation expansion in countries with high hydrological variability, to be implemented in the Colombian system for the 2027–2041 period. The method combines a Vector Autoregressive (VAR) model, evaluated using daily series of inflows, demand, and exports for 2010–2024, with a two-stage stochastic optimization model. Unlike previous methods, the model maintains the contemporaneous covariance calculated when creating the planning scenarios. Hydrological inflows cause demand (p = 0.0003), which reflects the possible effect of the El Niño–Southern Oscillation (ENSO). Their contemporaneous correlation of −0.1018 comes from the dry months, which coincide with the months of higher consumption. The main contribution of this research lies in measuring the cost of disregarding this dependence. This is performed using two similar optimization problems that differ only in whether they maintain the covariance. In the base case, the plan based on the independence hypothesis adds 45% more capacity, replaces wind power with solar power, and requires an additional 950 million dollars compared with the scenario that maintains the correlation, representing 1.97% of the system's optimal cost.

Keywords: stochastic generation expansion planning; vector autoregressive model; scenario generation; contemporaneous dependence; hydropower dependence; Colombia

1. Introduction

It is well known that long-term operations, in addition to the planning of the different energy systems, face challenges due to emerging factors, such as the global energy transition, together with the impact of extreme climate events that occur in different regions of the world [1–3]. There are countries where the majority of energy production is generated from water sources, which depend directly on water availability and climatic conditions determined by rainfall and river flows. However, there are climate phenomena that modify water availability in these regions [4,5]. In the case of Colombia, the El Niño and La Niña phenomena significantly alter the amount of water available for energy generation in the different electrical grids across the country. This causes operational stress, as well as high volatility in electricity generation, with a high risk of energy deficits [6]. Colombia depends on hydropower for between 65% and 70% of its electricity generation, and it is well documented that there have been periods of drought that have been induced by the El Niño phenomenon, during which the country, due to extreme temperatures and

Check for updates

Academic Editor: Eugen Rusu

Published: 30 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

lack of rainfall, experienced large drops in the levels of its reservoirs [7]. This caused the country to use thermal generation as a backup and changed the commercial and international dynamics (exports to countries like Ecuador), even though this may have been more expensive than expected [8]. Over the years, the different governments have made policy and regulatory efforts to diversify the energy matrix through renewable and non-conventional energy sources, such as solar and wind energy, while ensuring that these developments remain aligned with the growth in electricity demand [9,10]. However, this is where a bottleneck arises. An analysis of the scientific literature shows that energy system planning is approached using different models, generally through mathematical optimization models that minimize costs while maximizing production. Despite this, most tools have methodological limitations. For example, they operate under the assumption that demand is represented as a static average value, treated as an exogenous and deterministic variable. They isolate climate impacts, such as typical days and drought days, thereby breaking the continuity of the system. Therefore, the complexity of the power system and the interdependencies within the region are significantly ignored, as well as the autocorrelation of the series and the statistical memory contained in each variable. For example, demand does not react in a linear or immediate manner when there is a production surplus; instead, its response follows different dynamics and may exhibit cumulative effects over time.

In the Colombian context, different energy planning models have been developed with the objective of determining the optimal combination of energy technologies under demand, investment cost, and operational constraints. However, these types of models have specific limitations. For instance, the uncertainty associated with the variables that determine the energy balance is not incorporated through a jointly estimated dependence structure based on the historical behavior of the system.

If we talk about more recent studies, different investigations have addressed the planning of systems that depend on hydropower energy, but the joint dependence between hydrology and demand is not modeled. An example is Guamán et al. [11], which quantifies hydrological uncertainty and does so through Monte Carlo simulations, which are calibrated by taking into account the relationships between precipitation, inflows, and generation; however, it is important to highlight that they study demand as an uncorrelated exogenous forecast. Studies such as Parra-Rodas et al. [12] conduct research where they model capacity expansion with different scenarios that are exploratory and that do not consider the low water availability caused by the El Niño phenomenon. In addition, there is evidence showing that this dependence does exist. For example, Pulgarín-Morales et al. [13] report that in El Niño years, inflows fall by up to 31%, but on the other hand, national demand increases by 10%. Furthermore, other studies such as Jiménez-Sáenz et al. [14] show in their research that the different effects found are especially heterogeneous. This is where none of the works takes into account this dependence structure, which comes from the historical behavior of the entire system, nor do they develop or transfer it to the different scenarios of the expansion model.

There is no study in the Colombian context that models the dynamics of hydrological inflows, electricity demand, and electricity exports and uses a covariance matrix to generate multivariate scenarios within a generation expansion model. It is well established that these variables exhibit dynamic behavior and are likely to be correlated; however, existing studies do not account for this type of behavior and include some of these variables without considering their covariance structure, and, in some cases, demand is treated as an assumption rather than as an extrapolation derived from its historical behavior.

Therefore, considering this limitation in the state of the art of energy research in Colombia [15,16], we propose to develop a methodological framework that integrates

two main components: a Vector Autoregressive (VAR) model and a two-stage stochastic optimization model that can minimize costs under multivariate scenarios that maintain the contemporaneous covariance of the variables. The main contribution of this research lies in measuring the cost of disregarding this dependence, which is validated through a case study based on the Colombian power system and which can be implemented in other similar regions with a high dependence on hydropower.

2. Materials and Methods

2.1. Data and Case Study

The different data used in this research are collected from the Colombian National Interconnected System, from the SINERGOX platform of XM, the Colombian power system operator [17], where hydrological inflows, national demand, and exports are extracted; these are presented at a daily frequency from January 2010 to December 2024, with approximately 5479 observations for each series. It should be noted that the only null observation, on 31 December 2015, was imputed using the average of the five previous days, and it should be clarified that no extreme values that may have occurred due to phenomena such as El Niño and La Niña were removed, as these are precisely what the study aims to represent. The technical parameters come from the Mining and Energy Planning Unit (UPME)'s 2025–2039 Reference Generation Expansion Plan, such as the installed capacities as of 31 December 2024, as well as investment costs, committed projects, firm energy, thermal retirements, and unserved energy [18].

2.2. Econometric Framework

The methodology used to analyze this part of the research was the VAR model, which allows for examining the joint temporal dynamics between water inflows, national demand, and energy exports, especially to Ecuador. This specification was selected because it treats the three variables as endogenous and estimates their contemporaneous dependence in a covariance matrix, unlike approaches that include demand as an exogenous assumption. Therefore, to develop this model, it was necessary to follow a series of steps aimed at analyzing and transforming the data, as well as validating the employed model [19–22]. Furthermore, the exact model used is summarized in Equations (1) and (2).

$$\begin{bmatrix} h_t \\ d_t \\ x_t \end{bmatrix} = c + A_1 \begin{bmatrix} h_{t-1} \\ d_{t-1} \\ x_{t-1} \end{bmatrix} + A_2 \begin{bmatrix} h_{t-2} \\ d_{t-2} \\ x_{t-2} \end{bmatrix} + \cdots + A_p \begin{bmatrix} h_{t-p} \\ d_{t-p} \\ x_{t-p} \end{bmatrix} + \begin{bmatrix} e_{h,t} \\ e_{d,t} \\ e_{x,t} \end{bmatrix} \quad (1)$$

or, in compact notation:

$$y_t = c + \Sigma_{i=1}^{p} A_i y_{t-i} + e_t \quad (2)$$

2.3. Variable Transformation and Level Estimation

It is important to highlight that prior to estimating the VAR model, each variable was transformed in order to obtain stationary series without having to resort to differencing each one, since the first-difference specification produced physically inadmissible trajectories. Water inflows were transformed using the natural logarithm, exports were transformed using ln(1 + xt), and for electricity demand, a natural logarithm was initially applied and subsequently a linear trend was extracted to preserve the parameters of this behavior.

2.4. Scenario Generation

Once the VAR model was estimated and validated, trajectories were generated for water inflows, electricity demand, and energy exports. To do this, the residuals obtained from the VAR model were used, which in a way represent a shock that is not explained by

the temporal dynamics of the variables. To carry out this trajectory generation process, empirical bootstrap modeling of the residuals was specifically used, in which innovations were randomly selected from the historical sample of residuals. By applying this methodology, exactly 200 trajectories were created, directly related to the two configurations considered in this research, which, to some extent, preserved the dynamic structure established by the VAR; in Set B, full residual rows were resampled jointly, preserving the contemporaneous covariance, whereas in Set A, each component was resampled independently. Additionally, it is worth noting that the last 22 observed values were used as initial conditions.

$$\mathrm{E}[e_t] = 0,\ \mathrm{E}\left[e_t {e_t}'\right] = \Sigma,\ \mathrm{E}\left[e_t {e_s}'\right] = 0\ \forall\ t \neq s \tag{3}$$

where the covariance matrix

$$\Sigma = \begin{bmatrix} {\sigma_h}^2 & \sigma_{hd} & \sigma_{hx} \\ \sigma_{hd} & {\sigma_d}^2 & \sigma_{dx} \\ \sigma_{hx} & \sigma_{dx} & {\sigma_x}^2 \end{bmatrix} \tag{4}$$

For sampling, two configurations were considered: Set B is the correlated one, generated because the three components of each innovation are selected jointly by sampling complete rows from the generated residual matrix. For Set A, each component of the residuals was selected independently. It is worth noting that both configurations preserve the temporal dynamics constructed by the VAR model; however, only Set B maintained the dependency structure among the variables. Once this procedure was completed, the simulated trajectories were back-transformed into physical units and aggregated to a monthly frequency, matching the time scale of the optimization model. Here, the 200 trajectories of each set were reduced to 30 scenarios representing those trajectories, created using the k-means methodology, in which the trajectory closest to the centroid was selected for each cluster, assigning a probability proportional to the number of trajectories in that cluster. Following this, the scenarios were trimmed to the 2027–2041 horizon. Furthermore, the difference in correlation between demand and water inflows was observed and verified to serve as input for the optimization model.

2.5. Optimization Model Formulation

The model integrates three fundamental components, which are the following: investment costs (CAPEX), operational costs (OPEX), and penalties for unserved energy, all evaluated in terms of net present value (NPV), considering a discount rate of 9% and a base reference year of 2026.

$$\begin{aligned} \text{Min } z = & \sum_{i \in I_{\text{exp}}} \sum_{n \in N} \sum_{y \in Y_{\text{d}}} \frac{\text{Exp}_{i,n,y} \cdot \text{CAPEX}_i \cdot \pi_n}{(1+r)^{y-2026}} \\ & + \sum_{y \in Y_{\text{d}}} \frac{\Delta\text{Int}_y \cdot \text{CAPEX}^{\text{int}}}{(1+r)^{y-2026}} \\ & + \sum_{i \in I} \sum_{n \in N} \sum_{y \in Y} \frac{\text{OM}_i^{\text{fix}} \cdot \text{Cap}_{i,n,y}}{(1+r)^{y-2026}} \\ & + \sum_{s \in S} \sum_{t \in T} \frac{p_s \cdot \text{OC}_{t,s}}{(1+r)^{\text{year}(t)-2026}} \end{aligned} \tag{5}$$

where the operating cost of each period and scenario is

$$\text{OC}_{t,s} = \sum_{i \in I} \sum_{n \in N} \text{Gen}_{i,n,t,s} \cdot \text{OPEX}_i + \text{Unserved}_{t,s} \cdot \text{C}^{\text{ENS}} + \text{Im}_{t,s} \cdot \text{C}^{\text{imp}} - \text{Ex}_{t,s} \cdot P^{\text{exp}} \tag{6}$$

where

- $\text{Exp}_{i,n,y}$: Newly installed capacity of technology i in region n in year y (MW).

https://doi.org/10.3390/engproc2026152016

- CAPEX_i: Investment cost per unit of capacity of technology i (USD/MW).
- π_n: Regional cost premium (1.25 in Caribe, 1.00 in the remaining regions).
- ΔInt_y: Interconnection capacity added in year y (MW), with a unit cost CAPEXint of 660,000 USD/MW.
- OM_i^{fix}: Fixed operation and maintenance cost of technology i (USD/MW-year).
- $\text{Gen}_{i,n,t,s}$: Energy generated by technology i in region n in period t under scenario s (MWh).
- OPEX_i: Variable operational cost of technology i (USD/MWh).
- $\text{Unserved}_{t,s}$: Unserved energy in period t under scenario s (MWh).
- C^{ENS}: Cost of unserved energy (3514.43 USD/MWh, UPME Section 3).
- C^{imp}, P^{exp}: Import cost (120 USD/MWh) and export price (60 USD/MWh).
- p_s: Probability of scenario s, obtained from the scenario reduction stage.
- r: Discount rate (9%).
- $\text{year}(t)$: Corresponding year of period t.
- Y_{d}, Y: Set of decision years (2029–2041) and set of all years of the horizon (2027–2041).

Energy Balance Constraint.

$$\sum_{i\in I}\sum_{n\in N}\text{Gen}_{i,n,t,s}+\text{Unserved}_{t,s}+\text{Im}_{t,s}=\text{Demand}_{t,s}+\text{Ex}_{t,s}\ \forall t\in T,\forall s\in S \quad (7)$$

Generation Capacity Constraint

$$\begin{aligned}\text{Gen}_{i,n,t,s}\ &\le\ \text{Cap}_{i,n,t}\cdot\text{AF}_{i,n}\cdot H_t\quad \forall\, i\in I_{\text{exp}},\ n\in N,\ t\in T,\ s\in S\\ \text{Gen}_{\text{hydro},n,t,s}\ &\le\ \text{Cap}_{\text{hydro},n,t}\cdot 0.95\cdot H_t\quad \forall\, n\in N,\ t\in T,\ s\in S\end{aligned} \quad (8)$$

Dynamic Installed Capacity and Availability Factor

$$\text{Cap}_{i,n,t}=\begin{cases}\frac{\text{Cap}^{\text{base}}_{i,year(t)}}{|N|}+\sum_{y+\text{LT}_i\le\text{year}(t)}\text{Exp}_{i,n,y}, i\in I_{\text{exp}}\\ \frac{\text{Cap}^{\text{base}}_{i,year(t)}}{|N|},\ i=\text{hydro}\end{cases} \quad (9)$$

$$\text{Cap}^{\text{base}}_{i,y}=\text{Cap}^{0}_{i}+\sum_{y'\le y}\left(\text{Pipe}_{i,y'}-\text{Ret}_{i,y'}\right) \quad (10)$$

$$\text{AF}_{i,n}=\begin{cases}\text{af}_{i,n}, i\in I_{\text{exp}}\\ 0.95,\ i=\text{hydro}\end{cases} \quad (11)$$

Reservoir Balance, Firm Energy and Interconnection

$$\begin{aligned}\text{Vol}_{t,s}&=\text{Vol}_{t-1,s}+\text{Inflow}_{t,s}-\tfrac{1}{10^3}\sum_{n\in N}\text{Gen}_{\text{hydro},n,t,s}-\text{Spill}_{t,s}\\ \text{Vol}_{-1,s}&=V_0=0.65\cdot V^{\max}\end{aligned} \quad (12)$$

$$\text{Vol}_{t,s}\ \ge\ \gamma_t\cdot V^{\max},\gamma_t=0.45\text{ if month}(t)\ \in\ [\text{Dec, Jan, Feb, Mar}],\ 0.30\text{ otherwise}\quad \text{Vol}_{t_{\text{end}},s}\ \ge\ 0.30\cdot V^{\max} \quad (13)$$

$$\begin{aligned}\text{FirmCap}_y&=\sum_{i\in I}\text{Cap}^{\text{base}}_{i,y}\cdot\text{ENFICC}_i+\sum_{i\in I_{\text{exp}}}\sum_{n\in N}\sum_{y'+\text{LT}_i\le y}\text{Exp}_{i,n,y'}\cdot\text{ENFICC}_i\\ 8760\cdot\text{FirmCap}_y\ &\ge\ 0.90\cdot\sum_{s\in S}p_s\sum_{t\in T_y}\text{Demand}_{t,s}\quad\forall\, y\in Y_{\text{d}}\end{aligned} \quad (14)$$

$$\begin{aligned}\text{Int}_t&=\text{Int}_0+\sum_{y\in Y_{\text{d}}:y\le\text{year}(t)}\Delta\text{Int}_y\\ \text{Ex}_{t,s}\ &\le\ \text{Int}_t\cdot H_t,\ \text{Im}_{t,s}\ \le\ \text{Int}_t\cdot H_t,\ \text{Ex}_{t,s}\ \le\ 1.5\cdot X_{t,s}\\ \text{Int}_0+\sum_{y\in Y_{\text{d}}}\Delta\text{Int}_y\ &\le\ \text{Int}^{\max}\end{aligned} \quad (15)$$

To guarantee the mathematical and physical validity of the model, different non-negativity constraints were applied to avoid the decision variables having negative values that do not make sense. In practical terms, this means that it is not possible to generate negative energy, have negative deficits, or build negative capacity, which guarantees that the obtained solutions are coherent with the reality of the electrical system. The complete formulation is a two-stage stochastic linear program, solved with the HiGHS solver (https://highs.dev, accessed on 27 September 2026) through the Pyomo interface (https://www.pyomo.org, accessed on 27 September 2026).

3. Results

3.1. VAR Specification and Out-of-Sample Validation

Before running the VAR model, it was necessary to verify that the transformed series met all the conditions and requirements for the specification of a VAR model. When the transformation was performed, the Augmented Dickey–Fuller (ADF) tests confirmed the stationarity of the series. Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests, working with a fixed bandwidth, confirmed stationarity for inflows and demand, with exports remaining at the limit, as can be observed in Table 1. The Bayesian information criterion (BIC) selected 22 lags out of a maximum of 30. Thus, the estimated model meets the stability conditions.

Table 1. Stationarity tests and persistence of the transformed series.

Series	ADF p	KPSS p (L = 365)	AR (1) ρ	Half-Life (Days)
Inflows	0.0000	0.0684	0.913	7.6
Demand	0.0000	0.1000	0.462	0.9
Exports	0.0000	0.0445	0.924	8.8

It is important to clarify that the predictive performance of the VAR model must be evaluated prior to feeding it into the optimization model; therefore, a validation was conducted using a sample spanning from 2010 to 2024. Here, the sample was split into two datasets: one for training, covering the period from 2010 to 2022, and another for testing, running from 2023 to 2024; the VAR model was estimated on the training sample, and forecasts were generated for the test period. Overall, the results were evaluated using root mean square error (RMSE) and normalized RMSE (NRMSE) metrics at a monthly frequency. As shown in Table 2, demand exhibited the best predictive performance, with a monthly RMSE of 160 GWh and an NRMSE of 15.2%. Water inflows, along with exports, presented higher relative errors. A comparison was made between the VAR model specifically with an autoregressive integrated moving average (ARIMA) (2, 0, 1), as well as a seasonal naive forecast using the same data series. The VAR model outperforms both models in the three proposed series, with the greatest improvement observed in demand, where the seasonal naive forecast doubles the forecast error; this enabled a clear assessment of the VAR model's performance while evaluating its ability to reproduce the observed dynamics before applying the model to stochastic scenario generation.

Table 2. Out-of-sample validation of the VAR model against benchmark forecasting models (2023–2024).

Model	Inflows RMSE (GWh)	Inflows nRMSE (%)	Demand RMSE (GWh)	Demand nRMSE (%)	Exports RMSE (GWh)	Exports nRMSE (%)
VAR (proposed)	2322	30.2	160	15.2	118	38.9
ARIMA (2, 0, 1)	2347	30.5	164	15.7	119	39.3
Seasonal naive	2525	32.8	296	28.3	133	44.0

3.2. Validation of the Scenario Dependence Structure

As expected, the results reveal a contrast between the two sets. The first is that the correlation between demand and inflows is −0.1440 in Set B. In Set A it is −0.0658, representing a difference of approximately −0.0782, in addition to a t-statistic of around −4.26. This result indicates that there is a statistically significant difference. Moreover, the value for Set B is consistent with the correlation of −0.1018 estimated from the VAR residuals. What can be concluded here is that the use of complete rows of these residuals allows the dependence structure that has been estimated to be transferred to the set of scenarios. In contrast, it can also be observed that Set A presents a residual correlation, but it is considerably lower, since independent resampling weakens the association between these variables. It can also be observed that the rows corresponding to the equivalence presented in Table 3 are equally important. The mean annual inflows and the mean annual demand are the same in both sets. The retained scenarios and the temporal dynamics inherited from the execution of the VAR model are also very similar. This is where it can be concluded that both sets maintain the same marginal distributions, in addition to the same growth trajectories and the same autocorrelation structure, and that the only difference between them is the contemporaneous dependence among the three variables. This is important to highlight because any differences in the investment decisions obtained in Section 3.4 can be attributed primarily to this single factor.

Table 3. Statistical characterization of the two scenario sets.

Parameter	Value
Horizon	1 January 2027 to 1 December 2041 (180 months)
Demand–inflow correlation	
Set B	−0.1440
Set A	−0.0658
Difference	−0.0782
T	−4.26
Significant	Yes
VAR residual correlation	−0.1018
Annual demand	106.1 TWh/year
Annual inflows	59.3 TWh/year

3.3. Optimal Expansion Plan Under Correlated Scenarios

The main result, obtained from the joint optimization of the correlated scenario set, determines a total expansion of 9228 MW over the entire planning horizon, distributed across the three technologies and three of the four macro-regions. It should be noted that the interconnection is expanded to 1016 MW. The system reaches a firm capacity of 13,029 MW, with an expected unserved energy of 6777 GWh over the entire 15-year evaluation period. This represents approximately 0.4% of the total accumulated demand. The complete composition of the expansion plan is presented in the following table (Table 4).

It can be observed that the Caribe region reaches up to 96% of its potential for wind energy production, and, on the other hand, the Orinoquía region reaches 100% of its potential for solar energy production. This indicates that the expansion is limited by the availability of the resource, and not by the cost of each technology.

https://doi.org/10.3390/engproc2026152016

Table 4. Optimal capacity expansion under the correlated scenario set (Set B), 2029–2041.

Region	Wind (MW)	Solar (MW)	Thermal (MW)	Biomass (MW)	Total (MW)	% of Regional Potential
Caribe	5773	0	0	0	5773	96%
Andina	0	0	0	0	0	0%
Pacífica	0	0	455	0	455	91%
Orinoquía	0	3000	0	0	3000	100%
Total	5773	3000	455	0	9228	

3.4. Cost of the Independence Assumption

The central result of this research lies in comparing these two sets, since it is necessary to isolate the effect of the independence assumption. Both plans are evaluated under the same set of scenarios. For example, the correlated plan installs 9228 MW, including 5773 MW of wind power generation in the Caribe region (Table 5). The plan based on independence installs 13,410 MW, which is 45% more and changes the technological composition toward solar power generation, with 10,463 MW. Of these, 7463 MW are located in the Caribe region and only 2947 MW correspond to wind power generation. The independence assumption increases investment. It also modifies the technology and location of the expansion. Under the correlated scenario set, the independence-based plan reaches an NPV of 49.27 billion USD against 48.32 billion for the correlated plan, a difference of 0.95 billion USD, which is equivalent to 1.97% of the optimal system cost.

Table 5. Expansion mix and system indicators under the two scenario sets.

Indicator	B (Correlated)	A (Independent)	Difference	Difference %
Total new (MW)	9228	13,410	4182	45.3200
Firm capacity (MW)	13,028	13,193	165	1.2700
Interconnection (MW)	1015	1398	383	37.7300
Expansion mix (MW):				
Thermal	455	0.0000	−455	
Solar	3000	10,463	7463	
Wind	5773	2947	−2826	
Biomass	0.0000	0.0000	0.0000	

The difference between the two expansion plans is due to the temporal profile of the resources. When independence is assumed, the model does not identify that demand increases when water inflows decrease, and it assumes a uniform deficit throughout the evaluated period. This is where solar power generation is favored. When the estimated dependence is considered, the stress is concentrated in the dry season and specifically during the night hours, when wind power generation in La Guajira is more valuable. This suggests that the error is one of diagnosis rather than magnitude. When the evaluation is performed under the correlated set, the plan based on independence costs 950 million dollars more, equivalent to 1.97% of the optimal net present cost. This result could be referred to as the cost of the independence assumption, which can be considered a main result of the comparison between these two scenarios.

3.5. Robustness and Sensitivity Analysis

To calculate the cost of the independence assumption, three different resampling seeds were used, where in all three cases it was positive, first with a mean of 2.98% and moving within ranges between 1.79% and 5.18% (Table 6). The magnitude has some dispersion,

since the reference NPV is between values of 35.1 and 54.2 billion, varying according to the seed. This may indicate that 30 scenarios do not stabilize the objective function. It is important to highlight that both plans are evaluated on the same set for each seed, and this does not compromise the validity of the estimated cost. What this does, in some way, is force the reporting of the range within which it varies.

Table 6. Cost of the independence assumption by resampling seed.

Seed	New Capacity B (MW)	New Capacity A (MW)	NPV B (bn USD)	NPV A (bn USD)	CIA (bn USD)	CIA (%)
20260101 (base)	9228	13,410	48.32	49.27	0.95	1.97
777	9044	9226	35.14	35.77	0.63	1.79
2024	11,256	9044	54.19	57.00	2.81	5.18
Mean	—	—	—	—	1.46	2.98

4. Conclusions

This research was able to confirm that the inflows from hydropower production cause Colombian electricity demand (p-value of 0.0003). Both variables, hydrological inflows and electricity demand, present a contemporaneous correlation of −0.1018, since dry seasons coincide with those of higher consumption. The proposed VAR model outperformed the univariate ARIMA model and the seasonal naïve forecast in accuracy in the three series evaluated out of the sample.

The main contribution in terms of methodological development is to specifically quantify the cost of assuming correlational independence between hydrology, demand, and exports, and this was achieved by comparing two expansion plans obtained with a set of scenarios that differ specifically in whether they preserve the covariance matrix over time, evaluating them on the same reference set; in this way, the design of the methodology isolates the effect of the assumption and expresses it as an economic and monetary-level indicator, which can be differentiated from previous works, which incorporate uncertainty without jointly estimating its dependence structure.

If the base case is analyzed, the plan constructed under independence installs 13,410 MW, unlike the 9228 MW of the plan that takes into account the correlation between the variables (i.e., 45% additional capacity), shifting the investment from wind energy towards solar energy. When the uncorrelated set is evaluated economically, it costs 950 million dollars more, which is equivalent to 1.97% of the optimal net present value. The error is oriented towards the diagnosis and not towards the magnitude, as it is not recognized that the stress is concentrated in the dry season and during nighttime hours. Therefore, the model underestimates the value of wind energy in the La Guajira region.

The estimated cost was positive in the different resampling seeds that were evaluated, with a mean of 2.98% and ranges varying between 1.79% and 5.18%. The NPV varies between 35.1 and 54.2 billion dollars depending on the seed, which does not completely stabilize the objective function. This is why a range is reported and not a specific value. The results correspond to the Colombian system. For its extension to other countries that depend on hydrology, it should be verified that there is a statistical relationship between hydrology and demand. Furthermore, if this relationship is weak or even positive, the cost of the independence assumption may be lower or even have a different sign.

Author Contributions: Conceptualization, C.D.C. and L.M.; Methodology, C.D.C. and L.M.; Software, C.D.C.; Validation, L.M. and M.A.-O.; Formal analysis, C.D.C. and L.M.; Writing—original draft preparation, M.A.-O.; Writing—review and editing, L.M. All authors have read and agreed to the published version of the manuscript.

Funding: The research has been developed in the framework of the project PAID-06-25, financed by the Vice-Rectorate for Research of the Universitat Politècnica de València.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. These data can be found in the SINERGOX platform of XM at https://sinergox.xm.com.co (accessed on 27 September 2026) and in the UPME 2025–2039 Reference Generation Expansion Plan [18].

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Perera, A.T.D.; Nik, V.M.; Chen, D.; Scartezzini, J.-L.; Hong, T. Quantifying the Impacts of Climate Change and Extreme Climate Events on Energy Systems. *Nat. Energy* **2020**, *5*, 150–159. [CrossRef]
2. Perera, A.T.D.; Hong, T. Vulnerability and Resilience of Urban Energy Ecosystems to Extreme Climate Events: A Systematic Review and Perspectives. *Renew. Sustain. Energy Rev.* **2023**, *173*, 113038. [CrossRef]
3. Castro Hernandez, C.D.; Montuori, L.; Alcázar-Ortega, M.; Olczak, P. Extreme Climate Events and Energy Market Vulnerability: A Systematic Global Review. *Appl. Sci.* **2025**, *15*, 6210. [CrossRef]
4. Wasti, A.; Ray, P.; Wi, S.; Folch, C.; Ubierna, M.; Karki, P. Climate Change and the Hydropower Sector: A Global Review. *WIREs Clim. Change* **2022**, *13*, e757. [CrossRef]
5. Kemarau, R.A.; Harun, S.N.; Sa'adi, Z.; Mohd Hanafiah, M.; Sakawi, Z.; Norzin, M.A.F.; Wan Mohd Jaafar, W.S.; Anak Suab, S.; Eboy, O.V.; Abdul Maulud, K.N. Transforming Hydropower: An in-Depth Systematic Review of Climate Change Impacts. *Renew. Sustain. Energy Rev.* **2025**, *219*, 115890. [CrossRef]
6. Restrepo-Trujillo, J.; Moreno-Chuquen, R.; Jiménez-García, F.; Flores, W.; Chamorro, H. Scenario Analysis of an Electric Power System in Colombia Considering the El Niño Phenomenon and the Inclusion of Renewable Energies. *Energies* **2022**, *15*, 6690. [CrossRef]
7. Arias-Gaviria, J.; Van Der Zwaan, B.; Kober, T.; Arango-Aramburo, S. The Prospects for Small Hydropower in Colombia. *Renew. Energy* **2017**, *107*, 204–214. [CrossRef]
8. Henao, F.; Dyner, I. Renewables in the Optimal Expansion of Colombian Power Considering the Hidroituango Crisis. *Renew. Energy* **2020**, *158*, 612–627. [CrossRef]
9. Arias-Gaviria, J.; Carvajal-Quintero, S.X.; Arango-Aramburo, S. Understanding Dynamics and Policy for Renewable Energy Diffusion in Colombia. *Renew. Energy* **2019**, *139*, 1111–1119. [CrossRef]
10. Pupo-Roncallo, O.; Campillo, J.; Ingham, D.; Hughes, K.; Pourkashanian, M. Large Scale Integration of Renewable Energy Sources (RES) in the Future Colombian Energy System. *Energy* **2019**, *186*, 115805. [CrossRef]
11. Guamán, W.; Benalcazar, P.; Cordova-Garcia, J.; Torres, M. An Integrated Framework for the Optimal Expansion of Hydro-Dependent Power Systems under Water-Resource Uncertainty. *Energy Convers. Manag. X* **2025**, *28*, 101297. [CrossRef]
12. Parra-Rodas, J.F.; Arias-Gaviria, J.; Olaya, Y.; Larsen, E.; Arango-Aramburo, S. Quantifying the 2050 Scenarios for Colombia's Electricity Sector Transition. *Renew. Energy* **2026**, *256*, 124040. [CrossRef]
13. Pulgarín-Morales, L.; Krueger, T. ENSO Effects on Hydropower and Climate Adaptation Strategies in Four Colombian Case Studies. *J. Hydrol. Reg. Stud.* **2025**, *60*, 102599. [CrossRef]
14. Jiménez-Saenz, C.; Pulgarin-Morales, L.; Krueger, T. The Water-Energy Nexus under ENSO Variability in Four Colombian Hydropower Basins. *Int. J. River Basin Manag.* **2026**, *24*, 445–456. [CrossRef]
15. Collins, S.; Deane, J.P.; Poncelet, K.; Panos, E.; Pietzcker, R.C.; Delarue, E.; Gallachóir, B.P.Ó. Integrating Short Term Variations of the Power System into Integrated Energy System Models: A Methodological Review. *Renew. Sustain. Energy Rev.* **2017**, *76*, 839–856. [CrossRef]
16. Kachirayil, F.; Weinand, J.M.; Scheller, F.; McKenna, R. Reviewing Local and Integrated Energy System Models: Insights into Flexibility and Robustness Challenges. *Appl. Energy* **2022**, *324*, 119666. [CrossRef]
17. XM S.A. E.S.P. Reporte de Reservas Hídricas Del SIN: Capacidad Útil de Energía de Los Embalses 2026. Available online: https://sinergox.xm.com.co/ (accessed on 27 September 2026).
18. *Unidad de Planeación Minero Energética Plan Indicativo de Expansión de La Generación 2025-2039*; UPME: Bogotá, Colombia, 2025.
19. Narayan, P.K.; Narayan, S.; Prasad, A. A Structural VAR Analysis of Electricity Consumption and Real GDP: Evidence from the G7 Countries. *Energy Policy* **2008**, *36*, 2765–2769. [CrossRef]
20. Esmaeili, P.; Rafei, M. Dynamics Analysis of Factors Affecting Electricity Consumption Fluctuations Based on Economic Conditions: Application of SVAR and TVP-VAR Models. *Energy* **2021**, *226*, 120340. [CrossRef]

21. Silva, S.; Soares, I.; Pinho, C. The Impact of Renewable Energy Sources on Economic Growth and CO2 Emissions—A SVAR Approach. *Eur. Res. Stud. J.* **2012**, *XV*, 133–144. [CrossRef]
22. Valencia-Herrera, H.; Santillán-Salgado, R.J.; Venegas-Martínez, F. On the Interaction among Economic Growth, Energy-Electricity Consumption, CO2 Emissions, and Urbanization in Latin America. *Rev. Mex. Econ. Finanz.* **2020**, *15*, 745–767. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.