

# Capacity planning of renewable energy systems using stochastic dual dynamic programming

J. Hole<sup>a,b,c,\*</sup>, A.B. Philpott<sup>a</sup>, O. Dowson<sup>a</sup>

<sup>a</sup>*Electric Power Optimization Centre, University of Auckland, New Zealand*

<sup>b</sup>*Norwegian University of Science and Technology, Norway*

<sup>c</sup>*Norwegian Water Resources and Energy Directorate, Norway*

---

## Abstract

We present a capacity expansion model for deciding the new electricity generation and transmission capacity to complement an existing hydroelectric reservoir system. The objective is to meet a forecast demand at least expected cost, namely the capital cost of the investment plus the expected discounted operating cost of the system. The optimal operating policy for any level of capacity investment can be computed using stochastic dual dynamic programming. We show how to combine a multistage stochastic operational model of the hydro system with a capacity expansion model to create a single model that can be solved by existing open-source solvers for multistage stochastic programs without the need for customized decomposition algorithms. We illustrate our method by applying it to a model of the New Zealand electricity system and comparing the solutions obtained with those found in a previous study.

**Keywords:** OR in energy; stochastic dual dynamic programming; capacity expansion; multistage stochastic programming

---

---

\*Corresponding author

Email address: [jarand.hole@ntnu.no](mailto:jarand.hole@ntnu.no) (J. Hole)

Preprint submitted to Elsevier

December 12, 2023

## 1. Introduction

Electricity systems around the world are transitioning to technologies with zero or near-zero carbon emissions (Clarke et al., 2022). The most popular sources of renewable energy (wind and solar) are either intermittent or diurnal and so some form of flexibility is required to match supply and demand when the wind is not blowing or the sun is not shining. One option is thermal peaker plants, which can quickly ramp up and down to balance supply and demand. Another option is to store and discharge energy from some form of storage, which may be medium-term storage in the form of hydro-reservoirs, or short-term storage in the form of batteries. Since future wind and solar generation are uncertain, the optimal operation of storage is the solution to a stochastic control problem of some complexity. A third option is to invest in interconnectors, through which neighbouring power systems can share excess electricity and flexibility.

In addition to the intermittency challenges, isolated hydro-dominated energy systems like Iceland, Brazil, and New Zealand, face a “dry-year” risk, in which a sustained period of low inflows leaves insufficient water in the reservoirs to meet demand. To compensate, the system must either be willing to periodically shed demand at high cost (and social interruption), build and operate thermal peakers as a back-up source of energy, or over-build intermittent renewable capacity to ensure sufficient supply. In each case, extra transmission may be needed to move power from the newly built generation to demand loads. The main contribution of this paper is a model which can be used to correctly size the investment and analyse the operation of each option.

A key challenge when analysing the operation of a hydro-dominated system is the trade-off between the value of using the water stored in a reservoir in the current period, compared with the value of keeping the water for a future period. Most commonly, the operation of the system is modelled as a multistage stochastic program and solved using stochastic dual dynamic programming (SDDP) (Pereira & Pinto, 1991). One such example is the Brazilian energy system, which has used SDDP to plan its operation for over 25 years (Maceira et al., 2008, 2018). Because of the use of both hydro and thermal generation, the problem is often called the *hydro-thermal scheduling problem*.

A number of authors have augmented multistage stochastic operational models with capacity expansion decisions. As a brief sampling of the literature, we point to the following works (Newham, 2008; Rebennack, 2014; Wu et al., 2016; PSR-Inc., 2023; Bruno et al., 2016; Thomé

et al., 2019; Lara et al., 2020; Bødal et al., 2022; Jacobson et al., 2023). The literature can be categorized into two broad streams, which we refer to as *static* and *dynamic* investment models.

In *static* investment models, the capacity investment and operation are treated as a two-stage problem, in which the agent chooses a level of new capacity investment, and then evaluates the operation of the energy system with the investment complete. Static models focus on the ideal end-state of the energy system, but not the timing or sequence of decisions that are required to get there. In its simplest form, a static model enumerates all potential investments and evaluates their capital and operating costs before choosing the best. More sophisticated static models treat the operation of the energy system as the subproblem of a two-stage model that has first-stage capacity investment decisions (modelled by binary or continuous variables) optimized using a form of decomposition. The second-stage is solved, often using SDDP, to guide the solution of the first stage. A unifying feature of static models is that they decompose the investment and operational decisions, and that their operational problems model a finite time horizon. The use of a finite-horizon model to evaluate the steady-state operational cost of the system forces models to use excessively long horizons to mitigate end-of-horizon effects. Moreover, because these models use decomposition, the authors each code a customized algorithm to solve their particular problem.

*Dynamic* investment models co-optimize the capacity investment and operation of the energy system over time. In contrast to static models, changes in capacity can occur more than once over the time horizon. Dynamic models are usually formulated as monolithic multistage stochastic programs, typically with integer decisions for the investments. Examples of such models are Dominguez et al. (2016), Liu et al. (2017), and Backe et al. (2022). Because of the difficulty in solving large instances, various approximations or heuristics are used to generate sub-optimal solutions. A unifying feature of dynamic models is that they model a finite time horizon.

In this paper, we propose a static capacity investment model that co-optimizes the strategic level investment of new capacity to minimize the steady-state operational cost of how the energy system should be operated in the face of uncertain inflows and quantity of variable renewable energy. Our model is based on the policy graph modelling framework of Dowson (2020) and can be solved using existing open-source multistage stochastic programming solvers such as SDDP.jl (Dowson & Kapelevich, 2021) without the need for the practitioner to code their own specialized decomposition algorithm. The single, co-optimized, infinite-horizon model formulation is

the main novel contribution of this paper. We apply our model to a case study that investigates the capacity expansion of variable renewable energy in the New Zealand electricity system. We consider different investment scenarios, and to verify that the method works, we compare our optimal solutions with those computed by enumeration by [Philpott & Downward \(2023\)](#). The ability to solve our model without the need to implement custom decomposition algorithms is particularly important, because it means that our approach can easily be applied by other researchers to other energy systems and case studies. Thus, this paper also serves as a demonstration of the flexibility of the policy graph modelling approach and of SDDP.jl.

The paper is laid out as follows. In the next section we describe the proposed method for the integration of capacity investments in SDDP.jl. In [Section 3](#) we describe how we model wind investment in SDDP.jl. In [Section 4](#) we give a short description of the New Zealand case study, and compare the results from our investment model and previous work in [Section 5](#), before we conclude in [Section 6](#). [Appendix A](#) gives the full formulation of the SDDP stage problem and [Appendix B](#) provides parameters for the generation plants modelled in the case study. More detailed results from the case study are presented in [Appendix C](#) and [Appendix D](#).

## 2. Modelling investments using a policy graph

In this section we introduce the hydro-thermal scheduling problem, explain how to model the finite- and infinite-horizon variants as a policy graph, and expand the policy graph to include investment decisions.

The classical version of the hydro-thermal scheduling model is a finite-horizon discrete-time stochastic optimal control problem. The *stages* of this problem are indexed  $1, 2, \dots, T$ . The *states* of this problem are reservoir levels measured at the end of each stage (denoted  $X$ ) and *actions* (denoted  $U$ ) are water releases through generating turbines, flows through spill and river arcs, dispatchable non-hydro generation, and load reduction. In each stage, each reservoir  $r$  experiences a random inflow  $\omega_r(t)$ . The upper-case notation used for  $X$  and  $U$  indicate that these are also random variables.

The general formulation of this hydro-thermal scheduling problem (HTP) is as follows:

$$\begin{aligned}
\mathbf{HTP} : \min \quad & \mathbb{E}[\sum_{t=1}^T C_t(X(t), U(t))] \\
\text{s.t.} \quad & X(t) = f_t(X(t-1), U(t), \omega(t)) \quad t = 1, 2, \dots, T \\
& X(0) = \bar{x} \\
& U(t) \in \mathcal{U}_t(x(t-1), \omega(t)) \quad t = 1, 2, \dots, T \\
& X(t) \in \mathcal{X}_t \quad t = 1, 2, \dots, T.
\end{aligned}$$

Here,  $C_t(X(t), U(t))$  is the cost of meeting demand (possibly with load shedding) in stage  $t$  from dispatchable generation. The finite horizon ignores actions after stage  $T$ , but future costs and constraints on  $X(T)$  can be modelled by suitable choices of  $C_T(X(T), U(T))$  and  $\mathcal{X}_T$ . Note that these exogeneous modelling choices will affect the optimal solution of **HTP**. We revisit this issue below.

The transition function  $f_t(X(t-1), U(t), \omega(t))$  maps the incoming state  $X(t-1)$  at the beginning of stage  $t$  to a random outgoing state  $X(t)$  at the end of stage  $t$ . For example:

$$X(t) = f_t(X(t-1), U(t), \omega(t)) = X(t-1) + B \cdot U(t) + \omega(t),$$

where the matrix  $B$  maps the control flows  $U(t)$  to the change in reservoir level. The vector of inflows  $\omega(t)$  is assumed to have a finite probability distribution (drawn from the sample space  $\Omega(t)$ ) that is independent of that in any previous stage. Observe that random variables  $X$  and  $U$  are required to be measurable with respect to the history of the random inflow process, so they must satisfy *non-anticipative* constraints that we have suppressed in the formulation.

The model **HTP** can be defined by a *policy graph* as described in Dowson (2020). The policy graph defines the dynamic structure of the decision problem we are modelling, showing how actions affect the state variables and how realizations of the random variables are revealed over time. The policy graph that represents **HTP** is shown in Figure 1.



Figure 1: The policy graph structure for **HTP**.

There are  $T$  stages with subproblems  $\mathbf{SP}_t$ , where each  $t$  is a time step when decisions are made. The random variable  $\omega(t)$  is shown as squiggly lines in the policy graph. Each subproblem  $\mathbf{SP}_t$  has the following form:

$$\begin{aligned}
\mathbf{SP}_t(x(t-1), \omega(t)) : \min_{u(t), x(t)} \quad & C_t(x(t), u(t), \omega(t)) \\
\text{s.t.} \quad & x(t) = f_t(x(t-1), u(t), \omega(t)) \\
& u(t) \in \mathcal{U}_t(x(t-1), \omega(t)) \\
& x(t) \in \mathcal{X}_t.
\end{aligned}$$

The need to choose a suitable terminal cost function  $C_T$  can be mitigated by solving the discounted infinite-horizon problem:

$$\begin{aligned}
\mathbf{HTP} - \infty : \min \quad & \mathbb{E}[\sum_{t=1}^{\infty} \rho^t C_t(X(t), U(t))] \\
\text{s.t.} \quad & X(t) = f_t(X(t-1), U(t), \omega(t)) \quad t = 1, 2, \dots \\
& X(0) = \bar{x} \\
& U(t) \in \mathcal{U}_t(x(t-1), \omega(t)) \quad t = 1, 2, \dots \\
& X(t) \in \mathcal{X}_t \quad t = 1, 2, \dots
\end{aligned}$$

Infinite-horizon discounted cost models can be represented by policy graphs that contain cycles. An example is shown in Figure 2. In each step, the probability of transitioning from node  $t$  to  $t+1$  (or from node  $T$  to node 1) is  $\rho$  where  $\rho < 1$ . Most commonly,  $T$  is chosen so that one loop of the graph represents one year, and the nodes within each loop allow for seasonality in demand and inflows over the course of each year. Importantly, the subproblems  $\mathbf{SP}_t$  do not change from the finite-horizon model, only the structure of the graph changes.

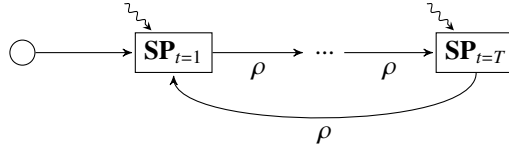


Figure 2: The policy graph structure for  $\mathbf{HTP} - \infty$ .

We can extend the  $\mathbf{HTP} - \infty$  problem to include an investment variable  $u_{inv}$  as follows:

$$\begin{aligned}
\mathbf{INV} - \mathbf{HTP} - \infty : \min \quad & c_{inv}^\top u_{inv} + \mathbb{E}[\sum_{t=1}^{\infty} \rho^t C_t(X(t), U(t))] \\
\text{s.t.} \quad & X(t) = f_t(X(t-1), U(t), \omega(t)) \quad t = 1, 2, \dots \\
& X(0) = \bar{x} \\
& U(t) \in \mathcal{U}_t(x(t-1), u_{inv}, \omega(t)) \quad t = 1, 2, \dots \\
& X(t) \in \mathcal{X}_t(u_{inv}) \quad t = 1, 2, \dots
\end{aligned}$$

Here  $u_{inv}$  is the vector of new capacity that is installed before the start of stage  $t = 1$ . The newly installed capacity modifies the feasibility set of the state variables  $\mathcal{X}_t$ , and it may also modify the feasibility set of the control variable  $\mathcal{U}_t$ . Here  $c_{inv}$  is the overnight investment cost  $I$  per unit of  $u_{inv}$  [\$/MW], assuming that the invested capacity has an infinite lifetime. If the lifetime is finite, say  $\tau$  years, we assume that the capacity is reinvested every  $\tau$  years and set:

$$c_{inv} = I + \sum_{i=1}^{\infty} I \cdot \beta^{i\tau} = \frac{I}{1 - \beta^\tau}.$$

The policy graph in Figure 2 can be amended to accommodate capacity expansion decisions at the beginning of the infinite operating horizon. This gives the policy graph shown in Figure 3. The decisions in  $\mathbf{SP}_{inv}$  determine the system capacities, and the decisions in  $\mathbf{SP}_t$  are operating decisions optimizing the control of the resulting system over an infinite horizon.

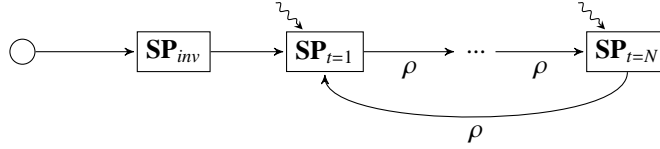


Figure 3: The policy graph structure for  $\text{INV} - \text{HTP} - \infty$ .

In order to pass the capacity decisions from  $\mathbf{SP}_{inv}$  into the cyclic policy graph, we add a state variable,  $x_{inv}$ , which is set by the investment control variable  $u_{inv}$ . Subproblem  $\mathbf{SP}_{inv}$  is as follows:

$$\begin{aligned} \mathbf{SP}_{inv} : \min_{u_{inv}} \quad & c_{inv}^\top u_{inv} \\ \text{s.t.} \quad & u_{inv} \geq 0 \\ & x_{inv}(0) = u_{inv} \end{aligned}$$

The new  $x_{inv}$  state variable changes the constraints of each subproblem  $\mathbf{SP}_t$ . For dispatchable plants, investment increases the maximum output in any dispatch period. The same holds for grid investments, where the maximum power flow on a transmission line is increased. The effect of investment on the output of intermittent capacity such as wind or solar power is less straightforward. In the next section, we describe a methodology that approximates wind generation from invested wind capacity for use in a medium term hydro-scheduling model.

### 3. Modelling renewable investments in the subproblems

In the previous section we showed how capacity investments can be modelled in SDDP.jl using a policy graph. In this section we show how to translate this investment decision into capacity constraints on generation in each stage problem. When generation comes from intermittent sources the constraints are stochastic. We illustrate how we approximate these constraints using wind as an example. The same method will apply for solar power.

In each stage problem, wind generation  $w_m(h, t)$  (MW from wind farm  $m$  in hour  $h$  for week  $t$  of the year) is a decision variable that contributes to the electricity supply with a short-run marginal cost equal to 0. The output of wind farm  $m$  in week  $t$  and hour  $h$  depends upon its capacity  $x_m^{inv}$ , but also the (random) amount of wind available at that time. This can be represented using by a random factor  $\lambda_{m,h,t}(\omega_t)$  that limits the output through:

$$w_m(h, t, \omega_t) \leq \lambda_{m,h,t}(\omega_t) \cdot x_m^{inv},$$

where the inequality constraint gives that curtailment of wind generation is possible at zero cost.

In our model we approximate the distribution of  $\lambda_{m,h,t}(\omega_t)$  by an empirical distribution from  $Y$  years of historical wind generation data. The wind farms for which these data are available are grouped into geographical regions  $r$ , and  $\lambda_{m,h,t}(\omega_t)$  is assumed to be the same value  $\lambda_{r,h,t}(\omega_t)$  for each wind farm  $m$  in region  $r$ . In any region  $r$  and hour  $h$  in week  $t$ ,  $\lambda_{r,h,t}(\omega_t)$  can take on one of  $Y$  values with equal probability. Each value in this distribution (corresponding to year  $y \in \{1, 2, \dots, Y\}$ ) is obtained by dividing the observed wind generation in year  $y$  in hour  $h$  in week  $t$  in region  $r$  by the installed wind capacity in that region. This model assumes wind generation in any hour is independent of that in the previous hour. This ignores the short-term temporal dependence of wind generation from hour to hour that might become important if investment in short-term storage were being considered. The standard approach (see e.g. [Liu et al. \(2017\)](#)) in this case is to use “representative days” of wind generation, which model the dynamics at the expense of introducing some perfect foresight into the decision making.

Since each weekly stage problem in the SDDP model is to be solved many times, hourly demand and hourly wind generation are approximated by piecewise constant load duration curves. The standard approach to estimating these assumes a level of wind capacity  $K_r$  for each region  $r \in \mathcal{R}$  and subtracts the wind generation  $\lambda_{r,h,t}(\omega_t)K_r$  from the demand  $D_r(h, t)$  in that region to

give a net demand:

$$\hat{D}_r(h, t, K_r, \omega_t) = D_r(h, t) - \lambda_{r,h,t}(\omega_t)K_r,$$

for region  $r$  in hour  $h$  in week  $t$  corresponding to the wind in historical year  $\omega_t$ . Here  $\lambda_{r,h,t}(\omega_t)$  is the load factor computed from historical wind generation in region  $r$  divided by capacity in year  $\omega_t$ . Then total system net demand  $\sum_{r \in \mathcal{R}} \hat{D}_r(h, t, K_r, \omega_t)$  is sorted to give a duration curve that is approximated by load blocks  $b = 1, 2, \dots, B$ , each having constant net load equal to average net load over the hours and historical years in its block.

It is easy to see that different choices of wind capacity  $K_r$  will produce different values of net demand. Moreover the duration curve obtained by sorting  $\sum_{r \in \mathcal{R}} \hat{D}_r(h, t, K_r, \omega_t)$  will sort the observations from different hours and historical years into a possibly different order for each choice of  $K_r$ . A brute-force approach to capacity optimization would enumerate all potential choices of  $K_r$ ,  $r \in \mathcal{R}$ , and solve an SDDP model with this choice. In our model, the values of  $K_r$ ,  $r \in \mathcal{R}$  are represented by state variables  $x_{inv}$ , which means that they are optimized by SDDP rather than enumerated ex-ante. This requires a different approach.

The first approximation that we make is to assume that  $K_r$  is a fixed proportion  $\alpha_r$  of system wind capacity  $K$ , a single state variable that will be optimized by SDDP. This implies that net demand in region  $r$  is:

$$\hat{D}_r(h, t, \alpha_r K, \omega_t) = D_r(h, t) - \lambda_{r,h,t}(\omega_t)\alpha_r K,$$

and total net demand for wind capacity  $K$  is:

$$\sum_{r \in \mathcal{R}} \hat{D}_r(h, t, \alpha_r K, \omega_t) = \sum_{r \in \mathcal{R}} D_r(h, t) - \sum_{r \in \mathcal{R}} \alpha_r \lambda_{r,h,t}(\omega_t) K.$$

Given a nominal level  $\bar{K}$  of wind capacity we compute  $\sum_{r \in \mathcal{R}} \hat{D}_r(h, t, \alpha_r \bar{K}, \omega_t)$  and sort to give a duration curve that is approximated by load blocks  $b = 1, 2, \dots, B$ . Each load block  $b$  for week  $t$  contains  $T(t, b)$  observations. We fix this number for each load block.

To model the load duration curve as  $K$  varies from  $\bar{K}$ , we make a second approximation. To understand how  $K$  affects the duration curve it is helpful to fix week  $t$  and write:

$$\begin{aligned} \hat{D}(h, t, K, \omega_t) &= \sum_{r \in \mathcal{R}} \hat{D}_r(h, t, \alpha_r \bar{K}, \omega_t) - \sum_{r \in \mathcal{R}} \alpha_r \lambda_{r,h,t}(\omega_t)(K - \bar{K}) \\ &= X(h, \omega_t) - W(h, \omega_t)(K - \bar{K}), \end{aligned}$$

where  $X(h, \omega_t) = \sum_{r \in \mathcal{R}} \hat{D}_r(h, t, \alpha_r \bar{K}, \omega_t)$  and  $W(h, \omega_t) = \sum_{r \in \mathcal{R}} \alpha_r \lambda_{r,h,t}(\omega_t)$ . Consider first the highest load block  $b = 1$  for week  $t$ , containing the  $T(1, t)$  largest values of  $X(h, \omega_t)$ . As  $K$  increases from  $\bar{K}$ , the indices  $(h, \omega_t) \in b$  will remain the same but the average load will decrease until a swap occurs when we reach a critical value  $K_1$ , such that for  $K > K_1$  and some index  $(h^i, \omega_t^i) \in b$ :

$$X(h^i, \omega_t^i) - W(h^i, \omega_t^i)(K - \bar{K}) < X(h^j, \omega_t^j) - W(h^j, \omega_t^j)(K - \bar{K})$$

for  $(h^j, \omega_t^j) \notin b$ , and we swap  $(h^j, \omega_t^j)$  into block 1 and  $(h^i, \omega_t^i)$  out of block 1. When this happens we will have:

$$X(h^i, \omega_t^i) - W(h^i, \omega_t^i)(K_1 - \bar{K}) = X(h^j, \omega_t^j) - W(h^j, \omega_t^j)(K_1 - \bar{K})$$

so  $\hat{D}(h, t, K, \omega_t)$  is continuous in  $K$ . (Observe that  $X$  and  $W$  are not continuous in  $K$ , since they can both jump at  $K_1$ .) Since wind investment reduces net demand, we have  $\hat{D}(h^i, t, K, \omega_t^i)$  is decreasing in  $K$ , and:

$$X(h^i, \omega_t^i) - W(h^i, \omega_t^i)(K_1 - \bar{K}) = X(h^j, \omega_t^j) - W(h^j, \omega_t^j)(K_1 - \bar{K})$$

implies:

$$W(h^i, \omega_t^i) > W(h^j, \omega_t^j).$$

This implies that  $\sum_{(h, \omega_t) \in 1} \hat{D}(h, t, K, \omega_t)$  is a convex decreasing piecewise linear function of  $K$ . This enables us to model the effect of wind in block  $b = 1$  using linear inequality constraints (which is required for SDDP). Unfortunately the above analysis does not apply to the other blocks. Indeed it can be shown that for the lowest demand block,  $B(t)$ , we have  $\sum_{(h, \omega_t) \in B(t)} \hat{D}(h, t, K, \omega_t)$  is a *concave* decreasing piecewise linear function of  $K$ , which cannot be represented in SDDP using linear inequality constraints.

Our approximation for week  $t$  takes the load blocks constructed for  $\sum_{r \in \mathcal{R}} \hat{D}_r(h, t, K, \omega_t)$  for five different values of  $K$  ranging from 1 to 5 GW, where each load block  $b$  is restricted to contain exactly  $T(t, b)$  observations. (We set  $Y = 1$  in our experiments meaning that  $\omega_t$  is ignored and  $\sum_{b=1}^B T(t, b) = 168$ .) The observations (hours) that are in block  $b$  for capacity choice  $K$  are denoted  $b(K)$ . The average extra wind generation in this block is

$$\begin{aligned} & \frac{\sum_{h \in b(K)} W(h, \omega_t)}{T(t, b)} K - \frac{\sum_{h \in b} W(h, \omega_t)}{T(t, b)} \bar{K} \\ &= \sum_{r \in \mathcal{R}} \alpha_r \left( \frac{\sum_{h \in b(K)} \lambda_{r,h,t}(\omega_t)}{T(t, b)} K - \frac{\sum_{h \in b} \lambda_{r,h,t}(\omega_t)}{T(t, b)} \bar{K} \right) \end{aligned}$$

Recall that  $\frac{\sum_{h \in b(K)} \lambda_{r,h,t}(\omega_t)}{T(t,b)} K$  is not continuous in  $K$ , since it can jump up when  $b(K)$  changes. We construct a continuous linear approximation by fitting a straight line through  $(K, \frac{\alpha_r K \sum_{h \in b(K)} \lambda_{r,h,t}(\omega_t)}{T(t,b)})$  for five different values of  $K$  ranging from 1 to 5 GW. The slope  $\mu_{r,b,t}$  of this line gives an estimate of the change in wind generation in region  $r$  for a unit increase in national wind capacity. We can then model wind generation  $w_r(t, b)$  in block  $b$  and week  $t$  in SDDP using the investment state  $x_{inv}$  and the constraint:

$$w_r(t, b) \leq \mu_{r,b,t} x_{inv}.$$

The approximation can be illustrated by showing how it applies to some regions in New Zealand, as a preview to the next section where we discuss the New Zealand system in more detail. Figure 4 shows the fitted linear functions representing wind generation for the Otago region of New Zealand for the first (peak), third and fifth load block. A linear regression gives a

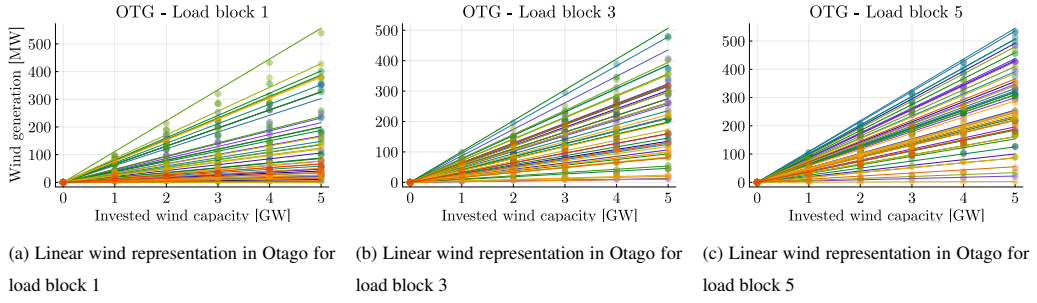


Figure 4: Otago wind generation (MW) versus  $K$  (GW), where load block sizes  $T(b, t)$  are determined using  $\bar{K}=2.5$  GW. Different colours correspond to different weeks of the year.

good representation of the wind generation in each load block at least at the five data points we have chosen. The figure also shows that for most weeks, more of the wind generation is assigned to the load blocks with less demand (the average generation across the weeks shifts upwards as the block number increases). In fact, for all the regions the yearly mean wind generation is always increasing going from the first block to block two, three, four and five. Figure 5 shows the data points and the corresponding linear regression for the Wellington region. This has a poorer fit, and a few of the estimated slopes in load block 1 are negative. They end up negative because  $W$  is discontinuous in  $K$  (possibly) jumping downwards when  $b(K)$  changes. The impact of this is strongest in the Wellington region because the value of  $K_r/K$  is large compared to other regions in the model.

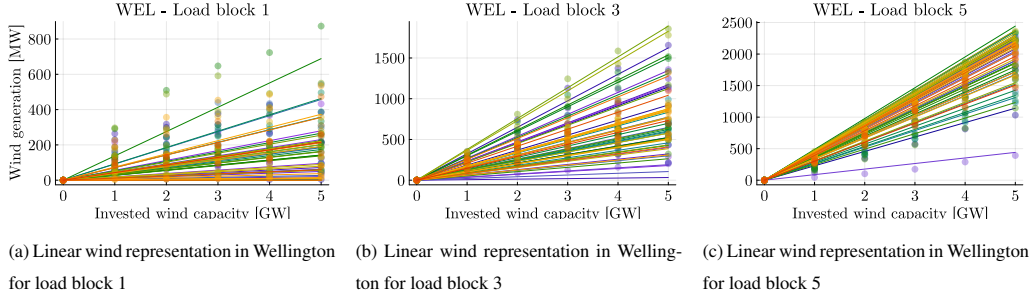


Figure 5: Wellington wind generation (MW) versus  $K$  (GW), where load block indices are determined using  $\bar{K}=2.5$  GW. Different colours correspond to different weeks of the year.

#### 4. A New Zealand case study

In this section, we show the results of applying our model to the New Zealand electricity system. We do this by amending JADE, an existing open-source SDDP model of the New Zealand electricity system ([Electric Power Optimization Centre, 2021](#)). JADE is written in Julia ([Bezan-son et al., 2017](#)) using the JuMP ([Lubin et al., 2023](#)) and SDDP.jl ([Dowson & Kapelevich, 2021](#)) packages, and it is distributed by the New Zealand Electricity Authority. For this paper, we extended JADE to add the investment decisions and a wind generation model as described in the previous sections.

The New Zealand electricity system is spread over two islands as shown in Figure 6. The details of this system can be found in, for example, [Philpott et al. \(2019\)](#) and so we only give a brief description here. As in [Philpott & Downward \(2023\)](#), JADE in the current paper approximates the full 250-node transmission network by an 11-node model as shown in the right-hand panel. Generation capacities and demand in each island for 2017 are also shown in Figure 6. The islands are joined by a HVDC cable of capacity 1050 MW (shown by the dashed line in Figure 6).

JADE has the ability to model arbitrarily many state variables but following [Philpott & Downward \(2023\)](#) we limit these to storage volumes in Lake Taupo at node WTO, storage volumes in Lakes Tekapo, Ohau and Pukaki at node CAN, and storage volumes in Lake Hawea and Lakes Manapouri-Te Anau at node OTG. The proposed Lake Onslow reservoir is located in node OTG. The details of the power plants associated with these reservoirs are given in [Appendix B](#), and the mathematical formulation of JADE is provided in [Appendix A](#).

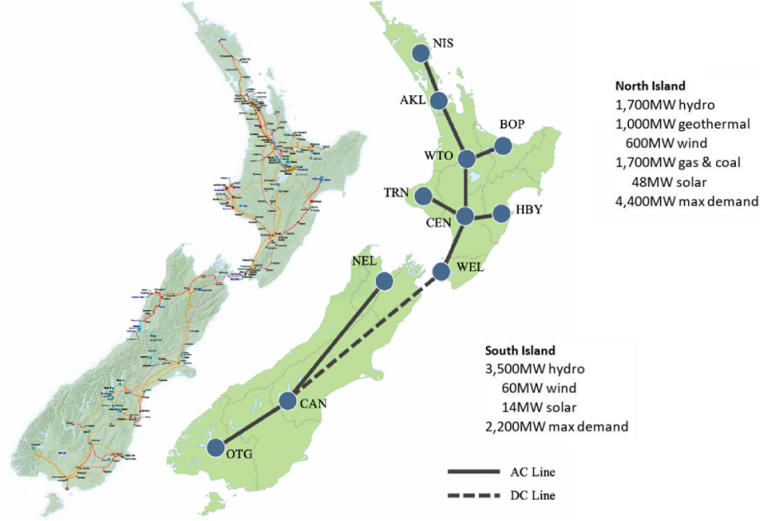


Figure 6: The New Zealand electricity system, showing major transmission lines and the 11 node approximation used in the JADE model.

Our use of JADE enables a straightforward comparison to be made with the wind investment decisions optimized by [Philpott & Downward \(2023\)](#) using enumeration in JADE. They considered three alternative scenarios for a fully renewable electricity system in 2035, in which all capacity decisions were determined up front apart from wind capacity that was optimized. In all three scenarios existing non-renewable electricity plants were closed. In the first case a large pumped hydro storage facility *Lake Onslow* (1.5 GW and 5 TWh), is built to balance the system. The second alternative includes a zero-emission green peaker plant, and the third case assumes existing renewable capacity. We refer to these three cases as *Onslow*, *Peakers* and *Wind only* from now on. The investments in new wind capacity for each case were then found by solving JADE a number of times with different capacity levels and selecting the capacity choice yielding a generation weighted average price (GWAP) equal to the LCOE for the new wind power plants. The results of these experiments provide a useful benchmark for the investments generated by our method.

Of the 11 regions shown in Figure 6, seven have consented wind farms. We allocate new wind capacity proportional to the existing capacity in these regions, as shown in Table 1. This means that the state variable  $x_{inv}$  is implemented as the national wind capacity  $K$ , and the regional

shares of this capacity are fixed to  $\alpha_r K$ .

Region	$\alpha_r$	Representative wind power plant	Capacity [MW]
CAN	0.0814	White Hill	58
CEN	0.0627	Tararua Stage 3	93
HBV	0.0914	Te Uku	28
OTG	0.1926	White Hill	58
TRN	0.000	Waipipi	133
WEL	0.5670	West Wind	142.6
WTO	0.0049	Te Uku	28

Table 1: Representative wind farms and  $\alpha_r$  for each region with wind.

Recall that the levelized cost of energy (LCOE) is defined as:

$$\text{LCOE} = \frac{\text{Lifetime cost of capacity}}{\text{Lifetime energy produced}}.$$

If we denote the overnight capital cost of 1 MW of wind by  $I$ , its lifetime by  $\tau$  years, its capacity factor by  $\eta$ , and  $H$  as the number of hours in a year (8760), then (ignoring maintenance costs):

$$\text{LCOE} = \frac{I}{H \cdot \eta \cdot (1 + \beta + \beta^2 + \dots + \beta^{\tau-1})}.$$

Thus:

$$I = \text{LCOE} \cdot H \cdot \eta \cdot \frac{1 - \beta^\tau}{1 - \beta}.$$

Following [Philpott & Downward \(2023\)](#), we assume  $\text{LCOE} = 65 \text{ NZD/MWh}$ . The national wind capacity factor  $\eta$  is estimated from the wind profiles to be 0.355, and we assume  $\tau = 20$ . When  $\beta = 0.9$ , this gives an overnight investment cost for our infinite-horizon model of  $I = 1.78 \times 10^6 \text{ NZD/MW}$ .

## 5. Results

In this section, we present the results of applying our model to the New Zealand case study briefly introduced in the previous section. The main results presented in this paper are the resulting invested wind capacities for different model runs with our altered model.

### 5.1. Comparison to Philpott & Downward (2023)

Figure 7 compares the investment decisions from our model with the levels published in Philpott & Downward (2023), which were found by re-solving JADE with different investment levels and choosing the best (shown by the horizontal, dotted lines). In each of the three cases, our model is trained for ten SDDP iterations, and then the investment decision for the national wind capacity is evaluated. This is repeated for a total of 1000 SDDP iterations, adding new cuts to the investment problem for ten consecutive forward and backward passes at the time. The plot shows that the endogenous investment decision approaches the benchmark investment levels as the model is trained with more iterations.

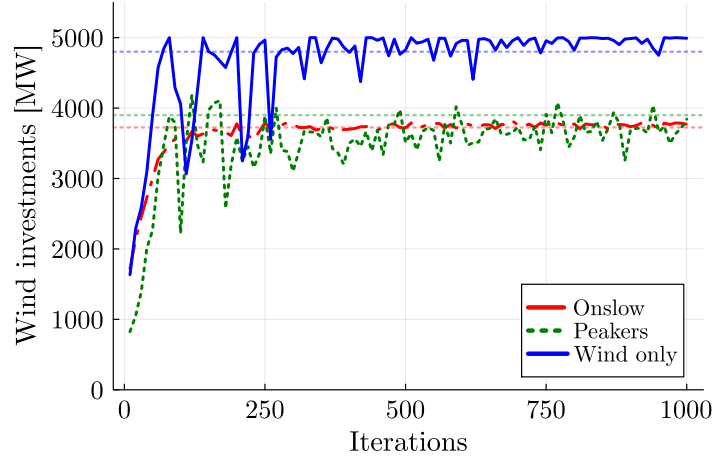


Figure 7: Investment decisions for new wind capacity in the three different cases against the number of SDDP training iterations. Horizontal, dotted lines are those found by Philpott & Downward (2023).

In general, the results in Figure 7 indicates that our method for integrating the investment decision works, with resulting wind investments close to the ones found in the previous study. We would not expect the optimal investments in our JADE model to match exactly the optimal investments in the JADE model used in Philpott & Downward (2023) for several reasons. First, the linear wind representation described in Section 3 is a simplification of the affect that wind investment will have on residual load, compared with Philpott & Downward (2023) where the load duration curve is re-computed for each level of investment. Second, our model represents tranches of load shedding as a proportion of original load, whereas Philpott & Downward (2023) treats these as proportions of residual load (after wind has been subtracted). Finally, Philpott &

Downward (2023) uses round numbers when enumerating different levels of wind investment, whereas the investment decision in our model is a continuous variable.

Despite some differences between the two versions of the model, the resulting investments are similar. Evaluating the policy every tenth iteration however, introduces some randomness to Figure 7. It is easy to see that the investments improve as more cuts are added to the policy, but the plot does not prove convergence in any way. Figure 8 shows the lower bound and the rolling mean of the simulation values from the forward pass for all of the three cases. We terminated the training loop of SDDP.jl after a fixed iteration limit of 1000 iterations.

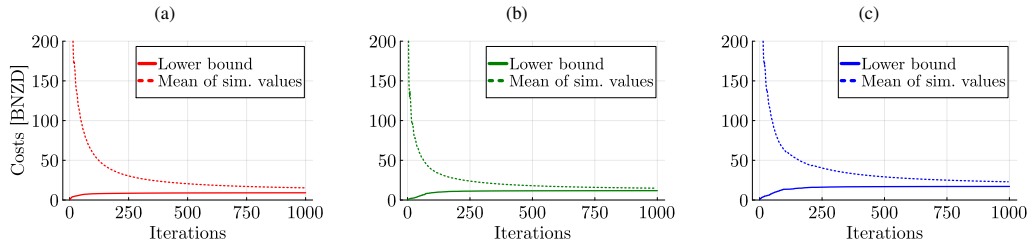


Figure 8: Plots of the lower bound and the rolling mean of simulation values from the forward pass for each of the three cases when the model is trained with 1000 iterations: (a) *Onslow*; (b) *Peakers*; and (c) *Wind only*

## 5.2. Incorporating HVDC and peaker capacity

One of the strengths of integrating the investment decisions in the model compared to the method from Philpott & Downward (2023) is that one can easily optimize more investment types, at the cost of increased computational effort. To illustrate this, we conduct an experiment where the wind capacity is split into North Island and South Island investments. In addition, we let the model optimize the capacity of the green peaker in the *Peakers* case, and the HVDC transmission capacity between the North Island and the South Island of New Zealand in all of the cases. Thus, the  $x_{inv}$  state variable is now a vector with four elements, increasing the number of subproblems solved in training the policy with 1000 iterations from 11.6 million to 15.4 million. Solving such a case by enumeration is much more difficult.

Table 2 summarizes the results from the extended investment model after training with 1000 iterations. The total wind investment is lower in all cases than the values in Figure 7. Multiple factors are contributing to this outcome. In our previous model we assumed that the capacity on

Investments and costs	<i>Onslow</i>	<i>Peakers</i>	<i>Wind only</i>
Wind - North Island [MW]	1315	1458	1997
Wind - South Island [MW]	0	0	0
HVDC capacity [MW]	643	0	379
Green peaker [MW]	-	837	-
CAPEX [MNZD]	4633	3671	5047
OPEX [MNZD]	2128	2219	3775
Total [MNZD]	6761	5890	8822

Table 2: Resulting investments and costs from the extended investment model. *CAPEX* is the cost of the deterministic investment node. *OPEX* is the SDDP lower bound less the CAPEX cost.

the HVDC connection between the South Island and the North Island was sufficient to meet all transfers. This enabled a comparison with [Philpott & Downward \(2023\)](#) as shown in Figure 7. The new model optimizes the HVDC capacity, with the existing capacity as the starting point. Thus, optimal wind investments are lower when wind investments needs to come with following grid investments to access storage in the South Island. In this case, shedding more load is cost efficient compared to extensive grid expansion. Also, as load centers are located on the North Island (and the capacity factor of the representative wind power plants on the North Island is higher than on the South Island), all wind investments are made in the North Island only.

Table 2 shows that increased HVDC capacity is most beneficial in the *Onslow* case, because the Lake Onslow pumped hydro storage is located in the South Island and requires grid capacity to provide its flexibility to the North Island.

In the *Peakers* case, the optimal value of peaking capacity determined by the new model is significantly higher than that assumed in [Philpott & Downward \(2023\)](#). The *Wind only* case has the highest CAPEX, because of the larger wind investments, and the highest OPEX, because of load shedding, of the three cases. The investment costs from building Lake Onslow (estimated in [MBIE: New Zealand Battery Project \(2022\)](#) to be 15700 MNZD) are not included in the CAPEX cost for the *Onslow* case. Thus the total cost for the *Onslow* case is 22461 MNZD, which is almost four times the cost for the cheapest alternative, namely the *Peakers* case.

Figure 9 shows the utilization of the HVDC cable in each investment case over the 31 historical inflow years 1990–2020. The resulting HVDC capacities are shown as the horizontal, dotted

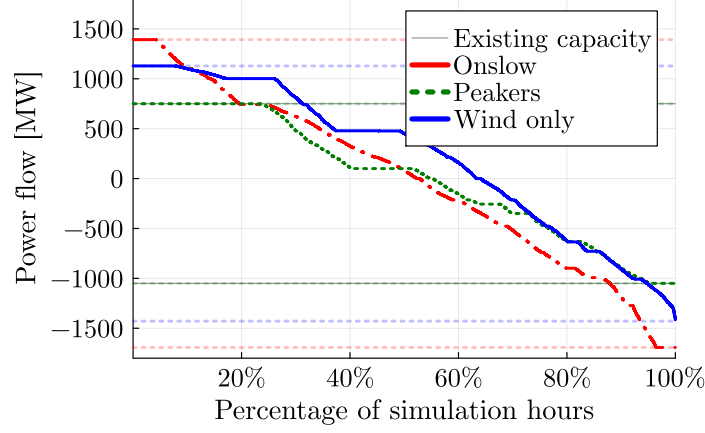


Figure 9: Flow-duration curve for flow on the HVDC line from WEL (North Island) to CAN (South Island) in the three different cases when simulated over the 31 historical inflow years 1990–2020. The horizontal, dotted lines shows the maximum capacity of the lines in each case, which is the sum of the existing capacity (the gray line) and the invested grid capacity.

lines, while the three plotted lines display the duration curve for the load flow on the HVDC connection with positive load flow defined as flow from the Wellington region (North Island) to the Canterbury region (South Island). The green, horizontal line is equal to the existing capacity (gray line) as now grid investments are made in the *Peakers* case. The  $x$ -axis gives the percentage of simulated hours where load flow is equal to or greater than  $y$  MW. In all cases the utilization of the transmission line goes both ways, with the largest flows in the *Onslow* case, where North to South flow is needed for pumping and reverse flow transports Onslow generation north. In the two other cases, HVDC investments are lower, and the full capacity from the North Island to the South Island is utilized more than the full reverse flow.

Figure 9 is complemented by Figure 10, which visualizes the utilization of the six (five without Lake Onslow) main reservoirs in the system. The figure shows that four of the five reservoirs are operated with lower storage levels when other sources of flexibility are in place, like Lake Onslow or a green peaker plant, hence reducing the risk of spillage.

Figure 11 plots trajectories of marginal water values in the hydro reservoirs. These values can be taken as an indication of spot energy prices when these are determined by energy constraints rather than capacity constraints. Marginal water values are very low in the *Wind only* case,

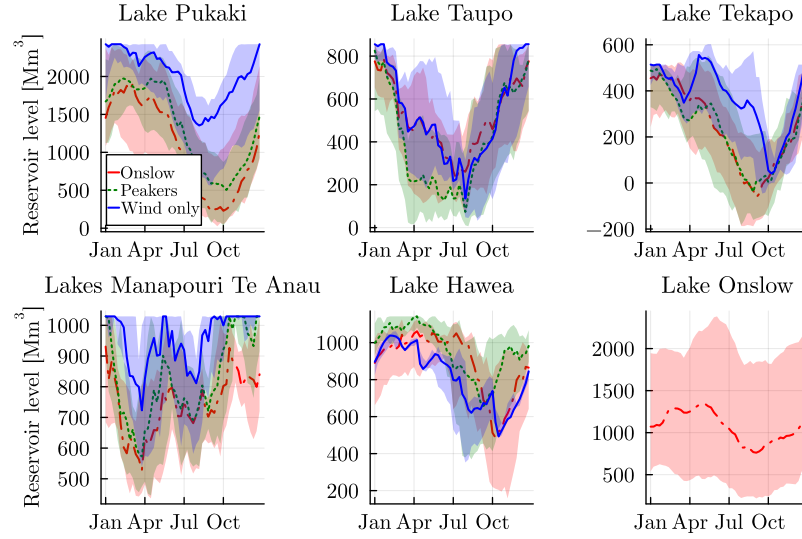


Figure 10: Hydro reservoir levels relative to consented minimum operating levels for the three different cases when simulated over the 31 historical inflow years 1990–2020. The lines are the median reservoir level and each coloured band depicts the 25th and 75th percentile levels. Some reservoirs can have negative storage when the reservoir level falls below the consented operating minimum (incurring a cost penalty).

and higher in the other two cases. In the *Wind only* case, prices are set by periods when the wind is not blowing and load must be shed because conventional capacity and/or grid capacity is insufficient, even though energy in reservoir storage is plentiful giving low marginal water values. In the other two cases, we have increased dispatchable capacity, and prices are set by anticipated energy shortages that are reflected by higher marginal water values. More detailed plots of the storage trajectories and marginal water values for the 31 historical inflow years from 1990–2020 can be found in [Appendix C](#) and [Appendix D](#).

When demand cannot be met, the system must shed load. Figure 12 shows the probability of shedding more than  $y$  MWh of load in a random week. In all three cases, more than 10 percent of the 1612 simulated weeks have some load shedding.

The *Wind only* case sheds load in a large proportion of weeks, in periods when the amount of wind generation is unable to meet demand even if all other generators are at full capacity. This means prices are set by shortage costs while marginal water values are generally lower (since generation or grid capacity constraints mean extra water cannot be used to offset this shortage). The prices in these periods provide revenue for wind investment; in other periods prices are zero.

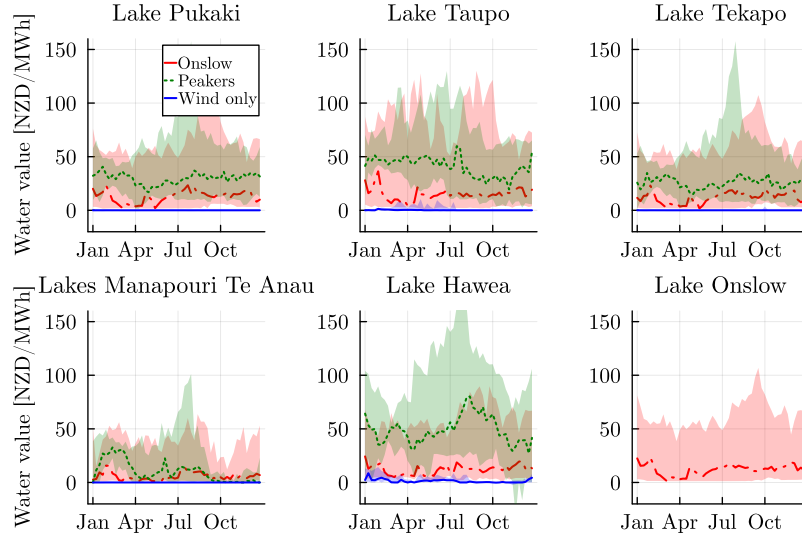


Figure 11: Marginal water values for different reservoirs in the three different cases when simulated over the 31 historical inflow years 1990–2020. The lines are the median marginal water value and each coloured band depicts the 25th and 75th percentiles.

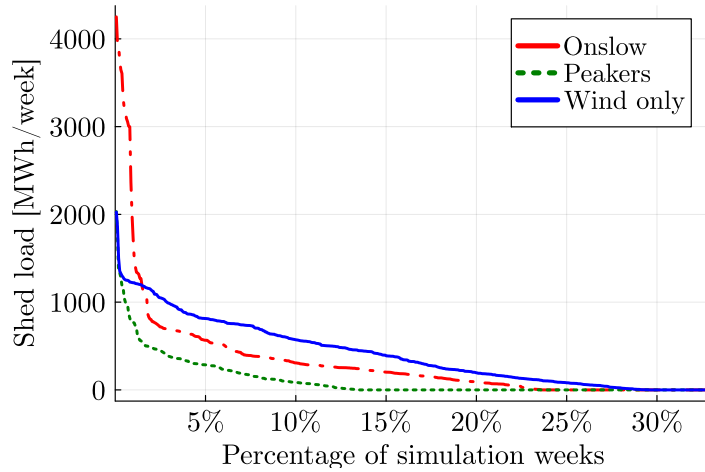


Figure 12: Duration curves of the weekly load shedding in the three cases when simulated over the 31 historical inflow years 1990–2020.

The *Peakers* case provides dispatchable generation capacity to reduce the number of shortage periods. The green trajectory in Figure 12 shows that this results in less load shedding from capacity constraints. Some of the investment that was made in the *Wind only* case is diverted to

peakers. The availability of more dispatchable plant makes reservoir operation less conservative making energy shortages more likely, and marginal water values increase.

The *Onslow* case also provides dispatchable generation capacity to reduce the number of shortage periods so the red trajectory in Figure 12 lies below the blue, at least for nearly every week in the simulation. Observe that there are a small number of weeks in the simulations when the load shedding is much higher. As observed in Philpott & Downward (2023), these cases occur when Onslow generation is required but the lake has been emptied by a dry period in the previous year. Marginal water values and load shedding become very high in these weeks.

## 6. Conclusions

In this paper, we have shown how capacity expansion investments can be included in an infinite-horizon operational SDDP model by augmenting the policy graph to include an investment node. We derived a linear representation of wind generation in different load blocks to link the investment decisions with the subproblem constraints. Additionally, by applying our method to the New Zealand energy system, we have shown that our proposed method is flexible, and that it can be used to analyse a variety of options for expansion within the same system. Because our model is based on the policy graph and SDDP.jl, it did not require us to code a customized decomposition algorithm, which makes our model easier to adopt and implement than alternatives in the literature.

The results from our case study show that, with the assumptions we made, a peaker plant located in the North Island is a more cost efficient supplement to wind investments than the *Lake Onslow* pumped hydro storage.<sup>1</sup>

## Acknowledgements

J. Hole was supported by the Norwegian Water Resources and Energy Directorate (NVE), the Department of Electric Energy at the Norwegian University of Science and Technology (NTNU) and the NZ MBIE Catalyst grant UOCX2117 - New Zealand-German Platform for Green Hydrogen Integration (HINT). The work is associated with the two Centres for Environment-friendly Energy Research FME CINELDI (RCN grant 257626) and FME NTRANS (RCN grant 296205).

---

<sup>1</sup>As we were finishing writing this article, the New Zealand Government announced that they will scrap the plans to build Lake Onslow (Inside Government NZ, 2023).

## References

- Backe, S., Skar, C., del Granado, P. C., Turgut, O., & Tomasgard, A. (2022). Empire: An open-source model based on multi-horizon programming for energy transition analyses. *SoftwareX*, 17, 100877.
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, 59, 65–98.
- Bruno, S., Ahmed, S., Shapiro, A., & Street, A. (2016). Risk neutral and risk averse approaches to multistage renewable investment planning under uncertainty. *European Journal of Operational Research*, (pp. 979–989).
- Bødal, E., Botterud, A., & Korpås, M. (2022). Capacity expansion planning with stochastic rolling horizon dispatch. *Electric Power Systems Research*, 205, 107729. doi:[10.1016/j.epsr.2021.107729](https://doi.org/10.1016/j.epsr.2021.107729).
- Clarke, L., Wei, Y.-M., De La Vega Navarro, A., Garg, A., Hahmann, A., Khennas, S., Azevedo, I., Löschel, A., Singh, A., Steg, L., Strbac, G., & Wada, K. (2022). Climate change 2022: Mitigation of climate change. In P. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, & J. Malley (Eds.), *Climate Change 2022: Mitigation of Climate Change* chapter Energy Systems. Cambridge, UK and New York, NY, USA: Cambridge University Press. doi:[10.1017/9781009157926.008](https://doi.org/10.1017/9781009157926.008).
- Dominguez, R., Conejo, A. J., & Carrion, M. (2016). Investing in generation capacity: A multi-stage linear-decision-rule approach. *IEEE Transactions on Power Systems*, 31, 4784–4794.
- Dowson, O. (2020). The policy graph decomposition of multistage stochastic optimization problems. *Networks*, 76, 3–23. doi:<https://doi.org/10.1002/net.21932>.
- Dowson, O., & Kapelevich, L. (2021). SDDP.jl: a Julia package for stochastic dual dynamic programming. *INFORMS Journal on Computing*, 33, 27–33.
- Electric Power Optimization Centre (2021). JADE.jl: A Julia DOASA Environment. <https://github.com/EPoC-NZ/JADE.jl>. Accessed: 2023-07-07.
- Inside Government NZ (2023). Minister sinks ‘hugely wasteful’ Lake Onslow hydro scheme. <https://insidegovernment.co.nz/minister-sinks-hugely-wasteful-lake-onslow-hydro-scheme>. [Accessed 2023-12-06].
- Jacobson, A., Pecci, F., Sepulveda, N., Xu, Q., & Jenkins, J. (2023). A computationally efficient benders decomposition for energy systems planning problems with detailed operations and time-coupling constraints. *INFORMS Journal on Optimization*, . doi:<https://doi.org/10.1287/ijoo.2023.0005>. [Ahead of print].
- Lara, C. L., Siirola, J. D., & Grossmann, I. E. (2020). Electric power infrastructure planning under uncertainty: stochastic dual dynamic integer programming (sddip) and parallelization scheme. *Optimization and Engineering*, 21, 1243–1281.
- Liu, Y., Sioshansi, R., & Conejo, A. J. (2017). Multistage stochastic investment planning with multiscale representation of uncertainties and decisions. *IEEE Transactions on Power Systems*, 33, 781–791.
- Lubin, M., Dowson, O., Dias Garcia, J., Huchette, J., Legat, B., & Vielma, J. P. (2023). JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation*, . doi:[10.1007/s12532-023-00239-3](https://doi.org/10.1007/s12532-023-00239-3).
- Maceira, M., Penna, D., Diniz, A., Pinto, R., Melo, A., Vasconcellos, C., & Cruz, C. (2018). Twenty years of application

- of stochastic dual dynamic programming in official and agent studies in brazil-main features and improvements on the newave model. In *2018 Power Systems Computation Conference (PSCC)* (pp. 1–7). IEEE.
- Maceira, M. E. P., Duarte, V., Penna, D., Moraes, L., & Melo, A. (2008). Ten years of application of stochastic dual dynamic programming in official and agent studies in brazil-description of the newave program. *16th PSCC, Glasgow, Scotland*, (pp. 14–18).
- MBIE: New Zealand Battery Project (2022). <https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/low-emissions-economy/nz-battery/>.
- New Zealand Electricity Authority (2023). Electricity Authority Wholesale Market Tools. <https://www.emi.ea.govt.nz/Wholesale/Tools/JADE>. [Accessed 2023-12-11].
- Newham, N. (2008). *Power system investment planning using stochastic dual dynamic programming*. Ph.D. thesis The University of Canterbury, New Zealand.
- Pereira, M. V. F., & Pinto, L. M. V. G. (1991). Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52, 359–375.
- Philpott, A., & Downward, A. (2023). *Using JADE to analyse pumped storage in New Zealand*. Technical Report Electric Power Optimization Centre.
- Philpott, A., Read, E., Batstone, S., & Miller, A. (2019). The new zealand electricity market: Challenges of a renewable energy system. *IEEE Power and Energy Magazine*, 17, 43–52.
- PSR-Inc. (2023). OptGen. <https://www.psr-inc.com/en/software/>. [Online; accessed 2023-11-22].
- Rebennack, S. (2014). Generation expansion plannign under uncertainty withh emissions quotas. *Electric Power Systems Research*, (pp. 78–85).
- Thomé, F., Perez, R. C., Okamura, L., Soares, A., & Binato, S. (2019). Stochastic multistage co-optimization of generation and transmission expansion planning. *arXiv preprint arXiv:1910.01640*, .
- Wu, Z., Zeng, P., & Zhang, X.-P. (2016). Two-stage stochastic dual dynamic programming for transmission expansion planning with significant renewable generation and nk criterion. *CSEE Journal of power and energy systems*, 2, 3–10.

## Appendix A. JADE formulation

JADE seeks a policy of electricity generation that meets demand and minimizes the expected cost of thermal generation fuel consumed plus any costs of load reduction. All data are deterministic except for weekly inflows that are assumed to be stagewise independent. The resulting stochastic dynamic programming model is defined as follows. Let  $x_j(t)$  denote the storage in reservoir  $j$  at the *end* of week  $t$ , and let the Bellman function  $C_t(\bar{x}, \omega(t))$  be the minimum discounted expected fuel cost to meet electricity demand in weeks  $t, t+1, \dots$ , when reservoir storage  $x_j(t-1)$  at the start of week  $t$  is equal to  $\bar{x}_j$  and the inflow to reservoir  $j$  in week  $t$  is known to be  $\omega_j(t)$ .

In JADE, a weekly discount factor  $\rho = \beta^{1/52}$  is used when going from any stage to the next, where  $\beta < 1$  is the annual discount factor. This implies that in each stage there is a probability of

$1 - \rho$  of transitioning from stage  $t$  to a zero-node 0 where  $C_0(\bar{x}, \omega(0)) = 0$ .

The Bellman function  $C_t(\bar{x}, \omega(t))$  for week  $t$  is the optimal solution value of the mathematical program:

$$\begin{aligned}
C_t(\bar{x}, \omega(t)) = & \\
\min \quad & \sum_{i \in \mathcal{N}} \sum_b T(b, t) \left( \sum_{m \in \mathcal{F}(i)} \phi_m f_m(b, t) + \sum_{l \in \mathcal{L}(i)} \psi_{lb} z_i(l, b, t) \right) \\
& + \rho \cdot \mathbb{E}[C_{t+1}(x(t), \omega(t+1))] \\
\text{s.t.} \quad & g_i(y(b, t)) + \sum_{m \in \mathcal{F}(i)} f_m(b, t) + \\
& \sum_{m \in \mathcal{H}(i)} \gamma_m h_m(b, t) + \sum_{l \in \mathcal{L}(i)} z_i(l, b, t) = D_i(b, t), \quad i \in \mathcal{N} \\
& x(t) = \bar{x} - S \sum_b T(b, t) (A h(b, t) + A s(b, t) - \omega(t)) \\
& 0 \leq f_m(t) \leq a_m, \quad m \in \mathcal{F}(i), i \in \mathcal{N} \\
& 0 \leq h_m(t) \leq b_m, \quad 0 \leq s_m(t) \leq c_m, \quad m \in \mathcal{H}(i) \\
& 0 \leq x_j(t) \leq r_j, \quad j \in \mathcal{J}, i \in \mathcal{N}, y \in Y.
\end{aligned}$$

This description uses the following indices:

Index	Refers to
$t$	index of week
$i$	node in transmission network
$b$	index of load block
$m$	index of plant
$j$	index of reservoir
$\mathcal{N}$	set of nodes in transmission network
$\mathcal{F}(i)$	set of green peaker plants at node $i$
$\mathcal{H}(i)$	set of hydro plants at node $i$
$\mathcal{L}(i)$	set of load types at node $i$
$\mathcal{J}$	set of reservoirs.

The parameters are:

Symbol	Meaning	Units
$\phi_m$	short-run marginal cost of peaker plant $m$	\$/MWh
$\psi_{lb}$	cost of shedding load type $l$ in load block $b$	\$/MWh
$\gamma_m$	conversion factor for water flow into energy	MWs/m <sup>3</sup>
$D_i(b, t)$	electricity demand in node $i$ in block $b$ , week $t$	MW
$T(b, t)$	number of hours in load block $b$ in week $t$	h
$S$	number of seconds per hour (3600)	
$a_m$	thermal plant capacity	MW
$b_m$	hydro plant capacity	m <sup>3</sup> /s
$c_m$	spillway capacity	m <sup>3</sup> /s
$r_j$	reservoir capacity	m <sup>3</sup>
$Y$	feasible set of transmission flows	
$A$	incidence matrix of river chain	

The variables are:

Symbol	Meaning	Units
$\rho$	weekly discount factor	
$f_m(b, t)$	generation of green peaker plant $m$ in load block $b$ in week $t$	MW
$z_i(l, b, t)$	shed load of type $l$ in load block $b$ in node $i$ in week $t$	MW
$x_j(t)$	storage in reservoir $j$ at end of week $t$	m <sup>3</sup>
$\bar{x}_j$	known storage in reservoir $j$ at start of week $t$	m <sup>3</sup>
$h(b, t)$	vector of hydro releases in block $b$ , week $t$	m <sup>3</sup> /s
$s(b, t)$	vector of hydro spills in block $b$ , week $t$	m <sup>3</sup> /s
$\omega(t)$	inflow (assumed constant over the week)	m <sup>3</sup> /s
$y(b, t)$	flow in transmission lines in load block $b$ in week $t$	MW
$g_i(y)$	sum of flow into node $i$ when transmission flows are $y$	MW

The water-balance constraints in the storage reservoirs at the end of week  $t$  are represented by:

$$x(t) = \bar{x} - S \sum_b T(b, t)(A h(b, t) + A s(b, t) - \omega(t)),$$

where  $x_j(t)$  is the storage in reservoir  $j$  at the end of week  $t$ ,  $s_j(b, t)$  denotes the rate of spill (in m<sup>3</sup>/second) in load block  $b$  in week  $t$ , and  $\omega_j(t)$  is the uncontrolled rate of inflow into reservoir

$j$  in week  $t$ . We multiply all of these by  $S$  to convert to  $\text{m}^3/\text{hour}$ , and then by  $T(b, t)$  to give  $\text{m}^3$  in each load block. All these are subject to capacity constraints. (In some cases we also have minimum flow constraints that are imposed by environmental resource consents.) The parameter  $\gamma_m$ , which varies by generating station  $m$ , converts flows of water  $h_m(t)$  into electric power. The same variables and constraints can be used to model pumping of water into a reservoir, except the value of the parameter  $\gamma_m$  is negative to reflect that energy is consumed as water is pumped into a higher reservoir.

## Appendix B. System data

We summarize here the features of the New Zealand system represented in JADE. A full data set can be downloaded from [New Zealand Electricity Authority \(2023\)](#). For our study, all existing fossil-fuel plant have been removed from the model.

Generator	Region	Capacity [MW]	Specific power (MW/cumec)
Arapuni	WTO	192	0.462
Aratiatia	WTO	78	0.284
Atiamuri	WTO	84	0.196
Karapiro	WTO	96	0.264
Maraetai	WTO	352	0.526
Matahina	BOP	80	0.595
Ohakuri	WTO	112	0.284
Rangipo	CEN	120	1.960
Tokaanu	CEN	240	1.750
Waikaremoana	HBV	140	3.535
Waipapa	WTO	54	0.139
Whakamaru	WTO	124	0.316

Table B.3: North Island hydro generation stations optimized in JADE model

## Appendix C. Storage trajectories

In Figures [C.15](#), [C.16](#), and [C.17](#), we plot the reservoir storage for the six largest reservoirs over the 31 years of historical simulation.



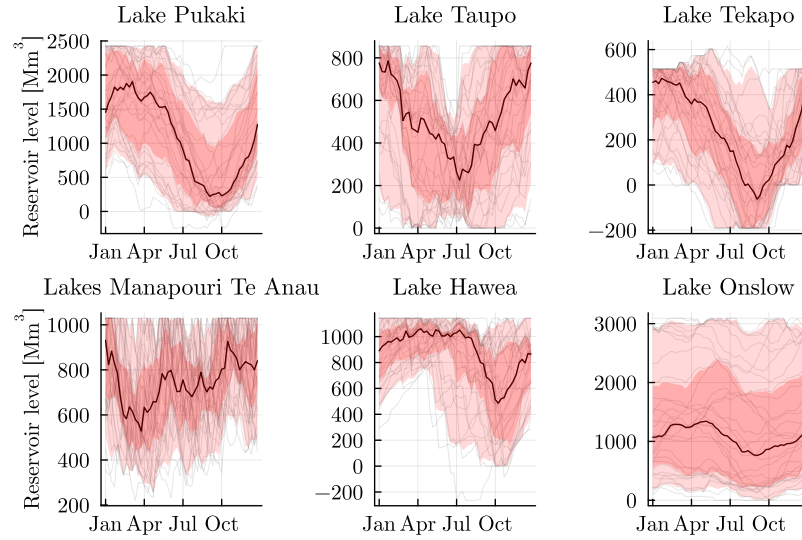


Figure C.15: Reservoir levels relative to the operating minimum for the historical inflow years 1990–2020 for the *Onslow* case. Light gray lines are the 31 trajectories from each simulated year, the solid line is the median, and the shaded bands are the 10–90 and 25–75 percentiles.

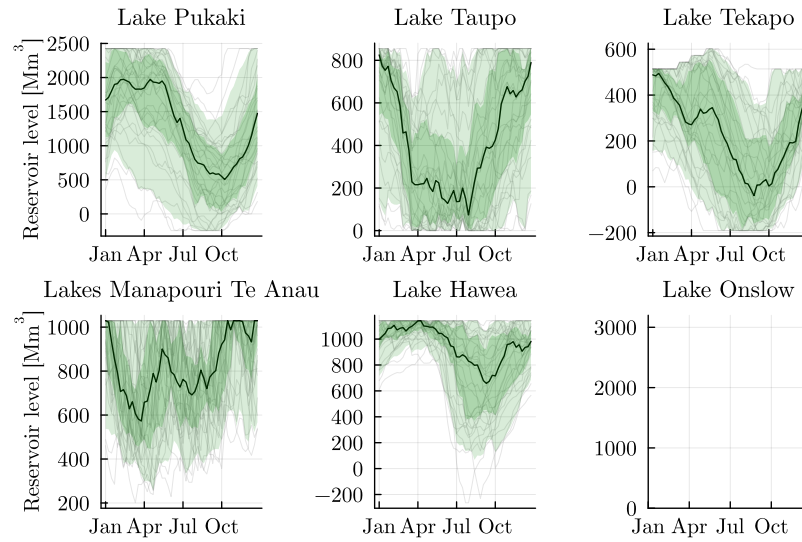


Figure C.16: Reservoir levels relative to the operating minimum for the historical inflow years 1990–2020 for the *Peakers* case. Light gray lines are the 31 trajectories from each simulated year, the solid line is the median, and the shaded bands are the 10–90 and 25–75 percentiles.

Generator	Region	Capacity [MW]	Specific power (MW/cumec)
Aviemore	CAN	220	0.310
Benmore	CAN	540	0.818
Clyde	OTG	464	0.518
Cobb	NEL	32	4.405
Coleridge	CAN	39	1.009
Manapouri	OTG	842	1.531
Ohau_A	CAN	264	0.501
Ohau_B	CAN	212	0.417
Ohau_C	CAN	212	0.417
Roxburgh	OTG	320	0.402
Tekapo_A	CAN	27	0.232
Tekapo_B	CAN	154	1.285
Waitaki	CAN	105	0.162
Onslow_Pump	OTG	1500	-7.027
Onslow_Gen	OTG	1500	5.417

Table B.4: South Island hydro generation stations optimized in JADE model

#### Appendix D. Marginal water values

In Figures [D.18](#), [D.19](#), and [D.20](#), we plot the marginal water values for the six largest reservoirs over the 31 years of historical simulation.

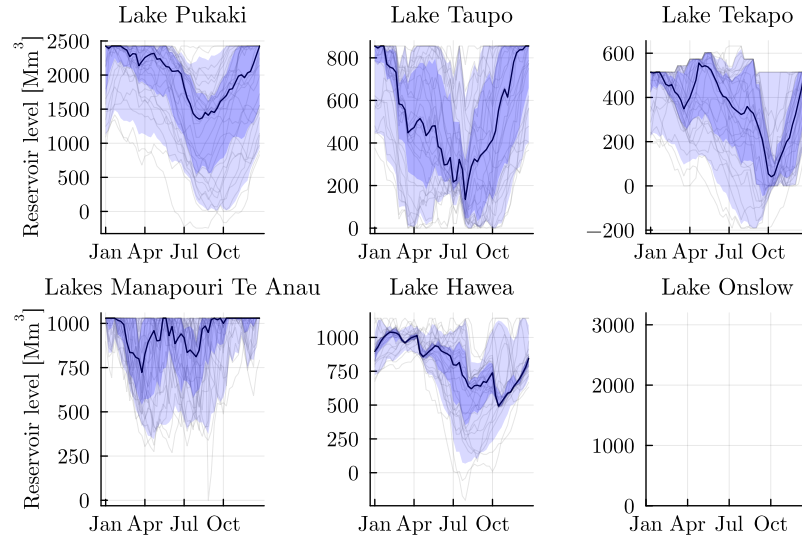


Figure C.17: Reservoir levels relative to the operating minimum for the historical inflow years 1990–2020 for the *Wind only* case. Light gray lines are the 31 trajectories from each simulated year, the solid line is the median, and the shaded bands are the 10–90 and 25–75 percentiles.

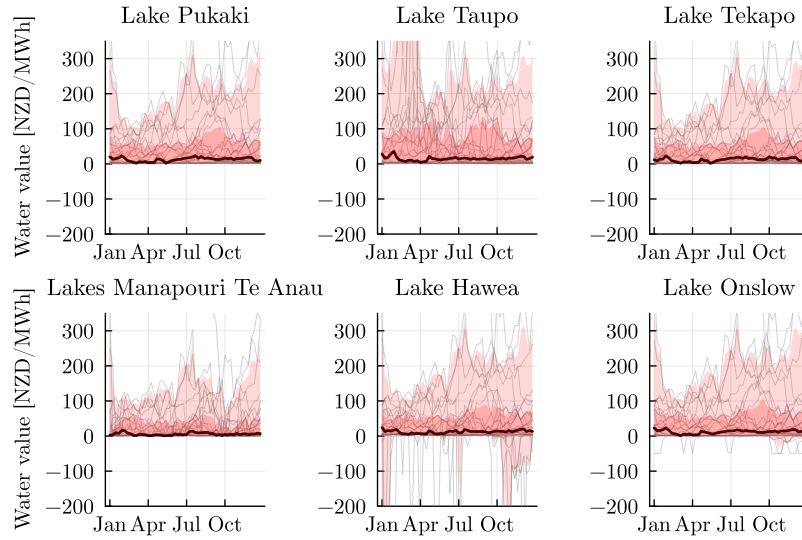


Figure D.18: Marginal water values for the historical inflow years 1990–2020 for the *Onslow* case. Light gray lines are the 31 trajectories from each simulated year, the solid line is the median, and the shaded bands are the 10–90 and 25–75 percentiles.

Region	Generation (MW)
AKL	39.54
WTO	864.97
BOP	240.23
CEN	30.16
TRN	59.91
CAN	15.80
NEL	22.24
OTG	40.23

Table B.5: Small fixed generation (including geothermal) for each region where this exists.

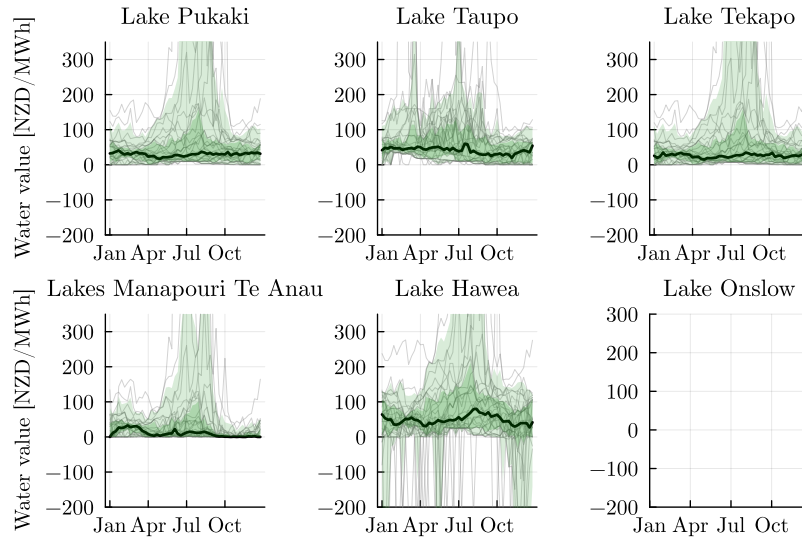


Figure D.19: Marginal water values for the historical inflow years 1990–2020 for the *Peakers* case. Light gray lines are the 31 trajectories from each simulated year, the solid line is the median, and the shaded bands are the 10–90 and 25–75 percentiles.

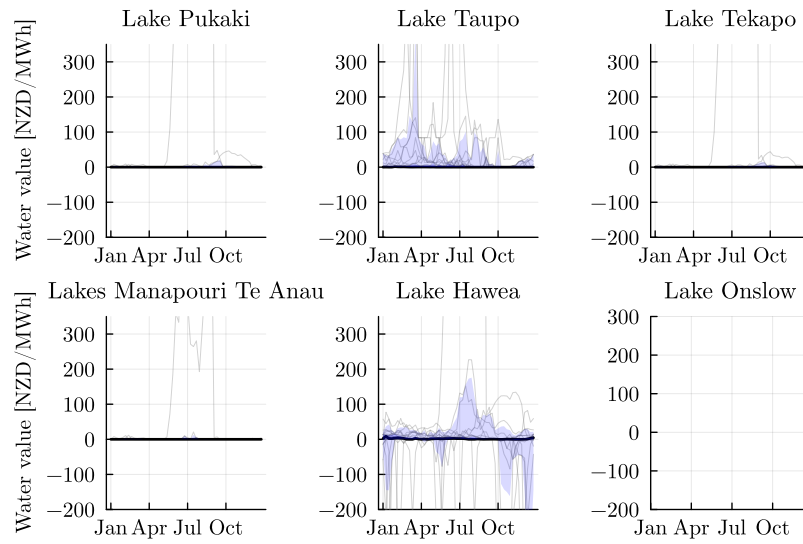


Figure D.20: Marginal water values for the historical inflow years 1990–2020 for the *Wind only* case. Light gray lines are the 31 trajectories from each simulated year, the solid line is the median, and the shaded bands are the 10–90 and 25–75 percentiles.