Using JADE to analyse pumped storage in New Zealand

A.B. Philpott & A. Downward
Electric Power Optimization Centre
University of Auckland

March 28, 2023

Abstract

JADE is a hydrothermal scheduling system developed in the JuMP modeling language and distributed as the Julia package JADE.jl. It leverages the stochastic dual dynamic programming package SDDP.jl to compute optimal water release policies over a calendar year. This paper describes the application of JADE to study a proposed pumped storage facility (Lake Onslow) in New Zealand's South Island in a target year of operation (2035). We compute optimal levels of wind investment with and without the Onslow storage, and show that the value of Onslow accrues from both energy storage (to cover dry winters) and capacity (to deal with windless peak demand periods). The Onslow solutions are then compared with solutions that can access dispatchable peaking plant.

1 Introduction

The New Zealand Government has set a target of 50% of total energy consumption to come from renewable sources in 2035, and has stated an aspiration to meet a target of 100% renewable electricity by 2030 [11]. One impediment to achieving this aspiration is the so-called "dry-winter" problem faced by the New Zealand electricity system. Dry winters with low winter rainfall and snowfall occur from time to time in New Zealand. These winters yield lower than expected inflows into New Zealand's hydroelectric reservoirs, resulting in the risk of an energy shortage. Because New Zealand cannot import electricity, energy shortages must be dealt with locally, either by dispatching backup generation or reducing load.

The market response to the risk of an energy shortage is to build thermal plant that can perform a hydro-firming role. The revenues that will be earned by these plants come from future electricity prices that vary with inflows and so are random. In principle, price distributions can be estimated from hydro-thermal scheduling models that optimize the trade-offs between thermal and hydro generation with uncertain inflows. If 100% of electricity generation is renewable then there will be no thermal plant to firm hydro. Firming must come from "overbuilt" renewable (wind) generation, demand response or from increased storage capacity, or some combination of these. Determining the costs and benefits of different firming approaches is currently under investigation by the New Zealand Battery Project [10].

Since future inflows to reservoirs are uncertain, the expected costs of load reduction (or thermal hydro-firming) must be estimated from stochastic optimal control models. The exact solution of these models is only possible (e.g. by dynamic programming [2]) for small instances with few state variables. In practice, the number of state variables required to represent typical electricity system models is beyond this, so methods of approximate dynamic programming are used to solve them. The most popular method of this type used in the energy sector is the Stochastic Dual Dynamic Programming (SDDP) algorithm of Pereira and Pinto [14].

The software we have developed to optimize hydro-firming is called JADE. The development of JADE arose out of research in reservoir optimization being carried out in the Electric Power Optimization Centre at the University of Auckland. The first code developed by this group was called Dynamic Outer Approximation Sampling Algorithm (DOASA). This solves the optimization problem using a

re-sampling version of SDDP as described in [15]. DOASA was originally implemented in AMPL as a general purpose implementation, which was later redeveloped in C++ and configured for the New Zealand electricity system under the name EMI-DOASA [16]. EMI-DOASA is about 20 times faster than the AMPL version.

The advent of the Julia programming language [3] and the JuMP mathematical programming package [6] enabled the development of a Julia package SDDP.jl by Dowson and Kapelevich [5]. The Julia package JADE.jl combines SDDP.jl with a JuMP model of the New Zealand electricity system that can be easily modified to focus on specific research questions related to the operation of this system. A default version of JADE is maintained by the New Zealand Electricity Authority with an up-to-date data set of historical inflows, demand, and infrastructure capacities. JADE.jl runs marginally slower than EMI-DOASA while gaining the flexibility afforded by the JuMP modeling language.

One of the leading contenders in the NZ Battery project is a scheme to increase storage in Lake Onslow in the lower South Island by pumping water from the Clutha river. This stored water can then be used to generate electricity during a dry winter. The water also can be used in a classical pumped-storage mode to generate electricity in peak demand periods when intermittent renewable generation (wind and solar) is insufficient to meet load. The capital costs of building such a scheme have recently been estimated to be (NZD)$15.7 billion [10].

The main aim of this paper is to demonstrate the use of JADE to compute the benefits of the Lake Onslow proposal to the electricity system. Although these benefits can be compared with the capital cost of this project, to see if it is worthwhile pursuing, we do not attempt a complete analysis here. Our intention is to showcase the insights that a SDDP model like JADE can provide. The expected operational benefits computed by our JADE model use the NZ CCC demand forecast for 2035 under their "Demonstration" scenario, assuming also that the Tiwai Point aluminium smelter will still operate or be replaced by other industrial load of the same amount. Other scenarios predict more load from plug-in electric vehicles, and different industrial electricity load, attenuated by smelter closure, A full analysis using JADE would repeat what we describe here for these alternative scenarios.

Our model assumes a central planning approach in which the total expected cost of operating the electricity system is minimized. This could be contrasted with a simulation model that computes electricity market outcomes arising from offers of energy from generators. Such a market would correspond to our model if all agents were to offer generation at its short-run marginal cost and were risk-neutral. In practice, generators (and loads) are risk-averse and so one might expect their actions to reflect this. Some of the risks faced by agents can be traded with counter parties through derivative instruments of various types. In theory, a complete market for trading risks when these are modeled using coherent risk measures leads to an equivalent risk-averse central planning problem [7], which raises the possibility of exploring the actions of risk-averse agents using a risk measure in SDDP. We leave this for future work.

The paper is laid out as follows. In the next section we describe the stage model that forms the weekly optimal dispatch problem in the 52-week steady-state discounted dynamic program. In section 3 we provide the assumptions and parameter choices made for our Lake Onslow study. These involve the physical design characteristics of the scheme, the approximation to the New Zealand transmission network, assumptions about electricity demand in 2035, and our inflow model. Section 4 presents the results of running JADE and simulating the (approximately) optimal dispatch policies that this yields. Finally, Section 5 draws conclusions.

2 JADE

2.1 JADE and SDDP

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 expected fuel cost to meet electricity

demand in weeks $t, t+1, \ldots$, 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 $\omega_j(t)$. We assume $\mathbb{E}_{\omega(T+1)}\left[C_{T+1}(\bar{x}, \omega(T+1))\right]$ is a known function of $\bar{x}$ that defines the expected future cost at the end of stage T when $x(T) = \bar{x}$. Then the Bellman function $C_t(\bar{x}, \omega(t))$ is the optimal solution value of the mathematical program:

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

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
ϕ_m	short-run marginal cost of peaker plant m	\$/MWh
ψ_{lb}	cost of shedding load type l in load block b	\$/MWh
γ_m	conversion factor for water flow into energy	MWs/m^3
$D_i(b,t)$	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^3/s
c_m	spillway capacity	m^3/s
r_j	reservoir capacity	m^3
Y	feasible set of transmission flows	
A	incidence matrix of river chain	

The variables are:

Symbol	Meaning	Units
$f_m(b,t)$	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^3
$\bar{x}_j$	known storage in reservoir j at start of week t	m^3
$h(b,t)$	vector of hydro releases in block b, week t	m^3/s
$s(b,t)$	vector of hydro spills in block b, week t	m^3/s
$\omega(t)$	inflow (assumed constant over the week)	m^3/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.

Here 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^3/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 m^3/hour, and then by $T(b,t)$ to give 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 γ_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 γ_m is negative to reflect that energy is consumed as water is pumped into a higher reservoir.

The (blocked) node-arc incidence matrix A represents the collection of river valley networks that make up the hydroelectric system, where each row of A aggregates controlled flows that leave a reservoir by spilling or generating electricity and subtracts those that enter a reservoir from upstream. In other words row j of $Ah(b,t) + As(b,t)$ gives the total controlled flow out of the reservoir (or river junction) represented by row j, this being the release and spill of reservoir j minus the sum of any immediately upstream releases and spill.

2.2 Steady-state optimization

JADE applies SDDP.jl to a JuMP model of the New Zealand electricity system as described above. SDDP.jl has the ability to model this system over an infinite horizon using a *PolicyGraph* description of a multistage stochastic optimization problem [4]. Whereas a finite-horizon implementation of SDDP has a linear policy graph with final stage T and an associated terminal cost function $C_{T+1}(x(T+1))$, it is possible to link stage T to stage 1 by adding an edge to the policy graph with an annual discount factor d that will apply to the future cost in stage 1 (of the next year). However, to give a smooth transition between years we discount costs in each stage of JADE by an equivalent weekly factor $d^{1/52}$.

The training iterations of SDDP occur as they would in the finite horizon case except that the future cost at stage $T+1$ is now modeled using the discounted future cost at the start of stage 1. Almost sure convergence of this process to an ε-optimal solution can be demonstrated by applying convergence results for long finite horizon models [15], but obtaining reasonable policies typically takes many more iterations than a finite-horizon training. The outcome of this training is a policy that can be simulated over sampled or historical inflow sequences to yield steady-state trajectories as shown in Figure 2. Here the value of national storage at the right-hand end of a trajectory is the starting value for the following year. In each week these values represent a probability distribution of national energy levels in steady-state operation of the JADE policy.

3 Lake Onslow study

The design and operation of the proposed Lake Onslow pumped-storage facility have not yet been finalized. Some preliminary estimations of design parameters can be found in [1] and [9]. Here we

summarize these parameters and the assumptions that we have made for design decisions that have yet to be determined.

3.1 The reservoir and power station

There have been two proposed reservoir options: Lake Onslow and Lake Onslow plus Manorburn Basin. We have chosen to restrict attention to filling only Lake Onslow by building a dam where it meets the Teviot River. Our study assumes a maximum stored volume of 3,100 Mm^3 amounting to stored energy of 5,000 GWh [10].

Two tunnel options for pumping and release of water into Lake Onslow that have been considered in previous research [1] are:

1. the Upper Roxburgh Dam connects to Lake Onslow via a 24km tunnel; or
2. the Lower Clutha River connects to Lake Onslow via a 15km tunnel.

Our study has assumed the first of these options, to be modelled by two arcs, a release arc ($m = r$) and a pumping arc ($m = p$). We assume a constant pumping efficiency of 77%, so $\gamma_r = -0.77\gamma_p$.

3.2 The transmission network

JADE requires load duration curves for the study year for each week of this year and each region of the model. The full network used in Transpower's economic dispatch model SPD is too large to represent in JADE and so we approximate this network by 11 regions as shown in Figure 1. Observe that the approximate network is radial (i.e. contains no loops), which results in some loss in accuracy (e.g. in ignoring the transmission line from Waikato (WTO) to Hawkes Bay (HBY).)

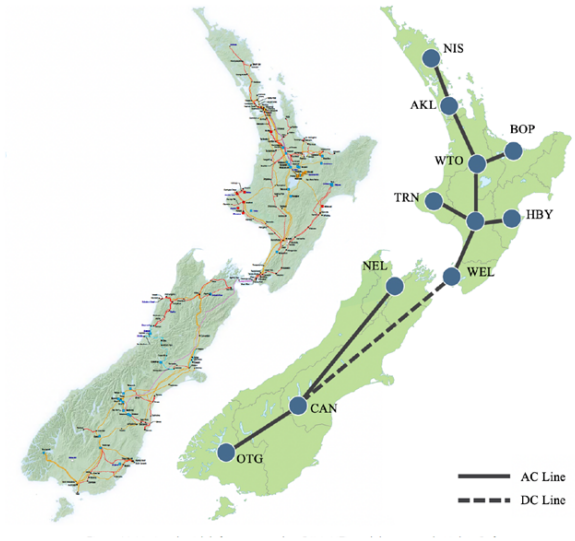


Figure 1: SPD network on left is approximated by an 11-node transmission system

In order to model transfers of power between regions we identify transmission lines that form *cuts* that separate the SPD network into the regions shown. Each line has an implicit direction, so some recorded flows will be negative. The list of lines used in our model is shown in Table 1.

From	To	Lines
NIS	AKL	BRB_HPI1.1
AKL	WTO	PAK_WKM1.2; PAK_WKM2.2; OTA_WKM1.1; OTA_WKM2.1; DRY_HLY1.1
WTO	BOP	KIN_TRK1.2; KIN_TRK2.2; ATI_TRK1.1; ATI_TRK2.1
WTO	CEN	ARI_ONG.2
CEN	TRN	BPE_WGN1.1; BPE_WGN2.1; BPE_BRK1.1; BPE_BRK2.1
CEN	HBY	DVK_WPW1.2; DVK_WPW2.2
WEL	CEN	HAY_LTN1.1; MGM_WDV1.1; MGM_WDV.1
WEL	CAN	HAY_BEN1.1; HAY_BEN2.1
CAN	NEL	ISL_KIK1.1; ISL_KIK2.3; ISL_KIK3.3; COL_OTI1.1; COL_OTI2.1
CAN	OTG	NSY_ROX.1
AKL	NIS	HPI_MDN1.1; HEN_MPE2.1; HEN_MPE1.1
WTO	AKL	OHW_OTA2.1; OHW_OTA1.1; HLY_OTA2.1
BOP	WTO	KAW_OHK.1
CEN	WTO	TKU_WKM1.1; TKU_WKM2.1; RPO_WRK1.1
CEN	WEL	BPE_HAY1.1; BPE_HAY2.1; BPE_PRM_HAY1.1; BPE_PRM_HAY2.1; BPE_WIL1.2
CAN	WEL	BEN_HAY1.1; BEN_HAY2.1
OTG	CAN	CYD_TWZ1.2; CYD_TWZ2.2

Table 1: SPD lines joining regions in 11-node model, shown in direction of positive flow.

3.3 Regional load estimates for 2020

The load duration curves for each region are based on the observed load values for 2020. We compute the generation in each half-hour trading period of 2020, by running vSPD[1] with the GDX file for the trading periods in each day[2]. This provides us with 366 days, each containing 48 periods (apart from daylight savings). For each trading period we record the generation of each offering plant and the sent and received flows in each transmission line in Table 1.

Let $r \in R$ denote the index of a region. For each region r we create a list of generators $g \in r$ that we wish to model in this region. The 2020 generation in a given trading period p is denoted $q_g(p)$. In each region r there is a set of nodes $n \in \mathcal{N}_r$, where $\cup_r \mathcal{N}_r = \mathcal{N}$. For every $r \in R, s \in S$ we identify all lines in the set

$$\mathcal{L}(r,s) = \{l \in \mathcal{L} : l = (m,n), m \in \mathcal{N}_r, n \in \mathcal{N}_s\}.$$

Observe that these are directed arcs so $\mathcal{L}(r,s) \cap \mathcal{L}(s,r) = \emptyset$. We adopt the convention that the node names are ordered (e.g. are integers) so that $m < n$ for every directed line (m,n).

The flow on arc $l = (m,n)$ in period p is denoted $f_l(p)$. We identify two values with $f_l(p)$ namely $f_l^+(p)$ denoting the flow measured at the start node m of the directed line, and $f_l^-(p)$ denoting the flow measured at end point n. The loss in flow in line l is then

$$h_l(p) = f_l^+(p) - f_l^-(p).$$

Observe that $h_l(p) > 0$, even if $f_l^+(p) < 0$. In that case, positive flow is sent from n to m, so $-f_l^-(p) > -f_l^+(p)$, and $f_l^-(p) < f_l^+(p) < 0$.

For each line we compute

$$\begin{aligned} f_l^+(p) &= f_l(p) + \frac{\alpha}{2} h_l(p) \\ f_l^-(p) &= f_l(p) - \frac{\alpha}{2} h_l(p), \end{aligned}$$

where we choose $\alpha = 0$ for a model with lossless flow and $\alpha = 1$ to represent losses. Then we compute

[1] vSPD is the New Zealand Electricity Authority GAMS implemetation of SPD, available for download from www.emi.ea.govt.nz/Wholesale/Tools/vSPD.

[2] GDX files for each historical day are downloadable from https://www.emi.ea.govt.nz/Wholesale/Datasets/FinalPricing/GDX.

the total generation in a region r to be

$$q_r(p) = \sum_{g \in r} q_g(p).$$

Now consider two adjacent regions r and s. The net flow leaving r for s is

$$F_{rs}(p) = \sum_{l \in \mathcal{L}(r,s)} f_l^+(p) - \sum_{l \in \mathcal{L}(s,r)} f_l^-(p)$$

and the net flow leaving s for r is

$$F_{sr}(p) = \sum_{l \in \mathcal{L}(s,r)} f_l^+(p) - \sum_{l \in \mathcal{L}(r,s)} f_l^-(p).$$

The sum of these two flows is

$$\begin{aligned} F_{rs}(p) + F_{sr}(p) &= \sum_{l \in \mathcal{L}(r,s)} f_l^+(p) - \sum_{l \in \mathcal{L}(s,r)} f_l^-(p) + \sum_{l \in \mathcal{L}(s,r)} f_l^+(p) - \sum_{l \in \mathcal{L}(r,s)} f_l^-(p) \\ &= \sum_{l \in \mathcal{L}(r,s)} h_l(p) + \sum_{l \in \mathcal{L}(s,r)} h_l(p) \end{aligned}$$

which is the transmission loss in transfers between r and s.

The demand in region r in period p is now the generation in r in period p minus the total net transfer out of r to other regions:

$$d_r(p) = q_r(p) - \sum_{s \in R} F_{rs}(p).$$

Observe that

$$\begin{aligned} \sum_{r \in R} d_r(p) &= \sum_{r \in R} q_r(p) - \sum_{r \in R} \sum_{s \in R} F_{rs}(p) \\ &= \sum_{r \in R} q_r(p) - \sum_{r \in R} \sum_{s > r} \left(\sum_{l \in \mathcal{L}(r,s)} h_l(p) + \sum_{l \in \mathcal{L}(s,r)} h_l(p) \right) \\ &= \sum_{r \in R} q_r(p) - \alpha(\text{flow loss}). \end{aligned}$$

3.4 Regional load estimates for 2035

The load duration curves for future years are estimated based on the demand $q_r(p), r \in R$ for periods p in a *base* year, chosen in this study to be 2020. Since 2020 is a leap year we remove the February 29 records if the future year to be estimated is not a leap year. The set of periods is then P, where $|P| = 17520$ or 17568 depending on the future year. This means that the number of hours H in the base year and future year are the same. We let h be the number of hours in each period (0.5), so $H = |P|\, h$.

The growth factor for demand in a future year relative to the base are based on base-year national demand and forecasts for total national annual demand in the future year taken from the New Zealand Climate Change Commission 2021 Draft Advice Spreadsheets [8]. We denote

$$\begin{aligned} I_k &= \text{Total industrial demand in year } k \text{ (GWh)} \\ C_k &= \text{Total commercial and residential demand in year } k \text{ (GWh)} \\ S_k &= \text{Total solar energy generation in year } k \text{ (GWh)} \\ G_k &= \text{Total geothermal energy generation in year } k \text{ (GWh)} \\ V_k &= \text{Total plug-in electric vehicle load in year } k \text{ (GWh)} \end{aligned}$$

where $k = 0$ indicates the base year. We have that $I_k + C_k = T_k$, the total electricity demand (GWh) in year k. We define increases in these quantities (GWh) by

$$\begin{aligned} \Delta I_k &= I_k - I_0 \\ \Delta C_k &= C_k - C_0 \\ \Delta G_k &= G_k - G_0 \\ \Delta V_k &= V_k - V_0 \\ \Delta S_k &= S_k - S_0. \end{aligned}$$

We let P_r =population of region r in the base year and define its proportion

$$\rho_r = \frac{P_r}{\sum_{r \in R} P_r}, \quad r \in R.$$

The increase in industrial load (MWh) in region r is assumed to be distributed by population, so is estimated by $\rho_r 1000 \Delta I_k$. This is then divided by H to give a flat increase of

$$i_r(p) = \rho_r 1000 \Delta I_k / H$$

MW in each period p. Similarly the increase in plug-in electric vehicle load is

$$v_r(p) = \rho_r 1000 \Delta V_k / H.$$

The increase in geothermal generation ΔG_k is allocated to Waikato (WTO) and Bay of Plenty (BOP) regions in the same ratio $w : (1 - w)$ as 2020 geothermal generation. This gives a flat increase in MW of

$$g_W = w 1000 \Delta G_k / H$$

for Waikato and

$$g_B = (1 - w) 1000 \Delta G_k / H$$

for Bay of Plenty.

The increase $s_r(p)$ in solar generation (MW) in region r is estimated using a given annual solar profile σ_r for each region. Here we take $\sigma_r(p)$ to be the 2008 historical solar radiation figures for Auckland, Wellington and Christchurch respectively from [12], assigned to the appropriate region $r \in R$.[3]

$$s_r(p) = \frac{\sigma_r(p) 1000 \Delta S_k}{h \sum_{p \in P} \sum_{r \in R} \sigma_r(p)}.$$

The increase $c_r(p)$ in commercial and residential demand (MW) in region r is estimated as follows. We first compute the 2020 commercial and residential demand $\hat{C}_0(p)$ (MW) from vSPD.

$$\hat{C}_0(p) = \sum_{r \in R} d_r(p) - 1000 I_0 / H.$$

This gives a scale factor

$$\beta = \frac{1000(C_k - C_0)}{h \sum_{p \in P} \hat{C}_0(p)}.$$

Then

$$c_r(p) = \beta \rho_r \hat{C}_0(p).$$

The forecast load $\hat{d}_r(p)$ in region r and period p for the future year can now be computed to be

$$\hat{d}_r(p) = d_r(p) + c_r(p) + i_r(p) + v_r(p) - s_r(p).$$

Note that this accounts for generation from solar irradiation which is subtracted from the demand.

[3] Regions CEN,TRN,HBY,WEL were assigned Wellington radiation figures, and the other North Island regions assigned Auckland radiation figures. The South Island regions were all assigned Christchurch radiation figures.

3.5 Building load duration curves

JADE requires load duration curves for each each region and each week of the year to be studied. This load will be met by generation from renewable energy sources, namely wind farms, solar panels, geothermal energy and hydroelectric generation, and possibly some peaking plant (which may be configured to burn biofuel or hydrogen). Solar generation is assumed to be embedded, i.e. not dispatched by SPD. Geothermal energy is taken as fixed baseload quantity distributed between WTO and BOP regions. Hydroelectric generation is either from storage reservoirs or run-of-river plant. We discriminate between small hydroelectric facilities for which inflow information is not available and larger plant for which inflows are recorded in the Electricity Authority EMI database. The small plant are assumed to operate at a fixed average generation level throughout the year.

Wind generation requires special attention because of its intermittency. Since the wind might not blow during peak demand periods we cannot uniformly subtract average wind generation from the load. Our approach is to correct the demand $\hat{d}_r(p)$ by subtracting forecast wind generation $\hat{w}_r(p)$, and then to build a load duration curve for the national load

$$n(p) = \sum_{r \in R} \left(\hat{d}_r(p) - \hat{w}_r(p) \right).$$

Here $\hat{w}_r(p)$ is estimated using forecast wind capacity expansion E and actual wind generation observed in 2021. The capacity increment E is allocated to regions in proportion to the capacity of new consented wind farms in that region, so region r has capacity increase $\alpha_r E$ where $\sum_{r \in R} \alpha_r = 1$.

We construct wind data for each region based on historical wind generation for the calendar year 2021. Wind generation data are available from the EMI database for ten New Zealand wind farms[4]. Each wind farm in this database is located in a region $r \in R$ as shown in Table 2.

Windfarm	Region	Rated capacity (MW)
Mahinerangi	OTG	36
MillCreek	WEL	59.8
Tararua Stage 1	CEN	31.7
Tararua Stage 2	CEN	36.3
Tararua Stage 3	CEN	93.0
Te Apiti	CEN	90.8
Te Rere Hau	CEN	48.5
Te Uku	WTO	64.4
Turitea	CEN	222
Waipipi	TRN	133
WestWind	WEL	142.6
White Hill	OTG	58

Table 2: Wind farms modeled for 2021 and rated capacity from NZWEA [13].

For each region r with current or consented new wind capacity we selected a representative wind farm (in the region or a neighbouring one) as shown in Table 3. Wind farms were selected only if their 2021 generation data set was complete.

Given historical wind generation $w_r(p)$ for the representative wind farm for region r, and its rated capacity W_r, we can compute the ratio $\frac{w_r(p)}{W_r}$ that can be multiplied by current and forecast extra wind capacity in region r to give forecast wind generation with total capacity expansion E to be

$$\hat{w}_r(p) = \frac{w_r(p)}{W_r}(E_r(2021) + \alpha_r E)$$

[4]Downloadable from https://www.emi.ea.govt.nz/Wholesale/Datasets/Generation/Generation_MD

Region	α_r	Representative Windfarm	Capacity W_r (MW)
CAN	0.0814	White Hill	58
CEN	0.0627	Tararua Stage 3	93
HBY	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 3: Representative wind farms for regions.

where $E_r(2021)$ is the capacity of wind farms in region r in 2021, as shown in Table 4.

Region	$E_r(2021)$	$E\alpha_r$
CAN	0.0	81.4
CEN	522.0	62.7
HBY	0.0	91.4
OTG	94.0	192.6
TRN	133.0	0.0
WEL	202.4	567
WTO	64.4	4.9

Table 4: Locations of wind capacity (MW) for 2021 and for $E = 1000$.

The final step in producing load duration curves for each week i, is to reorder the series $n(p)$ for periods p in each week i by decreasing magnitude. We then construct a piecewise constant approximation of the resulting decreasing curve with B pieces, where each piece $b = 1, 2, \ldots, B$ contains a set $P(i, b)$ of trading periods for week i. This is done using a lot-sizing algorithm that minimizes the weighted error between the curve and its approximation. To enable peaks to be represented accurately, the weights are chosen heuristically so that $T(i, b) = |P(i, b)|$ is smaller when b contains peak periods. Given $P(i, b)$ for each week i, we assign block b the average demand (MW)

$$D(i,b) = \frac{\sum_{p \in P(i,b)} \left(\hat{d}_r(p) - \hat{w}_r(p) \right)}{T(i,b)}$$

and record the "hours-per-block" $T(i, b)/2$. Note that $\hat{w}_r(p)$ depends on E, so we construct a demand file and hours-per-block file for each value of E that we wish to study.

3.6 Demand response and lost load

We assume that $\mathcal{L}(i)$ and the prices for lost load and demand response, as a proportion of positive nodal net-demand, is the same at each node $i \in \mathcal{N}$. These proportions and prices are given in Table 5 below.

4 Results

JADE was used to estimate the value of Lake Onslow operating in a steady state in 2035. Demand duration curves were computed using 2035 forecasts of I_k, C_k, S_k, G_k,and V_k made by the Climate Change Commission for their 2021 Draft Advice "Demonstration" scenario [8]. In contrast to the Climate Change Commission data, we assume that Tiwai Point Aluminium smelter is still operating in 2035 (or is replaced by some industrial base load at the same level of 570 MW).

Proportion	Price ($ / MWh)
0.025	530
0.025	740
0.050	3180
0.150	5290
0.750	10580

Table 5: Prices for demand response and involuntary lost load for load types in $\mathcal{L}(i)$.

Except where assumed otherwise, all gas and coal thermal generation is assumed to have been shut down. JADE is then run in infinite-horizon mode with stagewise independent inflows (adjusted in variance using a DIA factor of 4 sampled from historical years 1990-2019). A discount factor of 0.90 is assumed. Each run of JADE assumes a different amount E of extra wind investment, distributed over the regions as described above. Different values of E have been tested to locate the value where GWAP for wind is approximately equal to the LCOE for wind. This rather ad-hoc search procedure leads to an assumed level E of wind investment that can be tested using simulation of JADE policies over historical years 1932-2020.

Three case studies have been investigated:

(a) **Wind-only**: 4800MW of wind capacity is built across the country in the proportions outlined in Table 4.

(b) **Onslow**: Onslow is built with 5000GWh of storage capacity and 1500MW of pumping and generation capacity. 3725MW of wind capacity is built.

(c) **Green Peakers**: 500MW of peaking plant is available in Waikato, the fuel for this plant costs $160/MWh. 3900MW of wind capacity is built.

The line capacities specified in JADE for each case are given in Table 6; these account for the different locations of the investments for each case. The line capacities are chosen so as to relieve congestion. We see that the necessary transmission capacities from CAN to WEL, between WEL and CEN and between CAN and OTG are largest for the Onslow case, whereas the capacities between WTO and CEN and from WEL to CAN are largest for the Wind-only case. The green peakers case generally required the least transmission capacity of the three cases. These observations are consistent with what one might expect, given the location of generation and demand.

Line	(a) Wind-only	(b) Onslow	(c) Green Peakers
NIS ↔ AKL	300	270	300
AKL ↔ WTO	2300	2500	2300
WTO ↔ BOP	500	500	500
WTO ↔ CEN	1900	1800	1600
CEN ↔ TRN	450	450	450
CEN ↔ HBY	300	300	300
WEL ↔ CEN	1700	1900	1600
WEL → CAN	1400	1600	1200
CAN → WEL	1800	2600	1700
CAN ↔ NEL	250	250	250
CAN ↔ OTG	1100	2300	1000

Table 6: JADE line capacities for each case (MW).

4.1 Hydro storage trajectories

Since the investments by 2035 will have a significant effect on water valuation in reservoirs, JADE is trained using the steady-state mode with a discount factor of 0.9 for each case study. These are evaluated by running a continuous simulation from 1932 to 2020. This means that the reservoir levels at the end of a year continue into the beginning of the next.

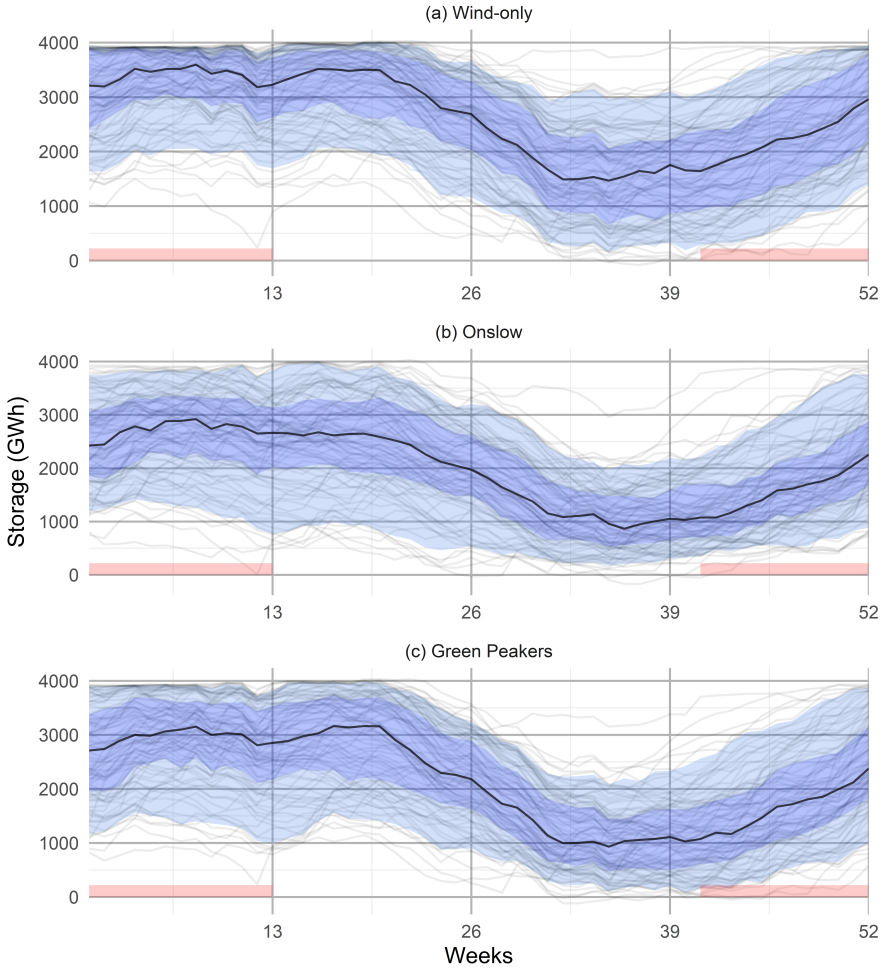


Figure 2: Hydro-storage trajectories in existing reservoirs, simulated from 1932 to 2020, for each of the cases.

We can see in Figure 2 how the distribution of storage in the existing reservoirs at the beginning of the year is different for each case. Note that the central blue band represents the 25^{th} to 75^{th} percentile of the distribution each week, and the light blue represents the 5^{th} to the 95^{th} percentile. The red boxes represent variable contingent storage in Lake Tekapo. Furthermore, the black line gives the median of the storage of all the years, and the grey lines are the storage trajectory for each of the years.

In the Wind-only case, storage is maintained at a higher level, increasing the amount of spill, whereas with Onslow the storage at the beginning of the year is centred lower than the other cases. The reason for this is shown in Figure 3, where we see that Lake Onslow, whose storage exceeds the capacity of all the current reservoirs combined, is almost full at the start of 25% of years within the simulation; this allows the other reservoirs to be maintained at a lower level, lowering the risk of spill. This is confirmed by Figure 4, which shows distribution of spill each year (in GWh).

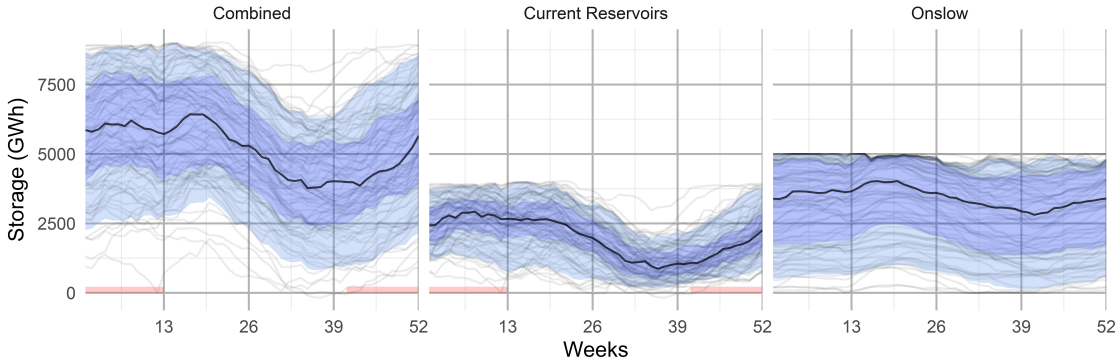


Figure 3: Hydro-storage trajectories simulated from 1932 to 2020, in case (b) Onslow.

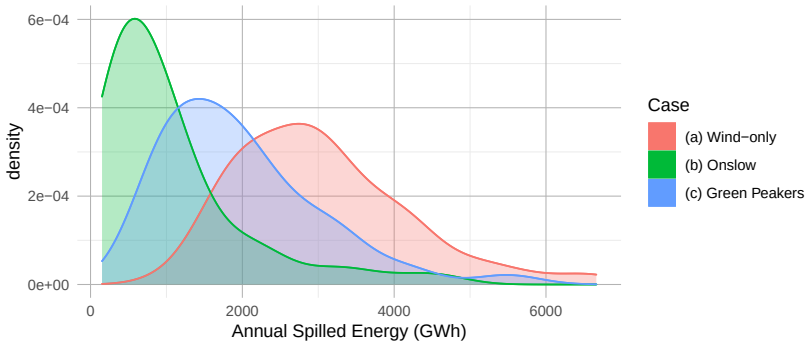


Figure 4: Distributions of annual spilled energy (in GWh) for each case.

4.2 Marginal Water Values

As shown in Figure 5 and Figure 6 (enlarged), the distribution of marginal water values varies by case. The Wind Only and Green Peaker cases look similar in Figure 5, but we can see from Figure 6 that the Green Peaker 25th-percentile marginal water values are persistently higher. This is consistent with the storage trajectories in Figure 2 that are lower for the Green Peaker case.

The distribution of marginal water values for the Onslow case is centred lower than the other two, accounting for the extra storage in Onslow. However we can observe from Figure 5 that the maximum marginal water values are much higher for Onslow than the other cases. In years where inflows are low in consecutive years such as 1977 and 1978, Onslow levels become very low. Green peakers or more wind capacity attenuates the effect of this energy shortage that is more keenly felt in the Onslow case when these are absent.

4.3 Prices

We now consider the prices from the simulation of the three cases. Various weighted-average price metrics are shown in Table 7. These are *time-weighted average price* (TWAP), *load-weighted average price* (LWAP), and *generation-weighted average price* (GWAP) which varies by generator. The national TWAP and LWAP prices are lowest for the Onslow case, and these prices are also similar to the GWAP price for wind generation, and generation from Onslow. However, the cost of pumping is much lower, at about $18/MWh. The wind GWAP for the Wind Only and Green Peakers cases are similar at $65-67, this is approximately the estimated LCOE for wind generation. However, the LWAP average prices are much higher in the other cases with $115/MWh for the Green Peakers case, and $181 for the Wind Only case. The TWAP prices are about 20% lower in each case. The GWAP for the Green Peakers is $591/MWh, this is much higher than the $160/MWh SRMC for the plant, but given the plant's low capacity factor, peaking plant makes an average annual operating surplus of $315,000 on each MW of capacity.

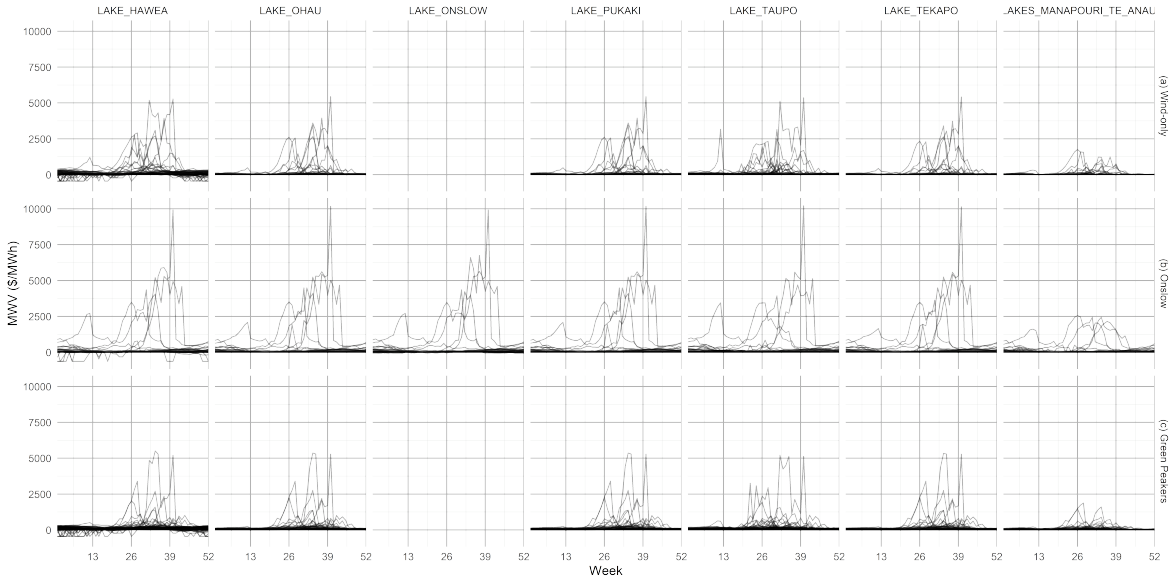


Figure 5: Marginal water value simulated from 1932 to 2020 for each case.

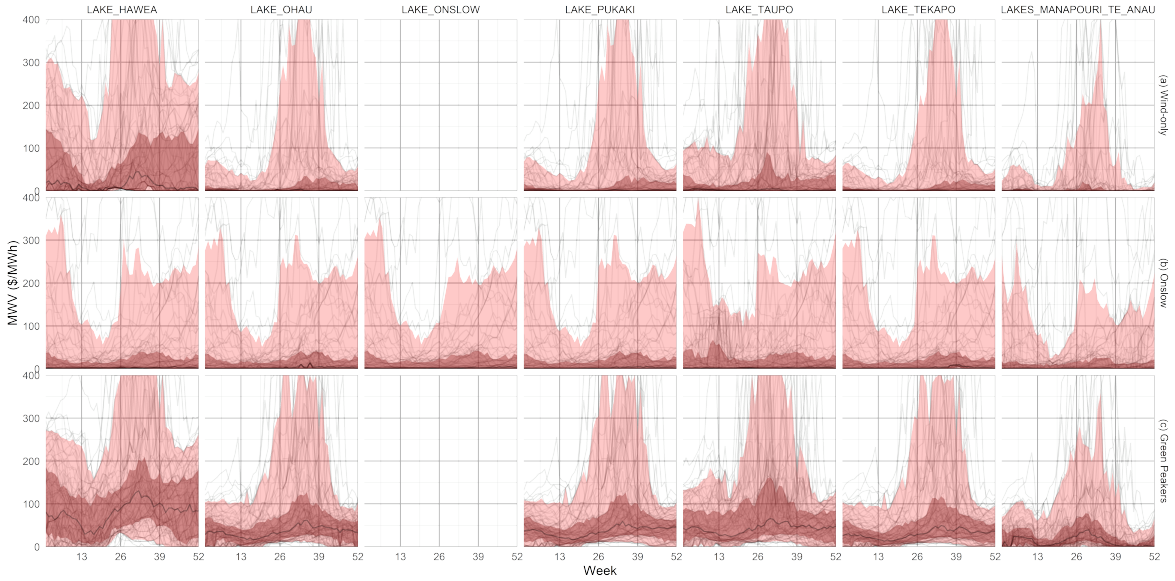


Figure 6: Marginal water value simulated from 1932 to 2020 for each case (y-axis truncated to \$0–\$400).

Metric	(a) Wind-only	(b) Onslow	(c) Green Peakers
TWAP	145.1	72.4	99.0
LWAP	181.0	75.8	116.0
Wind GWAP	65.0	67.1	65.7
Hydro GWAP	237.1	64.0	126.5
Peaker GWAP	NA	NA	590.6
Pump GWAP	NA	18.3	NA
Release GWAP	NA	67.6	NA

Table 7: Price metrics ($ / MWh).

Finally, the GWAP earned by the hydro plants varies considerably over the three cases. In the wind-only case, the hydro plants provide backup capacity when the wind isn't available, leading to a GWAP ($237.2/MWh) that exceeds the LWAP ($181.0/MWh). For the case with Onslow, the hydro plants' GWAP ($64.0/MWh) is lower than the LWAP ($75.8/MWh). For the case with green peakers, the hydro GWAP ($126.5/MWh) is only slightly higher than the LWAP ($116.0).

In appendix A1, we plot the weekly national energy storage. Figures 16. 18, and 20 show how LWAP prices vary for the sequence of 89 simulated years for each of the cases. This shows how the prices vary throughout the years and how sequences of wet and dry years can lead to low and high prices, respectively. It is interesting to note, for example, in the period from 2005 to 2008, shown for all three cases in Figure 7, a sequence of dry and moderate years leads to extremely high prices in 2008 for all three cases, but the Onslow case fares worst of all, since limited wind has been built, and there are no peaking plants.

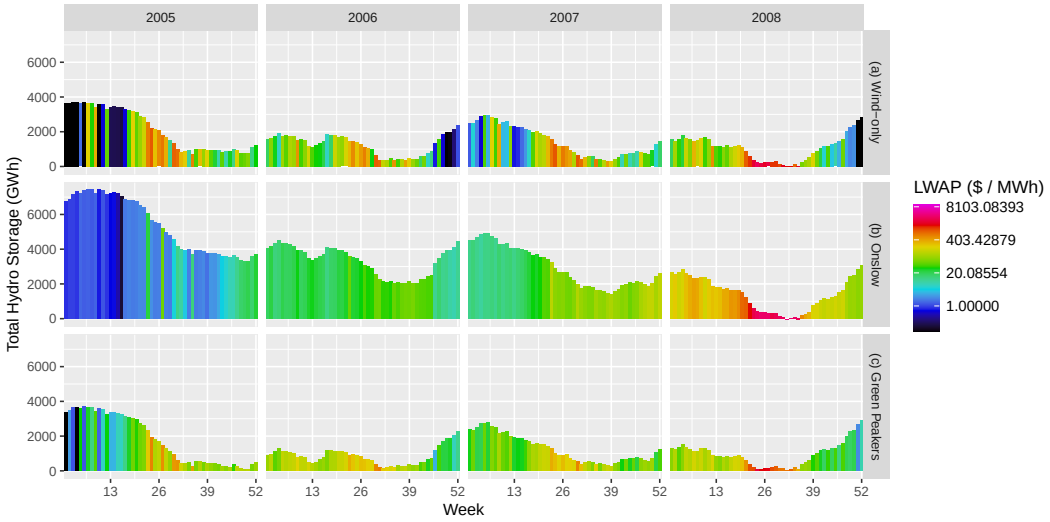


Figure 7: Storage and LWAP for each case from 2005 to 2008.

Recall that for the investment in new wind generation capacity, we have found the capacity such that the GWAP for wind (over the sequence of 89 historical years) approximately matches the LCOE for wind investment. However, it is important to realise that the GWAP will vary each year, which presents a risk for potential investors; this variability is depicted in figures 8 and 9. To create these plots the steady-state hydro policies were simulated for a sequence of 1000 uniformly sampled hydrological years.

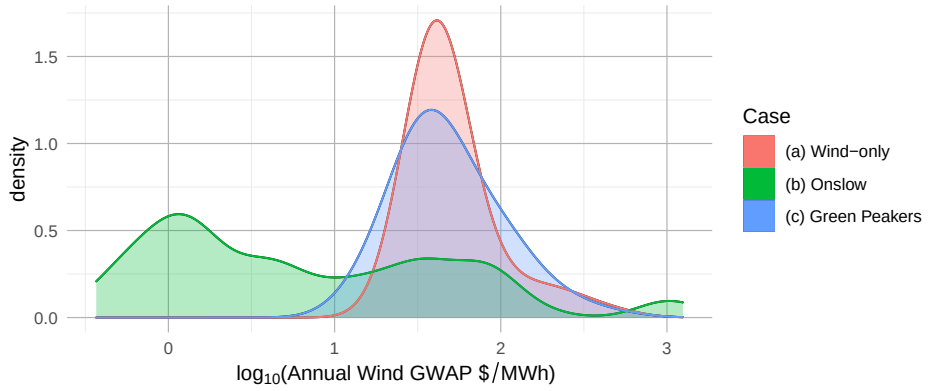


Figure 8: Smoothed distributions of annual GWAP for wind (transformed using $\log_{10}$).

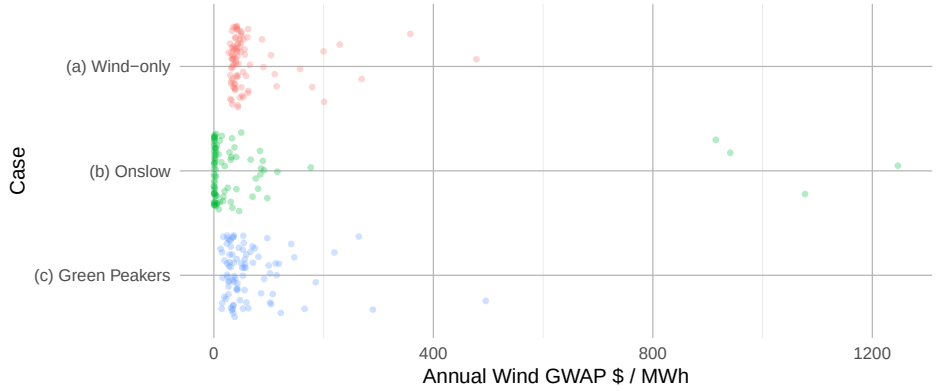


Figure 9: Annual GWAP for wind. Each dot represents a single year from the sequence of 89 years.

4.4 Lost load

Simulating the steady-state hydro policies over the hydrological years from 1932-2020 yields the distribution of lost load shown in Figure 10. These show that Onslow generally reduces energy shortages in comparison with the other cases. In some rare circumstances, the Onslow case will experience some very high lost load.

4.5 Costs

Simulation over the sequence of 89 historical inflows gives an average annual load shedding cost of \$314.4m for the wind-only case, \$96.4m for the Onslow case and \$81.8m for the Green Peakers case; in addition the Green Peakers case incurs, on average, an additional \$81.8m per year in fuel costs. All these costs, of course, vary depending on the hydrological conditions each year, as shown in Figure 11.[5] The wind-only case builds 1075MW more wind than the Onslow case. Assuming a capacity factor for wind of 0.4, and converting the previously assumed \$65 / MWh into an annual cost, gives:

$$\$65/\text{MWh} \times 8760 \text{ hours} \times 1075\text{MW} \times 0.4 = \$244.8\text{m}.$$

On the other hand, the annualised cost of Onslow using a discount factor of 0.9 over 60 years can be computed as follows:

$$\$15.7\text{b} \times \frac{1-0.9}{1-0.9^{60}} = \$1.573\text{b}.$$

[5]These costs do not include non-financial costs, such as violations of flow-rate consents or utilisation of contingent storage.

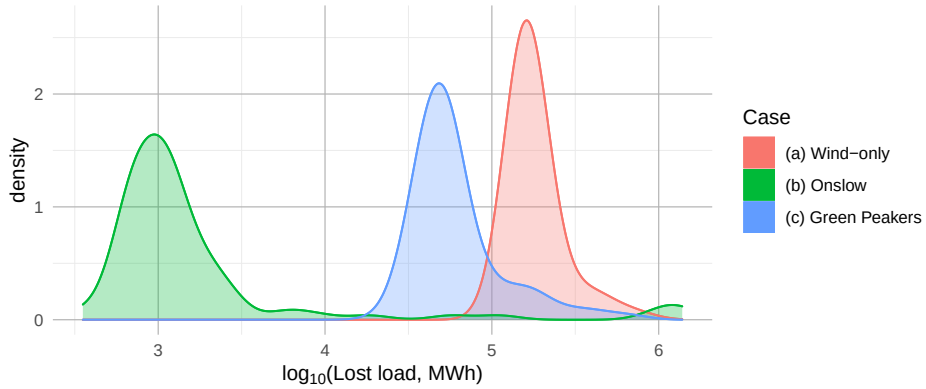


Figure 10: Smoothed distributions of annual lost load (in MWh) for each case (transformed using $\log_{10}$).

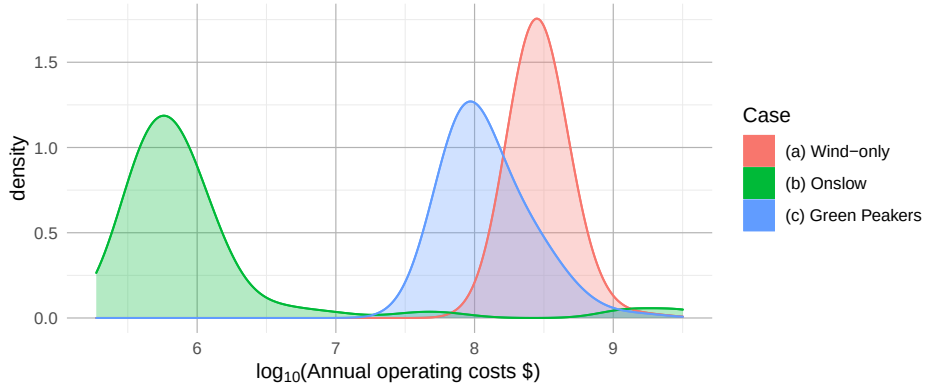


Figure 11: Smoothed distributions of annual load shedding + fuel costs (in \$) for each case (the data has been transformed using $\log_{10}$).

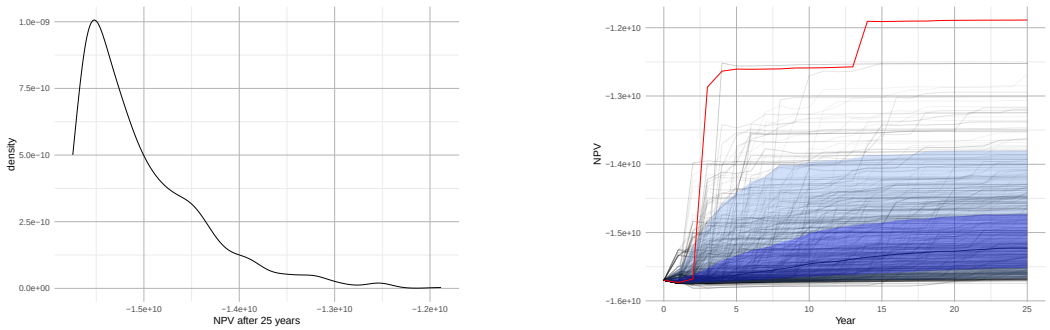
Under these assumptions the Onslow case would be \$1.11b more expensive per annum (on average) than the wind-only case, and it would have increased cost volatility, as shown in Figure 11.

Of course it would be possible to assume a lower discount rate. However, if we only changed the discount rate for Onslow, leaving all other costs the same as above, we would need to assume a discount rate of 2.5% in order for the Onslow case to reach parity (in expectation) with the wind-only case.

4.6 Onslow NPV

Given Onslow's assumed round-trip efficiency of 77%, we can compute the average annual operating profit for Onslow in each year. Since the order of wet years and dry years is important when computing the NPV for Onslow, we have simulated 1000 sequences of 25 years selected uniformly (and with replacement) from the years 1932 to 2020. Using a discount factor of 0.9, the simulated NPV varies considerably between sequences of hydrological years. However, no sequence recovers the estimated \$15.7b cost of constructing and commissioning Onslow, with only 1 out of the 1000 sequences recovering \$4b. The distribution of the NPVs after 25 years are shown in Figure 12a. Figure 12b shows the NPV over time, with the dark- and light-blue shaded regions representing the 25^{th} to 75^{th} and the 5^{th} to 95^{th} percentiles, respectively. This negative NPV means that (given all the assumptions of the model) the Lake Onslow pump-storage proposal would not be commercially viable in a perfectly competitive market.

It is instructive to understand the sequence for which Onslow attains the highest NPV (shown in red



(a) (b)

Figure 12: NPV distribution for Onslow: (a) after year 25 and (b) over time $ simulated over 1000 sequences of 25 hydrological years.

in Figure 12b). This sequence begins with two moderate years, allowing Onslow to fill partially, and then an extremely dry year (1932), where Onslow did not start with enough energy to avoid significant load shedding; this is shown in figures 13 and 14. Later in the sequence Onslow is emptied another time (over the 1950 and 1971 hydrological years). We see that Onslow makes a significant profit both times, since load-shedding is required to manage the energy shortfall, with prices rising sharply.

It is important to note that prices are low in most years; however, since the wind capacity is optimized to recover its capital costs from the spot market in expectation, the wind capacity is chosen such that a small proportion of years will experience an energy shortage with elevated prices. The majority of the capital costs will be recovered in the dry years where Onslow's starting storage is low. Due to the large, but rare, shortage events that we observe, there are significant financial (price) and physical (energy shortfall) risks for market participants.

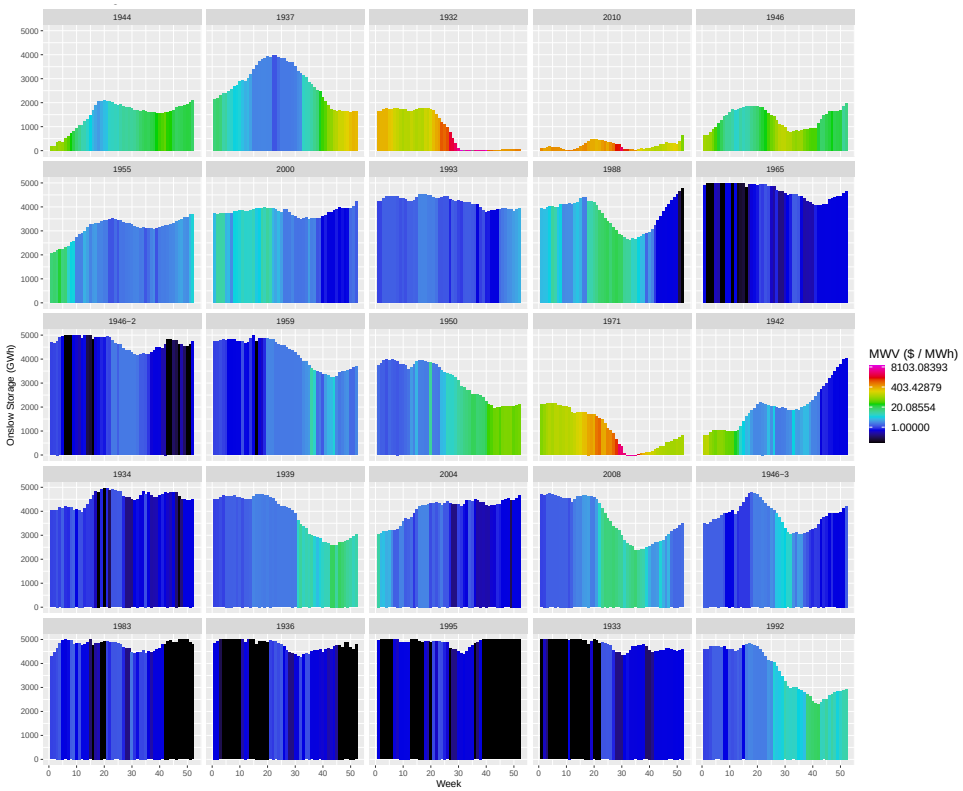


Figure 13: Onslow storage (GWh) and the corresponding marginal water value ($/MWh) over the most profitable sequences of hydrological years sampled.

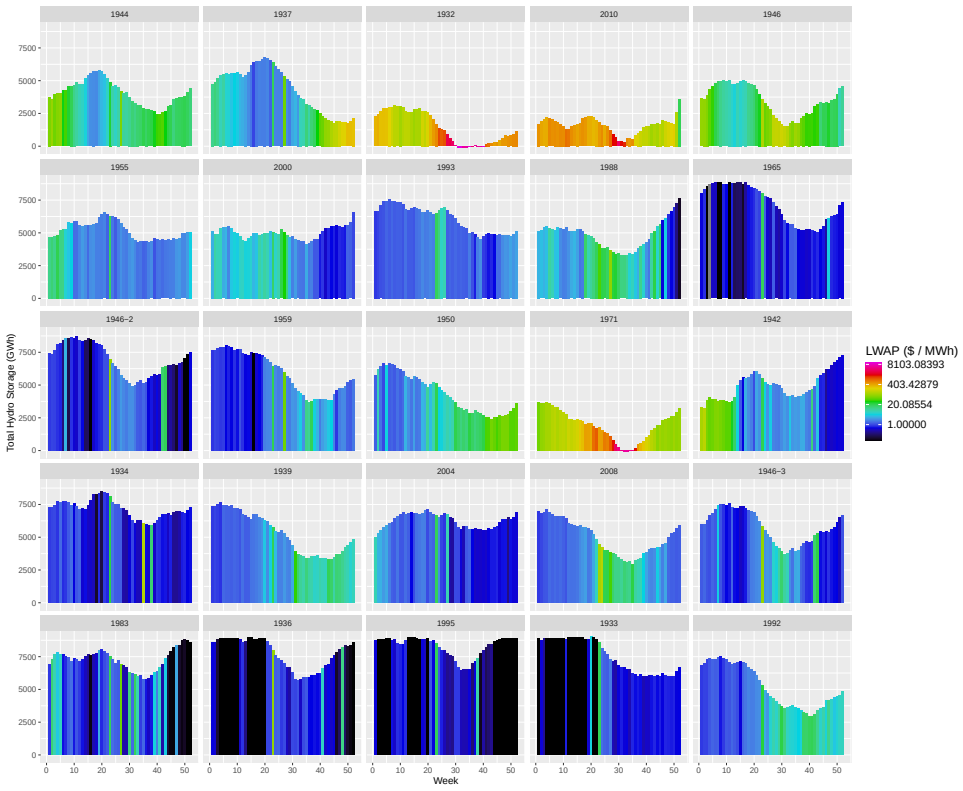


Figure 14: National storage (GWh) and the corresponding LWAP ($/MWh) over the most profitable sequences of hydrological years sampled.

5 Conclusions

This paper has described some experiments with a steady-state reservoir optimization model of the New Zealand electricity generation and transmission system. These experiments have been designed to evaluate the operation of a proposed pumped-storage scheme at Lake Onslow in the year 2035 and compare it with two counterfactual cases that operate without the scheme. The experiments we report on are limited in scope, and should ideally be repeated for other future years (e.g. 2050) and different assumptions about demand growth and risk aversion, to enable confident estimates to be made of the range of NPVs that might come from different investment options. Notwithstanding these limitations, the experiments show some interesting results that we now summarize.

5.1 Onslow decreases wholesale electricity prices

As shown in Table 7 the presence of Lake Onslow results in lower LWAPs than counterfactual cases. GWAPs for all generators are also lower than in counterfactual cases. Although this appears attractive from a distributional perspective, several observations are worth making.

1. The prices of course do not cover the capital cost of Onslow that must be recovered from other means.
2. The models assume that only wind investment is optimized. Capacity of existing renewable generating plant is assumed to remain in the market. It is possible that some of these plants have GWAPs lower than their LCOE, so they would either seek some capacity remuneration outside the wholesale energy market or be shut down.
3. More transmission capacity is needed to handle increased demand, with the largest transmission increase needed in the Onslow case.

5.2 Onslow depresses wind investment

The three models we have tested have different levels of additional wind investment. Wind capacity has been increased in each case to the point at which the GWAP covers the levelized cost of wind investment. With Onslow, this break-even point is at 3725 MW; without Onslow or peakers it is 4800 MW; and without Onslow but including peakers it is 3900 MW. The presence of Onslow has resulted in a lower level of wind investment than counterfactual models. As shown in Figure 8 the wind GWAP also has higher variance for the Onslow case than for the counterfactuals, so one might expect an even lower break-even level of wind investment in this case. This result runs counter to the intuitive conjecture that Onslow will provide a greater incentive for wind investment as pumping will provide a floor on wind GWAP. Figures 8 and 9 show that this floor can be very low especially in years where Onslow has plenty of water. The extra storage of Onslow lowers marginal water values in many (wet) years as shown by the plots in Appendix A.2; wind will not cover costs in the wettest of these years. In some years (e.g. 1978 and 2008 in A.2) Onslow experiences consecutive dry years and marginal water values become very high. Wind generation during these periods earns very high rents.

Because it lowers marginal water values, Onslow decreases wholesale electricity prices in most periods apart from those where lake levels are very low and wind is not blowing when prices are set by shortage. In contrast, the presence of green peakers can help reduce prices in peak periods irrespective of reservoir levels, but results in higher marginal water values since it has no Onslow storage. Given the same wind investment, GWAPs for wind will be lower in the Onslow case which incentivizes higher levels of wind investment when there are peakers. Observe that water values and hence GWAPS become even greater for the counterfactual without peakers or Onslow, leading to even greater incentive for wind investment.

5.3 Onslow reduces shortages

Figure 10 shows that Onslow generally reduces energy shortages in comparison with the other cases. However the figure shows that Lake Onslow does not eliminate the dry-winter energy problem if levels of wind investment are chosen to cover their LCOE. There will be inflow sequences (e.g. in consecutive dry years) where extreme energy shortages and very high prices will occur.

5.4 Further experiments

The results described in this paper are a first step towards a comprehensive analysis of the costs and benefits of the Lake Onslow scheme and its alternatives. A striking feature of our results is the subtlety of the interplay between the different components of the market, sometimes resulting in counterintuitive conclusions that arise from the effects of random inflows and wind in conjunction with a steady-state operating model. A stochastic optimization model like JADE is essential for understanding the resulting tradeoffs between different investment options.

We propose to continue this study with a suite of experiments that will enrich our understanding of this problem. We briefly describe these here.

1. The investment in wind has been calibrated so that the expected price over the simulated 89 years, approximately covers the annualised cost of wind capacity. It is important to note that this assumes that investors are risk-neutral with a fixed discount rate, which might be unrealistic. Moreover, the level of risk that wind investors face is not uniform across the cases investigated; for example, with Onslow built, the majority of years have low prices and a small number of years have very high prices. In addition, for the Onslow case, the expected price is very sensitive to the amount of wind. This occurs because Onslow provides significant inter-year storage, smoothing water values between years, meaning that the water value becomes tied to the long-run balance between energy coming into the system (wind, and inflows) and the demand. Since the first demand-response tranches shed load at $530, the average prices quickly change from close to $0 to near to $530, as the wind capacity is decreased. These risks associated with investment should be studied comprehensively.

2. In this study, for each case, the transmission grid has been augmented to ensure no congestion in 99.9% of periods. Depending on the cost of transmission line upgrades, it may be cheaper to allow for congestion to occur more often. This can be investigated using the JADE model, but some additional assumptions about costs and capacities of transmission upgrades must be made.

3. The steady-state distribution of water values has been computed using a discount factor of 0.9. The choice of this value has been made for convenience. A comprehensive analysis of the industry should enable us to provide a more defensible estimate of the true discount factor to be used in comparing net present values.

4. Reserve capacity has not been considered in this study. JADE does not currently include reserve in its weekly representation of a competitive spot market. However, from this study we see that, due to the shut down of thermal plant, shortages of generation capacity may become more frequent, and an accurate estimate of reserve prices may be needed when determining the optimal investment in new capacity.

5. A final investment decision on Lake Onslow is not expected until 2026, so given a 9-year build time it will not be available for pumping until after 2035. The 2035 study carried out here should be repeated for estimated demand in 2040 and 2050.

Acknowledgements

We would like to acknowledge the contributions of University of Auckland Engineering Science students Benjamin Best, Jamie Chen, Ryan Delaney, Cal Roughan, Connor Roulston, Michael Soffe and Harry Thurman in developing preliminary versions of the models described in this report, and acknowledge Oscar Dowson for discussions on SDDP.jl.

References

[1] W.E. Bardsley. Note on the pumped storage potential of the Onslow-Manorburn depression, New Zealand. *Journal of Hydrology (New Zealand)*, 44(2):131–135, 2005.

[2] D.P. Bertsekas. *Dynamic programming and optimal control: Volume I*. Athena Scientific, 2012.

[3] J. Bezanson, A. Edelman, S. Karpinski, and V.B. Shah. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017.

[4] O. Dowson. The policy graph decomposition of multistage stochastic optimization problems. *Networks*, 76:3–23, 2020.

[5] O. Dowson and L. Kapelevich. SDDP.jl: a Julia package for stochastic dual dynamic programming. *INFORMS Journal on Computing*, 33(1):27–33, 2021.

[6] I. Dunning, J. Huchette, and M. Lubin. JuMP: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320, 2017.

[7] M.C. Ferris and A.B. Philpott. Dynamic risked equilibrium. *Operations Research*, 70(3):1933–1952, 2022.

[8] He Pou a Rangi Climate Change Commission: Data and modelling. *https://www.climatecommission.govt.nz/get-involved/sharing-our-thinking/data-and-modelling*, 2021.

[9] M.K. Majeed. *Evaluating the potential for a multi-use seasonal pumped storage scheme in New Zealand's South Island*. PhD thesis, The University of Waikato, 2019.

[10] MBIE: New Zealand Battery Project. *https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/low-emissions-economy/nz-battery/*, 2022.

[11] MBIE: Terms of Reference for Energy Strategy. *https://www.mbie.govt.nz/dmsdocument/25373-terms-of-reference-new-zealand-energy-strategy*, 2022.

[12] National Institute of Water and Atmospheric Research (NIWA): Solarview. *http://solarview.niwa.co.nz/*, 2008.

[13] New Zealand Wind Energy Association. *https://www.windenergy.org.nz/operating-&-under-construction*, 2023.

[14] M.V.F. Pereira and L.M.V.G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1-3):359–375, 1991.

[15] A.B. Philpott and Z. Guan. On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters*, 36(4):450–455, 2008.

[16] A.B. Philpott and G. Pritchard. EMI-DOASA 4.0. *https://www.epoc.org.nz/doasa.html*, 2018.

A Total energy storage

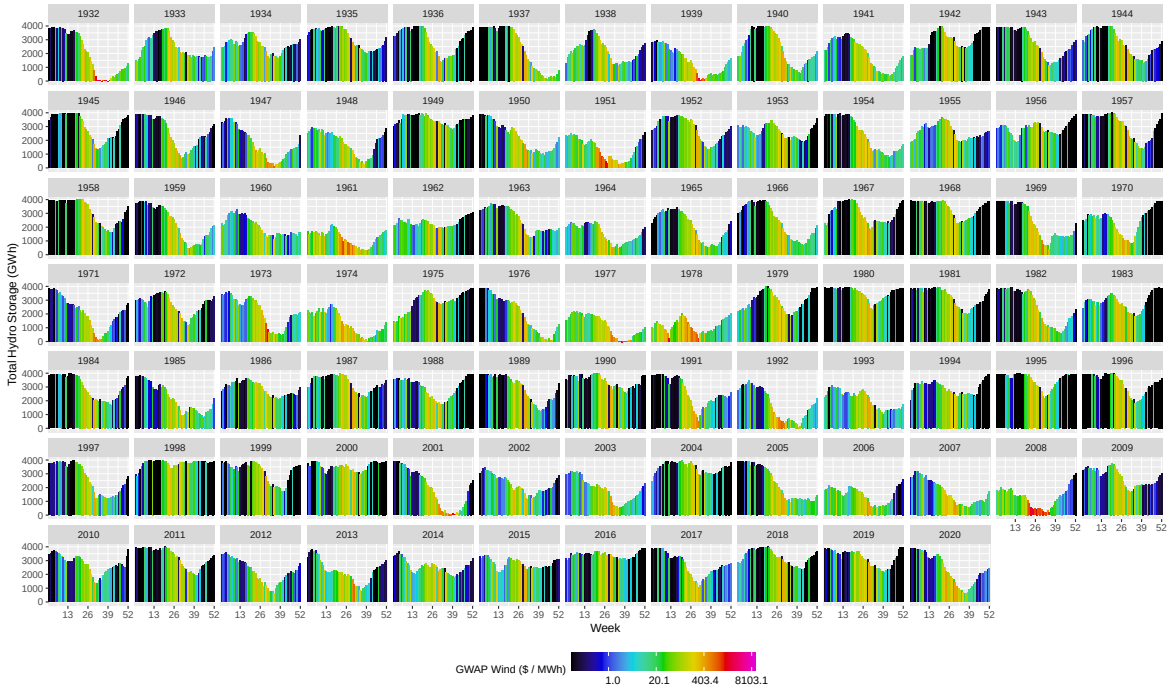


Figure 15: Total storage (GWh) over 89 historical hydrological years for the wind-only case, coloured by the weekly GWAP for wind generation.

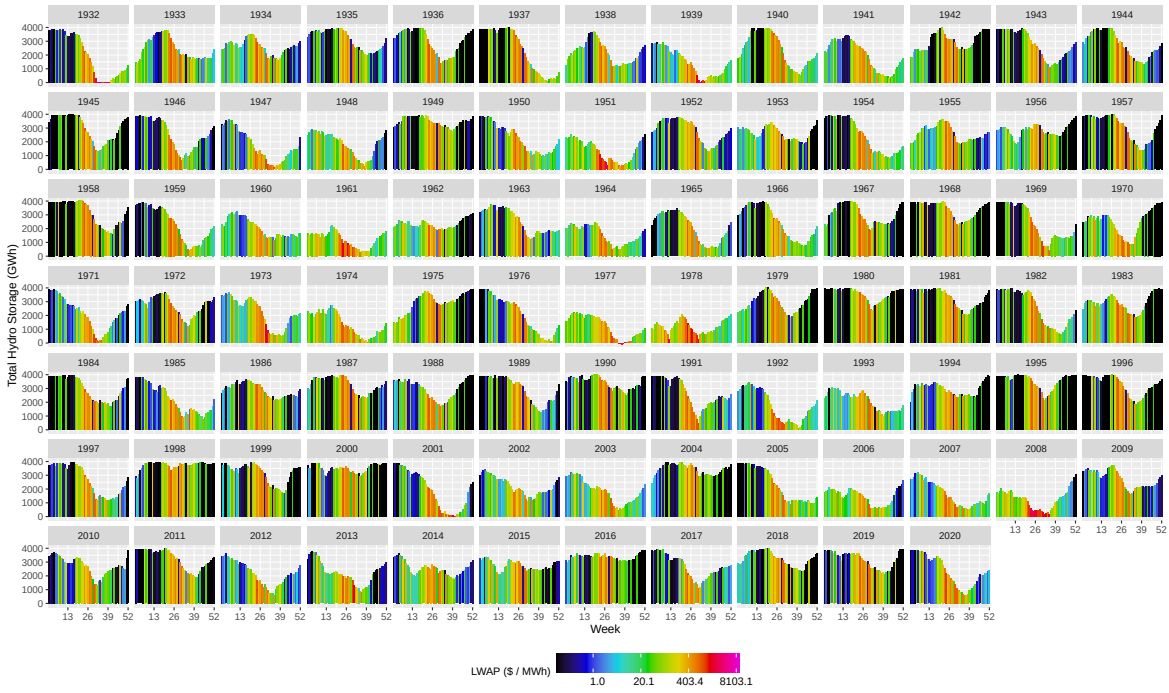


Figure 16: Total storage (GWh) over 89 historical hydrological years for the wind-only case, coloured by the weekly LWAP.

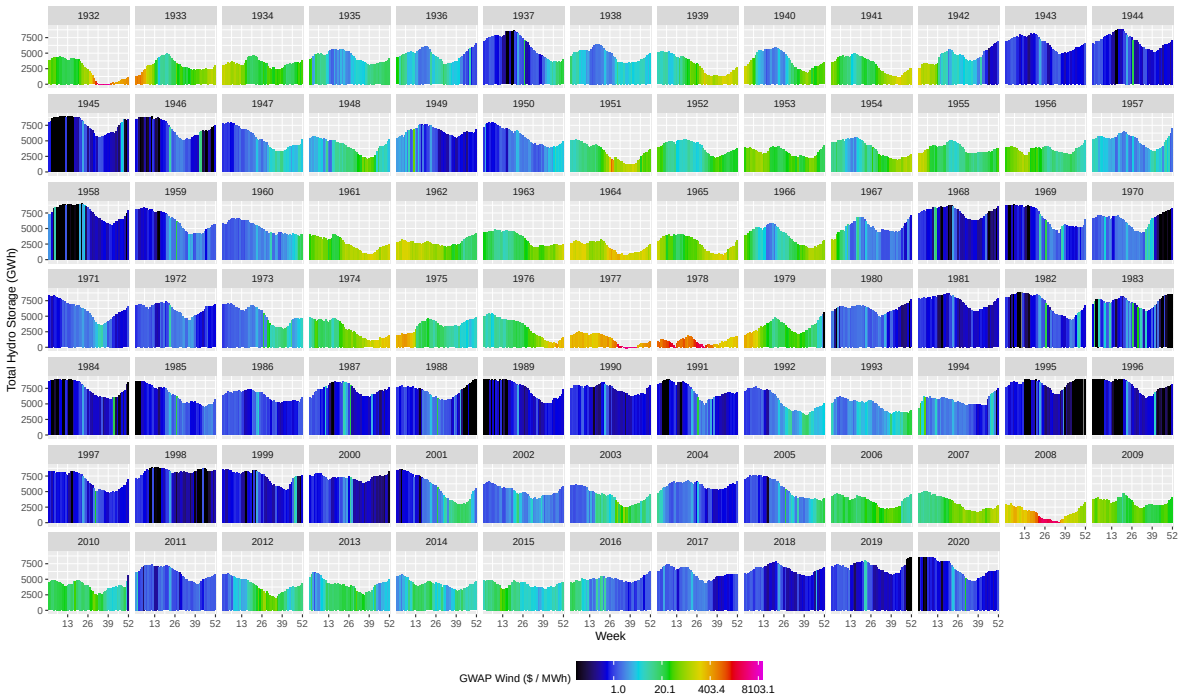


Figure 17: Total storage (GWh) over 89 historical hydrological years for the Onslow case, coloured by the weekly GWAP for wind generation.

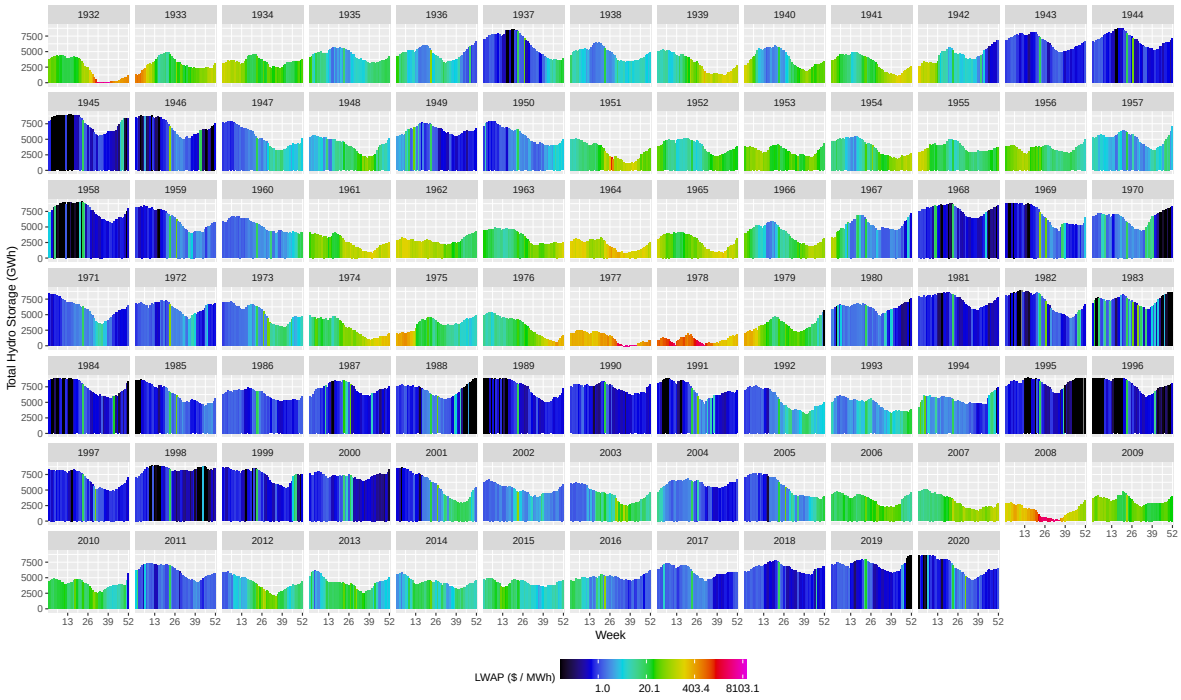


Figure 18: Total storage (GWh) over 89 historical hydrological years for the Onslow case, coloured by the weekly LWAP.

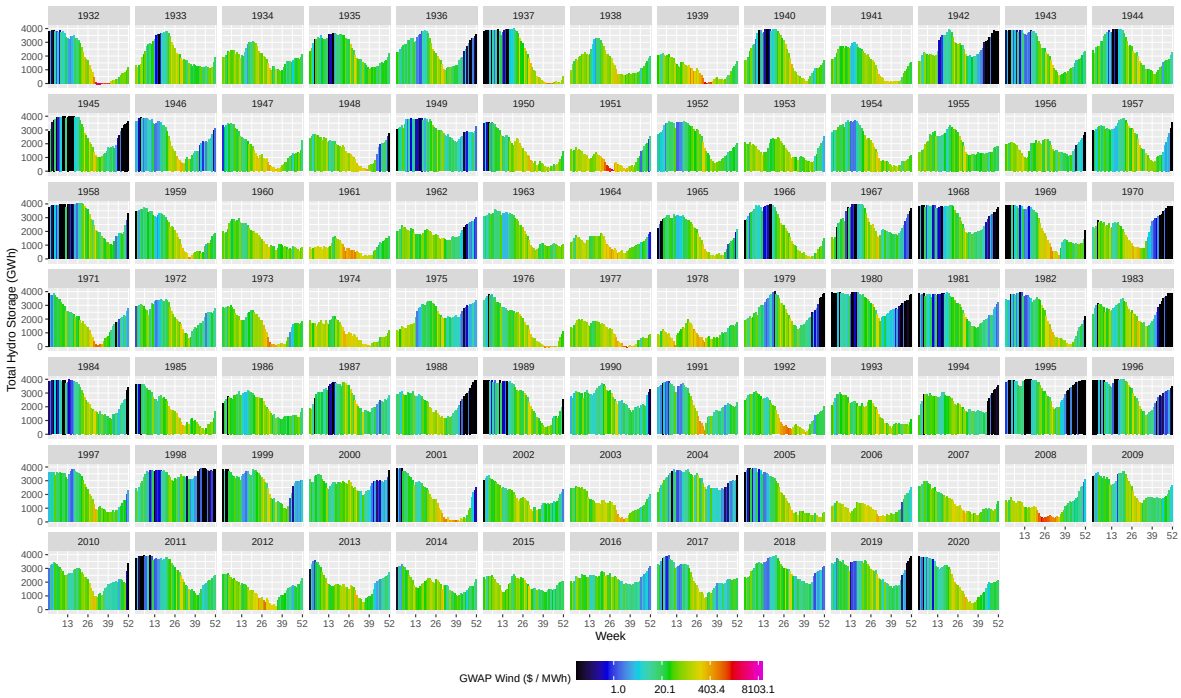


Figure 19: Total storage (GWh) over 89 historical hydrological years for the green peakers case, coloured by the weekly GWAP for wind generation.

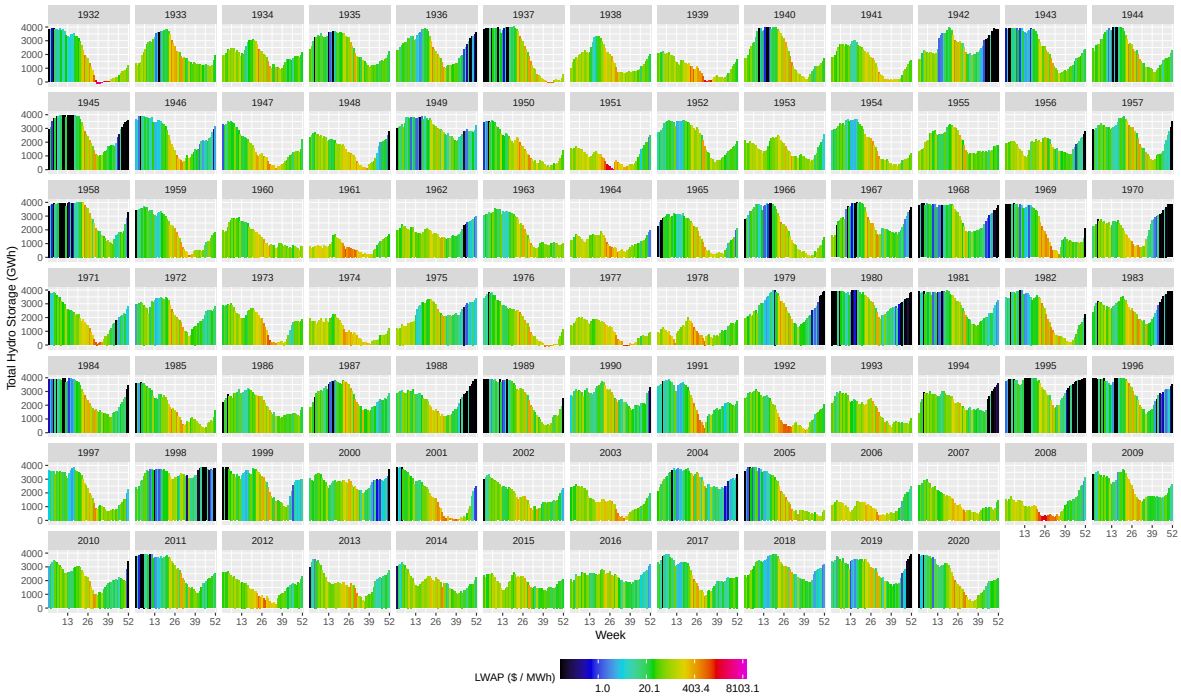


Figure 20: Total storage (GWh) over 89 historical hydrological years for the green peakers case, coloured by the weekly LWAP.

B Onslow Operations

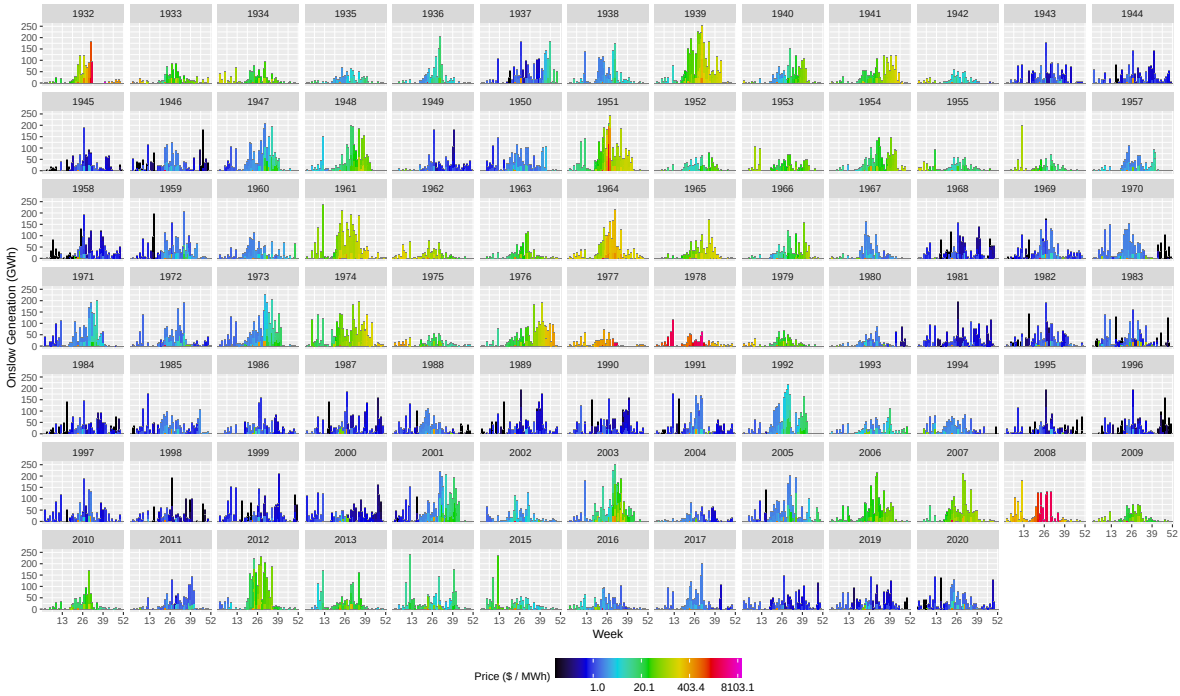


Figure 21: Onslow generation (in GWh produced) over 89 historical hydrological years, coloured by price received.

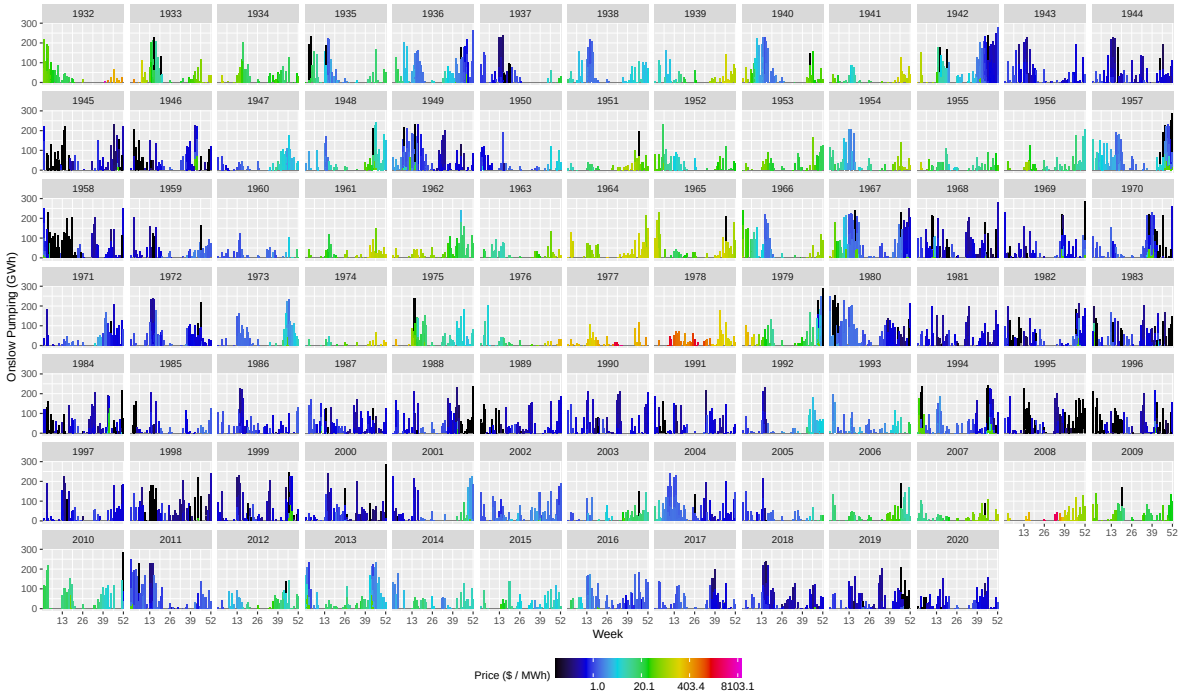


Figure 22: Onslow pumping (in GWh consumed) over 89 historical hydrological years, coloured by price paid.

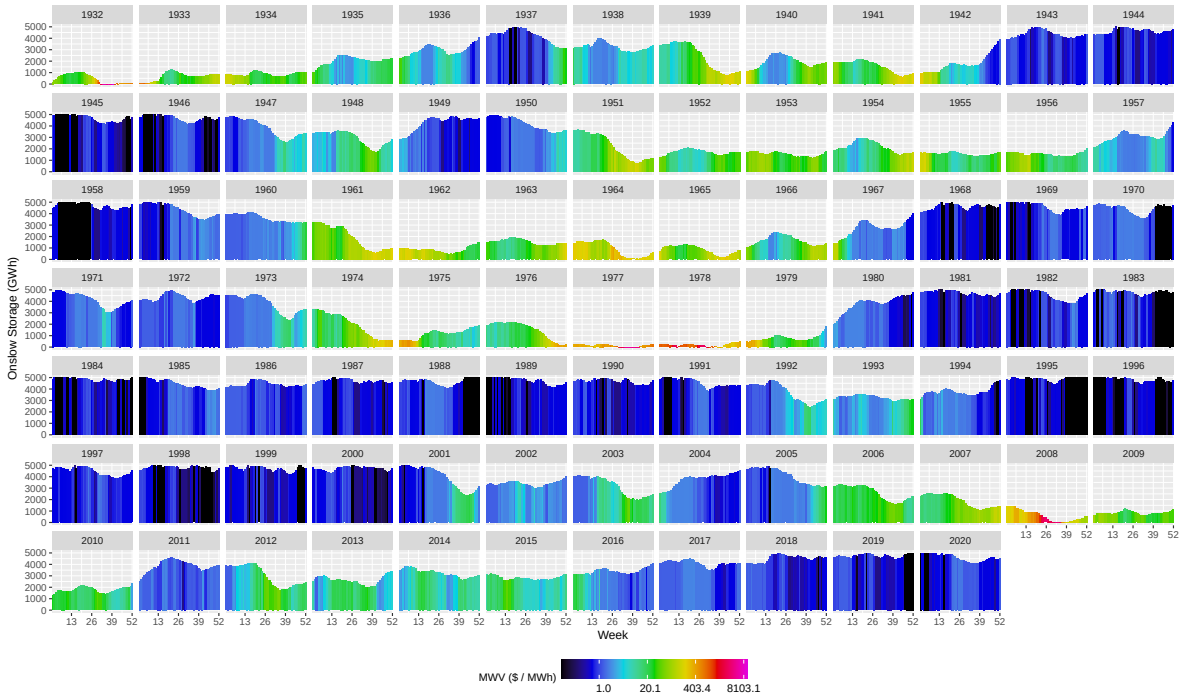


Figure 23: Onslow storage (in GWh) over 89 historical hydrological years, coloured by the marginal water value at Onslow.

C Green Peaker Operations

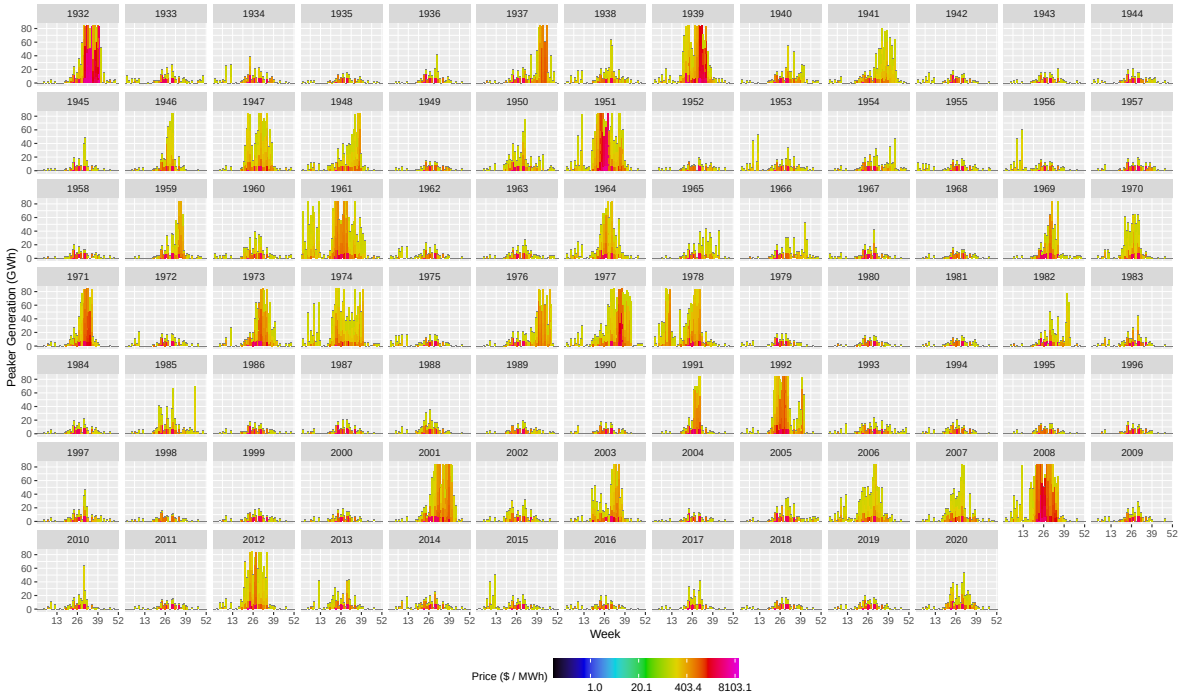


Figure 24: Green peaker generation (in GWh) over 89 historical hydrological years, coloured by price received.

D Load-weighted Average Prices

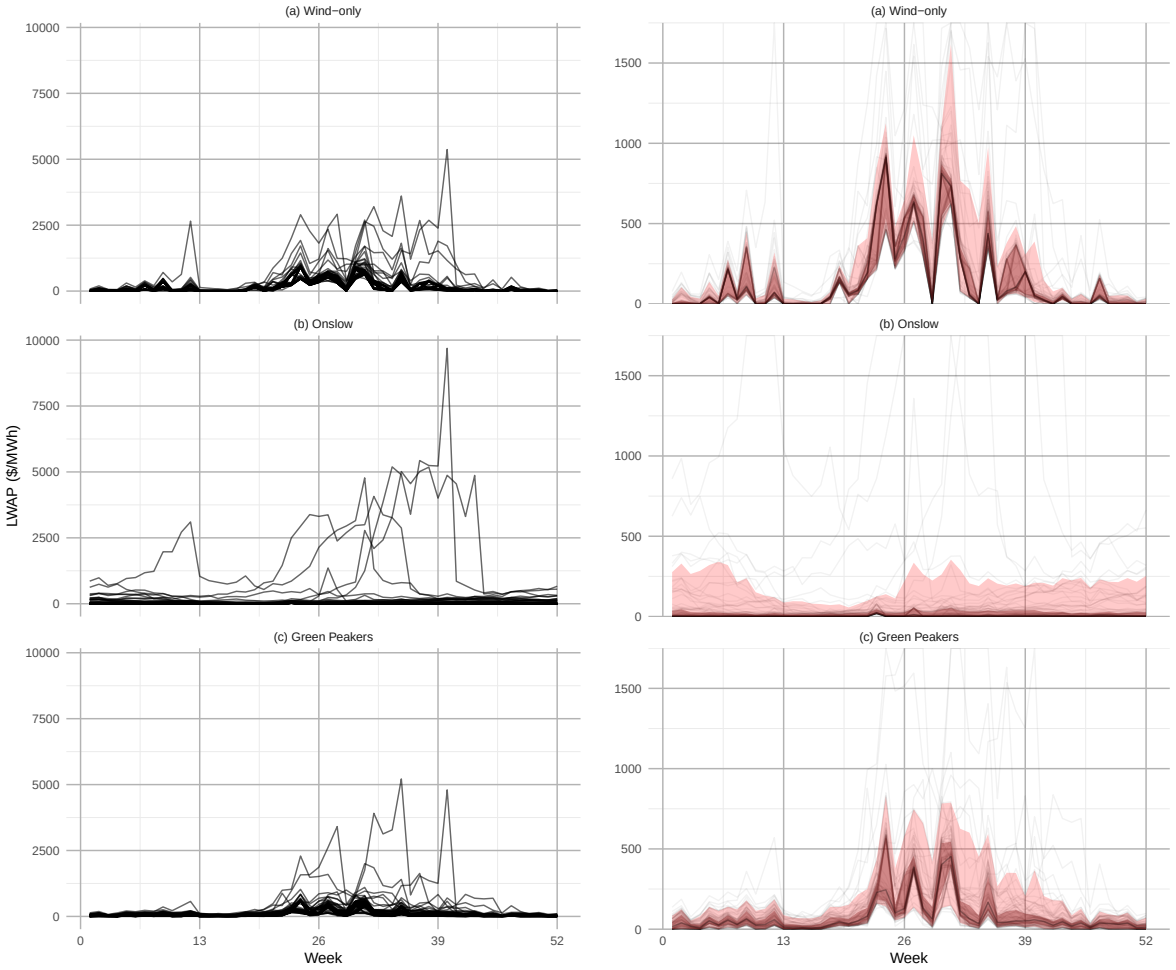


Figure 25: LWAP ($/MWh) simulated from 1932 to 2020 for each case, on the left the full price range is shown; on the right, prices are truncated at $1600/MWh.

Figure 26: LWAP ($/MWh) simulated from 1932 to 2020 for each case.

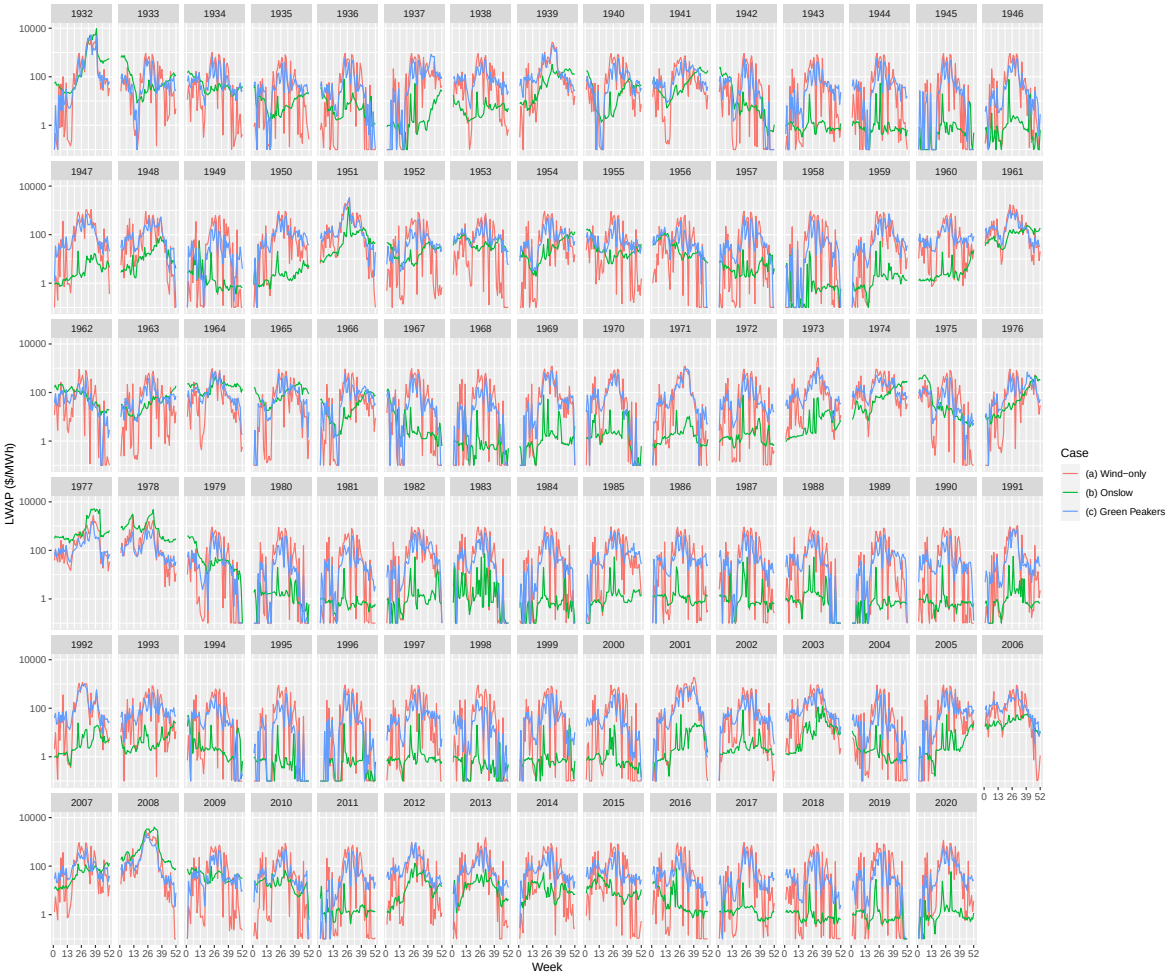


Figure 27: LWAP ($/MWh) simulated from 1932 to 2020 for each case (shown on a log scale).