Submitted to *Operations Research*
manuscript (Please, provide the manuscript number!)

100% renewable electricity with storage

Michael Ferris
Computer Sciences Department and Wisconsin Institute for Discovery, University of Wisconsin, Madison, WI 53706, ferris@cs.wisc.edu

Andy Philpott
Electric Power Optimization Centre, University of Auckland, New Zealand, a.philpott@auckland.ac.nz

We formulate and compare optimization and equilibrium models of investment in renewable generation under uncertainty and risk. The main focus of the paper is a suite of social planning models that compute optimal generation capacity investments for a hydro-dominated electricity system. The models allow for investments in hydro, geothermal, solar, wind, and thermal plant, as well as battery storage for smoothing load profiles. The aim of each model is to optimize the risk-adjusted cost of capacity expansion and operation in a two-stage stochastic program that incorporates uncertain seasonal hydroelectric energy supply and short-term variability in renewable supply. The models are applied to data from the New Zealand electricity system and used to estimate the costs of moving to a 100% renewable electricity system by 2035. Outputs from the models are compared with those from competitive equilibrium models with differing ownership structures to explore the impact of incomplete markets for risk. We also explore the outcomes obtained when applying different forms of CO_2 constraint that limit respectively non-renewable capacity, non-renewable generation, and CO_2 emissions on average, almost surely or in a chance-constrained setting.

Key words: renewable electricity, carbon price, climate change, risk aversion

History: This paper was first submitted on May 9, 2019.

1. Introduction

In this paper we study investment in renewable generation under uncertainty and risk. There has been much attention devoted to capacity planning models to design electricity systems that will deliver (close to) 100% renewable electricity. In most electricity systems renewable energy comes

from either wind power or solar energy. Since this is intermittent, the planning problem must account for uncertainty. Renewable energy can also come from hydroelectricity and geothermal power. Inflows to hydroelectricity reservoirs are uncertain, albeit at a different time-scale, but there is some control over their release. Run-of-river hydroelectric plants on the other hand convert uncertain inflows into uncertain levels of energy. Geothermal power is more predictable (although exploration involves significant uncertainties), but it is a base-load energy source. Finally nuclear power is typically treated as non-renewable, but can be a useful technology if greenhouse gas emissions are to be reduced.

When the generation of electricity comes from sources with uncertain fuel supply, some care must be taken in defining 100 % renewable. The strictest definition would admit no non-renewable generation under any realization of the uncertainty. One would expect this to be very expensive in terms of both capital investment and operating cost. In New Zealand, the Government is seeking an electricity system that is 100% renewable in a normal hydrology year. This could be interpreted in a number of ways. One option is to model this as a chance constraint: i.e. the probability of any year having 100% renewable generation is at least $1-\alpha$, where α is chosen appropriately. We show in our calculations that such a policy is not guaranteed to have low emissions when averaged over all possible hydrological years, while being more expensive in terms of capital and operating costs.

In this paper we develop a suite of models for investment in renewable generation that incorporate 100 % renewable constraints in a variety of ways. We study two forms of storage that might make such a constraint easier to satisfy. Within a day, battery storage can be used to transfer energy between time periods. This can be used to accommodate more intermittent energy (wind and solar) that otherwise would be wasted. Over a longer time horizon, hydroelectric reservoir storage can be used to transfer energy from season to season to account for low seasonal inflows. We provide mathematical models for both these phenomena, and demonstrate their usefulness on an example based on New Zealand data.

In most industrialized countries, electricity is supplied by large generating companies operating in competitive markets. Since these companies will make the investments in new renewable energy,

the plans computed from models such as the one in this paper should be aligned with the profits made by these investments. If the investors are risk neutral then a risk-neutral social plan can be shown to correspond to the competitive equilibrium.

This is not necessarily true if investors are risk averse or behave strategically. However, if electricity companies are price takers with similar risk measures, and a complete market of instruments for trading risk is available, then there is a (theoretical) correspondence between a competitive equilibrium and a risk-averse social plan, where the risk measure of the social planner emerges from those of the investors (see Ralph and Smeers 2015, Ferris and Philpott 2018). So there is an argument that an optimal social plan, with an appropriately chosen risk measure could produce the same investments as an equilibrium in a perfectly competitive market with risk averse agents where all the risks are priced.

1.1. Conventional capacity expansion models

Classical electricity capacity planning models (Stoft 2002, Joskow 2006) use a *screening curve* (Figure 1) to rank generation options by their long-run marginal cost (LRMC), thus finding the best option to serve the production profile for each additional demand unit. The screening curve shows the annual total cost per MW capacity plotted against the number of annual operating hours. The total cost is a combination of fixed and variable cost based on the number of production hours in a year. A minimum cost for each capacity factor can be found by combining the screening curve with the *load duration curve* (LDC), here approximated by 10 load blocks with piecewise constant demand. The projection produces the least-cost capacity combination that can serve the load profile. For example, to supply the part of the LDC that has higher capacity factor (i.e. running most of the year), base load is the least cost option. As the number of operating hours decreases, the plants that are less expensive to build but more costly to run start to become more economical. For a small number of hours at the tip of the duration curve, high variable cost peakers are the most economical.

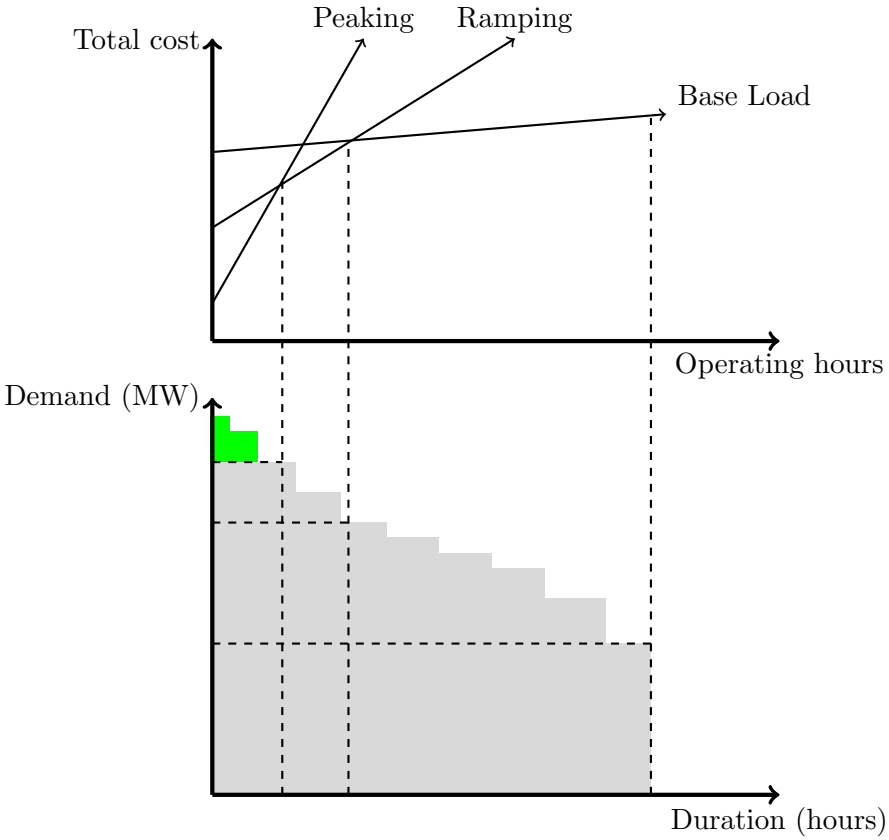


Figure 1 The screening curve: how capacity is traditionally planned in electricity systems.

Observe from Figure 1 that the optimal peaking plant capacity does not cover all of the peaking demand. There is a green area of peak load that will not be met. Since we assume that load is inelastic, the load in this triangle will be shed. When this happens electricity price will take the value of lost load (VOLL). In a market setting, the difference between VOLL and the short-run marginal cost of peaking generation provides the necessary rent to cover the fixed cost of the peaking plant (as well as contributing to rent on all the other plants).

In simple cases, the screening curve solution can be found by inspection. When the LDC is piecewise constant, it is the solution to the linear program

$$\text{LP: } \min \textstyle\sum_{k\in\mathcal{K}} M_k z_k + \sum_{b\in\mathcal{B}} H_b \left(\sum_{k\in\mathcal{K}} C_k y_{k,b} - V(d_b - r_b)\right)$$

$$\begin{aligned} \text{s.t.} \quad & 0 \le y_{k,b} \le z_k, && k\in\mathcal{K}, b\in\mathcal{B}, \\ & 0 \le r_b \le d_b, && b\in\mathcal{B}, \\ & d_b \le \textstyle\sum_{k\in\mathcal{K}} y_{k,b} + r_b, && b\in\mathcal{B}, \end{aligned}$$

where

- $k \in \mathcal{K}$ denotes different generating technologies;
- $b \in \mathcal{B}$ indexes load blocks where H_b denotes the number of hours in block b, and $\sum_{b\in\mathcal{B}} H_b$ gives the number of hours in a year;
- the variable z_k is the capacity invested in technology k;
- variable cost for technology k is defined as C_k respectively;
- annual fixed cost for technology k is defined as M_k;
- the load in load block b is d_b, and value of lost load is V;
- $y_{k,b}$ denotes the production (MW) of technology k in each hour in load block b;
- r_b denotes the load shed (MW) in each hour in load block b.

The model LP can be enhanced by including different locations $i \in \mathcal{N}$ and transmission variables f that transfer power between locations. The variables f satisfy constraints $f \in \mathcal{F}$, that might represent the constraints of a DC-Load flow model, for example.

Formulations like LP appeared as early as the 1950s (Masse and Gibrat 1957), although the basic formulation has been extended in the past two decades to include operational constraints, and some supply-side uncertainties such as plant outages and technological changes were added. Demand distributions are still represented by load duration curves, or their discretized versions (De Jonghe et al. 2012). A number of authors (see e.g. Bishop and Bull 2008) have extended LP to include binary capital planning decisions (so z_k takes on discrete values) and planning over multiple years to accommodate growing demand.

In this paper we explore the extent to which uncertainty and risk affects the screening-curve approach. With uncertain supply from renewable energy, Figure 1 no longer represents the optimal capital planning problem which becomes a stochastic linear program. There are different levels of detail one can apply to this problem. For example, Wu et al. (2017) and Khazaei and Powell (2017) use dynamic programming subproblems to model optimal thermal ramping investments that are needed with increased wind and solar generation.

Our interest is in markets with a lot of stored hydroelectricity (such as New Zealand), that provides plenty of ramping capacity. We focus here on supply adequacy when reservoir inflows, wind, and solar energy are uncertain or variable. To do this we shall use New Zealand as a case study and build up a new screening curve stochastic optimization model for capacity planning that accommodates this.

Our ultimate goal is to expose the effects of different assumptions on the outcomes of the models. Imposing regulatory constraints (such as dictating 100% renewable capacity) without careful thought can lead to unforeseen outcomes. Our model shows that one should focus on constraining CO_2 emissions: the equilibrium outcomes then give the mix of generation plant that one might expect to see.

1.2. Relation to previous work

Not suprisingly, models for studying the decarbonization of energy systems are receiving considerable attention in the literature. Many of these models (for example Graf and Marcantonini 2017) focus on the intermittency of renewables and the effect of this on backup thermal generation and/or storage. The investment paths of these models are either prescribed in advance or simulated by estimating net present values of candidate investments at each stage and then incrementing the model by one time step with selected investments in place.

Our model is closer in spirit to the classical system planning models such as MARKAL (Fishbone and Abilock 1981) and its modern implementation in the TIMES system (Loulou and Labriet 2008, Loulou 2008). Other similar planning models are ReEDS (Short et al. 2011) and GEM (Bishop

and Bull 2008). Our model extends these to include uncertainty in operations. In its simplest form, this gives a two-stage model in which stage one invests in capacity and stage 2 operates this in different states of the world. A multistage version would invest in capacity over several stages, and in each stage operate the system subject to the realized uncertainty in operating conditions.

A number of authors have developed models similar to ours. In the United States, Boffino et al. (2018) study the effect of emissions reduction in ERCOT, the Texas electricity market. Their model includes wind variation, but Texas has no hydro reservoirs, so the ERCOT model does not model uncertainty in energy supply, a criterion that is critical for New Zealand. Similar models to that in Boffino et al. (2018) have emerged for Europe, for example the EMPIRE model for capacity expansion developed by Skar et al. (2014). Similar to our model, EMPIRE restricts capacities of generators using a stochastic availability factor (for e.g. wind and run-of-river plant), but their treatment of hydro storage is more simple than ours, constraining total reservoir hydro generation for each reservoir by a seasonal energy constraint.

In the following sections we describe our model as it is applied to the New Zealand electricity system. To our knowledge there are no stochastic optimization models of the decarbonization of the New Zealand system. Mason et al. (2010, 2013) examine New Zealand historical generation, and estimate levels of renewable generation capacity needed to replace historical thermal generation. Their results are calibrated to historical outcomes in the years 2005-2010. Although the study included a very dry year (2008) any future year with a different (and unknown) hydro inflow sequence and different demand levels might require more renewable capacity than indicated in Mason et al. (2013). Our model explicitly includes future demand forecasts and stochastic variation in inflows and wind.

The contributions of our paper are as follows.

1. We describe how to incorporate uncertain seasonal hydroelectricity energy supply into a two-stage model.
2. We demonstrate the effect of including risk aversion into the social planning solution using a coherent risk measure.

3. We illuminate the differences between applying different forms of CO_2 constraint. In particular we investigate constraints on non-renewable capacity, non-renewable generation, CO_2 emissions, and a chance constraint that limits the frequency of years that use non-renewable generation.
4. We create both the first stochastic system optimization model and the first stochastic equilibrium of the New Zealand electricity system, explicitly targeted to CO_2 emission reduction.

The paper is laid out as follows. In the next section we describe deterministic and stochastic social planning models for investment in electricity systems with stored hydro. In section 3 we describe three approaches to model the reduction in CO_2 emissions. Section 4 describes the calibration of our model to the New Zealand electricity system, and section 5 presents some selected results from applying this calibrated optimization model to investigate some specific questions arising from the New Zealand Government policy of a 100% renewable electricity system by 2035. Section 6 introduces competitive models of storage and provides contrasting results for the equilibrium model compared to the optimization setting. Section 7 concludes the paper with a discussion of results.

2. Social planning model for investment

2.1. Deterministic social planning model

The first model we consider will be a simple version of LP for the New Zealand wholesale electricity market based on data collected for 2017. In this model there is no storage available (i.e. no stored hydro or batteries).

We solve

$$
\begin{aligned}
\text{P: min} \quad & \psi = \textstyle\sum_{k\in\mathcal{K}} \left(K_k(x_k) + L_k(z_k)\right) + Z \\
\text{s.t.} \quad & Z = \textstyle\sum_{b\in\mathcal{B}} H_b \left(\sum_{k\in\mathcal{K}} C_k y_{k,b} - V(d_b - r_b)\right), \\
& 0 \le x_k \le u_k, && k \in \mathcal{K}, \\
& 0 \le z_k \le x_k + U_k, && k \in \mathcal{K}, \\
& 0 \le y_{k,b} \le z_k, && k \in \mathcal{K}, b \in \mathcal{B}, \\
& 0 \le r_b \le d_b, && b \in \mathcal{B}, \\
& d_b \le \textstyle\sum_{k\in\mathcal{K}} y_{k,b} + r_b, && b \in \mathcal{B}.
\end{aligned}
$$

This is the same model as LP except that here we account for existing capacity using the parameter U_k, and we split the annual fixed cost M_k into a capital investment cost K_k and an annual maintenance cost $L_k(z_k)$ on existing and new capital. Note that we could also adjust the cost data C_k to account for a constraint on renewables, perhaps using a carbon tax, but we will defer that discussion until later in this paper.

2.2. Stochastic planning model data

We extend the above analysis to account for uncertainty. The uncertainty will manifest itself in various ways that we will endeavour to model in a two-stage stochastic modeling framework. Our approach to modeling uncertainty closely follows that of Kok et al. (2018). Since the types of uncertainty we consider have different effects at distinct time-scales, we need two sets of parameters to deal separately with the short-term uncertainty in wind and run-of-river generation, and the medium-term uncertainty in reservoir storage. We do this by introducing a scenario index $\omega \in \Omega$ that represents a random future state of the world. In general, each state of the world ω is a vector of random outcomes. In our model this vector is two dimensional where the first component ω_1 corresponds to a certain type of year, and the second component ω_2 relates to parameter variation in each season t of a year. The effect of these outcomes on the model is represented for each region i by the set of random parameters $\mu_{k,i,b,t}(\omega), \nu_{k,i,t}(\omega), k \in \mathcal{K}, b \in \mathcal{B}$.

The parameter $\mu_{k,i,b,t}(\omega)$ for technology k in region i and season t denotes a proportional reduction in its capacity in load block b and random event ω. Here μ is used in constraints on the generation (recourse variables) $y_{k,i,b,t}(\omega)$ that hold in each of the scenarios ω:

$$y_{k,i,b,t}(\omega) \leq \mu_{k,i,b,t}(\omega) z_{k,i}.$$

Note that $z_{k,i}$ is now indexed by both region i and technology k. Some examples will help illustrate the model.

If $k = 1$ is run-of-river hydroelectricity generation then $\mu_{k,i,b,t}(\omega_1, \omega_2) = \mu_{1,i,b,t}(\omega_1)$, a parameter that depends only on the type of year being experienced that scales down the nominal capacity of

the generators in region i to reflect the inflows that occur in season t of that year. For run-of-river plant with intra-day flexibility, the capacity factor $\mu_{1,i,b,t}(\omega_1)$ can depend on the load block, being greater than average in peak times and less in offpeak times, while averaging out to the value corresponding to the season and type of year. We represent this by scaling a base value $\hat{\mu}$ for the plant by a block dependent parameter α, so $\mu_{1,i,b,t}(\omega_1) = \hat{\mu}_{i,t}(\omega_1)\alpha_{i,b,t}$. Values of α can be estimated from historical dispatch records.

If $k = 2$ is wind power, then we might assume that the distribution of wind in a given load block does not depend on that block. (Of course in some cases this will not be true, for example a sea breeze might occur in coastal town in periods close to the evening system peak.). The simplest model assumes that the wind contributes uniformly to the hours in a load block in a region i and season t with a constant load factor $\mu_{2,i,t}(\omega_1)$ that might have several realizations estimated from historical wind generation in each load block. For small amounts of wind this approach is sufficient, but as wind capacity grows, the model becomes less realistic. Historical wind generation data contain many periods with no wind at all, so assuming an average load factor will give a smoother picture than reality, especially in peak periods when a suddden lack of wind requires peaking plant to be dispatched. We approximate this by including scenarios ω_2 with $\mu_{2,i,b,t}(\omega_1, \omega_2) = 0$ for the peak load block ($b = 1$), having probabilities $p_t(\omega_2)$ estimated from the frequency of historical no-wind hours in the load block. Observe that p_t depends on t, which allows it to vary with season.

If $k = 3$ is solar power then we might assume for example that the insolation depends only on the time of day. (Of course in some cases it will depend on random weather.) Then $\mu_{3,i,b,t}(\omega) = \mu_{3,i,b,t}$ depends only on the load block, region and season. For a technology like stored hydro, say $k = 4$, $\mu_{4,i,b,t}(\omega) = 1$, unless ω_2 corresponds to an outage event with probability α_k, when $\mu_{4,i,b,t}(\omega_1.\omega_2) = 1 - \alpha_k$, the probability that the station is at full capacity.

The parameter $\nu_{k,i,t}(\omega)$ for technology k denotes a proportional reduction in its total annual energy production in region i in season t in a year of type ω. If, as before, $k = 4$ is generation from stored hydroelectric power then the event ω could correspond to lower than average reservoir

inflows over a year. In this dry-year event we can still run the reservoir hydro-station turbines at 100% of their capacity, but not for the whole year. In New Zealand, $\nu_{k,i,t}(\omega)$ for stored hydro is between 0.4 and 0.6, which is the range of capacity factors for hydro stations. If $k=5$ is thermal plant then $\nu_{5,i,t}(\omega)$ can model random fuel stockpile levels, but typically we assume $\nu_{5,i,t}(\omega)=1$. Some values of $\mu(\omega)$ and $\nu(\omega)$ estimated from New Zealand historical data are given in Appendix A below.

2.3. Hydroelectric storage

We now complete the model by including storage and transmission variables. Electricity systems with stored hydroelectricity transfer energy from seasons with high inflows (e.g. from snowmelt) to seasons with low inflows. In New Zealand this generally correponds to a transfer of energy from a wet summer to a potentially dry winter. This transfer is profitable not only because winter supply of energy is smaller than summer, but energy demand in winter tends to be higher than in summer. The most realistic models for optimizing this transfer of energy use stochastic dynamic programming to optimize reservoir releases when inflows are uncertain. We approximate such a model by adding a time index $t=0,1,2,\ldots,T-1$ to the model, where typically t will denote a season of the year (so $T=4$). We let s_t denote the water transferred from the end of time interval t to the beginning of time interval $(t+1)$ mod T, where we interpret $t=-1$ as the period $T-1$ of the previous year. Each time interval t can be broken up into load blocks $b\in\mathcal{B}_t$, each having H_b hours. The total number of hours in each time interval is then $\sum_{b\in\mathcal{B}_t} H_b$.

The hydroelectric storage equations use $\nu_{k,i,t}(\omega)\sum_{b\in\mathcal{B}_t} H_b z_{ki}$ which is a measure of the total energy available for hydro generation in time interval t in scenario ω. Here $\nu_{k,i,t}(\omega)$ can be estimated from the total possible energy $W_t(\omega)$ that could be produced from stored hydro in time interval t in scenario ω. Given a historical year ω we let

$$\begin{aligned} W_t(\omega) = {} & \text{historical hydro generation in } (t,\omega) + \text{ energy stored at the end of } t \\ & - \text{ energy stored at the start of } t, \end{aligned}$$

and set

$$\nu_{k,i,t}(\omega) = \frac{W_t(\omega)}{\sum_{b\in\mathcal{B}_t} H_b z_{k,i}}.$$

In this case, the generation variables are now indexed additionally by t, and the controlling constraint on hydroelectric storage s_t has the following form:

$$\sum_{b\in\mathcal{B}_t} H_b y_{k,i,b,t}(\omega) \leq \nu_{k,i,t}(\omega) \sum_{b\in\mathcal{B}_t} H_b z_{k,i} - s_{i,t} + s_{i,t-1}\,, \tag{1}$$

along with other physical constraints modeled using $s \in \mathcal{H}$. Observe that (1) does not define $s_{i,t}$ uniquely and for any i we may add or subtract a constant from $s_{i,t}$ for all t and remain feasible (as long as s remains in $\mathcal{H}$).

Observe that $\nu_{k,i,t}(\omega)$ can take values larger than 1, if season t in scenario ω corresponds to high inflows, a large part of which are retained as reservoir storage at the end of season t. The energy constraint (1) is accompanied by a generation capacity constraint

$$y_{k,i,b,t}(\omega) \leq \mu_{k,i,b,t}(\omega) z_{k,i},$$

where $\mu_{k,i,b,t}(\omega) = 1$, so (1) may not be binding if there are very large inflows and $s_{i,t}$ is at its capacity (meaning some inflows will be spilt).

In our model, reservoir storage decisions $s_{i,t}$ are chosen in a first stage (for the end of each season) to make (1) feasible for all ω. The operating policy will then be required to drive the storage through these points. This could be overly restrictive, for example, if we had large inflows in season t, in year ω, when it makes sense to choose $s_{i,t-1}$ to be lower just for this year. On the other hand allowing $s_{i,t-1}$ to freely anticipate future inflows removes the need for the variable entirely. We compromise by allowing $s_{i,t}$ to vary around a set point $\bar{s}_{i,t}$ by a limited amount $\hat{s}$. Thus (1) becomes

$$\sum_{b\in\mathcal{B}_t} H_b y_{k,i,b,t}(\omega) \leq \nu_{k,i,t}(\omega) \sum_{b\in\mathcal{B}_t} H_b z_{k,i} - s_{i,t}(\omega) + s_{i,t-1}\,(\omega),$$

and

$$s_{i,t}(\omega) \leq \bar{s}_{i,t} + \hat{s},$$

$$s_{i,t}(\omega) \geq \bar{s}_{i,t} - \hat{s}.$$

2.4. Battery storage

Battery storage is denoted by indices $k \in \mathcal{S} \subseteq \mathcal{K}$, where installed energy capacity in region i is denoted z_{ik} (MWh), and charging in region i and season t by power variables $g_{k,i,b,b',t}(\omega)$ where $b' \neq b$ denotes a load block in which the power charged in b will be discharged. Note that both b and b' are blocks in $\mathcal{B}_t$, where t is the time period (season). The term

$$\sum_{b' \neq b} g_{k,i,b,b',t}(\omega)$$

is the extra power needed in load block b that will charge the battery k for later discharge. Different battery types with the same storage capacity have different maximum charging rates (and costs). Given battery type $k \in \mathcal{S}_1$, denote by $\theta_k z_k$ the maximum rate of charge. Thus if z_k doubles (by installing twice as many batteries) then the charging rate doubles. This gives a charging constraint

$$\sum_{b' \neq b} g_{k,i,b,b',t}(\omega) \leq \theta_k z_{k,i}, \quad b \in \mathcal{B}_t,\ k \in \mathcal{S}_1,$$

where i is the region and t is the time period (season).

Let the number of days in each time period t be denoted D_t. We assume that each day in a season has the same number of hours in each load block, and the battery operates in the same way on each day. Then $\frac{H_b}{D_t}$ is the number of hours in load block b in any given day in time period t. This means that that total amount that the battery is charged in any block $b \in \mathcal{B}_t$ during a day cycle is

$$\frac{H_b}{D_t} \sum_{b' \neq b} g_{k,i,b,b',t}(\omega).$$

In the worst case, the battery will need to be charged in contiguous hours (e.g. overnight) to be discharged in later peak hours. This assumption means that the total amount that a battery will be charged in a day will involve adding the charge over all blocks b where $\sum_{b' \neq b} g_{k,i,b,b',t}(\omega) > 0$. The battery energy capacity z_{ik} (MWh) then restricts the choice of g, with the constraint

$$\sum_{b} \frac{H_b}{D_t} \sum_{b' \neq b} g_{k,i,b,b',t}(\omega) \leq z_{k,i}, \quad k \in \mathcal{S}.$$

The amount of energy stored in the battery for discharge in block b on a given day will be $\sum_{b' \neq b} \frac{H_{b'}}{D_t} g_{k,i,b',b,t}(\omega)$. Upon discharge this will yield the energy

$$\eta_k \sum_{b' \neq b} \frac{H_{b'}}{D_t} g_{k,i,b',b,t}(\omega)$$

where η_k is the roundtrip efficiency of the battery. This energy will be spread over the hours in block b on the given day. The power contribution (MW) in each of these hours is then

$$\frac{\eta_k \sum_{b' \neq b} \frac{H_{b'}}{D_t} g_{k,i,b',b,t}(\omega)}{\left(\frac{H_b}{D_t}\right)}$$

which simplifies to

$$\eta_k \sum_{b' \neq b} \frac{H_{b'}}{H_b} g_{k,i,b',b,t}(\omega).$$

2.5. Transmission constraints

The final part of the model are constraints on the electricity transmission. Transmission from region i to region j in load block b is denoted by $f_{i,j,b,t}(\omega)$, and the net flow arriving at region i from other regions is

$$\sum_j \Big((1 - \frac{\alpha_{j,i}}{2}) f_{j,i,b,t}(\omega) - (1 + \frac{\alpha_{i,j}}{2}) f_{i,j,b,t}(\omega) \Big) \tag{2}$$

where $\alpha_{i,j}$ is the proportional loss in energy incurred by transmission in the line between i and j. The transmission flows can be restricted by additional constraints $f \in \mathcal{F}$ that model Kirchhoff's voltage constraints from DC-load flow, for example.

The formula (2) gives a total net supply of power in load block b at region i defined by

$$\begin{aligned} q_{i,b,t}(\omega) = & \sum_{k \in \mathcal{K}_i} y_{k,i,b,t}(\omega) \\ & + \sum_j \left((1 - \tfrac{\alpha_{j,i}}{2}) f_{j,i,b,t}(\omega) - (1 + \tfrac{\alpha_{i,j}}{2}) f_{i,j,b,t}(\omega) \right) \\ & - \sum_{k \in \mathcal{K}_i} \sum_{b' \neq b} g_{k,i,b,b',t}(\omega) \\ & + \sum_{k \in \mathcal{K}_i} \eta_k \sum_{b' \neq b} \frac{g_{k,i,b',b,t}(\omega) H_{b'}}{H_b}, \end{aligned}$$

where $\mathcal{K}_i \subseteq \mathcal{K}$ are the technologies available at region i. This constraint assumes that all destination blocks b' occur in the same day as b. This won't necessarily be the case. Some load blocks

(representing annual system peaks) have only a few hours in them. One could restrict b' to load blocks that are not too far from b, so shifted load can be assumed to come from the same block. Note that other data such as K_k, L_k, C_k, μ_k, ν_k and D_t can be extended to be location dependent in a straightforward way.

2.6. Demand response

One mechanism for reducing load in peak times is to shift load. With appropriate incentives, consumers can be assumed to not purchase electricity in a peak period for a given activity by deferring it to a period with lower total demand. Typically these shifts of demand are within a day, where a consumer chooses to do their laundry for example in off-peak periods where prices are lower. Our model of batteries could be used to represent this form of demand shifting.

A different issue arises when the system faces an energy constraint. In the New Zealand case study in Section 4, this corresponds to a dry winter in which hydro reservoir inflows are low. To deal with such an uncertain energy shortage, it is not enough to shift load out of peak periods, electricity must substituted or foregone. With strongly interconnected systems, substitutes for local electricity can come from imported power. With isolated systems like New Zealand the substitution must come from industrial users of electricity who reduce production in New Zealand and increase it in other countries where energy supplies are (temporarily) more plentiful.

We model industrial load reduction over a protracted period in each location i using variables $y_{k,i,b,t}(\omega)$, where $k \in \mathcal{I}_i$ indicates a particular type of industrial entity in location i that can shut down to save energy. This incurs an operating loss for the industry of $C_k y_{k,i,b,t}(\omega)$. The maximum amount that load can be reduced at this cost is $U_{k,i}$. The model is flexible enough to provide more of this option (possibly) with fixed costs $K_k(x_{k,i}) + L_k(z_{k,i})$, and increasing operating losses with tranches of increasing marginal cost C_k.

This modeling feature is not intended to capture peak shaving (which is accomplished by the battery model). We therefore add constraints to preclude shutdowns cycling over a short period (e.g. between peak and off-peak periods within a day). These constraints take the form

$$y_{k,i,b,t}(\omega) = \bar{y}_{k,i,t}(\omega), \quad b \in \mathcal{B}(t),$$

$$\bar{y}_{k,i,t}(\omega) \leq z_{k,i}.$$

This means that in a feasible solution, every load block in a season t will have the same load reduction. If total shutdown of a plant is required then the second constraint could be modeled by variables $\delta \in \{0,1\}$ with

$$\bar{y}_{k,i,t}(\omega) \leq U_{k,i}\delta.$$

Alternatively we can interpret a fractional value of δ as a shut over a fraction of a season which reduces load in all load blocks uniformly.

The constraints presented here assume that the industrial load can anticipate the uncertain outcome ω. With no anticipation we would obtain

$$y_{k,i,b,t}(\omega) = \bar{y}_{k,i,t}, \quad \omega \in \Omega,\, b \in \mathcal{B}(t),$$
$$\bar{y}_{k,i,t}(\omega) \leq z_{k,i},$$

which is arguably too restrictive as it would require plants to plan to shut in advance, and go through with this even if electricity supply turned out to be plentiful. A practical compromise selects uncertain outcomes that can reasonably be anticipated when they start to take effect (e.g a dry winter) and relaxes the nonanticipativity constraints over these dimensions only. Thus if $\Omega = \Omega_1 \times \Omega_2$, where $\omega_1 \in \Omega_1$ denotes a year climate outcome, and $\omega_2 \in \Omega_2$ denotes other random outcomes that cannot be anticipated then we obtain

$$y_{k,i,b,t}(\omega_1,\omega_2) = \bar{y}_{k,i,t}(\omega_1), \quad (\omega_1,\omega_2) \in \Omega, \quad b \in \mathcal{B}(t),$$
$$\bar{y}_{k,i,t}(\omega_1) \leq z_{k,i}.$$

2.7. Stochastic planning model

Since in the most sophisticated version of our model, we will be endowing each agent with a risk measure, we will introduce coherent risk measures directly into our model. For each agent a, risk is modeled by a mapping ρ from random outcomes to real numbers. The risk measure is assumed to be *coherent*, in the sense that it satisfies the axioms of Artzner et al. (1999).

Some care is needed in writing down the cost of the social planner summed over seasons $[0, T]$. In each season t they observe a random outcome $\omega_t = (\omega_1, \omega_{2t})$ that defines what sort of year it is (ω_1) and a seasonal random outcome (ω_{2t}) that is independent of previous seasons (e.g. the model for random wind events described above). The operational decisions $y_{k,i,b,t}(\omega_1, \omega_{2t})$ in this model are assumed to be independent of ω_{2t} in previous seasons, but depend on ω_1 that is the same across all seasons. Thus in each season t there are $\mid \Omega_1 \mid\mid \Omega_{2t} \mid$ (possibly) different decisions for $y_{k,i,b,t}(\omega)$. This means that in each season t there are $\mid \Omega_1 \mid\mid \Omega_{2t} \mid$ possibly different outcomes for the system cost $Z_{i,t}(\omega_1, \omega_{2t})$. Like $y_{k,i,b,t}$ in different seasons t_1 and t_2, the cost outcomes $Z_{i,t_1}(\omega_1, \omega_{2t_1})$ and $Z_{i,t_2}(\omega_1, \omega_{2t_2})$ are independent. However in evaluating the total annual cost using a coherent risk measure, in general we need to compute the random variable $\sum_{t\in[0,T]} Z_{i,t}(\omega)$, which entails sums for all possible sequences $\omega_{2t}, t \in [0, T]$. Observe that if the risk measure is expectation then

$$\mathbb{E}[\sum_{t\in[0,T]} Z_{i,t}(\omega)] = \sum_{t\in[0,T]} \mathbb{E}[Z_{i,t}(\omega)]$$

and so the total capital and operating cost $\psi(\omega)$ can be expressed more simply as a sum over seasons of $\mathbb{E}[Z_{i,t}(\omega_1, \omega_{2t}]$. This is not possible for general risk measures.

The stochastic social planning model is as follows. We seek a solution that minimizes risk-adjusted capital and operating costs.

$$
\begin{aligned}
\text{P: min} \quad & \rho(\psi) \\
\text{s.t.} \quad & \psi(\omega) = \textstyle\sum_i \left(\sum_{k \in \mathcal{K}_i} \left(K_k(x_{k,i}) + L_k(z_{k,i}) \right) + \sum_{t \in [0,T]} Z_{i,t}(\omega) \right) \\
& Z_{i,t}(\omega) = \textstyle\sum_{b \in \mathcal{B}(t)} H_b \left(\sum_{k \in \mathcal{K}_i} C_k y_{k,i,b,t}(\omega) - V \left[d_{i,b,t}(\omega) - r_{i,b,t}(\omega) \right] \right), \\
& x_{k,i} \leq u_{k,i}, \\
& z_{k,i} \leq x_{ki} + U_{k,i}, \\
& y_{k,i,b,t}(\omega) \leq z_{k,i}, \\
& y_{k,i,b,t}(\omega) \leq \mu_{k,i,b,t}(\omega) z_{k,i}, \\
& \textstyle\sum_{b \in \mathcal{B}_t} H_b y_{k,i,b,t}(\omega) \leq \nu_{k,i,t}(\omega) \sum_{b \in \mathcal{B}_t} H_b z_{k,i} - s_{i,t} + s_{i,t-1}, \\
& r_{i,b,t}(\omega) \leq d_{i,b,t}(\omega), \\
& q_{i,b,t}(\omega) = \textstyle\sum_{k \in \mathcal{K}_i} y_{k,i,b,t}(\omega) + \sum_j \left((1 - \frac{\alpha_{j,i}}{2}) f_{j,i,b,t}(\omega) - (1 + \frac{\alpha_{i,j}}{2}) f_{i,j,b,t}(\omega) \right) \\
& \qquad\qquad - \textstyle\sum_{k \in \mathcal{K}_i} \sum_{b' \neq b} g_{k,i,b,b',t}(\omega) + \sum_{k \in \mathcal{K}_i} \eta_k \sum_{b' \neq b} \frac{g_{k,i,b',b,t}(\omega) H_{b'}}{H_b}, \\
& d_{i,b,t}(\omega) \leq q_{i,b,t}(\omega) + r_{i,b,t}(\omega), \\
& \textstyle\sum_b H_b \sum_{b' \neq b} g_{k,i,b,b',t}(\omega) \leq z_{ki} D_t, \\
& \textstyle\sum_{b' \neq b} g_{k,i,b,b',t}(\omega) \leq \theta_{k,i} z_{k,i}, \\
& f \in \mathcal{F}, \\
& s \in \mathcal{H}.
\end{aligned}
$$

The first-stage decisions in this model are the capacity decisions x and z, and the decisions $s_{i,t}$ for each time interval t that determine how much energy will be transferred by storage at i from period t to period $(t+1)$ mod T. The risk measure is denoted here by ρ and is a system-wide functional accounting for all the risks modeled using ω.

As described above, using different choices of the data μ and ν, we can generate a set of models that will add uncertainty in wind and run-of-river hydro generation. The capital plans that result will be different. Further changes in data lead to another class of models that add uncertainty in stored hydro generation, but still seek a solution that minimizes expected capital and operating costs. The capital plans that result will be different again.

3. Constraints on renewables

All of the models of the previous section can be used to study the effect of adding a constraint on non-renewable generation. Four forms of this constraint will be considered. The first three of these non-renewable constraints hold in expectation or in every outcome ω, while the fourth is a chance-constraint formulation.

3.1. Capacity constraint

This is a first stage regulation that limits total non-renewable (coal, gas and diesel) capacity:

$$\sum_i \sum_{k \in \mathcal{N}_i} z_{k,i} \leq E, \tag{3}$$

where $\mathcal{N}_i$ represents the nonrenewable sources at location i. It is independent of scenarios ω.

Instead of imposing the constraint explicitly in the optimization problem, we can introduce a Lagrange multiplier σ for the constraint and move this into the objective $\psi(\omega)$:

$$\psi(\omega) = \sum_i \left(\sum_{k \in \mathcal{K}_i} (K_k(x_{k,i}) + L_k(z_{k,i})) + \sum_{k \in \mathcal{N}_i} \sigma z_{k,i} + \sum_{t \in [0,T]} Z_{i,t}(\omega) \right) - \sigma E.$$

The resulting effect of the Lagrangian form of the constraint is to simply replace the $L_k(z_{k,i})$ term by $L_k(z_{k,i}) + \sigma z_{k,i}$ for each $k \in \mathcal{N}_i$. Thus the capacity constraint essentially amounts to an increase in maintenance cost in the non-renewable technologies. Clearly σ and E are intimately related and we can exogenously parameterize the optimization using either σ or E.

3.2. Generation constraint

The second renewable constraint is a second stage regulation that limits expected non-renewable generation:

$$\mathbb{E}[\sum_{b \in \mathcal{B}_t} H_b \sum_i \sum_{k \in \mathcal{N}_i} y_{k,i,b,t}(\omega)] \leq E, \quad \forall t. \tag{4}$$

A more restrictive *almost sure* constraint is to impose the regulation in each scenario ω:

$$\sum_{b \in \mathcal{B}_t} H_b \sum_i \sum_{k \in \mathcal{N}_i} y_{k,i,b,t}(\omega) \leq E, \quad \forall t, \omega.$$

Introducing Lagrange multipliers for either of these constraints leads to an adjusted cost optimization parameterized by those multipliers. In the almost sure form of the constraint, for example, the modification to $Z_{i,t}$ is as follows:

$$Z_{i,t}(\omega) = \sum_{b \in \mathcal{B}(t)} H_b \left(\sum_{k \in \mathcal{K}_i} C_k y_{k,i,b,t}(\omega) + \sum_{k \in \mathcal{N}_i} \sigma_t(\omega) y_{k,i,b,t}(\omega) - V\left[d_{i,b,t}(\omega) - r_{i,b,t}(\omega) \right] \right) - \sigma_t(\omega) E,$$

amounting essentially to a scenario and season increase in variable operating cost for non-renewables.

3.3. Emission constraint

The third constraint is a limit E on total emissions from all generation that emits CO_2 (including geothermal and CCS generation). Again, this can be expressed in expectation form:

$$\mathbb{E}[\sum_{b \in \mathcal{B}_t} H_b \sum_i \sum_{k \in \mathcal{N}_i} e_k y_{k,i,b,t}(\omega)] \leq E, \quad \forall t, \tag{5}$$

or in every scenario ω:

$$\sum_{b \in \mathcal{B}_t} H_b \sum_i \sum_{k \in \mathcal{N}_i} e_k y_{k,i,b,t}(\omega) \leq E, \quad \forall t, \omega.$$

Here e_k denotes an emissions factor for output from plant of type k. Although the data H, e and E could vary by location i and/or season t with minor changes to this constraint, we assume in all our experiments that this is not the case.

The constraint E on emissions from electricity generation can be thought of as a regulatory intervention. In practice these emissions will be traded off against emissions in the rest of the economy (or the world if carbon credits are traded internationally). The shadow price σ thus represents a carbon price, which could be imposed on the electricity sector, or emerge from a general equilibrium model with many sectors apart from electricity. It can be implemented via a modification to $Z_{i,t}$ similar to that given above:

$$Z_{i,t}(\omega) = \sum_{b \in \mathcal{B}(t)} H_b \left(\sum_{k \in \mathcal{K}_i} C_k y_{k,i,b,t}(\omega) + \sum_{k \in \mathcal{N}_i} \sigma_t(\omega) e_k y_{k,i,b,t}(\omega) - V\left[d_{i,b,t}(\omega) - r_{i,b,t}(\omega) \right] \right) - \sigma_t(\omega) E.$$

We discuss later how this expression should be updated if the constraint was in expectation, or if the carbon tax does not vary by season t.

3.4. Chance-constraint on emissions

We denote the tonnes of CO_2 emissions in scenario ω by $J(\omega)$. We can impose a chance constraint on $J(\omega)$ of the form

$$\mathbb{P}(J(\omega) > 0) \leq \beta.$$

Thus if we were to choose $\beta = 0.5$, and ω denotes potential inflow scenarios (possibly sampled from history), then this constraint would restrict annual emissions to zero in at least 50% of these scenarios. We model this as a mixed integer program using a "big-M" constraint by simply adding the following constraints to the stochastic planning model above:

$$\begin{aligned} J(\omega) &= \textstyle\sum_{t\in[0,T]} \sum_{b\in\mathcal{B}(t)} H_b \sum_i \sum_{k\in\mathcal{N}_i} e_k y_{k,i,b,t}(\omega) \\ J(\omega) &\leq M\delta(\omega) \\ \textstyle\sum_\omega \mathbb{P}(\omega)\delta(\omega) &\leq \beta, \\ \delta(\omega) &\in \{0,1\}. \end{aligned}$$

In our experiments we used $\omega \in \Omega_1$ instead of the full generality of $\omega \in \Omega$. We do not derive a Lagrangian form of this problem since the primal problem now involves binary variables.

4. New Zealand Case Study

The New Zealand Labour Party and Green Party of Aotearoa New Zealand Confidence and Supply Agreement of 2017 (as reproduced in New Zealand ICCC Terms of Reference 2019) states that the "Government will: Request the Climate Commission to plan the transition to 100% renewable electricity by 2035 (which includes geothermal) in a normal hydrological year" The social planning model P has been implemented using New Zealand data as part of a project seeking to understand mechanisms by which this goal can be achieved. The data set for the experiment is provided in the online companion to the paper. Selected parameter values are displayed in tables in this section. We also refer the reader to the tables (5-12) in Appendix A to the paper. We give a brief summary here of how the parameters of the model were estimated.

The model has three regions (i =SI, HAY, NI) representing the South Island, lower North Island and Upper North Island respectively. These three regions are joined by nominal transmission lines

having capacity 1200MW for SI-HAY and 1000 MW for HAY-NI. We model four seasons ($t = 0, 1, 2, 3$) representing the calendar months January-March, April-June, July-September, October-November.

Demand data for the model at the three locations are estimated from half-hourly metered load for every grid exit point in the national transmission system. These data are archived at Electricity Market Information System (2019). The data are adjusted for wind and photovoltaic generation that has not been recorded, and aggregated into regional demand for each half-hour period in 2017. Wind generation data from each region for years 2005-2017 is similarly aggregated. The total metered demand over the three regions is then sorted from maximum to minimum to give a load-duration curve, and the wind generation data in each year is sorted to give the same order of trading periods as the load. This enables us to identify peak periods in which wind generation is absent and estimate the probability of no wind in a peak load block. From the demand and wind data we create 10 load blocks for each season. The number of hours in each load block is shown in Table 5.

For each load block we estimate a level of demand for 2035 by applying a load growth factor to the residential and commercial components that are expected to increase with population and economic growth. (A large aluminium smelter is excluded from these growth estimates.) The level of demand in each block in 2035 is then increased by projected increases in load from electric vehicles and industrial electrification. The result is a set of tables of projected load (MW) in each block in each region in each season in 2035 with load blocks shown in Table 6. The parameters $\mu_{k,i,b,t}$ where k represents photovoltaic solar generation are given in Table 7; the parameters $\mu_{k,i,b,t}$ where k represents wind generation are given in Table 8; the parameters $\mu_{k,i,b,t}$ where k represents run-of-river generation are given in Table 9; and the parameters $\nu_{k,i,t}$ where k represents hydro reservoir generation are given in Table 11.

The existing capacities of generation technologies in each region were sourced from the Electricity Authority generation database available at Electricity Market Information System (2019). For 2035

Existing Cap.	SI	HAY	NI
CCGT	0.0	0.0	403.0
CCS	0.0	0.0	0.0
DIESEL	0.0	0.0	155.0
DR	50.0	0.0	0.0
GEOT	0.0	0.0	892.7
HYDROr	840.0	0.0	687.0
HYDROs	2573.0	0.0	1051.0
OCGT	0.0	0.0	350.8
SLOWBATT	0.0	0.0	0.0
MEDBATT	0.0	0.0	0.0
FASTBATT	0.0	0.0	0.0
SOLAR	0.0	0.0	0.0
WIND	0.0	143.0	232.2

Pot. Increase	SI	HAY	NI
CCGT	0.0	0.0	2000.0
CCS	0.0	0.0	2000.0
DIESEL	0.0	0.0	0.0
DR	0.0	0.0	0.0
GEOT	0.0	0.0	542.0
HYDROr	130.5	0.0	0.0
HYDROs	0.0	0.0	0.0
OCGT	0.0	0.0	920.0
SLOWBATT	500.0	500.0	500.0
MEDBATT	500.0	500.0	500.0
FASTBATT	500.0	500.0	500.0
SOLAR	1000.0	1000.0	1000.0
WIND	5000.0	5000.0	5000.0

Table 1 Capacity of existing plant (MW) that will be available in 2035 and potential electricity capacity increases (MW) by technology and region.

we assumed that the coal/gas fired Rankine units at Huntly would be decommissioned as would the Stratford combined cycle plant. This gives a mix of existing capacities as shown in Table 1. Each run of the model proposes limits on new capacity to build of each technology. In our example runs we have chosen possible capacity additions as also shown in Table 1.

New capacity incurs an annualized fixed capital cost, and new and existing capacity incurs an annual operations and maintenance cost. The annualized capital cost is a capital recovery factor per MW of capacity that gives the before tax annual revenue that would be required to give an internal rate of return of 8% (after depreciation and tax) on the capital. Estimates of capital costs for thermal plant have been obtained from US Energy Information Administration (EIA) (2019),

and costs for CCGT with Carbon Capture and Storage (CCS) were sourced from Rubin and Zhai (2012).

The values of K_k, C_k, and L_k assumed in our models are shown in Table 2. All costs are measured in 2018 New Zealand dollars.

Costs	K_k (NZD/MW/Yr)	C_k (NZD/MWh)	L_k (NZD/MW/Yr)
CCGT	138000	70	45000
CCS	242717	75	45000
DIESEL	110400	232	15000
DR	0	1000	0
GEOT	430000	1	150000
HYDROr	430000	6	0
HYDROs	516000	6	0
OCGT	110400	93	15000
SLOWBATT	48364	0	5000
MEDBATT	163879	0	6000
FASTBATT	314788	0	7000
SOLAR	110400	2	35000
WIND	178000	12	20000

Table 2 Capital costs, variable costs and maintenance costs

5. Computational results

This section of the paper describes the results of some computational experiments with the social planning model under a number of assumptions. Experiment 1 studies the effect of thermal capacity reduction on average CO_2 emissions. Experiment 2 uses "business-as usual" forecasts of 2035 electricity demand and studies the effect on total investment and expected operating cost of tightening the constraints either on non-renewable generation or on average CO_2 emissions. The experiment also compares the output from a risk-neutral model with results from one that optimizes a risk-averse coherent risk measure. Experiment 3 assumes more electricity demand in 2035 (due to electric vehicle growth and industrial electrification) and studies the effect on expected system cost of tightening constraints either on non-renewable generation or on average CO_2 emissions. Experiment 4 compares different probabilistic versions of the constraint on CO_2 emissions.

5.1. Experiment 1

The first experiment we carried out tested a conjecture that decreasing the capacity of thermal plant could increase CO_2 emissions. This effect was first noticed in a more detailed model developed by Fulton (2018), solved using a version of the SDDP algorithm. In our model we relaxed the CO_2 reduction constraints and amended Table 1 so that each region had 1200MW of CCGT, no OCGT or DIESEL, and no changes in generation capacity were allowed. The value of $\hat{s}$ was set to 0 for simplicity. This mix of generation gave annual average CO_2 emissions of 4409 kt. We then reduced the CCGT capacity to 700MW in each region, while all other capacities remained unchanged. The model then gave annual average CO_2 emissions of 4428 kt.

The average energy generated by geothermal plant is the same in each run, so the difference in average emissions results from CCGT generation. The average energy generated by CCGT plant in each season in these two cases is shown in Table 3. One can see that the CCGT generation in the 700MW case is higher overall, and is higher on average in the first half of the year.

The optimal values of the first-stage storage levels $\bar{s}_{i,t}$ are not unique. It is easy to see that one can add or subtract a constant from all of them and remain feasible as long as the values remain

GWh	0	1	2	3	Total
700 MW	2675	2834	2018	2375	9903
1200 MW	2410	2636	2407	2399	9853

Table 3 Average CCGT generation in season t

between their bounds. To make a comparison, we have normalized the results so that both runs have the same storage at the end of period 2. The resulting figures are shown in Table 4a (for 1200 MW) and Table 4b (for 700 MW).

GWh	0	1	2	3
SI	1809	1526	0	943
NI	327	0	374	741
Total	2136	1526	374	1683

(a) CCGT capacity 1200MW

GWh	0	1	2	3
SI	1736	1526	0	957
NI	593	390	374	691
Total	2328	1916	374	1648

(b) CCGT capacity 700MW

Table 4 Reservoir storage at the end of season t

Observe that with lower thermal capacity (Table 4b) the reservoir levels at the end of period 3 are about the same but the model gives a higher average reservoir volume 2328 GWh at the end of period 0 than in the higher thermal capacity case (Table 4a with 2136 GWh). This increase in reservoir levels requires additional thermal generation to attain the higher storage level, which is put in place as a hedge against the chance of a dry winter. The additional storage hedge is not needed when 1200MW of CCGT is available to be used in case of low inflows.

Clearly this conclusion requires the use of the stochastic model to provide value for hedging, and using the water levels as a first stage variable enables the planning model to capture the observed effect. More generation (of the lower capacity units) leads to higher emissions. This effect, which can be observed over a wide range of data values, is examined in more detail in Fulton (2018).

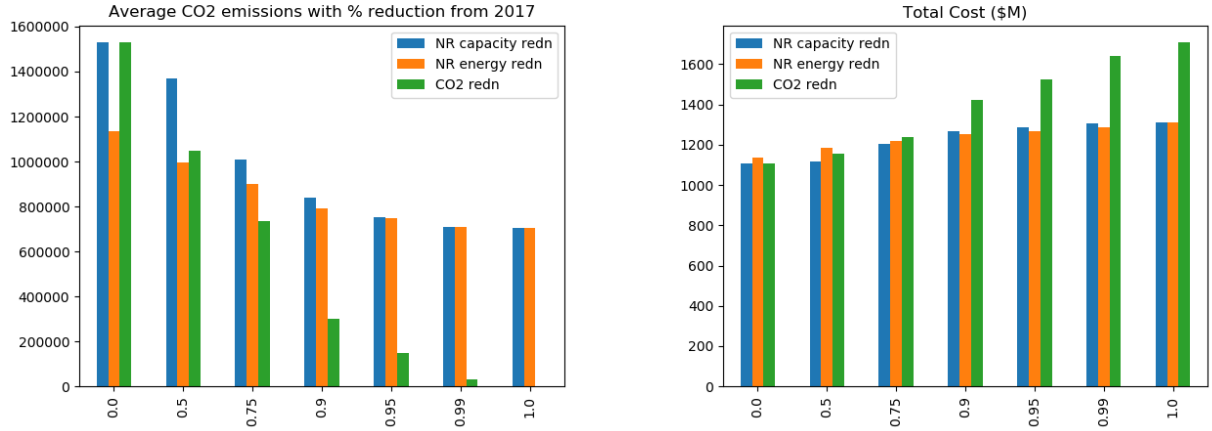
5.2. Experiment 2

In the second experiment, we investigated the effect of increasing the level of the renewable constraints on the model. The alternative formulations of these constraints ((3), (4), and (5)) impose bounds E on non-renewable capacity, non-renewable energy and average CO_2 emissions respectively. The constraints are made progressively more restrictive by a parameter θ that increases from 0 to 1. The right-hand side E of each constraint ((3), (4), and (5)) is replaced by $(1-\theta)\bar{E}$, where $\bar{E}$ denotes the 2017 level of the appropriate quantity (including pre-determined closures in the case of (3).

It is important to be clear what θ measures, especially when the constraint (5) is imposed. In this case the right-hand side E of constraint (5) is replaced by $(1-\theta)\bar{E}$, where we $\bar{E}$ equals the 2017 level of emissions from electricity generation (three million tonnes). Here a value of $\theta = 0.5$ does not mean a 50% renewable electricity system, but a system that emits on average 50% of the electricity CO_2 emissions of 2017 (which comes from about 15% of generated electricity in 2017). So, in terms of average emission levels, a given value of θ amounts to an $(85+15\theta)$-%-renewable electricity system.

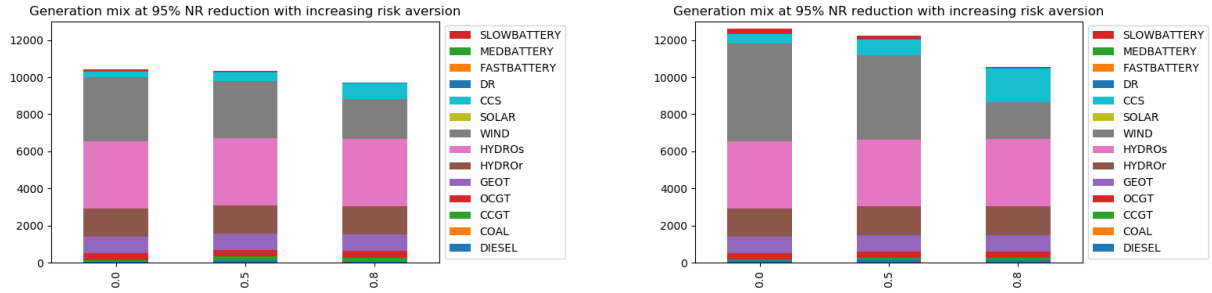
In Figures 2a and 2b, θ is shown on the horizontal axis, and capacity constraint (3) from Section 3.1 is depicted by the "NR capacity redn" bars, the generation constraint (4) from Section 3.2 is depicted by the "NR energy redn" bars, and the emission constraint (5) from Section 3.3 is depicted by the "CO2 redn" bars. The y-axis for Figure 2a is kt of carbon and in Figure 2b it depicts the annual cost in (2018)$B NZ. Since (renewable) geothermal and CCS emit some CO_2 (so renewable is not the same as no carbon emission), Figure 2a shows that a value of $\theta = 1$ yields modest reductions in actual CO_2 emissions if we impose only the capacity or energy reduction constraints (3) or (4). When θ represents percentage reductions in actual CO_2 emissions (including those from geothermal and CCS) the green bars in Figure 2b show that the cost increases are fairly modest (27%) up to $\theta = 0.95$, but are around 45% for zero carbon emissions ($\theta = 1$).

We finish this example by giving the results from optimizing a risk-averse system risk measure, namely $(1-\lambda)\mathbb{E}[Z] + \lambda \mathrm{AVaR}_{0.90}(Z)$. The example constrains the amount of non-renewable energy



(a) Carbon emissions (b) Costs

Figure 2 Increasing θ on constraints



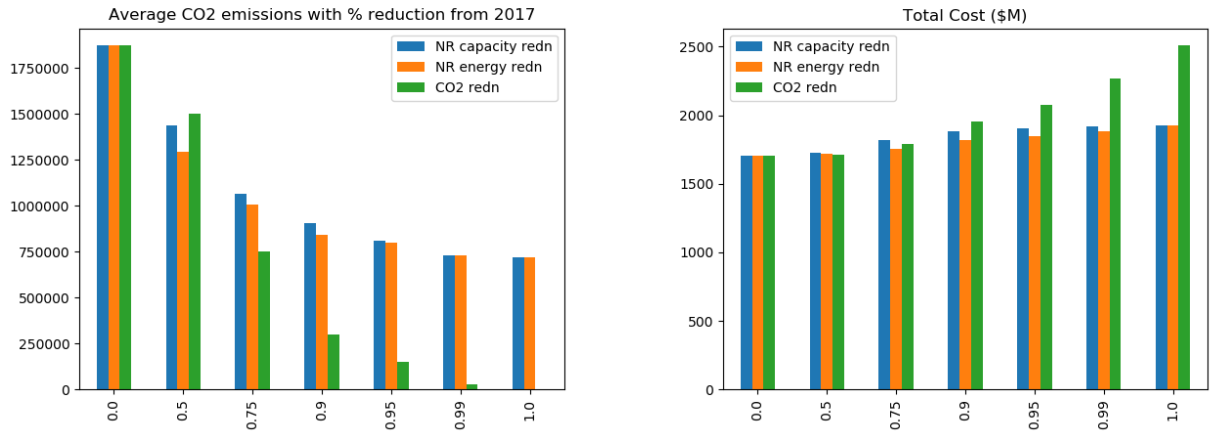
(a) Original load (b) Increased load

Figure 3 Risk aversion modelled using $(1-\lambda)\mathbb{E}[Z]+\lambda\text{AVaR}_{0.90}(Z)$, for $\lambda=0, 0.5, 0.8$.

generated to be 5% of 2017 levels, i.e. $\theta=0.95$. The optimal capacity mixes with three choices of λ are given in Figure 3. The amount of wind capacity installed decreases as risk aversion increases, and since this is replaced by (dispatchable) CCGT plant, much of which has CCS, the total amount of capacity needed drops.

5.3. Experiment 3

The next experiment we carried out compared the expected cost of meeting targets on nonrenewable capacity as compared with meeting targets on nonrenewable energy, but with increased forecast electricity load (see Table 12 in Appendix A) arising from electric vehicles and conversion of industrial process heat from gas and coal to electricity.



(a) Carbon emissions (b) Costs

Figure 4 Increasing θ on constraints (increased load)

The bars denoted "CO2 redn" in Figure 4 show that the cost of actually reaching zero CO_2 emissions (without geothermal or CCS) increases substantially as we approach the limit. For completeness, in the emission constraint case shown in Figure 5, we split the costs out into four bars depicting the investment cost, the cost of maintenance, the expected operating cost, and the expected cost of lost load.

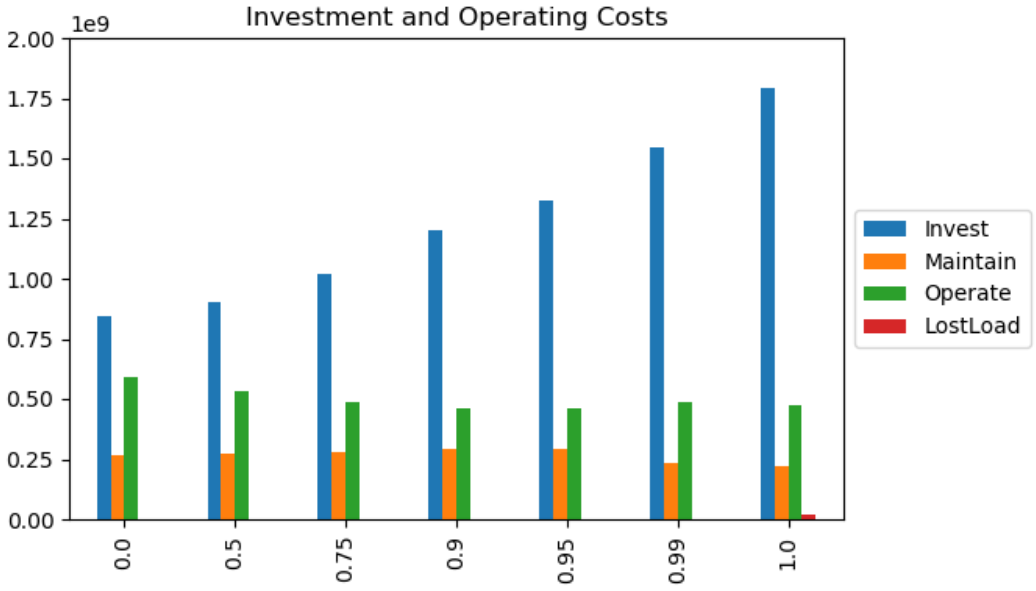
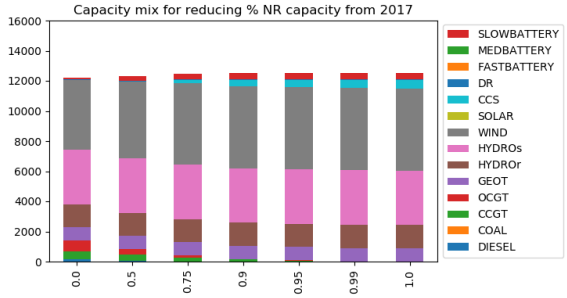
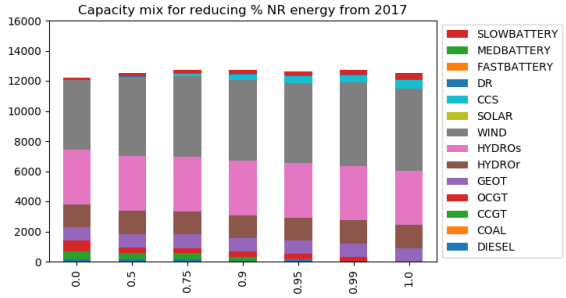


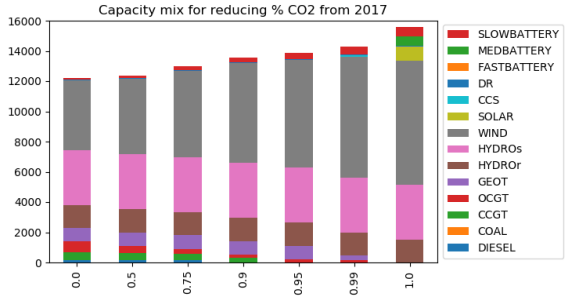
Figure 5 Increasing θ on constraints (increased load)



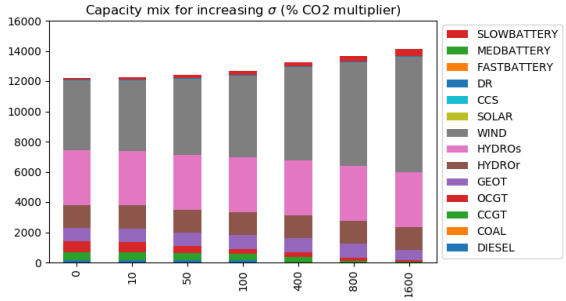
(a) Capacity reduction



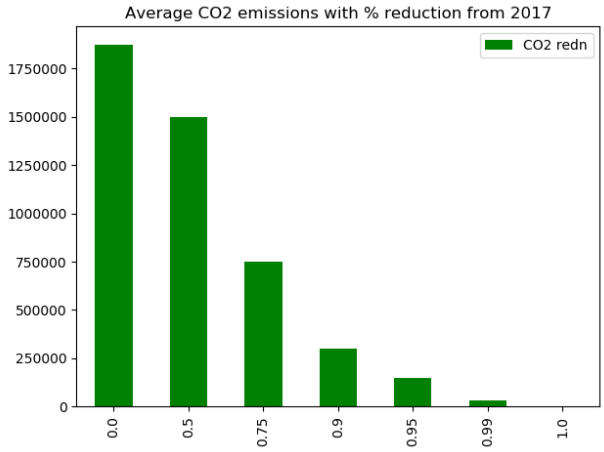
(b) Energy reduction



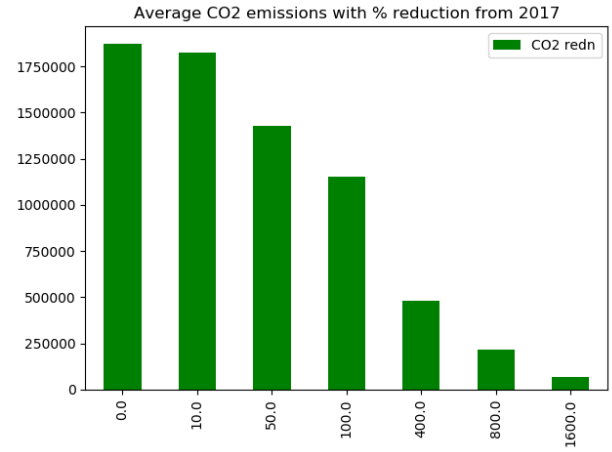
(c) Emission reduction



(d) Carbon tax



(e) Emission reduction



(f) Carbon tax

Figure 6 Mix of technologies (increased load)

The mix of technologies used in the solutions (for the increased load case) is shown in Figure 6. The capacity and energy reduction mixes are almost indistinguishable, although the energy reduction capacity choice does retain more non-renewable capacity (Figures 6a and 6b) at high levels of non-renewable energy reduction. The larger plants are kept open and used sparingly to

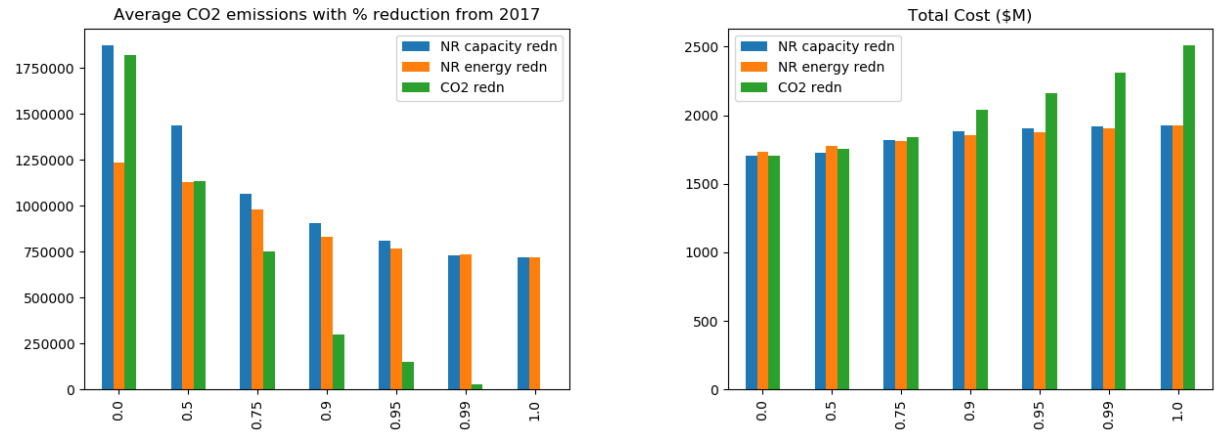
provide peaking support in dry years. The mix for emission reduction is more diverse as the level of emission becomes small. Geothermal is removed (since it is a CO_2 emitter) and replaced by large amounts of wind coupled with some solar and battery capacity. We have modeled three different forms of battery. It is interesting to note that CCS comes into the mix when $\theta = 0.99$, but cannot be present at the 100% level.

As mentioned earlier in the paper, we are able to replace the parametric constraint on emissions by a carbon tax. The figures 6c and 6d could be made identical by suitable choices for the carbon tax, but we give a representative set of values only in Figure 6d. Note that a value of the carbon tax of around \$400 per tonne leads to similar capacity mix as the $\theta = 0.90$ system. Note however that Figures 6e and 6f show that a carbon tax of \$400 per tonne will not yield the depth of CO_2 reductions attained by the $\theta = 0.90$ system, and a tax of closer to \$800 per tonne is needed to ensure that the non-renewable plant that is installed is run less often so as to attain the desired reduction in CO_2 to meet the $\theta = 0.90$ target.

5.4. Experiment 4

In this experiment, we investigate the effect of using different probabilistic versions of the constraint on non-renewable energy.

5.4.1. Almost sure constraint Figure 7 shows the results for meeting the constraint in the almost-sure sense. Since capacity expansion is a first-stage decision, the "NR capacity redn" outcomes are the same as those shown in Figure 4 for constraining expected emissions. Significant differences between Figure 4 and Figure 7 for the "NR energy redn" and "CO2 redn" outcomes are observed only at relatively low levels of CO_2 reduction. The "NR energy redn" bar reduces height significantly for the 0.0 average reduction case. In Figure 4 this bar measures average non-renewable energy production when this is constrained to be below 2017 levels, whereas the orange bar in Figure 7 measures average non-renewable energy generation when this generation is constrained to be below 2017 levels *in every scenario*. The latter is more restrictive and will give a lower average.



(a) Carbon emissions (b) Costs

Figure 7 Increasing θ on constraints, almost sure case (increased load)

In fact, there is a single year, 2005, in which the emissions are significantly higher than all the others in the average case, but is compensated for by reduced emissions in other years. Similar (albeit less dramatic) differences are seen for the 0.5 average reduction case, and then these differences disappear as the emissions constraints become stricter, and non-renewable energy reductions in every scenario become necessary. When the average reduction factor is 1.0, the constraints are identical and the outcomes shown in Figure 4 and Figure 7 are the same.

5.4.2. Chance constraint The next results compute the cost of meeting a chance constraint (with original and increased load data). The chance constraint requires that the system has zero CO_2 emissions in 7 out of 13 inflow scenarios. The optimal capacity mix for the original load data is shown in Figure 8a and the optimal capacity mix for the increased load data is shown in Figure 8b. The mix of capacities in Figure 8b is commensurate with the mix shown in Figure 6c at a renewable level of 100%. The expected cost of the chance-constrained solution with original load is $ 1.58 B NZ, while the cost for the chance-constrained solution with increased load is $ 2.36 B NZ. One can compare these with the costs for "CO2 redn" in Figure 2b and Figure 4b. It is tempting to suppose that the chance-constrained solution delivers nearly 100% reduction in CO_2 emissions at about the same cost as the 100% figures shown in Figure 2b and Figure 4b. However Figure 8c

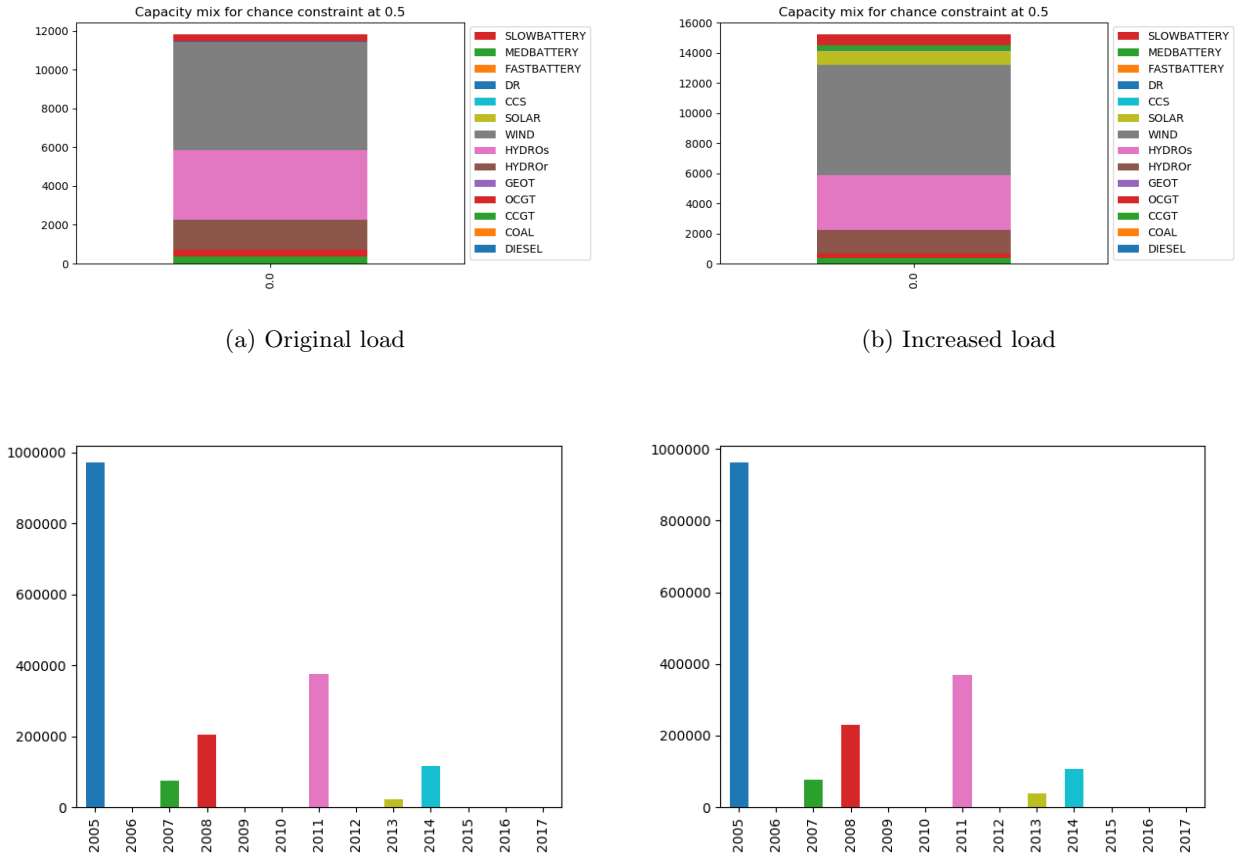
and Figure 8d show the realized reductions in CO_2 in each scenario from the chance-constrained capacity choices shown in Figure 8a and Figure 8b. As expected, there are nonzero CO_2 emissions in 6 out of the 13 scenarios, and although the system is 100% renewable in the other 7 scenarios, the average level of CO_2 emissions in both cases amounts to about 0.138 Mt, or approximately 95.4% reduction in CO_2 emissions from 3 million tonnes in 2017.

It is also intriguing to see that the average level of CO_2 emissions (about 0.138 Mt) is the same for the original load and the increased load. The CCGT and OCGT plant capacity choices are approximately the same in both cases. There is no constraint on using these technologies in 6 out of 13 scenarios, which will correpond to years with low hydro inflows. So these technologies will be frequently used to their capacities in these years (and not at all in other years), giving similar levels of emissions.

6. Competitive models of storage

The last class of models consider a competitive setting in which agents maximize profit as price takers, where electricity prices are determined by a market clearing agent. In theory these models should yield the same solution as the second class of models, and this is verified numerically. Endowing agents with coherent risk measures gives a different result. We examine the competitive equilibrium in the case when markets for risk are incomplete, and when they are complete.

We simplify the exposition in this section in a number of ways, all of which are for clarity and ease - the additional constructs from above can be added with only an increase in the complexity of the notation. We capture all the constraints except the emission and balance of supply and demand constraints into a single set $\mathcal{X}$. The forms of the objective are simplified into functions of the first-stage and second stage decision variables. The resulting form of the stochastic planning model (P) from Section 2.7, combined with the renewable emission constraint from Section 3.3 is



(a) Original load

(b) Increased load

(c) Original load, emissions in each scenario

(d) Increased load, emissions in each scenario

Figure 8 Chance-constraint on zero emissions in 50% of years

then:

$$\begin{aligned} \text{P:}\ \min\ \rho(\psi)\ \ \text{s.t.}\ \ & (x,y,z,q,r) \in \mathcal{X} \\ \psi(\omega) \ &= \ \mathcal{C}(x,z) + \mathcal{Z}(y,q,r,\omega) \\ 0 \ &\leq \ q(\omega) + r(\omega) - d(\omega) \\ 0 \ &\leq \ E - \mathbb{E}[\mathcal{E}(y,\omega)]. \end{aligned}$$

Competition arises in this model from the fact that generation and demand may have distributed ownership. Introducing an additional index a for such entities, the stochastic planning model then

becomes:

$$\begin{aligned} \text{P: } \min\ \rho(\psi) \ \text{ s.t. } \ & (x_a, y_a, z_a, q_a, r_a) \in \mathcal{X}_a, \forall a \\ \psi(\omega) \ &= \ \textstyle\sum_a \left(\mathcal{C}_a(x_a, z_a) + \mathcal{Z}_a(y_a, q_a, r_a, \omega)\right) \\ 0 \ &\leq \ \textstyle\sum_a \left(q_a(\omega) + r_a(\omega) - d_a(\omega)\right) \\ 0 \ &\leq \ E - \textstyle\sum_a \mathbb{E}[\mathcal{E}(y_a, \omega)]. \end{aligned}$$

In this model, the prices on the supply and emission constraints are determined endogenously from the solution. Observe that if we explicitly introduce Lagrange multipliers, $\pi(\omega)$ for the constraint

$$0 \leq \textstyle\sum_a \left(q_a(\omega) + r_a(\omega) - d_a(\omega)\right)$$

and σ for the emission constraint

$$0 \leq E - \textstyle\sum_a \mathbb{E}[\mathcal{E}(y_a, \omega)]$$

we can derive an equivalent Lagrangian version of P:

$$\begin{aligned} \text{L: } \min\ \rho(\psi) \ \text{ s.t. } \ & (x_a, y_a, z_a, q_a, r_a) \in \mathcal{X}_a, \forall a \\ \psi(\omega) \ &= \ \textstyle\sum_a \Big(\mathcal{C}_a(x_a, z_a) + \mathcal{Z}_a(y_a, q_a, r_a, \omega) \\ &\quad + \pi(\omega)\left(d_a(\omega) - q_a(\omega) - r_a(\omega)\right) + \sigma(\omega)\mathbb{E}[\mathcal{E}(y_a, \omega)]\Big) \\ &\quad - \sigma(\omega) E, \end{aligned}$$

where

$$\begin{aligned} 0 \leq \sum_a \left(q_a(\omega) + r_a(\omega) - d_a(\omega)\right) \perp \pi(\omega) \geq 0, \\ 0 \leq E - \sum_a \mathbb{E}[[\mathcal{E}(y_a, \omega)] \perp \sigma(\omega) \geq 0. \end{aligned}$$

If the risk measure is separable across agents (for example if the agents are risk neutral) then we can decouple the Lagrangian version of P into a MOPEC (Kim and Ferris 2019), where each agent a then solves

$$\begin{aligned} \text{S}(a)\text{: } \min\ \rho_a(\psi_a) \ \text{ s.t. } \ & (x_a, y_a, z_a, q_a, r_a) \in \mathcal{X}_a \\ \psi_a(\omega) \ &= \ \mathcal{C}_a(x_a, z_a) + \mathcal{Z}_a(y_a, q_a, r_a, \omega) \\ &\quad + \pi(\omega)\left(d_a(\omega) - q_a(\omega) - r_a(\omega)\right) + \sigma\mathbb{E}[\mathcal{E}(y_a, \omega)] \end{aligned}$$

and the prices, production and purchases satisfy the market clearing conditions

$$0 \leq \sum_a (q_a(\omega) + r_a(\omega) - d_a(\omega)) \perp \pi(\omega) \geq 0,$$
$$0 \leq E - \sum_a \mathbb{E}[\mathcal{E}(y_a, \omega)] \perp \sigma \geq 0.$$

(A MOPEC is a collection of parametric optimization problems coupled with a parametric complementarity or variational inequality.)

The model S(a) imposes an ex-ante emissions price σ on expected emissions $\mathbb{E}[\mathcal{E}(y_a, \omega)]$ for agent a. The cost of this is incurred in every scenario, and so it alters their operational disbenefit $\psi_a(\omega)$ by a constant. It is arguably more natural to imagine the agent paying σ for each tonne of CO2 emitted, so

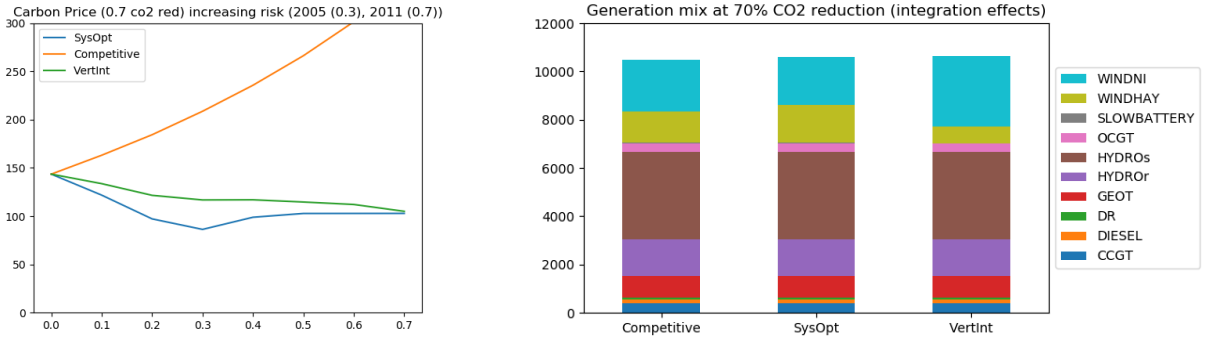
$$\hat{\psi}_a(\omega) = \mathcal{C}_a(x_a, z_a) + \mathcal{Z}_a(y_a, q_a, r_a, \omega) + \sigma\mathcal{E}(y_a, \omega),$$

and then minimizing risk-adjusted disbenefit $\rho_a(\hat{\psi}_a)$. This corresponds to a model in which σ is an exogeneous CO2 price to be set for example by an environmental regulator or by a cap and trade market involving sectors outside the electricity system.

If the agents are risk neutral ($\rho_a = \mathbb{E}$) then the models are equivalent. In the general case where the agents are not risk neutral, the two models for emission charges are not equivalent. Furthermore, in the absence of a complete market for risk (i.e. where there is a full set of Arrow-Debreu securities), then as detailed in Ferris and Philpott (2018), neither model of emission pricing gives a MOPEC that corresponds to the system optimization problem P given above. It is clear that when some of the agents represent supply and others represent demand sectors, then the form of $\mathcal{C}_a$ and $\mathcal{Z}_a$ are quite different (sometimes even vanishing for specific agents) and this gives rise to richer forms of models and interaction effects.

We illustrate some of these effects using a simplification of our model to two inflow scenarios (years 2005 and 2011 with probabilities 0.3 and 0.7) and original load data. We impose a CO2 target of 900 kt (corresponding to $\theta = 0.7$).

The results are shown in Figure 9a. This plots the value of σ in equilibrium for three different



(a) Carbon prices with increasing λ (b) Ownership

Figure 9 Affects of increasing risk aversion carbon price and investment

industry structures and levels of risk aversion λ ranging from 0 (risk neutral) to 0.7, where risk aversion is modeled using $(1-\lambda)\mathbb{E}[Z]+\lambda\text{AVaR}_{0.90}(Z)$. The industry structures were a socially planned system where all technologies are operated as a state monopoly (SysOpt), a "competitive" system with five competing companies called WIND/SOLAR, HYDRO, THERMAL, BATTERY/DR, GEOTHERMAL/CCS (Competitive) and a "vertically-integrated" system with five companies called WIND/THERMAL, HYDRO, SOLAR, BATTERY/DR, GEOTHERMAL/CCS (VertInt). The vertically integrated system enables some risk pooling between wind generation and thermal plant.

As shown in Figure 9a, the equilibrium CO2 price (σ) is the same value for all industry structures when agents are risk neutral ($\lambda=0$). As risk aversion of the agents increase, the risk-averse system optimal solution requires less expensive CO2 prices to meet the ex-ante emissions constraint. This is similar to the solution for the vertically integrated industry structure. However in the competitive case with more players, there are fewer opportunities to pool risk and so as risk aversion increases the CO2 price needed to meet the emissions constraint increases.

Figure 9b picks the three points on the graphs corresponding to $\lambda=0.3$ and compares the capacity choices. Surprisingly, despite the large differences in CO2 prices for these cases, the optimal capacity choices are very similar. Wind generation expansion is shifted between locations NI and HAY but the total amount remains similar. In the vertically integrated case, more wind is built

in NI closer to existing thermal plant that are co-owned. One can see that higher carbon charges are needed in the competitive case to induce optimal operations of the wind and thermal plant to meet the expected emissions target without risk pooling.

7. Conclusion

This paper has presented a two-stage stochastic programming model for planning the expansion of electricity generation to achieve specific renewable-energy targets. Different policy prescriptions can be modeled using different formulations of the objective function of this model, each of which is shown to yield different capacity mixes and generation outcomes. This serves to illuminate the advantages and drawbacks of different policy choices.

Versions of our model have been applied to a case study using New Zealand data. This study is not intended to compute recommended policy choices, but rather to serve as an illustration of the power of our models. The models show that there are important differences between policies that mandate nonrenewable capacity limits and those that mandate nonrenewable energy (and CO_2) levels. The latter class of model is more aligned with the underlying objective of reducing greenhouse gas emissions. Similarly chance-constrained versions of our model yield less desirable emission outcomes than those that focus on constraining average emissions.

Since the models we are considering have long time horizons over which uncertain effects will become realized gradually, there is considerable value in developing a multistage version of our model. One approach uses multihorizon scenario trees (see Kaut et al. 2014), in which investment decisions are made at nodes of a coarse scenario tree that models long-term uncertainty (such as load and technology advances) and operational decisions are made at the nodes of this tree, subject to short-term uncertainty (as modeled by the scenarios described in this paper). Developing such a model for the New Zealand electricity system is the subject of a forthcoming companion paper (Downward et al. 2019).

Acknowledgments

This project has been supported in part by the US Department of Energy. The paper was completed while the authors were participating in the Mathematics of Energy Systems thematic programme at the Isaac Newton Institute at the University of Cambridge. The authors wish to acknowledge the support of the INI, and the second author acknowledges the support of the New Zealand Marsden Fund under contract UOA1520. The results of the paper are based on the research of the authors and do not represent any official view of the New Zealand Interim Climate Committee or the New Zealand Government.

References

Artzner P, Delbaen F, Eber JM, Heath D (1999) Coherent measures of risk. *Mathematical finance* 9(3):203–228.

Bishop P, Bull B (2008) The future of electricity generation in New Zealand. *13th annual conference of the New Zealand Agricultural and Resource Economics Society, Nelson, New Zealand*, 28–29.

Boffino L, Conejo A, Sioshansi R, Oggioni G (2018) A two-stage stochastic optimization planning framework to deeply decarbonize electric power systems. Technical Report downloadable from https://u.osu.edu/sioshansi.1/ramblings/, Department of Integrated Systems Engineering, Ohio State University.

De Jonghe C, Hobbs B, Belmans R (2012) Optimal generation mix with short-term demand response and wind penetration. *IEEE Transactions on Power Systems* 27(2):830–839.

Downward A, Ferris M, Philpott A (2019) Multistage models for 100 percent renewable electricity. Technical report, Electric Power Optimization Centre (EPOC), Auckland, NZ.

Electricity Market Information System (2019) Technical Report Downloaded from https://www.emi.ea.govt.nz , New Zealand Electricity Authority.

Ferris MC, Philpott AB (2018) Dynamic risked equilibrium. *Operations Research* Submitted.

Fishbone L, Abilock H (1981) MARKAL , a linear-programming model for energy systems analysis: Technical description of the bnl version. *International Journal of Energy Research* 5(4):353–375.

Fulton B (2018) *Security of Supply in the New Zealand Electricity Market* (BE Honours Project Report, Department of Engineering Science, University of Auckland).

Graf C, Marcantonini C (2017) Renewable energy and its impact on thermal generation. *Energy Economics* 66:421–430.

Joskow P (2006) Competitive electricity markets and investment in new generating capacity. Technical report, AEI-Brookings Joint Center Working Paper 06-14.

Kaut M, Midthun K, Werner A, Tomasgard A, Hellemo L, Fodstad M (2014) Multi-horizon stochastic programming. *Computational Management Science* 11(1-2):179–193.

Khazaei J, Powell W (2017) SMART-Invest: a stochastic, dynamic planning for optimizing investments in wind, solar, and storage in the presence of fossil fuels. the case of the PJM electricity market. *Energy Systems* 1–27.

Kim Y, Ferris MC (2019) Solving equilibrium problems using extended mathematical programming. *Mathematical Programming Computation* URL `http://dx.doi.org/10.1007/s12532-019-00156-4`, online first.

Kok C, Philpott A, Zakeri G (2018) Value of transmission capacity in electricity markets with risk averse agents. Technical Report www.epoc.org.nz/papers/TransmissionPaperOperationsResearch.pdf, EPOC Working Paper.

Loulou R (2008) ETSAP-TIAM: the TIMES integrated assessment model. Part II: Mathematical formulation. *Computational Management Science* 5(1-2):41–66.

Loulou R, Labriet M (2008) ETSAP-TIAM: the TIMES integrated assessment model Part I: Model structure. *Computational Management Science* 5(1-2):7–40.

Mason I, Page S, Williamson A (2010) A 100% renewable electricity generation system for New Zealand utilising hydro, wind, geothermal and biomass resources. *Energy Policy* 38(8):3973–3984.

Mason I, Page S, Williamson A (2013) Security of supply, energy spillage control and peaking options within a 100% renewable electricity system for New Zealand. *Energy Policy* 60:324–333.

Masse P, Gibrat R (1957) Application of linear programming to investments in the electric power industry. *Management Science* 3(2):149–166.

New Zealand ICCC Terms of Reference (2019) Technical Report Downloaded from https://www.iccc.mfe.govt.nz/who-we-are/terms-of-reference , New Zealand Interim Climate Change Committee.

Ralph D, Smeers Y (2015) Risk trading and endogenous probabilities in investment equilibria. *SIAM Journal on Optimization* 25(4):2589–2611.

Rubin E, Zhai H (2012) The cost of carbon capture and storage for natural gas combined cycle power plants. *Environmental science & technology* 46(6):3076–3084.

Short W, Sullivan P, Mai T, Mowers M, Uriarte C, Blair N, Heimiller D, Martinez A (2011) Regional Energy Deployment System (ReEDS). Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States).

Skar C, Doorman G, Tomasgard A (2014) The future European power system under a climate policy regime. *2014 IEEE International Energy Conference (ENERGYCON)*, 318–325 (IEEE).

Stoft S (2002) *Power System Economics* (IEEE press Piscataway).

US Energy Information Administration (EIA) (2019) Capital cost estimates for utility-scale electricity generation. *Downloaded from https://www.eia.gov* .

Wu A, Philpott A, Zakeri G (2017) Investment and generation optimization in electricity systems with intermittent supply. *Energy Systems* 8(1):127–147.

Appendices

Appendix A: Data used in New Zealand model

	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	10	50	100	200	300	300	300	300	300	300
1	10	50	124	200	300	300	300	300	300	300
2	10	50	148	200	300	300	300	300	300	300
3	10	50	148	200	300	300	300	300	300	300

Table 5 Number of hours in each load block b in each season $t = 0, 1, 2, 3$.

SI	b1	b2	b3	b4	b5	b6	b7	b8	b 9	b10
0	2194	2152	2113	2055	1989	1913	1845	1752	1620	1473
1	2302	2263	2219	2146	2042	1942	1833	1710	1587	1447
2	2311	2290	2254	2178	2073	1959	1857	1755	1634	1478
3	2188	2189	2170	2096	1999	1924	1873	1778	1670	1498
HAY	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	583	570	550	541	524	502	469	417	355	328
1	744	722	707	665	613	580	545	474	411	364
2	760	735	720	681	633	597	563	500	429	376
3	630	591	570	552	535	512	473	422	369	343
NI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	3502	3465	3428	3335	3227	3027	2797	2504	2119	1905
1	4196	4106	3912	3656	3427	3264	3012	2635	2250	1981
2	4310	4160	4011	3810	3583	3380	3172	2805	2350	2079
3	3692	3578	3480	3382	3255	3092	2856	2550	2164	1955

Table 6 Estimated 2035 demand in each load block b in each season $t = 0, 1, 2, 3$, in each region i.

SI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	0.158	0.205	0.338	0.318	0.311	0.214	0.253	0.131	0.015	0.003
1	0.016	0.030	0.038	0.067	0.123	0.121	0.077	0.078	0.006	0.001
2	0.008	0.028	0.050	0.084	0.121	0.149	0.162	0.064	0.006	0.001
3	0.409	0.370	0.371	0.375	0.248	0.218	0.230	0.157	0.045	0.017
HAY	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	0.158	0.205	0.338	0.318	0.311	0.214	0.253	0.131	0.015	0.003
1	0.016	0.030	0.038	0.067	0.123	0.121	0.077	0.078	0.006	0.001
2	0.008	0.028	0.050	0.084	0.121	0.149	0.162	0.064	0.006	0.001
3	0.409	0.370	0.371	0.375	0.248	0.218	0.230	0.157	0.045	0.017
NI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	0.246	0.271	0.348	0.321	0.277	0.185	0.240	0.129	0.013	0.002
1	0.025	0.039	0.050	0.083	0.133	0.143	0.099	0.079	0.008	0.001
2	0.011	0.042	0.081	0.103	0.137	0.152	0.171	0.077	0.010	0.001
3	0.374	0.342	0.340	0.309	0.224	0.202	0.234	0.155	0.036	0.012

Table 7 Estimated parameters $\mu_{3,i,b,t}$ for photovoltaic solar generation as a proportion of capacity.

SI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	0.25	0.18	0.14	0.28	0.37	0.37	0.41	0.37	0.34	0.38
1	0.34	0.18	0.14	0.28	0.37	0.37	0.41	0.37	0.34	0.38
2	0.35	0.17	0.25	0.26	0.27	0.30	0.37	0.36	0.32	0.35
3	0.22	0.11	0.16	0.24	0.29	0.32	0.37	0.33	0.28	0.37
HAY	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	0.35	0.15	0.17	0.36	0.51	0.48	0.46	0.47	0.35	0.54
1	0.30	0.15	0.17	0.36	0.51	0.48	0.46	0.47	0.35	0.54
2	0.44	0.31	0.38	0.42	0.44	0.45	0.53	0.47	0.41	0.61
3	0.32	0.11	0.19	0.35	0.40	0.43	0.48	0.50	0.35	0.48
NI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	0.28	0.25	0.30	0.40	0.49	0.46	0.48	0.52	0.40	0.55
1	0.52	0.25	0.30	0.40	0.49	0.46	0.48	0.52	0.40	0.55
2	0.36	0.31	0.34	0.35	0.34	0.37	0.45	0.41	0.33	0.45
3	0.19	0.13	0.23	0.28	0.38	0.45	0.41	0.42	0.31	0.48

Table 8 Estimated parameters $\mu_{k,i,b,t}$ for wind generation for the scenario when wind blows in block 1. The same table applies when wind does not blow in block 1, except that the column headed b1 is replaced by zeroes.

SI	0	1	2	3	NI	0	1	2	3
2005	0.611	0.471	0.448	0.397	2005	0.400	0.338	0.380	0.413
2006	0.431	0.487	0.437	0.644	2006	0.360	0.510	0.621	0.384
2007	0.492	0.506	0.492	0.612	2007	0.335	0.230	0.512	0.350
2008	0.486	0.413	0.399	0.572	2008	0.209	0.294	0.685	0.378
2009	0.464	0.594	0.531	0.549	2009	0.275	0.292	0.506	0.464
2010	0.495	0.601	0.536	0.579	2010	0.303	0.335	0.603	0.416
2011	0.618	0.539	0.441	0.553	2011	0.405	0.503	0.511	0.402
2012	0.394	0.439	0.522	0.572	2012	0.436	0.432	0.515	0.385
2013	0.498	0.554	0.656	0.592	2013	0.201	0.260	0.339	0.390
2014	0.555	0.601	0.673	0.630	2014	0.209	0.349	0.396	0.348
2015	0.532	0.632	0.598	0.611	2015	0.168	0.413	0.527	0.348
2016	0.532	0.680	0.592	0.647	2016	0.269	0.330	0.545	0.483
2017	0.562	0.381	0.500	0.513	2017	0.356	0.564	0.682	0.450

Table 9 Parameters $\hat{\mu}_{i,t}$ for run-of-river generation. Since there is no hydro generation possible in the HAY region these data are not estimated.

SI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	1.305	1.305	1.305	1.000	1.000	1.000	1.000	0.946	0.946	0.946
1	1.383	1.383	1.383	1.000	1.000	1.000	1.000	0.922	0.922	0.922
2	1.415	1.415	1.415	1.000	1.000	1.000	1.000	0.904	0.904	0.904
3	1.263	1.263	1.263	1.000	1.000	1.000	1.000	0.939	0.939	0.939
NI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	1.546	1.546	1.546	1.000	1.000	1.000	1.000	0.903	0.903	0.903
1	1.680	1.680	1.680	1.000	1.000	1.000	1.000	0.861	0.861	0.861
2	1.488	1.488	1.488	1.000	1.000	1.000	1.000	0.887	0.887	0.887
3	1.447	1.447	1.447	1.000	1.000	1.000	1.000	0.897	0.897	0.897

Table 10 Run-of-river flexibility parameters $\alpha_{i,t,b}$ as estimated from historical run-of-river hydro generation in each region . The sum of $\alpha_{i,b,t}$ weighted by the hours in each block equals the number of hours in season t. The parameters $\mu_{k,i,b,t}$ for $k =$run-of-river plant are computed to be $\mu_{k,i,b,t} = \alpha_{i,b,t}\hat{\mu}_{i,t}$.

SI	0	1	2	3	NI	0	1	2	3
2005	1.141	0.575	0.593	0.637	2005	0.365	0.321	0.451	0.494
2006	0.851	0.729	0.615	1.144	2006	0.370	0.480	0.578	0.466
2007	0.862	0.622	0.594	0.938	2007	0.342	0.306	0.551	0.427
2008	0.893	0.471	0.652	1.140	2008	0.204	0.376	0.754	0.486
2009	0.839	0.832	0.691	0.920	2009	0.286	0.319	0.535	0.480
2010	0.924	0.808	0.680	1.115	2010	0.271	0.407	0.629	0.419
2011	0.803	0.652	0.492	0.957	2011	0.385	0.472	0.499	0.531
2012	0.723	0.517	0.707	0.975	2012	0.422	0.392	0.607	0.451
2013	0.774	0.685	0.774	1.064	2013	0.224	0.375	0.420	0.461
2014	0.759	0.922	0.588	0.988	2014	0.215	0.391	0.453	0.402
2015	0.789	1.016	0.579	0.980	2015	0.213	0.467	0.579	0.361
2016	1.129	0.894	0.613	0.822	2016	0.316	0.393	0.591	0.475
2017	0.723	0.517	0.707	0.975	2017	0.422	0.392	0.607	0.451

Table 11 Parameters $\nu_{k,i,t}$ for stored hydro generation. Since there is no hydro generation possible in the HAY region these data are not estimated.

SI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	2550	2506	2442	2346	2255	2159	2076	1932	1789	1630
1	2579	2538	2497	2410	2285	2169	2024	1871	1736	1585
2	2588	2565	2532	2442	2316	2186	2048	1916	1784	1615
3	2544	2543	2499	2388	2265	2171	2104	1959	1839	1655
HAY	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	787	771	742	713	681	649	610	528	463	429
1	913	887	876	825	760	718	668	577	510	456
2	929	900	890	841	780	736	686	602	528	468
3	834	791	762	724	693	659	614	533	476	444
NI	b1	b2	b3	b4	b5	b6	b7	b8	b9	b10
0	4355	4315	4223	4046	3876	3641	3376	2967	2551	2315
1	4891	4798	4604	4312	4029	3839	3513	3059	2644	2352
2	5005	4852	4703	4466	4185	3955	3673	3229	2743	2450
3	4544	4428	4274	4093	3904	3707	3435	3013	2597	2365

Table 12 Increased estimated 2035 demand in each load block b in each season $t = 0, 1, 2, 3$, in each region i.