

EMI-DOASA

Andy Philpott and Geoffrey Pritchard
Stochastic Optimization Limited
www.sol.co.nz

November 28, 2013

Abstract

Dynamic Outer Approximation Sampling Algorithm (DOASA) is a version of SDDP for hydro-thermal scheduling. This document gives an overview of the application of this algorithm to a hydro-thermal scheduling model of the New Zealand electricity system, as implemented as part of the EMI system maintained by the Electricity Authority.

1 Introduction

Dynamic Outer Approximation Sampling Algorithm (DOASA) is an optimization methodology for hydro-thermal scheduling and water valuation. This document describes the DOASA algorithm and its implementation in the code EMI-DOASA distributed by the Electricity Authority of New Zealand. The DOASA code developed by Stochastic Optimization Ltd¹ is based on the Stochastic Dual Dynamic Programming (SDDP) algorithm of Pereira and Pinto [6]. The aim of DOASA (like SDDP) is to construct a policy of releasing water from reservoirs and generating from thermal plant to minimize the expected cost of meeting electricity demand over a fixed time horizon. The implementation of this algorithm, as distributed by the Electricity Authority, divides a year up into 52 weekly stages and is restricted to the New Zealand electricity system. At the end of each week we may observe the state of the system which is represented by a set of reservoir levels in seven lakes (Taupo, Benmore, Pukaki, Tekapo, Ohau, Hawea, Manapouri) measured in cubic metres above some minimum level datum. The

¹DOASA is the result of several years of research in the Electric Power Optimization Centre at the University of Auckland with contributions by Andy Philpott, Vitor de Matos, Geoff Pritchard, and Ziming Guan. The SOL implementation was developed by Andy Philpott, Vitor de Matos, and Geoff Pritchard.

operating policy for the system is defined implicitly by a cost-to-go function or Bellman function. This function represents the expected future cost of operating an optimal policy from the end of the current week till the end of the planning horizon if the system is in a given state, i.e. there is a certain amount of water stored in each reservoir. The optimal decision of what water to release and what thermal generation to dispatch is then determined by solving a single-stage optimization with known inflows that minimizes the sum of thermal generation and shortage cost in that week plus the expected cost to go from the levels of water in each reservoir that result at the end of the week.

The DOASA algorithm performs many of the same steps as the SDDP algorithm as described in [6]. To avoid confusion, we should note that SDDP is the trade name of the software distributed by the Brazilian company Power Systems Research (<http://www.psr-inc.com.br>). It is also used by other software providers to refer to the algorithmic technique used by their software. The name DOASA was chosen by researchers at the Electric Power Optimization Centre to avoid this confusion.

The first versions of DOASA were described by Philpott and Guan in [7]. At this time (2007) the main differences between DOASA and other versions of SDDP were in the sampling procedures. As opposed to SDDP that created fixed sample paths of 200 inflow realizations to be used throughout the algorithm, DOASA resampled the inflow sequences randomly at every iteration. This resampling is necessary to ensure the almost sure convergence of the method. DOASA also samples the inflow sequences once in each iteration rather than working with a batch of 200. This proved to give better policies when the number of iterations is small². The implementation of DOASA described here has both of these features.

The random variables in DOASA are the inflows to the hydro systems³. These are sampled from the historical record, and are assumed to be stagewise independent. Since the samples are taken from all locations in a given week in a randomly chosen year, they preserve the spatial dependence between inflows in the sample. Some implementations of SDDP also model stagewise dependence (i.e. from week to week) in inflows using periodic autoregressive moving average (PARMA) models. These models relate variations in the current week's inflows to those observed in previous weeks. This approach multiplies the number of coefficients needed in the cutting planes by the number of lags in the PARMA model. Although this will improve the accuracy of the cutting plane approximation at the places where

²Some recent work in a forthcoming paper by de Matos and Philpott shows that for a large number of iterations, there are benefits of using more than one inflow sequence in each iteration.

³It is often assumed that all randomness in this method must be confined to the inflows. Although this assumption is made for the SOL implementation described here, it is not a serious restriction for the method and other random parameters (with finitely many realizations) can be included, albeit with increases in computation time.

it is computed, the increase in dimension of the problem can entail a significant increase in the number of cutting planes required to yield a good policy. For this reason we have chosen in the SOL implementation to keep the number of reservoirs with cutting plane coefficients to a minimum (7) and assumed stage-wise independence for inflows. Even a simple system like this takes hundreds of iterations to converge to a policy that is close to optimal (for the model we have created). The policy can easily be tested by simulation on inflow sequences with very complicated stagewise dependence, and in many cases it performs well.

The document is laid out as follows. In the next section we describe the medium-term hydro-scheduling problem in mathematical terms. This is the optimization problem that DOASA is designed to solve. Section 3 describes the input files for DOASA and what they contain, and section 4 describes the outputs delivered by the program. The final section gives a mathematical description of the steps of the algorithm used by DOASA for solving this problem.

2 The medium-term hydro-scheduling model

The DOASA model seeks a policy of electricity generation that meets demand and minimizes the expected fuel cost of thermal generation. All data are deterministic except for weekly inflows that are assumed to be stagewise independent. This results in a large-scale stochastic dynamic programming model which is defined as follows. Let $x_j(t)$ denote the storage in reservoir j at the *end* of week t , and let $C_t(x, \omega(t))$ be the minimum expected fuel cost to meet electricity demand in weeks $t, t+1, \dots, T$, when reservoir storage $x_j(t-1)$ at the start of week t is equal to \bar{x}_j and week t 's inflow is known to be $\omega_j(t)$. Here $C_t(x, \omega(t))$ is the optimal solution value of the mathematical program:

$$\begin{aligned}
P_t(\bar{x}, \omega): \quad & \min \quad \sum_{i \in \mathcal{N}} \sum_{m \in \mathcal{F}(i)} \phi_m \sum_b T(b, t) f_m(b, t) + \mathbb{E}[C_{t+1}(x(t), \omega(t+1))] \\
\text{s.t.} \quad & g_i(y(b, t)) + \sum_{m \in \mathcal{F}(i)} f_m(b, t) + \sum_{m \in \mathcal{H}(i)} \gamma_m h_m(b, t) = D_i(b, t), \quad i \in \mathcal{N}, \\
& x(t) = \bar{x} - S \sum_b T(b, t) (Ah(b, t) + As(b, t) - \omega(t)), \\
& 0 \leq f_m(t) \leq a_m, \quad m \in \mathcal{F}(i), \quad i \in \mathcal{N}, \\
& 0 \leq h_m(t) \leq b_m, \quad 0 \leq s_m(t) \leq c_m, \quad m \in \mathcal{H}(i), \\
& 0 \leq x_j(t) \leq r_j, \quad j \in \mathcal{J}, \quad i \in \mathcal{N}, \quad y \in Y.
\end{aligned}$$

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 thermal plants at node i
$\mathcal{H}(i)$	set of hydro plants at node i
\mathcal{J}	set of reservoirs

and the notation:

Symbol	Meaning	Units
ϕ_m	short-run marginal cost of thermal plant m	\$/MWh
$f_m(b, t)$	generation of thermal plant m in load block b 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
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)	
$y(b, t)$	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
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	node-arc incidence matrix	

Here the water balance constraints in the storage reservoirs at the end of week t are represented by

$$x(t) = \bar{x} - S \sum_b T(b, t)(Ah(b, t) + As(b, t) - \omega(t))$$

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 node-arc incidence matrix A represents the river-valley network, and 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.

3 Specific features of the DOASA code

In this section we look at some specific modelling assumptions that are made in our model.

3.1 Time

The DOASA model uses weekly stages. A calendar year is divided into 52 weeks. A plan year is typically a year of 52 weeks with the starting week chosen to be a particular week in the plan year and a fixed number of weeks used. Historical inflows are sampled from a file that records weekly inflows as described below. The New Zealand electricity system is represented as shown in Figure 1 below.

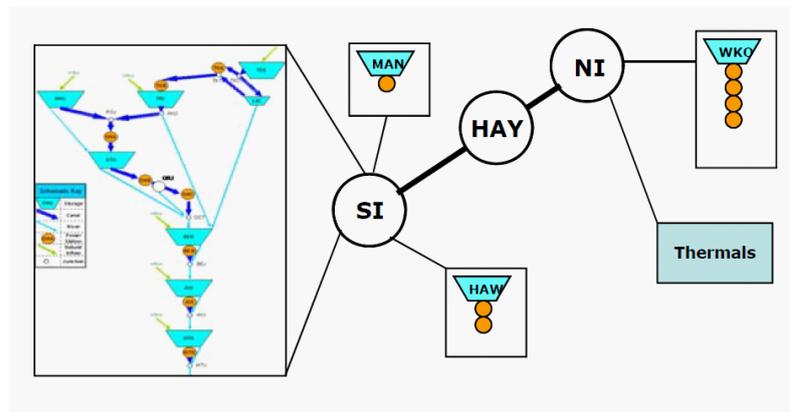


Figure 1: The 3 node transmission network and major generators in DOASA.

3.2 Load blocks

Weekly demand is represented by a load duration curve with three blocks. These are called peak, off-peak and shoulder. The load duration curve can be constructed by reordering national halfhourly demand from highest to lowest, and then arbitrarily determining which hours correspond to peak, off-peak and shoulder respectively. (This assumes that the same ordering applies to all nodes in the transmission system.) The total demand in MW in each node i is then averaged over these trading periods to get a total demand rate $D_i(b, t)$ for each block. The energy requirement in node i for each block b in week t will be its duration $T(b, t)$ times the average demand rate $D_i(b, t)$ for this block.

The choice of what data to include in demand is a delicate matter. Publicly available demand figures (e.g. those in the Centralized Data Set [1] (CDS)) make various assumptions about what embedded generation and demand is included. These must be carefully studied to ensure that demand is not overlooked or double counted.

The DOASA model aggregates demand to three locations (SI, HAY, NI) representing the South Island, the lower North Island, and the upper North Island, and allows transfers between these regions limited by line capacities. We emphasize here that *DOASA assumes that all line losses from transmission are zero*. This means that aggregating demand in each region will ignore the intra-regional losses, implying that the regional totals of historical demand will underestimate the true demand to be met by generation and net imports to SI, HAY, and NI. So some inflation of total demand is needed to account for these losses. If the aggregation is carried out geographically then ignoring line losses might also bias the generation mix in the dispatch towards geographically close (yet electrically distant) plant.

To overcome these effects, the user can adopt some form of demand scaling. One option is to apply a uniform scaling to demand, but this does not reflect the fact that demand is concentrated in some nodes (e.g. TWI which contains a large aluminium smelter) and not in others (e.g. WKM that meets the needs of a dispersed region in our model). The approach followed will depend on the purpose of using DOASA. If one is back testing on a historical year then the aggregate (into SI, HAY, and NI) of historical dispatch at the generators can be used as a proxy for the demand adjusted for losses. Of course this will beg the question of whether this was an optimal allocation of generation that DOASA is intended to compute, but it might serve as a starting point. Another option is to estimate historical demand using the approach of Philpott and Guan as described in [8]. If the user is applying DOASA to a future year then the estimates of aggregate demand and its distribution over load blocks will depend on the scenario that is being investigated.

In meeting demand, in case of supply shortages, load shedding (in MW) is allowed at high costs. The costs depend on the type of customers and amount of reduction (in \$/MWh). Load in each node is divided into three sectors to represent different types of customers, which are industrial, commercial and residential, and each sector has some distribution in each island. The default proportions are the proportions of consumption in 2003 adjusted to higher commercial and residential proportions in the North Island due to a denser population, and to a higher industrial proportion in the South Island due to an aluminium smelter. Although these proportions change over the years, for simplicity we have assumed they are constant. Each sector is then divided into three segments to represent the amount of reduction, namely 5%, 5% and 90%. The third segment represents unplanned interruption of power supply. The cost for load shedding is called the *value of loss load*, or *VOLL*, in the electricity industry. The VOLLs for the industrial sector are set to be lower than the other two and the VOLLs increase over segments in each sector. We assume that up to 10% reduction in load can be achieved at a relatively low cost, but the value of unplanned interruption is very high (\$10,000/MWh).

3.3 Inflows

The inflows $\omega(t)$ are sampled from historical inflow observations. Inflows are measured in cumecs as weekly averages. The hydro system assumes that seven reservoirs, Manapouri-Te Anau, Hawea, Ohau, Pukaki, Tekapo, Benmore and Taupo, can store water from week to week. The remaining reservoirs are treated as run-of-river plant with limited intra-week flexibility.

DOASA constructs a policy by solving a large number of stage problems, which are linear programs. In order for the algorithm to work, each stage problem must have a feasible solution. Otherwise the algorithm will terminate with an infeasible stage problem message. One source of infeasibility is a week in which inflows are so low that minimum flow constraints (such as those for Karapiro) must be violated. DOASA allows this violation in every stage, with a penalty cost per MWh of violation (LB flow penalty) that is defined in the run file. In a similar fashion, flood conditions sometimes require that river flows exceed their nominated upper bounds. These are also allowed with a penalty cost per MWh of violation (UB flow penalty) that is defined in the run file. The user can experiment with these to try and settle on good values. In practice some degree of violation is tolerated in extreme circumstances and so the penalties should reflect this reality. The simulation output indicates when these penalties are being incurred, and if this is deemed to be too often the user can increase the penalty costs to discourage this.

3.4 Dependent Inflow Adjustment

DOASA assumes stagewise independent inflows. In reality, reservoir inflows are not stagewise independent. This means that random sequences of consecutive inflows sampled from the assumed inflow distribution will have a different distribution from the real inflow sequences. This difference can lead to DOASA producing overly optimistic policies during dry periods, which will give underestimates of the true marginal water value during these periods.

Ideally the inflow distribution should be chosen to reflect the stagewise dependence properties of the actual inflow process. The exact nature of this dependence can only be deduced from historical observations, and so there are limitations on how close the model process will be to the true process. Furthermore, the true distribution can only be fitted from data, and some fitted distributions will be unsuitable for use in SDDP-type algorithms like DOASA. A common assumption is to use PARMA models, which requires linear autoregressive models to be fitted. Such assumptions will lead to marginal water value estimates that might differ from the true values.

An alternative approach adjusts the empirical inflow distribution to make it share some properties that we believe are present in the true distribution. There are different approaches to doing this. The approach followed by DOASA is called the *Dependent Inflow Adjustment* (DIA). DIA is applied to inflows when cutting planes are computed (and so it affects the policy and the marginal water values), but not in historical simulations, where the actual sequences are used.

DIA works on the inflow sequence for each reservoir separately as follows. Consider an historical inflow sequence $h(t, y)$ of inflows in week t of year $y = 1, 2, \dots, N$. Given such a sequence, denote the sample average inflow for week t of the year by

$$a(t) = \frac{\sum_y h(t, y)}{N}$$

and the one period standard deviation for week t of the year by

$$\sigma(t) = \sqrt{\sum_y \frac{(h(t, y) - a(t))^2}{N}}.$$

In DIA, the user selects an integer number of weeks (w) for which the inflow sequence will be adjusted. The number w is the same throughout the year, and loosely represents the horizon over which we wish to represent the total inflow distribution as well as possible. If $w = 0$ or 1 , then the inflows will remain the same as the original data, assumed to be stagewise independent.

If $w > 1$, then w consecutive weeks will be used to adjust the single period historical inflows. For each starting week t the total inflow over w weeks is

computed, to give for each t :

$$W(t, y) = h(t, y) + h(t + 1, y) + \dots + h(t + w - 1, y).$$

The average $m(t)$ and standard deviation $\rho(t)$ of $W(t, y)$ is then computed:

$$m(t) = \frac{\sum_y W(t, y)}{N}$$

$$\rho(t) = \sqrt{\frac{\sum_y (W(t, y) - m(t))^2}{N}}.$$

These values represent the mean and standard deviation from the random variable which is the sum of w consecutive inflows. An independent inflows assumption would give a standard deviation of about $\sqrt{w}\sigma(t)$ which will be smaller than $\rho(t)$ in general. To adjust the standard deviation, $h(t, y)$ is replaced by

$$g(t, y) = a(t) + \frac{(W(t, y) - m(t))}{\sqrt{w}} \quad (1)$$

Then the sum of w consecutive such inflows assumed to be independent has variance

$$w \frac{1}{N} \sum_y \left(\frac{W(t, y) - m(t)}{\sqrt{w}} \right)^2 = \rho(t)^2.$$

We require $g(t, y)$ to be non-negative, but some of the values obtained by (1) will be negative. These values are truncated to give

$$d(t, y) = \max\{0, g(t, y)\}.$$

Truncation implies that the averages of the $d(t, y)$ values will overestimate the true average inflows and so the inflows are scaled down to give a sequence with the average matching that of $a(t)$. In other words the final adjusted inflows we use are:

$$k(t, y) = \frac{Nd(t, y)a(t)}{\sum_y d(t, y)}.$$

Observe that truncation and scaling will reduce the variance slightly, so the sum of w consecutive $k(t, y)$ will have variance less than $\rho(t)^2$ if the sum involves weeks where truncation occurs.

3.5 Generators

In New Zealand, dispatches of some generators have limited control, such as those from cogeneration, geothermal plant, run-of-river hydro and wind. Although these have low marginal cost, their availability is subject to the vagaries of inflows and wind, and so we cannot centrally dispatch these in our model.

Our approach is to treat some of this generation as *embedded* (to be subtracted from demand⁴) and some to be *fixed* generation (to be dispatched at fixed levels). The remaining generation (thermal and hydro) is treated as variable. How to treat different generators is at the discretion of the user, but care must be taken to avoid double counting or missing generation. In DOASA fixed generation is defined in the file `fixed_stations.csv`. This file should include all cogeneration, geothermal plant, run-of-river hydro and wind that is not included in demand already, and is not already included in the files `thermal_stations.csv` and `hydro_stations.csv`. Some essentially run-of-river hydro plant is also included in `hydro_stations.csv`. These stations do not store water, and so generate the minimum of their capacity and the power that would result from using the observed inflow and upstream releases. Each generator has a nominal dispatch capacity (in MW), and a nominal conversion factor (in MW/cumec) to compute energy generated from flow through the generator at the average headwater levels.

3.6 Thermal fuel costs

Thermal generators run on different types of fuel, either coal, gas, or diesel, and have different heat rates (in GJ/MWh). Each station has a nominal capacity (in MW). We assume that a generator can supply any quantity of power up to its capacity at the fuel cost (in \$/MWh). The fuel cost is the product of heat rate and the wholesale cost for fuel (in \$/GJ). The offering thermal generators (using CDS nomenclature) are shown in Table 3.

⁴Note that the historical dispatches of embedded generators are aggregated into demands in the CDS. Their dispatches offset some demands and thus if the dispatches are larger than demand then negative values are presented. But as described in the CDS, some negative values have been set to zero, but no information is given on how much and where, and thus it is not possible to obtain all historical dispatches for embedded generators.

Generator	Station	Capacity	Heat rate	Fuel
GEN.Stratford.220KV	Stratford	387	7.3	Gas
GEN.Thermal.Huntly.gas	Huntly	430	6.8	Gas
GEN.Thermal.Huntly.main.g1	Huntly	260	10.5	Coal
GEN.Thermal.Huntly.main.g2	Huntly	260	10.5	Coal
GEN.Thermal.Huntly.main.g3	Huntly	260	10.5	Coal
GEN.Thermal.Huntly.main.g4	Huntly	265	10.5	Coal
GEN.Thermal.Huntly.Peak	Huntly	50	9.5	Gas
GEN.Thermal.NewPlymouth.110KV.g1	NewPlymouth	120	11	Gas
GEN.Thermal.NewPlymouth.110KV.g2	NewPlymouth	120	11	Gas
GEN.Thermal.NewPlymouth.220KV.g3	NewPlymouth	120	11	Gas
GEN.Thermal.Otahuhu.B	Otahuhu	396	7.05	Gas
Nl.Whirinaki.220KV	Whirinaki	159	11	Diesel

Table 3: Offering thermal generators and their locations, capacities, heat rates and fuel types.

Note that we assume all fuel can be purchased on demand at the fuel cost. In practice, natural gas is typically acquired under a take-or-pay contract that gives a different operating imperative from that faced by a purchaser with more flexibility. Similarly coal is typically used from a stockpile that is periodically restocked; in this setting, supply shortages can lead to high opportunity costs. DOASA ignores both of these restrictions.

Observe that some thermal generation units might be commissioned or become decommissioned in the period of interest. We model this by including a start year and start week for commissioning and end week and end year for decommissioning.

In practice, thermal generators are subject to ramping constraints which limit the change in dispatch between consecutive trading periods. There are also unit commitments for some thermal stations, e.g., a minimum dispatch of 110 MW for the generators in the Huntly station combined if any of them is dispatched, which may also limit their dispatch. The DOASA model ignores both of these features.

DOASA enables the user to de-rate generators (in MW) to model units that have been removed for planned maintenance. This is done by reducing the capacity of the generators by a certain number of MW in predefined weeks in the planning horizon. Users may choose these periods based on their own forecasts or make use of data in the POCP database [9]. The schedule in POCP defines the starting and end time of scheduled maintenance for generators⁵.

⁵We have observed from historical data that in some declared maintenance periods, the generators in POCP still offered or were dispatched energy. We have also observed that the reduction of capacities displayed in the schedules may not be consistent with the actual reduction in historical data.

3.7 Cuts

As stated above DOASA computes an optimal policy for minimizing the expected cost of thermal generation and shortages. The policy is defined by the optimal solution to the stage problem

$$\begin{aligned}
P_t(\bar{x}, \omega): \quad & \min \quad \sum_{i \in \mathcal{N}} \sum_{m \in \mathcal{F}(i)} \phi_m \sum_b T(b, t) f_m(b, t) + \mathbb{E}[C_{t+1}(x(t), \omega(t+1))] \\
\text{s.t.} \quad & g_i(y(b, t)) + \sum_{m \in \mathcal{F}(i)} f_m(b, t) + \sum_{m \in \mathcal{H}(i)} \gamma_m h_m(b, t) = D_i(b, t), \quad i \in \mathcal{N}, \\
& x(t) = \bar{x} - S \sum_b T(b, t) (Ah(b, t) + As(b, t) - \omega(t)), \\
& 0 \leq f_m(t) \leq a_m, \quad m \in \mathcal{F}(i), i \in \mathcal{N}, \\
& 0 \leq h_m(t) \leq b_m, \quad 0 \leq s_m(t) \leq c_m, \quad m \in \mathcal{H}(i), i \in \mathcal{N}, \\
& 0 \leq x_j(t) \leq r_j, \quad j \in \mathcal{J}, \quad y \in Y.
\end{aligned}$$

where $C_{t+1}(x(t), \omega(t+1))$ denotes the expected future cost of an optimal policy applied in weeks $t+1, t+2, \dots, T$.

To solve $P_t(\bar{x}, \omega)$ to determine the dispatch in any week we need a model for $C_{t+1}(x(t), \omega(t+1))$. DOASA's model is based on the SDDP technique of Pereira and Pinto [6]. The expected future cost function is recorded by storing an outer approximation defined by K cutting planes or *cuts*, giving

$$\mathbb{E}[C_{t+1}(x(t), \omega(t+1))] \approx \max_{k=1,2,\dots,K} \{\alpha_{t+1,k} - \beta_{t+1,k}^\top x(t)\}.$$

This is illustrated in Figure 2.

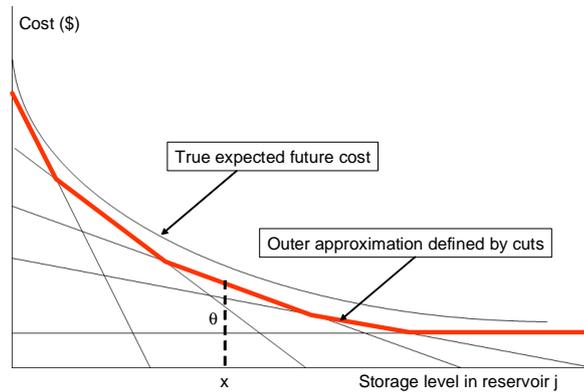


Figure 2: The cuts give the future cost θ at a particular state. Here state variable x_j is the level in reservoir j .

This gives a policy defined to be the optimal solution of the approximate stage problem

$$\begin{aligned}
\hat{P}_t(x, \omega): \quad & \min \quad \sum_{i \in \mathcal{N}} \sum_{m \in \mathcal{F}(i)} \phi_m \sum_b T(b, t) f_m(b, t) + \theta_{t+1} \\
\text{s.t.} \quad & g_i(y(b, t)) + \sum_{m \in \mathcal{F}(i)} f_m(b, t) + \sum_{m \in \mathcal{H}(i)} \gamma_m h_m(b, t) = D_i(b, t), \quad i \in \mathcal{N}, \\
& x(t) = x - \sum_b T(b, t) (Ah(b, t) + As(b, t) - \omega(t)), \\
& 0 \leq f_m(t) \leq a_m, \quad m \in \mathcal{F}(i), \quad i \in \mathcal{N}, \\
& 0 \leq h_m(t) \leq b_m, \quad 0 \leq s_m(t) \leq c_m, \quad m \in \mathcal{H}(i), \\
& \theta_{t+1} + \beta_{t+1, k}^\top x(t) \geq \alpha_{t+1, k} \quad k = 1, 2, \dots, K, \\
& 0 \leq x_m(t) \leq r_m, \quad m \in \mathcal{H}(i), \quad i \in \mathcal{N}, \quad y \in Y.
\end{aligned}$$

In summary, the optimal policy is stored as a set of K cutting planes for each stage. These are recorded as an intercept $\alpha_{t+1, k}$, and a vector $\beta_{t+1, k}$ of marginal values, defining one value for each reservoir. Given these numbers the (approximately) optimal dispatch for week t when the inflow vector for that week is ω , can be found by solving $\hat{P}_t(x, \omega)$.

4 Input files for DOASA

The input data for DOASA are provided by comma-separated files or csv files. In this section we define each of these files and explain what it contains. The input files are all contained in the directory `Input\input_csv_files`.

4.0.1 run.csv

The key file in this directory is `run.csv`. This presents the parameters for the run to be solved. These are:

Run name	text to be given to the name of the run
Save output in	text in giving the name of the directory for output (defaults to Output)
System	"Input/input_csv_files/index.csv"
Problem start year	integer giving the year of the study, e.g. 2005
Problem start week	integer giving the first week of the study (usually 1)
Number of weeks	integer giving the number of weeks in the study
Use saved cuts from	text in double quotes giving the path to cuts files or "" if no previously computed cuts available
Maximum iterations	integer giving how many new cuts to compute per stage, typically in order of thousands
Sample start year	integer giving the first year from which inflows will be sampled
Sample end year	integer giving the last year from which inflows will be sampled
Inflow correlation length	integer giving the number of weeks to use in the DIA procedure for inflows
Simulation type	text in double quotes indicating the type of simulation to run: "none" "Monte Carlo" or "historical"
Simulation sample size	Number of samples to simulate either in "Monte Carlo" or "historical"
Random seed	integer that seeds the sampling. Choosing this to be the same in different runs will replicate results.
LB flow penalty	penalty in \$/MWh on river flows going below their lower bounds.
UB flow penalty	penalty in \$/MWh on river flows going above their upper bounds.

It is important to bear in mind that DOASA has two modes, policy generation and simulation. Although a simulation can follow a policy generation using a single `run.csv` file, it is good practice to separate these into two different applications of the software. In other words, compute a policy and save the cuts and then load these to simulate the policy with a new `run.csv` file.

When computing a policy DOASA samples inflows randomly from years in the flow sequences between Sample start year and Sample end year (inclusive). There is no mechanism to select, say, only dry years from a contiguous set of years (although one could construct a synthetic inflow file and rename the dry years to be consecutive). The Sample start year and Sample end year are also used in MonteCarlo simulation. They are not used in historical simulation. All simulations begin in Problem start week of Problem start year. When historical simulations are carried out, "Simulations sample size" inflow sequences are

selected from the file `inflows.csv`. starting at the latest possible sequence.

For example suppose Simulations sample size=3. Thus if a 75 week sequence is to be simulated using historical data starting on June 1, 2006 and the inflow file runs up to December 31, 2011, then DOASA will run 3 simulations with the 75 week historical inflow sequences beginning in June 1, 2010, June 1, 2009, and June 1, 2008. Note that the June 1, 2011 does not have enough recorded data to make this eligible.

4.0.2 index.csv

The file `index.csv` defines locations for all of the input files used by DOASA. This structure need not be changed if DOASA uses the same input file names and directories in each run.

4.0.3 demand.csv

Demand is represented by three load blocks in each week. The values in each column are the number of MWh in that block. This means that they give $T(b,t)D_i(b,t)$. The demand file has six comma separated columns. The node refers to the transmission system and is either NI, HAY or SI in the EA implementation which is restricted to these three nodes.

NODE	YEAR	WEEK	peak	shoulder	offpeak
NI	2005	1	91682.52	119621.6	56945.59
NI	2005	2	152152.7	95206.62	56240.92

4.0.4 fixed_stations.csv

Some stations have fixed generation in MW. These are listed in this file as shown. This generation (when scaled by $T(b,t)$) is subtracted from demand in the corresponding node.

STATION	NODE	YEAR	WEEK	peak	shoulder	offpeak
Aniwhenua	NI	all	all	15	15	15
Patea	NI	all	all	14	14	14
Wheao	NI	all	all	13	13	13
Highbank	SI	all	all	11	11	11
Waipori	SI	all	all	22	22	22

4.0.5 hours_per_block.csv

In each week t , each demand block b has a certain number of hours $T(b,t)$ (adding up to 168). These are listed in this file:

% Hours per block:				
YEAR	WEEK	peak	shoulder	offpeak
2005	1	43	74	51
2005	2	69	54	45

4.0.6 hydro_arcs.csv

This file gives river reaches that are used for transferring water and do not have stations on them. Minimum flow rates and maximum flow rates in cumecs are given.

ORIG	DEST	MIN_FLOW	MAX_FLOW
Lake_Wanaka	Lake_Dunstan	na	na
Lake_Hawea	Lake_Dunstan	na	na
Clyde_tail	Lake_Roxburgh	50	1000
Roxburgh_tail	SEA	250	850
Lake_Taupo	Lake_Aratiatia	na	na
Karapiro_tail	SEA	148	na
Lake_Tekapo	Lake_Scott	na	850
Lake_Scott	Tekapo_canal_head	na	na
Lake_Scott	Lake_Benmore	na	850
Tekapo_canal_head	Tekapo_canal_tail	na	131
Lake_Pukaki	Lake_Benmore	na	3400
Lake_Pukaki	Pukaki_Ohau_canal_junction	na	560
Lake_Ohau	Pukaki_Ohau_canal_junction	na	200
Lake_Ohau	Lake_Ruataniwha	8	560
Lake_Ruataniwha	Lake_Benmore	na	1740
Ohau_BC_canal	Lake_Benmore	na	na
Waitaki_tail	SEA	150	na

4.0.7 hydro_junctions.csv

This file gives a list of names of small lakes and river junctions, excluding the main reservoirs. A sample of these is given below.

Lake_Aratiatia
Lake_Ohakuri
Lake_Atiamuri
Lake_Whakamaru
Lake_Maraetai
Lake_Waipapa
Lake_Arapuni

4.0.8 hydro_stations.csv

This file describes the hydro stations and how they are linked together by arcs in the river chains. Each station has a capacity in MW and a specific power parameter that defines the conversion rate of cumecs to MW. Bounds can also be specified to limit the spill.

GENERATOR	HEAD_WATER_FROM	TAIL_WATER_TO	POWER_SYSTEM	CAPACITY	SPECIFIC_POWER	SPILLWAY_MAX_FLOW
Arapuni	Lake_Arapuni	Lake_Karapiro	NI	196.7	0.439847649	na
Aratiatia	Lake_Aratiatia	Lake_Ohakuri	NI	78	0.273437947	na
Atiamuri	Lake_Atiamuri	Lake_Whakamaru	NI	84	0.206469423	na
Karapiro	Lake_Karapiro	Karapiro_tail	NI	100	0.266255522	na
Maraetai	Lake_Maraetai	Lake_Waipapa	NI	360	0.505804564	na
Ohakuri	Lake_Ohakuri	Lake_Atiamuri	NI	112	0.281025016	na
Waipapa	Lake_Waipapa	Lake_Arapuni	NI	55	0.144558945	na
Whakamaru	Lake_Whakamaru	Lake_Maraetai	NI	100	0.316127853	na
Aviemore	Lake_Aviemore	Lake_Waitaki	SI	230.2	0.327542644	5400

4.0.9 inflows.csv

This file describes the inflows to the various points of the system. Inflows are defined in cumecs for 27 locations in the system. Each location is either a hydro junction as above or is a storage reservoir. Inflows are defined as a rate of inflow in cumecs assumed constant over each week. These data are taken directly from the inflow sequences provided in the CDS. Currently the inflow files provided with DOASA run from 1970 to 2007, but these can be updated with new information as it becomes available from the CDS⁶. The file structure for inflows is as below.

CATCHMENT		Lake_Aviemore	Lake_Benmore	Lake_Hawea	Lakes_Manapouri
INFLOW_REGION		SI	SI	SI	SI
YEAR	WEEK				
1970	1	22.6924	56.731	70.98714286	422.4994286
1970	2	16.05834286	40.14585714	61.96285714	274.3777143
1970	3	13.88491429	34.71228571	61.15985714	387.8231429
1970	4	22.24822857	55.62057143	136.2862857	356.9377143
2007	49	13.79262857	34.48157143	40.37628571	139.0532857
2007	50	19.99554286	49.98885714	105.566	471.1974286
2007	51	24.15171429	60.37928571	94.63428571	250.3771429
2007	52	16.2299	40.57475	60.490875	403.68275

4.0.10 lost_load.csv

This file gives the amount of load that can be shed in each node by sector and load segment, as well as a matching value of lost load. There are three sectors industrial, commercial, and residential. In each load segment in each node the sector makes up a certain proportion of demand. For example in the HAY region we have assumed that industrial load is 34%, commercial load is 27% and residential load is 39% of the load amounts in each load block (see demand.csv). The cost of reducing load of each type is defined by the COST column. Thus the cost of reducing commercial load in HAY is \$2000/MWh for the first 5%, then \$4000/MWh for the next 5% then \$10000/MWh for the rest of the commercial load in this region. So if in some week the regional HAY load was 1000MW in

⁶As of April 2012, inflow data in the CDS are provided up to June 2010.

offpeak, 2000MW in shoulder, and 3000 MW in peak, we would have 270MW commercial load in offpeak, 540 MW in shoulder, and 810 MW in peak. We could then shed 13.5MW in offpeak, 27 MW in shoulder, and 40.5MW in peak times all at a system cost of \$2000/MWh. The next 13.5MW in offpeak, 27 MW in shoulder, and 40.5MW in peak times of commercial load to shed costs \$4000/MWh, and shedding any more costs \$10000/MWh.

The load shedding model is intended to reduce energy consumption in dry periods rather than responding to excess demand in peak periods. This is why the cost of shedding is the same across the different load segments. The full table is given below.

NODE	ISLAND	SECTOR	SEGMENT	PROPORTION	BOUND	COST
NI	NI	industrial	low	0.34	0.05	1000
NI	NI	industrial	medium	0.34	0.05	2000
NI	NI	industrial	high	0.34	0.9	10000
NI	NI	commercial	low	0.27	0.05	2000
NI	NI	commercial	medium	0.27	0.05	4000
NI	NI	commercial	high	0.27	0.9	10000
NI	NI	residential	low	0.39	0.05	2000
NI	NI	residential	medium	0.39	0.05	4000
NI	NI	residential	high	0.39	0.9	10000
HAY	NI	industrial	low	0.34	0.05	1000
HAY	NI	industrial	medium	0.34	0.05	2000
HAY	NI	industrial	high	0.34	0.9	10000
HAY	NI	commercial	low	0.27	0.05	2000
HAY	NI	commercial	medium	0.27	0.05	4000
HAY	NI	commercial	high	0.27	0.9	10000
HAY	NI	residential	low	0.39	0.05	2000
HAY	NI	residential	medium	0.39	0.05	4000
HAY	NI	residential	high	0.39	0.9	10000
SI	SI	industrial	low	0.58	0.05	1000
SI	SI	industrial	medium	0.58	0.05	2000
SI	SI	industrial	high	0.58	0.9	10000
SI	SI	commercial	low	0.15	0.05	2000
SI	SI	commercial	medium	0.15	0.05	4000
SI	SI	commercial	high	0.15	0.9	10000
SI	SI	residential	low	0.27	0.05	2000
SI	SI	residential	medium	0.27	0.05	4000
SI	SI	residential	high	0.27	0.9	10000

4.0.11 reservoirs.csv

This file defines the seven reservoirs used in the DOASA representation of the New Zealand system. The capacity and initial state are measured in cubic metres, and represent the capacity of the operating range, and the initial volume above minimum operating level.

RESERVOIR	INFLOW_REGION	CAPACITY	INI_STATE
Lake_Benmore	SI	423451076	394155969.1
Lake_Hawea	SI	1378764328	874260828
Lakes_Manapouri_Te_Anau	SI	1501878016	393126882.9
Lake_Ohau	SI	57245218.56	23125453.79
Lake_Pukaki	SI	2425440000	570043000
Lake_Taupo	NI	848624230	692149298.4
Lake_Tekapo	SI	823190000	351950000

4.0.12 station_outages.csv

DOASA models outages for planned maintenance (as well as possibly random shutdowns) as known reductions in capacity in MW as shown in the table `station_outages.csv`.

YEAR	WEEK	Stratford_220KV	Huntly_e3p	Huntly_main_g1	Huntly_main_g2	Huntly_main_g3	Huntly_main_g4	Huntly_peaker
2005	1	0	0	0	0	0	265	0
2005	2	0	0	0	0	0	143.8452381	0
2005	3	0	0	0	0	57.26190476	0	0
2005	4	0	0	0	0	0	0	3.273809524
2005	5	0	0	64.26785714	0	0	0	0

4.0.13 terminal_water_value.csv

In the absence of a terminal water value, DOASA would leave all reservoirs empty or nearly empty in the final week of the plan year. To avoid this we can provide a national water value surface to be provided to the total national stored energy (in MWh). This is expressed as a marginal water value surface as shown below.

STORED_ENERGY	VALUE
1000	137.2183986
1500	85.97183215
2000	55.85992946
2500	33.68992837
3500	18.99872191

Here, the first 1000MWh of water stored in all reservoirs is worth \$137 per MWh, the next 500 MWh is worth \$86/MWh and so on. Observe that water values for each reservoir are output by DOASA, but there are no "WaterValues" for the final week because that week's values are exogenously set in the table shown above.

4.0.14 thermal_fuel_costs.csv

Data for thermal stations can be obtained from various sources⁷. Thermal fuel costs have been discussed above. These are specified in a table as shown below.

⁷These data are available from [3, 2, 5]

YEAR	WEEK	coal	diesel	gas
2005	1	4	18.68	4.49
2005	2	4	18.68	4.49
2005	3	4	18.68	4.49
2005	4	4	18.68	4.49

4.0.15 thermal_stations.csv

Data for thermal stations (apart from fuel cost) are defined in the file `thermal_stations.csv`. This file defines the fuel used, the heat rate of the station and its capacity in MW. Commissioning and decommissioning dates can also be included if these are material.

GENERATOR	NODE	FUEL	HEAT_RATE	CAPACITY	START_YEAR	START_WEEK	END_YEAR	END_WEEK
Stratford_220KV	NI	gas	7.3	387	0	0	0	0
Huntly_e3p	NI	gas	6.8	430	2007	23	0	0
Huntly_main_g1	NI	coal	10.5	260	0	0	0	0
Huntly_main_g2	NI	coal	10.5	260	0	0	0	0
Huntly_main_g3	NI	coal	10.5	260	0	0	0	0
Huntly_main_g4	NI	coal	10.5	265	0	0	0	0
Huntly_peaker	NI	gas	9.5	50	2004	23	0	0
NewPlymouth_110KV_g1	NI	gas	11	120	0	0	2007	38
NewPlymouth_110KV_g2	NI	gas	11	120	0	0	2007	38
NewPlymouth_220KV_g3	NI	gas	11	120	0	0	2007	38
Otahuhu_B	NI	gas	7.05	396	0	0	0	0
NI_Whirinaki_220KV	NI	diesel	11	159	2004	22	0	0

4.0.16 transmission.csv

This table shows the transmission lines and their capacity in MW. DOASA is configured to solve with three nodes connected in a line as shown by the table below. It is possible to increase the capacity of these lines but the user cannot alter the connectivity. There is no loss in transmission. Power is assumed to be transmitted at a constant rate in each demand block.

FROM_NODE	TO_NODE	CAPACITY
NI	HAY	1000
HAY	NI	1000
HAY	SI	1040
SI	HAY	1040

4.1 InitialCuts

DOASA produces a policy defined by cutting planes. Since good policies typically require a large number of these (in the order of thousands) it makes sense to be able to use cuts computed in previous runs. Cuts from any run are saved in a set of csv files in a directory specified by the user (defaulting to `Output/<run`

`name>/Cuts`). To warm start DOASA using these cuts, the user moves the saved cuts from (`Output/<run name>/Cuts`) to `Input/InitialCuts`. At present this has to be done manually, but it may be possible to do it more elegantly via the EMI interface. Given a set of initial cuts DOASA will augment these with additional cuts. It is important to note that the overall problem being solved must be the same. Cuts valid for one problem can not be used to warm start another problem if the data (such as inflow sequences used or lost-load penalties) has changed.

The main use of the Initial Cuts feature is to enable simulation of the policy computed. Although DOASA can do both cut generation and simulation within a single run, it would be more usual for a run to be one or the other: either generating cuts to be used in later simulations, or running a simulation based on previously saved cuts. This enables the user to perform out-of-sample simulation (e.g. with completely new inflow samples) or simulation focusing on a particular plan year.

5 Output files for DOASA

The output from DOASA is saved in a directory that is specified by the user in the `run.csv` file. Without this, the directory defaults to `Output/<run-name>`. In the output directory there are six files and four possible directories. The six files are:

convergence.csv This file gives information relating to the convergence of DOASA. In each iteration DOASA computes a lower bound on the total expected cost (including all penalty costs) of an optimal policy. This lower bound is non-decreasing and usually improves strictly at every iteration. The algorithm can be terminated when the lower bound has not changed for several iterations. A more precise convergence test can be carried out by carrying out a Monte-Carlo simulation of a candidate policy. The estimated expected cost of this policy (and its standard error) gives a confidence interval for the true expected cost of the policy that can be used to see if this is close to the lower bound (see section 6 below).

profile.txt The profile file displays the computation time of a run and where DOASA was spending most of its time.

doasalog.txt The log file `doasalog.txt` displays the progress of forward and backward passes that are typically directed to the console.

SampledScenarios.csv The sampled scenarios give the scenarios created in the current run of DOASA. These scenarios are sequences of years sampled with equal probability from Sample start year to Sample end year as specified in the run file. DOASA generates a set of 200 such scenarios as a default. Each forward pass of DOASA then selects from these sequentially. If more than 200 iterations are required, then the code generates another set of scenarios.

MarkovStates.csv This output is relevant only to an enhanced version of DOASA that uses a Markov chain to represent stagewise dependence in inflows.

Transition Matrix.csv This output is relevant only to an enhanced version of DOASA that uses a Markov chain to represent stagewise dependence in inflows.

The output directories are as follows:

5.0.1 Archive

This contains a copy of all the input data for the run which allows it to be replicated if necessary. It also contains a file called **inflows-adjusted.csv** which contains the inflow values produced by the DIA procedure.

5.0.2 Cuts

This directory contains a set of csv files, one for each stage of the plan year. Each file (e.g. **BendersCuts_3_1.csv** for the third week) contains a row for each cut added. The first column in this row is the intercept α , and the next seven columns give the state coefficients of the cut (β). The last column should always be zero.

6552749.6	0.007448	0.003427	0.008188	0.014843	0.014843	0.012276	0.011671		0
5040392.3	0.00462	0.000896	0.006593	0.010153	0.010153	0.013425	0.00462		0
25656581	0.013261	0.029253	0.01443	0.026429	0.026429	0.019383	0.041161		0
35598981	0.015625	0.012009	0.016987	0.031139	0.031139	0.039294	0.048467		0
27586894	0.009214	0.006552	0.011933	0.018748	0.018808	0.013279	0.019887		0
195359274	0.013581	0.010818	0.000344	0.026808	0.027067	0.125419	0.042953		0
266349256	0.015227	0.067571	0.016564	0.030057	0.030348	0.010679	0.031433		0
268165491	0.011667	0.00572	0.000801	0.022843	0.024634	0.013659	0.038424		0
273648947	0.012338	0.008512	0.004521	0.024293	0.024634	0.015849	0.038424		0
277062737	0.012533	0.007683	0.010365	0.02471	0.024993	0.013531	0.038956		0
273660045	0.012353	0.008049	0.003477	0.024215	0.024715	0.015319	0.038449		0

Example of a cuts file **BendersCuts_3_1.csv**. The first column is the intercept α and the remaining columns define the values of β for each storage reservoir.

The state coefficients of the cuts play the role of marginal water values (in $\$/\text{m}^3$) in a particular reservoir when the cut is binding. For example, suppose the reservoir levels are defined by their initial state defined in the following table.

RESERVOIR	INFLOW_REGION	CAPACITY	INI_STATE
Lake_Benmore	SI	423451076	394155969.1
Lake_Hawea	SI	1378764328	874260828
Lakes_Manapouri_Te_Anau	SI	1501878016	393126882.9
Lake_Ohau	SI	57245218.56	23125453.79
Lake_Pukaki	SI	2425440000	570043000
Lake_Taupo	NI	848624230	692149298.4
Lake_Tekapo	SI	823190000	351950000

Then supposing that all reservoirs have their initial state at the end of week three, the estimate of future cost from the first 11 cuts is defined by the tenth cut with intercept 277062737. If we fix all reservoirs apart from Benmore at their initial levels then the cost to go as a function of Benmore storage is locally of the form shown in Figure 3. The slope of this line is $-\$0.012533/\text{m}^3$ as predicted by the value of β in the table above.

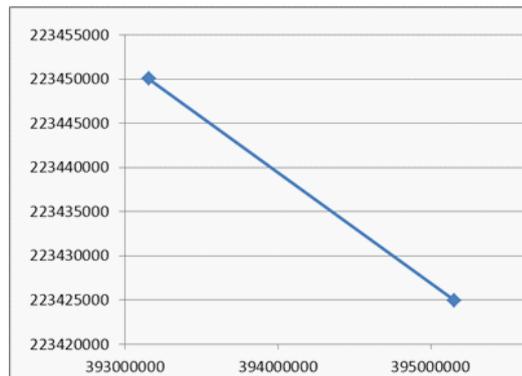


Figure 3: Cost-to go function for Benmore assuming that all reservoirs are at initial levels. The slope of the blue line is $-\$0.012533/\text{m}^3$

5.0.3 Water Values

Corresponding to each Benders cut file there is a file giving an estimate of national marginal water value in $\$/\text{MWh}$. This is computed using different levels of total stored energy in the New Zealand system (measured in GWh), divided amongst the storage reservoirs proportional to their capacity. To compute the energy in GWh, DOASA takes the specific power of each hydro station provided in `hydro_stations.csv` and converts this to a specific energy (MJ/m^3) of water at each point in the hydro system. Using these factors, the stored volumes in each reservoir (in m^3 above minimum levels) can be converted into stored energy (in GWh). The file below as an example shows that at the end of the week in

question, the marginal value of the first 58.83 GWh is \$1413.46/MWh, the next (517.29 - 58.83) GWh are each worth \$46.24/MWh, and so on.

Stored_energy	Water_value
58.83	1413.46
517.29	46.24
853.4	43.93
1344.24	42.02
1940.9	41.69
2434.04	41.15
3201.34	40.53
3978.6	39.55
4413.27	30.98

Example of Marginal Water Value file. The total energy is measured in GWh and the marginal water values are measured in \$/MWh.

5.0.4 Simulation

The simulation directory contains information obtained by simulating the policy. Simulation can be done using Monte Carlo simulation in which the inflows in each week are randomly sampled from historical inflows experienced in that week, or using actual historical inflow sequences. In the first case the policy is simulated using inflows that are assumed to be stagewise independent. This is useful for testing how close the policy is to an optimal policy under this assumption. The second case is useful for testing how the policy would perform in any historical year of inflows. The historical inflows have a higher degree of persistence than the Monte Carlo samples and so they give a better estimate of the expected thermal cost of the policy.

The simulation directory contains output files of various types. The rows correspond to scenarios (simulation realizations) and the columns to weeks. We will describe each in turn.

FlowLBCost.csv Cost (\$) from flows going below their lower bound in each week in each scenario.

FlowUBCost.csv Cost (\$) from flows going above their upper bound in each week in each scenario.

FutureCost.csv Expected future cost (\$) in each scenario at the end of each week.

LostLoadCost.csv Cost (\$) from lost load in each scenario at the end of each week.

PresentCost.csv The total cost (\$) in each stage in each scenario. This is the sum of FlowLBCost, FlowUBCost, LostLoadCost and ThermalCost.

SpilledEnergy_offpeak.csv Spilled energy in MWh in offpeak periods in each week in each scenario.

SpilledEnergy_peak.csv Spilled energy in MWh in peak periods in each week in each scenario.

SpilledEnergy_shoulder.csv Spilled energy in MWh in shoulder periods in each week in each scenario.

StoredEnergy.csv Stored energy in MWh in all reservoirs at the end of each week in each scenario.

ThermalCost.csv Thermal fuel cost (\$) in each stage in each scenario.

TotalCost.csv Total fuel cost over all stages plus expected future cost (\$) in each scenario.

Flow The flow files give flows in cumecs in every hydro arc of the system. Some of these correspond to flow through station turbines and some correspond to spill. The first row is the average over all simulations, and the remaining rows are individual realizations. The columns correspond to weeks.

Volume The volume files give volumes in cubic metres above minimum datum points for each of the seven storage reservoirs. The first row is the average over all simulations, and the remaining rows are individual realizations. The columns correspond to weeks.

6 Outer approximation algorithms

In this section we review the theory of outer approximation algorithms for solving multi-stage stochastic programs with stage-wise independence. This section is included for completeness and can be skipped. These methods are all variations of the Stochastic Dual Dynamic Programming (SDDP) algorithm proposed by [6] as a solution strategy for multi-stage stochastic linear programming. This algorithm can be proved to converge with probability 1 to an optimal policy ([4],[7]).

The class of problems we consider have T stages, denoted $t = 1, 2, \dots, T$, in each of which a random right-hand-side vector $b_t(\omega_t) \in \mathbb{R}^m$ has a finite number of realizations defined by $\omega_t \in \Omega_t$. We assume that the outcomes ω_t are stage-wise independent, and that Ω_1 is a singleton, so the first-stage problem is

$$\begin{aligned} z = \min \quad & c_1^\top x_1 + \mathbb{E}[Q_2(x_1, \omega_2)] \\ \text{s.t.} \quad & A_1 x_1 = b_1, \\ & x_1 \geq 0, \end{aligned} \tag{2}$$

where $x_1 \in \mathbb{R}^n$ is the first stage decision and $c_1 \in \mathbb{R}^n$ a cost vector, A_1 is a $m \times n$ matrix, and $b_1 \in \mathbb{R}^m$.

We denote by $Q_2(x_1, \omega_2)$ the second stage costs associated with decision x_1 and realization $\omega_2 \in \Omega_2$. The problem to be solved in the second and later stages t , given state x_{t-1} and realization ω_t , can be written as

$$\begin{aligned} Q_t(x_{t-1}, \omega_t) = \min \quad & c_t^\top x_t + \mathbb{E}[Q_{t+1}(x_t, \omega_{t+1})] \\ \text{s.t.} \quad & A_t x_t = b_t(\omega_t) - E_t x_{t-1}, \quad [\pi_t(\omega_t)] \\ & x_t \geq 0, \end{aligned} \tag{3}$$

where $x_t \in \mathbb{R}^n$ is the decision in stage t , c_t its cost, and A_t and E_t denote $m \times n$ matrices. Here $\pi_t(\omega_t)$ denotes the dual variables of the constraints. In the last stage we assume either that $\mathbb{E}[Q_{T+1}(x_T, \omega_{T+1})] = 0$, or that there is a convex polyhedral function that defines the expected future cost after stage T .

Outer approximation algorithms build a policy that is defined at stage t by a polyhedral outer approximation of $\mathbb{E}[Q_{t+1}(x_t, \omega_{t+1})]$. This approximation is constructed using cutting planes called Benders cuts, or just *cuts*. In other words in each t th-stage problem, $\mathbb{E}[Q_{t+1}(x_t, \omega_{t+1})]$ is replaced by the variable θ_{t+1} which is constrained by the set of linear inequalities

$$\theta_{t+1} + \bar{\pi}_{t+1,k}^\top E_{t+1} x_t \geq \bar{g}_{t+1,k} \quad \text{for } k = 1, 2, \dots, K, \tag{4}$$

where K is the number of cuts. Here $\bar{\pi}_{t+1,k} = \mathbb{E}[\pi_{t+1}(\omega_{t+1})]$, which defines the gradient $-\bar{\pi}_{t+1,k}^\top E_{t+1}$ and the intercept $\bar{g}_{t+1,k}$ for cut k in stage t , where

$$\bar{g}_{t+1,k} = \mathbb{E}[Q_{t+1}(x_t^k, \omega_{t+1})] + \bar{\pi}_{t+1,k}^\top E_{t+1} x_t^k.$$

The cuts are computed in a sequence of major iterations each consisting of a *forward pass* and a *backward pass*. In each forward pass, a set of N scenarios is sampled from the scenario tree and decisions are taken for each node of those N scenarios, starting in the first stage and moving forward up to the last stage. Different algorithms use different values of N . In DOASA the default is $N = 1$, whereas SDDP typically uses $N = 200$. In each stage, the observed values of the state variables x_t , and the costs of each node in all scenarios are saved.

At the end of the forward pass, a convergence criterion is tested, and if it is satisfied then the algorithm is stopped, otherwise it starts the backward pass, which is defined below. In the standard version of SDDP, the convergence test is satisfied when z , the lower bound on the expected cost at the first stage (called the *Lower Bound*), is statistically close to an estimate of the expected total operation cost (called the *Upper Bound*) obtained by averaging the cost of the policy defined by the cuts when applied to the N sampled scenarios. In this simulation the total operation cost for each scenario is the sum of the present cost ($c_t^\top x_t$) over all stages t . In DOASA we choose $N = 1$. This is more effective in producing good policies with a small number of iterations. The disadvantage is that the algorithm cannot easily compute an upper bound as it proceeds. For this reason we terminate DOASA after a fixed number of iterations that are chosen by the user a priori. Upon termination, an estimate of the expected cost of the candidate policy can then be computed by simulation.

If the convergence criterion is not satisfied (or in the case of DOASA we have not reached the iteration limit), a backward pass is carried out. This adds a cut to each stage problem, starting at the last stage and working backwards to the first. In each stage t we solve the next stage problems for all possible realizations (Ω_{t+1}). The values of the objective functions and dual variables at optimality are averaged over all realizations to define a cut that is added to all problems at stage t . In summary, outer approximation algorithms perform the following three steps repeatedly until the convergence criterion is satisfied.

1. Forward Pass

For $t = 1$ solve (2) and save x_1 and z ;

For $t = 2, \dots, T$ and $s = 1, \dots, N$,

Solve (3), where ω_t is defined by s , and save $x_t(s)$ and $Q_t(x_{t-1}, \omega_t)$.

2. Convergence Test (for SDDP at 90% confidence level)

Calculate the Upper Bound: $z_u = c_1^\top x_1 + \frac{1}{N} \sum_{s=1}^N \sum_{t=2}^T c_t^\top x_t(s)$

$$\sigma_u = \sqrt{\frac{1}{N} \sum_{s=1}^N \left(c_1^\top x_1 + \sum_{t=2}^T c_t^\top x_t(s) \right)^2 - z_u^2}.$$

Calculate the Lower Bound: $z_l = z$;

Stop if

$$z_u - \frac{1.96}{\sqrt{N}} \sigma_u < z_l < z_u + \frac{1.96}{\sqrt{N}} \sigma_u,$$

otherwise go to the Backward Pass.

3. Backward Pass

For $t = T, \dots, 2$, and $s = 1, \dots, N$,

For $\omega_t \in \Omega_t$, solve (3) using $x_{t-1}(s)$ and save $\pi_t(\omega_t)$ and $Q_t(x_{t-1}, \omega_t)$;

Calculate a cut (4) and add it to all nodes in stage $t - 1$.

7 Estimating marginal water values

We expect that one of the main uses of DOASA will be to estimate marginal water values. We digress briefly here to discuss the estimates that DOASA provides.

7.1 Independent inflows

We discuss first the case where the inflows are stagewise independent. DOASA terminates after a fixed set of iterations with a policy defined by cutting planes and a lower bound on the total expected cost of the optimal policy that assumes the independent inflows used in computing the cutting planes. A typical cutting plane has the form

$$\theta_{t+1} \geq \alpha_{t+1} - \beta_{t+1}^\top x_t,$$

where θ_{t+1} is an estimate of the optimal expected future cost $\mathbb{E}[Q_{t+1}(x_t, \omega_{t+1})]$ under the independence assumption we have made. If for some \bar{x}_t ,

$$\mathbb{E}[Q_{t+1}(\bar{x}_t, \omega_{t+1})] = \alpha_{t+1} - \beta_{t+1}^\top \bar{x}_t \tag{5}$$

then the vector β_{t+1}^\top defines the expected marginal value of water in each reservoir. If the volume of water \bar{x}_{ti} in reservoir i increases by one m^3 then the expected future cost $\mathbb{E}[Q_{t+1}(x_t, \omega_{t+1})]$ will decrease by $\beta_{t+1,i}$.

Using the correspondence between cuts and marginal water values to estimate marginal water values relies on two assumptions.

Assumption 1: We can compute $\mathbb{E}[Q_{t+1}(\bar{x}_t, \omega_{t+1})]$ exactly.

Assumption 2: The cuts that we compute at a point \bar{x}_t are exact at this point, i.e. (5) holds.

Neither of these assumptions is true for DOASA. In the case of Assumption 1, at any stage t and state \bar{x}_t , DOASA has at its disposal only a lower bound on $\mathbb{E}[Q_{t+1}(\bar{x}_t, \omega_{t+1})]$ under the independence assumption we have adopted. Recall also that $\mathbb{E}[Q_{t+1}(\bar{x}_t, \omega_{t+1})]$ is the *optimal* expected future cost under the independence assumption, that DOASA will never compute exactly as it converges to this value only in the limit.

Assumption 2 is only valid at stage $T - 1$. At earlier stages, the interpretation of the cut slopes as marginal water values needs some care. The cuts define

an approximate expected future cost function that we use to define a policy. The expected cost of operating this policy can be estimated by simulation with independent inflow sequences. The marginal water values can then be estimated from the results of different simulations (with different initial water values). Bear in mind that marginal water values are determined by optimal policies, and so these simulations will deliver correct marginal water values only to the extent that DOASA has delivered the optimal policy.

In summary, we can estimate marginal water values in two ways. The first method, implemented in DOASA, is to determine at x_t the slope of the cut that yields $\max_k \{\alpha_{t+1,k} - \beta_{t+1,k}^\top x_t\}$. The second estimate is obtained by simulating the policy obtained at termination of DOASA, for different initial storage levels, and estimating the changes in the sample average cost. Formally, for simulations $\nu = 1, 2, \dots, N$, we compute actions x_t^ν and the sample average cost

$$C(0) = \sum_{\nu=1}^N [c_1^\top x_1^\nu + c_2^\top x_2^\nu + \dots + c_T^\top x_T^\nu]$$

and then repeat the exercise using common random numbers and incrementing the storage in a reservoir by a small amount δ to give an estimate

$$C(\delta) = \sum_{\nu=1}^N [c_1^\top y_1^\nu + c_2^\top y_2^\nu + \dots + c_T^\top y_T^\nu].$$

We then estimate the expected marginal value of water for the DOASA policy to be

$$\hat{\pi} = \frac{C(0) - C(\delta)}{\delta}.$$

The DOASA policy is an approximation of the optimal policy that has corresponding expected cost $\Phi(0)$ and $\Phi(\delta)$ respectively. The true expected marginal value of water (or its lowest value if it is not unique) is

$$\pi = \lim_{\delta \searrow 0} \frac{\Phi(0) - \Phi(\delta)}{\delta}.$$

Now for some positive ε_0 and ε_δ we have

$$\begin{aligned} \mathbb{E}[C(0)] &= \Phi(0) + \varepsilon_0 \\ \mathbb{E}[C(\delta)] &= \Phi(\delta) + \varepsilon_\delta \end{aligned}$$

so

$$\begin{aligned} \mathbb{E}[\hat{\pi}] &= \frac{\mathbb{E}[C(0) - C(\delta)]}{\delta} \\ &= \frac{\Phi(0) - \Phi(\delta)}{\delta} + \frac{\varepsilon_0 - \varepsilon_\delta}{\delta} \\ &\approx \pi + \frac{\varepsilon_0 - \varepsilon_\delta}{\delta} \end{aligned}$$

In addition, since each $(c_1^\top x_1^\nu + c_2^\top x_2^\nu + \dots + c_T^\top x_T^\nu) - (c_1^\top y_1^\nu + c_2^\top y_2^\nu + \dots + c_T^\top y_T^\nu)$ is a sample from an i.i.d. random variable D say, the variance of the estimator $\hat{\pi}$ is

$$\sigma(\hat{\pi})^2 = \frac{\text{var}(D)}{N\delta^2},$$

which converges to zero as $N \rightarrow \infty$.

So there are two sources of error in the estimate of expected marginal water value by simulation: the optimization error $\frac{\varepsilon_0 - \varepsilon_\delta}{\delta}$ converges to zero as the number of iterations of DOASA (number of cuts) tends to ∞ , and the sampling error converges to zero as the number of simulations N tends to ∞ . Care must be taken in estimating marginal water values by simulation if DOASA has not converged. Then the magnitude of $\frac{\varepsilon_0 - \varepsilon_\delta}{\delta}$ can be much larger than $\sigma(\hat{\pi})$.

7.2 DIA inflows

When the inflows are known to have some stagewise dependent structure we adjust them using DIA and assume that they are independent. The cutting planes obtained under DIA are independent of previous weeks observations, when a truly dependent model should make the policy adapt to inflow history. So we should expect some error if estimating marginal water values from the cutting planes.

We could adopt the second approach and simulate the policy obtained under DIA assuming independent inflows with DIA. The estimator $\hat{\pi}$ described in the previous section can then be applied under the same assumptions. We note that this approach will give an estimate of the expected marginal water value under the assumption of DIA inflows. This estimate is not the same as the value obtained under a (true) stage-wise dependent model. However using DIA inflows makes the approximations of future cost better than those computed with original inflows. Thus the error $\frac{\varepsilon_0 - \varepsilon_\delta}{\delta}$ is likely to be smaller using DIA inflows.

7.3 Historical inflows

Finally, one can attempt to estimate marginal water values by simulating the DOASA policy (obtained either with assumed independent historical inflows, or with DIA inflows) with N historical inflow sequences. Typically N is no more than 80. These inflow sequences can be thought of as independent samples from the true inflow process. One problem with such an approach is that there are relatively few historical sequences that can be used, and to reduce sampling error we require \sqrt{N} to be large. Typically this means that the sampling error in the estimated expected marginal water values are significant. One way of reducing the sampling error is to build a Monte Carlo simulation model that accurately represents the inflow process and then sample synthetic inflow sequences from

such a high-fidelity inflow model. This model does not have to be convex to enable cut calculation, but can be quite nonlinear, as long as it computationally inexpensive enough to produce many synthetic sequences of inflows to reduce the sampling error.

References

- [1] Centralized Data Set for New Zealand Wholesale Electricity Market (downloadable from <http://www.ea.govt.nz/industry/monitoring/cds/>).
- [2] Denne, T., Bond-Smith, S., Hennessy, W., Coal prices in New Zealand markets, COVEC report for New Zealand Ministry of Economic Development, 2009 (downloadable from <http://www.med.govt.nz/upload/68784/coal-price-report.pdf>).
- [3] New Zealand Electricity Commission database (downloadable from <http://www.electricitycommission.govt.nz/opdev/modelling/gpas/May2007/-Generation>).
- [4] Girardeau, P., Philpott A.B. On the convergence of decomposition methods for multi-stage stochastic convex programs, (downloadable from <http://www.epoc.org.nz/publications.html>).
- [5] New Zealand Ministry of Economic Development - Energy Data, June 2009, (downloadable from http://www.med.govt.nz/templates/MultipageDocumentTOC_____21660.aspx).
- [6] Pereira M.V., Pinto L.M., Multi-stage stochastic optimization applied to energy planning, *Mathematical Programming* 52 (1991) 359-375.
- [7] Philpott, A.B., Guan, Z., On the convergence of stochastic dual dynamic programming and related methods, *Operations Research Letters* 36 (2008) 450-455.
- [8] Philpott, A.B., Guan, Z., Online companion for “Production inefficiency of electricity markets with hydro generation”, (downloadable from <http://www.epoc.org.nz/presentations/Modellingsummary.pdf>).
- [9] POCP database of planned outages (downloadable from <http://pocp.redspider.co.nz/>).