Value of information on the operation of a dual-purpose reservoir

THESIS REPORT

Khalif A Jusuf – 4342054 (TUD) – A0118952 (NUS)

Image of Salto Grande Reservoir (Lenilucho, 2010)

Contents

Acknowledgement

I would like to acknowledge the immeasurable help, support and advice provided by my supervisors Dr. Luciano Raso, and Dr. Susan Steele-Dunne

I would also like to acknowledge the help from the people who had provided technical, administrative, and personal support throughout this thesis. They include: Ruby Meilasari, Prof. dr. Nick van de Giesen, Dr. Ronald van Nooyen, ir. Willem M.J. Luxemburg, David Rodriguez Aguilera, Jelmar Schellingerhout, Otto de Kaizer, Jan Talsma, Dave de Koning, Carlos Sanchez, Markus Hrachowitz, Cecilia Shanti Dewi, Jason Chong Woon Siong, and Peter Hill

Lastly but most importantly, I would like to dedicate this thesis to the memory of Dr. ir. Peter-Jules van Overloop (1969 – 2015). I count myself very fortunate to have been his student. My thoughts are still with his beautiful family.

Executive Summary

Reservoirs are often operated to meet two or more objectives that sometimes may conflict. An example of this conflict is to keep storage low enough to maintain spare capacity for flood protection, but also to keep it high enough to maintain reliability of water / hydropower supply. Optimising the control decisions of reservoirs can provide a valuable contribution in effectively meeting those objectives. Those decisions require the procurement and interpretation of information from computer models and hydrological data. Any improvement in the availability or accuracy of these sources of information will naturally improve the quality of decision making in operating the water systems.

This thesis presents a method for quantifying the value of having different levels of information in operating a test case reservoir. The reservoir selected as a case study is Salto Grande Reservoir, located at the border of Uruguay and Argentina which is operated to maximise hydropower generation while minimising flood risk and damage.

Figure E-1 illustrates how the system is represented in this study. It is assumed that:

1) Salto Grande Reservoir can be represented as a water balance model where outflows (through hydroelectric turbine and spillways) are optimised

2) Upstream catchments are represented as a hydrological model (rainfall runoff model with routing)

3) Hydropower output and flood damage are represented as components in the objective function used to optimise the outflows

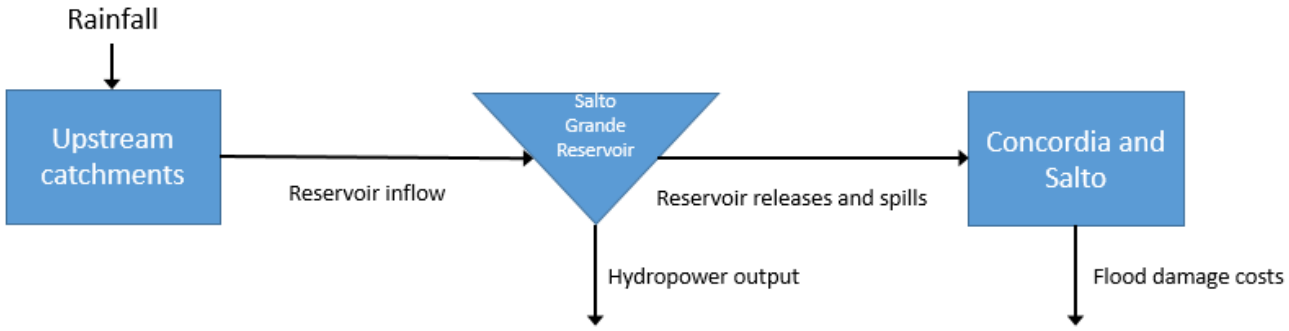


Figure E-1 Simplified physical conceptualisation of Salto Grande Reservoir and surroundings

The control method in this thesis emulates real time control and as such only accounts for short term objectives. Long term objectives such as seasonal water demands are ignored. The control has an optimisation horizon of 15 days, with 6 hour increments.

Information used in the control relates to the disturbance to the system (inflow to the reservoir), the state of the system (reservoir storage level), and controlled action (outflow from the reservoir). Inflow to the system can be observed and can also be forecasted. The availability and accuracy of the inflow forecast forms the basis of the different levels of information considered in this thesis.

The different levels of information considered are:

- Base case: where no forecast is available and only inflow and reservoir level observations are available

- 'Perfect model': where it is assumed that the hydrological model is able to replicate the dynamics of upstream catchments perfectly. Therefore, the only source of inflow forecast error is the rainfall forecast

product used. Two deterministic rainfall forecast products are used as inputs to the model, CPTEC and ECMWF

- 'Imperfect model': where it is assumed that the hydrological model parameterisation and initialisation is not robust. Therefore the hydrological model is also a source of inflow forecast error, in addition to the rainfall forecast product. The same rainfall forecast products are used as inputs

The value of those levels of information is done by using a multistage stochastic programming (MSP) on the control of reservoir outflows implemented using a decision tree. There are two key components in making a decision tree. An ensemble of future inflow trajectories, and a representation of information uncertainty along the horizon. In this thesis inflow trajectory ensembles are generated synthetically using an ARMA (2,1) model. Forecast uncertainty is represented by a covariance matrix of forecast errors. Measurement uncertainty is represented by the variance of measurement error. Measurement and forecast bias are neglected.

The levels of information will influence how uncertainty evolve along the optimisation horizon. Uncertainties constrain the decisions that can be made by a controller. If decisions are constrained, it is likely to be less optimal. Better availability and accuracy of information would result in uncertainty being resolved earlier in the optimisation horizon. The earlier uncertainties can be resolved, the earlier branches in the decision tree can bifurcate from each other, resulting in less constrained decisions.

To illustrate, Figure E-2 presents two decision trees. The one on the left was made using less information (no forecast) which results in a decision tree that is more constrained. The branching points in the left tree occur later than the tree on the right, which is made using more information (with forecast) and therefore less constrained.

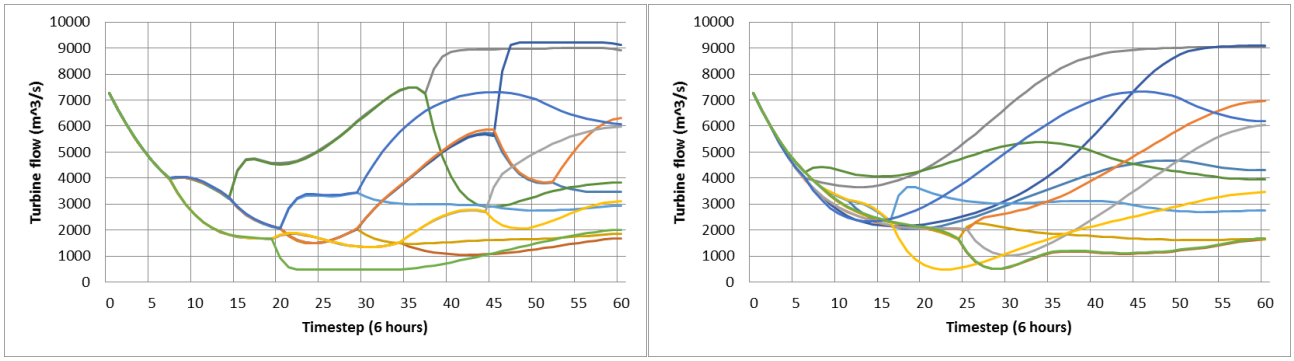


Figure E-2 Example of a more constrained (left) and less constrained (right) decision tree

Considering the above, there are 8 key steps to complete this analysis and quantify the value of the different levels of information:

- Step 1: Obtain weather forecasts and observed data
- Step 2: Run forecasted rainfall in the hydrological model
- Step 3: Verify forecast to quantify the variance of forecast error and produce the covariance matrix
- Step 4: For each observed inflow considered, synthetically produce an ensemble of all possible future trajectories using a synthetic streamflow generator
- Step 5: Reduce the future trajectory ensemble using scenario reduction algorithm
- Step 6: Generate tree for all levels of information and future trajectory ensemble
- Step 7: Utilise tree structure in optimising reservoir operations
- Step 8: Calculate value of information

Once the sequence of tasks above are applied to the test case, the value of information is quantified as the weighted average of the quasi-optimised objective function scores which is shown in Table E-1.

Table E-1 Average optimised objective function score for all scenarios and forecast products

	Base Case	'Perfect model'		'Imperfect model'	
		CPTEC	ECMWF	CPTEC	ECMWF
Objective function score	1371	1145	1145	1162	1162

The following can be concluded from the results above, and more detailed analysis presented in the report:

- **Better information does indeed result in better performance**
    - The 'perfect model' scenarios performs the best, followed by the 'imperfect model' and the base case.
- **'Imperfect model' is still very useful**
    - The performance of the imperfect model scenarios is not much worse than the perfect model, which implies that it is still useful despite the deficiencies in model parameterisation and the setting of initial conditions
- **Performance of the different forecast products are similar**
    - Initial forecast accuracy analysis indicates that the RMSE of ECMWF and CPTEC is similar. Any slight differences appear to be driven by the difference of bias of the two products
    - This operation of this system assumes that bias is removed, and once bias is removed, the variance of errors, the resulting decision tree, and the resulting weighted average score are very similar, if not identical

1. Introduction

1.1 Research objectives

Operational water management relates to the control of water systems such as reservoirs and canals that with the means at hand (i.e. constraints, and availability of controllers), aim to keep all stakeholders (e.g. farmers, residents, businesses) satisfied all the time (van Nooyen & van Overloop, 2014). Such operations must be conducted while the water system is subjected to a range of variability and disturbances such as floods and droughts.

In order to control water systems, water managers procure and interpret information from computer models and hydrological data (MDBA, 2015). These sources of information have costs attached to them. Streamflow and rainfall gauges need to be installed and maintained. Hydrological models need to be designed and calibrated. Weather forecast products may also have costs. The cost of such information can sometimes be quantified using industry price quotes and estimates (Melbourne Water, 2015; ECMWF, 2015; Global Water, 2015). There have also been works that quantify the cost of these sources of information with a more theoretical perspective (Goninon, et al., 1997; Zhu, et al., 2002; Boucher, et al., 2014).

Of particular interest in this thesis are weather and inflow forecasts, which allow water managers to anticipate the dynamics of the water systems that they are managing.

Considering the above, this thesis aims to contribute to understanding the role of information in operational water management by attempting to answer the following research question:

"To what extent do weather forecasts and rainfall runoff model accuracy influence the quality of decision making in reservoir operations"

To limit the complexity of this thesis, the research question above is simplified such that:

- The control of water systems considered in this thesis is limited to **reservoir releases**
- The reservoir considered is a **dual-purpose reservoir**, which must keep water level high enough for hydropower production but low enough to accommodate floods when it occurs.

Further simplifications are outlined in the methodology section of this thesis (Section 3).

To achieve its research objectives, the thesis uses a test case which:

- Has a real control problem: Has an objective, subject to disturbance, and can respond to those disturbances to pursue its objective,

And where:

- Information has costs: For example from installation and maintenance of gauges, or forecast product subscriptions
- Information has value: Better or more information would improve the performance of the control

1.2 Study area Salto Grande

1.2.1 Introduction to system

The test case chosen for this thesis is Salto Grande Reservoir, located in South America, on the Uruguay River which borders Uruguay and Argentina. Table 1 shows the key characteristics of the reservoir. Of particular interest is the size of the upstream catchment, relative to the reservoir storage capacity, which suggests a reservoir with a short hydraulic retention time. This indicates that the reservoir is likely to benefit from anticipatory control actions using forecasts, and refinements to its short term control regimes. Figure 1 is a map showing the location of the reservoir, key gauges, and subcatchment outlines. The cities of Concordia and Salto

are located on the banks of the Uruguay River, very closely downstream of the reservoir. Therefore excessive flows outflows and spills from the reservoir may result in flood damage in these cities.

Table 1 Key characteristics of Salto Grande Reservoir

Characteristics	Value (unit)	References
Minimum reservoir elevation	30.0 m above sea level	(Chaer & Monzon, 2008)
Minimum storage	2,805 GL	(Chaer & Monzon, 2008; Talsma, 2015b)
Maximum reservoir elevation	35.5 m above sea level	(Chaer & Monzon, 2008)
Maximum storage capacity	5,864 GL	(Chaer & Monzon, 2008; Talsma, 2015b)
Minimum reservoir outflow	0 m^3/s	(Talsma, 2015a)
Maximum turbine outflow	Approx. 8,400 m^3/s	(Chaer & Monzon, 2008)
Maximum spillway flow	Approx. 64,000 m^3/s	(ICOLD, n.d.)
Upstream catchment size	224,000 km^2	(ILEC, 1999)

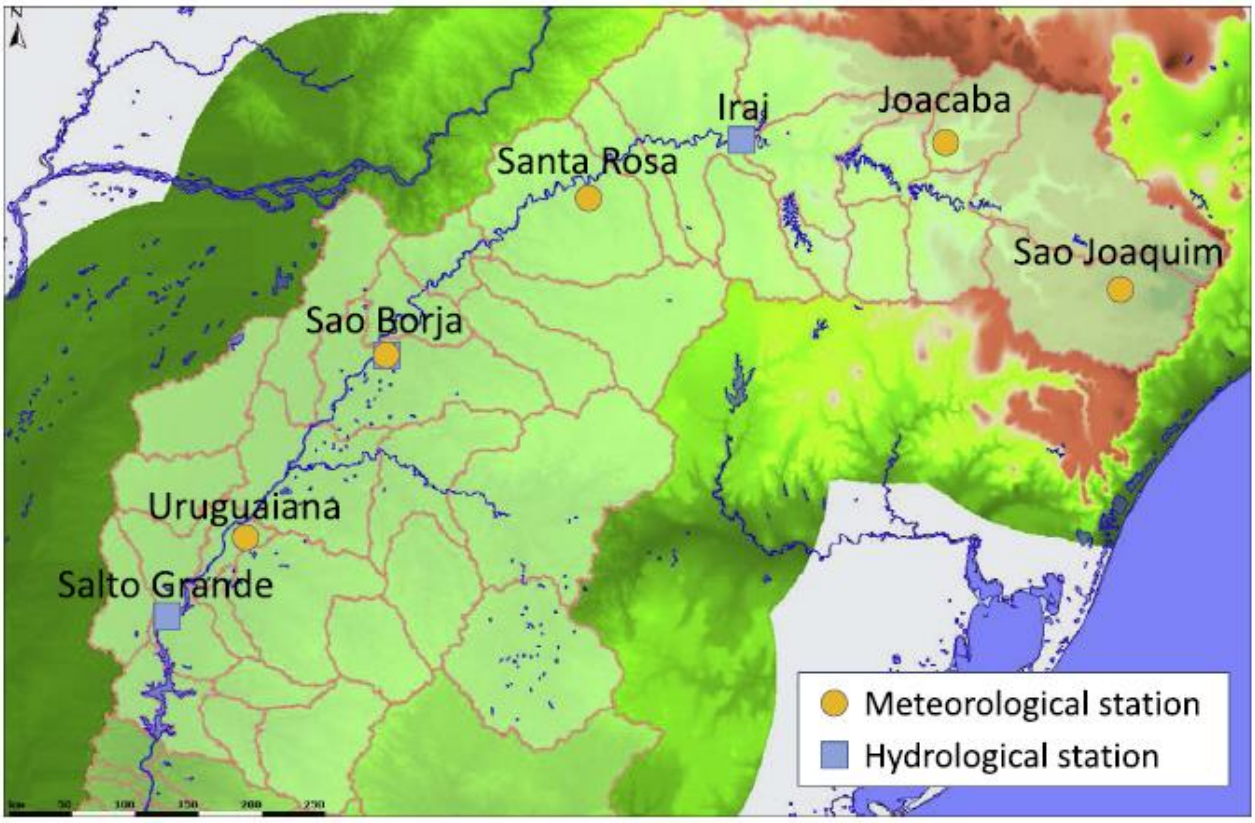


Figure 1 Location of Salto Grande Reservoir along with subcatchment outlines and key meteorological and hydrological stations (Raso, et al., 2014a)

1.2.2 Rationale

Salto Grande Reservoir was chosen because it meets the following qualitative criteria:

- Adequate amount of historic data and previous work particularly from Raso, et al. (2014a)
  - So that the case study in this thesis can be reflective of real world conditions
- Catchment upstream of the dam of control is sufficiently complex (e.g. many subcatchments or tributaries)
  - If the upstream catchment is too simple or small, it would limit the scope of considering different rainfall and streamflow gauges to be incorporated in the reservoir control

- Global retention time in the dam of less than 1 month
    - So that inflow changes has a more considerable impact on dam level trajectories
    - The reservoir should be small enough to benefit from better short term operations, particularly in response to flood risk
    - If the reservoir is too big, it's maximum capacity may be far above its normal operating levels, allowing it to accommodate large floods without releasing any water
- Hydrologically disconnected or sufficiently far from other dams to lessen the influence of the controls in these other dams
    - So that the analysis only deals with the control of the reservoir of interest, assuming the effect of any dams upstream and downstream to be negligible

1.3 Thesis outputs

Considering the research objectives and the test case selected, in the subsequent sections of this thesis, the following is presented:

- A review of operational water management, including the control methods used and timescales involved

- A review of monitoring and forecasting, its role in water systems operations and how to assess its quality and usefulness

Most importantly, the key output of this thesis is:

- A conceptualisation and demonstration of the use of decision trees to determine value of information, which examines three categories of information:

    - Weather forecasts

    - Reservoir inflow monitoring

    - Hydrologic model

The thesis builds on previous work on Tree-Based Model Predictive Control (Raso, et al., 2014a; Raso, 2013), a control method used to optimise operations of water systems, and tree structure generation (Raso, et al., 2013), to quantify the value of different forecast products and hydrologic model quality based on its effect on the water system operations.

2. Background and theory

2.1 Water systems operations

The following sections provide a brief review on how open water systems are operated. It approaches this topic in terms of how it relates to basic control theory, recent advances and approaches in prediction and optimisation, and distinctions between short- and long-term operations.

2.1.1 Basic type of controls

Figure 2 shows how the control of a water system can be structured, in the case of managing a water system to maintain a desired water level. The example illustrated in the figure has a combination of feedback and feedforward control loops.

Feedback control is when the controller looks at the state of the system (e.g. water level of a reservoir or canal), see how much it deviates from the desired state (e.g. a target level) and takes action on the system to correct this deviation (e.g. by releasing or retaining water) (van Overloop, 2006). An example of feedback control is open water systems is the use of PID (Proportional-Integral-Derivative) controllers which can be utilised to maintain a desired water level in canals (Shahverdi & Monem, 2011; Malaterre & Baume, 1998).

Feedforward control is when the controller observes the disturbance directly (e.g. storm event or lateral flows) and models how it will affect the state of the system (e.g. increase canal or reservoir level) (van Nooyen & van Overloop, 2014; van Overloop, 2006). A feedforward does not work well independently because disturbances can be overlooked by the measurement network, and the model representing how the disturbance influence the state of the system may not be accurate (van Nooyen & van Overloop, 2014; Stephens, et al., 2007). Therefore feedforward control is normally used alongside feedback control. This combination, as shown in Figure 2, takes advantage of both the error correction and disturbance anticipation characteristics of feedback and feedforward controls respectively.

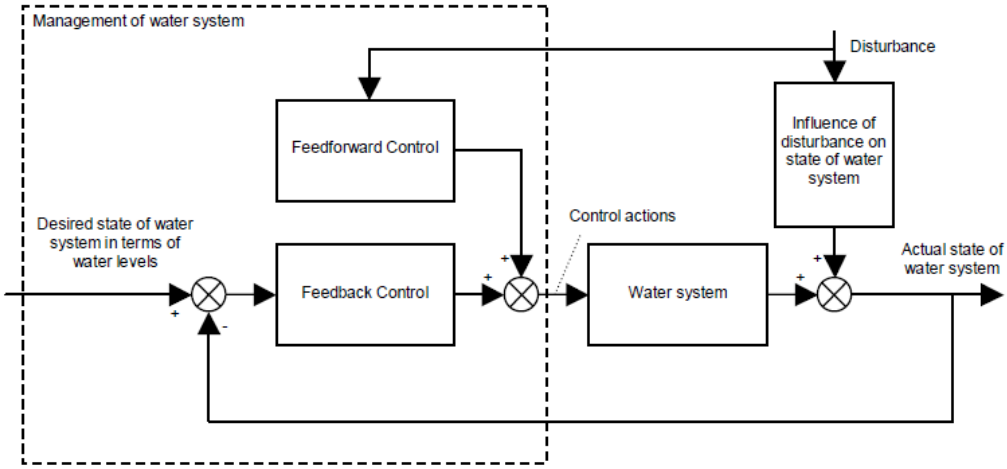


Figure 2 Structure diagram of a controlled water system (van Overloop, 2006)

2.1.2 Long term vs short term operations

Operation of water systems can be done with respect to both short term and long term goals. Long term operations tend to deal with seasonal changes such as irrigation, and crop growth patterns. The use of reservoir rule curves is an example of controlling water systems with respect to long term goals (Kangrang & Lokham, 2013). Short term operations tend to deal with daily or episodic changes in the system, such as emergency

flood events. Real time control methods such as model predictive control is an example of a control method that accounts for short term goals. (Lund, 1996)

The control method in this thesis emulates real time control and as such only accounts for short term objectives.

2.1.3 Model Predictive Control and Optimisation

There has been several publications on the application of Model Predictive Control (MPC) in the context of operating open water systems (van Overloop, 2006; Weijs, 2008; Delgoda, et al., 2012; Kearney, et al., 2011; Negenborn, et al., 2009).

Model predictive control is a real time control method that optimises a control strategy (e.g. reservoir outflows) within specified constraints (e.g. spillway or pumping capacity) in response to a receding horizon of forecasted disturbance (van Overloop, 2006). The first timestep of that control strategy is then adopted as the control action. This process is repeated each time a control action is to be taken. This mechanism is illustrated in Figure 3.

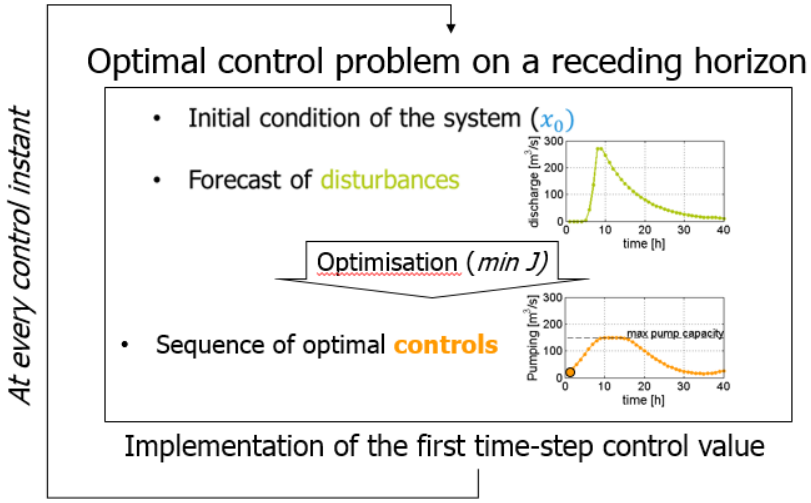


Figure 3 Illustration of the mechanism of Model Predictive Control (Raso & van Overloop, 2011)

The optimisation is usually such that it minimises or maximises a certain objective function. The objective function is usually a quantification of what is considered to be important by the operator. Examples of this include hydropower generation, upstream or downstream flood damage, environmental flow requirements, and recreational aspects.

A generic example of an objective function to be minimised in an MPC is shown in Equation 1. In the example, *N* is the optimisation horizon, *u* is the control action, *x* is the state, and *y* is the dependent variable. *h* is the hard constraint, *f* and *g* are arbitrary functions that govern the state and dependent variables.

Equation 1 Simplified generic MPC objective function (Schwanenberg & Becker, 2014)

$$\min_{u^k} \sum_{k=1}^{N} J(x^k, y^k, u^k)$$

subject to:

$$h(x^k, y^k, u^k, d^k) \leq 0, \qquad k = 1, \dots, N$$

$$x^k - f(x^{k-1}, y^k, u^k, d^k) = 0$$

$$y^k - g(x^k, u^k, d^k) = 0$$

There are numerous optimisation algorithms that can be used. One is linear programming, which allows quick computation but requires linear objective functions and constraints. Non-linear programming allows for non-linear objective function and constraints but tend to be slower to compute. There are also other algorithms such as Monte Carlo and Genetic Algorithms. Clemens (2001) and Luenberger & Ye (2008) provides more detail on the available optimisation algorithms.

The control method emulated in this thesis incorporates multi stochastic optimisation of an objective function. Further details on the optimisation algorithm and the cost function adopted in the thesis are available in Section 3.2.8.

2.1.4 Multistage stochastic optimisation and Tree-based MPC

Note in an MPC as described previously, the forecasted disturbance is deterministic, therefore if it is proven to be inaccurate, this can have significant ramifications to the performance of the system. Multi stage stochastic optimisation (MSO) is an attempt to address this problem by considering an ensemble forecast instead of a single trace forecast. The final objective function is a probability weighted sum of the objective function of each ensemble member or branch. This formulation is shown in Equation 2. *M* is the number of ensemble members and *p* is their respective probabilities.

Equation 2 Multistage stochastic optimisation objective function (Schwanenberg & Becker, 2014)

$$\min_{u^{j,k}} \sum_{j=1}^{M} p^j \left(\sum_{k=1}^{N} J\left(x^{j,k}, y^{j,k}, u^{j,k}\right) \right)$$

Raso, et al. (2013) identified a problem with MSO which is intractibility. The dimension of the problem increases exponentially as the number of timesteps or scenarios increases. This can result in prohibitively long computation times. To address this Raso, et al. (2013) proposes simplifying the ensemble in a form of a scenario tree. Ensemble forecast members tend to have small differences between each other at the early timesteps of the forecast horizon. Afterwards forecast uncertainty increases and the members diverge. A tree is generated from ensemble data by aggregating the scenarios with sufficiently small differences until a point in the horizon where they are deemed to be distinguishable from each other. Equation 3 shows the formulation of an MSO objective function defined using a scenario tree.

Equation 3 Multistage stochastic optimisation objective function defined using a tree (Raso, et al., 2013; Schwanenberg & Becker, 2014)

$$\min_{u^{x,t}} \sum_{x=1}^{M} p^j \left(\sum_{t=1}^{N} J(x^{x,t}, y^{x,t}, u^{x,t}) \right)$$

$$where\ u^{i,t} = u^{j,t}\ when \begin{cases} P(i) = j \\ t < B(i) \end{cases} i, j = 1, \dots, M$$

The formulation in Equation 3 presents a tree structure using a parent branch relationship. Such a relationship defines for a given ensemble member (i), their parent member (P), and the point in time (B) where that member diverges from its parent. Note that the very first branch has itself as its parent and a branching point occurring in the first timestep. Table 2 presents an example parent branch relationship.

Table 2 Example parent branch relationship (Raso, et al., 2014a)

i	1	2	3
P(i)	1	1	2
B(i)	1	3	5

A tree structure can also be presented as a nodal partition matrix (M), where variables are assigned for each member and timestep. In timesteps where member has not diverged from its parent, the variable of the member is the same as its parent, as shown in Equation 4. An example of a nodal partition matrix is shown in Equation 5, which has the same tree structure as the parent branch relationship example in Table 2. Figure 4 provides a visual illustration of the nodal partition matrix example.

Equation 4 Formulation of nodal partition matrix (Raso, et al., 2014a)

$$M(t,i) = M(t,j) when\ t < B(i), P(i) = j$$

$$i, j = 1, \dots, N$$

Equation 5 Example nodal partition matrix (Raso, et al., 2014a)

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 6 & 6 \\ 4 & 7 & 7 \\ 5 & 8 & 9 \end{bmatrix}$$

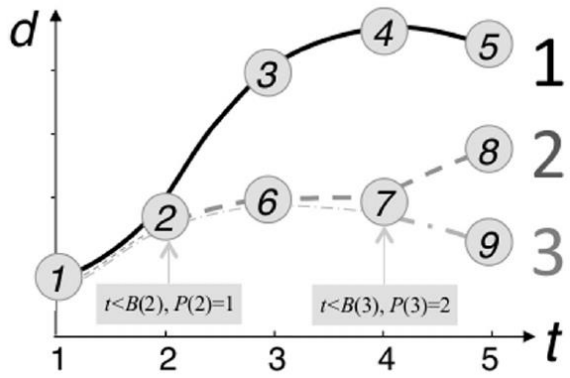


Figure 4 Visual illustration of the example parent branch relationship and nodal partition matrix (Raso, et al., 2014a)

The key feature of a tree structure is that it depends on the level and quality of information available. The better the information, the earlier an ensemble member can branch out from its parent. To illustrate this, consider this example:

- There are only two possible future scenarios, flow (x_1) and no flow (x_2)
- Whether or not there will be flow, is determined by rain, so probability of x_1 is equal to the probability of rainfall, $P(x_1) = P_{rain}$. Conversely, $P(x_2) = 1\text{-}P_{rain}$
- The controller of a storage has to release water from a storage to make room for this flow. Their aim is to keep the storage as full as possible without overtopping
- If the controller only has access to flow observations (as illustrated in Figure 5), they can only be sure that flow will/will not occur at timestep 2.5, resulting in the control action illustrated in Figure 6.

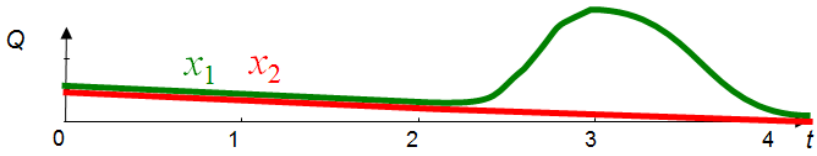


Figure 5 Example inflow observations received by a controller (pers. comm. Raso, May 2015)

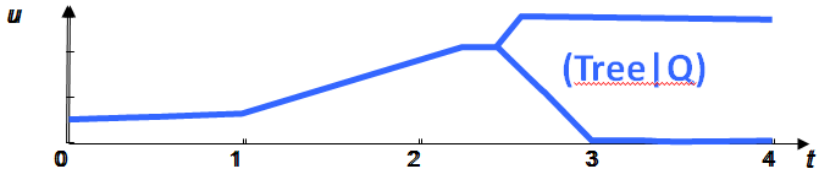


Figure 6 Example control decision given the availability of inflow observations (pers. comm. Raso, May 2015)

- However, if the controller also has access to rainfall observations (as illustrated in Figure 7), they will know that the rainfall that causes the flow event will/will not occur at timestep 1, resulting in the control action illustrated in Figure 8.

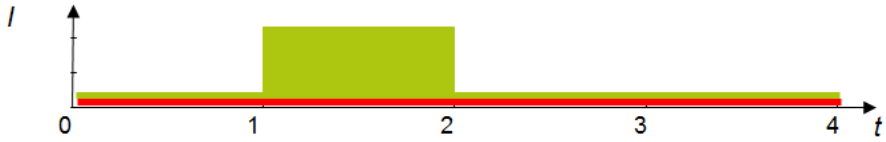


Figure 7 Example rainfall observations received by a controller (pers. comm. Raso, May 2015)

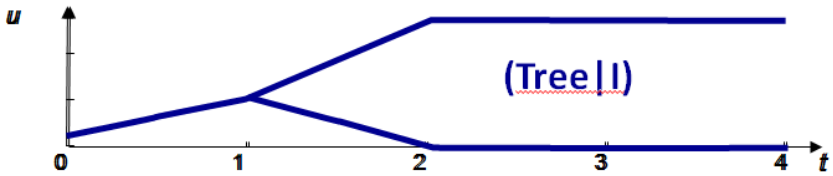


Figure 8 Example control decision given the availability of rainfall observations (pers. comm. Raso, May 2015)

- It is therefore apparent that having rainfall observations in addition to flow observations is more ideal. The controller is able to anticipate the flow/no flow event sooner, as illustrated by the

branching point occurring earlier. This allows more time for the controller to reduce the risk of storage overtopping or excessive drawdown of storage.

This is the premise of the information flow modelling approach developed by Raso, et al. (2013) which is summarised in Section 3.2.7. The types of information used in this thesis are observations and forecasts. Sections 2.2 and 2.3 provide an overview of those information, along with the uncertainties involved.

2.2 Measurements and observations

Measurements and observations are key aspects of operational water management. Literature frequently indicate that knowledge of the current state, disturbance, and the actuated control action are required in the control of water systems (van Overloop, 2006; Litirico & Fromion, 2009; Malaterre & Baume, 1998; Delgoda, et al., 2012).

In the context of this thesis, the state is considered as the depth or volume of the reservoir, disturbance as the inflow to the reservoir, and the control action is the reservoir release. Such information is obtained in practice via a monitoring system, such as a network of flow and depth gauges. Additional streamflow and rainfall gauges, particularly on the upstream catchments of the dam can also provide an indication of inflow into the reservoir.

Measurement uncertainty is something that needs to be accounted for, as they will affect the quality of decision making. In measuring streamflow, sources of uncertainty include (Sauer & Meyer, 1992):

- **Cross section error**: This refers to the error in measuring the depth and width of a water way. This can also refer to the error from the assumption that the vertical depth is representative of the mean depth in a segment of cross section (i.e. water surface is perfectly horizontal). This error influences flow measurement because the cross section of a waterway would indicate the flow in the waterway for each unit of flow velocity

- **Velocity measurement error:** This refers to the errors from flow velocity measurement devices, turbulence, assumptions in vertical and horizontal velocity distribution, and others

- **Computation error:** This refers to errors in calculating the depth and velocity of segments in a cross section to convert them to flow

- **Systematic errors:** This refers to errors coming from improperly calibrated or operated measurement devices. Unlike the other errors mentioned previously, this error is not random

- **Other uncertainties:** Uncertainties caused by weather effects, flow obstruction and other external factors

There have been many studies aimed at analysing and/or quantifying measurement uncertainty. The previous list was summarised from Sauer & Meyer (1992) which also provides methods and standard values that can be useful in quantifying streamflow measurment uncertainty. McMillan, et al. (2012) provides a thorough review and benchmark on the likely sources and magnitudes of uncertainty of measurements useful in water resources management. Di Baldassarre & Montanari (2009) also provides a framework for quantifying streamflow uncertainty with significant focus on rating curve errors. Other studies in this topic include Despax, et al. (2016) and Chen, et al. (2013).

The quantification of measurement uncertainty adopted in this thesis is explained in Section 3.2.1.

2.3 Forecasting

2.3.1 Role in water systems operations

Forecasting allows water managers to approximate how their water systems will behave in the future. This approximation allows them to adopt anticipatory operational decisions which may lead to better outcomes (van Andel, 2009). In an operational water management context, there are several types of forecasts that may be useful, which are listed below. Note that this list is not exhaustive:

- **Weather:** In a modern context. Weather forecasts are outputs of numerical weather models used to predict weather conditions in the future. In the context of water management commonly used weather forecasts are rainfall and temperature predictions. These can be short term, such as the weather predictions available to the public via television or newspapers, or long term such as rainfall and temperature outlooks under climate change (CSIRO, 2012)

- **Streamflow:** This is usually generated from rainfall runoff models, which utilises forecasted rainfall as input to predict streamflow. This can be short term such as the forecasting involved in flood warning services (Bureau of Meteorology, 2013), or long term such as streamflow outlooks under climate change (CSIRO, 2012)

- **Demand:** This is usually generated from models which takes into account demand patterns (e.g. seasonal and diurnal) which were observed in the past, and the influence of weather conditions such as temperature (White, et al., 2003). In the longer term, demand predictions can also take into account population growth and other socioeconomic factors. An example of this are the water supply and demand strategies which must be periodically prepared by urban water corporations in Victoria, Australia (DSE, 2011)

2.3.2 Forecast verification

Despite their potential benefits, water managers can be reluctant in using weather forecasts for short term operations. There are institutional and political reasons for this, but a key reason often cited by water managers is poor reliability of weather forecasts. While this can be due to valid concerns of forecast quality and reliability, there have been indications that water managers frequently don't understand the quality/skills of available weather forecasts or the level of forecast reliability that they would require. (Rayner, et al., 2005)

Therefore, the quality of forecasts is something that needs to be understood. Forecasts can be evaluated based on their accuracy, skill, and utility (Persson & Grazzini, 2005), which are explained further in the following subsections.

2.3.2.1 Accuracy and skill

Forecast accuracy and skill are related. Accuracy refers to the absolute measure in how forecasts can approximate what will happen in the future. Skill is a relative measure which compares the forecast against another reference, such as a different forecast product or method. (Persson & Grazzini, 2005)

To illustrate, Equation 6 shows an accuracy measure where the error of forecast (*f*) in approximating the actual observation (*x*) is measured using Mean Squared Error (MSE) for *n* instances. Equation 7 shows a measure of forecast skill, where a skill score is computed based on a comparison between the MSE of the forecast and the MSE of a reference. (Murphy, 1988)

Equation 6 Mean squared error (MSE)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(f_i - x_i)^2$$

Equation 7 Forecast skill score (SS)

$$SS = 1 - \frac{MSE_{forecast}}{MSE_{reference}}$$

There are other methods in calculating forecast skill and accuracy, and it depends of the nature of the forecast. A comprehensive summary is provided in WWRP/WGNE Joint Working Group on Forecast Verification Research (2015).

Quantification of forecast accuracy adopted in this thesis is summarised in Section 3.2.4. The results of the quantification is presented in Section 4.1.

2.3.2.2 Utility

Another way to assess the quality of a forecast is by assessing how useful it is in improving decision making. In other words, assessing forecast utility instead of just assessing their accuracy.

Such analysis of forecast utility requires two key components. **1)** An understanding of forecast accuracy and skill **2)** An understanding of how the forecast will be used, which requires an emulation of the decision making process involved.

A simple illustration of forecast utility is the cost-loss ratio in Persson & Grazzini (2005). This approach relates the ability of a forecast to estimate a probability of an event to the cost or risk of taking or avoiding preemptive action:

- Take *C* as the cost of taking protective action (e.g. evacuation),
- *L* as the loss or damage when no action is taken, and
- *p* is the probability of flood.
- A forecast that gives too many false alarms would result in unnecessary protective actions when $pL < C$
- whereas a forecast that misses events would result in preventable losses when $pL > C$
- Therefore it is ideal to have a forecast accurate enough in predicting *p* so that it prevents the decision maker in taking unecessary protective action when $pL < C$ and ensures that they do take protective action when $pL > C$

This thesis is ultimately an analysis of forecast utility, using the value of information approach outlined in Section 2.4

2.4 Value of information

Value of information can be expressed as in Equation 8. *I* refers to a situation where there is more information than I_0, in such a way that $I = I_0 + \Delta I$, where ΔI is the additional information made available. In a practical context, 'more information' can refer to better foresight or more accurate measurements.

Equation 8 Value of information from improvement of system performance

$$V(I|I_0) = V^*(I_0) - V^*(I)$$

The value of having ΔI in addition to the base level of information I_0 is therefore $V(I|I_0)$. It is calculated by subtracting the expected system performance of *I* and I_0 (i.e. $V^*(I)$ and $V^*(I_0)$ respectively). In estimating the

expected system performance, it is assumed that the information available to the system operator is used optimally.

Note that the problem illustrated in Equation 8 is a problem to be minimised, where a lower value is a more desirable outcome. An example of this is a cost function, with flood damage costs having a positive value. If it is a max problem, the subtraction is reversed.

For each information level, value of information can be estimated in terms of the expected system performance under all foreseeable conditions. This performance metric can be taken as the optimised objective function (J*) discussed in Section 2.1.4. In the context of reservoir outflow control, the foreseeable conditions are considered as all reservoir inflow (q_0), and level (v_0) observed by the controller. This is formulated in Equation 9 in continuous form. $p(q_0)$ and $p(v_0)$ are the probability distributions of the observed inflow and reservoir level respectively. In this thesis, the formulation above is discretised as the probability distributions are represented as histograms, this is further explained in Sections 3.2.5 and 3.2.8.

Equation 9 Expected level of performance for a given level of information

$$V^*(I) = \int_{q_0} \int_{v_0} J^*(I, q_0, v_0) \cdot p(q_0) \cdot p(v_0) \cdot dq_0 dv_0$$

The optimisation problem used to calculate the level of performance is shown in Equation 10. It is immediately apparent that the equation is based on Equation 3 but with reservoir inflows and level being explicitly identified. Equation 10 also specifies: (1) a tree structure to be used, which will be affected by the level of information, (2) a water balance model of the reservoir, (3) constraints to the optimisation such as $0 < v_{t,z} < v_{max}$, (4) assumption of equal initial reservoir level for all ensemble members, and (5) that a trajectory of reservoir inflow is given for each observed initial flow. In this thesis the trajectory is produced by synthetic streamflow generation discussed in Section 3.2.5.

Equation 10 Multistage stochastic optimisation problem with a tree structure in the context of reservoir operation

$$J^*(I, q_0, v_0) = \min_{u_{t,z}} \sum_{z=1}^{n} p(z) \cdot \left[\sum_{t=1}^{h-1} g_t(v_{t,z}, u_{t,z}, q_{t,z}) \right]$$

$where$:

$$(1)\ u_{i,t} = u_{j,t}\ when \begin{cases} P_I(i) = j \\ t < B_I(i) \end{cases} i, j = 1, \dots, n$$

$$(2)\ v_{t,z} = m(v_{t-1,z}, u_{t,z}, q_{t,z})$$

$$(3)\ c\left(v_{t,z}, u_{t,z}, q_{t,z}\right) \leq 0$$

$$(4)\ v_{0,z} = v_0 \forall\ z \in \{1, \dots, n\}$$

$$(5)\ q_{t,z} | \ q_0$$

3. Methodology

The following sections provide an outline of the methodology required to achieve the objectives and outputs set out in Section 1.3. It summarises the modelling scenarios involved, and the tasks required to complete those modelling scenarios, as well as relevant assumptions.

3.1 Conceptualisation

The Salto Grande reservoir and its surroundings is simplified conceptually in this thesis as a reservoir with a catchment upstream and a community downstream, as shown in Figure 9

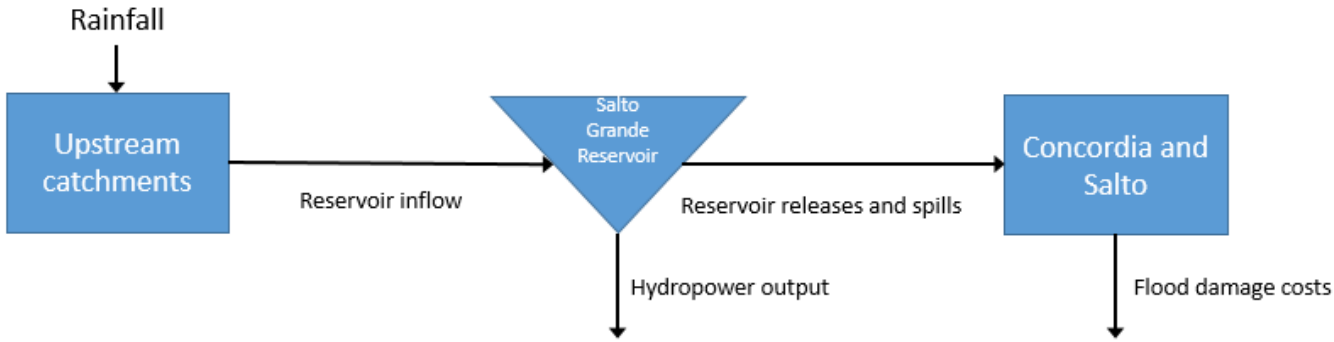


Figure 9 Simplified physical conceptualisation of Salto Grande Reservoir and surroundings

Note that there are significant assumptions inherent in such a conceptualisation:

- Rainfall only falls on the upstream catchments
    - The reservoir surface area, and the catchment between the dam wall and Concordia and Salto, is very small compared to the upstream catchment area. Therefore, to simplify the modelling, the additional runoff generated from rainfall falling directly into the reservoir, and from the catchment between the dam wall and Concordia and Salto are neglected. This implies that all floods at Concordia and Salto will be caused by reservoir spills and releases
- The only downstream communities considered are the cities of Concordia and Salto. Cities and towns further downstream are neglected
    - There are several cities and towns located on the banks of the Uruguay River further downstream of Concordia and Salto which may also be flooded in the event of very high flows in the Uruguay River. To simplify the modelling of flood damage, it is assumed that these cities and towns are located sufficiently far from the reservoir to not be affected by its controlled outflows and spills.
- Control effect of upstream dams are neglected
    - There are smaller dams located on the Uruguay River tributaries upstream of Salto Grande Reservoir. It is assumed that the effect of the controls of these reservoirs to be negligible in order to simplify the modelling of upstream catchments.

3.2 Modelling scenarios and elements

The modelling of the value of information in such a system shown in Figure 9 is conducted under three groups containing 5 different levels of information:

- Base case

This assumes that the controller has no access to any forecasted inflow. The controller only has access to inflow and storage level observations. The purpose of this is to provide a baseline level of information to illustrate the advantage the forecast scenarios

- Perfect model

This assumes that the rainfall runoff model used to convert forecasted rainfall ($P_{forecast}$) to forecasted inflow ($Q_{forecast}$) can perfectly model reality as observed in the observation network. This is done by assuming the observed inflow ($Q_{observed}$) is equal to the output of the rainfall runoff model run using gauged rainfall ($P_{observed}$) as inputs. The elements of this modelling and how they relate are shown in Figure 10

Two forecast products will be used resulting in two levels of information assuming a perfect model which is discussed in Section 3.2.2. The purpose of this scenario group is to compare weather forecast products in isolation without considering other forecast uncertainties. Note that this comparison is likely to exaggerate the differences between the two forecast products because in reality, other uncertainties will reduce any advantage a forecast product may have on another.

- Imperfect model

This assumes that the rainfall runoff model used to convert forecasted rainfall ($P_{forecast}$) to forecasted inflow ($Q_{forecast}$) is not perfectly accurate, as all models are in practice. This is done by obtaining the observed inflow ($Q_{observed}$) comes from the observation network. The elements of this modelling and how they relate are shown in Figure 11

Two forecast products will be used resulting in two levels of information assuming an imperfect model which is discussed in Section 3.2.2. The purpose of this scenario group is to compare weather forecast products in how they are likely to perform in practice, where rainfall runoff models does not provide a perfect representation of reality

The comparison between the 'perfect model' and 'imperfect model' scenario groups will also provide insights in quantifying the utility of having a very accurate (or 'perfect') model in operating a reservoir.

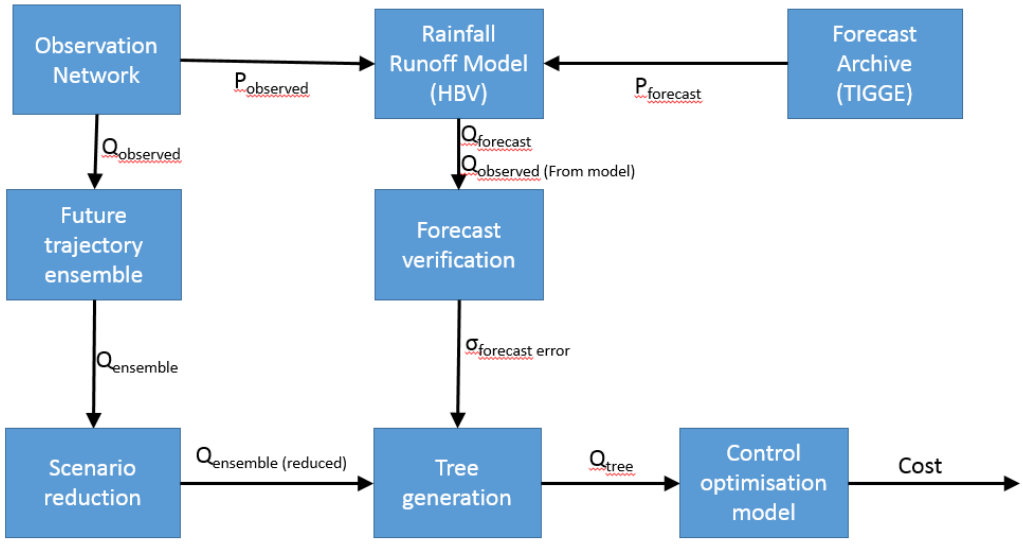


Figure 10 Modelling and optimisation elements and their relations (Perfect model)

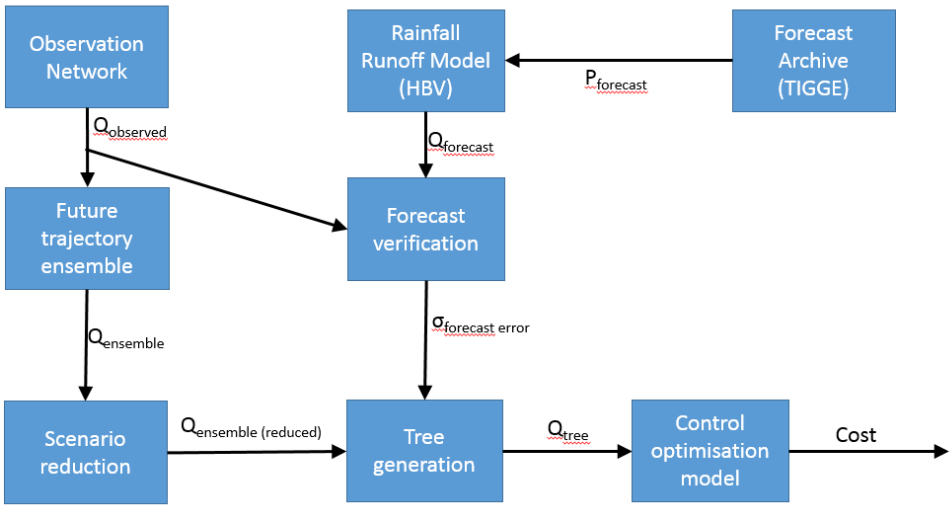


Figure 11 Modelling and optimisation elements and their relations (Imperfect model)

The following subsections further explain the elements and the information they deliver and receive.

3.2.1 Observation network

Observation network in this context of this thesis refers to the rainfall gauges used in the rainfall runoff model (see Figure 1), as well the measurement of inflow to the reservoir. All measured rainfall and inflow data were obtained from Raso (2013)

As discussed in Section 2.2, such observations have uncertainties associated with them, which must be analysed and when possible, quantified.

The observational uncertainty of interest in this thesis relates to the uncertainty of reservoir inflow measurement because reservoir inflows can be considered a disturbance that the reservoir outflow control is responding to.

The challenge is to provide a sufficiently robust quantification of this uncertainty, because as discussed in Section 2.2, there are many sources of uncertainty and the approaches used to quantify them.

In this thesis, the quantification of streamflow measurement uncertainty from Di Baldassarre & Montanari (2009) is adopted. Di Baldassarre & Montanari (2009) quantified the 95% confidence bands of streamflow measurements at the Po River basin in Italy. This confidence interval was back calculated to estimate a relationship between observed streamflow and the standard deviation of the observation, which is shown in Figure 12. The back calculation assumes the uncertainty is normally distributed and that the *68 – 95 – 99.7 rule* apply. The *68 – 95 – 99.7 rule* assumes that practically all values within the 95% confidence bands are within 2 standard deviations from the mean (Narasimhan, 1996).

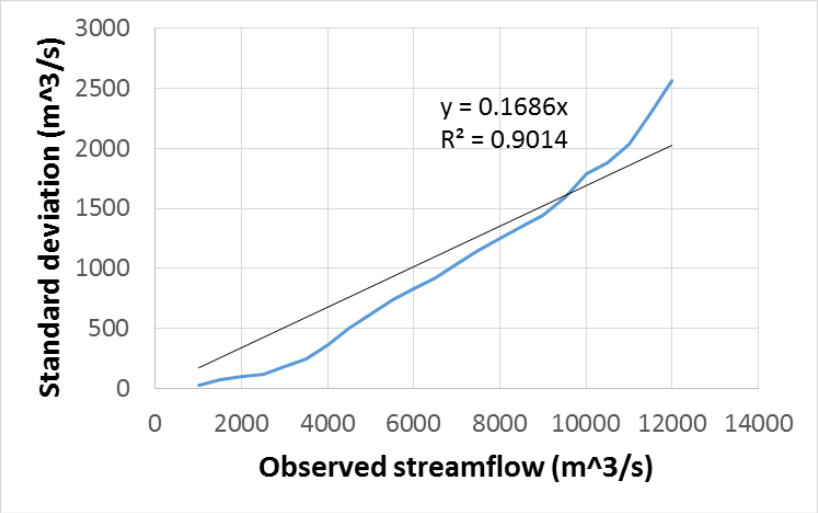


Figure 12 Standard deviation of measurement error as a function of observed streamflow (Di Baldassarre & Montanari, 2009)

A linear trendline is fitted to Figure 12. It's gradient provide an estimate of standard deviation of measurement error as a function of flow as shown in Equation 15. This standard deviation provides an indication of the spread of random errors for a given observation.

Equation 11 Standard deviation (a) and variance (b) of measurement error as a function of observed flow (Data from Di Baldassarre & Montanari (2009))

a) $\sigma_{observed} = 0.1686 \times Q_{observed}$

b) $Var_{observed} = {\sigma_{observed}}^2$

In the thesis, the standard deviation and variance are used in determining the bifurcation points of the decision tree which is discussed further in Section 3.2.7.

Note that Po River is located in a completely different region to Salto Grande, therefore there is a risk of this data not being applicable. However, due to a lack of similar studies available in the study area, and that the magnitude of flows observed at Po River is similar to Salto Grande, the uncertainty data from this study is assumed to be applicable.

3.2.2 Forecast archive (TIGGE)

TIGGE (THORPEX Interactive Grand Global Ensemble) is an archive that stores weather forecasts from several products released by a range of Numerical Weather Prediction (NWP) centres. It archives several weather forecasts including total precipitation, soil temperature, and soil moisture. (ECMWF, 2007)

For the purposes of this thesis, only the archived total precipitation forecasts ($P_{forecast}$ in Figure 10) are used as input to the HBV model.

There are two forecast products considered in this thesis. One is produced by the Brazilian agency CPTEC, which is currently used by the authorities at Salto Grande (J Talsma 2015, pers. comm. 6 Aug). The use of this forecast does not carry a cost. The other forecast product considered is produced by ECMWF. The authorities at Salto Grande Reservoir are considering the usage of this forecast product. The forecast from ECMWF carries a cost (J Talsma 2015, pers. comm. 6 Aug), which therefore can be weighed against its benefit.

For the two products, TIGGE provides archived ensemble forecasts, and control forecasts. TIGGE also archives deterministic forecasts exclusively for ECMWF. A control forecast is an unperturbed member of the ensemble

and represents the best approximation of current conditions and weather model physics. To produce ensemble forecasts, the initial conditions of the control forecasts are perturbed. (Buizza, et al., 2008)

In this thesis, the rainfall forecasts used as $P_{forecast}$ are deterministic, which is consistent with the type of forecasts used currently by the authorities at Salto Grande (J Talsma 2015, pers. comm. 5 Sep). For both CPTEC and ECMWF, it is assumed that the control forecast can be taken as a deterministic forecast. Rainfall forecasts are released every 12 hours (midday and midnight), and provide forecasts up to 15 days ahead in 6 hour increments (i.e. 60 timesteps with one timestep being 6 hours long).

3.2.3 Rainfall runoff model and routing

The rainfall runoff and routing modelling framework in this thesis was taken directly from Raso (2013). The rainfall runoff model is a semi-distributed HBV model (Bergström, 1995), which was calibrated over a period between 28/12/2009 and 28/06/2011, using observed rainfall from 5 gauges, and observed streamflow from 3 gauges. The rainfall runoff model, is coupled with a triangular unit hydrograph routing process. Figure 13 shows the layout of the model.

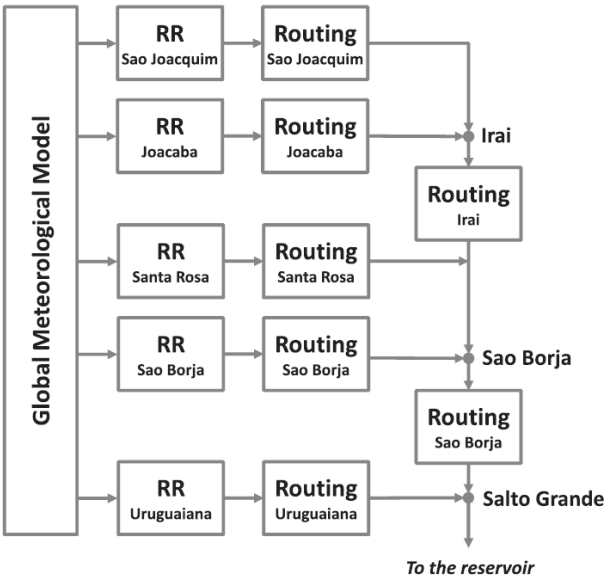


Figure 13 Structure of semi-distributed HBV model (Raso, et al., 2013)

The calibration of the model, as reported in Raso (2013), resulted in a Nash-Sutcliffe Efficiency Index of 0.74. No recalibration of model parameters are conducted in this thesis.

As implied in Figure 10 and Figure 11 there are two rainfall runoff model runs:

- Observed (Base case) run:
    - To produce observed streamflow data ($Q_{observed}$), in the 'perfect model' level of information, using observed rainfall ($P_{observed}$) as input.
    - To produce initial conditions for the forecast run
- Forecast run:
    - To produce forecasted inflow ($Q_{forecast}$), in the 'imperfect model' level of information, using forecasted rainfall ($P_{forecast}$) as input.

- The model is run for each instance that a forecast is released.
- The state of the HBV model under the observed run during the forecast release time is used to set the initial conditions of the model for each of the forecast run.
    - For the 'perfect model' scenario, this implies that not only the hydrological model calibration perfectly represents the system, but also that the initial conditions of the model perfectly captures the conditions of the system at the start of the forecast run
    - For the 'imperfect model' scenario, this implies that both the calibration of the hydrological model and the initialisation is not robust. This initialisation weakness is crucial. In reality initial conditions of such a model is likely to use actual measurements and observations instead of it being estimated from a model. Therefore this scenario carries a more pessimistic assumption of model accuracy than what is likely to occur in reality

In addition to the inputs mentioned above, there are also temperature and evaporation inputs required to run the model. It is assumed that this information is available to be used in the model at all times. The value of this information is therefore disregarded in this thesis. The temperature and evaporation data used in the model is also taken from Raso (2013) and is presented in Appendix B.

3.2.4 Forecast verification

Before calculating the value of information using optimised decision trees, forecast accuracy is calculated using the observed data from telemetry, the rainfall forecasts obtained from TIGGE, and the streamflow forecasts produced by the HBV model. Forecast verification in this thesis is conducted using three standard measures briefly explained in the following subsections. The results of the forecast verification are presented in Section 4.1.

3.2.4.1 Mean Error

Mean Error is a measure of forecast bias. It is shown in Equation 12. If Mean Error is positive, it is an indication that on average the forecast overestimates the observation. If Mean Error is negative, it is an indication that on average, the forecast underestimates the observation

Equation 12 Mean Error

$$Mean\ Error = \frac{1}{n}\sum_{i=1}^{n}(f_i - x_i)$$

Note that in this thesis. bias is neglected in generating decision trees. Uncertainties are only represented by covariance matrices of forecast error and the variance of measurement error.

3.2.4.2 Standard error

Standard error measures the spread of measurement and forecast errors, under the assumption that those errors are normally distributed. The higher this number, the higher the likelihood of large random errors occurring for a given observation or forecast. If only measurements are considered, the standard error can be taken as the variance which is a scalar value as shown in Equation 11.

If forecasts are also considered, the standard error is represented by a covariance matrix. A covariance quantifies how two random variables influence each other. A covariance matrix is a matrix of covariances of several random variables. Equation 13 shows an expression of the covariance of two random variables. Equation 14 shows an expression of a covariance matrix of three random variables.

Equation 13 Covariance of two random variables X, and Y (Dekking, et al., 2005)

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])$$

Equation 14 Covariance matrix of random variables X_1, X_2, and X_3

$$C = \begin{bmatrix} Cov(X_1,X_1) & Cov(X_1,X_2) & Cov(X_1,X_3) \\ Cov(X_2,X_1) & Cov(X_2,X_2) & Cov(X_2,X_3) \\ Cov(X_3,X_1) & Cov(X_3,X_2) & Cov(X_3,X_3) \end{bmatrix}$$

Note that a covariance of the same random variable is the same as the variance. Therefore the diagonal of a covariance matrix contains the variance of each of the random variables. In this thesis the random variables can be considered as the vectors of forecast errors at various timesteps. Therefore as the forecasts used in this thesis have 60 timesteps, the covariance matrix that represents their standard error has a dimension of 60 x 60.

3.2.4.3 Mean Squared and Root Mean Squared Error

Mean Squared Error (MSE) and Root Mean Squared Error are shown in. Root Mean Squared Error (RMSE) square roots MSE so that it's units is consistent with the units of the forecasts and observations. Both measures are a composite of bias and variance, in such a way that if there is no bias, MSE would be equal to the variance (Lebanon, 2010).

Equation 15 Mean Squared (a) and Root Mean Squared Error (b)

$$(a)\ MSE = \frac{1}{n}\sum_{i=1}^{n}(f_i - x_i)^2 \qquad (b)\ RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - x_i)^2}$$

3.2.5 Future trajectory ensemble

The reservoir inflow trajectory ensemble in this thesis is produced using synthetic streamflow generation. Synthetic streamflow generation uses stochastic processes to generate streamflow trajectories which mimic the statistical characteristics of historic data (Ochoa-Rivera, et al., 2002). Available synthetic streamflow generation methods can be simple, such as the Thomas-Fiering model, the autoregressive moving average (ARMA) and the autoregressive integrated moving average (ARIMA) models (Stedinger & Taylor, 1982). Those models only require historic streamflow data to be utilised. There are also more complex and advanced methods, such as methods that use artificial neural networks (Ahmed & Sarma, 2007; Ochoa-Rivera, et al., 2002; Zealand, et al., 1999) and the use of stochastic weather generators coupled with a hydrological model (Zahabiyoun, 1999; Ramadan, et al., 2012).

In this thesis an second order autoregressive, first order moving average (ARMA (2,1)) model is used to generate streamflow trajectories. The rationale behind this is:

- Simplicity. A more complex model is not desirable as the ensemble generation component is not the main focus of this thesis
- The model has the ability to produce hydrographs that are visually similar to the historic data based on spot checks
- Seasonality of the historic observations is unable to be depicted due to a limited dataset. The model assumes a stationary timeseries

The model is arbitrarily set to produce 2000 individual equiprobable streamflow trajectories. The model is expressed in Equation 16.

Equation 16 ARMA (2,1) model

$$Q_t = c + \varepsilon_t + \varphi_1 Q_{t-1} + \varphi_2 Q_{t-2} + \theta_1 \varepsilon_{t-1}$$

The parameters φ_1, φ_2, θ_1, as well as the variance of the error term ε_t are fitted to the log transformed observed reservoir inflow using the Maximum Likelihood Estimate method. The log transformation is to prevent the model from producing negative inflows. Therefore, the outputs of the ARMA (2,1) model have to be inverse log transformed before being used as a streamflow trajectory ensemble. Table 3 shows the fitted parameters. Note that ideally, fitting of such parameters should use data that covers a period of several years to adequately represent interannual variability. However, due to the limited availability of observed flow data, this is not possible in this thesis. The parameters are only fitted to 1 year and 5 months of 6 hourly observed data.

The applicability of the ARMA (2, 1) model structure and parameters is not fully validated in this thesis due to it not being the main focus of the work. For a more robust synthetic streamflow generation method, validation using approaches in Stedinger & Taylor (1982) or Lettenmaier & Burges (1982) can be considered.

Table 3 ARMA (2,1) model parameters

Parameters	*Value*
c	0.02560
φ_1	1.92322
φ_2	-0.92628
θ_1	0.91352
Variance of ε_t	4.79038e-05

As mentioned in Section 2.3.2.2, a streamflow trajectory ensemble is to be produced for each observed inflow. However, considering that there is 1 year and 5 months of 6 hourly observations, it would be highly impractical to generate ensembles and scenario trees for all of them. Therefore a 2D histogram is to be used to represent the observations. The histogram is 2D because of the second order autoregressive component of the model requires the first timestep of the trajectory to be calculated using the current observed flow and the observed flow 1 timestep before current (or the difference between current and 1 timestep before current). The histogram is shown in Figure 14.

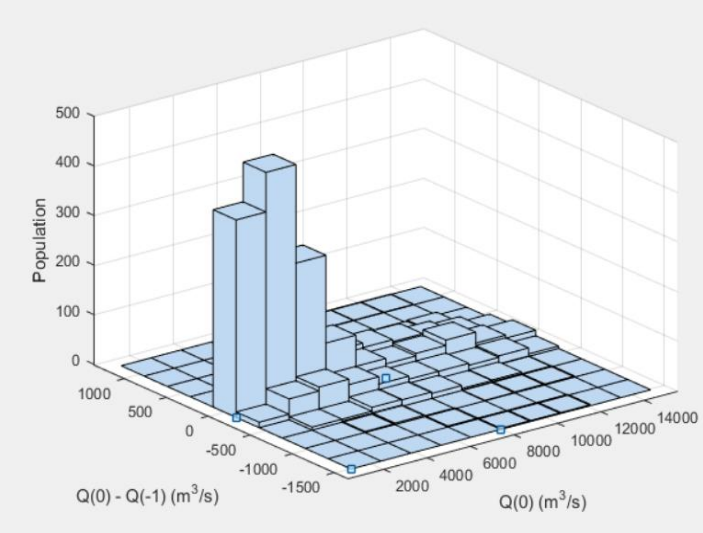


Figure 14 2D histogram of observed flows

The histogram has 55 populated bins. Therefore, there are 55 future trajectory ensembles to be produced by the ARMA (2, 1) model. These 55 are assigned scenario numbers which is shown in Table 4. The probability of each scenario is shown in

Table 5.

Table 4 Scenario numbers of future trajectory ensemble

Scenario numbers		Q(0) - Q(-1) (m^3/s)									
		-1377	-1103	-828	-554	-280	-5	269	543	818	1092
Q(0) (m^3/s)	1757					18	28				
	3136				10	19	29	38			
	4515			5	11	20	30	39			
	5894			6	12	21	31	40	47		
	7273		4		13	22	32	41	48		
	8652	1			14	23	33	42	49		
	10031	2		7	15	24	34	43	50		53
	11410	3		8	16	25	35	44	51		54
	12789			9	17	26	36	45	52		55
	14168					27	37	46			

Table 5 Prior probability of scenario

Probability (%)		Q(0) - Q(-1) (m^3/s)									
		-1377	-1103	-828	-554	-280	-5	269	543	818	1092
Q(0) (m^3/s)	1757					0.52%	18.75%				
	3136				0.28%	1.98%	22.52%	1.13%			
	4515			0.09%	0.28%	2.36%	13.14%	3.25%			
	5894			0.14%	0.66%	1.84%	4.80%	3.39%	0.14%		
	7273		0.09%		0.61%	1.13%	1.84%	2.31%	0.28%		
	8652	0.09%			0.47%	0.85%	1.32%	1.88%	0.24%		
	10031	0.09%		0.05%	0.19%	0.85%	1.55%	1.70%	0.09%		0.05%
	11410	0.09%		0.09%	0.24%	0.66%	2.07%	1.18%	0.19%		0.05%
	12789			0.09%	0.05%	0.80%	1.04%	0.85%	0.28%		0.05%
	14168					0.24%	0.75%	0.33%			

3.2.6 Scenario reduction

As mentioned in Section 3.2.5, the ensemble generation from the ARMA (2,1) model produces a future inflow trajectory ensemble of 2000 members with uniform probability. If the ensemble created in the ARMA (2,1) model is inputted directly to generate a tree, it would result in prohibitively long computation times, especially considering that there are 55 ensembles generated. Therefore, the number of scenarios has to be reduced.

The scenario reduction method follows the method in Raso, et al. (2013), which adopts an algorithm by Growe-Kuska, et al. (2003). The algorithm involves iteratively identifying the two closest ensemble members, deleting one of them, and assigning its probability to the other. The more neighbouring ensemble members, the more

they are deleted and their probabilities assigned, resulting in a higher probability member in the reduced ensemble.

In this thesis, the amount of members in the reduced ensemble is 12, which is driven by the need to limit excessive computation times during the control optimisation phase. Figure 15 provides an example of an ensemble before and after scenario reduction. Appendix E contains additional examples.

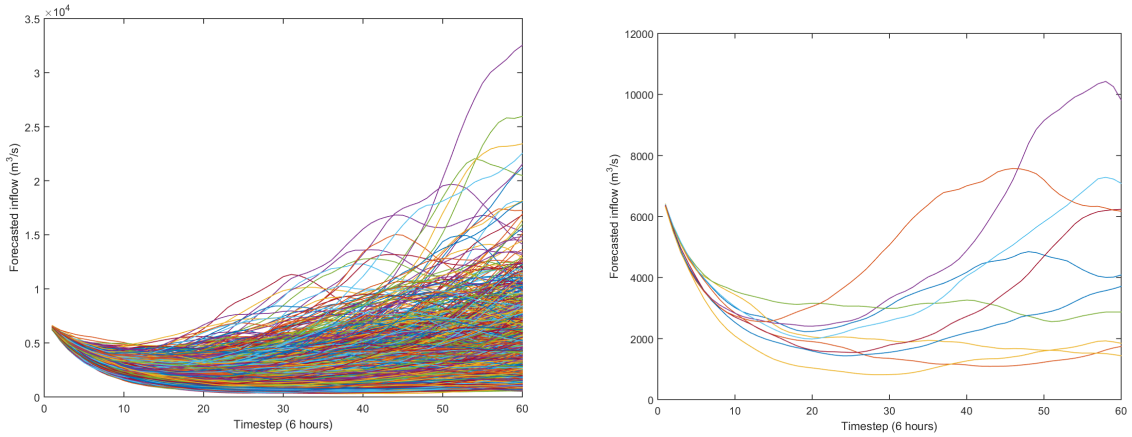


Figure 15 Example of an ensemble before (left) and after (right) being processed by the scenario reduction algorithm

3.2.7 Tree structure definition

The probabilities of both the pre- and post- reduced ensemble members can be considered as the prior probability, because the probability is taken at the initial timestep. As time moves forward along the forecast horizon, more information is obtained creating a new probability distribution which can be referred to as the posterior probability. When the information is continuous and scalar, the posterior probability can be calculated using the Bayes' rule shown in Equation 17. $\mathrm{p_t(x)}$ represents the prior probability of the ensemble member, and $p_{t+1}(x|I_t)$ represents the posterior probability distribution given an information level I_t.

Equation 17 Bayes' rule for continuous, one dimensional observations (Raso, et al., 2013)

$$p_{t+1}(x|I_t) = \frac{f_{I_t}(I_t|x)p_t(x)}{\sum_{j=1}^{N} f_{I_t}(I_t|x_j)p_t(x_j)}$$

Equation 17 cannot be applied directly in tree generation, because a tree is generated at the initial timestep when the aforementioned information is not yet available. To address this, Raso, et al. (2013) proposes using the Bayes' rule on average, at a given point in time, assuming a certain ensemble member is the occuring one. This is shown in Equation 18. x_k can be considered as the ensemble member that is assumed to be the occuring one. Observations and forecasts for that ensemble member will be taken from the distribution $f_{I_t}(I_t|x_j)$. The variance of the distribution is the standard error described in Section 0. This means that in this thesis, where a forecast has 60 timesteps, $f_{I_t}(I_t|x_j)$ is a multivariate distribution of 60 dimension. Integrating such functions as prescribed in Equation 18. is clearly complicated. Therefore in this thesis, the integration component in Equation 18 will be computed using Monte Carlo integration by generating random samples of observation vectors.

Equation 18 'On average' application of the Bayes' rule for a given information level and occurring ensemble member (Raso, et al., 2013)

$$E\left[p_{t+1}\left(x\middle|(I_t|x_k)\right)\right] = \int_{-\infty}^{\infty} p_{t+1}(x|I_t) \cdot f_{I_t}(I_t|x_k)dI_t$$

Equation 18 is to be applied for every timestep along the forecast horizon so that the probability dynamic of the occuring ensemble member x_k can be quantified. This is shown in Equation 19 where it defines how the probability of x_k will approach one as time moves forward towards the horizon, while the probability of other ensemble members will approach zero. Note that ϕ can be considered as a dependence function in which Equation 18 is rearranged so that its inputs are clearly presented.

Equation 19 Dynamic probability of ensemble members (Raso, et al., 2013)

$$(a)\ p_t(x) = p(x)\ \ for\ t = 1$$

$$(b)\ p_{t+1}(x) = \phi\big(p_t(x), f_{I_t}(I_t|x), x_k\big)\ \ for\ t = 1, \dots, H-1$$

This dynamic probability is key in determining the point in time where one ensemble member is distinguishable from another. This can also be referred to as the bifurcation point of a tree. This is done by comparing two ensemble members at a time, x_k which is the occuring one, and a different member x_j to calculate the probability of x_k in this reduced sample space. This is shown in Equation 20.

Equation 20 Probability of an occurring ensemble in a reduced sample space (Raso, et al., 2013)

$$p_t^r(x_k) = \frac{p_t(x_k)}{p_t(x_k) + p_t(x_j)}$$

From that, a distinguishability matrix can be produced. A distinguishability matrix is an n x n matrix, in which n is the number of ensemble members. It defines, between every pair of ensemble members, the time when the probability of the occuring member in the reduced sample space $p_t^r(x_k)$, exceeds a predefined confidence level p^*. This is shown in Equation 21. In this thesis, $p^*$is specified as 0.95.

Equation 21 Distinguishability matrix (Raso, et al., 2013)

$$D_{k,j}(p^*) = t|min_t[p_t^r(x_k) > p^*]$$

$$for\ \forall\ x_k \neq x_j$$

Once the distinguishability matrix has been created, a tree structure can be defined from it using the recipe prescribed in Raso, et al. (2013) below.

1. *Find the maximum value of D (maxD) and the member corresponding to its row and column, called 'row' and 'column' member*
2. *The maxD value and the column member are assigned to the row member as branching point and parent, respectively*
3. *Matrix D is reduced, removing row and column of' row member' (its row and its column)*
4. *The procedure from (1) to (3) is repeated until a single member remains*
5. *The last member is the root, its parent is itself and its branching point is 1*

(Raso, et al., 2013, p. 79)

For all 55 scenarios in Table 4, 5 trees are created for each of the levels of information described in Section 3.2 and forecast products mentioned in Section 3.2.2:

- Base case
- 'Perfect Model', CPTEC forecast product
- 'Perfect Model', ECMWF forecast product
- 'Imperfect Model', CPTEC forecast product
- 'Imperfect Model', ECMWF forecast product

Therefore, in total, there are 275 trees produced in this thesis.

3.2.8 Control Optimisation (reservoir) model

Each of the 275 trees, are to be put into a reservoir model, which will optimise the reservoir releases and produce a performance score, which can be used to compare the value of each information level. The reservoir model is taken from Raso, et al. (2014a) which programmed it using RTC-Tools (Schwanenberg & Becker, 2014). The discrete mass balance equation of the reservoir model is shown in Equation 22. v is the reservoir volume which can increase due to inflow (Q) and decrease due to the outflow (u). Outflow can be through the turbines, which generates electricity, or through the spillway, which can prevent upstream flooding if the turbine is at capacity. Δt is the timestep which is 6 hours in this thesis. Losses or gains due to climate or groundwater is ignored.

Equation 22 Reservoir mass balance (Raso, et al., 2014a)

$$v_t = v_{t-1} + \Delta t \,.\, (Q_t - u_t^{turb} - u_t^{spill})$$

A relationship between reservoir volume and height is defined using the rating curve shown in Equation 23.

Equation 23 Volume – height relationship of Salto-Grande Reservoir (Raso, et al., 2014a)

$$v_t = 1830\ {h_t^{salto}}^{4.2}$$

Flow through the turbines and spillway are constrained as defined in Equation 24.

Equation 24 Constraint of environmental outflows (Raso, et al., 2014a)

$$(a)\quad 0 \le u_t^{turb} \le u_t^{turb,max}$$

$$(b)\quad 0 \le u_t^{spill} \le u_t^{spill,max}$$

Maximum turbine flow ($u_t^{turb,max}$) and spillway flow ($u_t^{spill,max}$) are not constant. $u_t^{turb,max}$ would depend on the head difference of the dam ($h_t^{salto} - h_t^{downstream}$). The curve that governs that relationship is shown in Figure 16. $u_t^{spill,max}$ would depend on the reservoir level (h_t^{salto}) which relationship is presented in Figure 17. The water level downstream of the reservoir ($h_t^{downstream}$) itself is a function of the total outflow from the reservoir ($u_t^{turb} + u_t^{spill}$), which relationship is defined by the rating curve in Figure 18.

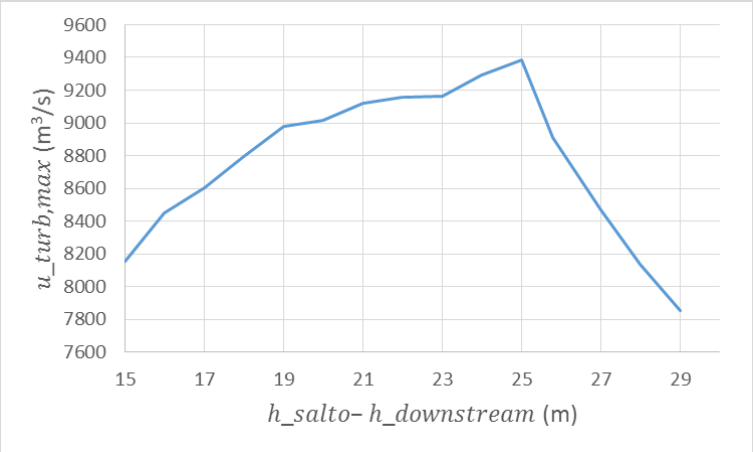


Figure 16 Maximum turbine flow ($u_t^{turb,max}$) and head difference ($h_t^{salto} - h_t^{downstream}$) relationship (Raso, et al., 2014a)

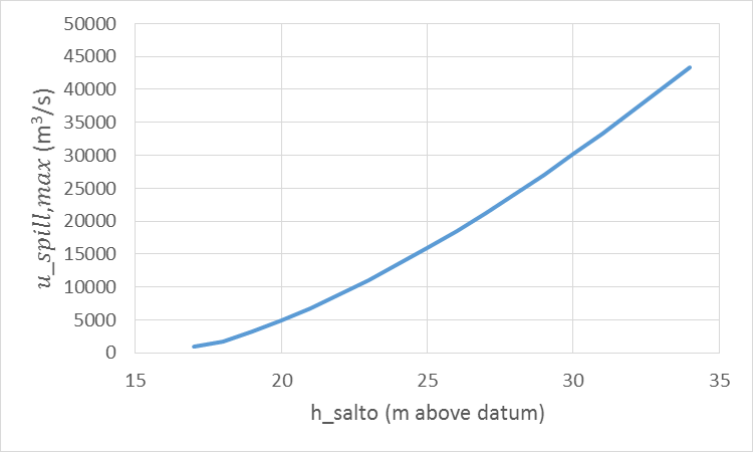


Figure 17 Spillway flow $(u_t^{spill,max})$ and reservoir level (h_t^{salto}) relationship (Raso, et al., 2014a)

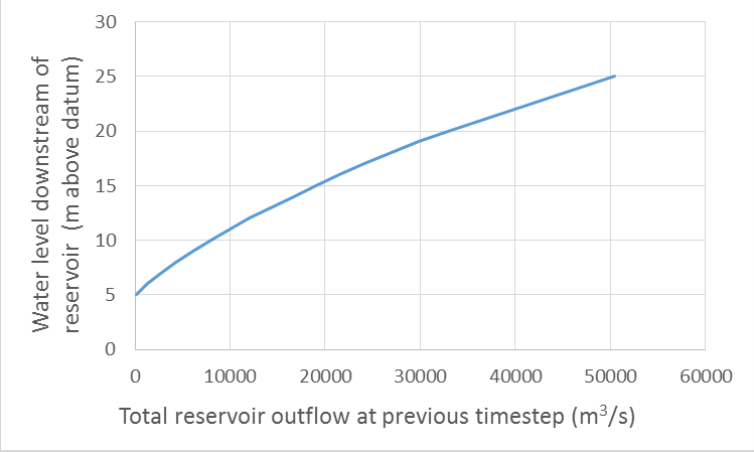


Figure 18 Water level downstream of reservoir $(h_t^{downstream})$ rating curve (Raso, et al., 2014a)

3.2.8.1 Objective function

As mentioned in the early sections of this thesis, Salto Grande Reservoir is considered to be a dual-purpose reservoir which is supposed to generate hydropower electricty while preventing upstream and downstream communities from being flooded. The objective function is therefore defined to reflect this pupose. The function of the optimisation is divided into 8 subfunctions. The scores of each subfunction (g_i) are then weighted based on their respective weights (w_i) and summed to form the final objective function score (J), as shown in Equation 25.

Equation 25 Weighing of objective subfunctions (Raso, et al., 2014a)

$$J = \sum_{t=1}^{h} \sum_{i=1}^{m} w_i \cdot g_{i,t}$$

The weights used for each subfunctions are summarised in Table 6.

Table 6 Subfunction formulas and weights

i	Subfunction	Weight (w_i)
1	Flood upstream	100
2	Over-emptying prevention	100
3	Energy production (head difference)	1
4	Energy production (spill prevention)	$5 * 10^{-5}$
5	Flood downstream	10^{-4}
6	Passing (environmental) flows	10^{-4}
7	Wear and tear (turbine)	$5 * 10^{-7}$
8	Wear and tear (spillway)	$5 * 10^{-7}$

The first subfunction (Equation 26) is put in place to penalise overflooding of upstream communities due to the overfilling of the reservoir. As seen in the equation it is assumed that the threshold for upstream flooding is when the reservoir water level exceeds 35 m above datum.

Equation 26 Subfunction 1: Flood upstream (Raso, et al., 2014a)

$$g_{1,t} = [\max(h_t^{salto} - 35,0)]^2$$

The second subfunction (Equation 27) attempts to prevent excessive reservoir drawdown, by penalising instances where the reservoir level falls below 30 m above datum. The first two subfunctions therefore imply that the controller imposes a reservoir level operating range of between 30 and 35 m above datum.

Equation 27 Subfunction 2: Over-emptying prevention (Raso, et al., 2014a)

$$g_{2,t} = [\min(h_t^{salto} - 30,0)]^2$$

The third (Equation 28) and fourth (Equation 29) subfunctions attempts to maximise energy production. The third subfunction does it by penalising reservoir levels below 35 m above datum, which indirectly rewards large head differences. The fourth subfunction does it penalising uneccesary spillway flow, which indirectly promotes turbine flow.

Equation 28 Subfunction 3: Energy production (head difference) (Raso, et al., 2014a)

$$g_{3,t} = [\min(h_t^{salto} - 35,0)]^2$$

Equation 29 Subfunction 4: Energy production (spill prevention) (Raso, et al., 2014a)

$$g_{4,t} = [u_t^{spill}]^2$$

The fifth subfunction (Equation 30) attempts to prevent downstream flooding by penalising excessive reservoir releases. It is assumed that the threshold of excessive reservoir release is 20,000 m^3/s.

Equation 30 Subfunction 5: Flood downstream (Raso, et al., 2014a)

$$g_{5,t} = [\max(u_t^{total} - 20000,0)]^2$$

The sixth subfunction (Equation 31) imposes an environmental flow requirement by penalising reservoir outflows lower than 500 m^3/s.

Equation 31 Subfunction 6: Passing (environmental) flows (Raso, et al., 2014a)

$$g_{6,t} = [\min(u_t^{total} - 500,0)]^2$$

The seventh and eighth subfunctions (Equation 32) attempts to prevent structural wear and tear by penalising drastic changes in turbine and spillway flow.

Equation 32 a: Subfunction 7: Wear and tear (turbine) and b: Subfunction 8: Wear and tear (spillway) (Raso, et al., 2014a)

$$(a)\ g_{7,t} = [u_t^{turb} - u_{t-1}^{turb}]^2$$

$$(b)\ g_{8,t} = [u_t^{spill} - u_{t-1}^{spill}]^2$$

3.2.8.2 Optimisation procedure

The model was and uses an interior-point (non linear) optimiser as its optimisation algoritm. Optimisation results would depend on the initial conditions of the state of the system (reservoir level), disturbance (inflow) and control action. The initial disturbance (Q_0) is specified as Q(0) of each scenario number in Table 4. For the control action, which is the reservoir outflow (u_0^{turb} and u_0^{spill}), the initial value is assumed to be the same as the inflow. It is also assumed that the initial turbine flow capacity ($u_0^{turb,max}$) is 8,400 m^3/s as approximated in Table 1. Therefore, if the inflow exceeds 8,400 m^3/s, it is assumed that the first 8,400 m^3/s is to flow through the turbines, with the remainder flowing through the spillway.

The initial reservoir level (h_0^{salto}) is determined using an analysis of the historical records of reservoir level. A plot of the historical reservoir level is available in Appendix A. The historical levels are then represented as a histogram with 3 bins, as shown in Figure 19. For every scenario and information level optimisation is to be done three times, once for each initial reservoir level specified in the histogram. The optimisation score is then weighed based on the probability of each bin.

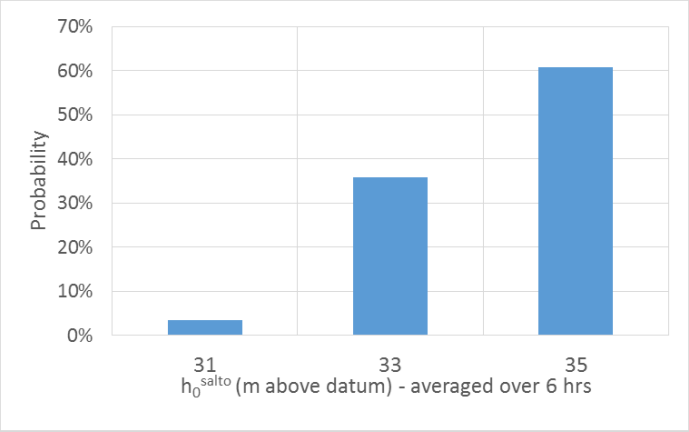


Figure 19 Histogram of historic reservoir level

Note that occasionally, the optimisation algorithm is not able to locate a local minima of the objective function within a feasible timeframe. This problem was not encountered in previous work by Raso, et al. (2014a) and was unable to be resolved during the course of the thesis. Therefore a time limit of 15 minutes is imposed on the optimisation algorithm. It is assumed that within 15 minutes, the algorithm is able to arrive at a 'quasi-optimal' result. This quasi-optimal result is then taken as the optimised objective function score. The prevalence of this problem, and with it the adoption of a quasi-optimised objective function score, is highest when the system is subjected to high flows.

3.3 Period of study

The period of study indicates the period in time, in which forecast and observed data are extracted. It is limited by the availability of measured data, in which the forecast uncertainty was derived from. The period of study in this thesis is between January 2010 and May 2011. These are the only complete months where rainfall and

streamflow data are available from the Salto Grande telemetry. The period is also within the calibration period of the HBV model as indicated in Raso, et al. (2014a).

3.4 Sequence of tasks

Considering the methodology presented in the previous subsections, the sequence of tasks required to calculate the value of information are:

- Step 1: Obtain weather forecasts and observed data
- Step 2: Run forecasted rainfall in the hydrological model
- Step 3: Verify forecast to quantify the variance of forecast error and produce the covariance matrix
- Step 4: For each observed inflow considered, synthetically produce an ensemble of all possible future trajectories using a synthetic streamflow generator
- Step 5: Reduce the future trajectory ensemble using scenario reduction algorithm
- Step 6: Generate tree for all levels of information and future trajectory ensemble
- Step 7: Utilise tree structure in optimising reservoir operations
- Step 8: Calculate value of information

4. Results and analysis

4.1 Forecast accuracy verification

4.1.1 Forecasted rainfall

Initially, Mean Error (ME) was calculated for forecasted rainfall obtained from TIGGE. The purpose is to obtain an understanding of the bias of the respective rainfall forecast products. ME of rainfall forecast is plotted against forecast leadtime for both CPTEC and ECMWF as shown in Figure 20.

It is immediately apparent that CPTEC on average, overestimates rainfall, particularly in larger leadtimes, whereas ECMWF tend to underestimate rainfall.

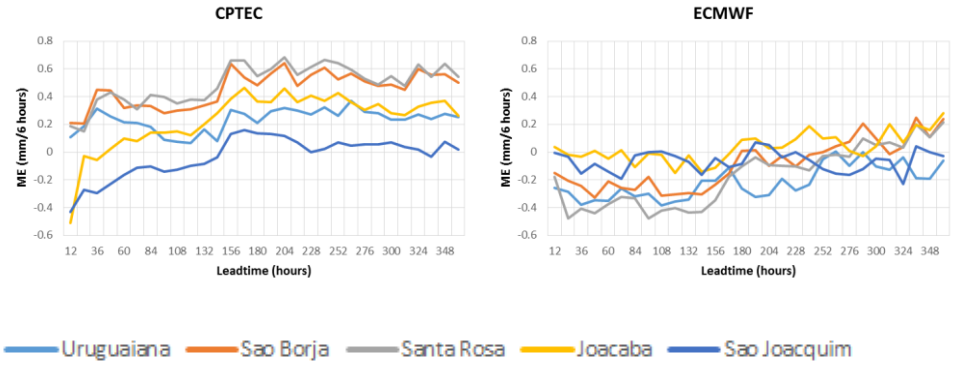


Figure 20 Forecasted rainfall ME comparison (Left: CPTEC, Right, ECMWF) for each rainfall sites of the HBV model

The next verification is the calculation of Root Mean Square Error (RMSE). The purpose is to obtain an understanding on which product is more likely to produce large random errors. RMSE of rainfall forecast is plotted against forecast leadtime for both CPTEC and ECMWF as shown in Figure 21.

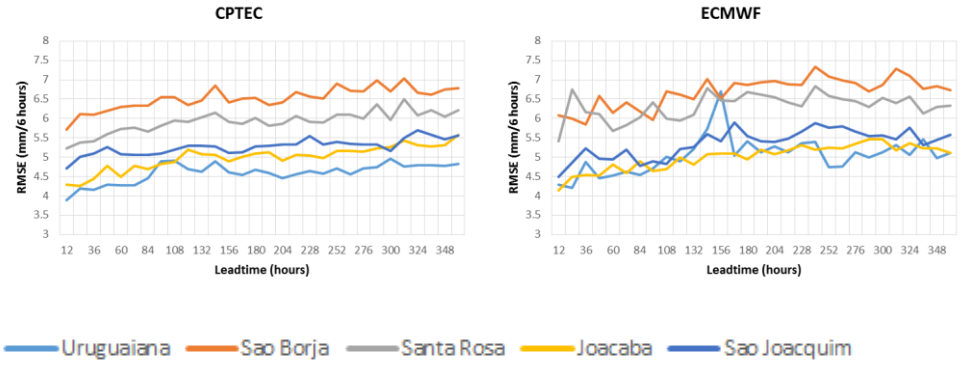


Figure 21 Forecasted rainfall RMSE comparison (Left: CPTEC, Right, ECMWF) for each rainfall sites of the HBV model

From the plot it is apparent that the RMSE of forecasted rainfall is more location dependent than forecast product dependent. Differences between forecast products are not as easily distinguishable as a comparison of ME. A closer analysis would reveal that ECMWF has a slightly higher RMSE when averaged across leadtimes. The correlation between RMSE and leadtime is also weaker for ECMWF. This suggests that ECMWF is slightly more likely to produce large random errors, and the prevalence of large random errors may not be strongly correlated to the forecast leadtime.

4.1.2 Forecasted reservoir inflows

The same calculations were also done to the forecasted reservoir inflows produced by the HBV model. The errors are calculated both against the 'perfect model' and 'imperfect model' observed inflows.

Figure 22 shows the mean error of forecasted inflow. The bias apparent in Figure 20 is also apparent in Figure 22, where ECMWF underestimates whereas CPTEC overestimates. There are two key things from Figure 22. One is that the magnitude of forecasted inflow bias can be very large, particularly in larger leadtimes, even though the magnitude of forecasted rainfall bias appear to be small. This is because errors in forecasted rainfall when inputted into the HBV model accumulate as the model runs through the forecast horizon. The other is that there are no major differences between the 'perfect model' and 'imperfect model' scenarios. This indicates that the HBV model itself has a very small bias. This is confirmed as the average HBV model error across all leadtimes is only 30 m^3/s.

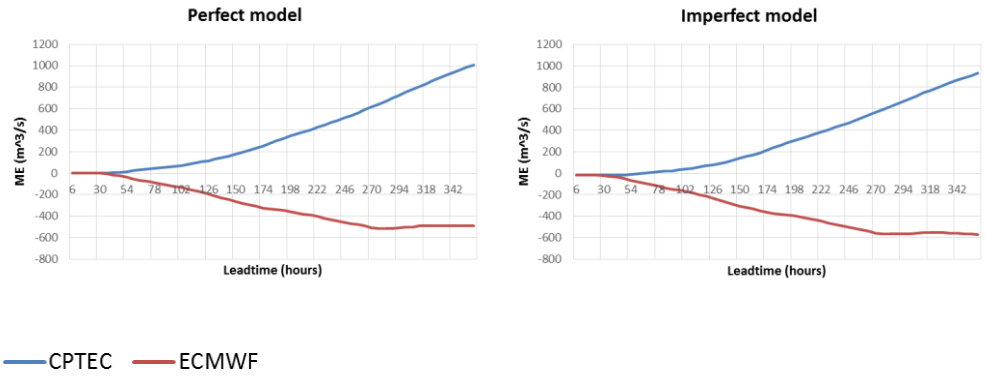


Figure 22 Forecasted inflow to Salto Grande ME comparison (Left: observed inflow from HBV using observed rainfall ('perfect model'), Right: observed inflow from telemetry ('imperfect model'))

Figure 23 shows the root mean square error (RMSE) of forecasted inflow. In this figure, the difference between 'perfect model' and 'imperfect model' scenarios are clearly apparent, showing the significant prevalence of larger random errors due to the HBV model at all leadtimes particularly in earlier leadtimes which highlights the weakness of the initialisation implemented in the 'imperfect model' scenario. Differences between the two forecast products are not as clearly apparent. No forecast product has a clear advantage in this measure as they have similar RMSE across all leadtimes and scenarios.

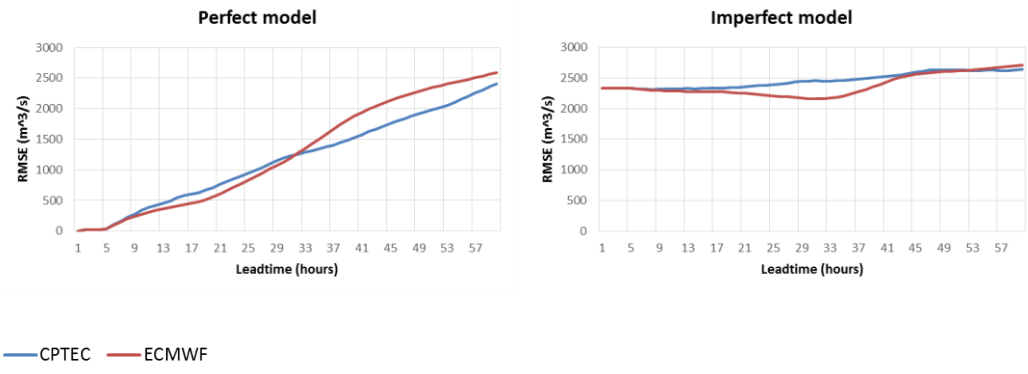
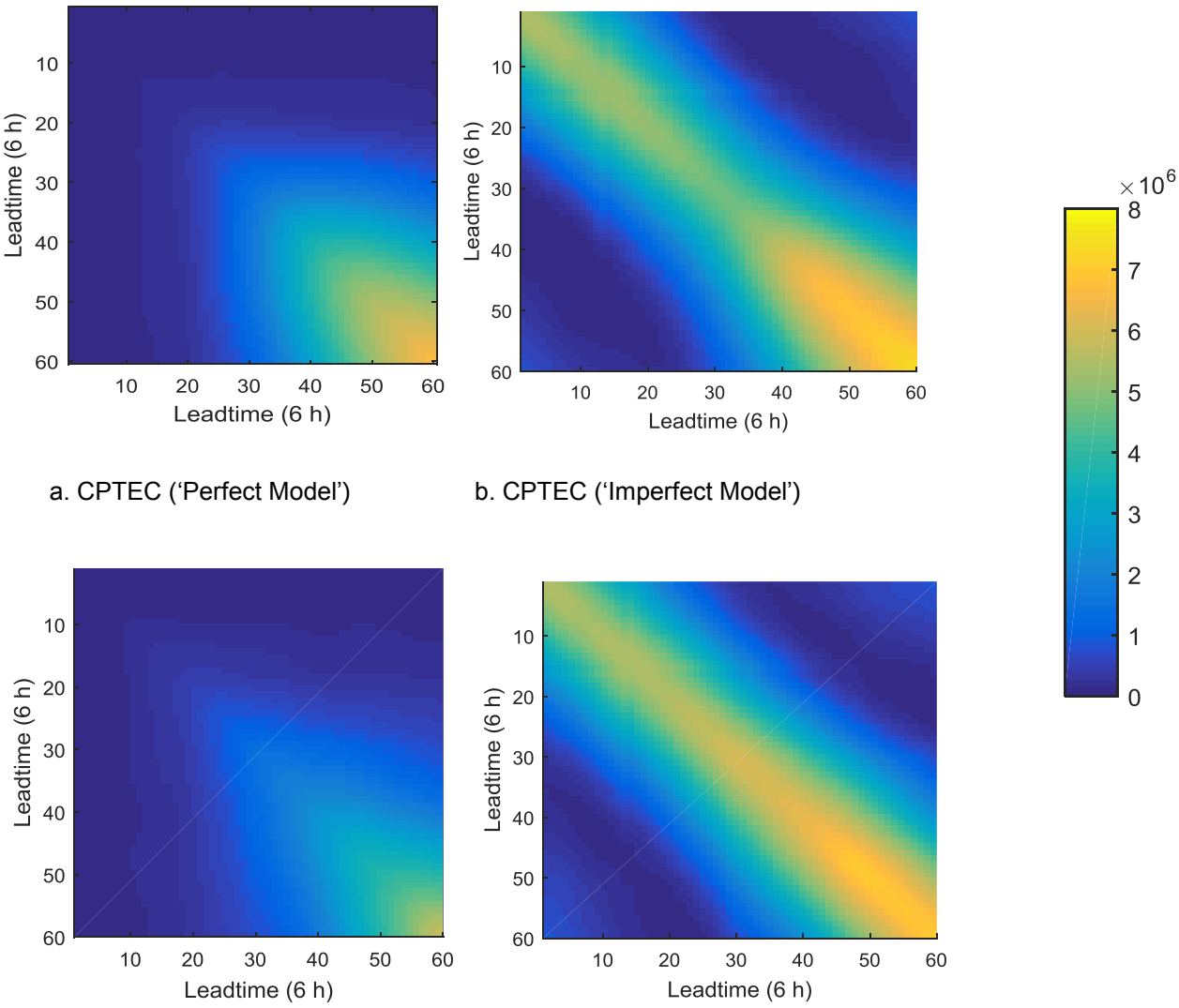


Figure 23 Forecasted inflow to Salto Grande RMSE comparison (Left: observed inflow from HBV using observed rainfall ('perfect model'), Right: observed inflow from telemetry ('imperfect model'))

As expected from the comparison of RMSE, the covariance matrix does not present a clear indication on which forecast product is more accurate. The differences in variance between products are small and not uniform across leadtimes and scenarios as shown Figure 24. As with the RMSE analysis, the difference between the 'perfect model' and 'imperfect model' scenario is noticeably significant.



a. CPTEC ('Perfect Model') b. CPTEC ('Imperfect Model')

c. ECMWF ('Perfect Model') d. ECMWF ('Imperfect Model')

Figure 24 Plot of covariance matrices for all forecast information levels

Considering the results of the forecast accuracy analysis, it is expected that there will be very small (if any) difference between the forecast products in the value of information calculation. The differences in variance is very small, therefore once bias is removed (which is assumed in this thesis), it is expected that the performance of the CPTEC and ECMWF will be very similar.

On the other hand, the difference between the 'perfect model' and 'imperfect model' scenarios are significant, particularly in the earlier timesteps. It is expected that this will have a noticeable effect in the value of information calculations.

4.2 Methodology verification

4.2.1 Tree generation algorithm

This section presents a verification of the tree generation algorithm described in Section 3.2.7 to confirm on whether having foresight obtained from a forecast does indeed make the branching point of a decision tree arrive earlier

A simple scenario is used where there are three possible disturbance (inflow) trajectories. These three possibilities are identical up until the 20th and 40th timestep, where they will diverge abruptly and significantly, as shown in Figure 25.

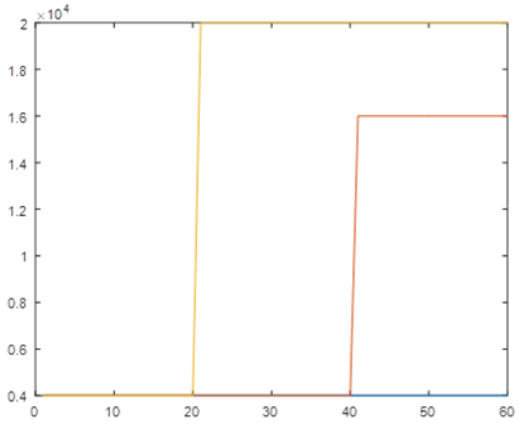


Figure 25 Dummy ensemble for verification

For the base case, the variance of measurement uncertainty is assumed to be a scalar with a value of 250,000. For the forecast scenario, the variance covariance matrix of uncertainty is an identity matrix with 250,000 on the diagonal, up to the 10th timestep, and 999,999,999, for the remaining timesteps.

The expected behaviour is that in the tree structure of the base case would bifurcate when the ensemble diverges, whereas in the forecast scenario, that bifurcation point would occur 10 timesteps earlier.

The result confirms this expectation, as seen in the distinguishability matrices and resulting parent-branch relationships shown in Table 7 to Table 10.

Table 7 Distinguishability matrix of the Base Case – Tree Generation Algorithm Verification

Ensemble member	1	2	3
1	N/A	40	20
2	40	N/A	20
3	20	20	N/A

Table 8 Distinguishability matrix of the Forecast Scenario – Tree Generation Algorithm Verification

Ensemble member	1	2	3
1	N/A	31	11
2	31	N/A	11
3	11	11	N/A

Table 9 Parent--Branch relationship of the Base Case – Tree Generation Algorithm Verification

Ensemble member	1	2	3
Parent	**1**	**1**	**1**
Branch	**1**	**21**	**41**

Table 10 Parent - Branch Relationship of the Forecast Scenario – Tree Generation Algorithm Verification

Ensemble member	1	2	3
Parent	**1**	**1**	**1**
Branch	**1**	**11**	**31**

4.2.2 Optimisation verification

This section presents the verification of the optimisation algorithm used in the thesis. The aim is to confirm if an earlier branching point will result in a lower (more optimal) objective function score.

This verification uses a simple scenario where the ensemble produced in scenario number 28 is reduced to form two ensemble members using the scenario reduction algorithm of this thesis. As seen in Figure 26, inflows of both ensemble members are generally very similar up to timestep 10, before diverging slightly between timestep 10 and 25, and more significantly from timestep 25 onwards.

The prior probability of the low flow (red line) and high flow (blue line) ensemble members are 73.7% and 26.3% respectively.

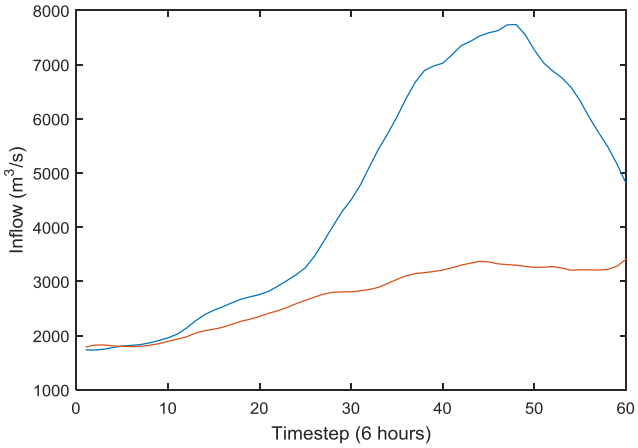


Figure 26 Reduced ensemble (2 member) of scenario 28 - Optimisation algorithm verification

Five tree structures are artificially created, each with a branching point of 1, 10, 20, 25, 30.

Optimisation was conducted assuming an initial reservoir level of 35 m above datum. The result is shown in Table 11, confirming that earlier the branching point, the more optimal the resulting control action.

Table 11 Objective function score for different branching points – Optimisation Verification

Branching point	Optimised objective function score
1	4.1801
10	4.1810
20	4.2704
25	4.5572
30	5.8345

The result is as expected. Because there is virtually no divergence between timestep 1 and 10, the resulting score of the two trees are very similar. The score increases slightly as the branching point increases to timestep 20, and further to timestep 25. The score increases more significantly with a branching point at timestep 30, where divergence between the two ensemble members becomes much larger.

The plots in Appendix C illustrates why this is the case. Poor scoring trees (i.e. branching point 25, and 30) are constrained in making more optimal controls between the initial timestep and the branching point. The more constraints exist, the worse the performance will become. An illustrative example of this is available in Section 2.1.4.

4.3 Value of information analysis

Table 12 shows the probability weighted quasi-optimised objective function score (V*) across all levels of information (base case, and all forecast scenarios). As expected, all forecast scenarios produce a lower objective function score, illustrating the advantage of having forecasts in the operation of a reservoir. The 'perfect model' scenarios also produce a higher score compared to the 'imperfect model' scenario illustrating the advantage of having a very accurate rainfall runoff model in producing inflow forecasts.

Table 12 Average optimised objective function score for all scenarios and forecast products (V*(I))

	Base Case	'Perfect model'		'Imperfect model'	
		CPTEC	ECMWF	CPTEC	ECMWF
Objective function score	1371	1145	1145	1162	1162

Appendix D shows the full record of the objective function scores generated in this analysis. Considering that there are a large number of trees produced and the even larger number of optimisations conducted, not all results can be included in this report. Only three scenarios are comprehensively documented in Appendix E. Scenarios numbers 4, 29, and 54 is chosen as a subset, the results of which are shown in Table 13.

Table 13 Optimised objective function score for a subset of scenarios (J*(I))

Scen.	Q(0) (m^3/s)	Q(0)-Q(-1) (m3/s)	Probability		Quasi-optimised cost function score				
			Histogram population	%	Base case	Perfect model		Imperfect model	
						CPTEC	ECMWF	CPTEC	ECMWF
4	7273	-1103	2	0.1%	22	20	20	22	23
29	3136	-5	478	22.5%	47	23	23	23	23
54	11410	1092	1	0.0%	83568	77932	77837	78426	78323

Scenario 2 represents a low likelihood occurrence of a fast receding flow, similar to the flows observed after the peak of a flood event. Because of the receding flow after an initial flow that is already below turbine capacity, flood risk and spillway use is very low. Therefore the advantage of having a forecast is not as apparent. In fact, the tree generated for 'imperfect model' (ECMWF) performs slightly worse than the base case. A reason for this is that although most of the bifurcation points of 'imperfect model' (ECMWF) occurs early, there are some bifurcation points occurring later than the base case. This may imply that the high variance of forecast error at the end of the horizon actually provides information that is misleading to the controller. Such information can be detrimental to the performance of the system. The 'perfect model' scenarios also have high variance of forecast error at the end of the horizon, but the detrimental effect of this appears to be outweighed by the benefit of accurate foresight early in the horizon.

Scenario 29 represents a high likelihood occurrence of a steady flow trajectory starting from an observed flow that is close to average. All forecast scenarios appear to perform better than the base case. There isn't a noticeable advantage of having a perfect model. This implies that the high accuracy foresight early in the horizon is not advantageous in such steady conditions.

Scenario 54 represents a fast rising extreme flood event from an initial flow that is already high. The advantage of having a forecast and a perfect model is clearly apparent in this scenario. The 'perfect model' and 'imperfect model' achieves objective function scores that are roughly 7.0% and 6.6% lower respectively from the base case.

Note that the scores for Scenario 54 are much higher than the scores of the other scenarios. This appears to be typical of scenarios with build up to high flows as shown in Appendix D. This is due to the design of the objective function, particularly the fourth subfunction (spill prevention, Equation 29). During high flow events spillway usage is generally high and sustained due to the limited turbine capacity. This would result in very high optimised scores. As a result, these high flow scenarios have a very large influence in determining the weighted average objective function score (V*(I)) despite their small probabilities. Scenarios with receding flows also exist with similar probabilities, but their objective function scores are several orders of magnitude smaller, making their influence to the weighted average score (V*(I)) negligible. This behaviour implies that the objective function as implemented in this thesis considers operational performance in managing floods more important than in managing droughts.

Note also that Scenario 54, along with other high flow scenarios, has a high prevalence of computation issues which result in quasi-optimal solutions being used (refer to Section 3.2.8.2). Therefore, if the optimisation routine is run again, it will result in slightly different objective function scores. Despite this issue, the remarks and findings above remain valid.

5. Conclusions and future work

5.1 Conclusions

The following can be concluded from the analysis conducted in this thesis:

- **Better information result in better performance**
    - This thesis argues that better information would result in earlier branching points of a decision or decision tree, which makes decision making less constrained and therefore increases operational performance
    - Section 3.2.7 describes how better information produces trees with earlier branching points. This is verified in Section 4.2.1
    - Section 2.1.4 illustrates how earlier branching points of a decision tree reduces the constraint to decision making. This is verified in Section 4.2.2
- **'Imperfect model' is still very useful**
    - As discussed in Section 3.2.3, the 'imperfect model' scenario in this thesis carries a more pessimistic assumption of model accuracy and robustness compared to what is likely to occur in practice
    - Despite the inaccuracy of model parametrisation and initialisation implied in the 'imperfect model' it still provides significant improvement in operational performance over the base case
    - The weighted average score of the 'imperfect model' scenario is only slightly higher than the 'perfect model' and much lower than the base case
    - This implies that the accuracy deficiency of the 'imperfect model' scenario only results in a marginally worse performance when the system is operated according to the objective function specified in this thesis
    - Note that as discussed in Section 4.3 the objective function used in this thesis appears to place significant importance in preventing floods, operational performance in receding or low flow conditions doesn't influence V*(I) as much as in rising or high flow conditions
    - In scenarios with receding flows, the J*(I) of the 'imperfect model' can be higher than the base case
    - Therefore the overall value of the 'imperfect model' scenario can be much worse if the objective function is altered so that receding flows become more important
- **Performance of the different forecast products are similar**
    - The initial forecast accuracy analysis presented in Section 4.1 indicates that the RMSE of ECMWF and CPTEC is similar. Any slight differences appear to be driven by the difference of bias of the two products
    - This operation of this system assumes that bias is removed, Once bias is removed, the variance of errors (refer to Figure 24), the resulting decision tree (as seen in Appendix E), and the resulting weighted average score (as seen in Table 12) are very similar, if not identical

5.2 Future work

5.2.1 Value of Perfect Information

The levels of information provided in the thesis are intended to emulate of the range of information likely to be available in practice. Future work should consider incorporating an analysis of 'perfect information'.

Raiffa & Schlaifer (1961) originally introduced the concept of Expected Value of Perfect Information (EVPI), which is the expected cost that the operator is willing to bear in order to utilise perfect information. It can be calculated as the difference in optimised objective function between the current level of information and perfect information (Raso, et al., 2014b).

Implementation of such analysis can be seen in van Andel (2009) and Raso, et al. (2014a). Raso, et al. (2014a) is particularly interesting because it uses a perfect forecast scenario as a benchmark to compare different model predictive control method, with a standard deterministic MPC chosen as a baseline. Therefore the other control methods can be compared by standardising their respective optimised objective function scores between 0% (benchmark, most desirable) and 100% (baseline, least desirable). Such comparison could have been done in this thesis as well, but a perfect forecast scenario was not included in the scope. A perfect forecast scenario could easily be included in future work by assuming that the branching points for all ensemble members occur at the first timestep.

5.2.2 Computation issues in the control optimisation process

As mentioned in Section 3.2.8.2, there is an issue with the optimisation algorithm adopted by this thesis where in some instances, it is unable to find a local minima of the objective function within a reasonable timeframe. As a result the algorithm was given a 15 minute time limit, where the result after it reaches that time limit is considered to be a quasi-optimal solution.

Future studies must consider the risk of such computation issues occurring. Potential solutions to consider include:

- **Linearising the objective function:** Linear programming algorithms can then be used, which can run much quicker than non-linear programming algorithms such as interior-point which is used in this thesis
- **Use a faster computer:** This thesis uses a 2015 model consumer grade laptop for all computations and model runs. The computations required for such study may warrant a faster computer
- **Refine optimisation method:** Due to the nature of the interior-point algorithm, outflows can be optimised down to very small increments ($<10^{-5}$ m^3/s). In reality, actuating such control action is unlikely to be that precise. Increasing the tolerance of the interior-point algorithm may be useful in reducing the time requirement of the optimisation process

5.2.3 A more tangible objective function

The objective function used in this thesis is directly taken from Raso, et al. (2014a). It was designed to induce a behaviour from the optimisation algorithm which maximises hydropower generation while minimising flood damage and structural wear and tear. However, the optimised objective function score produced does not represent any tangible value in reality, even though it was reported in this thesis.

Future work may consider implementing an objective function which utilises more tangible units. The ultimate aim is to be able to compare the value of information to the cost of obtaining that information. In order to do this, the objective function score must be quantified in financial terms or 'dollar value'

For example, the flood damage component may be linked to a simple damage model, which correlates flood height to an estimated financial cost. The hydropower generation component can also be refined by implementing a ‘Revenue per kW generated’ modifier. Hejazi & Cai (2011) provides an example in how such objective functions can be implemented.

6. Bibliography

Ahmed, J. A. & Sarma, A. K., 2007. Artificial neural network model for synthetic streamflow generation. *Water Resources Management,* 21(6), p. 1015–1029.

Bergström, S., 1995. The HBV model. In: V. P. Singh, ed. *Computer models of watershed hydrology.* Colorado: Water Resources Publications, pp. 443-476.

Bergström, S., 2014. *List of publications and documents on research on and applications of the HBV model,* s.l.: SMHI.

Birge, J. R. & Louveaux, F., 2011. *Introduction to Stochastic Programming.* 2 ed. s.l.:Springer Science & Business Media.

Boucher, M.-A., Ramos, M.-H. & Zalachori, I., 2014. *On the economic value of hydrological ensemble forecasts.* [Online]
Available at: http://hepex.irstea.fr/economic-value-of-hydrological-ensemble-forecasts/

Buizza, R., Leutbecher, M. & Isaksen, L., 2008. Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society 134*, pp. 2051-2066.

Bureau of Meteorology, 2013. *National Arrangements for Flood Forecasting and Warning,* Melbourne: Bureau of Meteorology, Commonwealth of Australia.

Chaer, R. & Monzon, P., 2008. *Stability conditions for a Stochastic Dynamic Optimizer for optimal dispatch policies in power systems with hydroelectrical generation.* Transmission and Distribution Conference and Exposition: Latin America, IEEE/PES.

Chen, Y.-C., Hsu, Y.-C. & Kuo, K.-T., 2013. Uncertainties in the Methods of Flood Discharge Measurement. *Water Resources Management,* Issue 27, pp. 153-167.

Clemens, F. H., 2001. *Hydrodynamic Models in Urban Drainage: Application and Calibration,* Delft: Delft University Press.

CSIRO, 2012. *Climate and water availability in south-eastern Australia: A synthesis of findings from Phase 2 of the South Eastern Australian Climate Initiative (SeaCI),* s.l.: CSIRO, Australia.

Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P. & Meester, L. E., 2005. *A Modern Introduction to Probability and Statistics: Understanding Why and How.* Delft: Springer.

Delgoda, D. K., Saleem, S. K., Halgamuge, M. N. & Malano, H., 2012. Multiple Model Predictive Flood Control in Regulated River Systems with Uncertain Inflows. *Water Resource Management.*

Despax, A. et al., 2016. Considering sampling strategy and cross-section complexity for estimating the uncertainty of discharge measurements using the velocity-area method. *Journal of Hydrology,* Issue 533, pp. 128-140.

Di Baldassarre, G. & Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth Systems Sciences,* Volume 13, pp. 913-921.

DSE, 2011. *Guidelines for the development of a water supply and demand strategy,* Melbourne: The State of Victoria Department of Sustainability and Environment.

ECMWF, 2007. *TIGGE - the THORPEX Interactive Grand Global Ensemble.* [Online]
Available at: http://tigge.ecmwf.int/
[Accessed 11 11 2015].

ECMWF, 2015. *Tariffs.* [Online]
Available at: http://www.ecmwf.int/en/forecasts/accessing-forecasts/payment-rules-and-options/tariffs

Global Water, 2015. *2015 GWI Price List.* [Online]
Available at: http://www.globalw.com/PriceList/GWI_Price_List.pdf

Goninon, T., Pretto, P., Smith, G. & Atkins, A., 1997. Estimating the Economic Costs of Hydrologic Data Collection. *Water Resources Management 11,* pp. 283-303.

Growe-Kuska, N., Heitsch, H. & Römisch, W., 2003. *Scenario reduction and scenario tree construction for power management problems.* Bologna, Power Tech Conference Proceedings, IEEE.

Hejazi, I. M. & Cai, X., 2011. Building more realistic reservoir optimization models using data mining – A case study of Shelbyville Reservoir. *Advances in Water Resources,* Volume 34, p. 701–717.

ICOLD, n.d. *Register of Dams.* [Online]
Available at: http://www.icold-cigb.org/GB/World_register/general_synthesis.asp?IDA=210
[Accessed 20 June 2016].

ILEC, 1999. *World Lake Database - International Lake Environment Committee Foundation..* [Online]
Available at: http://wldb.ilec.or.jp/Lake.asp?LakeID=SAM-12
[Accessed 20 10 2015].

Kangrang, A. & Lokham, C., 2013. Optimal Reservoir Rule Curves Considering Conditional Ant Colony Optimization with Simulation Model. *Journal of Applied Sciences,* Volume 13, pp. 154-160.

Kearney, M., Dower, P. M. & Cantoni, M., 2011. *Model predictive control for flood mitigation: A Wivenhoe Dam case study.* Melbourne, Engineers Australia, pp. 290-296.

Lebanon, G., 2010. *Bias, Variance, and MSE of Estimators,* Atlanta: Georgia Tech.

Lenilucho, 2010. *Represa Salto Grande.* [Art] (Wikimedia Commons).

Lettenmaier, D. P. & Burges, S. J., 1982. Validation of Synthetic Streamflow Models. *Developments in Water Science,* Volume 17, p. 424–444.

Litirico, X. & Fromion, V., 2009. *Modelling and Control of Hydrosystems.* London: Springer-Verlag.

Luenberger, D. & Ye, Y., 2008. *Linear and Nonlinear Programming.* 3rd ed. Springer Science+Business Media, LLC: New York.

Lund, J., 1996. *Developing Seasonal and Long-term Reservoir System Operation Plans using HEC-PRM,* Davis: Hydrologic Engineering Center U.S. Army Corps of Engineers.

Malaterre, P. O. & Baume, J. P., 1998. *Modeling and regulation of irrigation canals: existing applications and ongoing researches.* San Diego, IEEE Conference on System, Man and Cybernetics,.

McMillan, H., Krueger, T. & Freer, J., 2012. Benchmarking observational uncertainties for hydrology: rainfall. river discharge, and water quality. *Hydrological Processes,* Volume 26, pp. 4078-4111.

MDBA, 2015. *River Operations.* [Online]
Available at: http://www.mdba.gov.au/what-we-do/managing-rivers/river-murray-system/river-operations

Melbourne Water, 2015. *Customers and Prices > Hydrologic data.* [Online]
Available at: http://www.melbournewater.com.au/aboutus/customersandprices/pages/hydrologic-data.aspx

Murphy, A. H., 1988. Skill Scores based on the Mean Squared Error and their relationship to the Correlation Coefficient. *Monthly Weather Review,* December, Volume 116, pp. 2417-2424.

Narasimhan, B., 1996. *The Normal Distribution.* [Online]
Available at: http://statweb.stanford.edu/~naras/jsm/NormalDensity/NormalDensity.html
[Accessed 17 5 2016].

Negenborn, R., Overloop, P.-J. v., Keviczky, T. & Schutter, B. D., 2009. Distributed model predictive control of irrigation canals. *Networks and Heterogenous Media*, pp. 359-380.

Ochoa-Rivera, J. C., García-Bartual, R. & Andreu, J., 2002. Multivariate synthetic streamflow generation using a hybrid model based on artificial neural networks. *Hydrology and Earth System Sciences,* 6(4), pp. 641-654.

Persson, A. & Grazzini, F., 2005. *User Guide to ECMWF Forecast Products,* s.l.: ECMWF.

Raiffa, H. & Schlaifer, R., 1961. *Applied Statistical Decision Theory.* Boston: Harvard University.

Ramadan, H. H., Beighley, R. E. & Ramamurthy, A. S., 2012. Modelling streamflow trends for a watershed with limited data: case of the Litani basin, Lebanon. *Hydrological Sciences Journal,* 57(8), p. 1516.

Raso, L., 2013. *Optimal Control of Water Systems Under Forecast Uncertainty.* Delft: VSSD.

Raso, L., Schwanenberg, D., van de Giesen, N. C. & van Overloop, P. J., 2014a. Short-term optimal operation of water systems using ensemble forecasts. *Advances in Water Resources 71,* pp. 200-208.

Raso, L. et al., 2013. Tree structure generation from ensemble forecasts for real time control. *Hydrological Processes,* Issue 27, pp. 75-82.

Raso, L. & van Overloop, P.-J., 2011. *Model predictive control (MPC) for operational water management.* Vienna, TU Delft.

Raso, L., Werner, M. & Weijs, S., 2014b. *Maximising the Value of Information by Optimal Design: an application using least squares method for rating curve estimation.* s.l.:Unpublished Draft.

Rayner, S., Lach, D. & Ingram, H., 2005. Weather Forecasts are for Wimps: Why Water Resource Managers Do Not Use Climate Forecasts. *Climatic Change,* Volume 69, pp. 197-227.

Sauer, V. B. & Meyer, R. W., 1992. *Determination of Error in Individual Discharge Measurements,* Norcross: U.S. Geological Survey.

Schwanenberg, D. & Becker, B., 2014. *RTC-Tools: Reference Manual,* Delft: Deltares.

Shahverdi, K. & Monem, M. J., 2011. *Construction and Evaluation of PID Automatic Control System for Irrigation Canals at Laboratory Scale.* Tehran, ICID 21st International Congress on Irrigation and Drainage.

Stedinger, J. R. & Taylor, M. R., 1982. Synthetic Streamflow Generation 1. Model Verification and Validation. *Water Resources Research,* 18(4), pp. 909-918.

Stephens, A., Garcia, C., Kusumaatmaja, W. & Ng, M. Y., 2007. *FFControl.* [Online]
Available at: https://controls.engin.umich.edu/wiki/index.php/FFControl
[Accessed 17 11 2015].

Talsma, J., 2015a. *Salto Grande Historic Data and Telemetry (Deltares/Salto Grande)* [Interview] (7 8 2015a).

Talsma, J., 2015b. *Salto Grande storage rating curve* [Interview] (8 10 2015b).

van Andel, S.-J., 2009. *Anticipatory Water Management: Using ensemble weather forecasts for critical events,* Leiden: CRC Press/Balkema.

van Nooyen, R. & van Overloop, P.-J., 2014. *CIE5490 Operational Water Management Lecture 1 2014/15.* [Online].

van Overloop, P.-J., 2006. *Model Predictive Control for Open Water Systems.* Amsterdam: IOS Press.

Weijs, S. V., 2008. *Information content of weather predictions for flood-control in a Dutch lowland water system.* Toronto, Institute for Catastrophic Loss Reduction: 4th International Symposium on Flood Defence.

White, S. et al., 2003. *Urban Water Demand Forecasting and Demand Management: Research Needs, Review and Recommendations,* Sydney: Water Services Association of Australia.

WWRP/WGNE Joint Working Group on Forecast Verification Research, 2015. *Forecast Verification.* [Online] Available at: http://cawcr.gov.au/projects/verification/
[Accessed 4 April 2016].

Zahabiyoun, B., 1999. *Stochastic Generation of Daily Streamflow Data Incorporating Land Use and or Climate Change Effects ,* Newcastle Upon Tyne: University of Newcastle Upon Tyne.

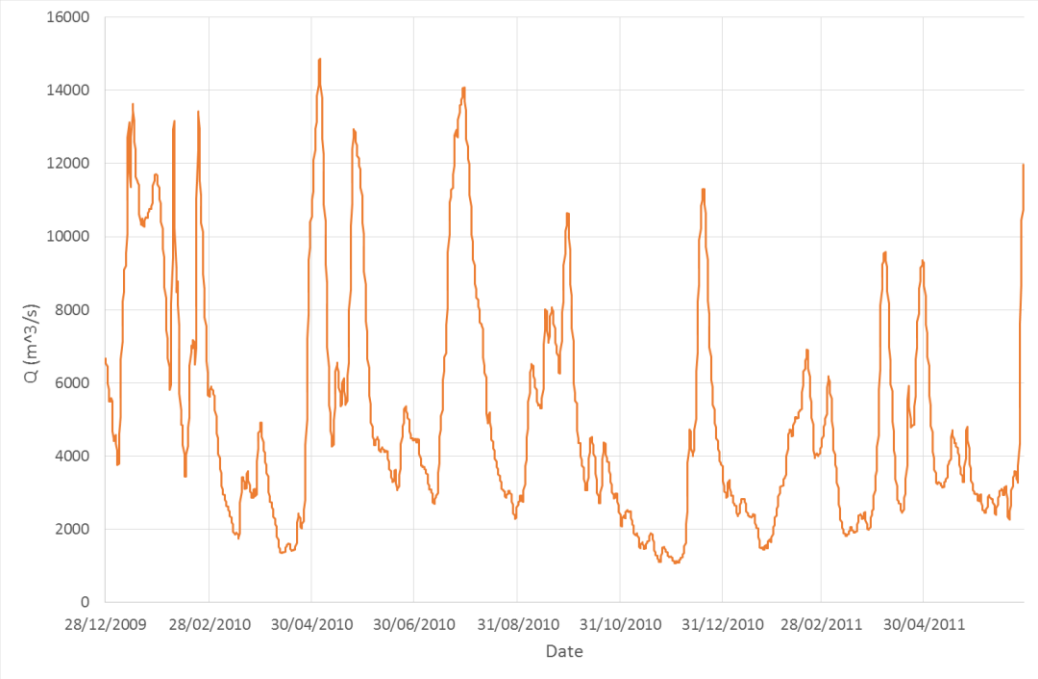
Zealand, C. M., Burn, D. H. & Simonovic, S. P., 1999. Short term streamflow forecasting using artificial neural networks. *Journal of Hydrology,* 214(1-4), pp. 32-48.

Zhu, Y. et al., 2002. The Economic Value of Ensemble-Based Weather Forecasts. *American Meteorological Society,* pp. 73-83.

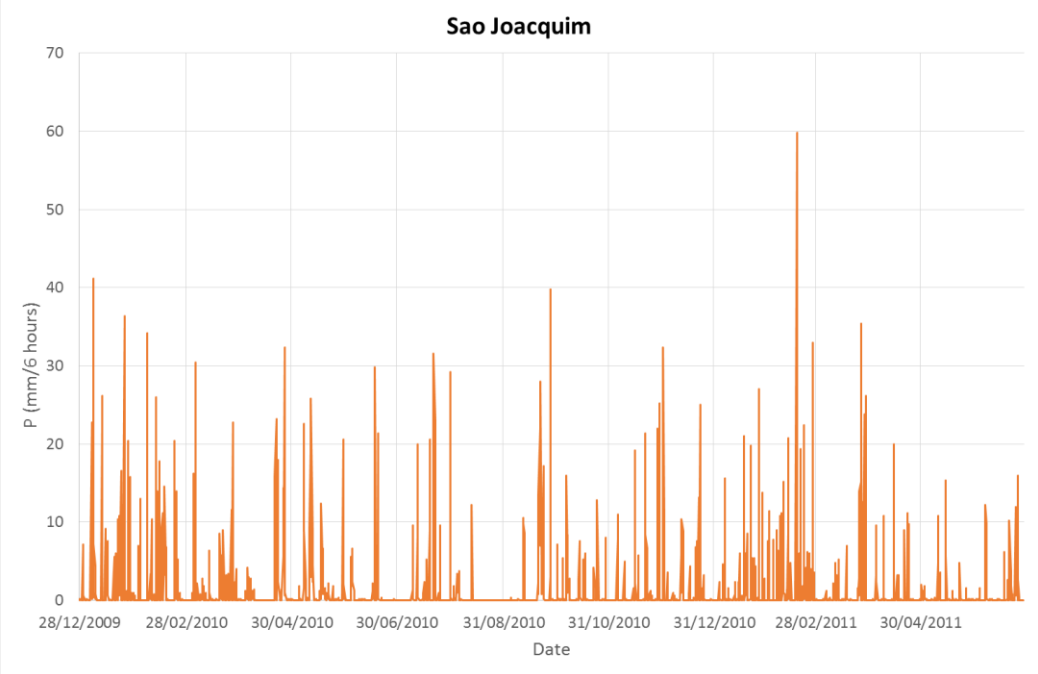
Appendix A. Measured historical streamflow, reservoir level and rainfall plots

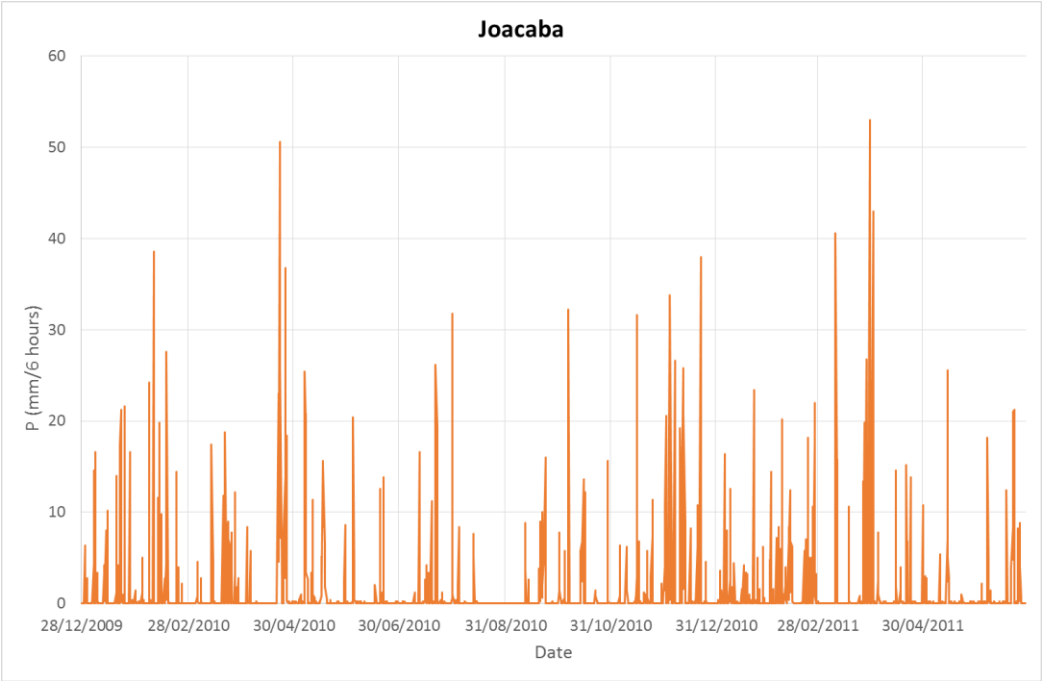
These are obtained from Raso (2013).

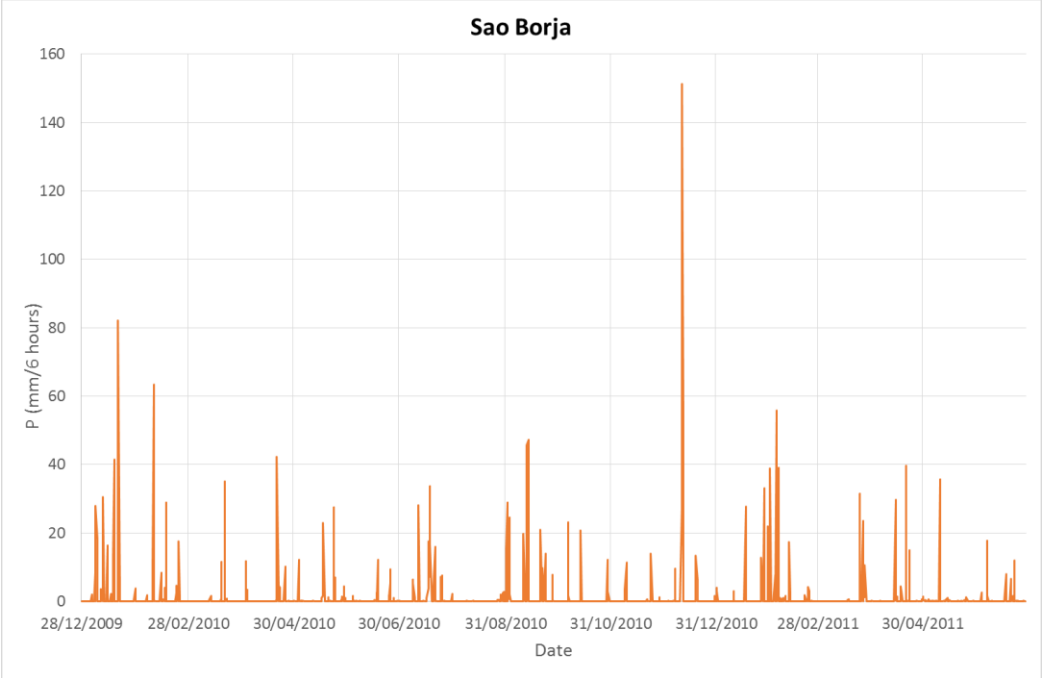
Streamflow (into reservoir)

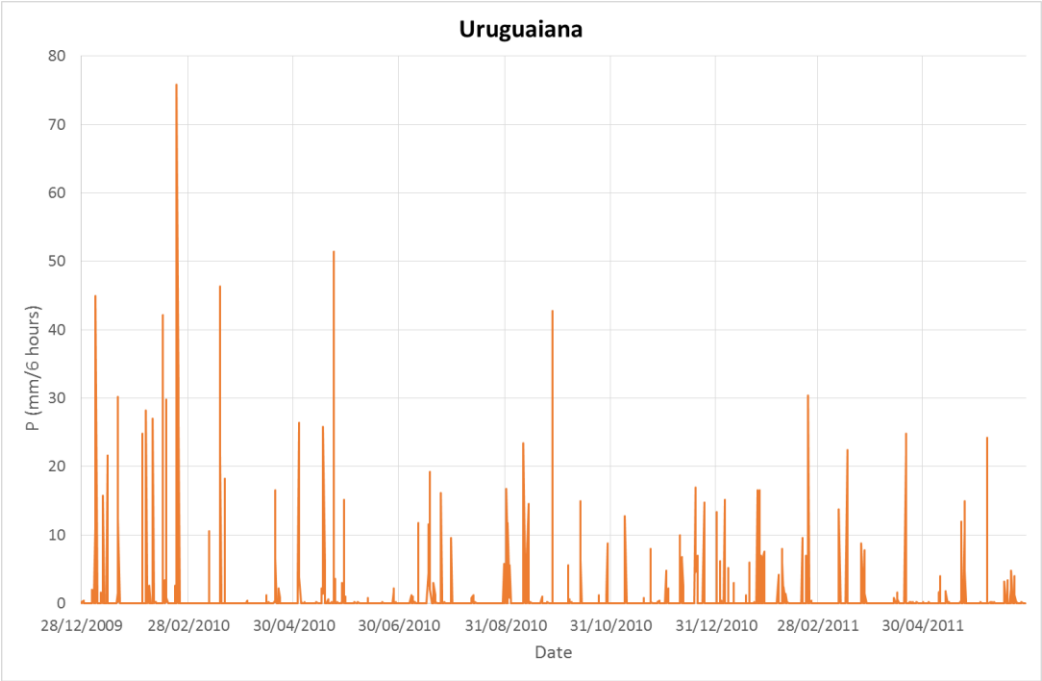


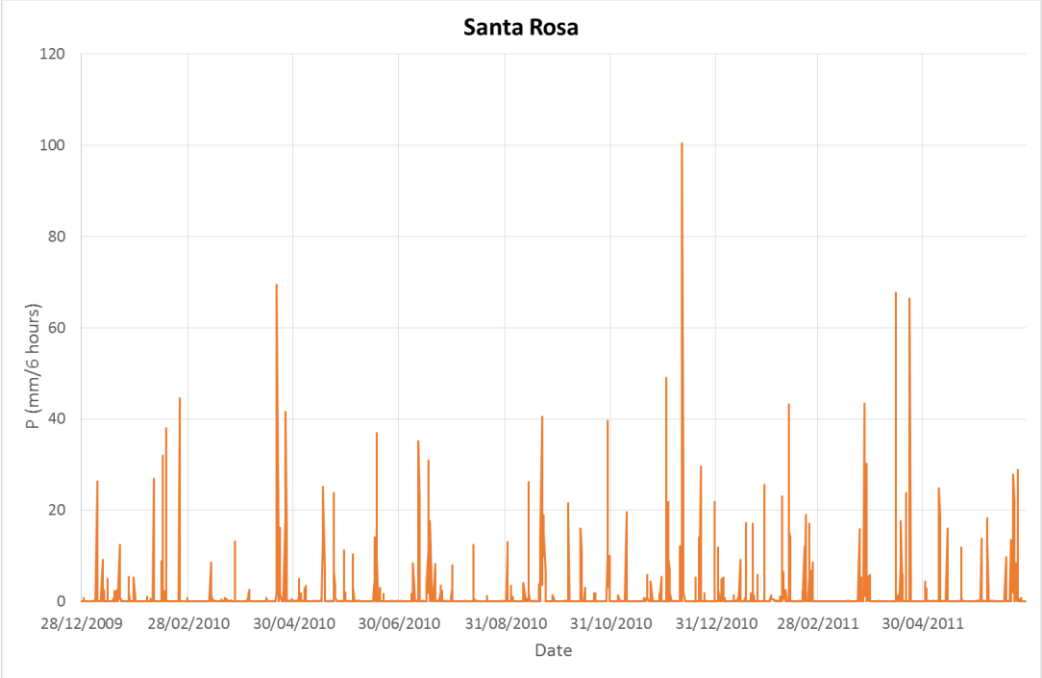
Rainfall

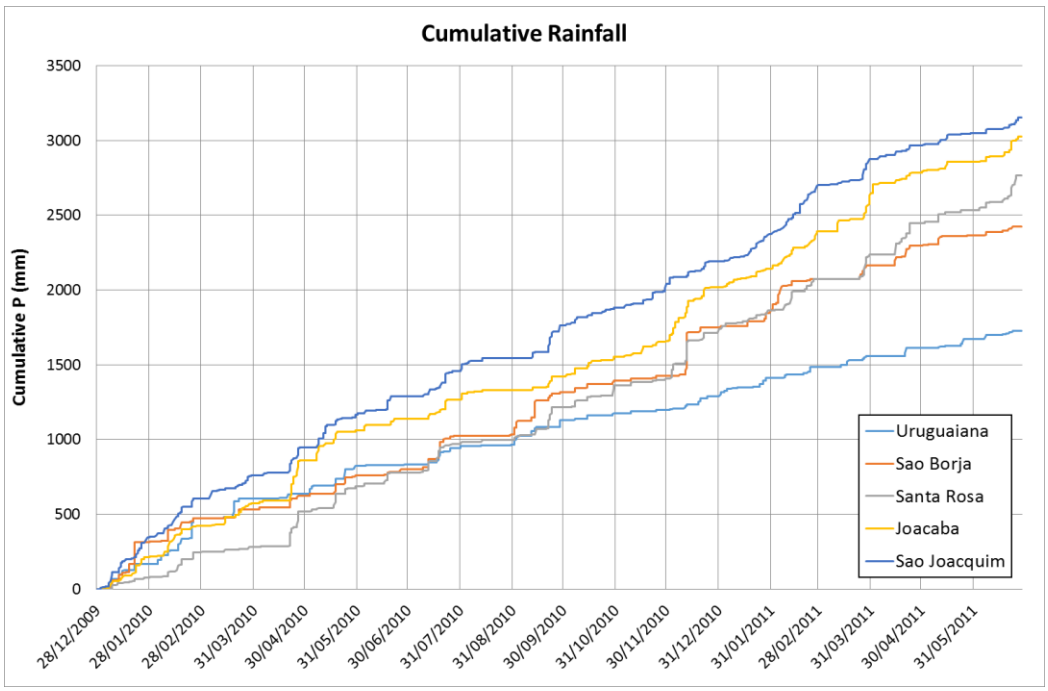



Joacaba
60
50
40
30
20
10
0
P (mm/6 hours)
28/12/2009
28/02/2010
30/04/2010
30/06/2010
31/08/2010
31/10/2010
31/12/2010
28/02/2011
30/04/2011
Date

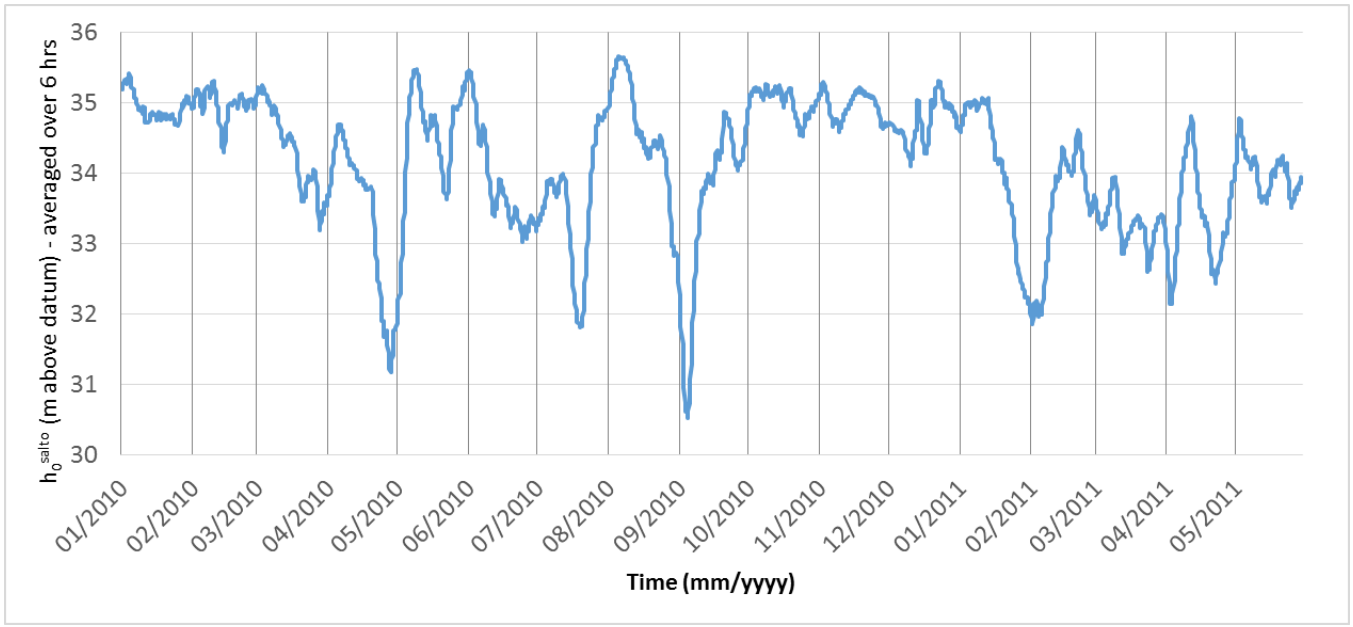

Sao Borja
160
140
120
100
80
60
40
20
0
P (mm/6 hours)
28/12/2009
28/02/2010
30/04/2010
30/06/2010
31/08/2010
31/10/2010
31/12/2010
28/02/2011
30/04/2011
Date


Uruguaiana
80
70
60
50
40
30
20
10
0
P (mm/6 hours)
28/12/2009
28/02/2010
30/04/2010
30/06/2010
31/08/2010
31/10/2010
31/12/2010
28/02/2011
30/04/2011
Date


Santa Rosa
120
100
80
60
40
20
0
P (mm/6 hours)
28/12/2009
28/02/2010
30/04/2010
30/06/2010
31/08/2010
31/10/2010
31/12/2010
28/02/2011
30/04/2011
Date



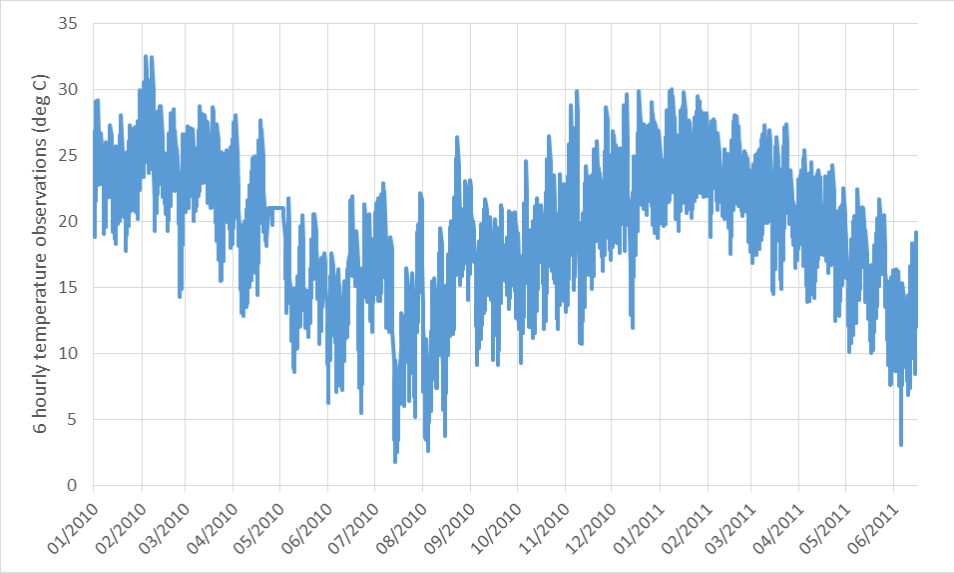
Reservoir level



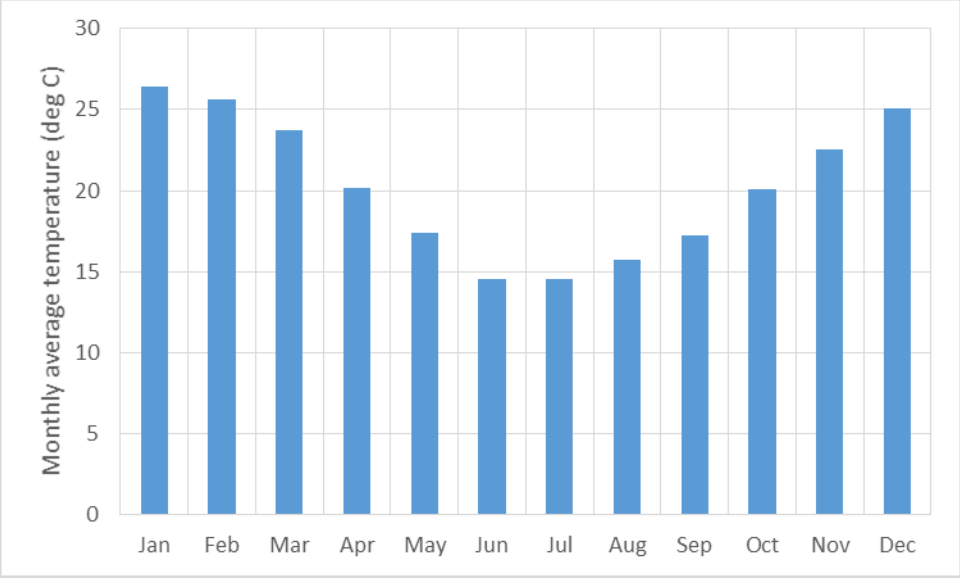
Appendix B. Temperature and evaporation data used in the HBV model

These are obtained from the data used in Raso (2013)

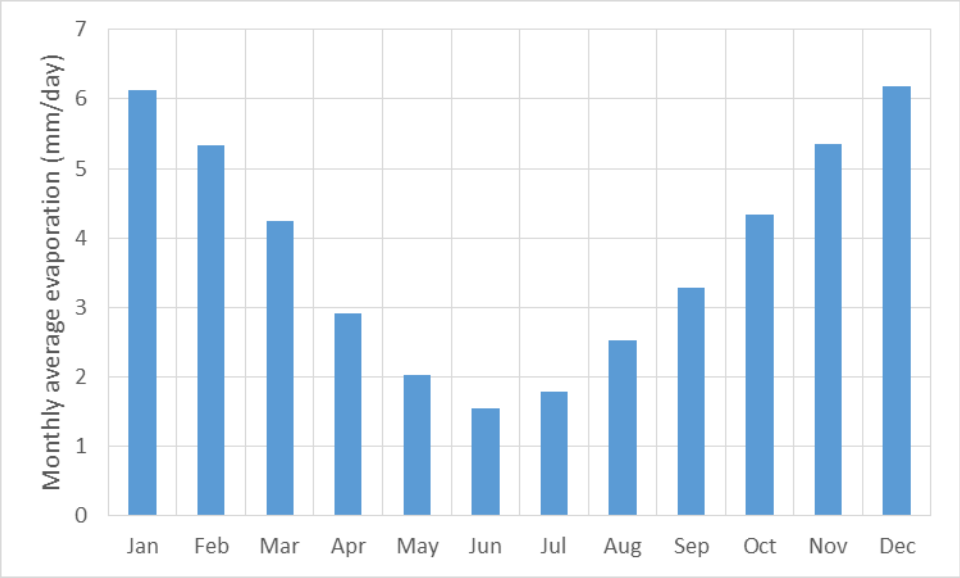
Temperature observations



Monthly average temperature data

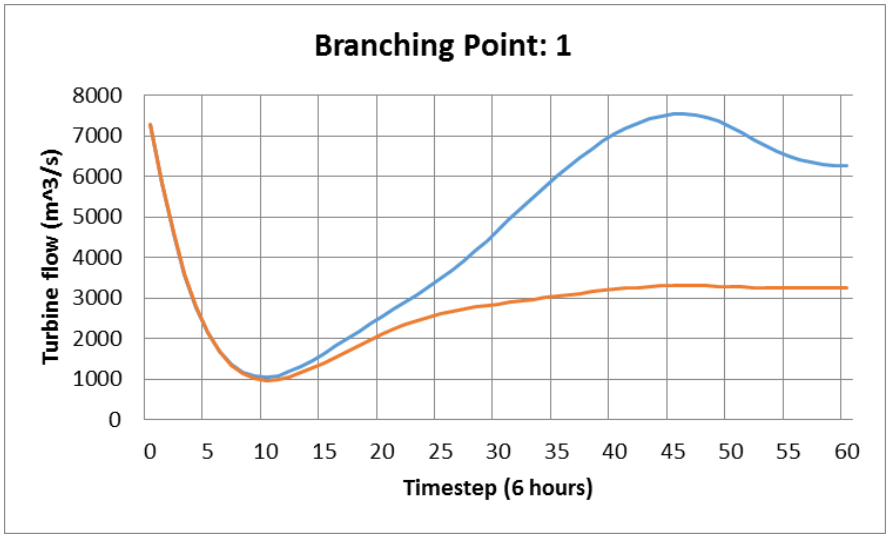


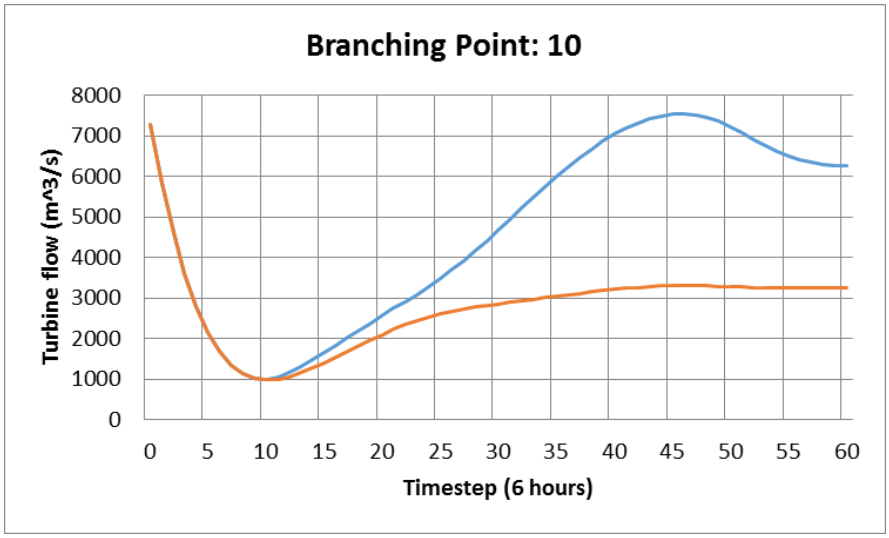
Monthly average evaporation

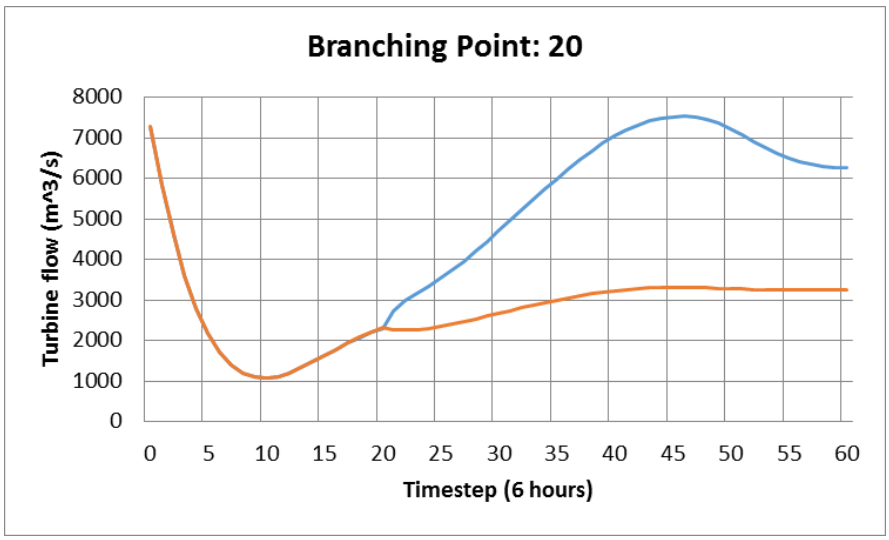


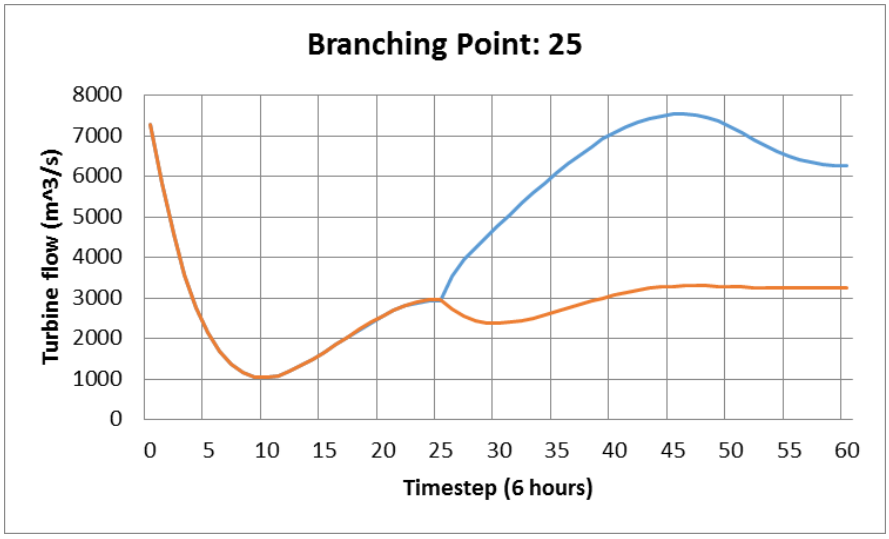
Appendix C. Optimisation Verification Plots

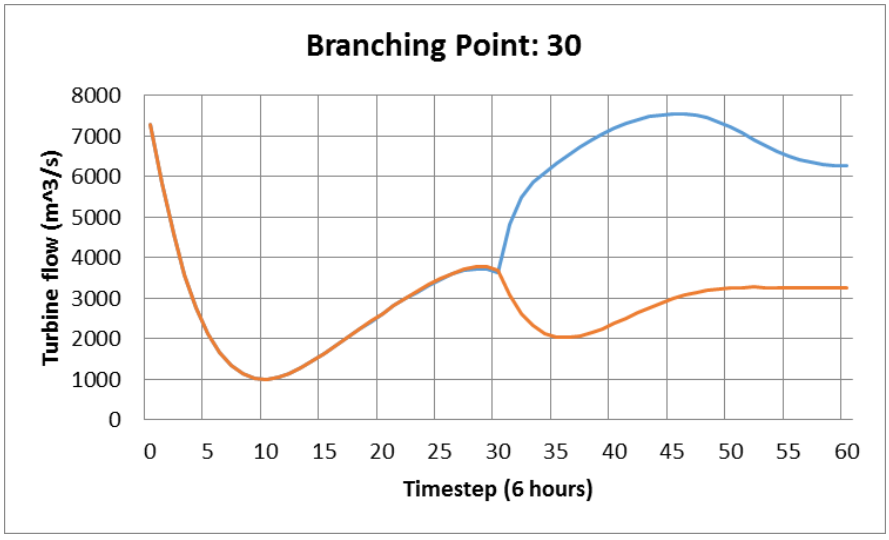
C.1 Turbine flow results



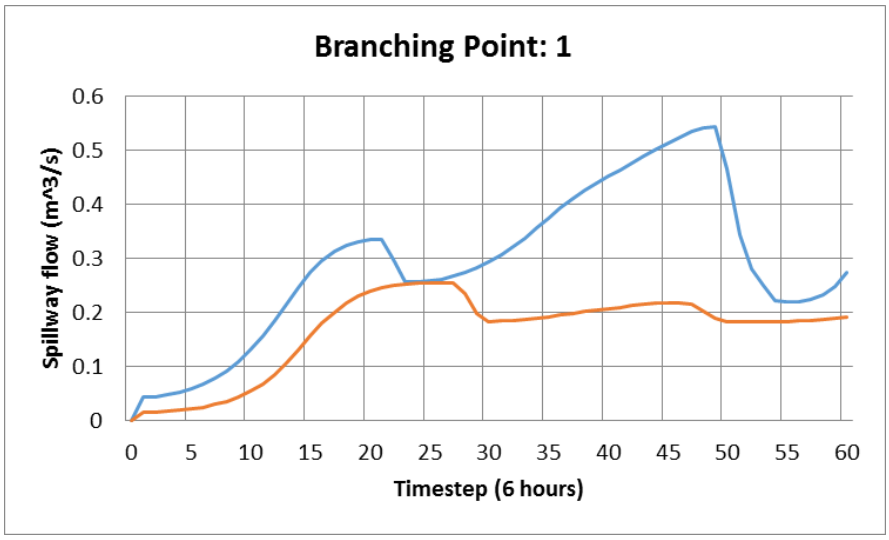


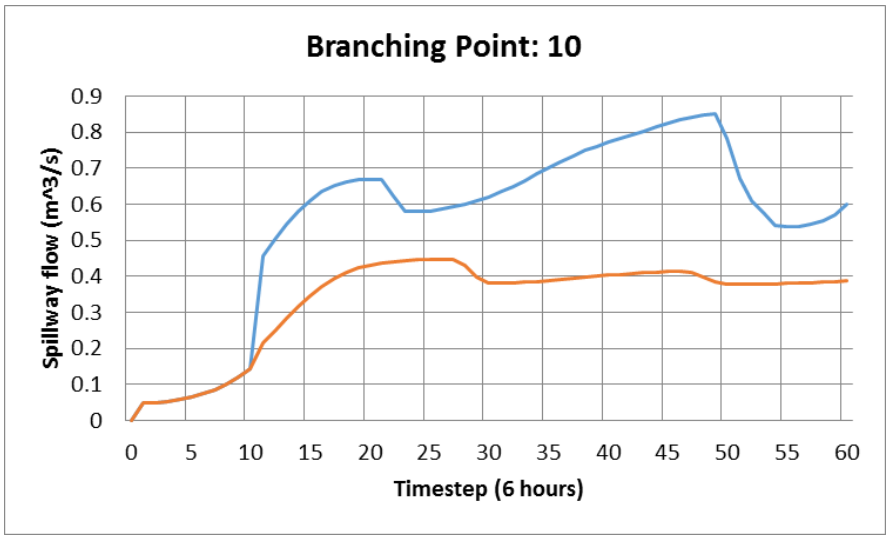

Branching Point: 20
Turbine flow (m^3/s)
8000
7000
6000
5000
4000
3000
2000
1000
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)

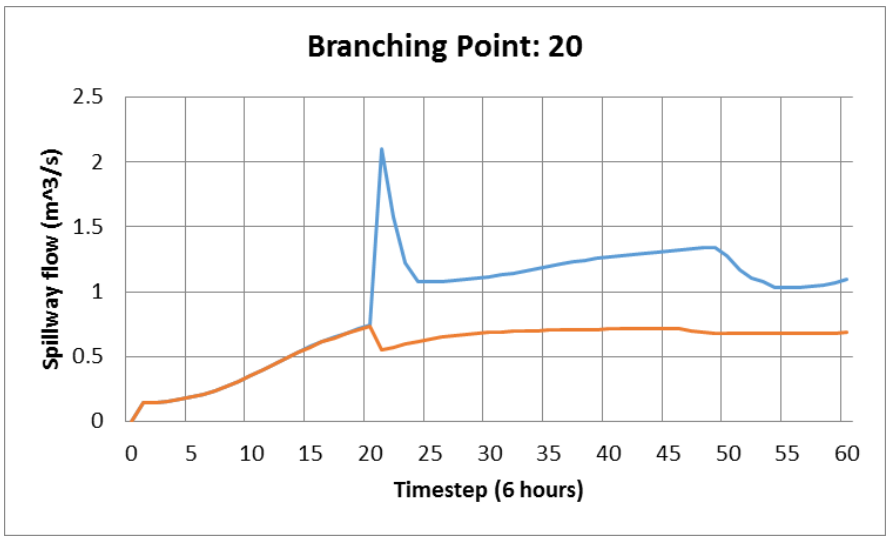

Branching Point: 25
Turbine flow (m^3/s)
8000
7000
6000
5000
4000
3000
2000
1000
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)

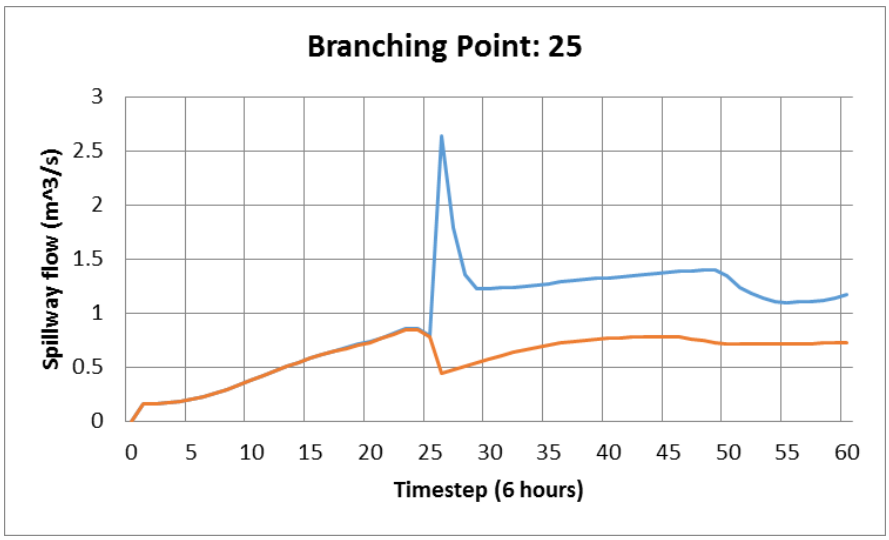

Branching Point: 30
Turbine flow (m^3/s)
8000
7000
6000
5000
4000
3000
2000
1000
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)

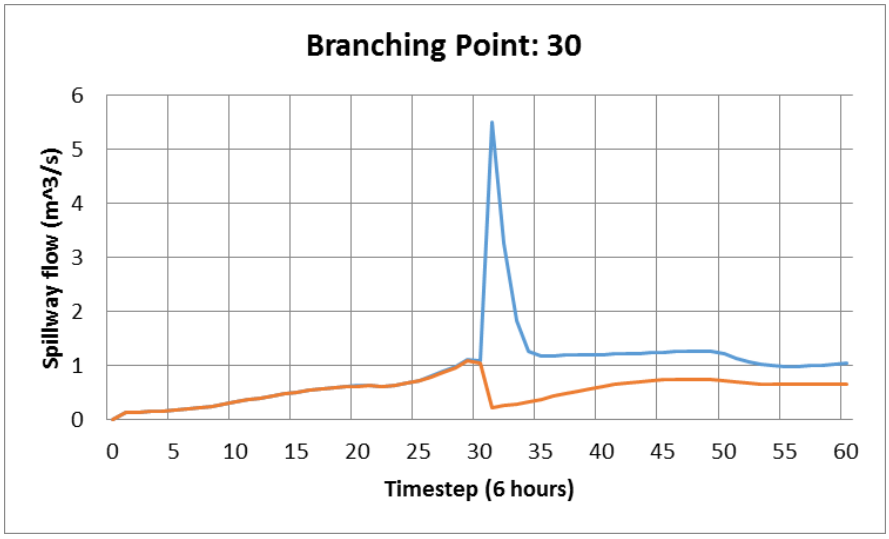
C.2 Spillway flow results



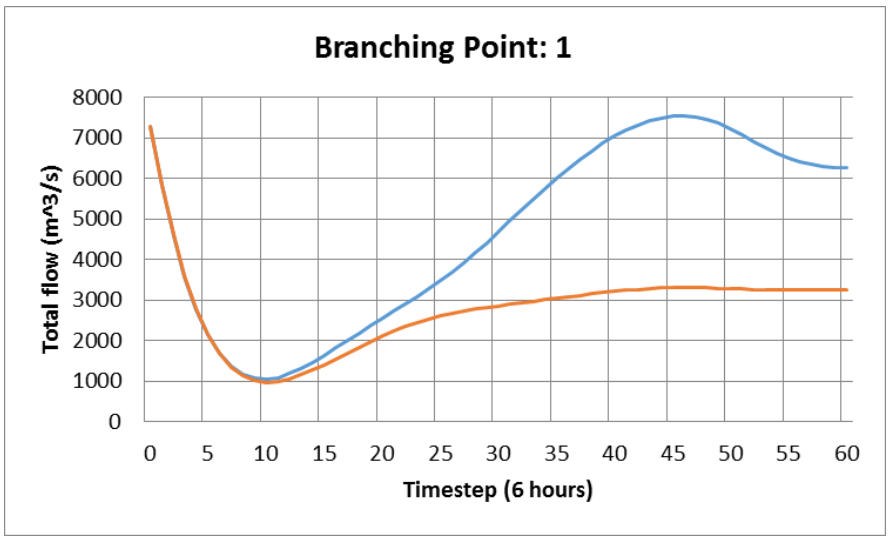


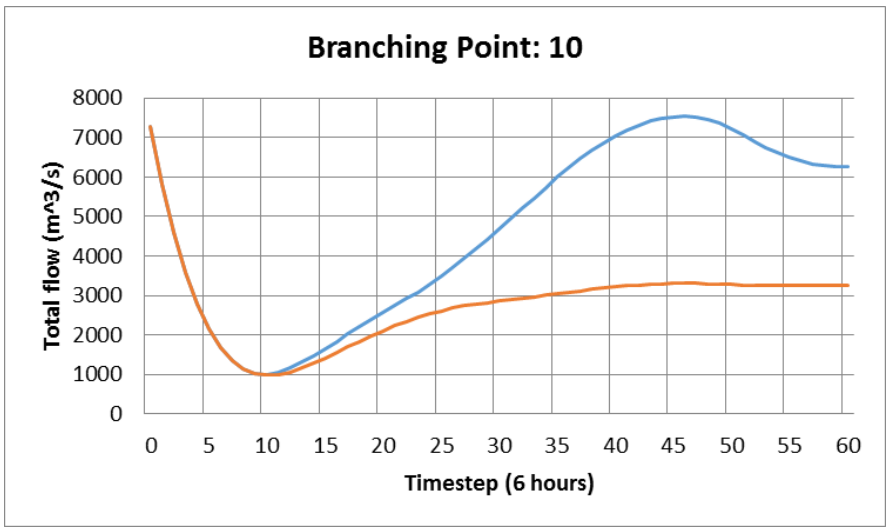


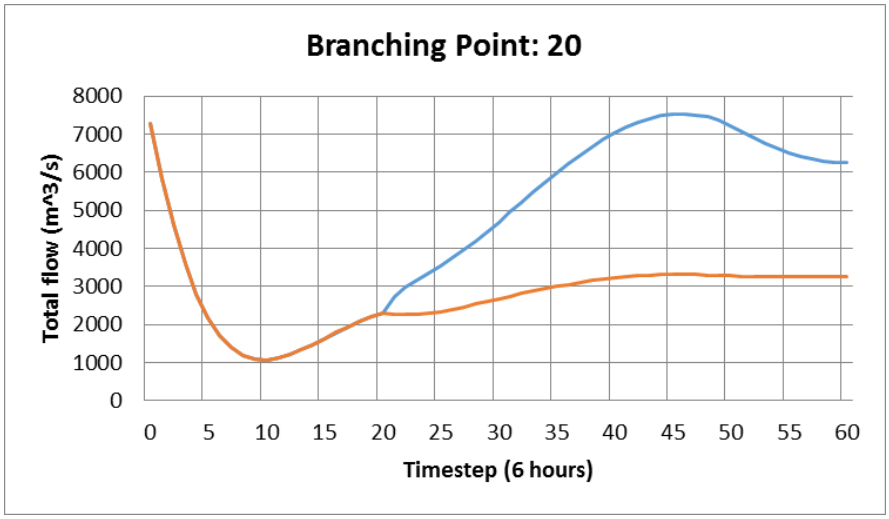


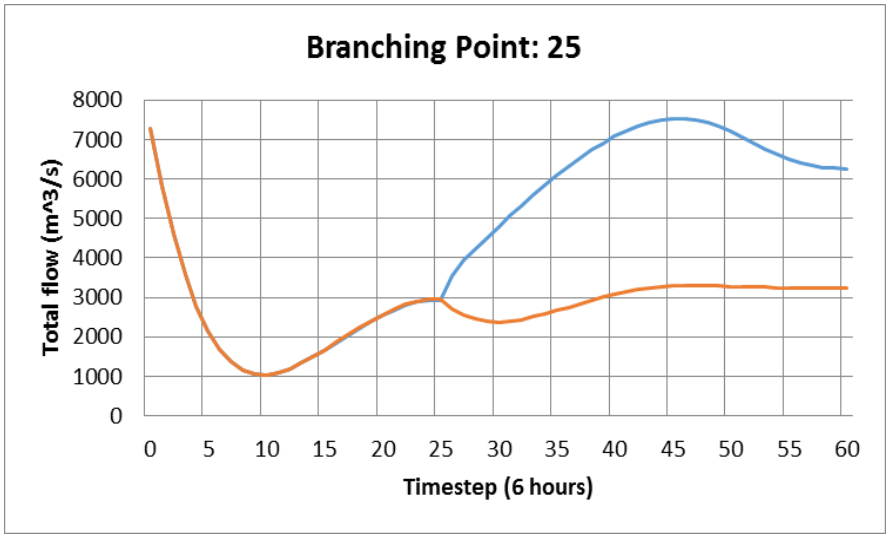


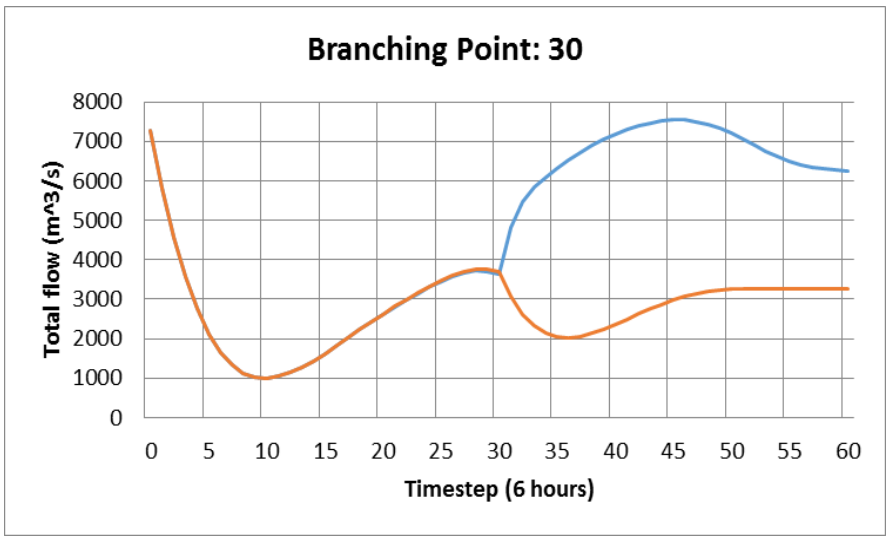
C.3 Total flow results



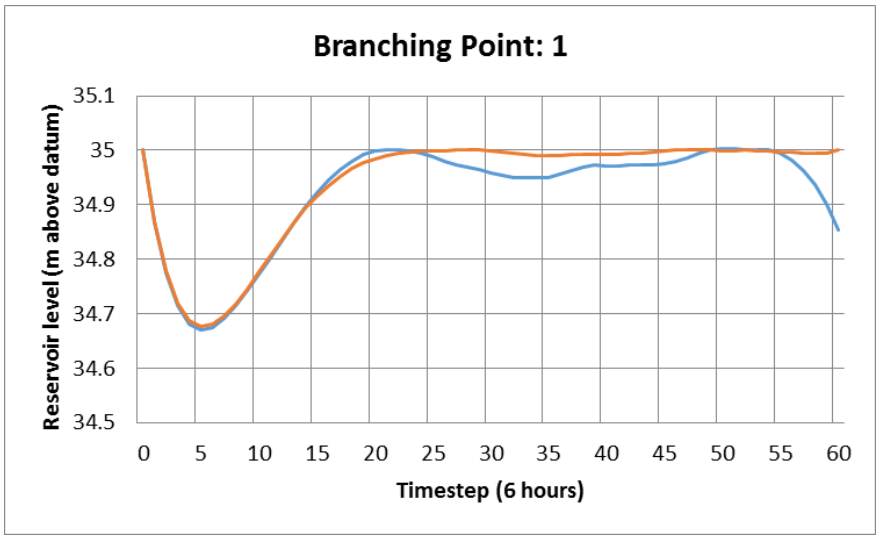

Branching Point: 10
Total flow (m^3/s)
8000
7000
6000
5000
4000
3000
2000
1000
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)

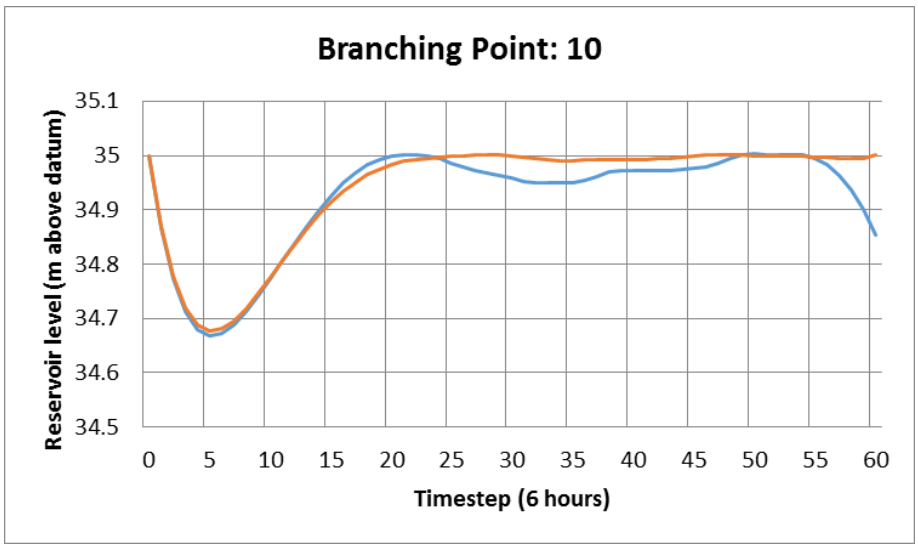

Branching Point: 20
Total flow (m^3/s)
8000
7000
6000
5000
4000
3000
2000
1000
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)

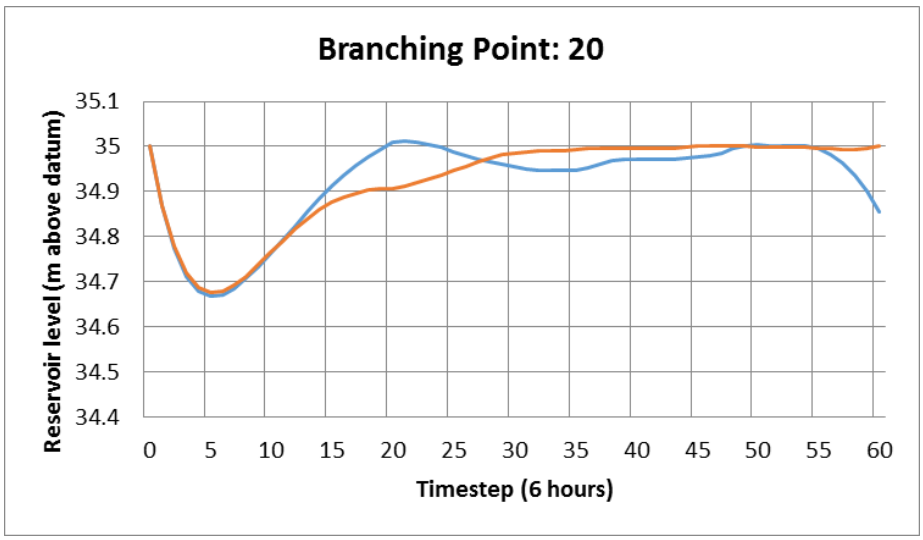

Branching Point: 25
Total flow (m^3/s)
8000
7000
6000
5000
4000
3000
2000
1000
0
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)

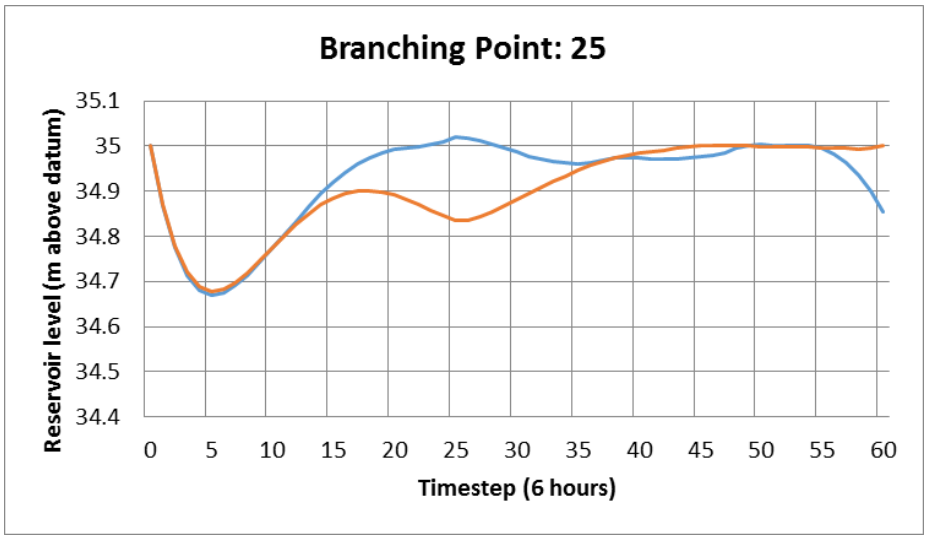


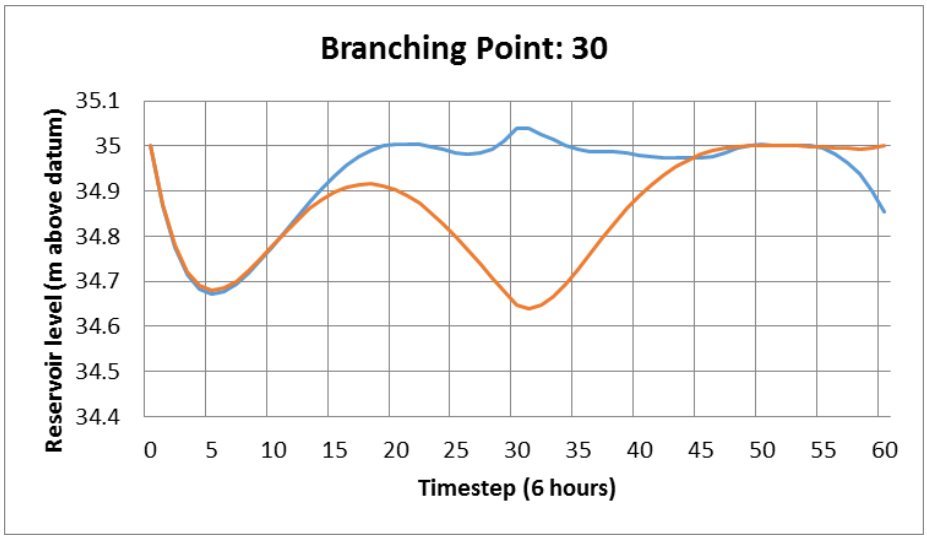
C.4 Reservoir level






Branching Point: 20
Reservoir level (m above datum)
35.1
35
34.9
34.8
34.7
34.6
34.5
34.4
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)


Branching Point: 25
Reservoir level (m above datum)
35.1
35
34.9
34.8
34.7
34.6
34.5
34.4
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)


Branching Point: 30
Reservoir level (m above datum)
35.1
35
34.9
34.8
34.7
34.6
34.5
34.4
0 5 10 15 20 25 30 35 40 45 50 55 60
Timestep (6 hours)

Appendix D. Record of optimisation function score

Scen.	Q(0) (m^3/s)	Q(0)-Q(-1) (m3/s)	Probability		Quasi-optimised cost function score				
			Histogram population	%	Base case	Perfect model		Imperfect model	
						CPTEC	ECMWF	CPTEC	ECMWF
1	8652	-1377	2	0.1%	24	22	22	34	35
2	10031	-1377	2	0.1%	154	149	149	155	155
3	11410	-1377	2	0.1%	474	470	470	471	471
4	7273	-1103	2	0.1%	22	20	20	22	23
5	4515	-828	2	0.1%	38	37	37	42	42
6	5894	-828	3	0.1%	27	22	22	27	27
7	10031	-828	1	0.0%	155	161	161	168	169
8	11410	-828	2	0.1%	479	470	470	470	470
9	12789	-828	2	0.1%	1058	1030	1030	1032	1032
10	3136	-554	6	0.3%	49	48	48	53	53
11	4515	-554	6	0.3%	54	31	31	35	35
12	5894	-554	14	0.7%	24	19	19	21	20
13	7273	-554	13	0.6%	23	17	17	17	17
14	8652	-554	10	0.5%	120	64	64	64	65
15	10031	-554	4	0.2%	304	241	243	245	245
16	11410	-554	5	0.2%	562	515	515	515	515
17	12789	-554	1	0.0%	1307	1203	1203	1214	1212
18	1757	-280	11	0.5%	66	66	66	76	73
19	3136	-280	42	2.0%	32	30	30	31	31
20	4515	-280	50	2.4%	50	29	29	31	35
21	5894	-280	39	1.8%	35	16	16	17	19
22	7273	-280	24	1.1%	46	19	19	20	20
23	8652	-280	18	0.8%	239	120	120	126	126
24	10031	-280	18	0.8%	470	362	366	368	365
25	11410	-280	14	0.7%	1636	1560	1562	1616	1562
26	12789	-280	17	0.8%	2627	2434	2438	2462	2455
27	14168	-280	5	0.2%	5734	5268	5387	5311	5314
28	1757	-5	398	18.7%	35	31	31	32	31
29	3136	-5	478	22.5%	47	23	23	23	23
30	4515	-5	279	13.1%	44	20	20	27	23
31	5894	-5	102	4.8%	67	45	45	126	130
32	7273	-5	39	1.8%	68	30	30	31	31
33	8652	-5	28	1.3%	862	574	570	584	587

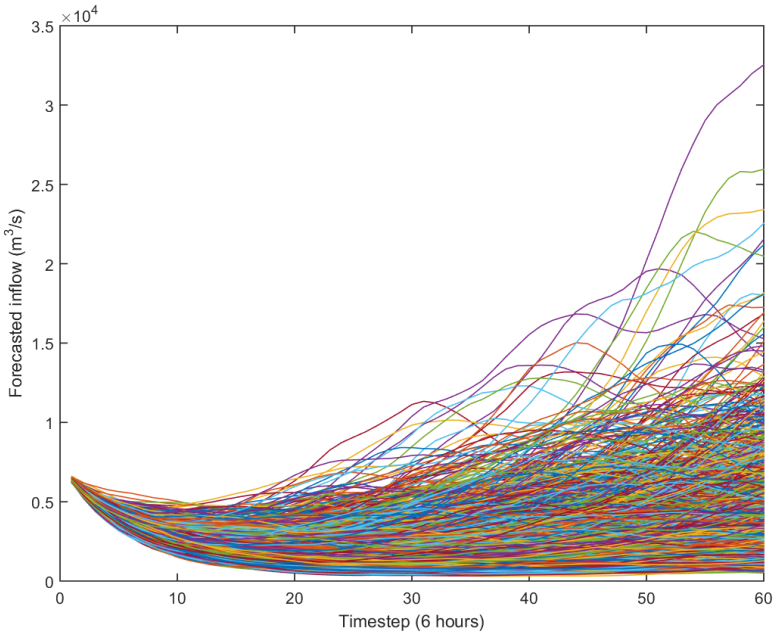
Scen.	Q(0) (m^3/s)	Q(0)-Q(-1) (m3/s)	Probability		Quasi-optimised cost function score				
			Histogram population	%	Base case	Perfect model		Imperfect model	
						CPTEC	ECMWF	CPTEC	ECMWF
34	10031	-5	33	1.6%	2270	1828	1829	1840	1836
35	11410	-5	44	2.1%	2973	2758	2753	2793	2778
36	12789	-5	22	1.0%	5813	4713	4713	4752	4802
37	14168	-5	16	0.8%	8401	7139	7218	7812	7653
38	3136	269	24	1.1%	1924	1418	1420	1351	1448
39	4515	269	69	3.3%	1518	1034	1025	1027	1056
40	5894	269	72	3.4%	1429	1060	1032	1068	1048
41	7273	269	49	2.3%	5400	4472	4483	4484	4506
42	8652	269	40	1.9%	2693	2177	2187	2209	2191
43	10031	269	36	1.7%	4839	4020	4009	4069	4071
44	11410	269	25	1.2%	12362	9057	9067	9129	9173
45	12789	269	18	0.8%	13071	11753	11787	11815	11889
46	14168	269	7	0.3%	18227	16222	16227	16337	16303
47	5894	543	3	0.1%	12396	10645	10603	10661	10702
48	7273	543	6	0.3%	13182	11505	11536	11687	11530
49	8652	543	5	0.2%	15643	13521	13541	13793	13609
50	10031	543	2	0.1%	19436	17035	17012	17196	17082
51	11410	543	4	0.2%	23347	19878	19964	20079	20038
52	12789	543	6	0.3%	27499	24324	24237	24239	24391
53	10031	1092	1	0.0%	93888	89194	89259	89278	89932
54	11410	1092	1	0.0%	83568	77932	77837	78426	78323
55	12789	1092	1	0.0%	90742	86350	86645	86469	86535
				Weighted total score	1371	1145	1145	1162	1162

Appendix E. Detailed results for a subset of scenarios

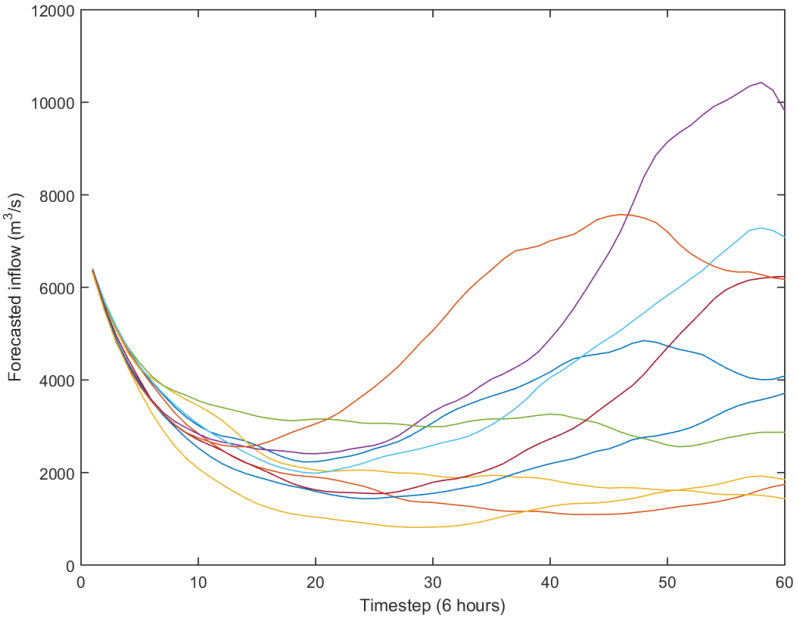
Scen.	Q(0) (m^3/s)	Q(0)-Q(-1) (m3/s)	Probability		Quasi-optimised cost function score				
			Histogram population	%	Base case	Perfect model		Imperfect model	
						CPTEC	ECMWF	CPTEC	ECMWF
4	7273	-1103	2	0.1%	22	20	20	22	23
29	3136	-5	478	22.5%	47	23	23	23	23
54	11410	1092	1	0.0%	83568	77932	77837	78426	78323

E.1 Scenario 4

E.1.1 Trajectory ensemble

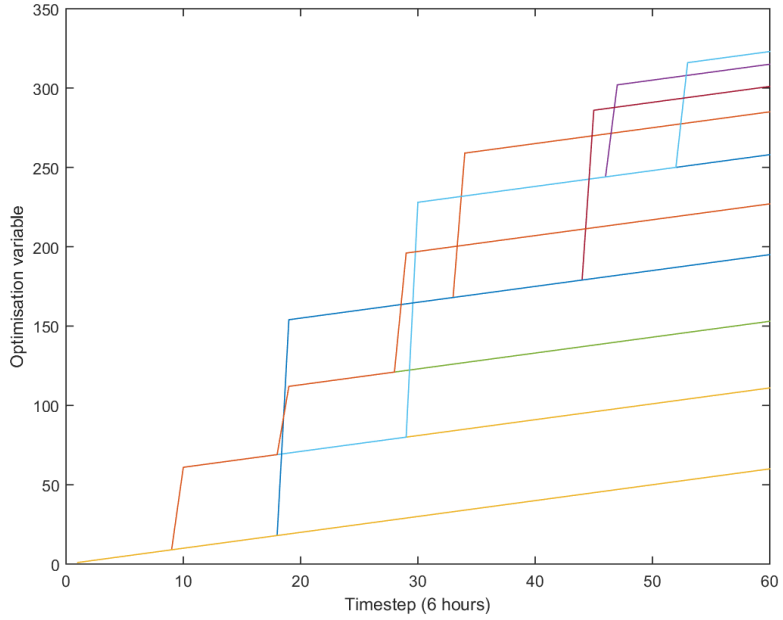


E.1.2 Reduced ensemble

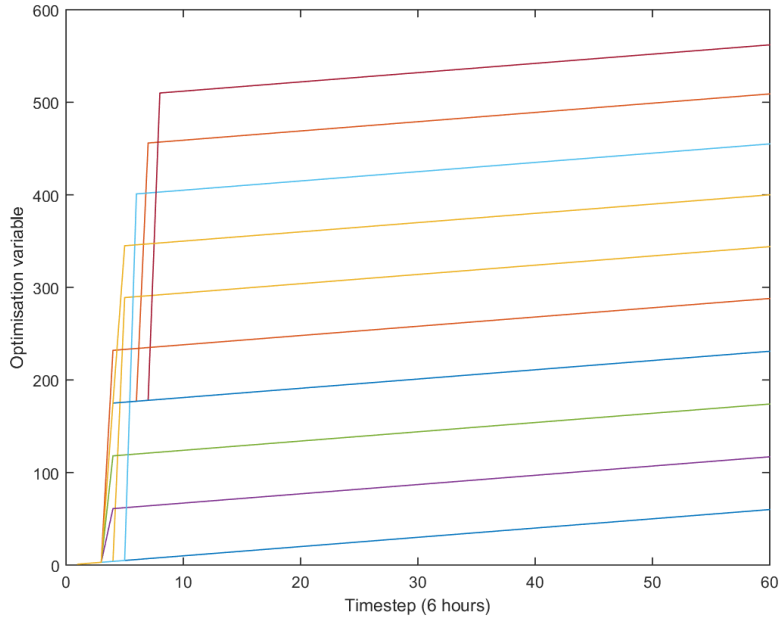


E.1.3 Nodal partition matrix

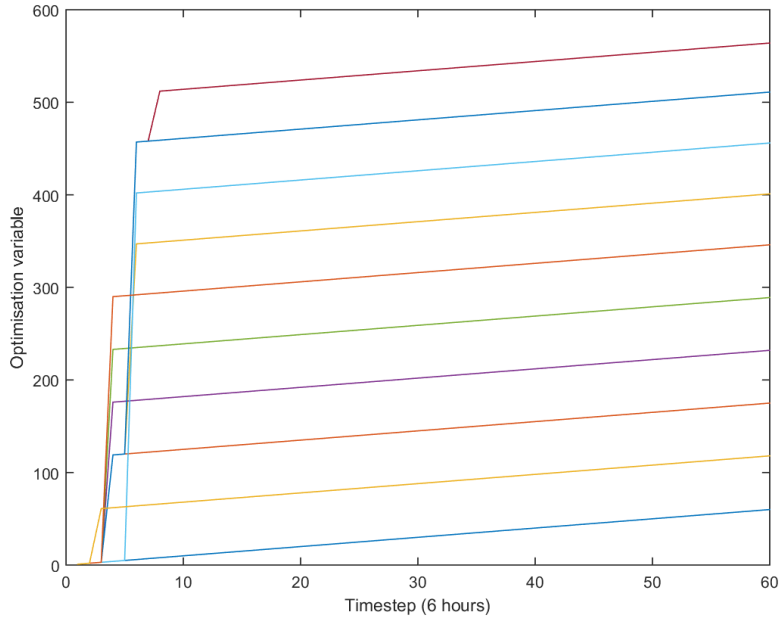
Base Case



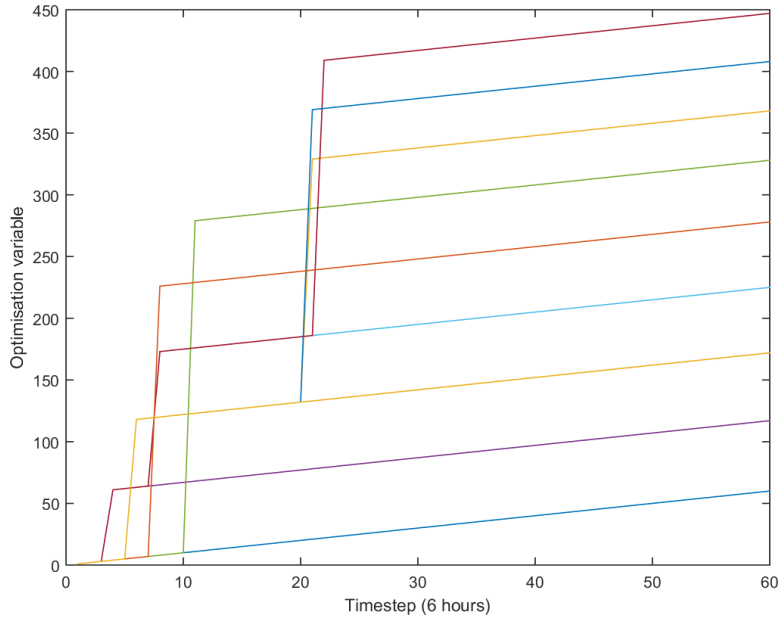
Perfect model – CPTEC



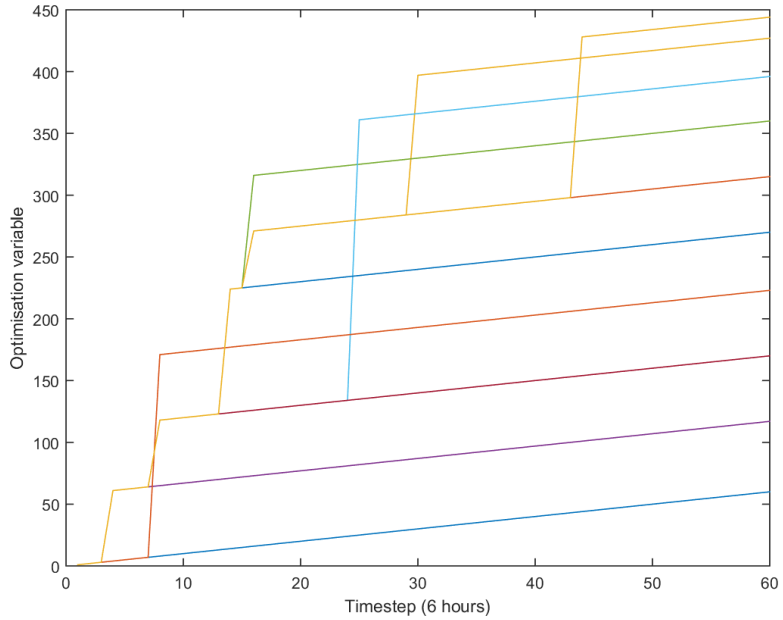
Perfect model – ECMWF



Imperfect model – CPTEC



Imperfect model – ECMWF

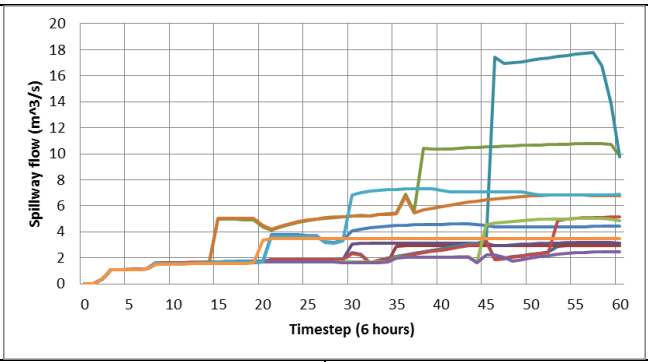


E.1.4 Results

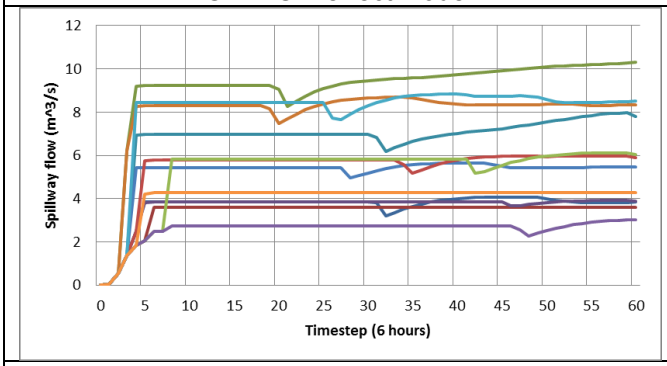
Scenario 4 Results		Quasi-optimised cost function score				
Initial reservoir level (m above datum)	Probability (%)	Base case	Perfect model		Imperfect model	
			CPTEC	ECMWF	CPTEC	ECMWF
31	3%	208.3	207.4	207.6	207.7	207.6
33	36%	37.2	34.8	35.1	36.1	36.5
35	61%	5.1	1.7	2.1	4.2	4.7

Spillway Flow - Initial level 31 m above datum

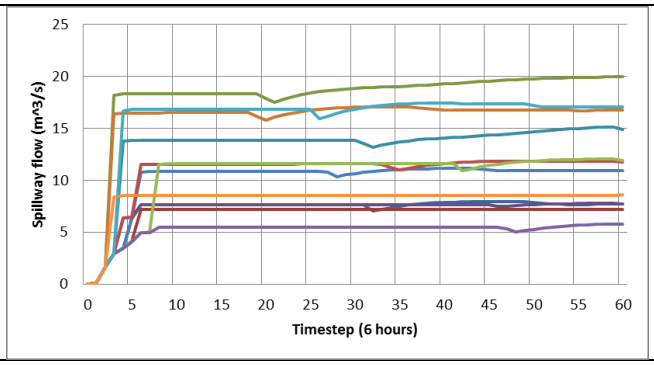
Base Case



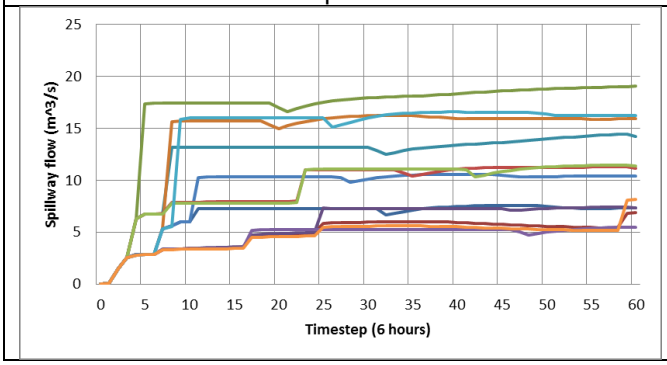
CPTEC 'Perfect Model'



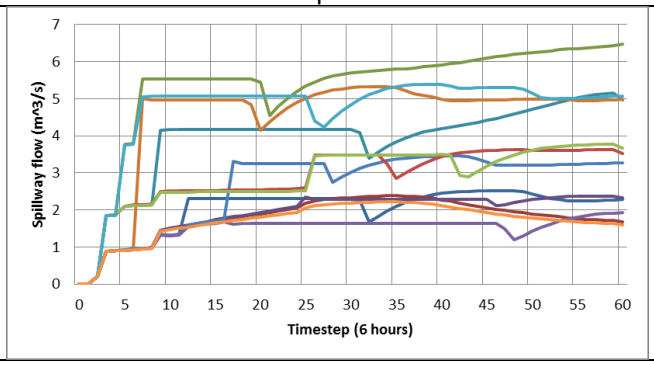
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

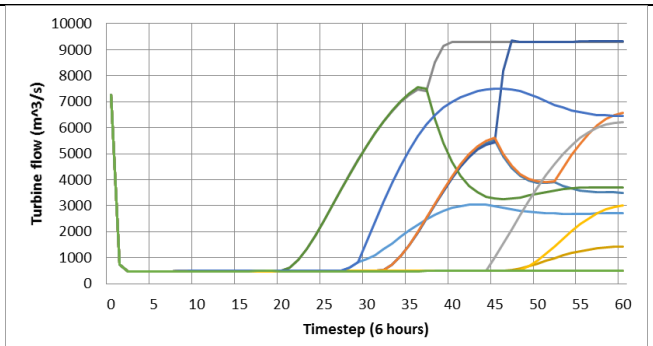


ECMWF 'Imperfect Model'

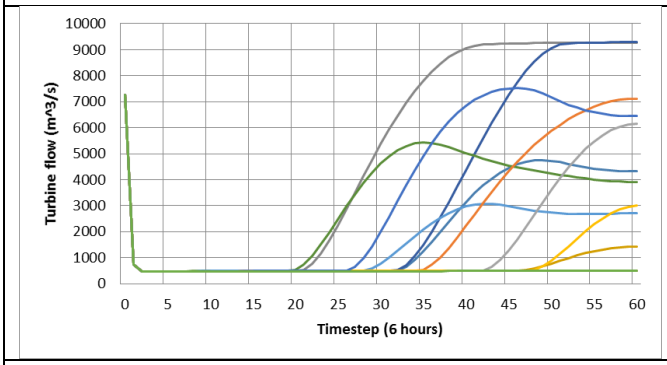


Turbine Flow - Initial level 31 m above datum

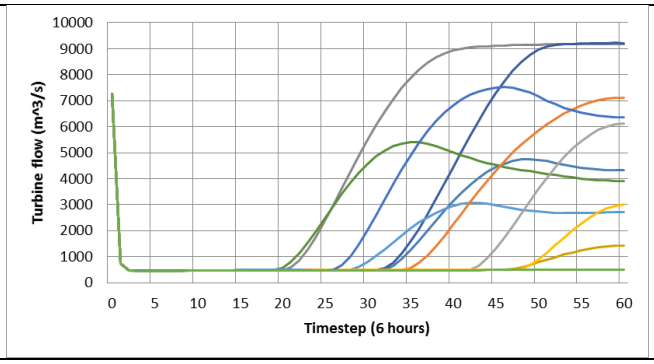
Base Case



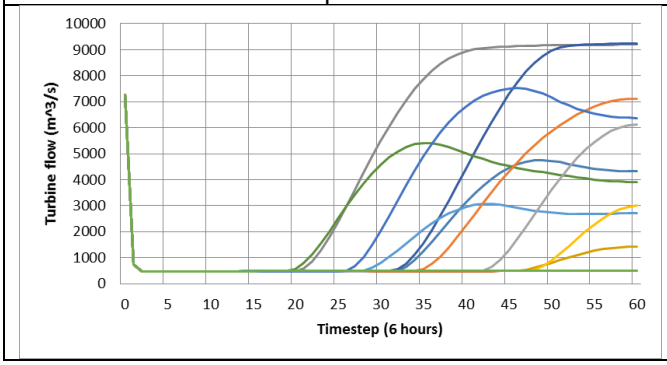
CPTEC 'Perfect Model'



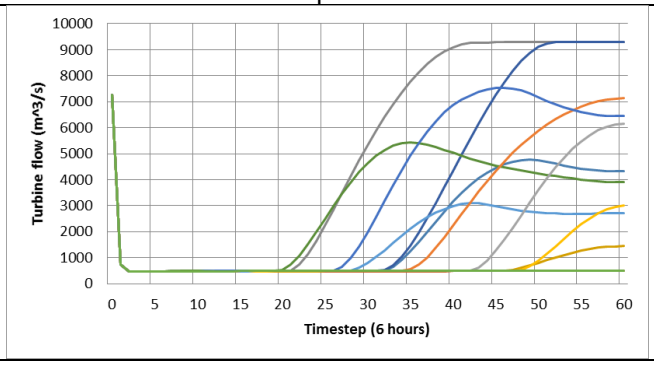
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

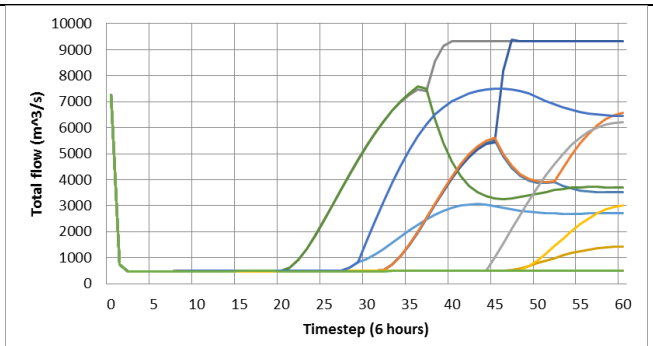


ECMWF 'Imperfect Model'

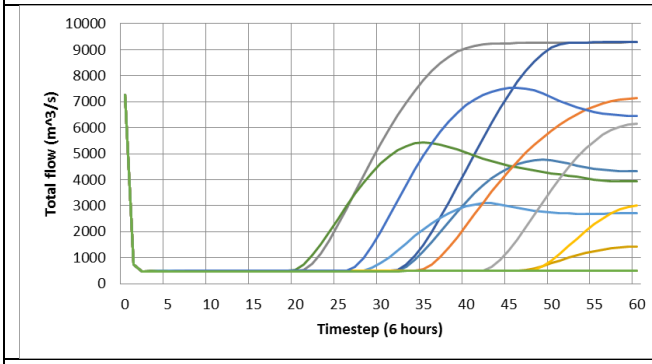


Total Flow - Initial level 31 m above datum

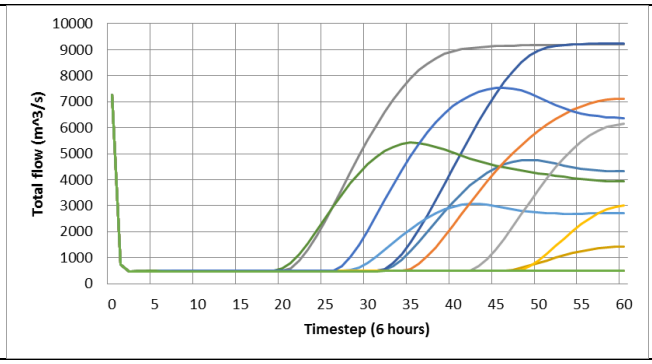
Base Case



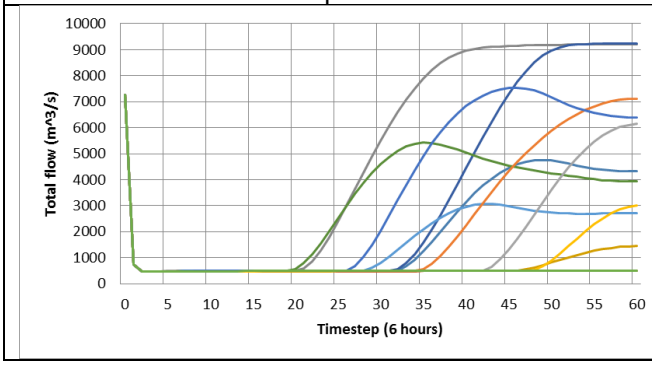
CPTEC 'Perfect Model'



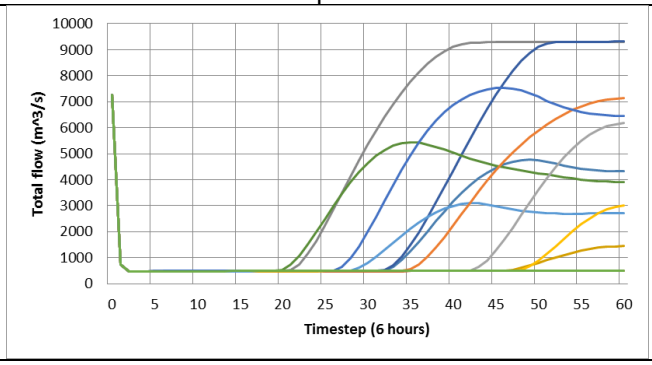
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

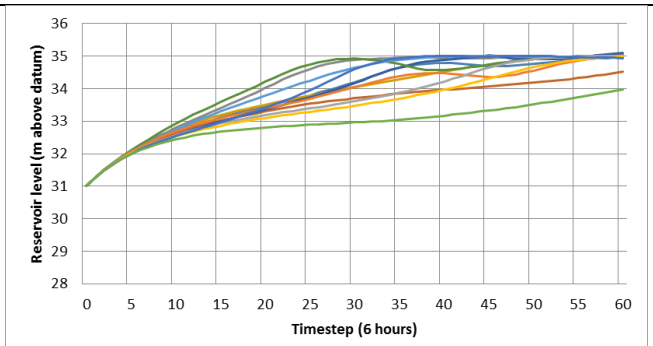


ECMWF 'Imperfect Model'

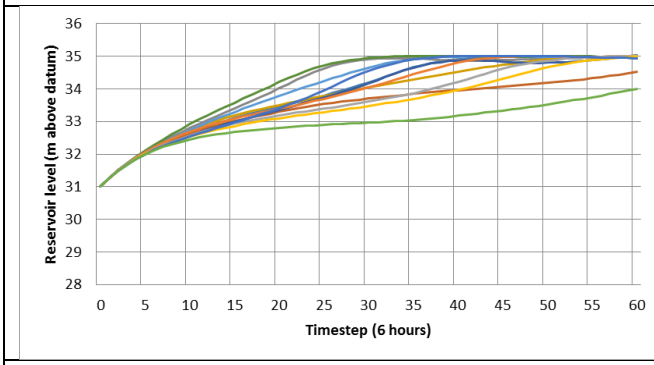


Reservoir level - Initial level 31 m above datum

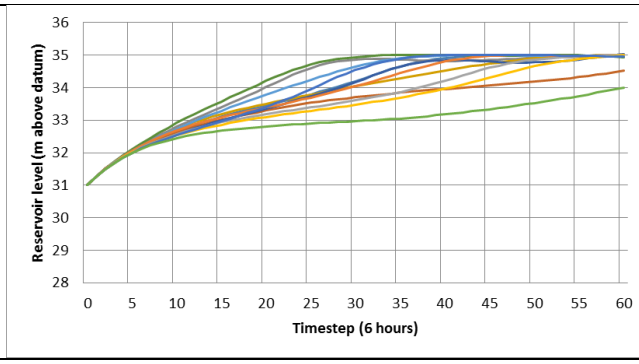
Base Case



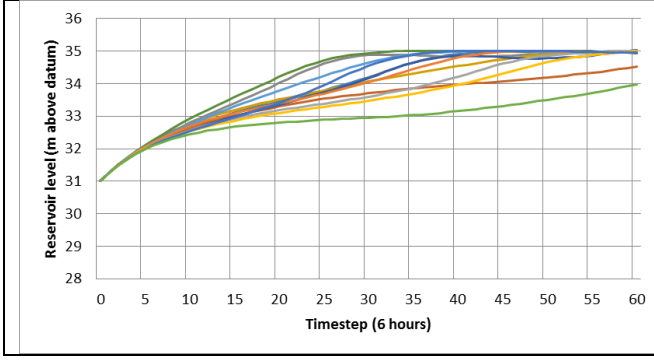
CPTEC 'Perfect Model'



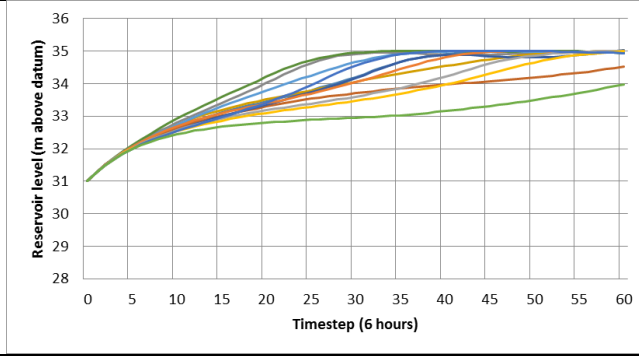
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

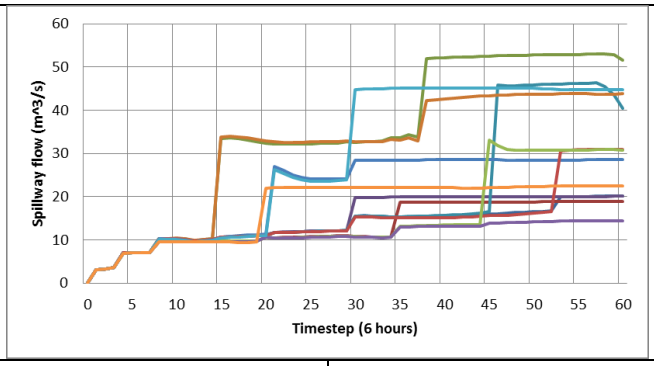


ECMWF 'Imperfect Model'

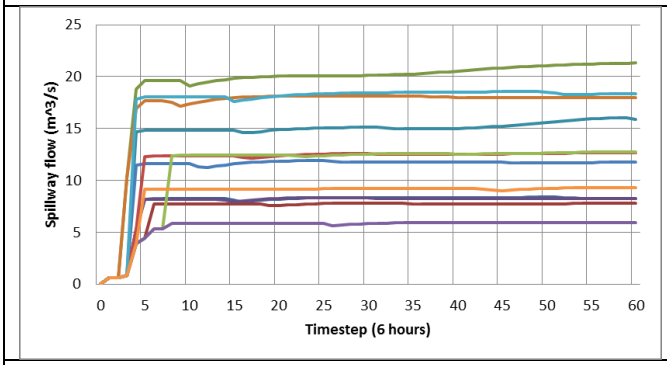


Spillway Flow - Initial level 33 m above datum

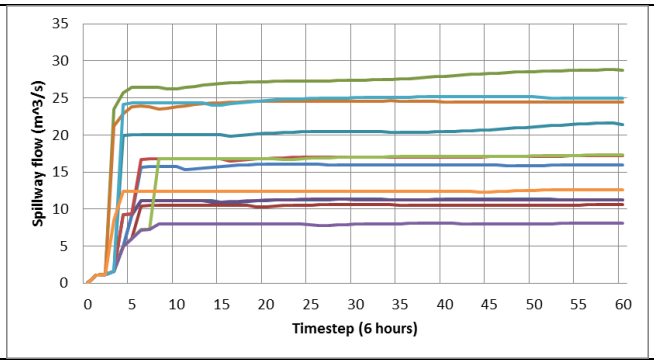
Base Case



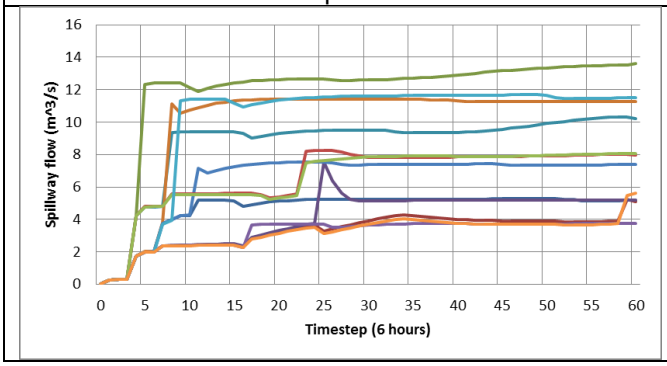
CPTEC 'Perfect Model'



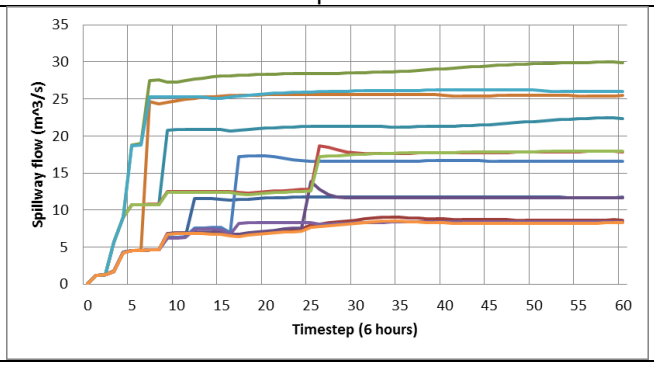
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

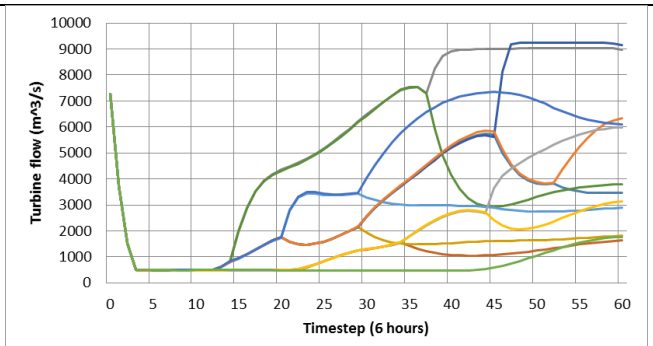


ECMWF 'Imperfect Model'

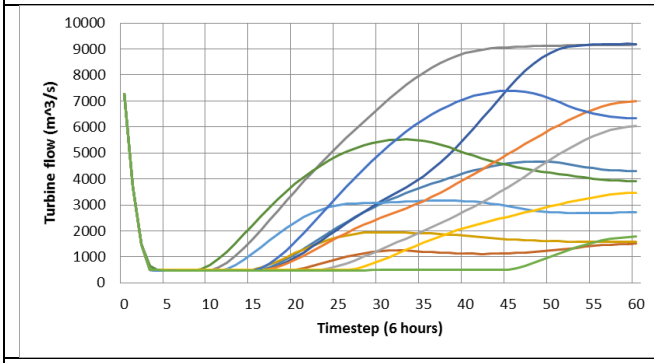


Turbine Flow - Initial level 33 m above datum

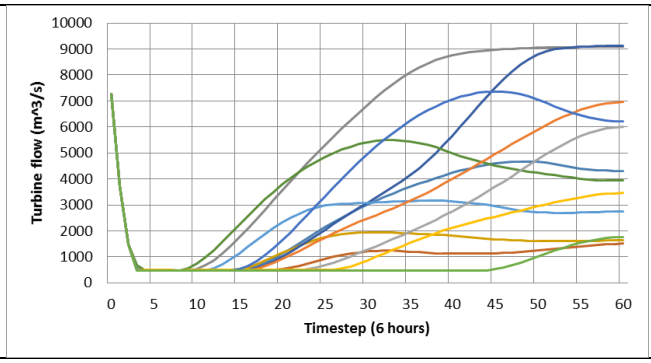
Base Case



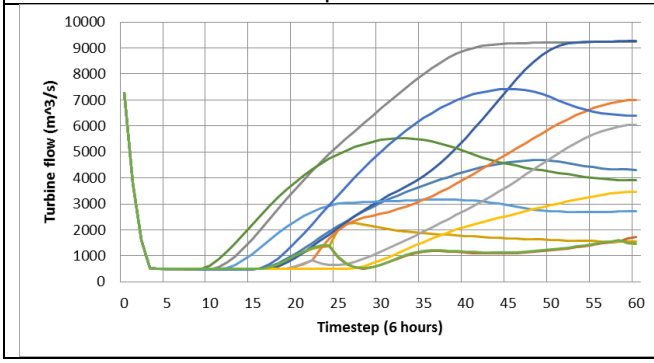
CPTEC 'Perfect Model'



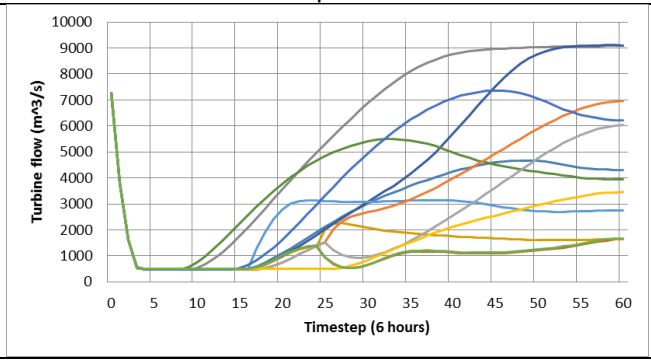
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

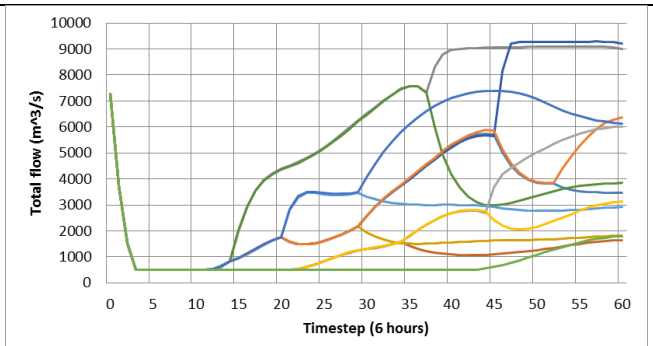


ECMWF 'Imperfect Model'

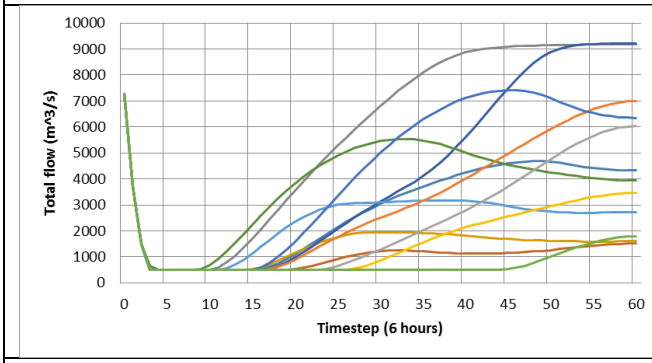


Total Flow - Initial level 33 m above datum

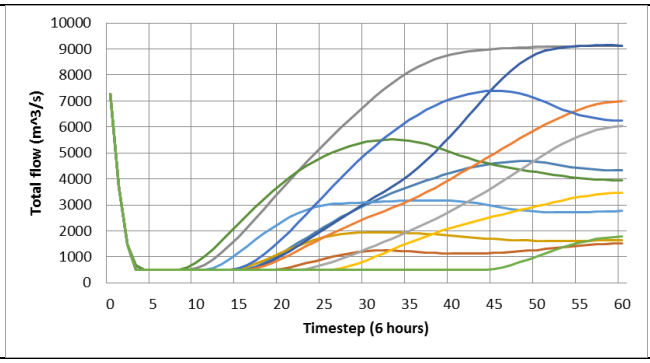
Base Case



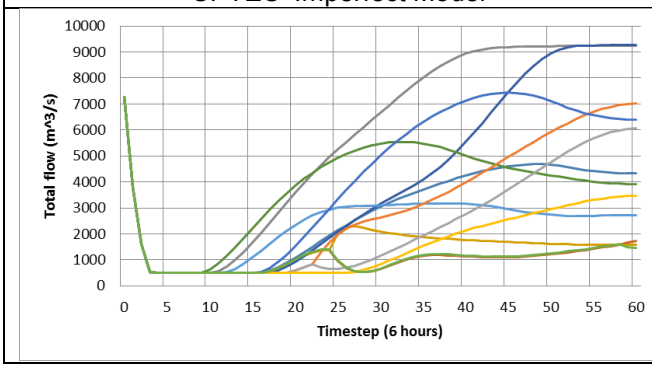
CPTEC 'Perfect Model'



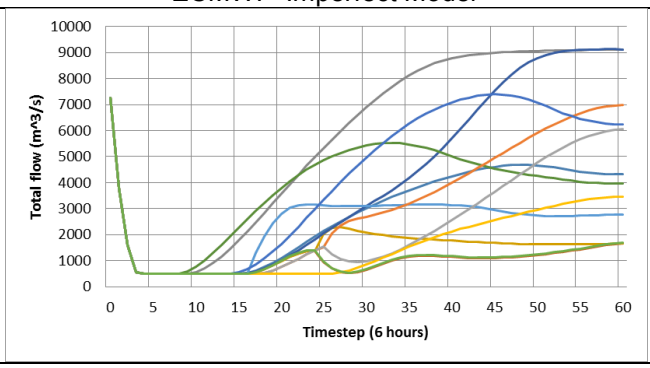
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

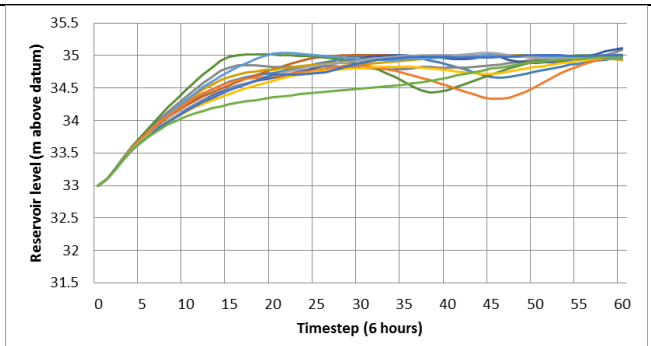


ECMWF 'Imperfect Model'

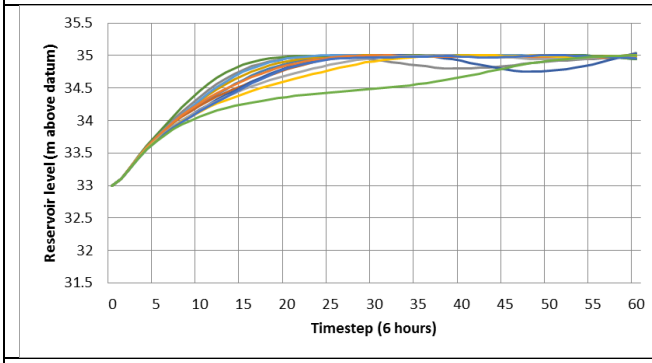


Reservoir Level - Initial level 33 m above datum

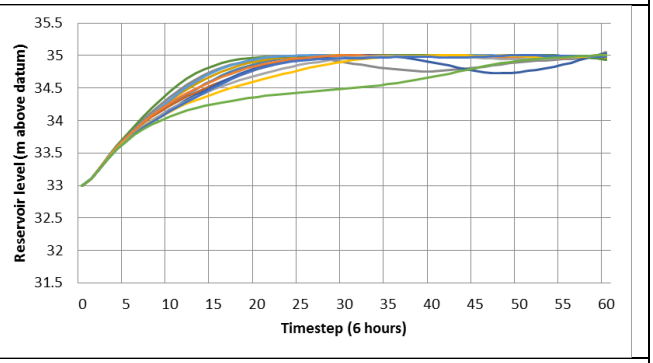
Base Case



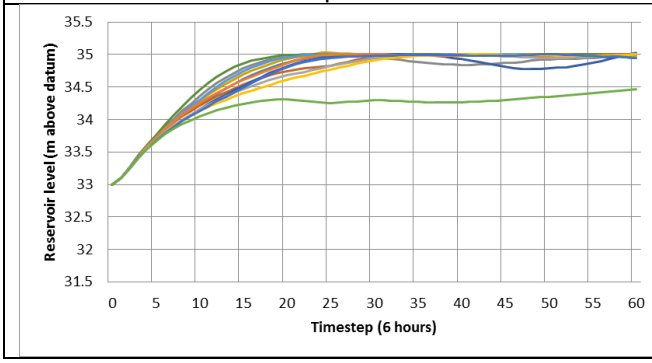
CPTEC 'Perfect Model'



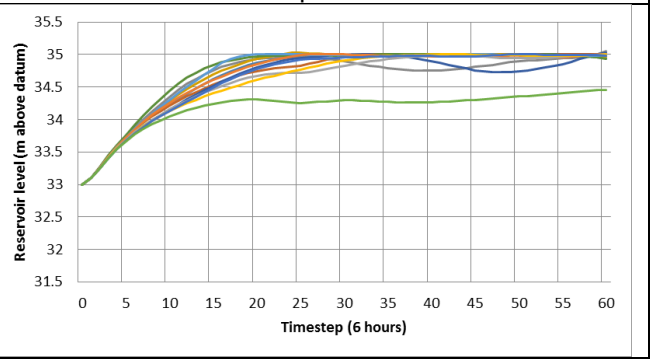
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

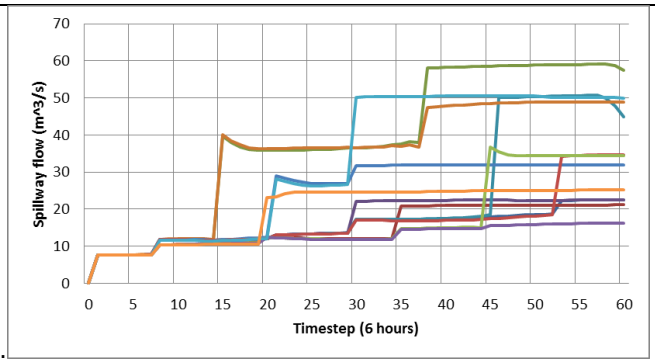


ECMWF 'Imperfect Model'

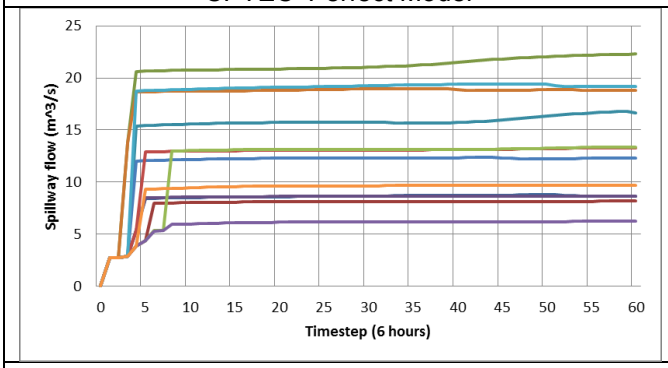


Spillway Flow - Initial level 35 m above datum

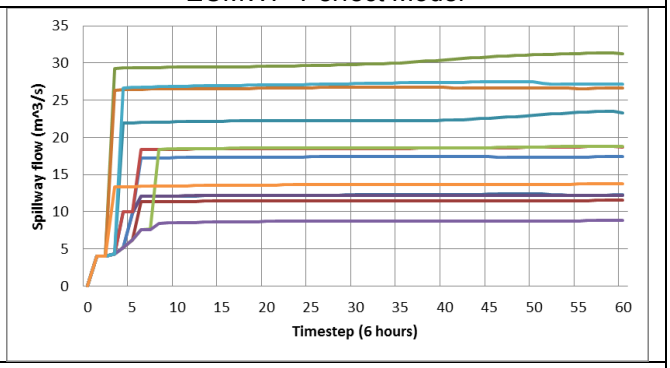
Base Case



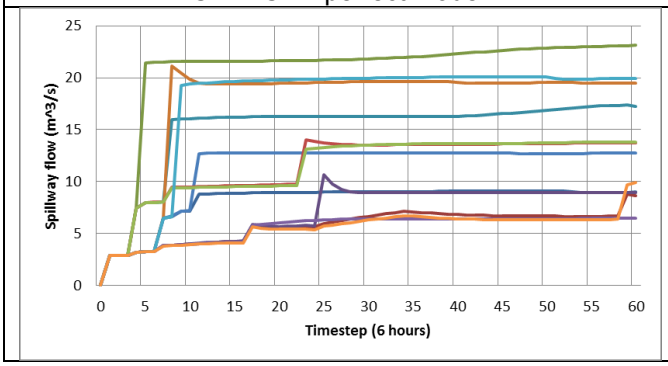
CPTEC 'Perfect Model'



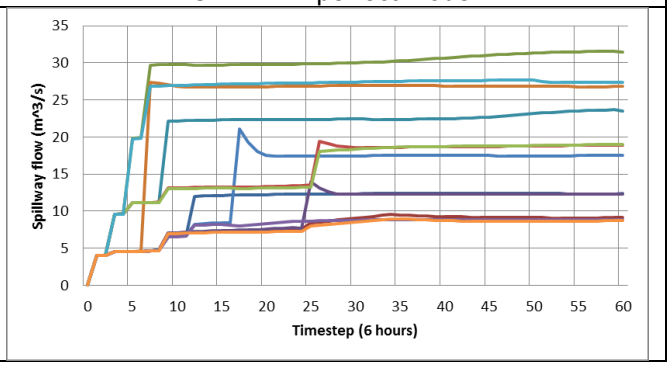
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

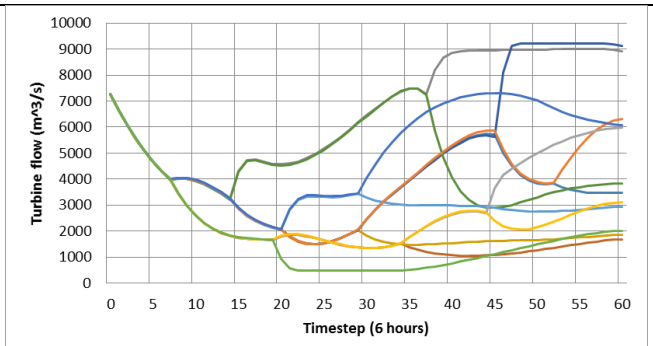


ECMWF 'Imperfect Model'

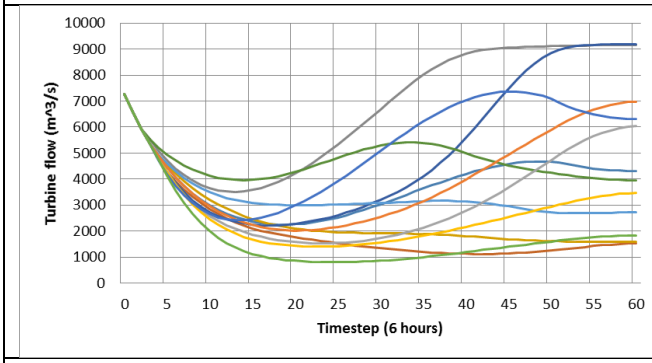


Turbine Flow - Initial level 35 m above datum

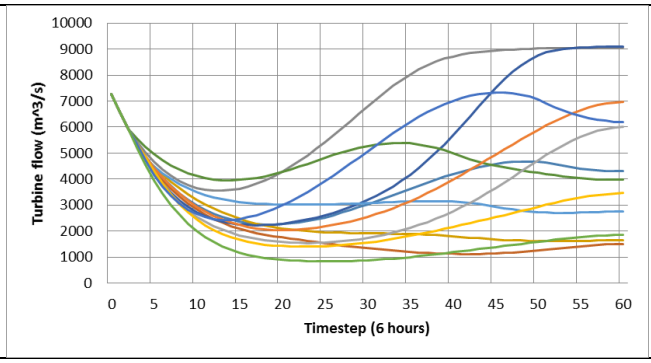
Base Case



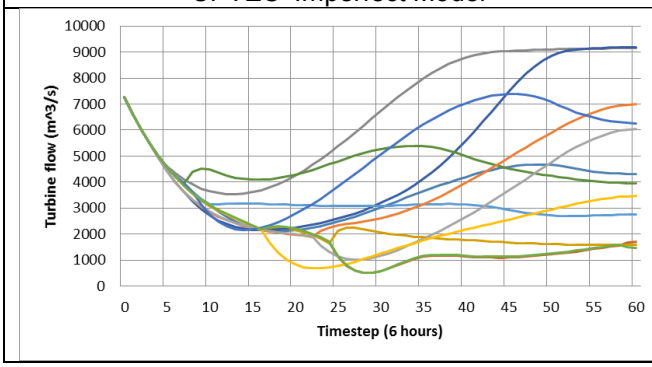
CPTEC 'Perfect Model'



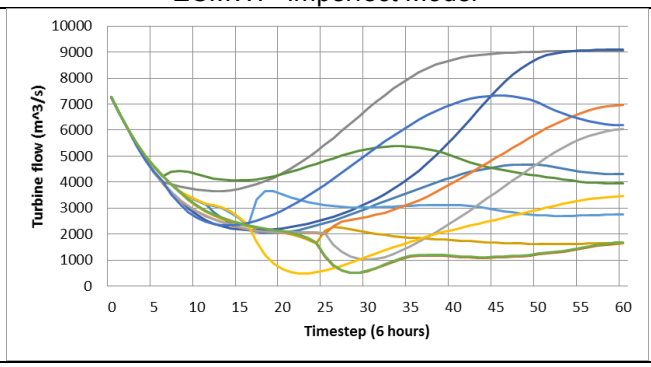
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

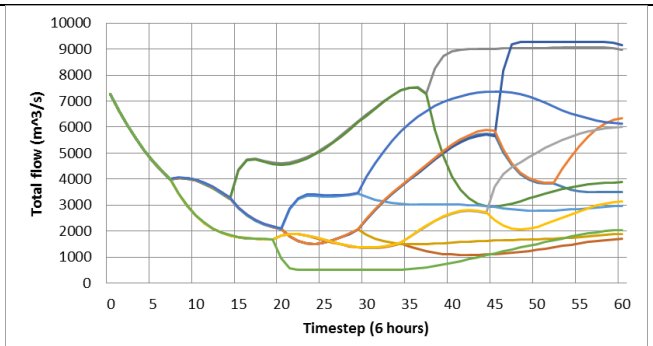


ECMWF 'Imperfect Model'

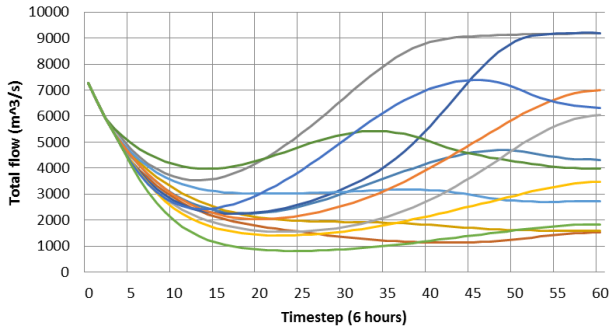


Total Flow - Initial level 35 m above datum

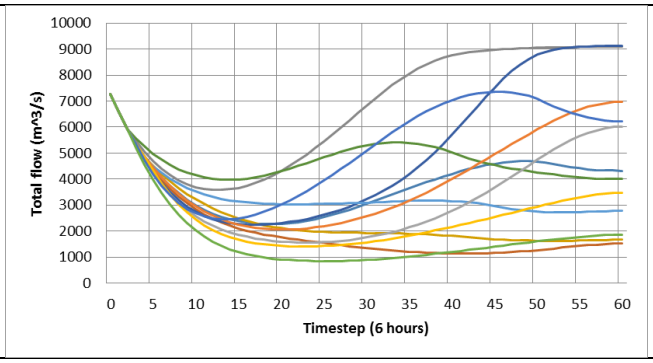
Base Case



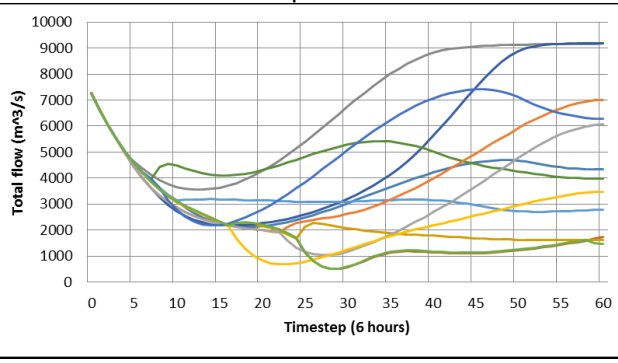
CPTEC 'Perfect Model'



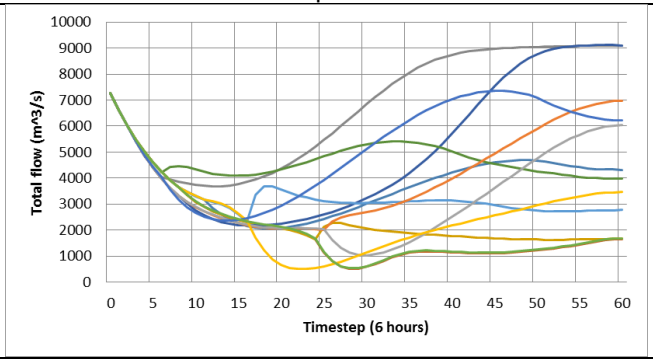
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

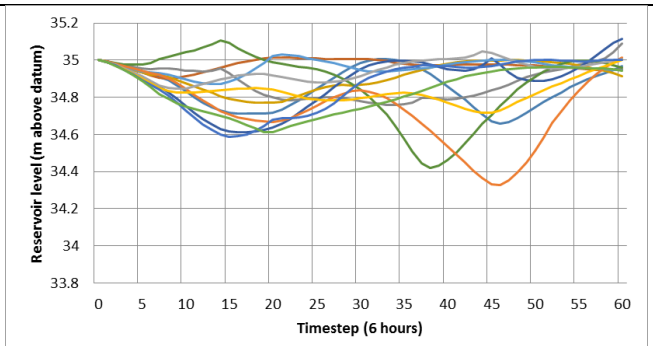


ECMWF 'Imperfect Model'

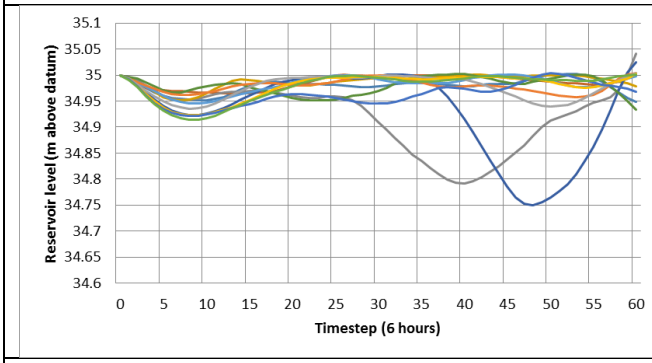


Reservoir Level - Initial level 35 m above datum

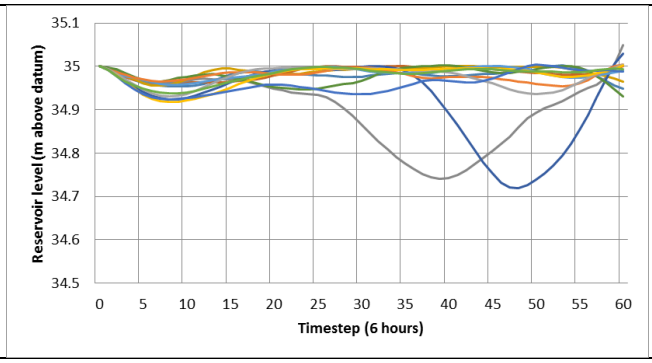
Base Case



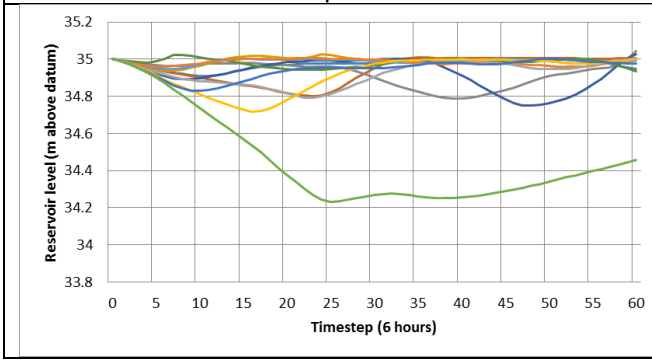
CPTEC 'Perfect Model'



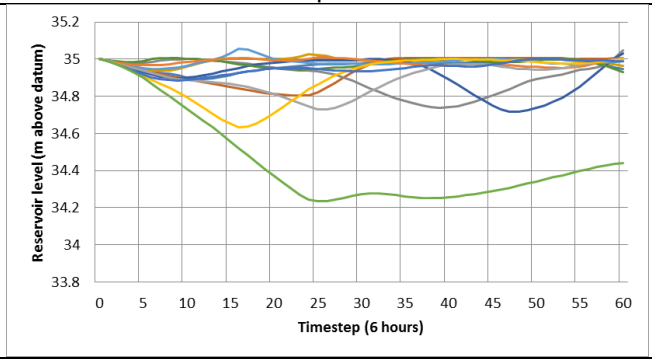
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

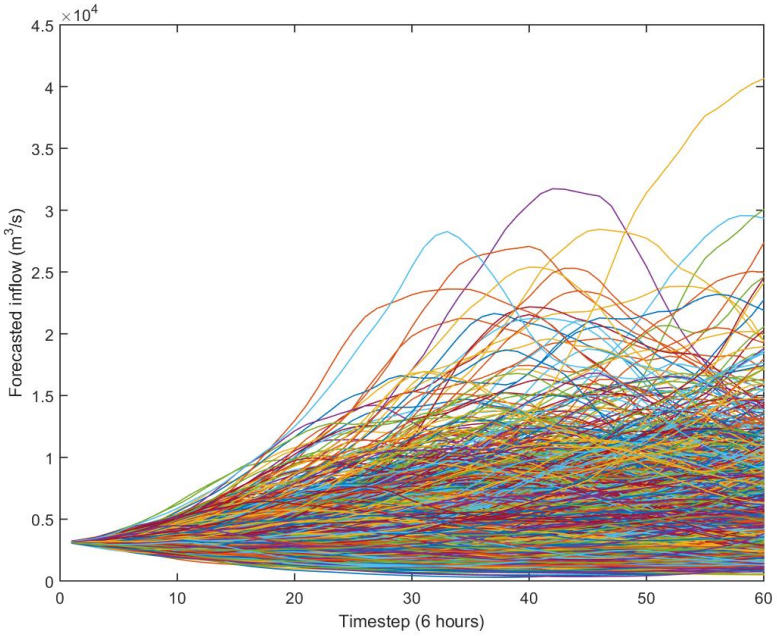


ECMWF 'Imperfect Model'

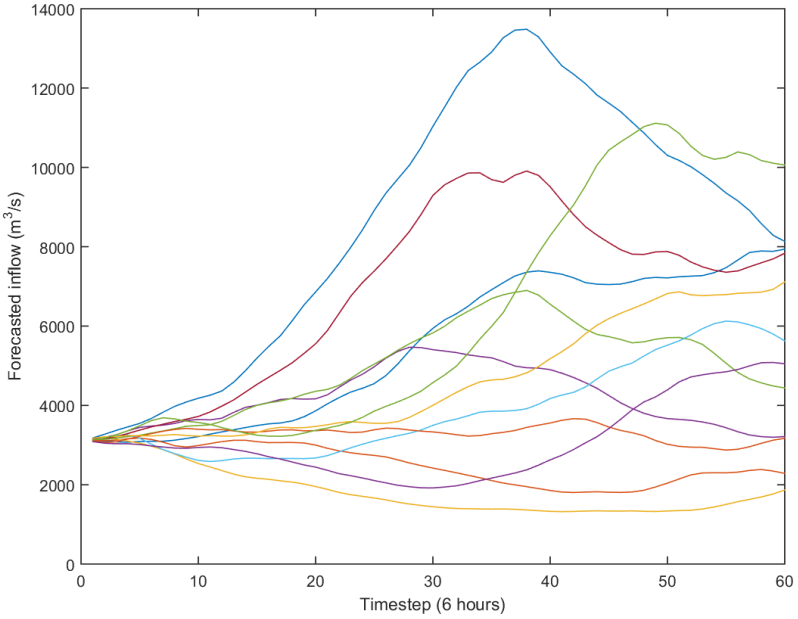


E.2 Scenario 29

E.2.1 Trajectory ensemble

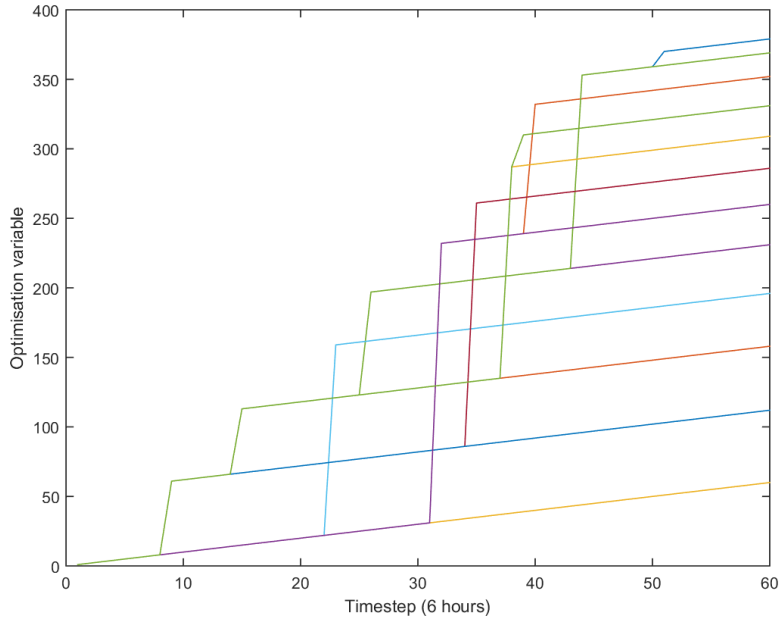


E.2.2 Reduced ensemble

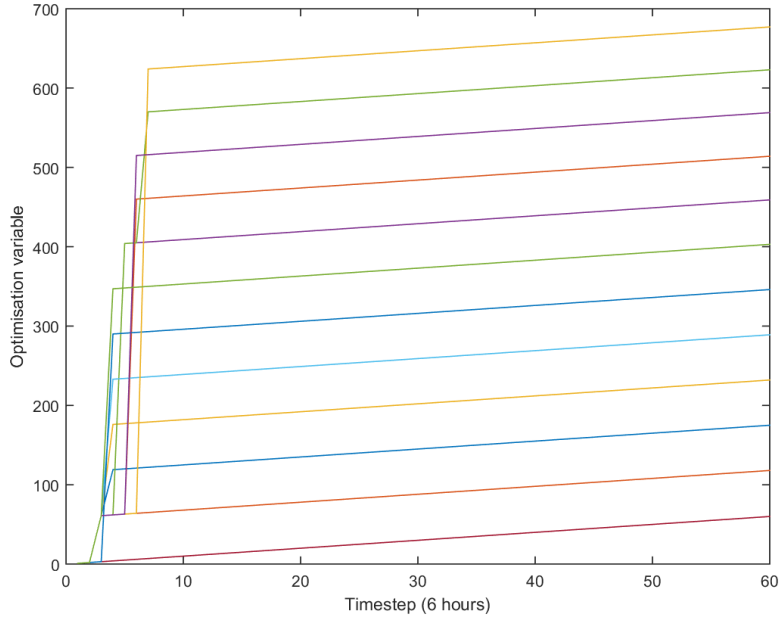


E.2.3 Nodal partition matrix

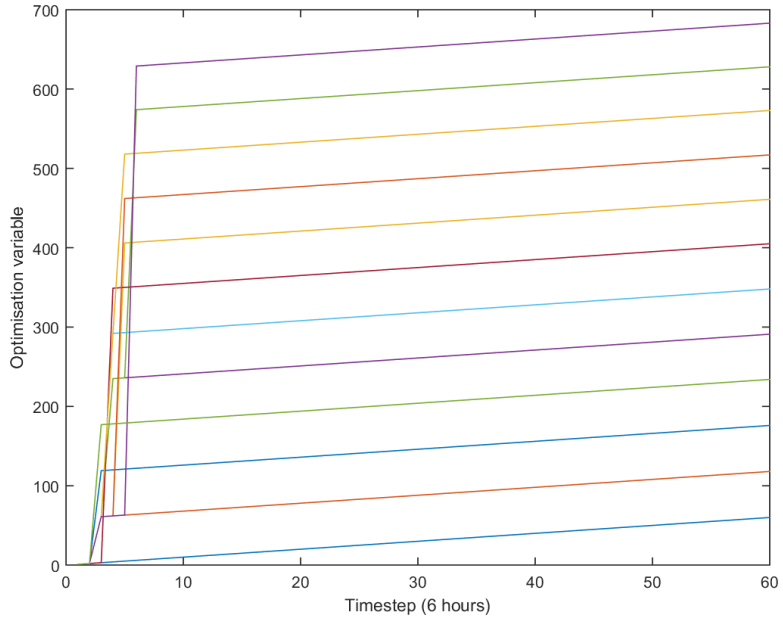
Base Case



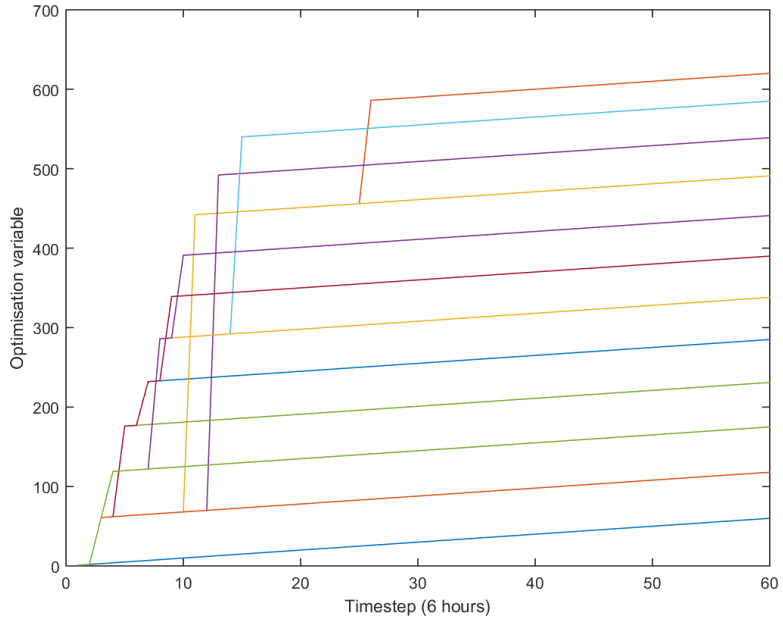
Perfect model – CPTEC



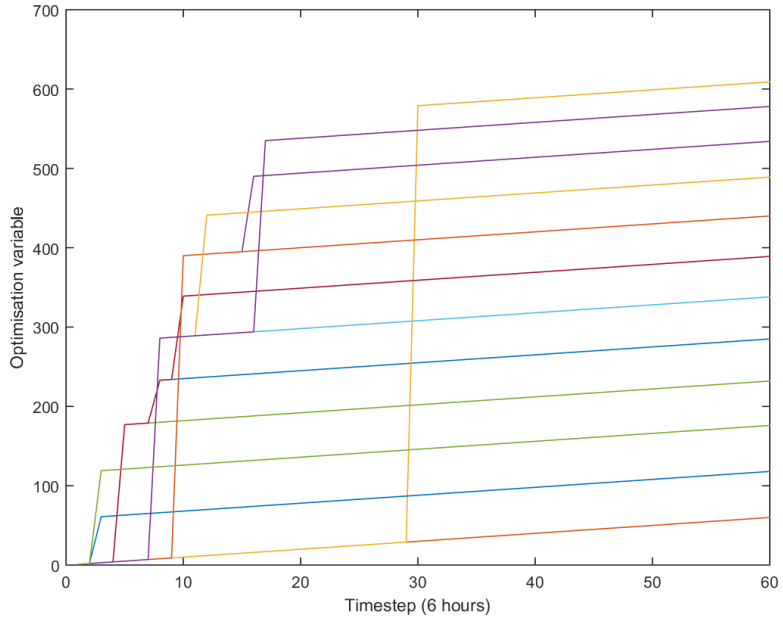
Perfect model – ECMWF



Imperfect model – CPTEC



Imperfect model – ECMWF

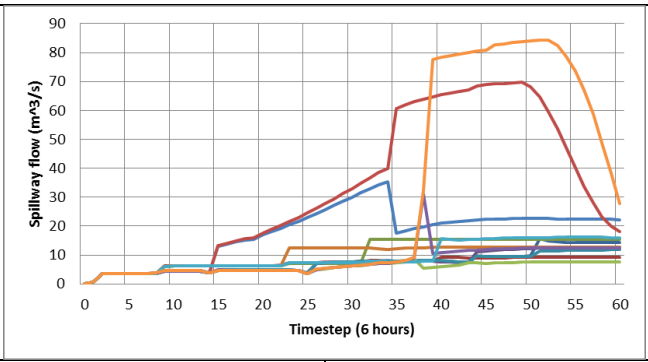


E.2.4 Results

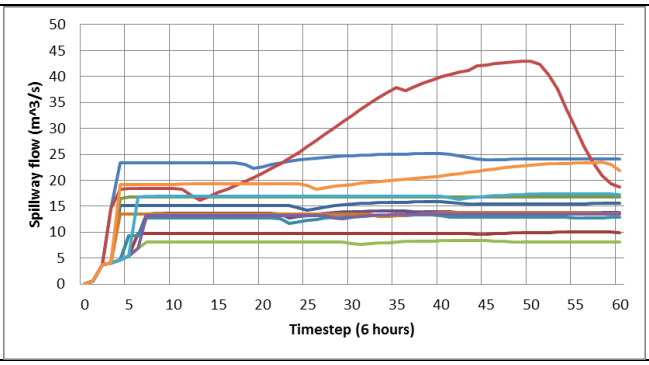
Scenario 29 Results		Quasi-optimised cost function score				
Initial reservoir level (m above datum)	Probability (%)	Base case	Perfect model		Imperfect model	
			CPTEC	ECMWF	CPTEC	ECMWF
31	3%	213.9	203.0	202.7	202.8	203.0
33	36%	60.1	37.3	38.3	37.6	37.6
35	61%	33.4	5.2	4.9	5.6	5.1

Spillway Flow - Initial level 31 m above datum

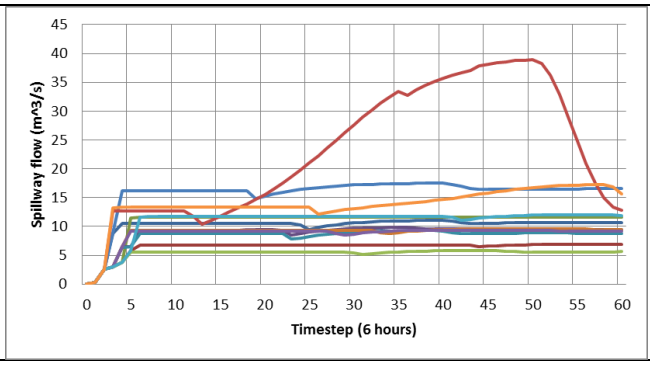
Base Case



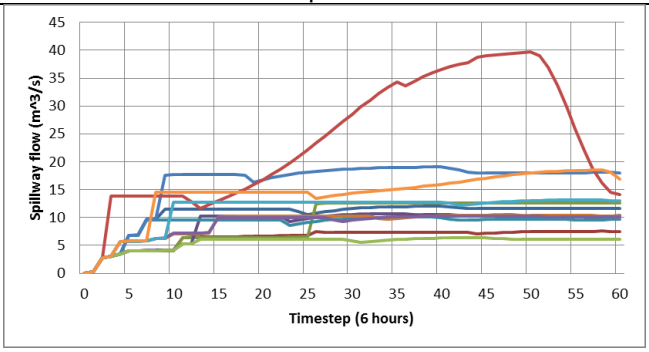
CPTEC 'Perfect Model'



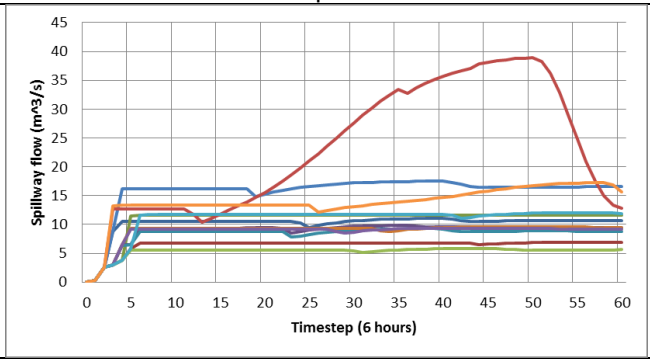
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

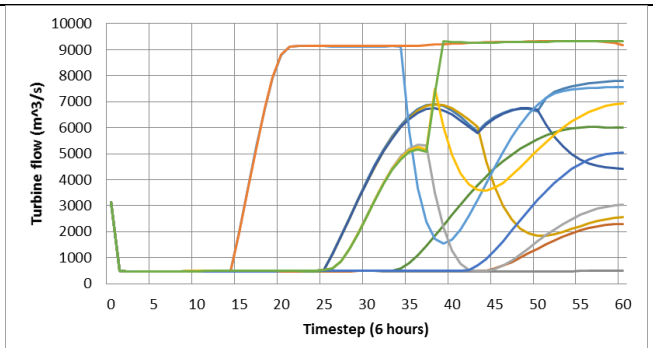


ECMWF 'Imperfect Model'

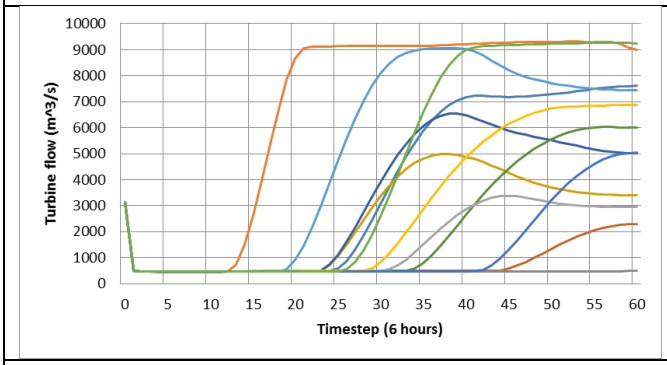


Turbine Flow - Initial level 31 m above datum

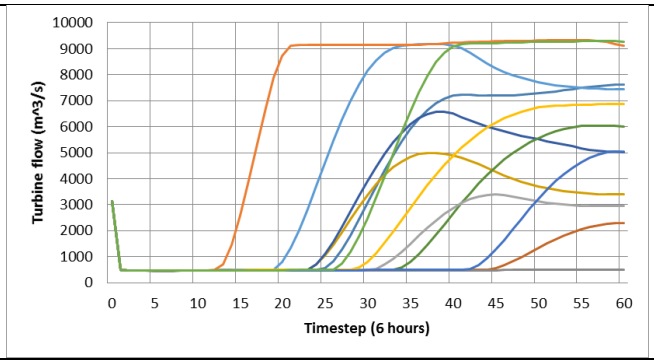
Base Case



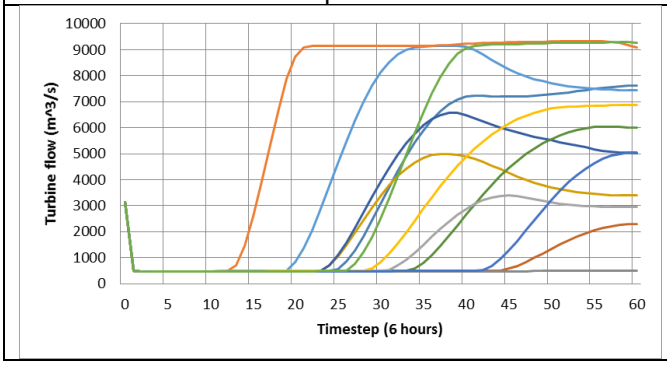
CPTEC 'Perfect Model'



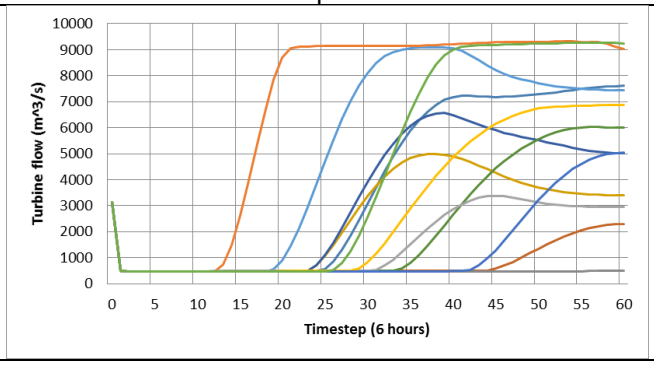
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

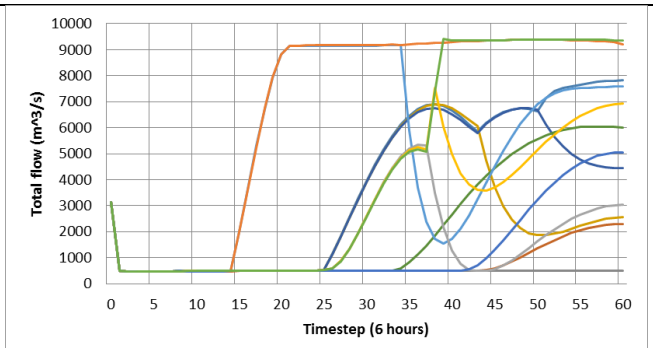


ECMWF 'Imperfect Model'

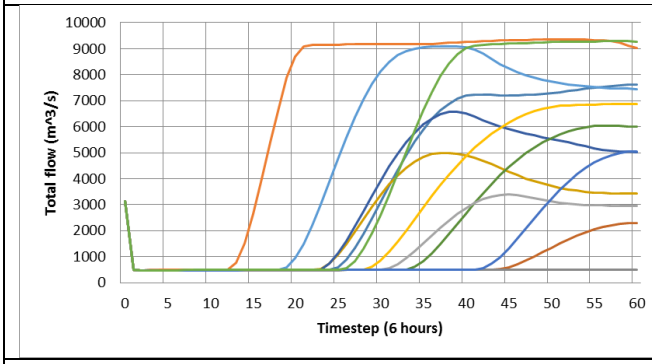


Total Flow - Initial level 31 m above datum

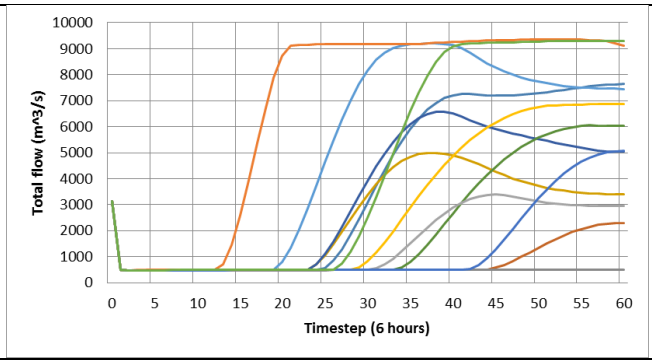
Base Case



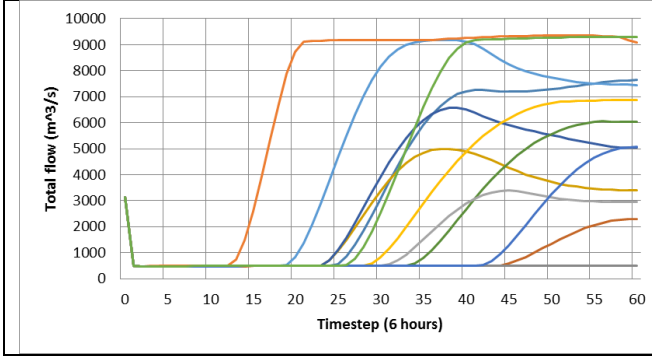
CPTEC 'Perfect Model'



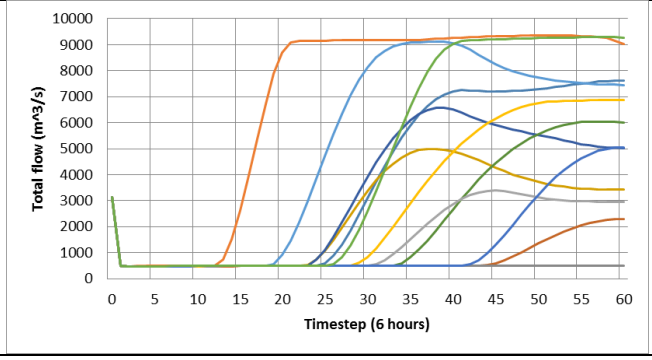
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

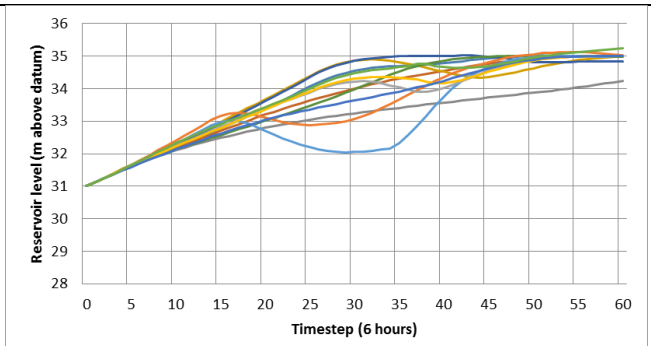


ECMWF 'Imperfect Model'

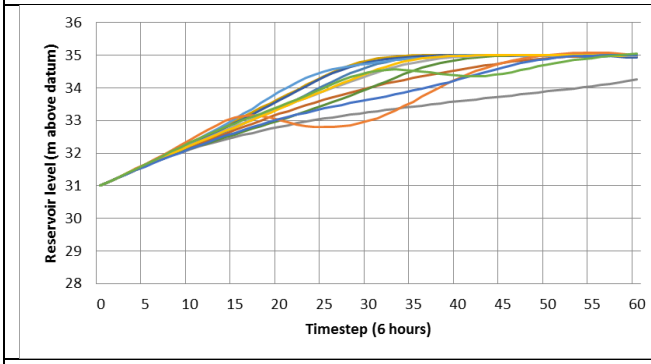


Reservoir Level - Initial level 31 m above datum

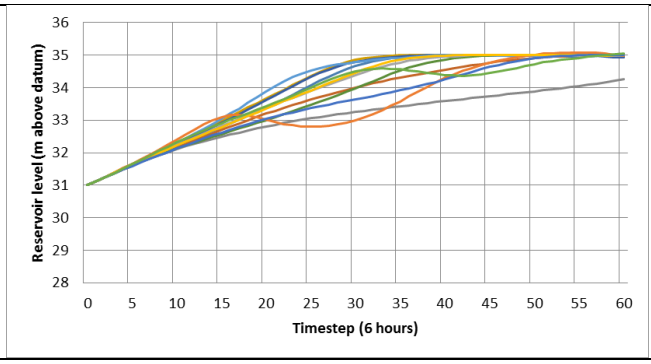
Base Case



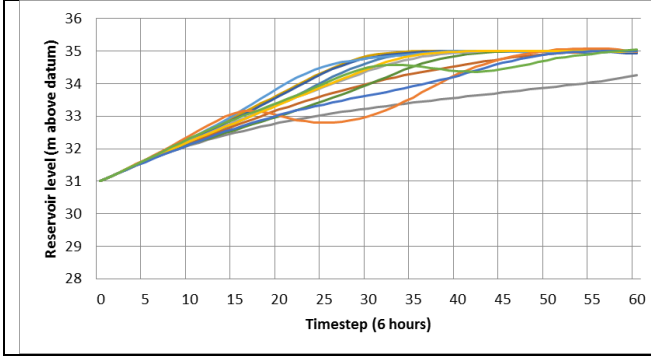
CPTEC 'Perfect Model'



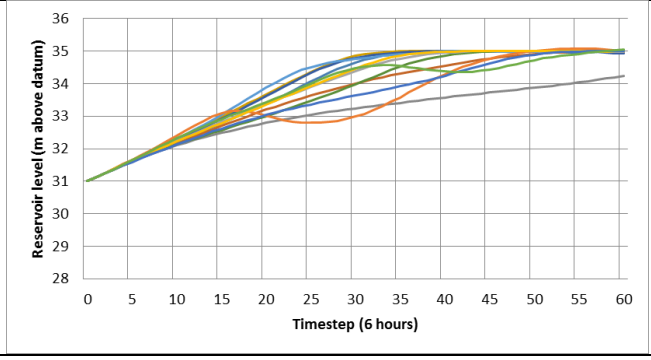
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

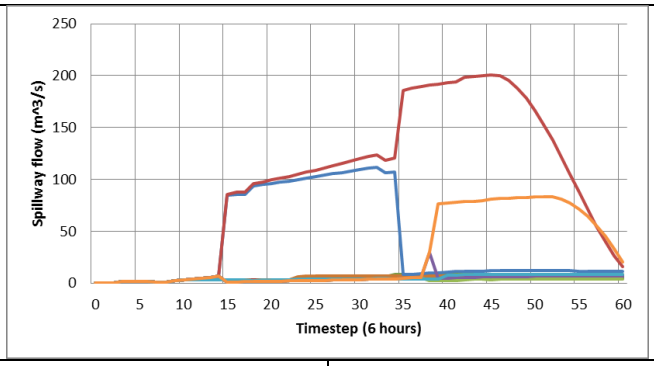


ECMWF 'Imperfect Model'

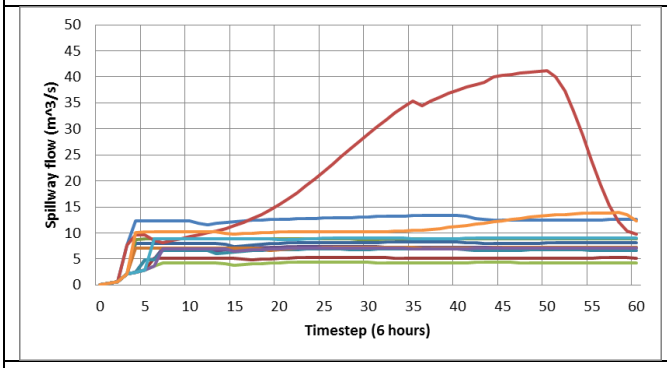


Spillway Flow - Initial level 33 m above datum

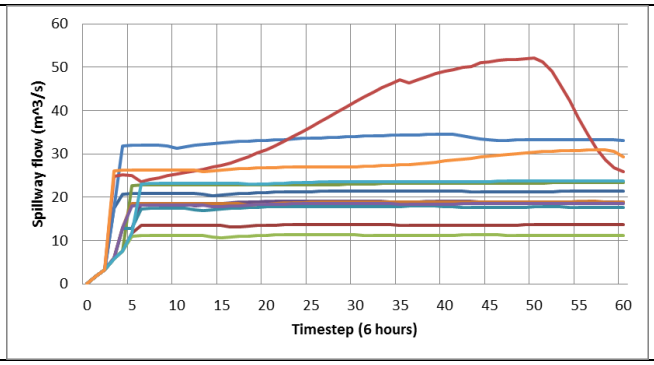
Base Case



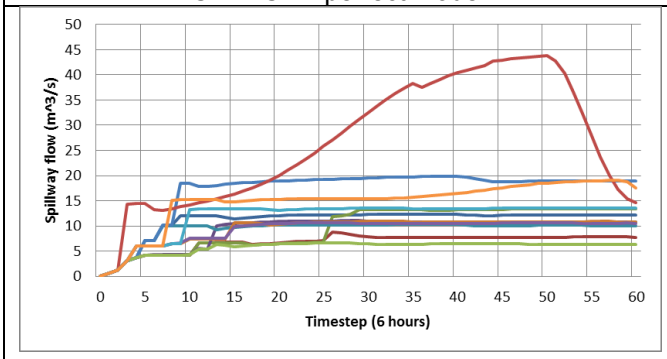
CPTEC 'Perfect Model'



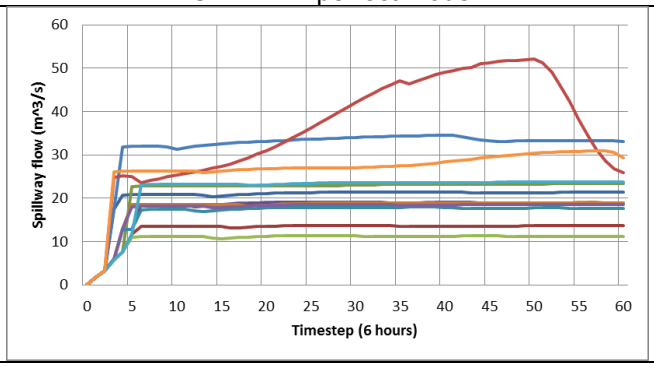
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

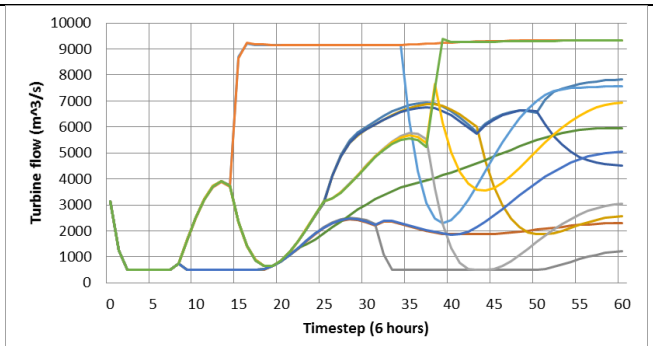


ECMWF 'Imperfect Model'

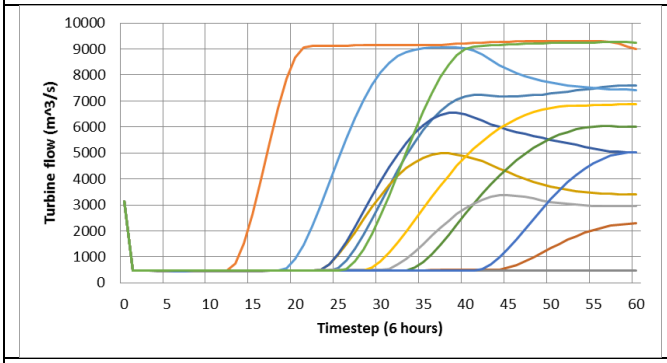


Turbine Flow - Initial level 33 m above datum

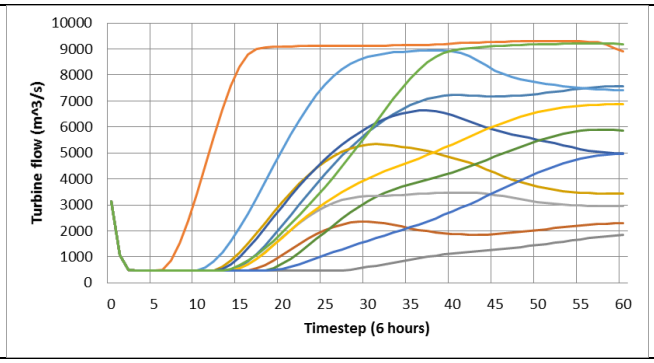
Base Case



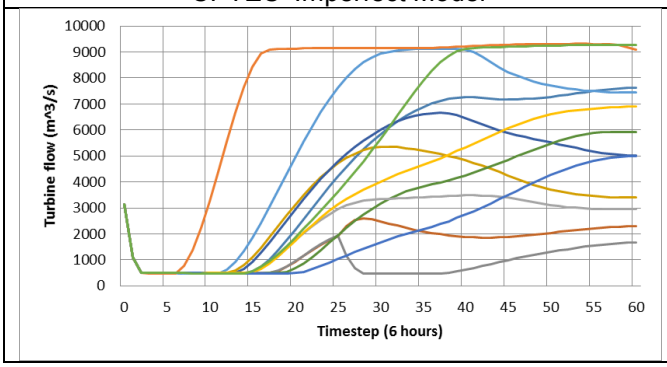
CPTEC 'Perfect Model'



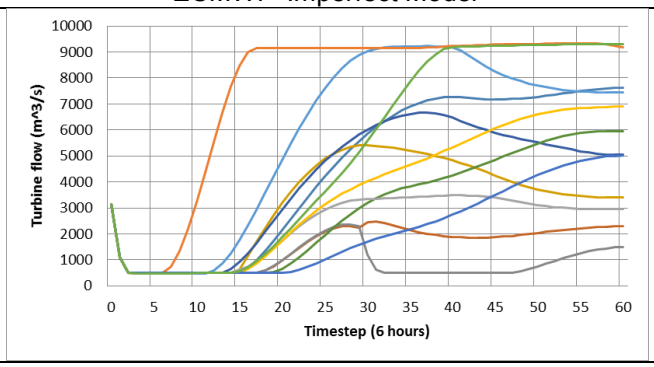
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

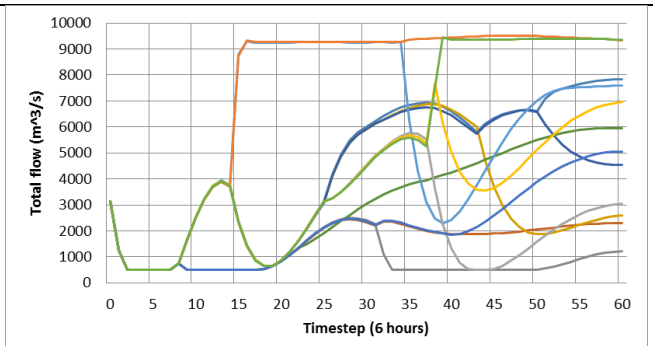


ECMWF 'Imperfect Model'

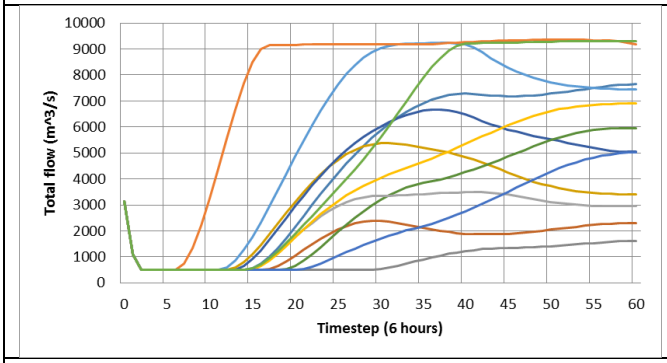


Total Flow - Initial level 33 m above datum

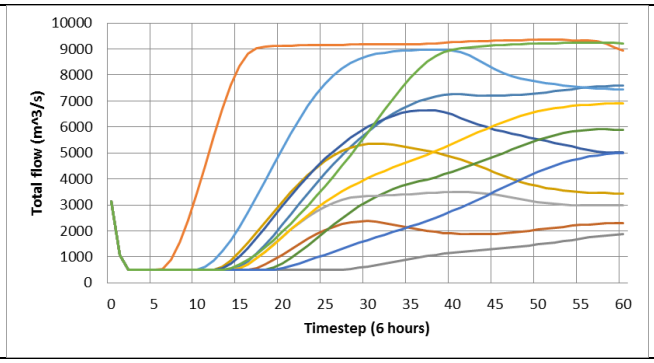
Base Case



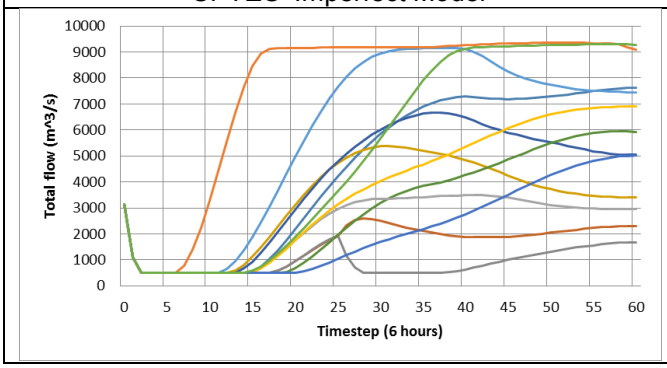
CPTEC 'Perfect Model'



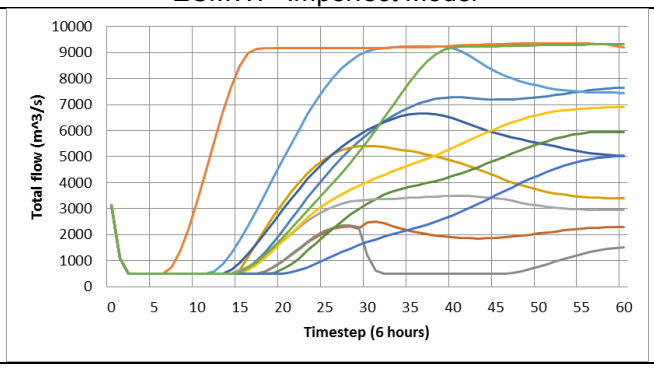
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

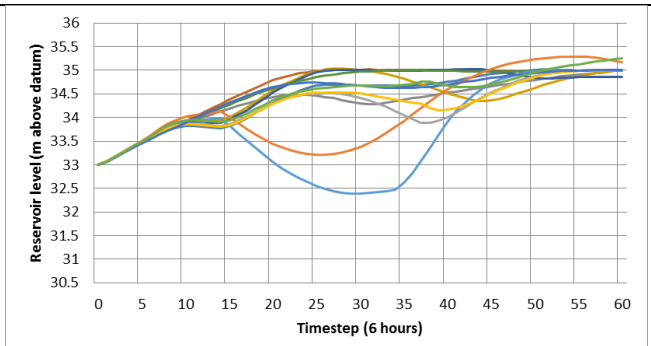


ECMWF 'Imperfect Model'

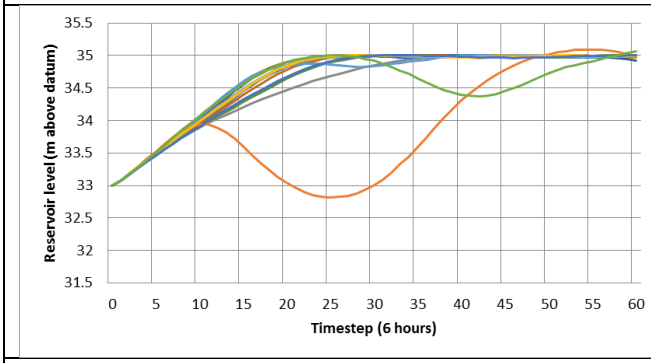


Reservoir Level - Initial level 33 m above datum

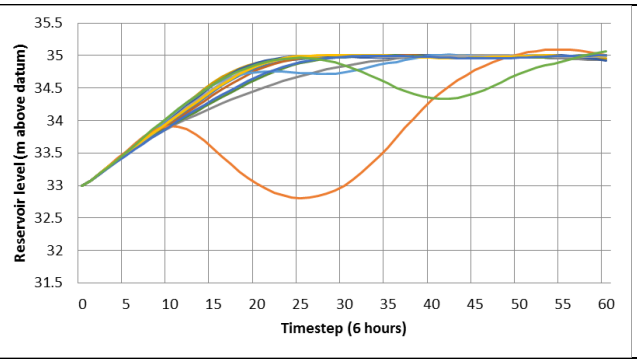
Base Case



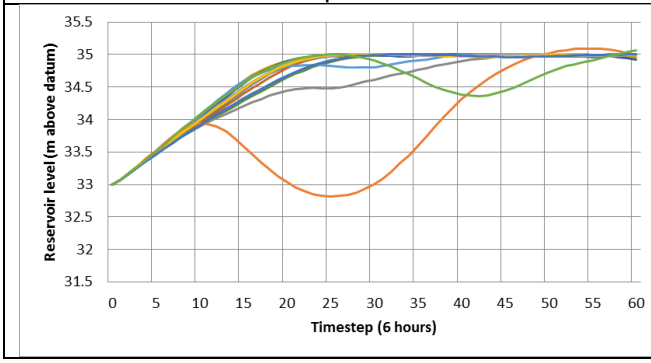
CPTEC 'Perfect Model'



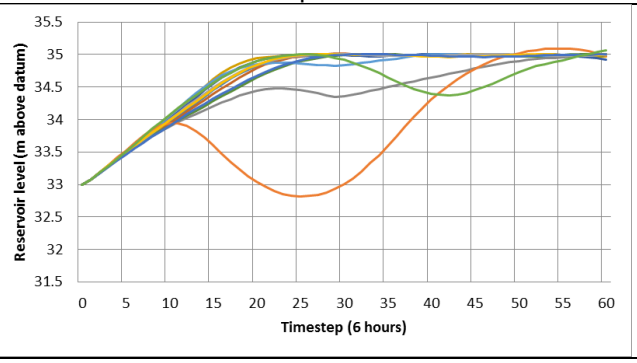
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

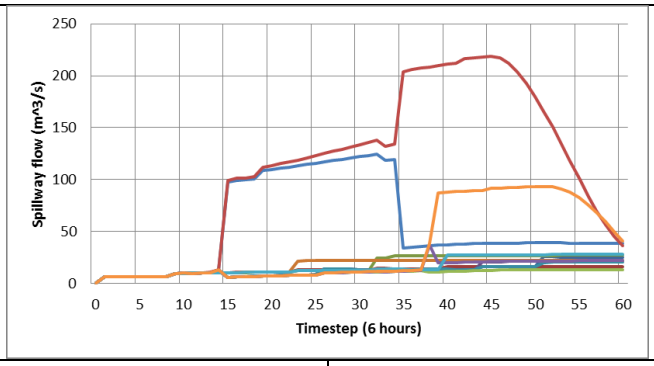


ECMWF 'Imperfect Model'

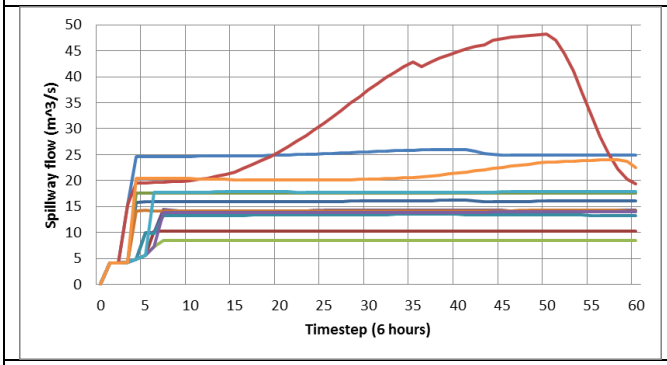


Spillway Flow - Initial level 35 m above datum

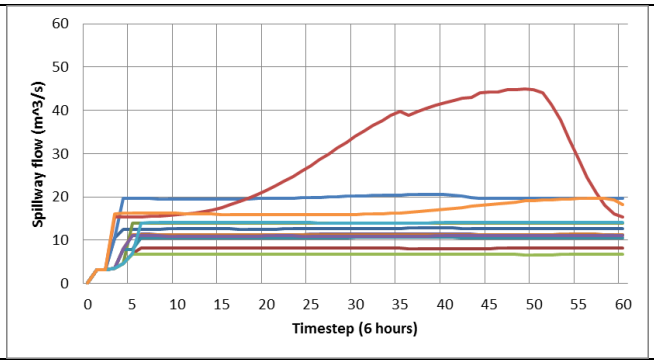
Base Case



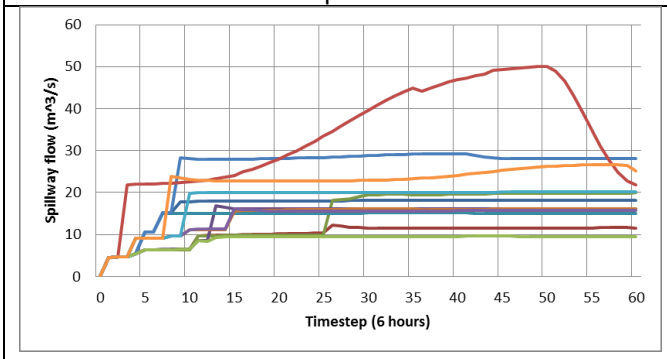
CPTEC 'Perfect Model'



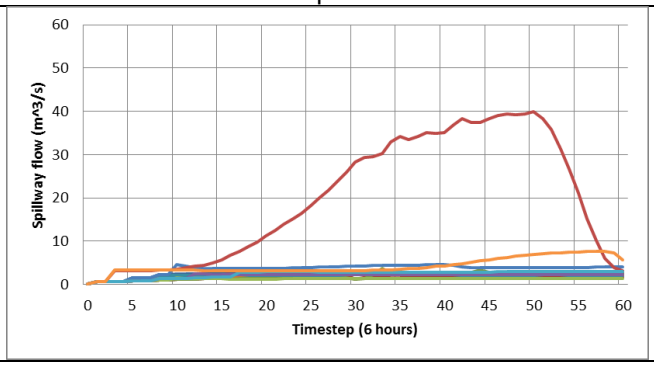
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

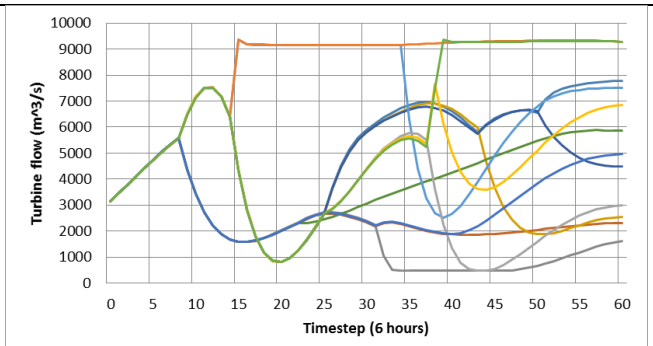


ECMWF 'Imperfect Model'

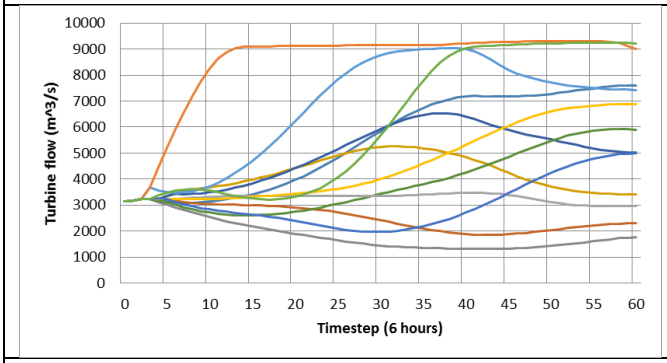


Turbine Flow - Initial level 35 m above datum

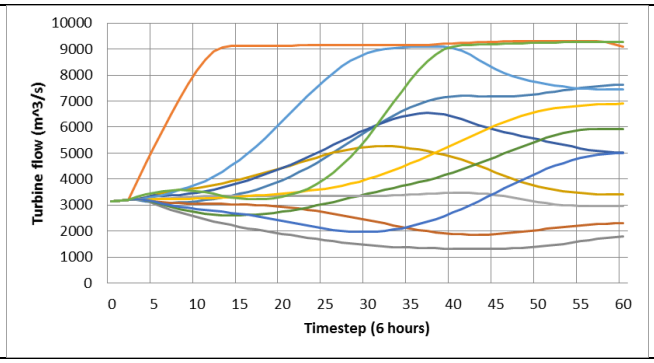
Base Case



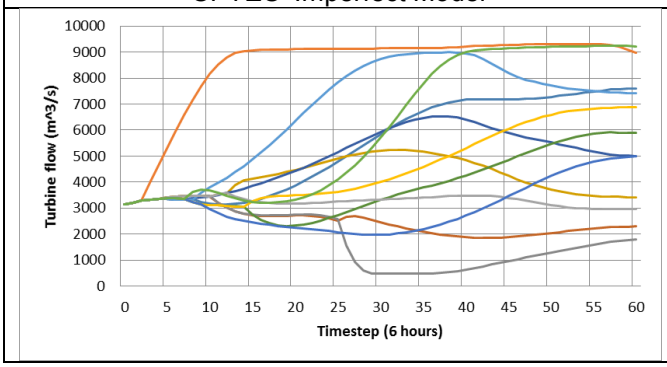
CPTEC 'Perfect Model'



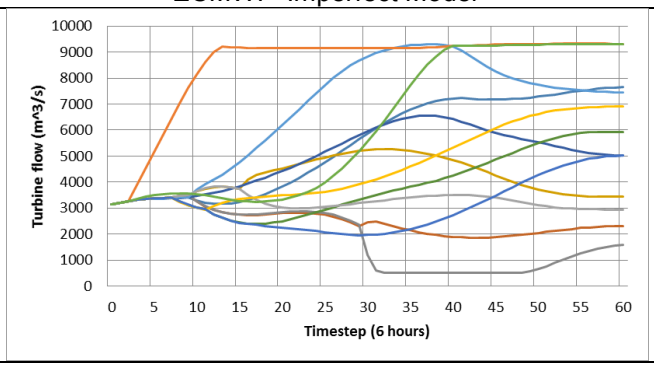
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

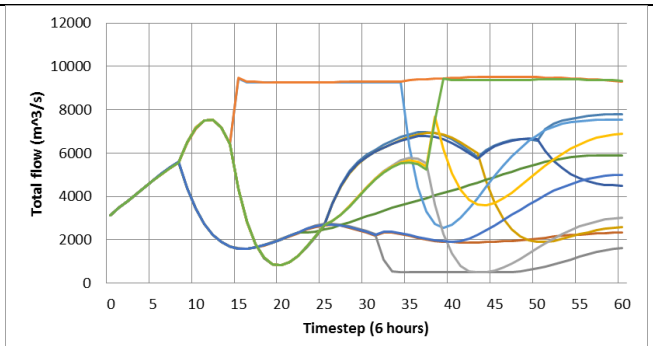


ECMWF 'Imperfect Model'

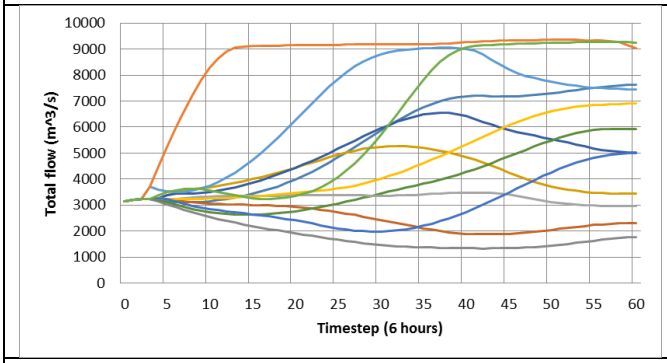


Total Flow - Initial level 35 m above datum

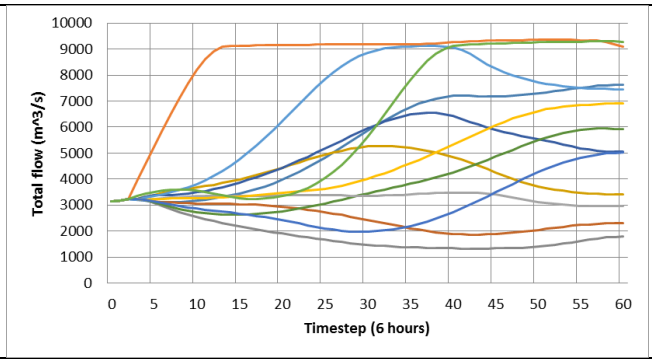
Base Case



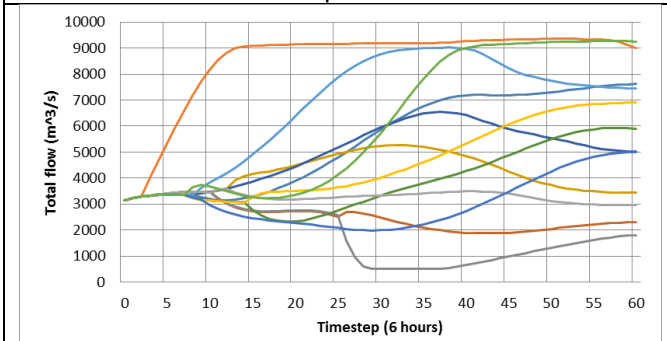
CPTEC 'Perfect Model'



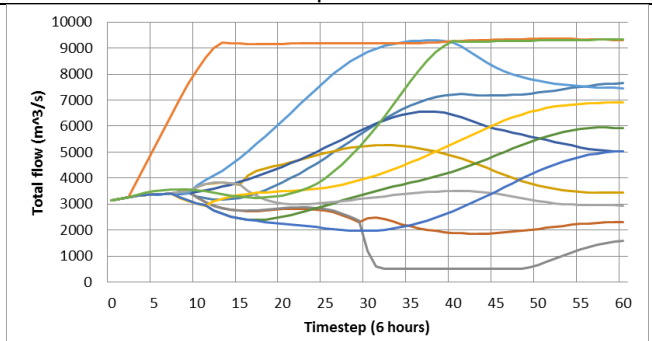
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

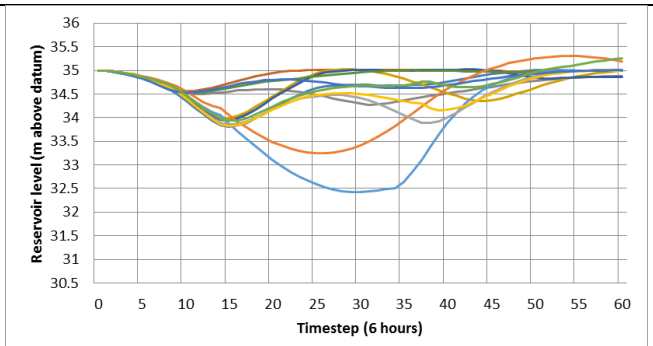


ECMWF 'Imperfect Model'

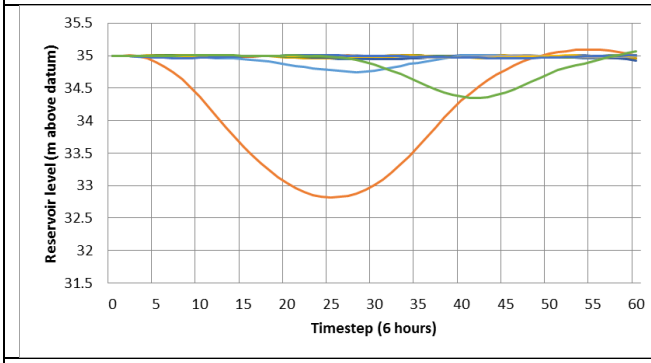


Reservoir Level - Initial level 35 m above datum

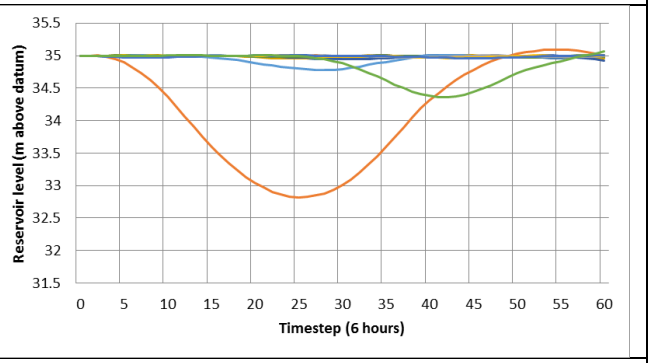
Base Case



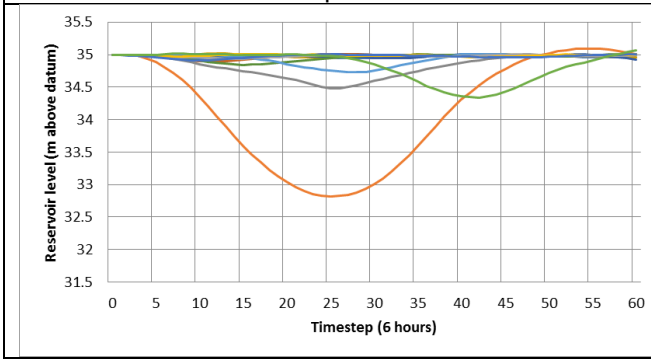
CPTEC 'Perfect Model'



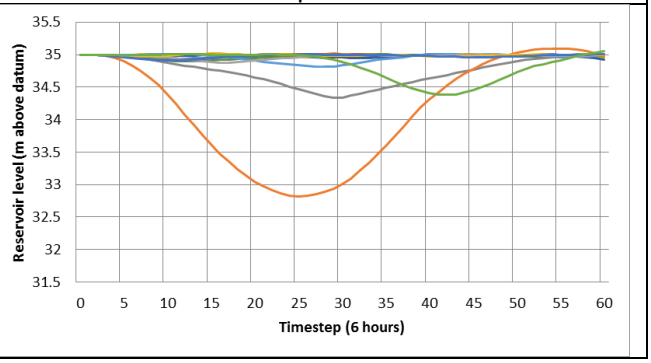
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

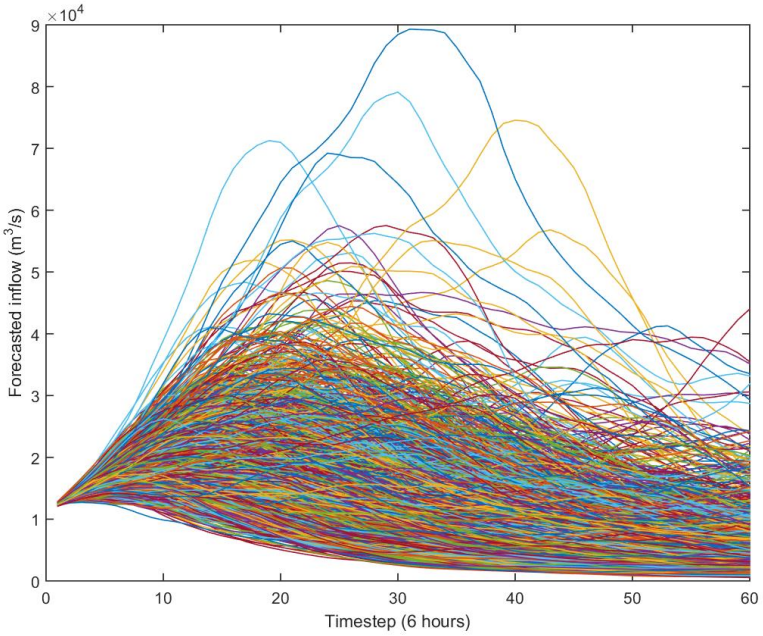


ECMWF 'Imperfect Model'

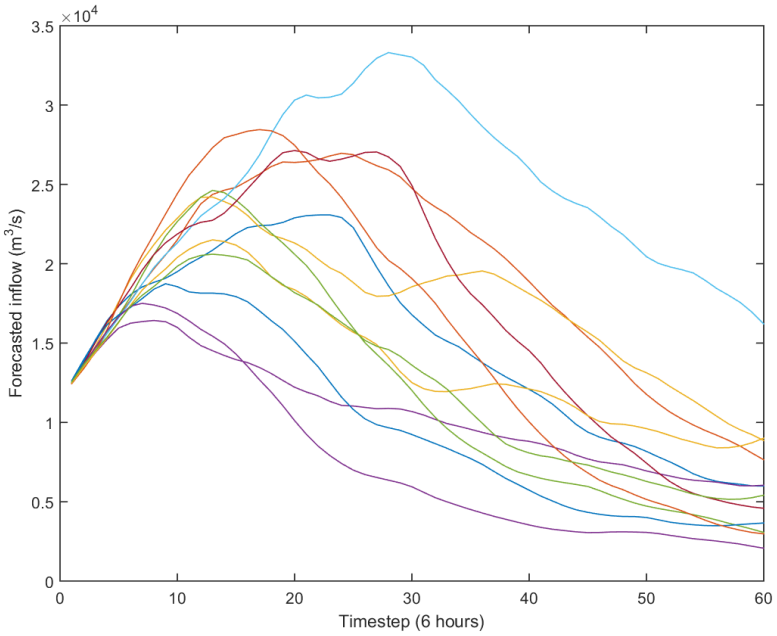


E.3 Scenario 54

E.3.1 Trajectory ensemble

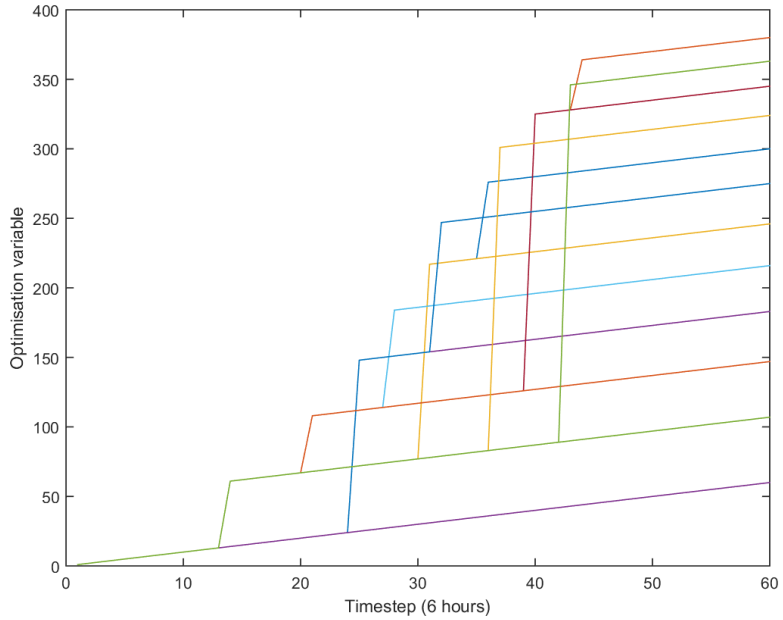


E.3.2 Reduced ensemble

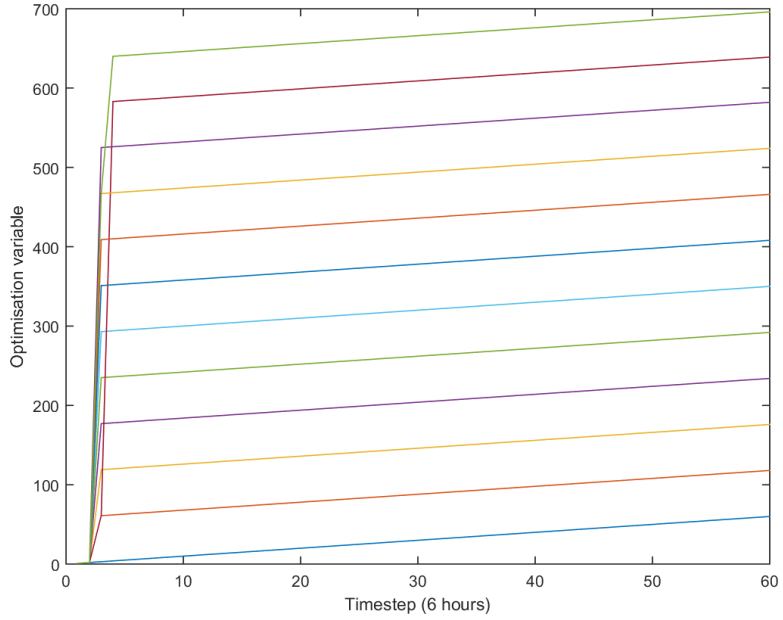


E.3.3 Nodal partition matrix

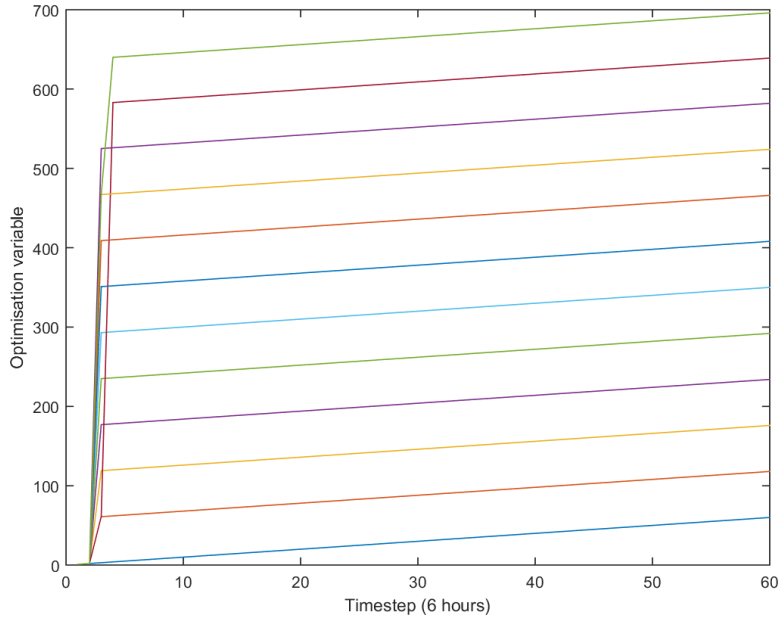
Base Case



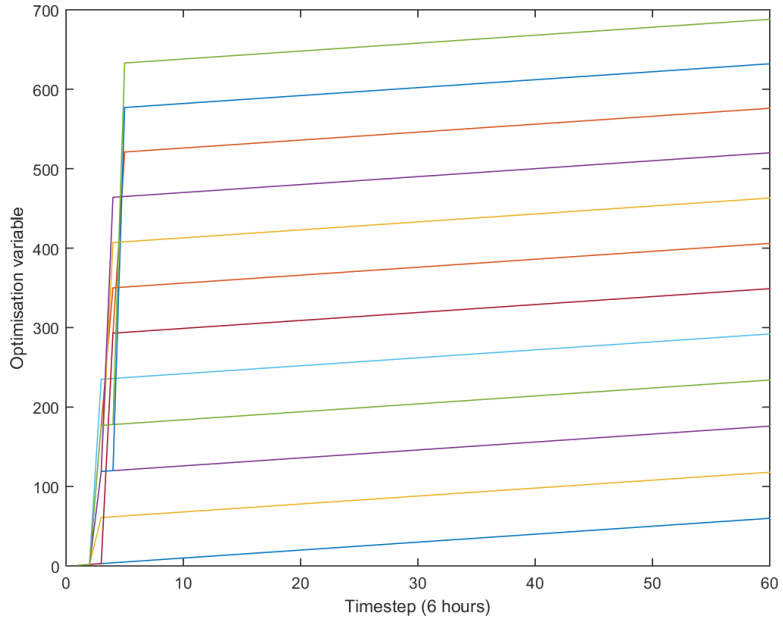
Perfect model – CPTEC



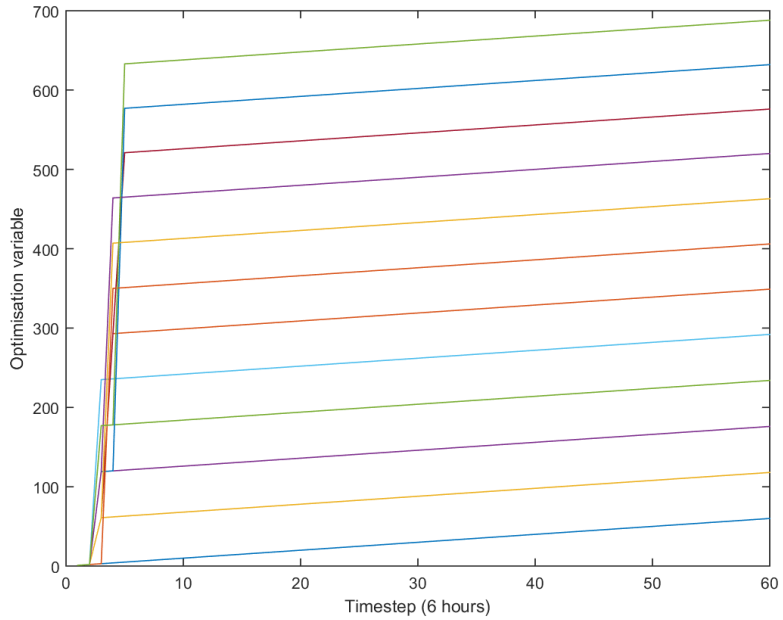
Perfect model – ECMWF



Imperfect model – CPTEC



Imperfect model – ECMWF

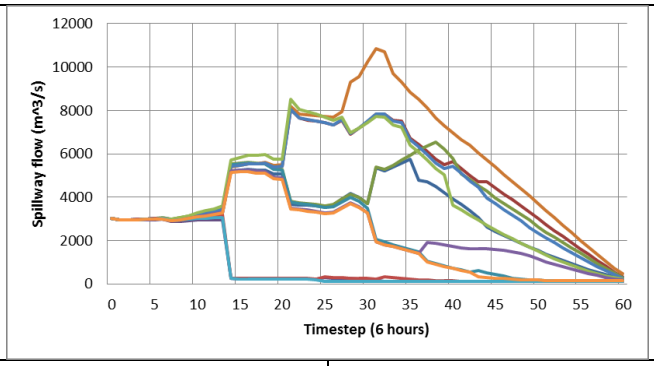


E.3.4 Results

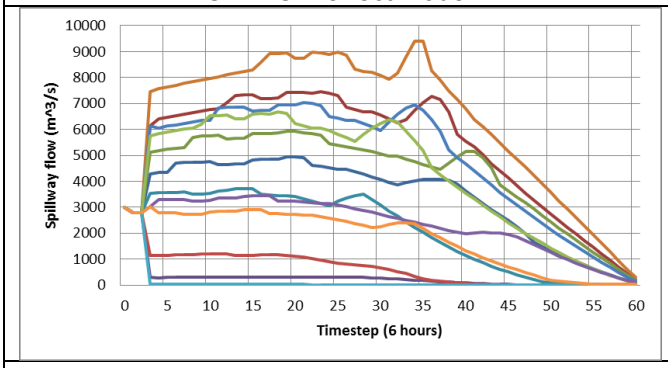
Scenario 54 Results		Quasi-optimised cost function score				
Initial reservoir level (m above datum)	Probability (%)	Base case	Perfect model		Imperfect model	
			CPTEC	ECMWF	CPTEC	ECMWF
31	3%	52359	46733	46924	46795	46871
33	36%	68508	62578	62401	63045	62594
35	61%	94283	88349	88433	88857	88504

Spillway Flow - Initial level 31 m above datum

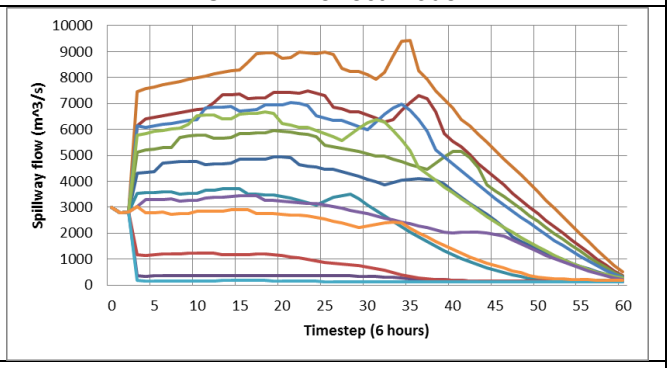
Base Case



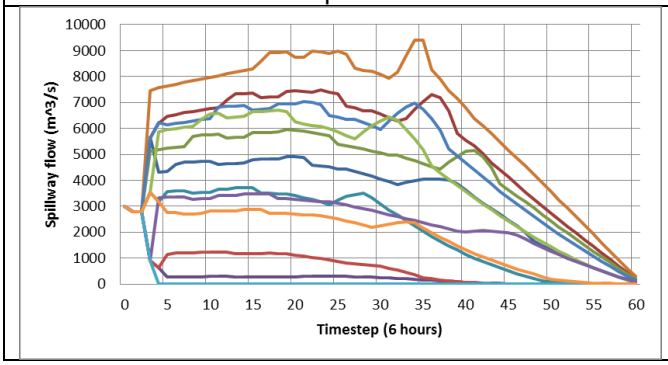
CPTEC 'Perfect Model'



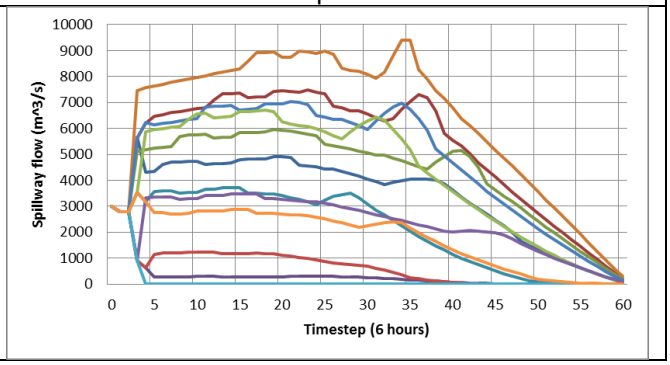
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

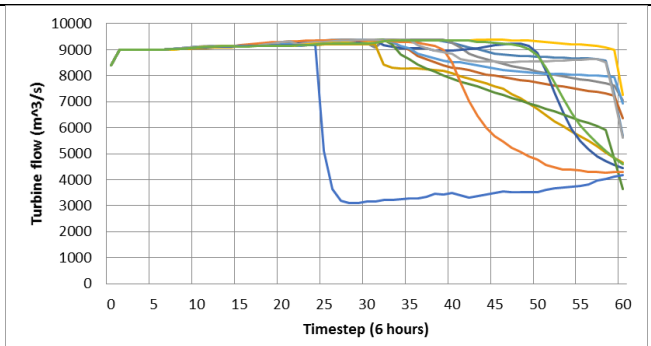


ECMWF 'Imperfect Model'

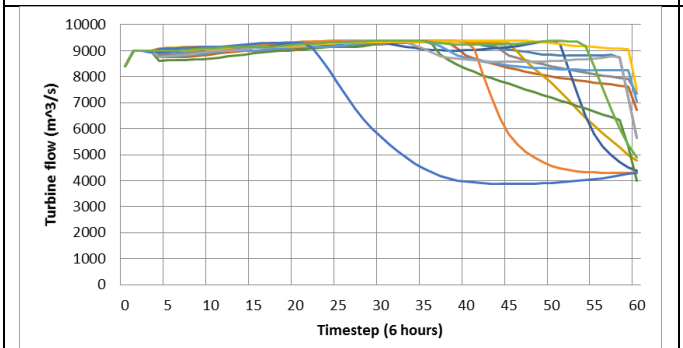


Turbine Flow - Initial level 31 m above datum

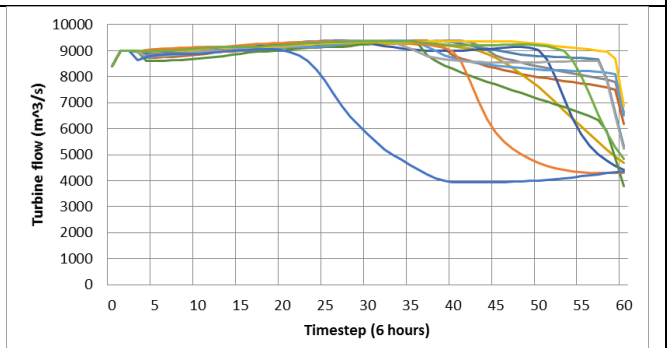
Base Case



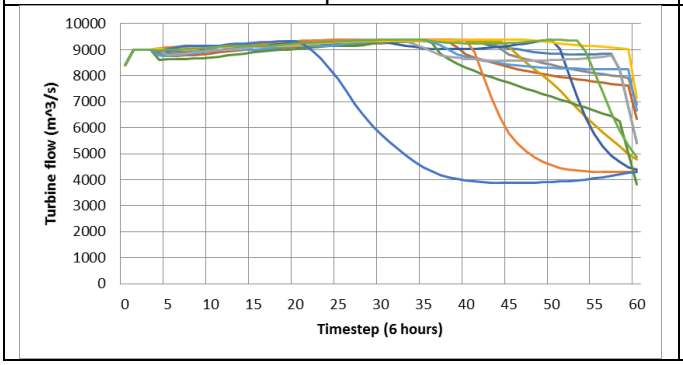
CPTEC 'Perfect Model'



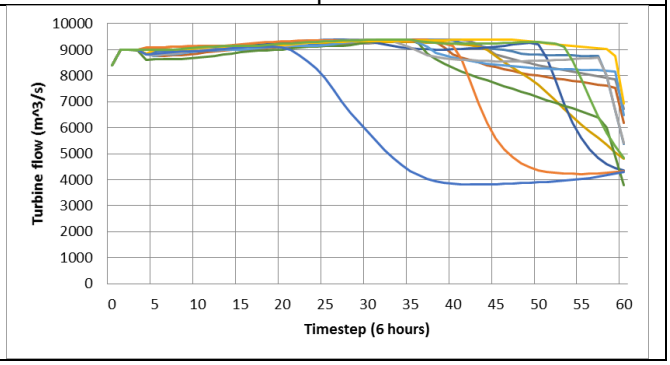
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

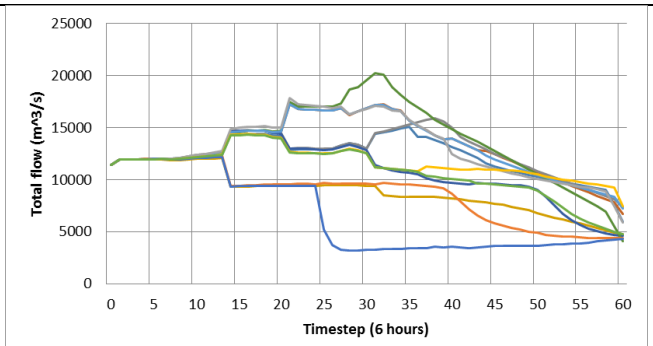


ECMWF 'Imperfect Model'

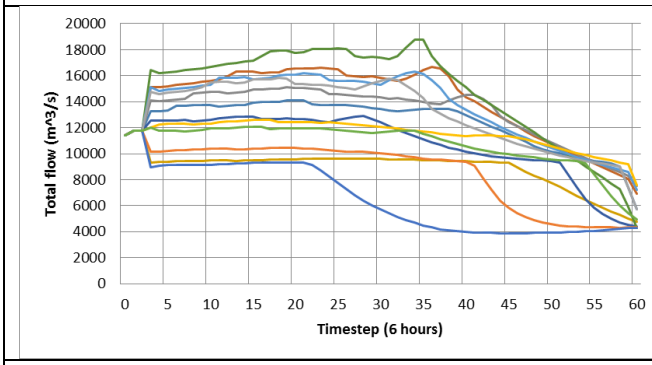


Total Flow - Initial level 31 m above datum

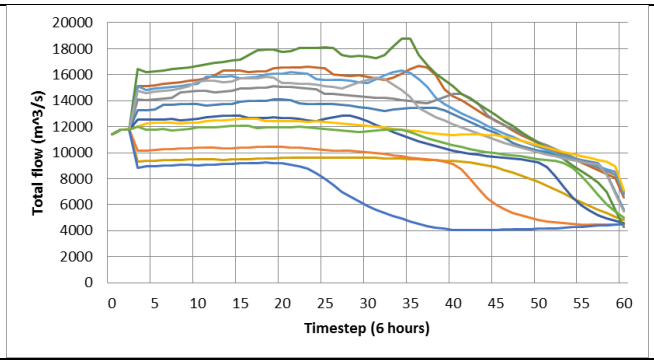
Base Case



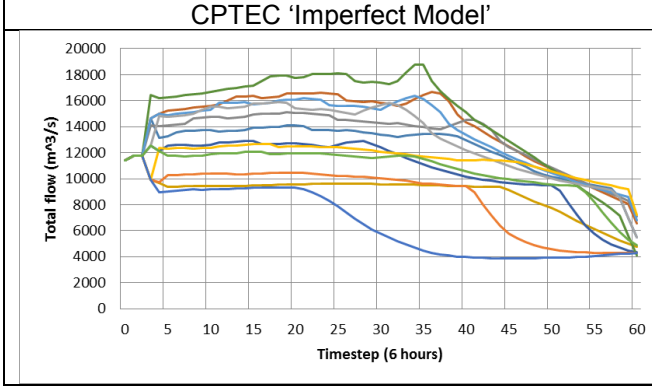
CPTEC 'Perfect Model'



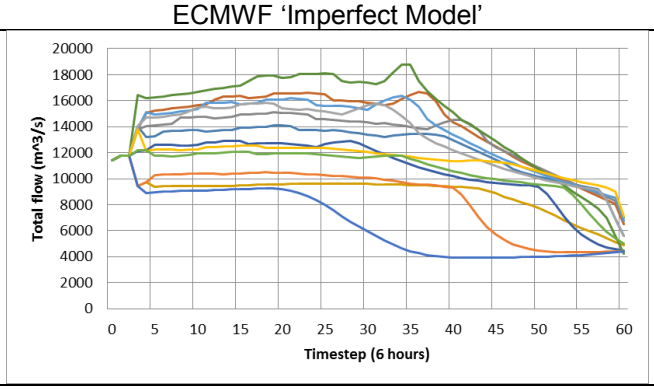
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

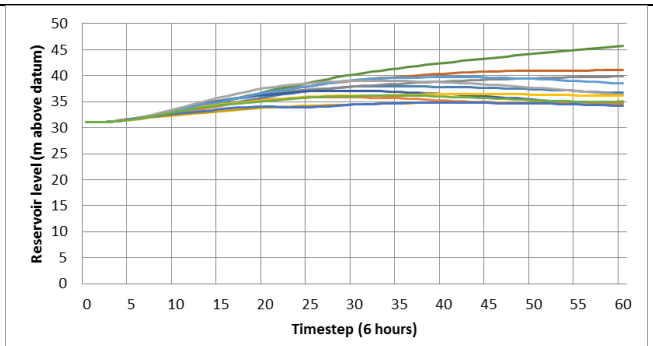


ECMWF 'Imperfect Model'

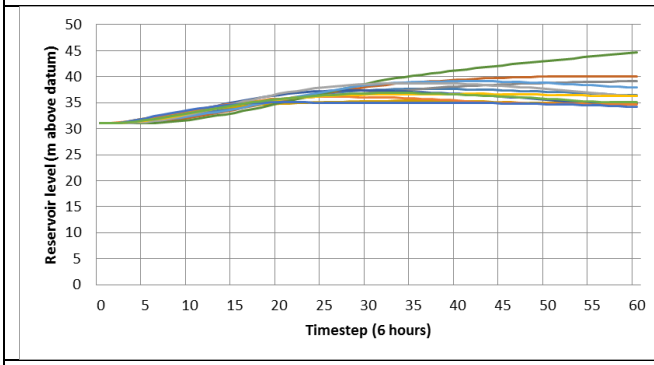


Reservoir Level - Initial level 31 m above datum

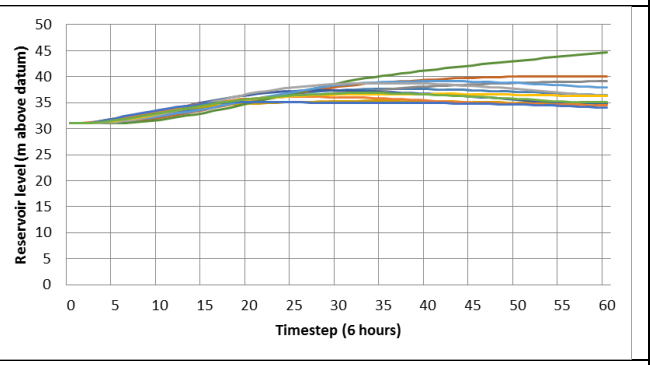
Base Case



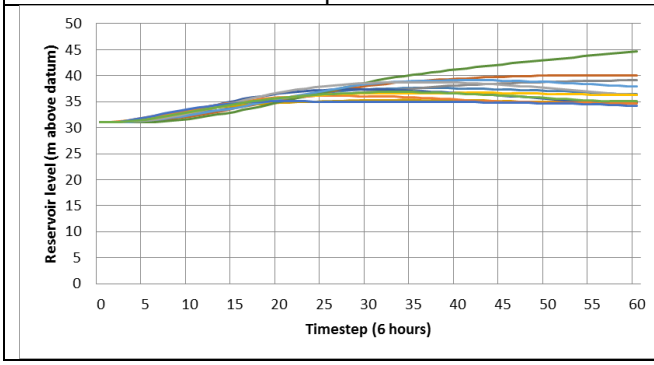
CPTEC 'Perfect Model'



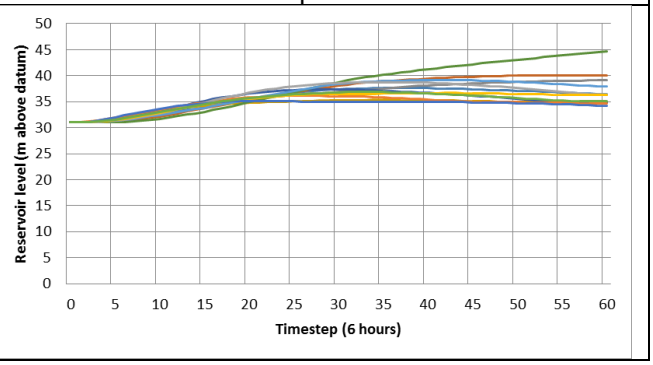
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

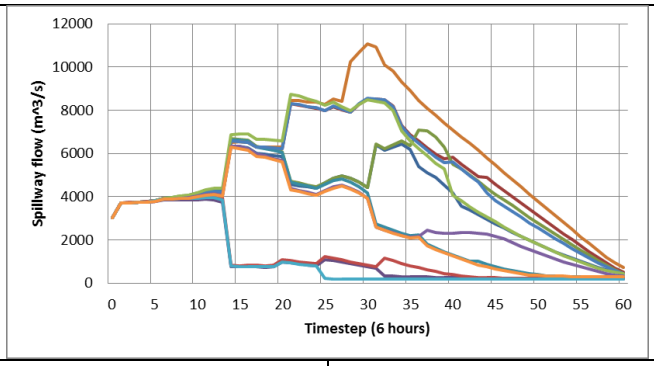


ECMWF 'Imperfect Model'

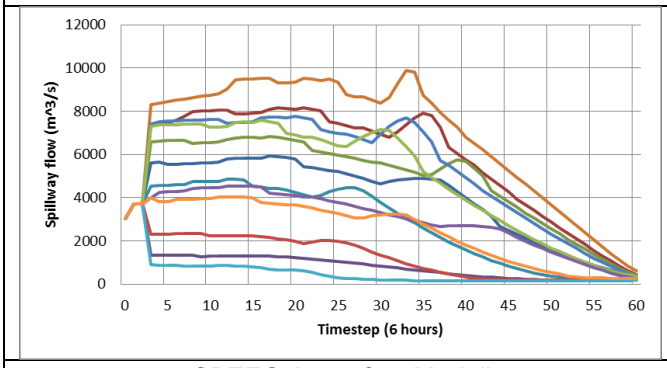


Spillway Flow - Initial level 33 m above datum

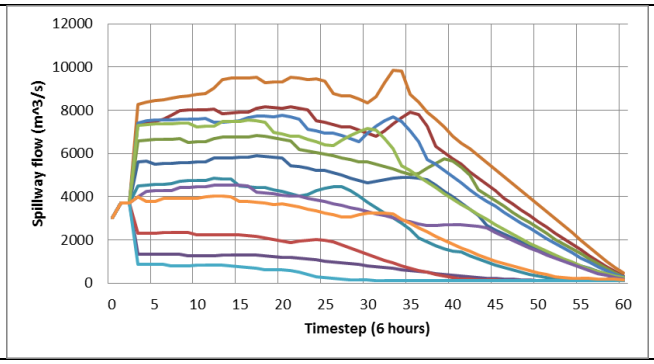
Base Case



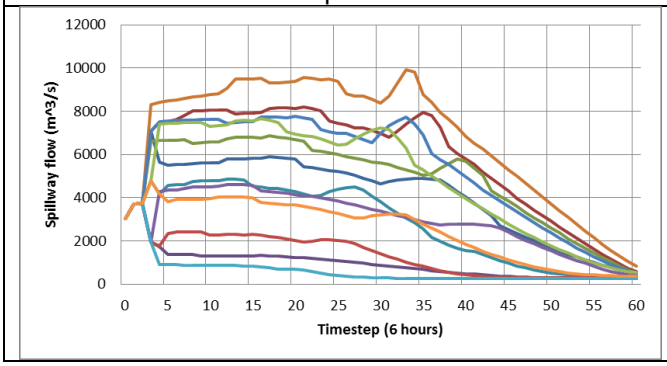
CPTEC 'Perfect Model'



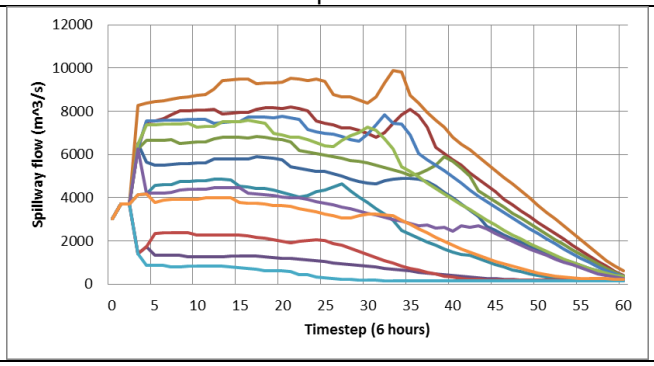
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

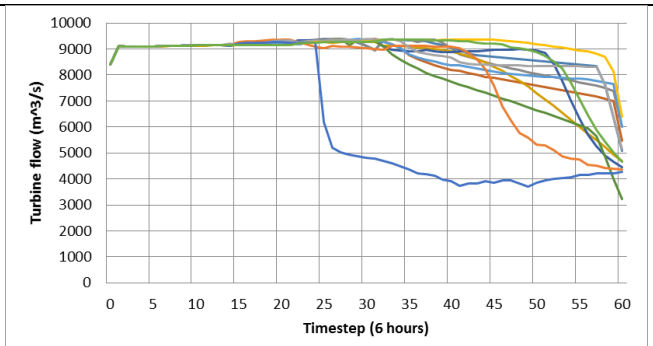


ECMWF 'Imperfect Model'

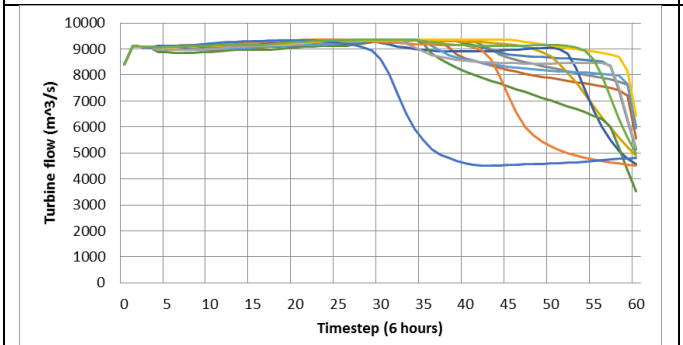


Turbine Flow - Initial level 33 m above datum

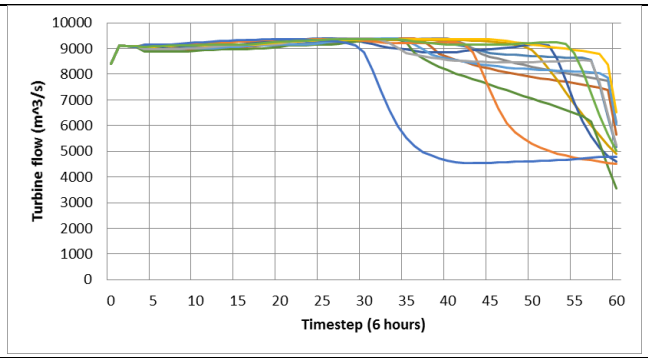
Base Case



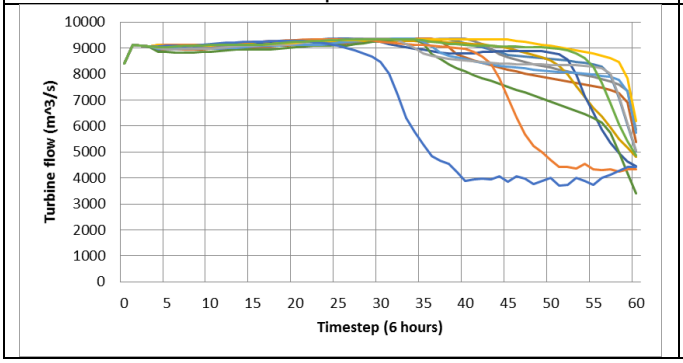
CPTEC 'Perfect Model'



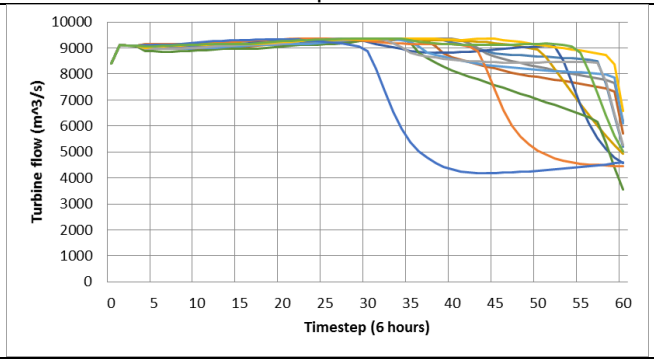
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

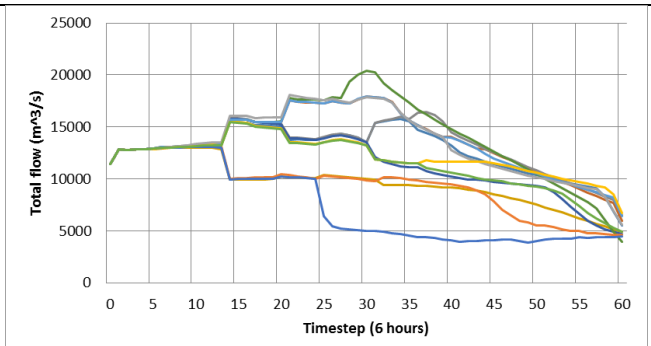


ECMWF 'Imperfect Model'

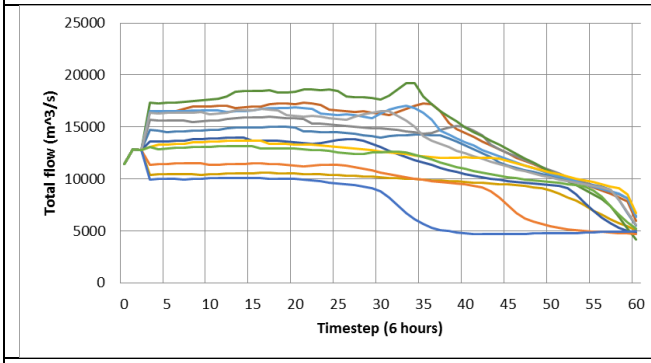


Total flow - Initial level 33 m above datum

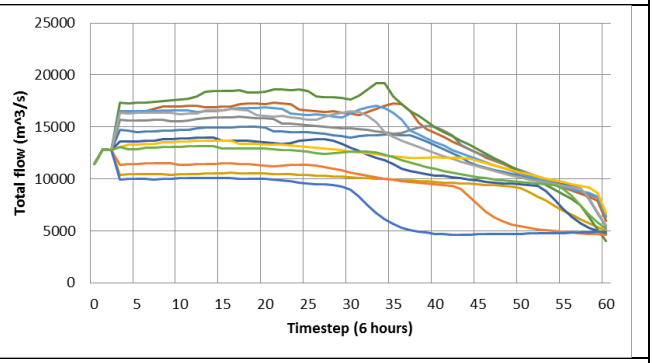
Base Case



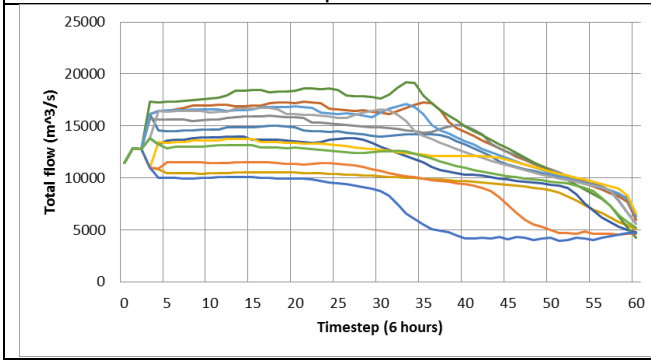
CPTEC 'Perfect Model'



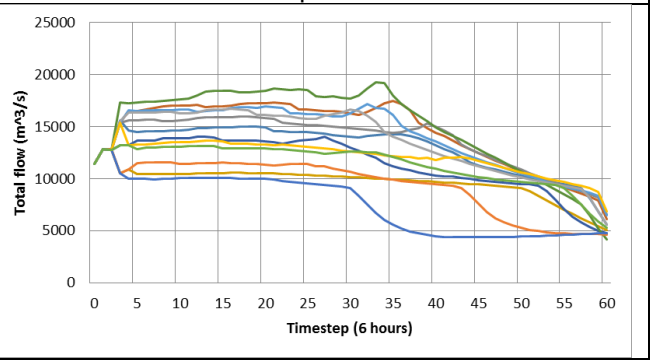
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

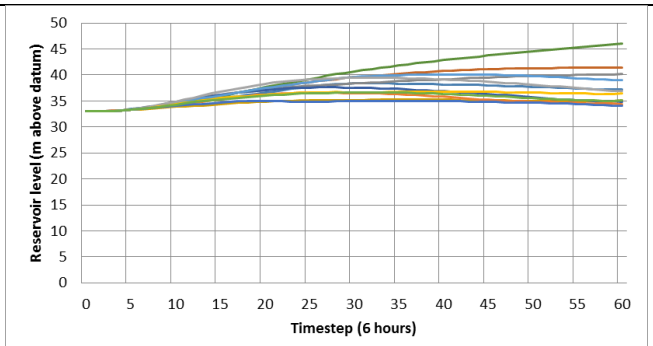


ECMWF 'Imperfect Model'

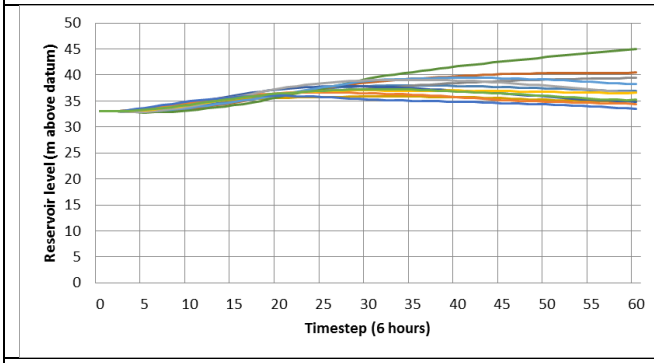


Reservoir Level - Initial level 33 m above datum

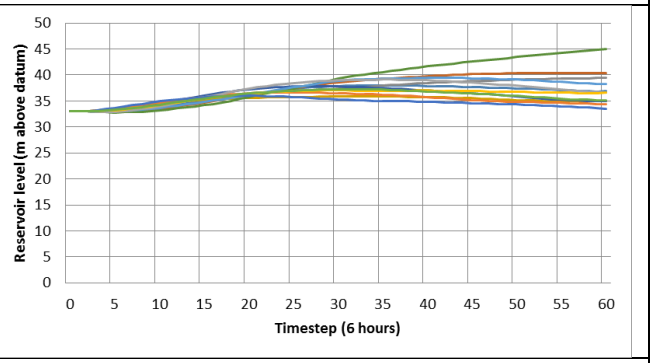
Base Case



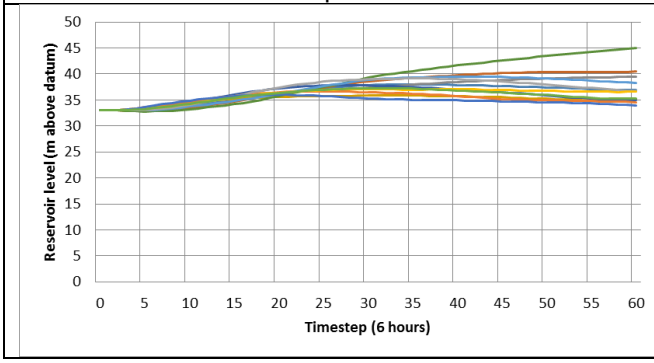
CPTEC 'Perfect Model'



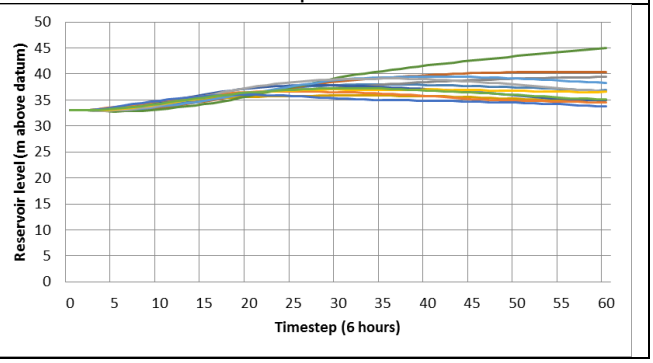
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

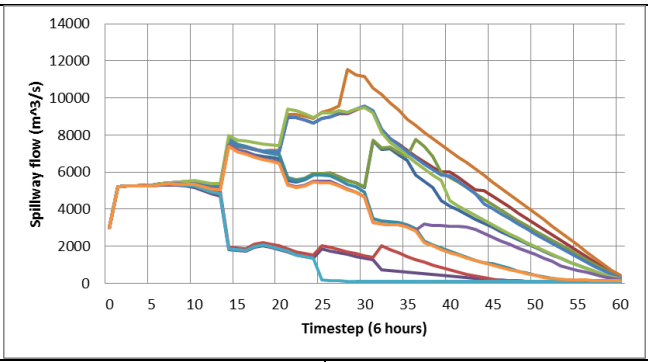


ECMWF 'Imperfect Model'

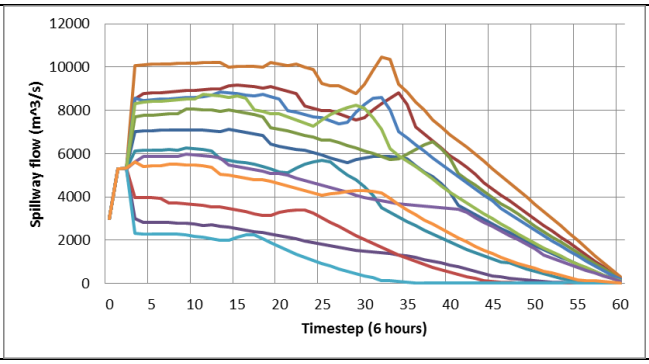


Spillway Flow - Initial level 35 m above datum

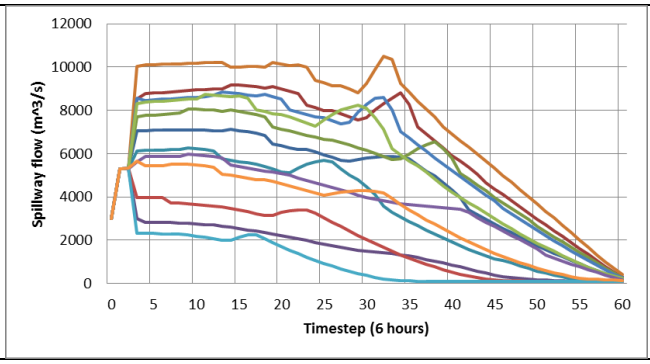
Base Case



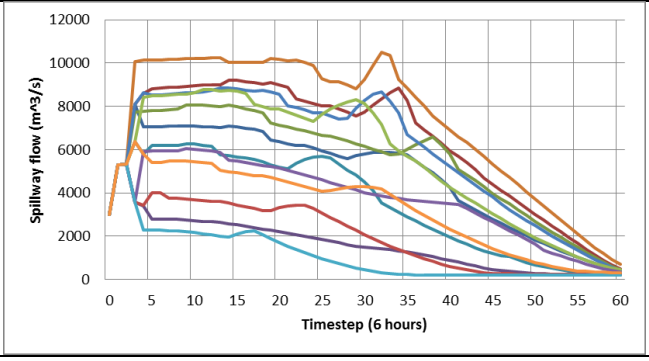
CPTEC 'Perfect Model'



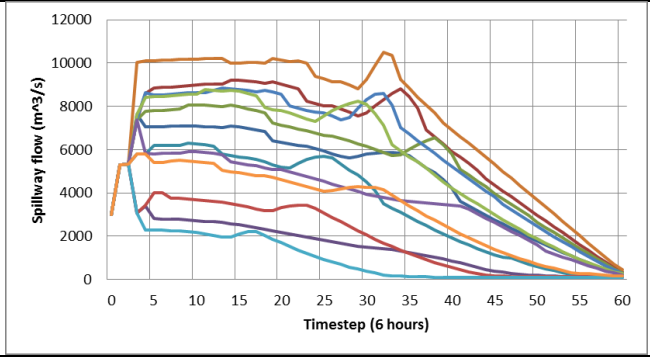
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

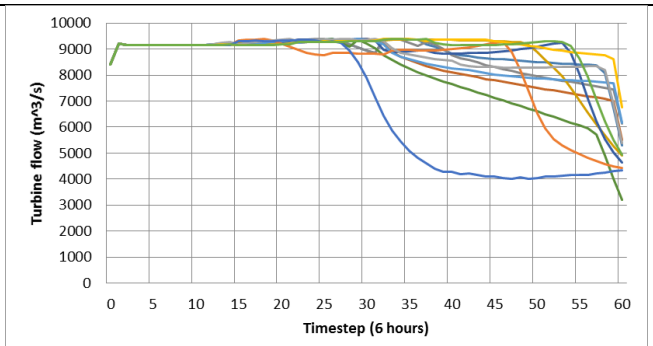


ECMWF 'Imperfect Model'

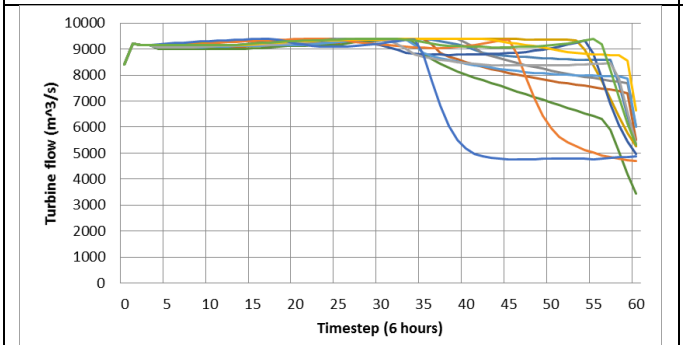


Turbine Flow - Initial level 35 m above datum

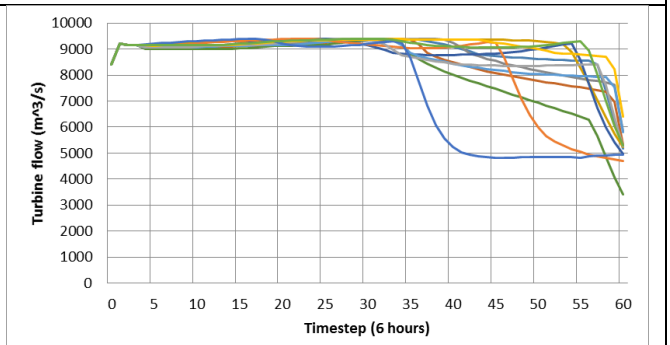
Base Case



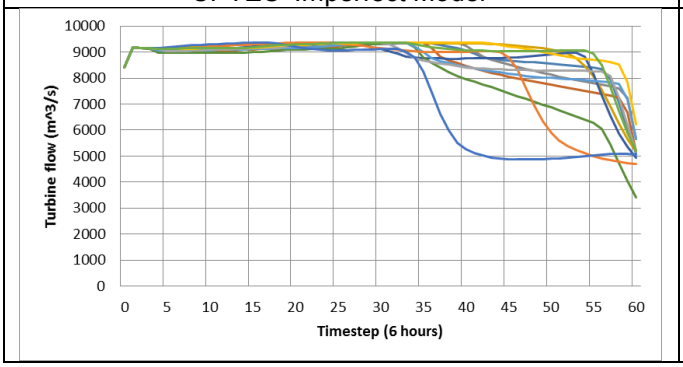
CPTEC 'Perfect Model'



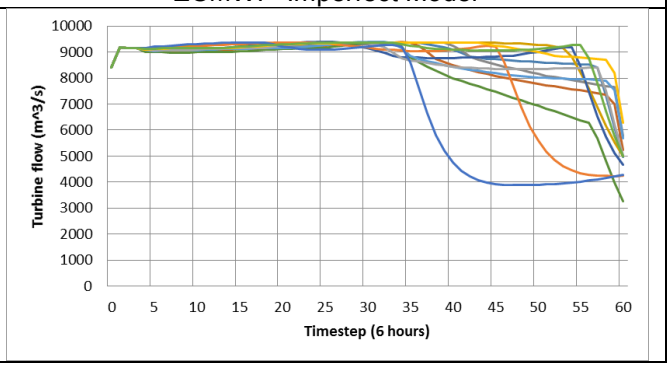
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

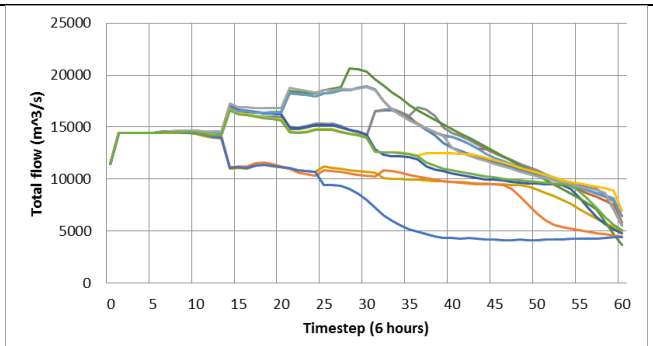


ECMWF 'Imperfect Model'

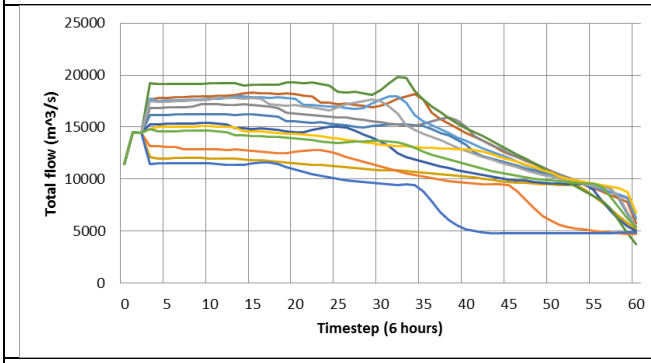


Total Flow - Initial level 35 m above datum

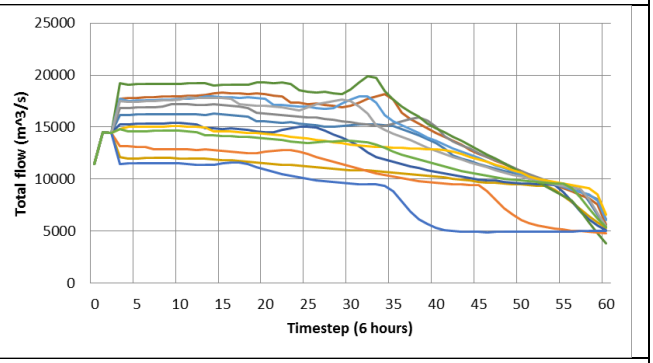
Base Case



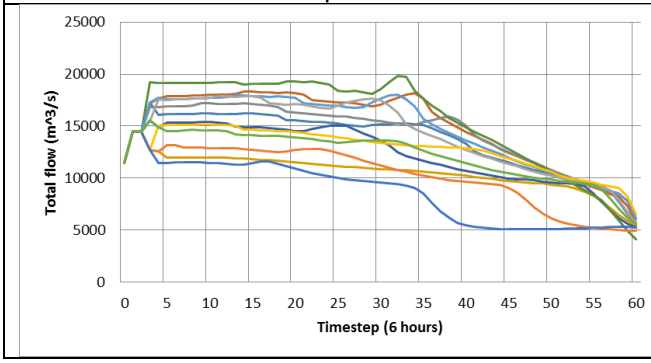
CPTEC 'Perfect Model'



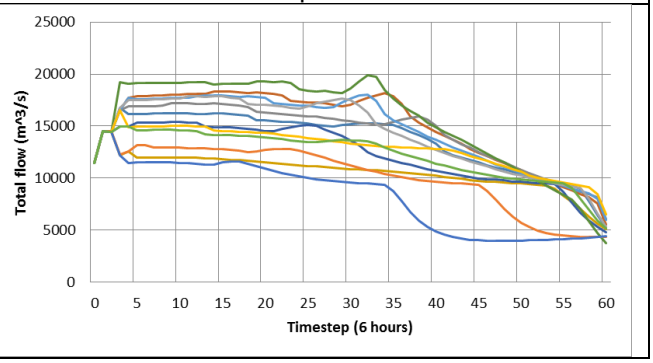
ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'

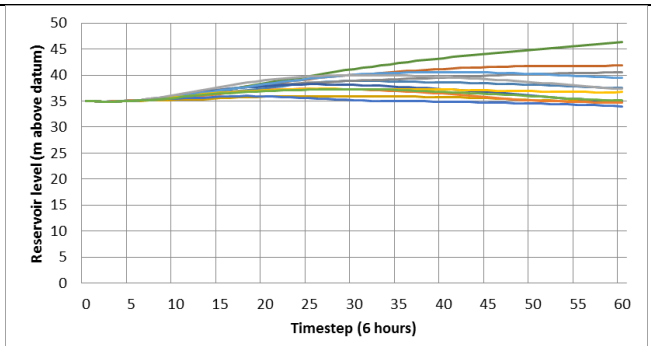


ECMWF 'Imperfect Model'

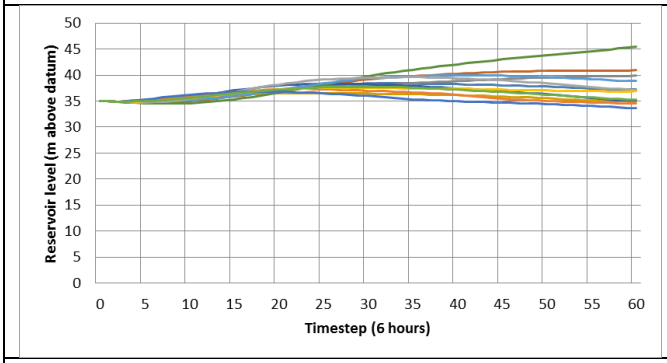


Reservoir level - Initial level 35 m above datum

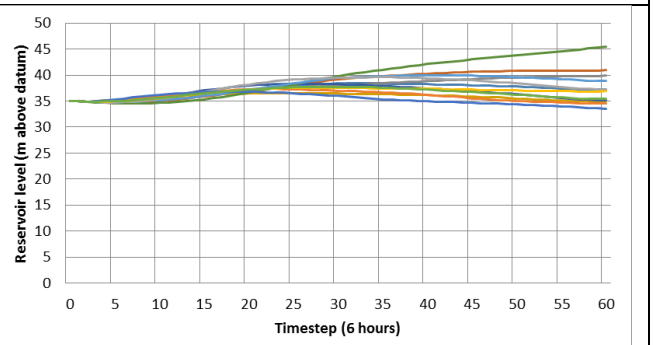
Base Case



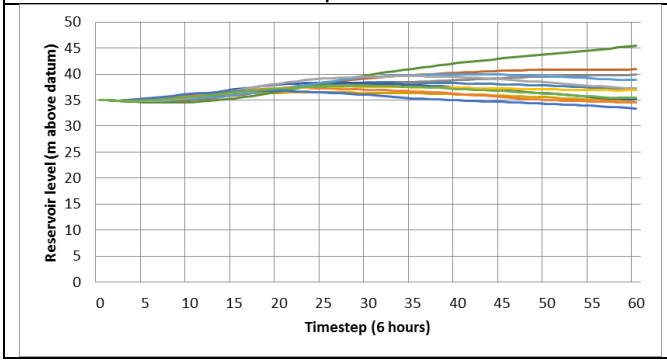
CPTEC 'Perfect Model'



ECMWF 'Perfect Model'



CPTEC 'Imperfect Model'



ECMWF 'Imperfect Model'