PEST hydrology calibration process for source catchments – applied to the Great Barrier Reef, Queensland

G.L. McCloskey [a], **R.J. Ellis** [a], **D.K. Waters** [a] **and J. Stewart** [b]

[a] *Department of Environment and Resource Management, Queensland Government;* [b] *Catchment Research Pty Ltd*
Email: Gillian.McCloskey@derm.qld.gov.au

Abstract: The rapid development in computer technology over the past decade has made automated calibration of models common practice. PEST, a model-independent parameter estimation tool, is one such tool used extensively in groundwater model calibration. PEST was linked to the eWater CRC Source Catchments water quantity and quality model to calibrate the parameter values for component SIMHYD rainfall runoff and Laurenson non-linear flow routing models. The calibration process was applied across 450,000 km^2 in the Great Barrier Reef (GBR) Queensland. Six models were built, with catchments ranging in area from 10,000 – 156,000 km^2. A consistent approach to model calibration was applied across all models. Unique SIMHYD parameter sets were generated for the three dominant Hydrologic Response Units (HRU): forest, grazing and agriculture.

A three-part objective function was used, incorporating daily and monthly flows, and exceedence time at 9 – 92 gauging locations throughout each project; all three objective function components were given equal weighting. The contributions of each gauge to the initial value of the objective function were equalised to remove bias caused through magnitude of flow volumes. A regionalisation approach, based on contributing areas to each observation gauge, was used. The purpose of regionalisation was to rationalise the number of parameter sets that PEST needed to derive during the simultaneous calibration process, and enabled only those identifiable parameters, in regions that were underperforming, to be adjusted. For example, if a HRU did not occur in a region, or comprised a very small area of that region, it could be 'turned off', so that those parameters most likely to impact calibration could be adjusted more quickly by PEST. In projects representing very large areas, a supervised singular value decomposition approach was taken, whereby PEST determined a defined set of parameter combinations most significantly influencing calibration, and these parameter combinations were adjusted during the calibration process. In all six GBR models, calibration was deemed sufficient when modelled volume was calculated within 20% of measured volume, the daily Nash-Sutcliffe was greater than 0.5, and the monthly Nash-Sutcliffe greater than 0.8, at each of the utilised gauging locations. The results of the hydrology calibration process ensure that the six GBR models are set up with a platform that is considered suitable to estimate runoff rates from the identified HRUs, from which to simulate runoff and erosion processes, and ultimately to estimate loads to the GBR. The design of this hydrology calibration was intended to provide stream network flow and storage characteristics reasonable enough to route constituent loads in a reasonable manner, but not to provide water quantity estimates of an accuracy deemed necessary for water supply and regulation needs.

Using PEST to parameterise the SIMHYD Rainfall Runoff and Laurenson non-linear flow routing models in Source Catchments was an efficient and repeatable methodology for simultaneous calibration across large catchments. Improvements to the process outlined above would include reconsidering the most appropriate objective functions, and using local and 'expert' knowledge to guide the parameter starting values and ranges.

Keywords: Source Catchments, hydrology calibration, SIMHYD, PEST

1. INTRODUCTION

Targets for improvements in water quality entering the Great Barrier Reef (GBR) have been set through the Reef Water Quality Protection Plan (Reef Plan). The adoption of improved agricultural practices that result in reductions in sediment, nutrient, and herbicides leaving the paddock is being encouraged through funding from the Reef Rescue Program. To measure and report on progress towards the targets set in Reef Plan and Reef Rescue, a program has been established that combines monitoring and modelling at paddock through to catchment and reef scales.

The program, which is a collaborative effort between State and Federal Governments, research organisations, Natural Resource Management (NRM) groups and agricultural industry, aims to provide evidence of links between land management activities, water quality, and reef health. The framework for the overall design involves monitoring and modelling a range of attributes including management practices and water quality at the paddock, sub-catchment, catchment, and marine scales. Further details on the P2R program are outlined in Carroll *et al,* 2011. The eWater CRC Source Catchment modelling framework is used to report end-of-catchment loads of key pollutants for each catchment in the GBR for the baseline condition (1986 to 2009) and for changes relative to the baseline every year thereafter (2010-2013).

Hydrology model calibration is a key initial step in the modelling process. Ellis *et al* (2009) described a method for coupling PEST and Source Catchments in a small GBR catchment. PEST is a model-independent parameter estimation tool that is commonly used in groundwater model calibration. Source Catchments is a water quality and quantity modelling framework, developed by eWater CRC. The results of Ellis *et al* (2009) indicate a significant improvement in calibrating a model in the Source Catchments environment (through the use of PEST), when compared with a secondary environment, such as the Rainfall Runoff Library. The main improvement being that the calibration process is able to consider all internal interactions in the model.

In this paper, the application of this methodology to six NRM regions in the GBR is described.

2. SOURCE CATCHMENTS MODEL SETUP FOR HYDROLOGY CALIBRATION

Six NRM regions have been modelled, extending from Cape York, to the Burnett-Mary. Table 1 provides a summary of each of these NRM regions. The challenges of applying a consistent approach to hydrological model calibration across a broad range of model catchment scales (9,130 – 156,610 km2), diverse land uses and climatic variability from Cape York in the north to Burnett Mary in the south of Queensland, is also highlighted in Table 1.

Table 1 Summary of the Great Barrier Reef NRM regions

Region	Area (km^2)	Major land uses	Climate
Cape York Peninsula (CYP)	42, 793	Grazing – 52% Nature Conservation – 45% Other – 3%	Distinct wet and dry season Rainfall: 800mm/yr – 2000mm/yr
Wet Tropics (WT)	21,710	Nature Conservation – 43% Grazing – 32% Sugarcane – 8% Other – 17%	Tropical Rainfall: 750 mm/yr – 5000 mm/yr
Burdekin (Burd)	140, 668	Grazing – 90% Nature Conservation 5% Other 5%	Sub-tropical Rainfall: 600mm/yr – 2000mm/yr
Mackay-Whitsunday (M-W)	9, 130	Grazing – 47% Sugarcane – 19% Nature Conservation – 16% Other – 18%	Humid, tropical Rainfall: 1000mm/yr – 3000mm/yr

McCloskey *et al.,* PEST hydrology calibration process for source catchments – applied to the Great Barrier Reef, Queensland

Fitzroy (Fitz)	156, 610	Grazing - 82 % Forest & Nature Conservation - 10% Cropping - 7%	Sub-tropical like conditions in north eastern areas; temperate in south west Rainfall: 500mm/yr – 1800mm/yr
Burnett-Mary (B-M)	52, 818	Grazing – 70% Forestry – 13% Conservation – 8% Other – 9%	Subtropical conditions Rainfall: 800mm/yr – 1800mm/yr

Regional models were built in the Source Catchments modelling framework, which comprise sub-catchments connected through a series of nodes and links which represent the stream network. The sub-catchment networks were developed using a hydrologically correct DEM (resolution 100m in most catchments, 270m in Burd region). A $30km^2$ stream identification threshold was used for the smaller areas of the WT and M-W, while a $50km^2$ threshold was used in the remainder of the regions.

The SIMHYD rainfall runoff model was used to generate runoff, and flow was routed through the nodes and link network via the Laurenson Flow Routing model. Daily rainfall and evapotranspiration files generated for each sub-catchment are used by the Source Catchments model to calculate daily runoff. Rainfall and potential evaop-transpiration (PET) inputs were derived from the DERM Silo Data Drill database (SILO, 2009).

Water storages were included in regional models where their capacity is greater than 10,000 ML. For small regions such as M-W, there are many smaller dams and weirs that do not individually meet the 10,000 ML criteria for model inclusion, however their combined capacity was deemed to have a significant impact on runoff characteristics, and they were therefore included in the model. Irrigation and town water supply extractions were represented in the models and applied as time series extractions at nodes. The timeseries data was supplied by the Qld State Government Water Resource Planning group. This data was generated from previous Integrated Quantity and Quality model (IQQM) model runs.

Table 2 Summary of GBR models composition

Region	No. of SC	No. of links	No. of nodes	Storages
CYP	546	547	547	0
WT	450	451	451	3
Burd	1568	1569	1569	5
M-W	191	192	192	5
Fitz	1976	1977	1977	3*
B-M	597	598	598	14

*Major storages were not initially included in the Fitzroy model due to computational constraints. Instead, the major storage (Fairbairne Dam) was included in the model using an inflow timeseries.

The Source Catchments modelling framework allows for individual rainfall, runoff, and stream routing parameters to be applied to each Functional Unit (FU) (i.e. landuse category), within each sub-catchment. Therefore it was essential to consolidate the number of FU's, to rationalise model run times for the hydrology calibration process, given the size of the modelled regions. A base set of 10 FU's, or land use categories were represented in each model initially (open and closed grazing, forestry, horticulture, irrigated cropping, dryland cropping, nature conservation, urban, water, and other) and applied to each region, with additional regional specific FU's also included where appropriate (for example, banana's in the Wet Tropics, or sugarcane). To minimise model run time, the 10 FU's were aggregated into three dominant Hydrological Response Units (HRUs): grazing (open or cleared), agriculture (cropping), and forest. HRU selection was based on a combination of soil type and land use.

To rationalise the number of model runs required for the hydrology calibration process, it was necessary to limit the number of stream gauges used in the automated calibration. For example, in the Burd (140, $668km^2$) out of a possible 110 gauges that could be used, 37 of the sites in key locations and having long term records were utilised. Gauges were selected based on the following criteria: length of record greater than five years, less than 80% missing data, site rating was of a suitable quality, location relative to upstream storages, and their relative contribution in terms of drainage area.

McCloskey *et al.,* PEST hydrology calibration process for source catchments – applied to the Great Barrier Reef, Queensland

In order to further rationalise the number of parameter sets derived during the calibration process, a regionalisation approach was adopted, based on contributing area. That is, the upstream area of each gauge and or the area between nested gauges was considered one region. Where there were ungauged areas, the nearest neighbour approach was taken (using SIMHYD parameters from the nearest sub-catchment), as opposed to a regionalisation approach (using parameters from those catchments with similar physical characteristics) (Chiew and Siriwardena, 2005; Post *et al*, 2007).

A 'PEST for Source Catchments' plug-in was written, which assisted with the preparation of input files required for the PEST process that implemented the regionalisation, as PEST does not explicitly account for the spatial structure of the Source Catchments project.

3. PEST

PEST (Doherty, 2009) is a model-independent, parameter estimation tool, which was linked to Source Catchments, to determine the parameter values for the SIMHYD rainfall runoff, and Laurenson flow routing component models. PEST operates largely via batch and instructional text files. DERM created a number of project specific tools to automate the compilation of these files, where possible. The PEST utility program TSPROC.exe (Time Series Processor) was also used to create the files used by PEST (PEST control file), and also to manipulate the modelled time series, and present the statistics to PEST for assessment (Stewart, 2010).

An equally weighted, three-part objective function was used. The three-part objective function incorporated monthly flows, exceedence values, and log transformed daily flows. The monthly flow volume component works to ensure modelled volumes match measured volumes over long periods, the exceedance values work to ensure the flow volumes are proportioned well into baseflows and event flows, while the log transformed daily flows work to obtain replication of the hydrograph shape (Stewart, 2010). Each of these were weighted equally, which overcomes inadvertent weighting of small value components in each of the observation groups.

That is, by using equal weighting, both large and small value components are given the same importance. Future research will explore different objective functions, and different weightings. For example, calibrating to high flows to improve load estimates.
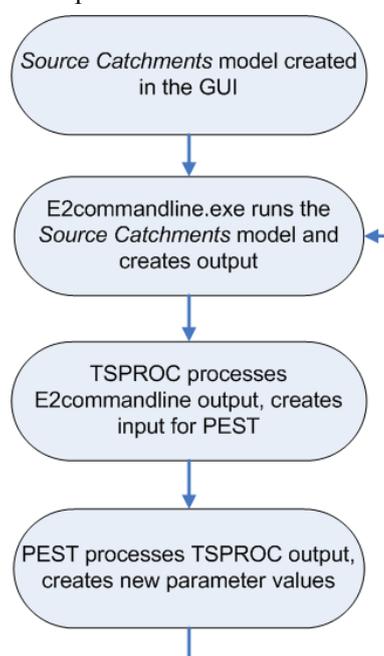
Figure 1 Source Catchments – PEST interaction (Stewart, 2010)

The interaction between PEST and Source Catchments is outlined in Figure 1, above. Final parameter sets were determined following numerous iterations, often between five and 10, which equates to upwards of 5,000 model runs in total. The final 'calibration' parameter set was achieved when the change in objective function between iterations showed minimal change and the 'fits' were satisfying to the modeller with minimal deviation from the 'preferred' values between iterations (Ellis *et al* 2009).

4. RESULTS

Model performance was assessed by calculating the Nash-Sutcliffe Coefficient of Efficiency (E) for daily and monthly flows, as well as the difference in total volumes between measured and modelled flows (% difference). The Nash–Sutcliffe coefficient (Nash and Sutcliffe, 1970) is commonly adopted for evaluating the goodness-of-fit for simulated hydrographs. The calibration at a gauge was deemed "acceptable" when daily E values were greater than 0.5, monthly E values were greater than 0.8, and the total modelled runoff volume for the run period was within ±20% of measured flows. Flow duration curves (R^2 between measured and modelled flow duration curves) were also produced, as an additional method of assessing model performance.

In the larger catchments (Fitz and Burd) it was necessary to use Singular Value Decomposition-Assist (SVD-A), to reduce the processing times for PEST. SVD-A is a utility program available with PEST which enables the user to determine combinations of parameters that are most likely to influence the calibration process. Modellers then set the number of 'super' parameters PEST uses, thereby dramatically reducing the total number of parameters, and model run times. Super-parameters are linear combinations of 'real' or base parameters. Further details of SVD-A is provided in Ellis (2009), Tonkin & Doherty (2005), and Doherty (2009). For example, for the Burd, the initial 851 available parameters (number of regions x number of FUs x 7 adjustable SIMHYD parameters + number of regions x 2 (routing parameters)) were reduced to 200 super parameters, reducing the run time for one iteration from seven days to only one (utilising Parallel PEST).

Table 3, below, provides an overview of the results of the PEST calibration for the six GBR models. Daily Nash-Sutcliffe values for four of the six regions are above 0.47, while five of the six regions achieved average monthly E values above the 0.8 target. Average volume differences were all within the ±10% target, with five of the regions being within 7% of the observed values.

Table 3 Summary PEST results for all GBR models

Region	No. of gauges used in calibration	Daily E average	Monthly E average	Average annual flow (ML)	Average % volume difference
CYP	18	0.58	0.84	17, 748, 740	+6.9
WT	21	0.67	0.89	21, 492, 487	-9.0
Burd	37	0.47	0.76	9, 169, 285	-2.02
M-W	9	0.81	0.91	5, 213, 367	+6.8
Fitz	84	0.24 (0.47)*	0.42 (0.82)*	6, 535, 179	-5%
B-M	32	0.24	0.81	2, 349, 534	+1.31

* Poor calibration at one gauge in the Fitzroy region has resulted in the average values being skewed. The values in parentheses are those averages with the outlier removed.

A brief summary of CYP, M-W, and Burd regions process and results, is provided below.

The CYP region, although data poor, calibrated well. For the 18 gauges, 83% of gauges met all three criteria, and 11% met two of three criteria. An issue in the CYP calibration process was the tendency for the model to generate high soil moisture storage capacity (SMSC) values, close to the maximum default value (500mm). The default maximum was increased to 750mm in some catchments. This may be due to the simplistic representation of the runoff generation process adopted by SIMHYD using only seven parameters. In wet tropical areas of north Queensland, there may be significant losses to groundwater, SIMHYD does not account for this loss hence forcing the SMSC to a large number to compensate for the excess water in the model. In the physical environment, many of the rivers in the southern CYP are large, sandy bed streams, where water is lost in the stream, to groundwater. Future work will assess the applicability of alternative rainfall runoff models within Source Catchments which account for groundwater losses.

Table 4, below, presents the mean SMSC values for all HRUs. The data indicates that land use categories in the model that were assigned forest SIMHYD parameters, had higher SMSC values than cleared grazing areas. This result supports the decision to split the catchments into different HRUs. Future work will explore the impact on model performance of representing different numbers of HRUs.

Table 4 CYP HRU mean SMSC values

HRU	Mean SMSC
Grazing	365
Forest	474
Agriculture	231

McCloskey *et al.,* PEST hydrology calibration process for source catchments – applied to the Great Barrier Reef, Queensland

Calibration of the large Burd region was a time-consuming process. Gauge selection criteria reduced the available number of gauges from 110 to 37 gauges. In terms of the daily and monthly E, 47% of gauges were classed as well parameterised. Sixty percent of the gauges met the volumetric error criteria. The volumetric error remained relatively constant as discharge increased, suggesting that for volume, suitable PEST calibration was achieved at both low and high flows (Figure 2). Time series plots of observed and modelled runoff at a range of scales across the Burd indicate an under prediction of daily peak flow, however the fit at a monthly and yearly scale was regarded as quite acceptable with E values greater than 0.7.
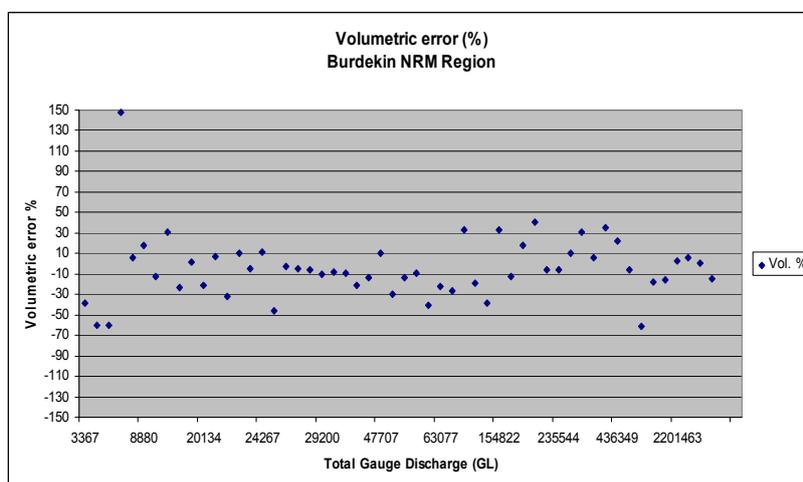
Figure 2 Volumetric error and total gauge discharge in the Burdekin NRM region

The results for the M-W region were excellent for the nine gauges used in calibration. Due to the size of the model run time was short, and enabled many PEST iterations to be undertaken to achieve these results. A key outcome of the M-W project was the importance of weir inclusion. As stated, initially only storages with a capacity of 10,000 ML or above were included in the models. As a result of this, four smaller weirs were excluded from the model. Throughout the calibration process, and from discussion with local hydrographers, it became evident that combined these four weirs were critical in terms of flow volume. Once they were incorporated into the model, the parameters were easily calibrated to an acceptable standard.

5. DISCUSSION

Undertaking hydrological calibration in a consistent manner across such an extensive area highlighted a number of modelling challenges. Each of the models varied in terms of quantity and quality of gauging station data, as well as their natural variability both within each individual region and across the GBR as a whole.

There are a number of reasons why a region may or may not have calibrated well, including data availability, storages and extractions, and climatic and landscape variability. Data availability was a key factor. For example, the data-poor CYP region has 18 gauging stations, many of which are now closed and limiting calibration performance. On the other hand, in the Burd, Fitz and B-M NRM regions, had good hydrological data, with extensive spatial and temporal hydrological records providing a range of options for improving model calibration. Satisfactory calibration results were achieved for all regions. To achieve this result required a significant investment of time by the modelling team (approximately six months).

The inclusion of water extractions in all regions was an important step in the calibration process given they were shown to have a significant impact on downstream flows in some areas. This also highlights the value of PEST as a parameter estimation tool, which enabled flow extractions to be incorporated in the simultaneous calibration process. Additional factors shown to influence calibration were climatic, such as the impact of the monsoon in the tropical regions.

By utilising PEST to estimate parameters for multiple HRUs in the same catchment simultaneously over many model runs, the distribution of equally likely parameter sets could be analysed statistically across different HRUs in particular sensitive parameters, such as the SMSC. These differences were observed across a number of model regions. Further work is required to explore the interaction between additional model complexity and versus improved calibration performance.

6. CONCLUSIONS

An automated hydrology calibration of the SIMHYD rainfall runoff model in Source Catchments was undertaken using PEST. Calibration was undertaken across six GBR regions using a three part objective function incorporating log transformed daily flows, monthly flows and flow exceedence values. The application demonstrates the flexible nature of the PEST software and its ability to be applied to a diverse range of catchments. The results showed that the PEST calibration performed extremely well across all six regions.

Careful consideration should be given to the user defined objective function, at the start of the PEST process. Future work will assess a range of alternative objective functions, including calibration to high flows to determine whether nutrient and sediment load estimates are improved. In addition, it is recommended that in future calibrations, local knowledge is sought at the outset to help set PEST starting parameters, as this will save processing time. Another key learning from the project was the importance of high performance computer power. 64-bit platforms are essential for modelling calibration exercises of this scale. To improve run times an additional option was used and is recommended: running PEST in parallel (enabling model runs to be undertaken on multiple computers in parallel) by slaving out model runs to other computers across a network, or the internet.

The calibration process described in this paper is an effective and efficient way of calibrating a range of models, representing a range of environments.

REFERENCES

Carroll, C., Waters, D., Searle, R.D., Vardy, S., Silburn, D.M., Attard, S., Thorburn, P.J., Milla, R., Halpin, N., Davis, A.M., Schmidt, M., Wilson, B., Clark, A., (2011). A Paddock to Reef Monitoring and Modelling framework for the Great Barrier Reef: Paddock and Catchment. Marine Pollution Bulletin, Special Issue: Catchments and Coral Reef (in press).

Chiew, F.H.S., and Siriwardena, L. (2005). Estimation of SIMHYD Parameter Values for Application in Ungauged Catchments. In Zerger, A., and Argent, R. M. (eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2005. http://www.mssanz.org.au/modsim05/papers/chiew_2.pdf

Doherty, J. (2009). PEST: Model-Independent Parameter Estimation. Watermark Numerical Computing, Brisbane, Australia. http://www.sspa.com/pest.

Ellis, R.J., Doherty, J., Searle, R.D., and Moodie, K. (2009). Applying PEST (Parameter ESTimation) to improve parameter estimation and uncertainty analysis in WaterCAST models. In Braddock, R. and Anderssen, B. (eds) MODSIM 2009 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, July 2009. http://www.mssanz.org.au/modsim09/I2/ellis.pdf

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part 1: A discussion of principles. J. Hydrol. 10, 282–290.

Post, D.A., Vaze, J., Viney, N., and Chiew, F. H. S. (2007). In Oxley, L. and Kulasiri, D. (eds) MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2007. http://www.mssanz.org.au/MODSIM07/papers/45_s40/Regionalisings40_Post_.pdf

SILO (2009). The SILO Data Drill. Queensland Department of Environment and Resource Management. Available from http://www.longpaddock.qld.gov.au/silo.

Stewart, J. (2010). Great Barrier Reef Catchment Modelling: Parameter Sensitivity and Uncertainty Analysis and Hydrologic Model Parameterisation Methodology. Prepared for the Queensland Department of Environment and Resource Management. BMT WBM Pty Ltd, Leichardt, NSW, Australia.

Tonkin, M.J., and Doherty, J. (2005). A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resour. Res., 41*, W10412, doi:10,1029/2005WR003995.

ACKNOWLEDGEMENTS