Climate and landscape controls on water balance model complexity over changing timescales

S. E. Atkinson

Centre for Water Research, Department of Environmental Engineering, University of Western Australia, Crawley, Western Australia, Australia

R. A. Woods

National Institute of Water and Atmospheric Research Ltd., Christchurch, New Zealand

M. Sivapalan

Centre for Water Research, Department of Environmental Engineering, University of Western Australia, Crawley, Western Australia, Australia

[1] A systematic approach is described for determining the minimum level of model complexity required to predict runoff in New Zealand catchments, with minimal calibration, at decreasing timescales. Starting with a lumped conceptual model representing the most basic hydrological processes needed to capture water balance, model complexity is systematically increased in response to demonstrated deficiencies in model predictions until acceptable accuracy is achieved. Sensitivity and error analyses are performed to determine the dominant physical controls on streamflow variability. It is found that dry catchments are sensitive to a threshold storage parameter, producing inaccurate results with little confidence, while wet catchments are relatively insensitive, producing more accurate results with more confidence. Sensitivity to the threshold parameter is well correlated with climate and timescale, and in combination with the results of two previous studies, this allowed the postulation of a qualitative relationship between model complexity, timescale, and the climatic dryness index (DI). This relationship can provide an a priori understanding of the model complexity required to accurately predict streamflow with confidence in small catchments under given climate and timescales and a conceptual framework for model selection. The objective of the paper is therefore not to present a perfect model for any of the catchments studied but rather to present a systematic approach to modeling based on making inferences from data that can be applied with respect to different model designs, catchments and timescales. *INDEX TERMS:* 1821 Hydrology: Floods; 1833 Hydrology: Hydroclimatology; 1836 Hydrology: Hydrologic budget (1655); 1860 Hydrology: Runoff and streamflow; 1866 Hydrology: Soil moisture; *KEYWORDS:* water balance, model complexity, timescale, landscape controls, climate controls, uncertainty

Citation: Atkinson, S. E., R. A. Woods, and M. Sivapalan, Climate and landscape controls on water balance model complexity over changing timescales, *Water Resour. Res.*, *38*(12), 1314, doi:10.1029/2002WR001487, 2002.

1. Introduction

[2] The paper is concerned with the choice of appropriate model complexity for prediction of streamflow responses of ungauged catchments at different timescales. At present there is inadequate guidance available to a priori choose a model of appropriate complexity for predictions of streamflow of a specified accuracy, given only the climatic, vegetative, topographic and soil characteristics, and a given timescale. Provided such a method exists, and a model of appropriate complexity can be chosen, the dominant physical controls on streamflow variability can then be properly investigated. These results can subsequently be used to concentrate the modelers' efforts on the best approaches for estimating the required parameter values, with a view to improving the accuracy of predictions and to reducing predictive uncertainty.

[3] The motivation for this paper is therefore not the development of a "perfect model", but a systematic method of model selection that makes tradeoffs between model complexity, accuracy and predictive uncertainty. Although integral to modeling strategies in catchment hydrology [*Garen and Burges*, 1981], the tradeoffs between model complexity, accuracy and predictive uncertainty have not been explored in a systematic manner in the past. A plethora of models with a wide-ranging complexity have been developed [*Chiew et al.*, 1993; *Singh*, 1995], but there is no consistent method available to select the most appropriate model for a given set of catchment conditions. *Chiew et al.* [1993] compared the relative accuracy of six predefined calibration-based models of increasing complexity, when

tested on a number of catchments with different climatic conditions, and concluded that the most accurate model was the most complex model, regardless of timescales and catchment characteristics. Although, in some cases simpler models were reported to make accurate enough predictions, the links between climate, timescale and catchment characteristics, and model accuracy and required model complexity, were not explored in any detail.

[4] Conceptual models dependent on calibration for their parameter values, such as the ones studied by *Chiew et al.* [1993], are used extensively in catchment hydrology. However, *Klemes* [1983] suggests that a common weakness of many calibration-based models is structural arbitrariness and overparameterization. As they require estimation of many parameters with little or no physical meaning, there is often no means to estimate them all a priori. Consequently, there is a large associated parameter uncertainty which, when propagated through the model, produces output errors too large to place confidence in the predictions. Because these calibrated model parameters often could not be linked to physically meaningful or measurable climate and landscape properties, the links between catchment characteristics, climate and timescale, and predictive uncertainty, could not be established, and hence the tradeoff between model complexity, accuracy and predictive uncertainty could not be fully investigated. This has also inhibited the identification of the dominant physical controls on streamflow variability at various time and space scales.

[5] The approach to model development adopted in this paper is inspired by the downward approach outlined by *Klemes* [1983]. Starting with the simplest model possible, model complexity is systematically increased in direct response to demonstrated deficiencies in the model predictions, and the process is continued until the required accuracy is achieved. For water balance modeling, the modeler often finds that a relatively simple model is sufficient to make predictions at large timescales, such as the annual timescale, whereas the model complexity needs to increase as the timescales decrease. The starting model used in this paper is a variant of the Manabe bucket model [*Manabe*, 1969; *Milly*, 1994]. It incorporates simple, but physically meaningful representations of hydrological processes, and requires parameter values and climatic inputs most of which can be realistically derived from a priori field measurements. Minimal calibration against selected streamflow records from a number of storm events (<10) each of approximately 5–10 days duration is required to estimate two of these parameters before simulations are run and predictions made.

[6] The ultimate extent of the model evolution reached in the downward approach presented here is dependent upon the tradeoff between model complexity, accuracy, and predictive uncertainty, which are assessed using systematic sensitivity and Monte Carlo error analyses. These analyses identify the dominant physical controls on streamflow variability at each timescale, and produce estimates of predictive accuracy and uncertainty, and hence the degree of model complexity required/available to capture observed streamflow variability. Between-catchment differences in model predictions are correlated to physically measurable catchment properties and climatic conditions. Although the model applied in this paper is specific to New Zealand catchments, the approach used has also been adopted with considerable success to catchments in
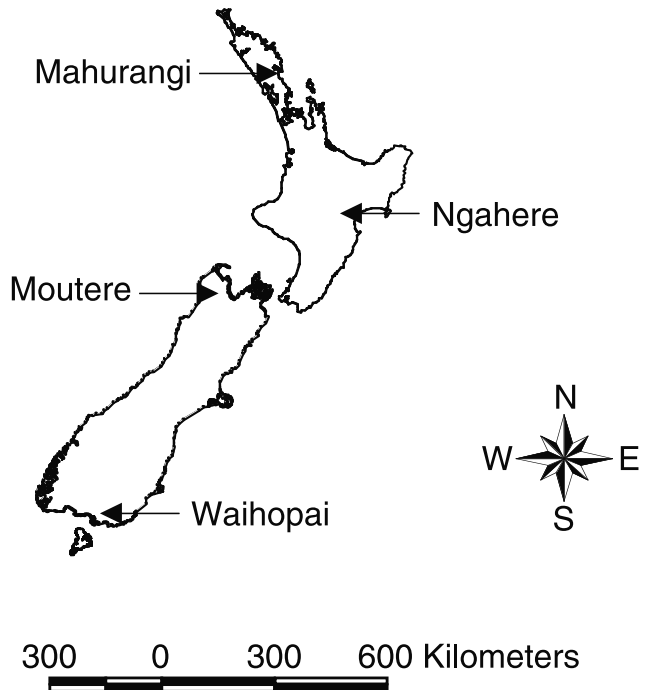
Figure 1.   Location of New Zealand catchments.

other climatic and physiographic regions [*Jothityangkoon et al.*, 2001; *Eder et al.*, 2002; *Farmer et al.*, 2002], and to other types of models as well. Based on the correlations obtained in all of these studies (including the present one), a qualitative relationship is formulated between climatic indices, timescale and required model complexity, which can provide useful guidance to model selection, development and intercomparison in other regions.

2. Study Sites

[7] Four New Zealand catchments were selected for this study because the climate is relatively moderate (rather than extreme), and the extensive data required to model streamflow was available through the National Institute of Water and Atmospheric Research (NIWA). The four catchments are Moutere, Waihopai, Mahurangi, and Ngahere, and as seen in Figure 1, these are well spaced to represent the different soil, climatic, vegetative and topographic conditions found in New Zealand. The relevant climatic, soils, vegetative and topographic information is presented in Table 1 including the length and period of datasets used.

3. Model Development

3.1. Model Components and Water Balance

[8] Our model development in fact commenced with the application of the classic *Manabe* [1969] "bucket" model, with the representation of the catchment as a single "bucket", as depicted in Figure 2a. Preliminary simulations with this rather simple model produced acceptable predictions of annual runoff in all four catchments (not presented here for brevity). The most accurate results were obtained in the very wet catchment with high annual rainfall (i.e., Ngahere). However, monthly and daily flow predictions using the model did not match the observations in any of the

Table 1. Mean Annual Rainfall, Potential Evaporation, and Streamflow as well as Relevant Soils, Vegetative, and Topographic Information

	Moutere	Waihopai	Mahurangi	Ngahere
Rainfall, P (mm/y)	958	1175	1628	2603
Potential Evaporation, Ep (mm/y)	1464	964	1315	959
Streamflow, Q (mm/y)	245	457	842	1728
Ep/P = dryness index (DI)	1.53	0.82	0.81	0.37
Length of complete dataset used in simulations (yrs)	13 (1972–1984)	30 (1962–1991)	15 (1983–1997)	23 (1974–1996)
Area (km^2)	0.07	150	47.5	0.52
Elevation (m)	100	50	150	1000
Porosity (ϕ)	0.540[a]	0.600[a]	0.570[b]	0.760[a]
Permanent Wilting Point (θ_{pwp})	0.230[a]	0.220[a]	0.350[b]	0.260[a]
Field Capacity (θ_{fc})	0.420[a]	0.500[a]	0.466[b]	0.650[a]
Vegetation	pasture	pasture	pasture/forest	forest
Mean Soil Depth (m)	0.370[a,c]	0.500[a]	0.870[b]	0.230[a]

[a] Source: *Department of Scientific and Industrial Research* [1968].
[b] Source: NIWA, Christchurch, New Zealand.
[c] Source: *Duncan* [1995].

catchments. It was therefore decided to include subsurface flow in addition to saturation excess runoff. In addition, the amount of vegetation cover has been observed to vary between catchments and so it was decided to include an explicit treatment of vegetation cover to enable between catchment inferences.

3.2. Bucket Representation

[9] Given these considerations, the Manabe bucket model was adapted to incorporate vegetation cover and a subsurface runoff generation mechanism. Figure 2b illustrates the operation of the new "leaky bucket" model, which represents the starting point of the analyses presented here. It introduces two storage thresholds governing water movement from the catchment. Soil moisture below the permanent wilting point is assumed to be immobile with respect to both evaporation and subsurface drainage, and "active water" is water that is stored above this threshold. Similarly, when the moisture content is below a field capacity threshold, then water does not drain through the subsurface flow pathway, but can be lost by evapotranspiration. The five parameters required to operate the "leaky bucket" model are f_c, S_{bc} (soil properties), M (vegetative property), a and b (conceptual soil and topographic parameters related to shallow subsurface flow). Their descriptions are provided in *Appendix*.

3.3. Bucket Capacity and Threshold Storage Parameter

[10] The bucket capacity (S_{bc}), is simply the storage capacity of the bucket, while the threshold storage parameter (f_c), determines the "threshold storage" (storage at field capacity, S_{fc}), controlling both runoff and evapotranspiration. The definitions for the bucket capacity, threshold storage parameter and threshold storage are provided in (A1–A3), along with equations describing these in terms of basic soil and vegetation properties, and showing a dependence on porosity (ϕ), moisture content at field capacity (θ_{fc}) and at the permanent wilting point (θ_{pwp}), and depth to an impervious clay layer (D).

3.4. Actual Evapotranspiration

[11] Actual evapotranspiration is a function of the "available water" (storage), and atmospheric conditions. Transpiration (E_{veg}) and bare soil evaporation (E_{bs}) are modeled explicitly and separately, as functions of soil properties and vegetative cover, as well as the dynamic value of soil moisture storage in the bucket. The rates of bare soil evaporation and transpiration are assumed to depend on the parameter M, representing the fraction of the catchment area covered by forests of any kind. Transpiration is modeled as occurring over a fraction M of the catchment, and bare soil evaporation occurs over the remaining fraction $1 - M$. Assuming all species have approximately the same transpiration rate, then $M = 1$ and $M = 0.1$ represent completely forested and completely pasture catchments, respectively. Pasture catchments are assumed to have 10% effective vegetation cover, due to the fact that the grasses in New Zealand are tall and relatively more deep rooted than usual, so not all evapotranspiration is the result of bare soil evaporation under pasture conditions. A catchment with short pasture would be modeled as $M = 0.05$. Equations (A4) and (A5) contain all of the relevant equations for estimating bare soil evaporation and transpiration. The combined effects of bare soil evaporation and transpiration are assumed to also implicitly include the effects of interception.

3.5. Discharge

[12] Daily discharge is produced via two mechanisms, surface runoff or quick flow via the saturation excess mechanism (Q_{se}), and subsurface runoff or delayed flow (Q_{ss}). Saturation excess runoff is produced when the storage
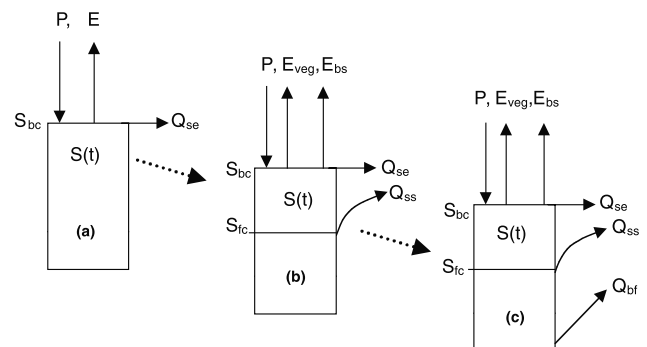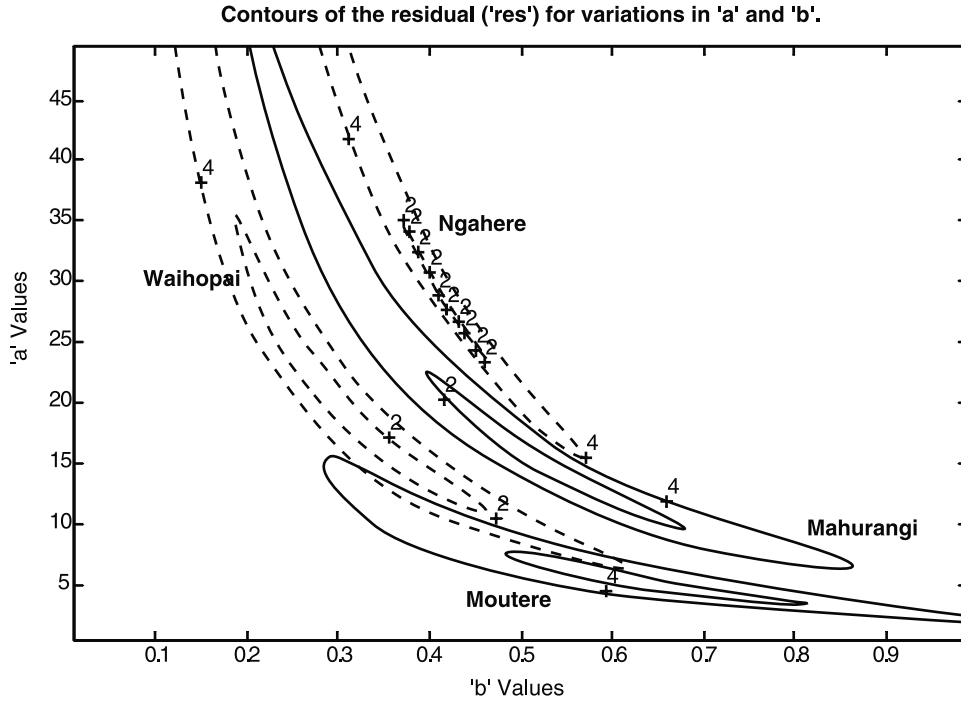

Figure 2. (a–c) Models used in downward approach.

Figure 3. Contours for residual "res" given variations in *a* and *b*.

exceeds the bucket capacity (S_{bc}). Similarly, subsurface flow is produced when the storage in the bucket exceeds the threshold storage S_{fc}, and is expressed through a non-linear storage-discharge relationship and governed by two parameters "*a*" and "*b*", as described in *Wittenberg and Sivapalan* [1999]. Equations (A6) and (A7) are the relevant equations.

3.6. Water Balance

[13] The resulting water balance equation for the catchment is given by:

$$S(t+1) = S(t) + P - E_{veg} - E_{bs} - Q_{se} - Q_{ss} \qquad (1)$$

Equation (1) is used to simulate the temporal variation of water storage and the various outgoing fluxes, at a daily time step, given daily inputs of rainfall (*P*), and potential evaporation (E_p), as well as estimates of the parameters; *M*, f_c, S_{bc}, *a*, and *b*. Appendix A contains all of the relevant definitions. To determine the initial storage (S_0) for each catchment at the beginning of simulations, the model was initialized by running the model for one year and then recording the soil storage at the end of the simulation.

3.7. Parameter Estimation

[14] The estimation of parameters *M*, f_c and S_{bc} is relatively straightforward, given the relevant equations in Appendix A, and the information presented in Table 1. The estimation of the nonlinear recession parameters *a* and *b* is not as straightforward. The estimation of these parameters involves a process of assembling together individual recession curves at various times of the year extracted from observed streamflow data to produce a master recession curve (or mean curve). A recession in this instance is defined as the period shortly after a runoff event (>1 day

after a rainfall event). The theoretical functional form of the recession curve corresponding to the nonlinear storage-discharge relationship (A7) is given by [*Wittenberg and Sivapalan*, 1999]:

$$Q_t = Q_o \left(1 + \frac{(1-b)Q_o^{1-b}}{ab} t \right)^{\frac{1}{b-1}} \qquad (2)$$

where Q_t is recession flow (mm/d) at time *t*, Q_o is the initial runoff in the mean time series, in mm/d. This theoretical relationship is fitted to the empirical master recession curve described above, through minimization of an appropriate objective function (calibration). The objective function we use, *res*, measures the quality of fit (a value of zero indicates perfect fit), and is given by (3) below:

$$res = \sum_{t=1}^{n} \left| Q_{pred}(t) - Q_{master}(t) \right| \qquad (3)$$

where *t* is the time in days, and Q_{pred} and Q_{mean} are the predicted and measured (master or mean) recession curves.

[15] When using (3) to minimize *res* and to solve for *a* and *b*, we found that there were numerous optimal solutions that produced equally accurate fits, in all four catchments. A sensitivity analysis was performed and the results (Figure 3) suggested that for *b* values between 0.4 and 0.6, all four catchments produced equally accurate recessions (low res). *Wittenberg and Sivapalan* [1999] produced similar results for arid and semi-arid catchments of southwest Western Australia, and concluded that it was reasonable to fix the exponent *b* at a mean or dominant value. A value of 0.5 was suggested, agreeing with the empirically estimated values for unconfined aquifers presented in the previous work, e.g., *Werner and Sundquist* [1951], and *Fukushima* [1988]. We have therefore, for convenience, adopted a value of 0.5 for the parameter *b*, and have assumed it to be the same for all

Table 2. Parameter Values for Moutere, Waihopai, Mahurangi, and Ngahere

Standard Model Parameters	Units	Moutere	Waihopai	Mahurangi[a]	Ngahere
M	dimensionless	0.1	0.1	0.44	1.0
f_c	dimensionless	0.61	0.74	0.53	0.78
S_{bc}	mm	115	190	190	115
a ($b = 0.5$)	$mm^{0.5}$ $d^{0.5}$	7.4	9.6	15.8	20.3
Tc_{bf}	d	10000[b]	336	88	137

[a] Areal average of S_{bc} for different soil types in catchment.
[b] No base flow at Moutere ($Tc_{bf} -> \infty$).

four catchments. Although fixing b at 0.5 is a pragmatic approach to reduce parameter interdependence and improve identifiability, it does not represent the optimum performance [*Campbell and Bates*, 2001]. Given that $b = 0.5$, (3) was subsequently used to estimate a for all catchments by minimization of the function *res*. Note that the effect of evaporation can be considered negligible during the recession periods, in zones of moderate climate such as those found in New Zealand [*Wittenberg and Sivapalan*, 1999]. Our sensitivity analyses have also shown that evaporation has negligible impact on the estimated recession parameters in these catchments.

[16] Table 2 presents estimates of the four parameters of the model for the four study catchments: the first three (M, f_c, S_{bc}) were estimated from field data (as described before), and the last (namely a) was estimated from the recession curve analyses (as described above). All of the parameters, including the parameters a and b, were estimated prior to the application of the leaky bucket model to the four catchments.

4. Results

[17] The results of model simulations are presented in the form of signature plots [*Jothityangkoon et al.*, 2001; *Farmer et al.*, 2002], and also in terms of the hydrographs of observed versus predicted streamflows, and through statistical summaries. Signature plots are statistical representations of the inter-annual, intra-annual (monthly) and daily streamflow variability of the catchment runoff response. They enable us to assess how well the model is able to predict the streamflow response of the catchment without the reliance on hydrograph fitting, and moreover, can give us considerable insight into catchment response.

[18] However, the hydrographs of observed and predicted streamflow have also been presented together to confirm if the matching of the signature plots is indeed reflected in an accurate reproduction of the observed hydrographs. This is because visual examination of the signature plots chosen provides little quantitative information about model performance in capturing individual storms, especially, streamflow recessions. Two numerical measures of fit, the correlation coefficients ($\rho_{x,y}$), and the Runoff Ratio (RR), are introduced to enable the comparison of the timing and magnitude (respectively) of the observed and predicted hydrographs. These are estimated using (4a) and (4b) given below:

$$\rho_{x,y} = \frac{\text{cov}_{x,y}}{\sqrt{\text{var}_x \cdot \text{var}_y}} \qquad (4a)$$

$$RR = \left(\frac{E[Q_{pred}]}{E[Q_{obs}]}\right) \qquad (4b)$$

where $\rho_{x,y}$ and $\text{cov}_{x,y}$ are the correlation coefficient and covariance, respectively, between the observed and predicted time series (x,y), and var_x and var_y are the variances of the individual time series. $\rho_{x,y}$ is used to measure how well the model mimics the time variability of the catchment. RR is the dimensionless runoff ratio and $E[Q_{obs}]$, $E[Q_{pred}]$ are the expected mean annual observed and predicted runoff respectively. RR is used to measure the bias in model estimates. Although there are many measures of model performance available, those selected here are deemed adequate to assess the general performance over the entire length of record.

4.1. Results From Preliminary Model Simulations

[19] Figures 4a, 4b, 4c, and 4d and Table 3 contain the signature plots, the observed versus predicted streamflow hydrographs, and the streamflow statistics, for all four catchments ($E[P]$ is the mean annual rainfall and $E[Q]$ is the mean daily runoff for the observed and predicted time series). The results suggest that the model predicts streamflow with reasonable accuracy at the annual and monthly timescales but fails to reproduce daily streamflows, especially during low flows. Failure to capture low flows is exemplified by the apparent "plunging" of the tail of each catchment's modeled flow duration curve and by the recession limbs of their streamflow hydrographs (Figure 4a, plots iii and iv; Figure 4b, plots iii and iv; Figure 4c, plots iii and iv; and Figure 4d, plots iii and iv), as well as by the low correlation coefficients corresponding to daily flows (Table 3).

[20] Low flows have little effect on the timing of annual and monthly flow volumes, but can have a significant effect on the timing of daily predictions (especially in summer). Correlation between observed and predicted daily flow is noticeably higher at Moutere than the remaining three catchments. This can be attributed to the very small catchment area and consequent lack of base flow. Waihopai, Mahurangi and Ngahere are all larger catchments with significant volumes of base flow, which should be represented in the model to improve predictions of daily flow. The Runoff Ratio (RR), was close to unity for all catchments ($RR \approx 1$), except in Moutere, which showed significant over-predictions at all timescales. Given the area of Moutere, the rainfall and pan evaporation measured on site are good estimates of the catchment average rainfall and evaporation so are unlikely to be the sources of over-predictions. These are more likely caused by one or more of the catchment parameters used in simulations. Poor timing and the poor low flow predictions do not seem to affect RR because it is more truly an annual statistic.

[21] Therefore it is suggested that the model as it stands could be used to predict annual and monthly flows, but would be inappropriate at the daily time step. Failure to capture low flows at the daily time step can be attributed to the operation of the bucket as depicted in Figure 2b and equation (A7), by the fact that streamflow reduces to zero when storage falls below S_{fc}. Since observed streamflow records suggest there is a persistent base flow component throughout the year at Waihopai, Mahurangi and Ngahere, introduction of base flow process, its parameterization, and the means to estimate the corresponding parameter value(s), are needed to enhance the model's capability to predict daily runoff accurately. With negligible base flow at Moutere, the
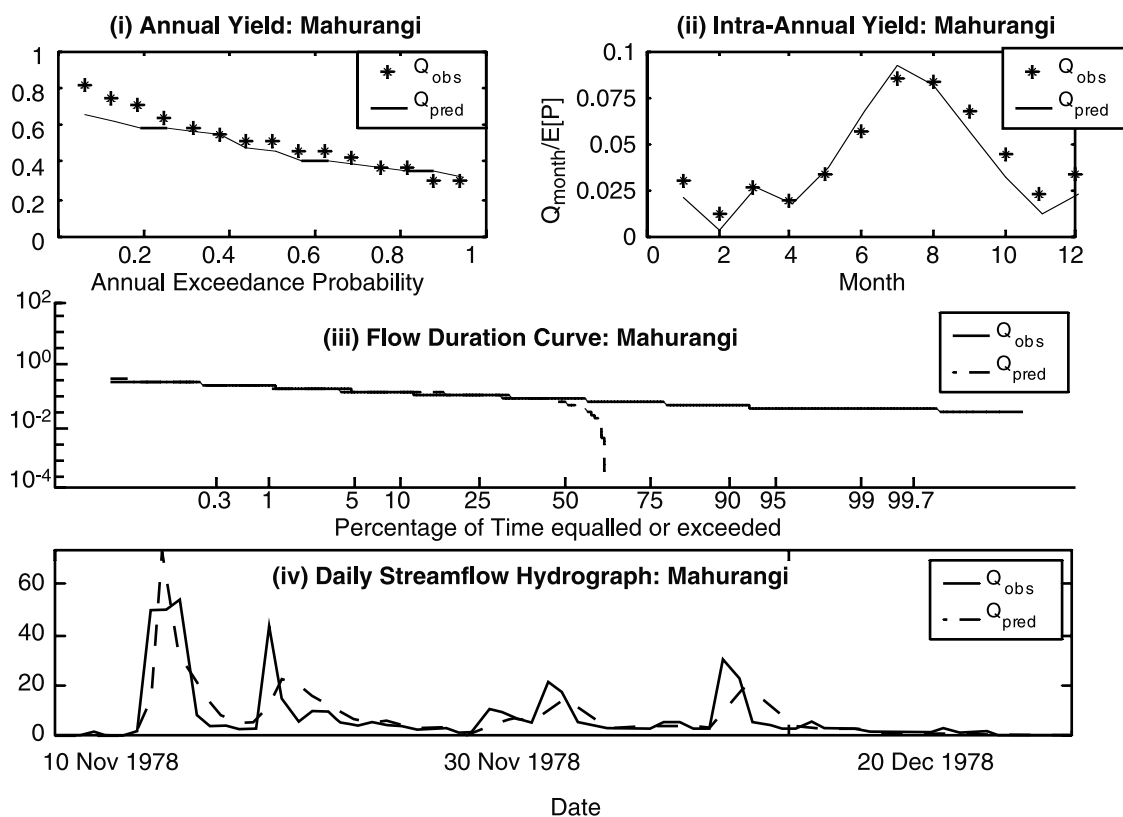
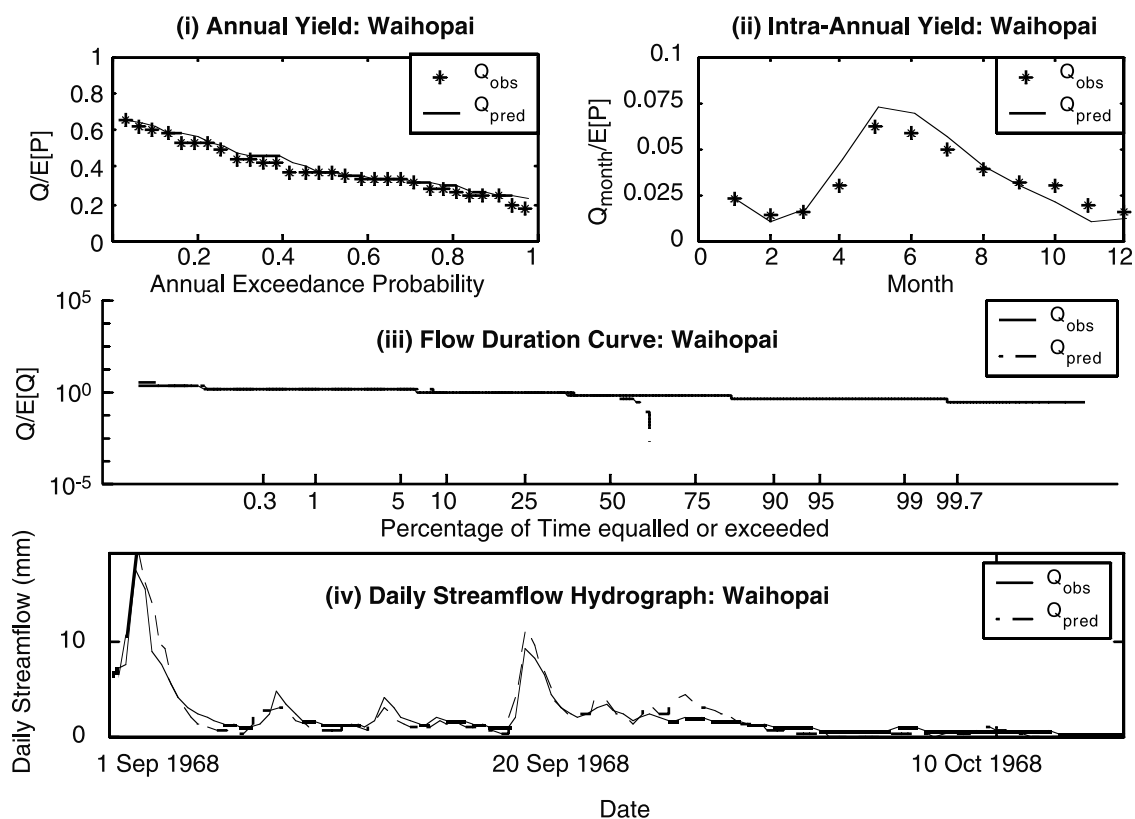**Figure 4a.** Signature plots for preliminary simulations: Mahurangi.

Figure 4b. Signature plots for preliminary simulations: Waihopai.
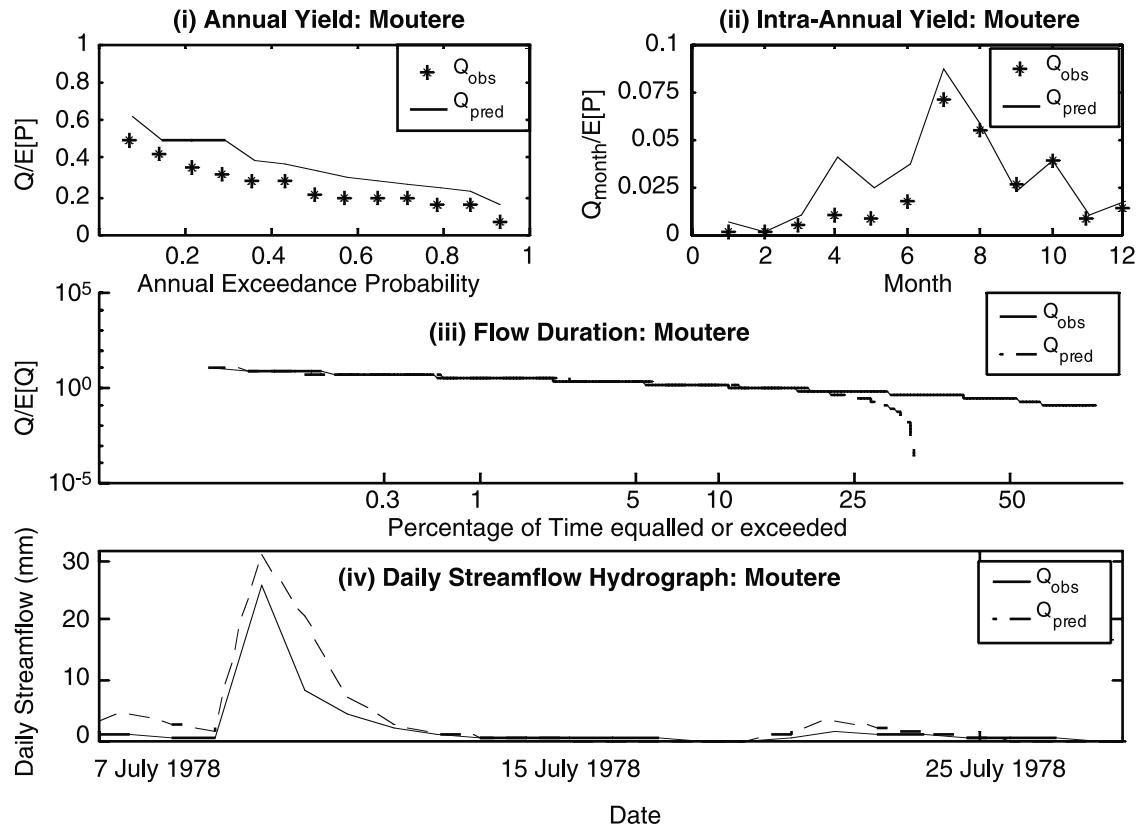
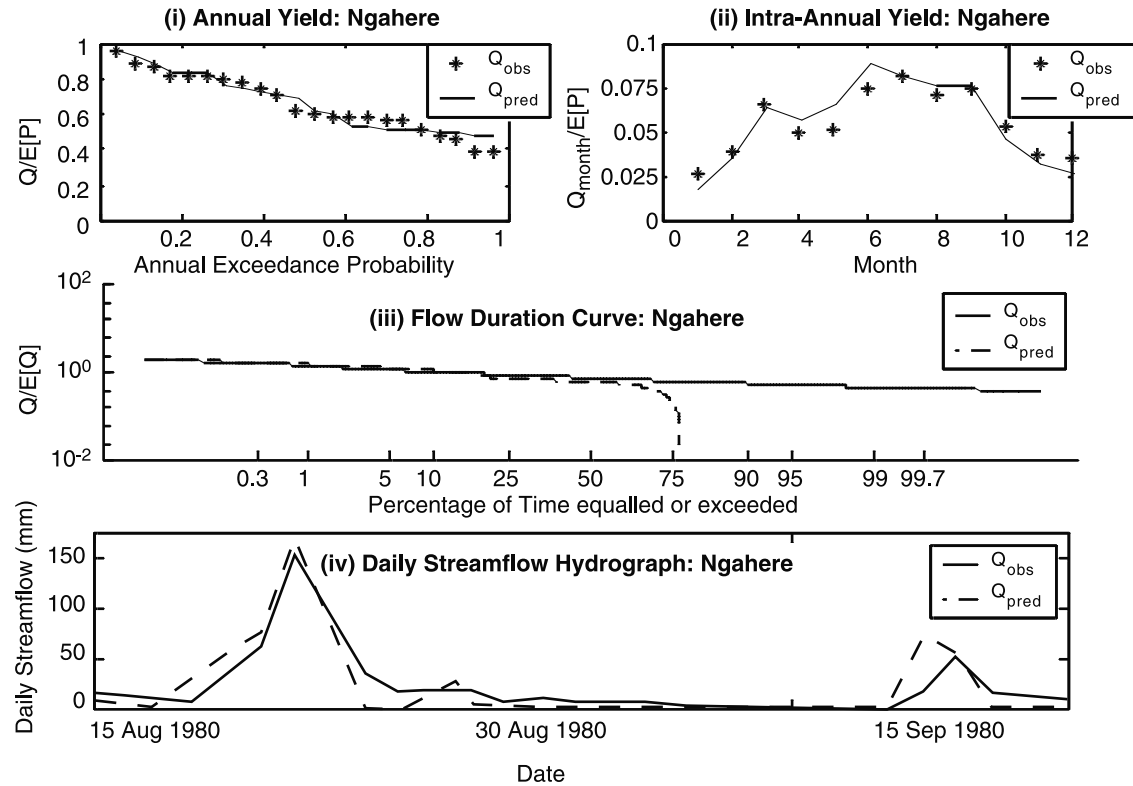**Figure 4c.** Signature plots for preliminary simulations: Moutere.

Figure 4d. Signature plots for preliminary simulations: Ngahere.

Table 3. Results From Model Simulation

	Moutere	Waihopai	Mahurangi	Ngahere
Annual ρ (x,y)	0.96	0.92	0.92	0.92
Monthly ρ (x,y)	0.89	0.92	0.94	0.92
Daily ρ (x,y)	0.86	0.81	0.80	0.78
RR^a	1.33	1.07	0.93	1.01

[a] A value of 1 signifies a perfect fit.

model must be extended so base flow can be turned off and on when required.

4.2. Introducing Base Flow

[22] A simple base flow mechanism is introduced by which water is slowly released from the bucket, but in such a way as to persist throughout the year. To limit model complexity we chose not to introduce an additional deep groundwater store, since this would require knowledge of the rates of leakage to the deeper store and the rate of discharge from this store, both of which will require considerable soil testing to estimate. Rather, we lump the effects of such leakage and discharge and, for simplicity, introduce the base flow discharge as a linear function of bucket storage. Being a linear response it introduces only one additional parameter, namely the base flow response time (Tc_{bf}). Figure 2c illustrates the operation of the revised model.

[23] Tc_{bf} is used the same way as the parameters controlling subsurface discharge (a and b) and takes the same form as (A7) when $b = 1$ (linear response). Base flow, calculated via (A8), is assumed not to be limited by field capacity, and operates on the total bucket storage. A similar calibration procedure described for parameters a and b is used to determine Tc_{bf}. The only difference is that estimation of

Tc_{bf} is based on the recession limbs of hydrographs below a predefined summer low-flow threshold or during "dry spells", and it is allowed to include effects of evapotranspiration [*Wittenberg and Sivapalan*, 1999]. Evapotranspiration is included because the focus is now on summer low flows, during which potential evaporation is as high as measured streamflow. The estimated base flow parameter values, Tc_{bf}, for the four catchments are presented in Table 2. The base flow parameter for Moutere was set to 10,000 to simulate the lack of base flow in very small catchments. Although the volume of base flow produced is minimal, it is still significant since it is able to generate low flows that were not captured by the previous model.

4.3. Results From Improved Model Simulations

[24] With the extended parameter sets for each catchment, model simulations were carried out and the results presented in Figures 5a, 5b, 5c, and 5d and Table 4. The signature plots, the observed versus predicted streamflow hydrographs, and the summary streamflow statistics all show marked improvements in the prediction of low flows, for all catchments except Moutere (which does not have base flow). On the other hand, the statistics contained in Table 4 suggest that the introduction of base flow has made very little change to the timing and magnitude of the predictions, at all three timescales. This is to be expected since the magnitude of low flows is negligible compared to flood peaks and the annual volumes of base flow would be negligible compared to annual volumes of total runoff.
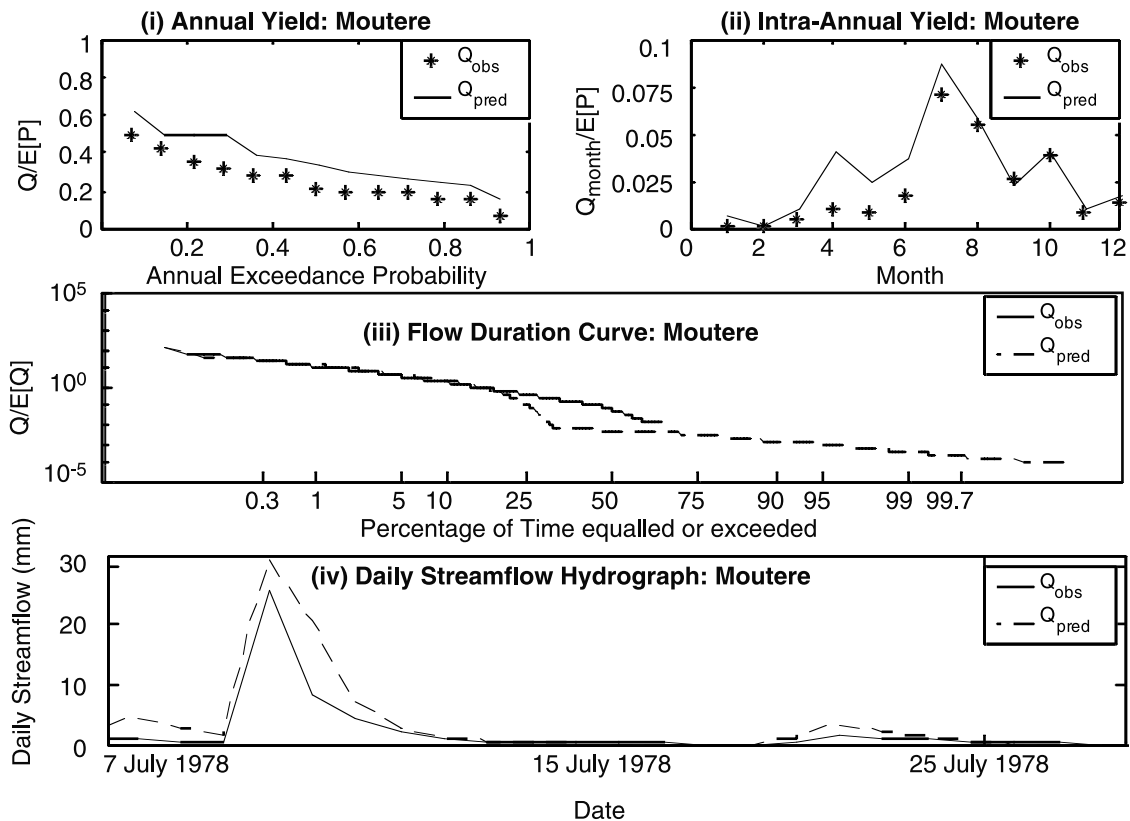
[25] Although the addition of base flow has helped to improve the accuracy of model predictions, this additional
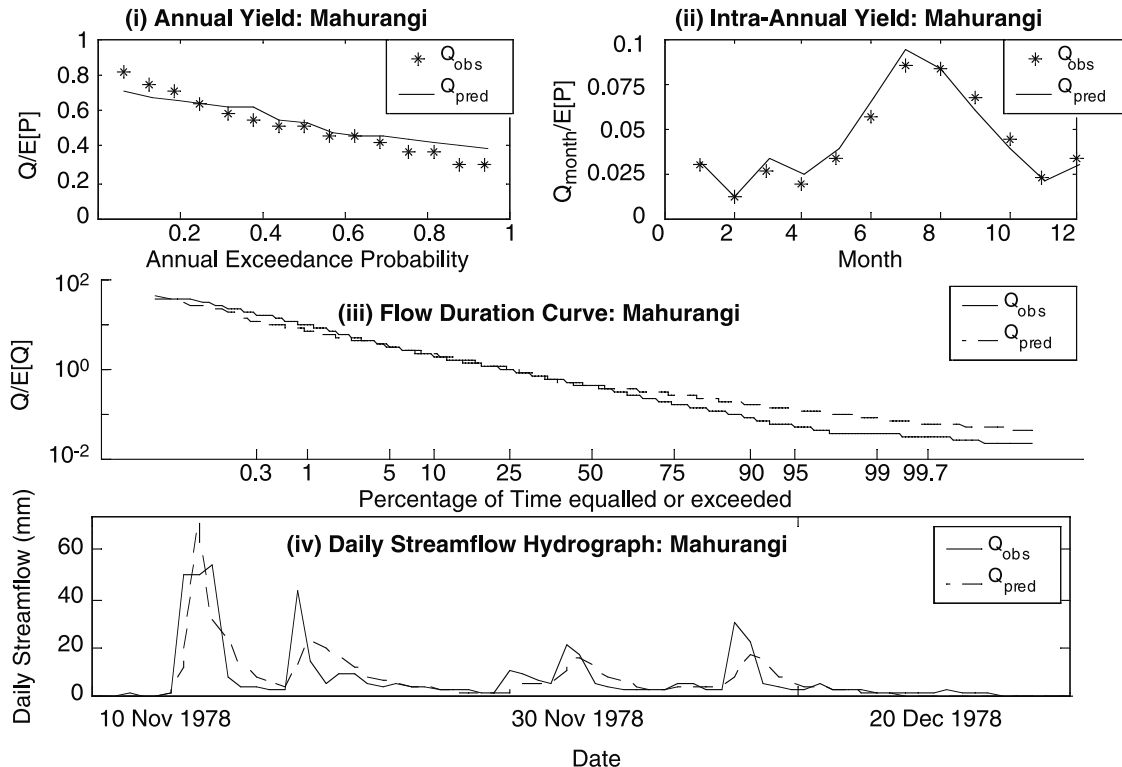
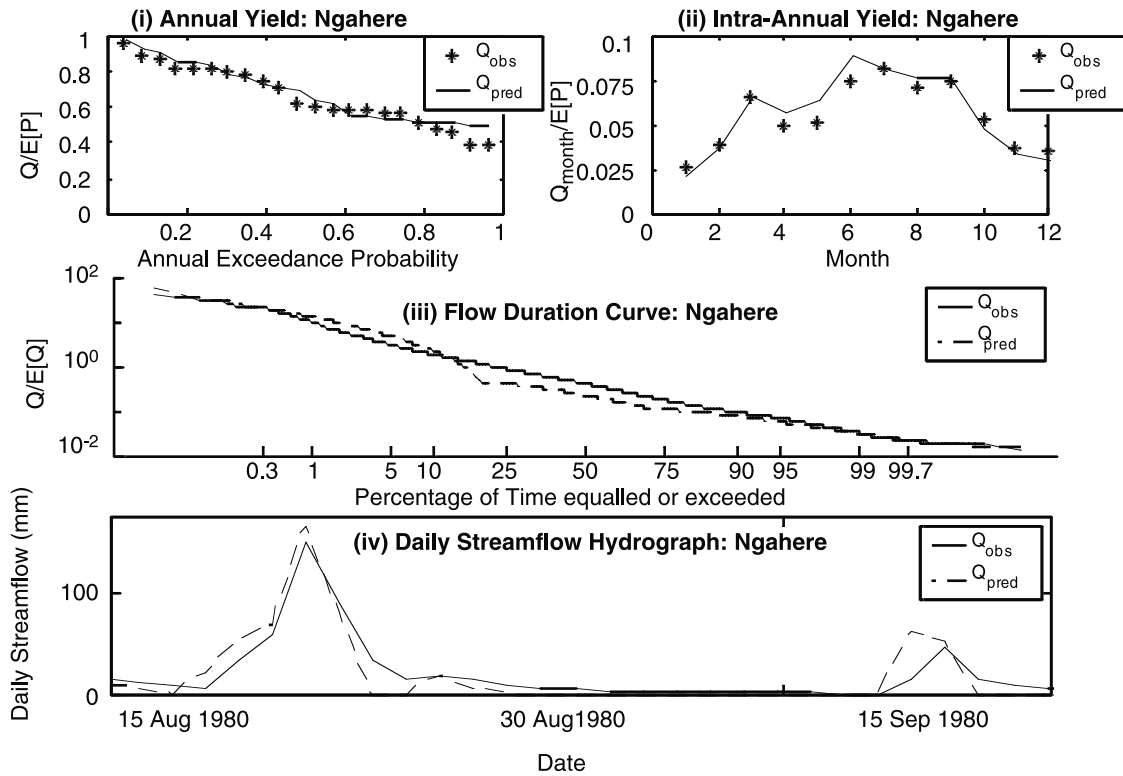**Figure 5a.** Signature plots for simulations with base flow: Mahurangi.
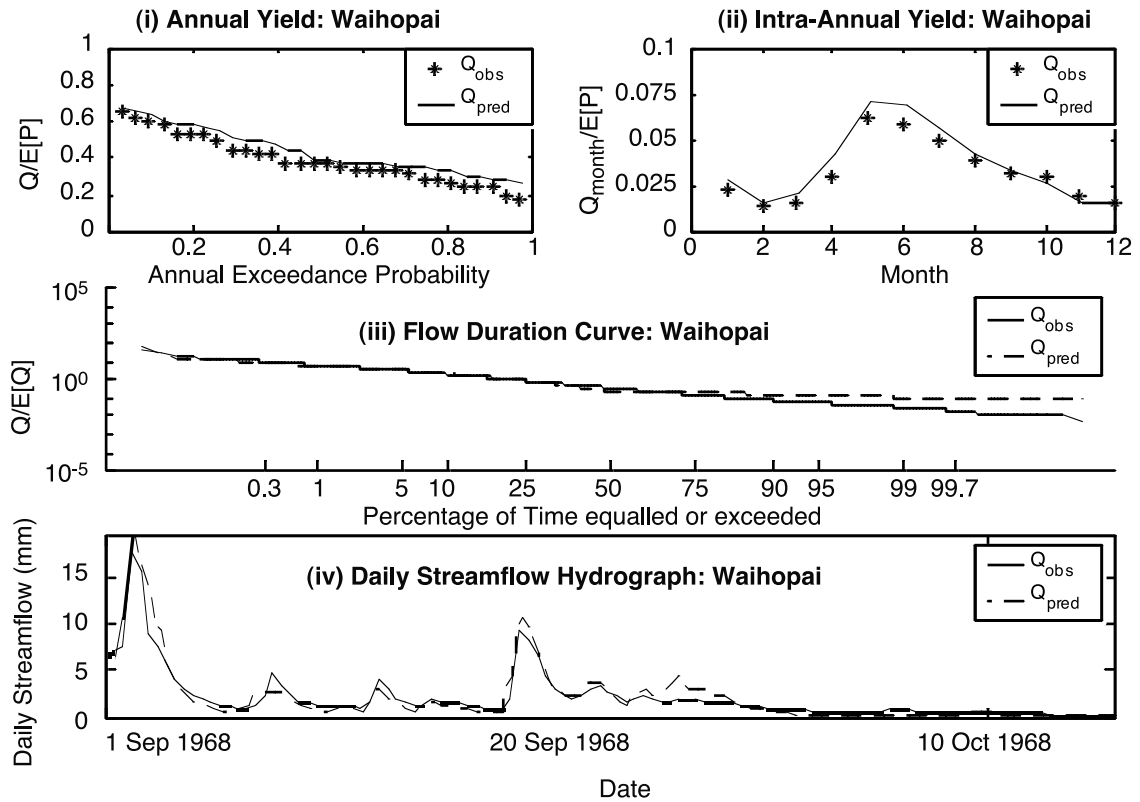
Figure 5b. Signature plots for simulations with base flow: Waihopai.

Figure 5c. Signature plots for simulations with base flow: Moutere.

Figure 5d. Signature plots for simulations with base flow: Ngahere.

complexity will increase the uncertainty of the predictions. Hence it is not clear whether the improved accuracy gained through additional complexity justifies the increased predictive uncertainty. The aim of this paper is to capture streamflow variability with confidence and sufficient accuracy, so that the model can be used to identify the dominant sources of variability; hence, it is important to keep the model as simple as possible so as to limit predictive uncertainty. At this stage, the current model predicts streamflow sufficiently accurately at all timescales, so the addition of further complexity is not deemed warranted. Therefore sufficient accuracy has been defined in this paper as predictions where $\rho_{x,y}$ is approximately ≥ 0.8, RR is close to 1, and there is little difference between the observed and predicted signature plots (as described by *Farmer et al.* [2002]). Sensitivity and error analyses were performed with the current version of the model to investigate the sensitivity of predictions to parameter variations, timescale and climatic conditions and to investigate the tradeoffs between model accuracy and predictive uncertainty.

5. Sensitivity and Error Analyses

5.1. Sensitivity Analysis

[26] Sensitivity analysis was identified as an effective means of determining the dominant physical controls on streamflow variability, as it provides insight into the operation of the model and explores its response to parameter variations. The sensitivity of annual, monthly and daily flow predictions was assessed at decreasing timescales, as well as between-catchment differences in model sensitivity, with a view to establishing a link between parameter sensitivity,

timescale and climate. This of course presupposes that the mechanisms controlling the water balance in all catchments at all timescales have been reasonably well represented in the model.

[27] It is important to first give a definition of "dominant physical control", which in this paper refers to a model parameter to which model predictions are extremely sensitive to, when varied over a realistic range of values (estimated from field measurements, observation and experience). The sensitivity analyses performed in this paper used a range of ±50% for all parameters shown in Table 2. In the circumstance where ±50% of the measured parameter value is infeasible (i.e., f_c and M), the range of values is set so they do not exceed the maximum and minimum allowable limits (1 and 0 respectively). The model was run repeatedly, with the starting nominal value for each parameter in Table 2 multiplied, in turn, by 0.5, 0.75, 1, 1.25 and 1.5, while keeping all other parameters constant at their nominal starting values. The resulting streamflow time series are assessed for the timing of annual, monthly and daily streamflows using the correlation coefficient $\rho_{x,y}$, and the magnitudes using RR. When varying each parameter in turn over the selected range, there is a spread of values for both $\rho_{x,y}$ and

Table 4. Results From Revised Model Simulations

	Moutere	Waihopai	Mahurangi	Ngahere
Annual ρ (x,y)	0.96	0.92	0.91	0.91
Monthly ρ (x,y)	0.89	0.93	0.95	0.93
Daily ρ (x,y)	0.86	0.81	0.77	0.80
RR[a]	1.33	1.14	1.05	1.03

[a] A value of 1 signifies a perfect fit.
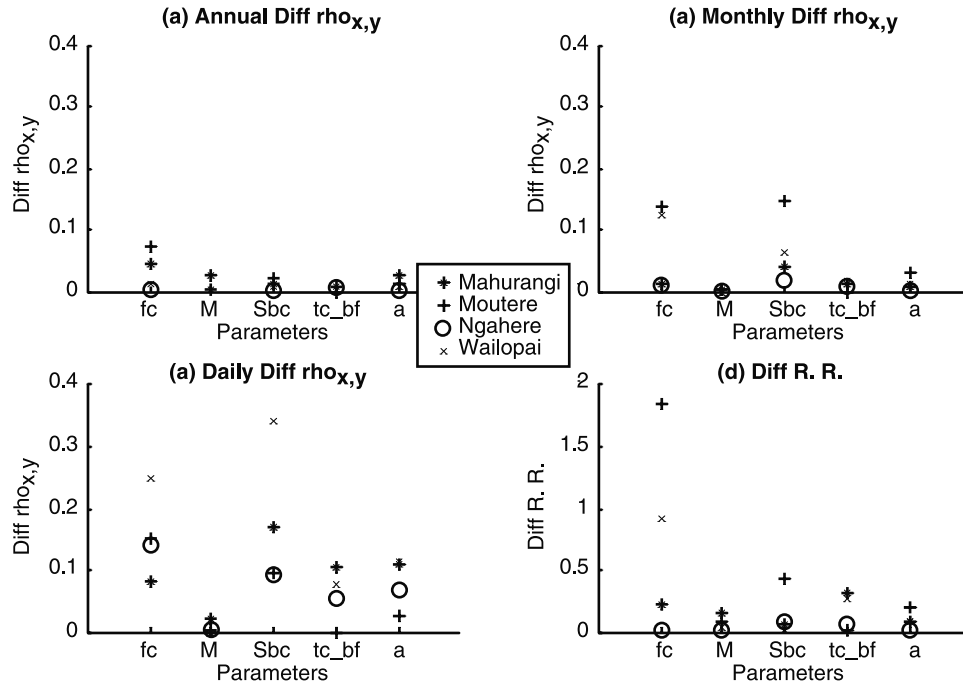
Figure 6. Sensitivity of (a) annual $\rho_{x,y}$, (b) monthly $\rho_{x,y}$, (c) daily $\rho_{x,y}$, and (d) runoff ratio to variations in parameter values. *Diff* $\rho_{x,y}$ and *Diff RR* are measures of sensitivity.
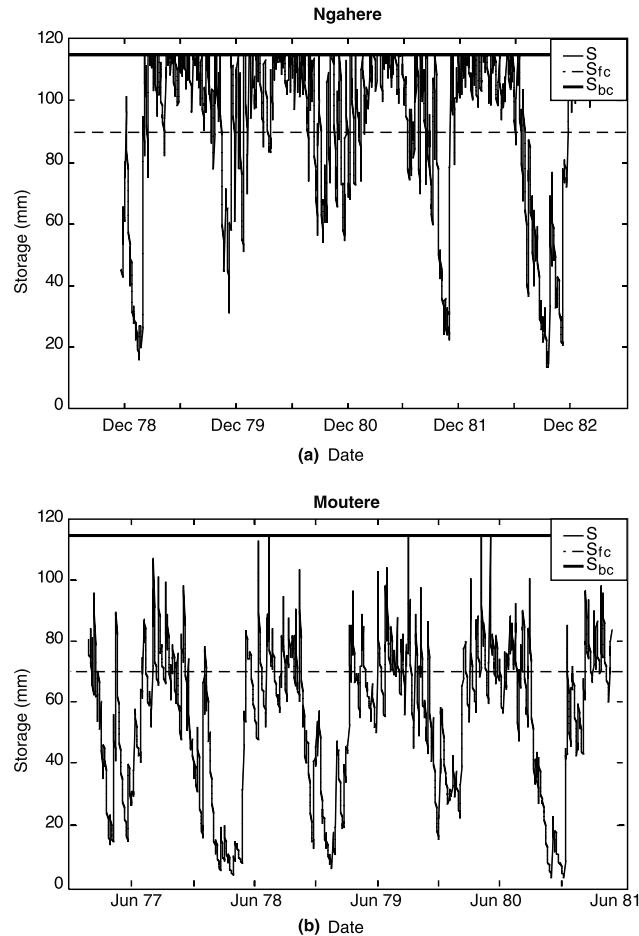
Figure 7. (a) Simulated daily storage fluctuations at Ngahere. (b) Simulated daily storage fluctuations at Moutere.

Table 5. Error Bounds

Parameter	Moutere	Waihopai	Mahurangi	Ngahere
M	±10	±10	±10	±10
Porosity (ϕ)	±15	±15	±5	±15
Field Capacity (θf_c)	±15	±15	±5	±15
Permanent Wilting Point (θ_{pwp})	±15	±15	±5	±15
Soil Depth (D)	±50	±50	±10	±50
Tc_{bf}	±30	±30	±30	±30
a (b = 0.5)	±30	±30	±30	±30

RR. Since we require some measure of the spread (which is a measure of the model sensitivity to the parameter values), the differences between the maximum and minimum estimates for $\rho_{x,y}$ and *RR* (*Diff* $\rho_{x,y}$ and *Diff RR*) are also estimated. These are presented in Figure 6 in separate plots.

[28] Although sensitivity to variations in rainfall and evaporation were also investigated (presented in a later section), we are more interested in the model response to variations in the physical parameter values.

5.2. Results of Sensitivity Analysis

[29] The results have been separated into the effects that variations in parameter values have on (1) the timing of annual, monthly and daily streamflow ($\rho_{x,y}$), and (2) the average magnitude of streamflow (*RR*). If the timing or magnitude of flow is highly sensitive to variations in parameter values, we would expect this to be reflected in large values of *Diff* $\rho_{x,y}$ or *Diff RR*, respectively.

5.2.1. Timing of Discharge

[30] For all parameters and all catchments Figures 6a–6c shows increasing sensitivity of the timing of flow with decreasing timescales. This increase corresponds to an increase in sensitivity, and needs to be explained physically. At annual timescales, the timing of streamflow is not important as almost all effective rainfall falling on a catchment in one year is discharged within the year, regardless of parameter values. At monthly timescales, timing appears to become more important in water-limited catchments such as Moutere, Waihopai and Mahurangi where there are distinctive months of high and low flows. As the timing of these high and low flows is dependent on antecedent soil moisture conditions, the correlation between observed and predicted monthly flows is sensitive to the parameters controlling storage (f_c and S_{bc}). This is reflected by the large values of *Diff* $\rho_{x,y}$ for these parameters in Figure 6b. In Figure 6c, the large scatter at daily timescales confirms that timing is very important in all catchments. The timing and magnitude of daily discharge is dependent upon the interaction of all physical processes governed by bucket storage, therefore runoff production and the correlation between observed and predicted hydrographs is sensitive to all of the parameters. However, of all the five parameters, Figure 6c suggests that the timing of daily streamflow is again most sensitive to f_c and S_{bc}.

5.2.2. Magnitude of Discharge

[31] In Figure 6d the magnitude of flow (*Diff RR*) in each catchment is found to be either most sensitive to f_c or relatively insensitive to all five parameters. For example, in Ngahere and Mahurangi *Diff RR* is insensitive to all five parameters while in Moutere and Waihopai it is extremely sensitive to f_c. In the catchments where *Diff RR* is extremely sensitive to f_c, there is a large risk of under-/over-prediction of streamflow because, according to (A2), (A3), (A6) and (A7), f_c controls the bulk of the discharge. Therefore sensitivity of *Diff RR* to variations in f_c can be used to help explain between-catchment differences in runoff response.
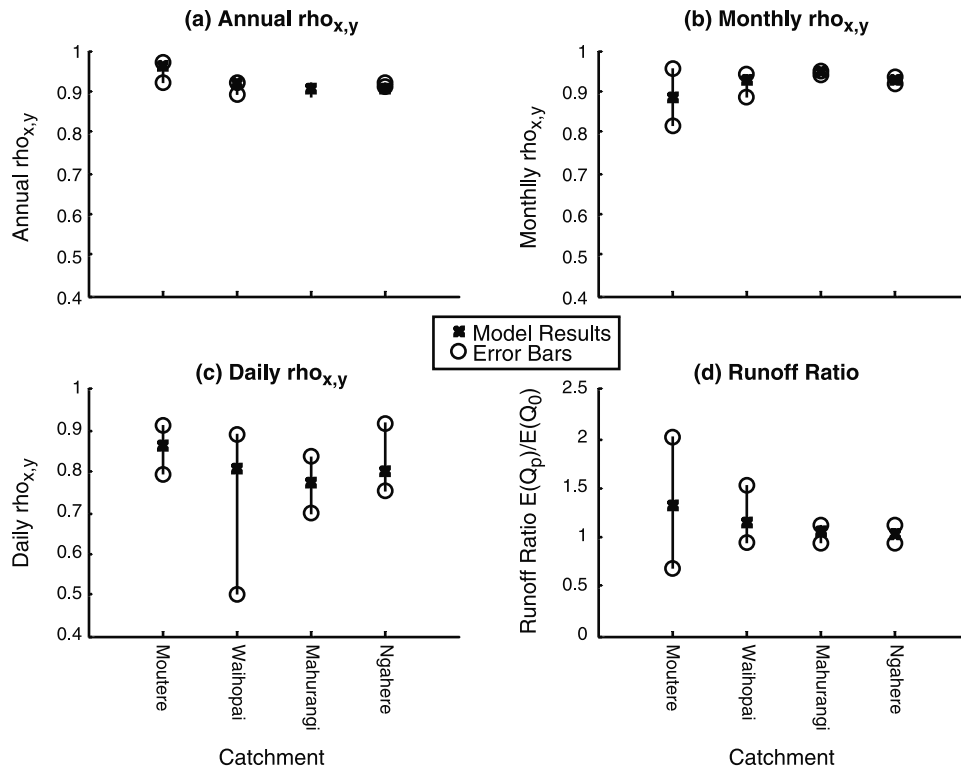
Figure 8. Accuracy (crosses) and error bars (circles) for predictions of (a) annual $\rho_{x,y}$, (b) monthly $\rho_{x,y}$, (c) daily $\rho_{x,y}$, and (d) the runoff ratio given uncertainty in parameter values.

Table 6a. Results of Error Analyses With Respect to the Runoff Ratio (*RR*)

	Moutere	Waihopai	Mahurangi	Ngahere
DI	1.53	0.82	0.81	0.37
Runoff ratio	1.33	1.14	1.05	1.03
Deviation (%): Upper Error Bar[a]	51.88	32.46	7.62	5.82
Deviation (%): Lower Error Bar[a]	−48.88	−17.54	−9.52	−1.94

[a] Deviation from original simulation results (RR).

To explain these between-catchment differences we use a measure of catchment dryness called the climatic dryness index (DI), presented in (5) below as the dimensionless ratio of mean annual potential evaporation to the mean annual rainfall.

$$D.I. = \frac{E[E_P]}{E[P]}. \tag{5}$$

[32] The between-catchment differences can then be explained through a link between DI and the sensitivity of *Diff RR* to f_c. Presented in Table 1 are the DI values for each catchment, which classify Moutere as a dry catchment (high DI), and Ngahere as wet catchment (low DI). There is strong correlation between the DI and the sensitivity of the *Runoff Ratio* (*RR*) to f_c. The correlation is also supported by evidence that the timing of discharge is dictated by f_c, especially in dry catchments. This trend between catchment dryness and sensitivity to f_c can be explained with respect to moisture storage.

[33] In wet, energy limited catchments, effective rainfall dominates the water balance so storage usually lies between S_{fc} and S_{bc}, evaporation remains close or equal to the potential rate for most of the year, and runoff is generated continuously via surface and subsurface processes. With such a high storage, f_c has little control on runoff production, most of which is generated via the saturation excess mechanism. This is the case for Ngahere where storage fluctuates between S_{fc} and S_{bc} for the majority of the year as depicted in Figure 7a. Drier catchments are water limited, receiving less rainfall than potential evaporation. Storage controls runoff production in these catchments, since it tends to fluctuate above and below S_{fc} most of the year (Figure 7b). As described in (A7), this limits discharge, resulting in streamflow predictions that are highly sensitive to f_c. Therefore this parameter can, in dry catchments, control both the timing and magnitude of monthly high/low flows, and the total annual runoff, and can account for the differences between observed and predicted signature plots shown in Figures 5a, 5b, 5c, and 5d.

[34] Although not an integral part of this study, sensitivity of the timing and magnitude of flow to variations in rainfall

and evaporation was also investigated. The results in drier catchments showed that although there was noticeable sensitivity to ±10% variations in rainfall, the timing and magnitude of flow were still most sensitive to soil parameters. In addition, the use of a ±50% range of values in the sensitivity analysis results in the routing parameters (a, Tc_{bf}) changing over a much wider range than the remaining parameters. Additional sensitivity analysis using smaller error bounds for the routing parameters was performed and found to have little effect in terms of importance of parameters. However, the choice of error bounds is something that requires further investigation.

[35] The sensitivity analyses have thus shown that accuracy generally decreases with decreasing timescale, and that the accuracy of estimated soil parameters (namely, f_c) would dictate the accuracy of model simulations. In particular, the accuracy of predictions of streamflow at all timescales is less dependent on the accuracy of soil parameters (namely, f_c) in wet catchments (Ngahere) than in drier catchments (Moutere). To compare these trends in model accuracy against trends in predictive uncertainty (which is a measure of confidence in model predictions), systematic error analyses were performed using carefully assessed bounds of uncertainty for each of the parameter values.

5.3. Error Analysis

[36] The parameter values used in the model simulations are "best guessed" estimates, so all of them possess their own error bounds, based on the reliability of the sources of information. It is possible that the correct but unknown values may be, to some degree, different from those used in the simulations.

[37] The error bounds for the parameter values for each New Zealand catchment are presented in Table 5, and within these bounds, the parameters are assumed to follow the uniform distribution. The error bounds used in simulations were based on the quality of the sources of information, for each catchment, from which they were derived. The uniform distribution was selected because it is easy to use, and since we are mostly interested in notional estimates of predictive uncertainty only. The error bounds for M, D, Tc_{bf} and a are estimated directly, while error bounds for the parameters f_c and S_{bc} were estimated via the error estimates of the basic catchment properties ϕ, θ_{fc}, θ_{pwp}, and D. It should be also noted that the error bounds for soil properties are much greater at Moutere, Ngahere and Waihopai than they are for Mahurangi, because the source of information for the latter is much more comprehensive.

[38] The effect of these parameter uncertainties on model outputs was tested using the Monte Carlo technique. A random number generator was used to generate numerous combinations of randomly chosen parameter values within

Table 6b. Results of Error Analyses With Respect to the Correlation Coefficient $(\rho_{x,y})$

Timescale	Deviation, %	Moutere DI = 1.53	Waihopai DI = 0.82	Mahurangi DI = 0.81	Ngahere DI = 0.37
Annual[a]	upper error bar	1.04	0.00	1.10	1.10
	lower error bar	−4.17	−2.17	−2.20	0.00
Monthly[a]	upper error bar	6.74	1.08	0.00	1.08
	lower error bar	−7.87	−4.30	−1.05	−1.08
Daily[a]	upper error bar	5.81	9.87	9.09	15.00
	lower error bar	−8.14	−38.27	−9.09	−6.25

[a] Deviation from original simulation results $(\rho_{x,y})$.

Table 7a. Model Complexity: Annual

DI	Catchment	Model A	Model B	Model C	Model D
0.37	Ngahere	X			
0.70	Snobs Creek[a]		X		
0.81	Mahurangi		X		
0.82	Waihopai		X		
0.85	Broken River[a]		X		
0.90	Bass River[a]		X		
1.40	Stones Creek[a]			X	
1.53	Moutere				XX[b]
2.10	Bullock[a]				X

[a] Taken from *Farmer et al.* [2002].
[b] XX: extrapolated from observed trends.

the chosen error bounds, and these are used with the model to produce multiple realizations of model predictions for the same climatic inputs. (This is distinct from the sensitivity analysis of the previous section, where we varied one parameter at a time.) Specifically, 100 simulations were performed for each catchment and the results presented in Figure 8 and Tables 6a and 6b. The errors in the timing and magnitude of flow were assessed using $\rho_{x,y}$ and *RR*. Crosses indicate model predictions using the "best guessed parameter estimates", while the bars represent the error bars for each catchment. The larger the error, the less confidence we can place in the results. Effects of uncertainties in the rainfall and potential evaporation inputs were also tested. The error bounds for the input uncertainties used in these tests (rainfall and potential evaporation) were ±10% (above and below the mean) for each input. However, once again these results will not be discussed in detail because we are mainly interested in parameter uncertainty and the consequent impact on model predictive uncertainty.

5.3.1. Timing of Discharge

[39] Figures 8a–8c and Table 6b suggest that the error associated with the timing of predicted streamflow becomes progressively larger as we move from annual to daily timescales. These results suggest that, for a given model, we can place progressively less confidence in model predictions as we move to shorter timescales. This was to be expected, since the sensitivity of flow to variations in parameter values also became progressively larger at decreasing timescales.

5.3.2. Magnitude of Discharge

[40] To assess the level of confidence in the magnitude of flow predictions, the error bars for RR are presented graphically and numerically in Figure 8d and Table 6a. Larger error bars tell us that we can have less confidence in the

results of simulations, when using the best-guessed parameter values. Assuming a reasonable correlation between observed and predicted streamflow has been achieved for all four catchments ($\rho_{x,y} \approx 0.8$), our results imply that we can have relatively little confidence in the results of simulations at Moutere (RR $\gg$ 1), moderate confidence in the results at Waihopai (RR > 1), and good confidence in the results at Mahurangi and Ngahere (RR $\approx$ 1). However, smaller errors do not necessarily mean that we have more accurate results, as these are achieved when RR is close to one. When ranking all catchments in order of decreasing accuracy and decreasing confidence, we find that the catchments were in the same order, and this order corresponds to the order of the dryness index (DI). Table 6a and Figure 8d present these results and show the correlation between accuracy, confidence and DI.

[41] The effect of the 10% uncertainty in rainfall and potential evaporation on model predictions was found to be minimal, increasing the error bars for each catchment by roughly equal amounts. Based on the results of sensitivity analyses, the largest source of error in model simulations remains the uncertainty surrounding the value of f_c. The results of this analysis suggest that it is possible to establish a link between accuracy, confidence and DI Dry catchments produced relatively "less accurate" results with "little confidence", while relatively "more accurate" results were produced with "good confidence" in wetter catchments.

6. Discussion

[42] Combined sensitivity and error analyses have assessed model sensitivity to parameter variations, timescales and catchment dryness. The results have shown that, with decreasing timescales, the simple single bucket water balance model produces progressively less accurate results with increasingly poorer confidence. This is a consequence of the fact that the model becomes increasingly more sensitive to the estimated parameter values.

[43] Of all the parameters, sensitivity and error analyses have identified f_c as a dominant control on streamflow variability in these New Zealand catchments at annual, monthly and daily timescales. The results of sensitivity and error analysis suggest that the degree of dominance is well correlated with the level of confidence in the results (size of error bars), accuracy (assessed by $\rho_{x,y}$ and *RR*), and climatic dryness index (DI). Simulations using the derived model have shown that streamflow prediction in dry catchments is sensitive to the soil parameter f_c, with relatively
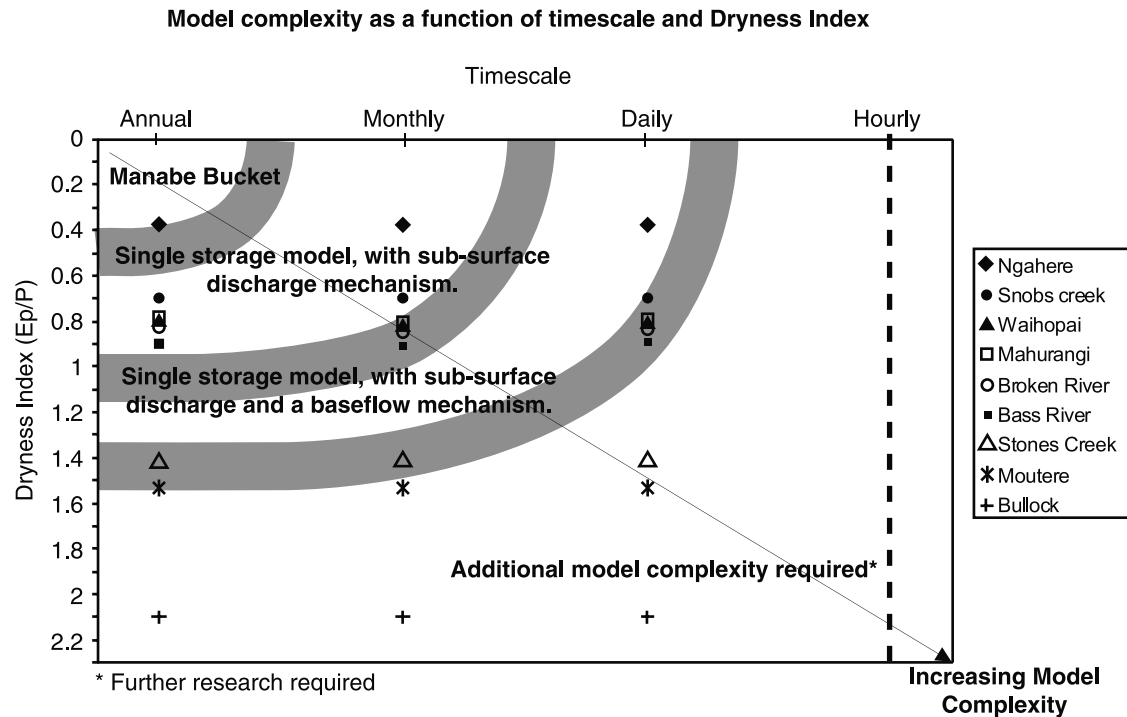
Table 7b. Model Complexity: Monthly

DI	Catchment	Model A	Model B	Model C	Model D
0.37	Ngahere		X		
0.70	Snobs Creek[a]		X		
0.81	Mahurangi		X		
0.82	Waihopai		X		
0.85	Broken River[a]			X	
0.90	Bass River[a]			X	
1.40	Stones Creek[a]			X	
1.53	Moutere				XX[b]
2.10	Bullock[a]				X

[a] Taken from *Farmer et al.* [2002].
[b] XX: extrapolated from observed trends.

Table 7c. Model Complexity: Daily

DI	Catchment	Model A	Model B	Model C	Model D
0.37	Ngahere			X	
0.70	Snobs Creek[a]			X	
0.81	Mahurangi			X	
0.82	Waihopai			X	
0.85	Broken River[a]				X
0.90	Bass River[a]				X
1.40	Stones Creek[a]				X
1.53	Moutere				XX[b]
2.10	Bullock[a]				X

[a] Taken from *Farmer et al.* [2002].
[b] XX: extrapolated from observed trends.

Model complexity as a function of timescale and Dryness Index

Figure 9. Hypothetical relationship between model complexity, timescale and climate characteristics.

less accurate results and with less confidence, while simulations in wet catchments are less sensitive to f_c, producing more accurate results and with greater confidence.

[44] Therefore accurate measurements of soil properties (namely, f_c) are required for predictions of flow in dry catchments but this is less important for predictions in wet catchments. A slight exception to this trend was observed for the Waihopai and Mahurangi catchments whose DI value are similar, yet they failed to display the same *RR*, or error bounds, or sensitivity to f_c. Analysis of Figures 6d and 8d shows Waihopai to have less accurate predictions (indicated by *RR*), larger output errors and greater sensitivity to f_c. The most likely cause of this is the difference in the sources of soil information. The source at Mahurangi is extensive, and more reliable than the information used at Waihopai, resulting in reduced parameter uncertainty, and reduced predictive uncertainty and sensitivity to parameter variations. Therefore the link established between DI, accuracy, error/confidence and sensitivity only applies when the sources of soil information are the same or of a similar quality. When all catchments have measured soil properties to the accuracy required to achieve similar levels of confidence in the results, the necessity of further increases of model complexity can be investigated. Detailed soil information was not available at all catchments in this case, so additional model complexity was not introduced, although this can be hypothesized.

[45] The results of simulations, sensitivity analysis and error analysis can also be used to identify the conditions under which additional model complexity is warranted. Simulation in wet catchments such as Ngahere suggest that model parameters have very little influence on the magnitude of streamflow produced at annual, monthly and daily timescales. Accurate predictions with confidence can be made at annual and monthly timescales using very basic models (such as the Manabe bucket), because the parameters

and processes they control have little effect on the water balance. The influence of model parameters (namely, f_c), does become important in the timing of daily streamflow and in the capturing low flow events, so the current model complexity is required to predict daily flows. Conversely, simulations in dry catchments such as Moutere suggest that the soil parameters have a stronger influence on the timing and magnitude of streamflow produced at annual, monthly and daily timescales. Therefore accurate predictions, with confidence, require the maintenance of the current model complexity at all three timescales, provided accurate soils information is available. Although detailed soils information may help improve the results, further complexity may be required to achieve daily predictions of acceptable accuracy.

[46] The combined results suggest that there are clear differences in the complexity required to model streamflow to a predefined accuracy and confidence, between wet and dry catchments, and with changing timescales. To summarize this trend a series of tables are presented listing the model complexity required for accurate predictions at annual (Table 7a), monthly (Table 7b), and daily (Table 7c) timescales for catchments with varying catchment dryness. All four New Zealand catchments have been considered in these tables along with an additional five Australian catchments studied by *Farmer et al.* [2002], who found a similar relationship between model complexity and timescale, but did not investigate its connection with catchment dryness in detail. The models of increasing complexity compared in Tables 7a, 7b, and 7c include: model A, the simple bucket model (Figure 2a) described by *Manabe* [1969]; model B, single storage model with a nonlinear subsurface discharge mechanism (Figure 2b); model C, single storage model with a nonlinear subsurface discharge mechanism and linear base flow (Figure 2c); model D, complex models. The complexity required has not been defined but is needed for accurate predictions. Suggested complexities may include an unsaturated zone,

routing, distributed representation, interception, hillslope representation and nonlinear base flow.

[47] "Appropriate model complexity" was identified when simulations produced predictions with $\rho_{x,y}$ approximately ≥ 0.8, RR close to 1, and there was little difference between the observed and predicted signature plots (as described by *Farmer et al.* [2002]). Analysis of the tables suggests the required model complexity increases with decreasing timescales, and with increasing catchment dryness (DI). This trend has been expressed schematically in Figure 9 in the form of 2-D picture linking required model complexity to the climatic dryness index (DI) and timescale. At this stage, the contours drawn into the schematic plot are merely qualitative indicators. Types of model complexity are added in, where appropriate, to indicate the type of processes that are considered important under the different combinations of DI and timescale. The admittedly qualitative relationship shown in Figure 9 assumes homogeneity of catchment characteristics, and roughly similar sized catchments. Although the model complexity needed for accurate predictions of hourly streamflow has been investigated recently by *Atkinson* [2002] using data from Mahurangi catchment, it has been left out of Figure 9 for reasons of brevity.

[48] The qualitative relationship shown in Figure 9 is a useful guideline that can aid the *a priori* selection of model complexity. It can help the modeler to choose a simple but effective model, avoiding the use of unnecessary parameterizations that may eventually cloud the physical meaning of the model predictions. With a model thus selected, the dominant controls on flow variability can then be identified and measures taken to accurately measure the appropriate model parameters to ensure acceptably accurate predictions are achieved, with sufficient confidence.

7. Conclusions

[49] Using a downward approach, and starting with a simple single bucket water balance model with physically meaningful parameter values, we have systematically added further process complexity (as required), with a view to making accurate predictions of streamflow at the annual, monthly and daily timescales in four New Zealand catchments. The results of sensitivity and error analyses suggest that the dominant control on streamflow variability is the soil parameter f_c, which is related to field capacity. For this model, it was found that more accurate predictions can be made with better confidence in wetter catchments, using relatively simple models, since such catchments tend to be less sensitive to f_c under the prevailing energy-limited conditions. Under water limited conditions, such as those found in dry or arid catchments, it is much harder to accurately predict streamflow with confidence, because the model is much more sensitive to f_c, so further model complexity may be required. More detailed soils information is therefore required to achieve accurate predictions.

[50] The results of sensitivity and error analyses enabled a qualitative conceptual relationship to be established between model complexity, measure of climatic wetness (dryness index), and the timescale. This relationship, presented by Figure 9, suggests that required model complexity increases with decreasing timescale, and increasing dryness index. It is suggested that the development of this relation-

ship, and its physical interpretation, is a step forward toward the development of a consistent method of a priori model selection that incorporates just the minimum level of complexity needed to predict streamflow with good confidence and acceptable accuracy. This way, unnecessary model complexity and parameterization are avoided, thus allowing us to explore the physical controls of streamflow variability in natural catchments, and then to spend more of our efforts toward estimating the critical model parameters either in the field or from other data sources.

Appendix A: Model Equations

Bucket capacity

$$S_{bc} = D(\phi - \theta_{pwp}) \tag{A1}$$

Threshold storage parameter and threshold storage

$$f_c = \frac{(\theta_{fc} - \theta_{pwp})}{(\phi - \theta_{pwp})} \tag{A2}$$

$$S_{fc} = S_{bc} \cdot f_c \tag{A3}$$

Transpiration

$$E_{veg} = ME_p \qquad \text{if } S(t) \geq S_{fc} \tag{A4}$$

$$E_{veg} = ME_p \left(\frac{S(t)}{S_{fc}} \right) \qquad \text{if } S(t) < S_{fc} \tag{A5}$$

Bare soil evaporation

$$E_{bs} = E_p(1 - M) \qquad \text{if } S(t) \geq S_{bc}$$

$$E_{bs} = E_p(1 - M) \left(\frac{S(t)}{S_{bc}} \right) \qquad \text{if } S(t) < S_{bc} \tag{A6}$$

Saturation excess

$$Q_{se} = 0 \qquad \text{if } S(t) \leq S_{bc}$$
$$Q_{se} = S(t) - S_{bc} \qquad \text{if } S(t) > S_{bc} \tag{A7}$$

Subsurface runoff

$$Q_{ss} = 0 \qquad \text{if } S(t) < S_{fc}$$

$$Q_{ss} = \left(\frac{S(t) - S_{fc}}{a} \right)^{\frac{1}{b}} \qquad \text{if } S_{fc} < S(t) < S_{bc} \tag{A8}$$

$$Q_{ss} = \left(\frac{S_{be} - S_{fc}}{a} \right)^{\frac{1}{b}} \qquad \text{if } S(t) > S_{bc}$$

Base flow

$$Q_{bf} = \left(\frac{S(t)}{Tc_{bf}} \right) \qquad \text{for all } S(t) \tag{A9}$$

Water Balance

$$S(t+1) = S(t) + P - E_{veg} - E_{bs} - Q_{se} - Q_{ss} \tag{A10}$$

Notation

S_{bc}	bucket capacity, the maximum storage of the bucket model, (mm).
D	soil depth, measured as the depth to an impervious layer promoting lateral subsurface runoff, (mm).
ϕ	porosity (dimensionless).
θ_{fc}	field capacity (dimensionless).
θ_{pwp}	permanent wilting point (dimensionless).
f_c	threshold storage parameter (dimensionless) ($0 < f_c < 1$).
S_{fc}	threshold storage (mm).
P	rainfall (mm/d).
E_{veg}	transpiration (mm/d).
E_{bs}	bare soil evaporation (mm/d).
E_p	potential evaporation (mm/d).
$M\%$	forest cover (dimensionless) ($0 \leq M \leq 1$).
$S(t)$	storage (mm/d).
$S(t+1)$	storage the next day (mm/d).
Q_{ss}	subsurface runoff (mm/d).
Q_{se}	saturation excess runoff (mm/d).
a	nonlinear discharge parameter ($\text{mm}^{1-b}\text{day}^{b}$).
b	nonlinear discharge parameter (dimensionless),
Q_{bf}	base flow (mm/d).
Tc_{bf}	base flow response time (days).
$\rho_{x,y}$	correlation between observed and predicted streamflow (dimensionless).
RR	runoff ratio (dimensionless).
Diff $\rho_{x,y}$	measure of spread in $\rho_{x,y}$ for variations in parameter values (dimensionless).
Diff RR	measure of spread in RR for variations in parameter values (dimensionless).
DI	dryness index.

References

Atkinson, S. E., Climate, soil, vegetation controls on streamflow variability with changing timescales and implications for model complexity, M.Sc. thesis, 110 pp., Cent. for Water Res., Univ. of Western Aust., Crawley, Australia, 2002.

Campbell, E. P., and B. C. Bates, Regionalization of rainfall-runoff model parameters using Markov Chain Monte Carlo samples, *Water Resour. Res.*, *37*(3), 731–739, 2001.

Chiew, F. H. S., M. J. Stewardson, and T. A. McMahon, Comparison of six rainfall-runoff modelling approaches, *J. Hydrol.*, *147*, 1–36, 1993.

Department of Scientific and Industrial Research, Soils of New Zealand, part 3, *Soil Bur. Bull.*, *26*(3), 1968.

Duncan, M. J., Hydrological impacts of converting pasture and gorse to pine plantation, and forest harvesting, Nelson, New Zealand, *J. Hydrol. N. Z.*, *34*(1), 15–41, 1995.

Eder, G., M. Sivapalan, and H. P. Nachtnebel, Modeling of water balances in Alpine catchment through exploitation of emergent properties over changing time scales, *Hydrol. Processes*, in press, 2002.

Farmer, D., M. Sivapalan, and C. Jothityangkoon, Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: Downward approach to water balance analysis, *Water Resour. Res.*, *38*, doi:10.1029/2001WR000328, in press, 2002.

Fukushima, Y., A model of river flow forecasting for a small forested mountain catchment, *Hydrol. Processes*, *2*, 167–185, 1988.

Garen, D. C., and S. J. Burges, Approximate error bounds for simulated hydrographs, *J. Hydraul. Div. Am. Soc. Civ. Eng.*, *107*(HY11), 1519–1534, 1981.

Jothityangkoon, C., M. Sivapalan, and D. L. Farmer, Process controls of water balance variability in a large semi-arid catchment: Downward approach to hydrological modeling, *J. Hydrol.*, *254*, 174–198, 2001.

Klemes, V., Conceptualization and scale in hydrology, *J. Hydrol.*, *65*, 1–23, 1983.

Manabe, S., Climate and the ocean circulation, 1, The atmospheric circulation and the hydrology of the Earth's surface, *Mon. Weather Rev.*, *97*(11), 739–774, 1969.

Milly, P. C. D., Climate, soil water storage, and the average annual water balance, *Water Resour. Res.*, *30*(7), 2143–2156, 1994.

Singh, V. P., *Computer Models of Watershed Hydrology*, 1130 pp., Water Resour. Publ., Highlands Ranch, Colorado, 1995.

Werner, P. W., and K. J. Sundquist, On the groundwater recession curve for large watersheds, *IAHS Publ.*, *33*, 202–212, 1951.

Wittenberg, H., and M. Sivapalan, Watershed groundwater balance estimation using streamflow recession analysis and baseflow separation, *J. Hydrol.*, *219*, 20–33, 1999.

S. E. Atkinson, and M. Sivapalan, Centre for Water Research, Department of Environmental Engineering, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6907, Australia. (sivapalan@cwr.uwa.edu.au)

R. A. Woods, National Institute of Water and Atmospheric Research Ltd., P.O. Box 8602, Christchurch, New Zealand. (r.woods@niwa.cri.nz)