## Monthly Precipitation-Evapotranspiration-Runoff Model (PERM)

PERM was developed specifically to estimate streamflow from GCM monthly precipitation and temperature times-series. This Supplementary Material describes PERM, its calibration and application to GCM output. The nature of the input data and the model's application necessitated a simple conceptual model based on a monthly time-step with a minimum number of lumped parameters to be optimised. The structure of PERM is shown in Figure 2 of the paper with its five parameters highlighted in bold.

*Model structure*

As observed in Figure 2 monthly precipitation is either added to the interception store (if the monthly mean daily temperature is > 0°C) or accumulated in a snow pack (if the monthly mean daily temperature is ≤ 0°C). If the interception store is full (*Imax*) any excess precipitation is subject to an infiltration function in which a surface runoff component (designated as *PAreaF*) is dependent on the contents of the soil moisture store. When precipitation is accumulated as snow, there is no evaporation for that month from the snow pack or evapotranspiration from the soil moisture store. The snow pack continues to accumulate as long as the monthly mean daily temperature is ≤ 0°C. The snow pack begins to melt when the monthly mean daily temperature is > 0°C. Snowmelt is partitioned into two components, a runoff and soil moisture infiltration component based on the parameter *Melt*. When the maximum capacity of the soil moisture store is exceeded, saturation excess runoff (*SMF*) occurs. When a snow pack is not present evaporation from the interception store occurs as a linear function of temperature. Any remaining energy is then used for evapotranspiration from the soil moisture store, which is estimated either as a linear function of the available soil moisture (water limited case) or as a linear function of mean monthly daily temperature (energy limited case). The algorithms representing these soil moisture or energy limiting conditions are given in Figure 2. Baseflow from the soil moisture store is simulated as a linear recession of the water content in the store.

The model therefore estimates runoff from four sources – a so-called partial area runoff, snowmelt runoff, saturation excess runoff and baseflow. Since the model is being calibrated using monthly/annual data the routing of streamflow is considered to be unnecessary in this study.

Two other hydrologic processes were considered for inclusion in PERM – impervious area and deep recharge. As we are dealing with non-urban catchments in the project, although some will have small areas of urbanisation, it was decided the model would not include an impervious parameter. With respect to PERM, deep seepage is defined as loss of water from the soil moisture store that recharges an underlying aquifer. Based on extensive reviews of recharge in Australia, initially by Petheram et al. (2002) and, more recently, by Crosbie et al. (2010) several empirical relationships have been developed between recharge, precipitation, vegetation and soil information. A large global study but only for semiarid and arid regions was conducted by Scanlon et al. (2006) who also developed several generalised relationships relating recharge to mean annual precipitation. Based on these three reviews the average maximum effect could be equivalent to 5% of mean annual precipitation. If deep seepage were a significant component of a catchments hydrology then its mean annual water balance may show unusual or implausible values. The 699 global catchments used in this analysis, from Peel et al. (2010), were identified as having 'plausible' long-term water balances, in that their mean annual runoff was less than mean annual precipitation and actual evapotranspiration (precipitation - runoff) was less than potential evapotranspiration. From the results of these reviews and taking into account the model time-step, the available data and noting the model parameters are optimised by calibration, it was concluded that incorporating imperious area and deep seepage would yield little benefit to the modelling exercise.

Finally, as a simple conceptual model PERM does not represent vegetation processes that may change and impact the quantity and variability of future runoff under enhanced $CO_2$ conditions, namely changes in herbivore predation, wildfire, stomatal conductance, leaf area index and species composition and distribution (see Peel, 2009).

*Input time-series data for PERM calibration*

A data set of 699 catchments was chosen from a larger data set of 1221 unregulated rivers with 10 or more years of continuous historical annual and monthly flows and concurrent monthly rainfall and temperature (see Peel et al., 2010 for details of selection process). Historical details regarding the development of this data set are given in McMahon et al. (1992, 2007) and Peel et al. (2001, 2004, 2010). Annual data and our analyses are based on water-years, not calendar years, and the streamflow data have been converted to mean annual runoff depth (by dividing mean annual flow volume by catchment area) and in units of mm. This unit allows direct comparison with annual precipitation.

Figure 3 shows the distribution of the 699 catchments with 430 located in the northern hemisphere and 269 in the southern hemisphere. Catchments vary in area from 130 $km^2$ (90th percentile) to 74,400 $km^2$ ($10^{th}$ percentile). The catchments exhibit a wide range of Köppen climate types (Peel et al., 2007) with the dominant climate type (≥75% of catchment area covered) being tropical for 94 catchments, arid for 42, temperature for 293, cold for 207, polar for 2, and a mixture (all climate types cover <75% of catchment area) for 61 catchments. Peel et al. (2010) provides details on how the concurrent monthly precipitation and temperature data for each catchment were collated.

*Calibration of PERM*

PERM is run on a monthly time-step using observed precipitation and temperature data and calibrated against observed annual runoff. At the beginning of each simulation, the snow accumulation store is set to zero and the interception and the soil moisture stores are set to half of their maximum capacities. The first twelve months of the period to be modelled are run twice, once to initiate the model stores and then for calibration purposes. Results from the first run through the first twelve months are ignored and are not used in the calibration.

The entire observed monthly runoff record is used in the calibration. The five model parameters are optimised to reduce an objective function defined as the sum of squared differences between the estimated and observed annual runoff (Equation A1).

$$OBJ = \sum_{i=1}^{n}(EST_i - REC_i)^2 \qquad \text{(A1)}$$

where $OBJ$ is the objective function, $EST_i$ is the estimated annual runoff for year $i$, $REC_i$ is the observed annual runoff, and n is the number of years of observed annual runoff.

Since the calibrated model will be used to simulate runoff for a climate change impact assessment based on annual summary statistics, extra effort is made to ensure the calibrated model is able to reproduce those observed runoff summary statistics. To this end penalties are applied during the calibration process to the objective function if the estimated annual runoff mean and or coefficient of variation differ from the observed values as outlined below.

- If the estimated and observed mean annual runoff differ by;
  - More than 5% then *OBJ* = *OBJ* x *5*
  - More than 10% then *OBJ* = *OBJ* x *25*
  - More than 20% then *OBJ* = *OBJ* x *125*
- If the estimated and observed annual coefficient of variation differ by;
  - More than 5% then *OBJ* = *OBJ* x *5*
  - More than 10% then *OBJ* = *OBJ* x *25*
  - More than 20% then *OBJ* = *OBJ* x *125*

The penalty of 5, and multiples thereof, was selected based on trial and error.

The mean annual runoff penalty is designed to ensure the estimated and observed mean annual runoff do not differ greatly. In general a good calibration using the objective function described in Equation A1 should produce little difference in the estimated and observed mean annual runoff. However, any calibration using this objective function will lead to an estimate of runoff with lower variability than observed (Gupta et al., 2009). Therefore, the annual coefficient of variation penalty is designed to constrain the model to reproduce the inter-annual variability of the observed runoff.

An automatic pattern search optimisation method is used to calibrate the model (Hooke & Jeeves, 1961; Monro, 1971) with 10 different parameter sets used as starting points to increase the likelihood of finding the global optimum of parameter values. The K-fold cross validation method described by Efron & Tibshirani (1993) is used to cross validate the calibrated model. The observed runoff is divided into K roughly equal parts (in this case K = 3). PERM is then calibrated against two parts of the observed runoff. The calibrated parameters are then used to estimate the runoff of the remaining part. This process is repeated K times, so that all parts are estimated once. The quality of the calibration can be verified by comparing the calibration against the cross validation estimated runoffs.

The median and range of calibrated PERM parameter values are set out in Table A1 along with the parameter limits. As most precipitation-runoff models are operated at a daily or shorter time step, little guidance was available in the literature regarding the limits of the five parameters. Values were based on experience of the authors in operating a range of hydrologic models. Limits of $K$ and $Melt$, which are proportional variables, were set at 0 and 1. These limits were reached by 88 and 115 catchments respectively, which suggests PERM is not modelling the melt and baseflow processes very well. For the other parameters, the lower limit was set to zero except for $Smax$ for which we adopted 50 mm. The upper limit for $Smax$ was 2000 mm which may have been too low given this limit was adopted for 16% of catchments. The upper limit of parameter $a$, which is the depth of snowmelt or evapotranspiration per °C per month, was set at 200 mm (Maidment, 1992, page 7.25). The lower limit of $a$ is set to zero. The maximum limit for $a$ was not reached for any catchment and $IC$ was reached in 19 catchments suggesting the limit set is satisfactory.

*Model performance*

Although PERM was calibrated for 699 catchments only a sub-set of 17 catchments were selected for later analyses (Figure 3 of the paper). Figure A1 shows exceedance curves of four metrics of model performance for the 699 catchments – Nash-Sutcliffe efficiency (NSE) for calibrated and verified annual flows, annual $R^2$ (between modelled and observed flows), and monthly NSE values (Nash & Sutcliffe, 1970). The NSE estimates the proportion of recorded runoff variance that is described by the model. The median value of NSE between modelled and observed annual runoffs across the 699 catchments is 0.60 and, for comparison, the equivalent correlation coefficient squared ($R^2$) is 0.63. Only 20% of catchments have an annual NSE ≥ 0.8, while 66% of catchments have an annual NSE ≥ 0.4. The median NSE between modelled and observed monthly runoffs is 0.33.

An objective of this study is to minimise the uncertainty in future runoff due to a poor hydrologic model calibration. Thus a sub-set of catchments were selected for further analysis based on the following criteria.

- Annual NSE > 0.8;
- Annual K-fold validation NSE > 0.60;
- Monthly NSE > 0.6;
- Catchment area > 1,000 $km^2$;
- Modelled mean annual runoff within ± 5% of observed value; and
- Modelled annual coefficient of variation within ± 10% of observed value.

The limitations set out above are consistent with a survey by Chiew & McMahon (1993). They found that NSE values > 0.6 are considered satisfactory subject to inspection of graphical plots of results and E > 0.8 are considered acceptable so long as the mean flow is estimated to within 10%. For a monthly model as used in this project, a cut-off value of monthly NSE of 0.6 and mean annual runoff < ±5% is not inconsistent with the Chiew & McMahon's recommendation. Forty-five catchments met the above criteria. Many of these catchments are clustered spatially, so in order to minimise redundant information and present results for a manageable set of catchments we selected a sub-set of 17 catchments from the 45 catchments that met the selection criteria that are well distributed world-wide. The locations of the final 17 catchments selected for subsequent analysis are shown in Figure 3.

**Supplementary Material References**

Chiew, F.H.S., McMahon, T.A., 1993. Assessing the adequacy of catchment streamflow yield estimates. *Journal of Hydrology*, **31**, 665-680.

Crosbie, R.S., Jolly, I.D., Leaney, F.W., Petheram, C., 2010. Can the dataset of field based recharge estimates in Australia be used to predict recharge in data-poor areas? *Hydrology and Earth System Sciences*, **14**, 2023-2038.

Efron, B., Tibshirani, R.J., 1993. *An Introduction to the Bootstrap*. Chapman & Hall, New York, 436 pp.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, **377**, 80-91.

Hooke, R., Jeeves, T.A., 1961. Direct search solution of numerical and statistical problems. *Journal Association Computing Machines*, **8** (2), 212-229.

Maidment, D.R., 1992. *Handbook of Hydrology*, McGrawHill.

McMahon, T.A., Finlayson, B.L., Haines, A.T., Srikanthan. R., 1992. *Global Runoff - Continental Comparisons of Annual Flows and Peak Discharges*. Catena Verlag, Cremlingen – Destedt, Germany.

McMahon, T.A., Vogel, R.M., Peel, M.C. and Pegram, G.G.S., 2007. Global hydrology – Part 1; Characteristics of annual and monthly streamflows. *Journal of Hydrology*, **347**, 243-259.

Monro, J.C., 1971. *Direct search optimisation in mathematical modelling and a watershed model application.* National oceanic Atmospheric Administration, National Weather

Service, US Dept. of Commerce, NOAA, Silver Spring, MD, Tech. Memo. NWS HYDRO-12, 52.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models, 1. A discussion of principles. *Journal of Hydrology*, **10**, 282-290.

Peel, M.C., 2009. Hydrology: catchment vegetation and runoff. *Progress in Physical Geography*, **33**(6), 837-844.

Peel, M.C., Finlayson, B.L. and McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth Systems Sciences* **11**, 1633-1644.

Peel, M.C., McMahon, T.A., Finlayson, B.L., Watson, F.G.R., 2001. Identification and explanation of continental differences in the variability of annual runoff. *Journal of Hydrology*, **250**, 224-240.

Peel, M.C., McMahon, T.A., Finlayson, B.L., 2004. Continental differences in the variability of annual runoff – update and reassessment, *Journal of Hydrology*, **295**, 185-197.

Peel, M.C., McMahon, T.A., Finlayson, B.L., 2010. Vegetation impact on mean annual evapotranspiration at a global catchment scale. *Water Resources Research*, **46**, W09508, doi:10.1029/2009/2009WR008233.

Petheram, C., Walker, G., Grayson, R., Thierfelder, T., Zhang, L., 2002, Towards a framework for predicting impacts of land-use on recharge: 1. A review of recharge studies in Australia. Australian *Journal of Soil Science*, **40**, 397-417.

Scanlon, B.R., Keese, K.E., Flint, A.L., Flint, L.E., Gaye, C.B., Edmunds, W.M., Simmers, I., 2006, global synthesis of groundwater recharge in semiarid and arid regions. *Hydrological Processes*, **20**, 335-3370.

**Tables A1** Median values of model parameters and their ranges for 699 catchments

| | Smax (mm) | a (mm/°C/month) | K | Melt* | IC (mm) |
|---|---|---|---|---|---|
| Limits | 50 - 2000 | 0 - 200 | 0 - 1 | 0 - 1 | 0 - 300 |
| Median | 838 | 6.8 | 0.20 | 0.34 | 47.2 |
| $25^{th}$ percentile | 368 | 4.0 | 0.07 | 0.11 | 16.2 |
| $75^{th}$ percentile | 1602 | 18.0 | 0.43 | 0.62 | 102 |
| Limits: number of catchments at limits | 50: 23#<br>2000: 114 | 0: 6<br>200: 0 | 0: 37<br>1: 51 | 0: 103<br>1: 12 | 0: 98<br>300: 19 |

* 179 catchments exhibited monthly melt runoff

#limits: number of catchments with model parameter values at limit shown

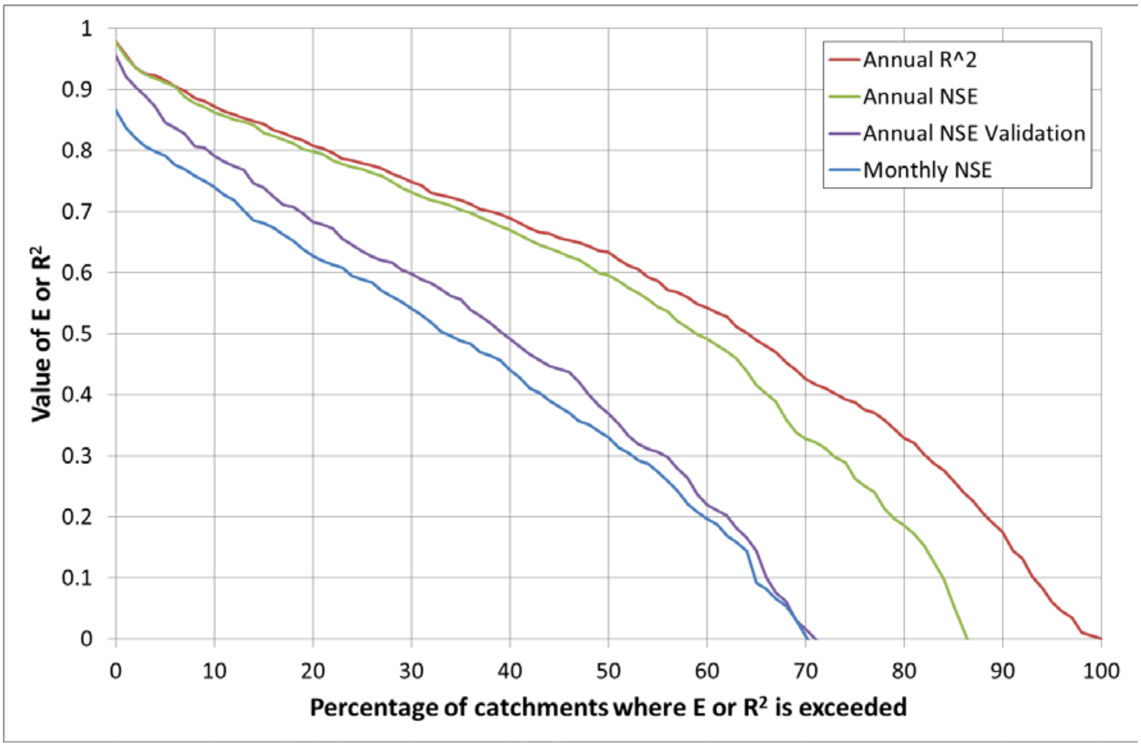


**Figure A1** Percentage of catchments with Nash and Sutcliffe efficiency (NSE) or $R^2$ (between observed and modelled flows) values greater than or equal to a given value