Integrating Site- and Regional-scale Data in Assessing the Hydrological Impact of Afforestation Using Rainfall-runoff Curves

Ashley J. B. Greenwood
Richard G. Benyon
Patrick N. J. Lane

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference
Integrating Site- and Regional-scale Data in Assessing the Hydrological Impact of Afforestation Using Rainfall-runoff Curves

Ashley J. B. Greenwood\textsuperscript{1,2}, Richard G. Benyon\textsuperscript{2}, Patrick N. J. Lane\textsuperscript{1,2}

\textsuperscript{1}CRC for Forestry, Sandy Bay Tasmania, Australia
\textsuperscript{2}Department of Forest and Ecosystem Science, University of Melbourne, Parkville Victoria, Australia

ashleyjg@student.unimelb.edu.au

Abstract: Following the procedure of Greenwood et al. [2011], a re-parameterised form of the hyperbolic tanh function of McMahon et al. [1992] and the well-known USDA model are used to explore the regional hydrological impacts of afforestation in Australia. Quantile regression is used to analyse a large, long-term, mean annual rainfall-runoff dataset. According to Budyko [1974], variability in regional mean annual runoff is driven by the regional radiation-water balance or potential evapotranspiration. The mean annual dataset is taken to represent a comprehensive description of the regional hydrology, supported by Australia’s distinctive regional hydro-climatology. Regional plantation and pre-conversion ‘grassland’ runoff relationships are identified by statistical association with high and low potential evapotranspiration rainfall-runoff quantiles using site-scale experimental data. Both models gave similar results, however the standard USDA model provided a superior fit to the regional data than the re-parameterised tanh. The result underlines the importance of regional context, that is, the representativeness of the regional data and the chosen model, in providing confidence in an assessment, over the quality of site data. A case study demonstrates how good regional context can be used to transparently scrutinise the appropriateness of site data for use in regionalisation. The approach has applicability to any distinctive hydro-climatic region that can be reasonably described by rainfall-runoff functions.

Keywords: forest hydrology; rainfall-runoff curves; quantile regression

1 INTRODUCTION

The hydrological impacts of land-cover conversion of grassland to plantation forestry are likely to become an issue of even greater interest to policy makers in coming years. International studies predict reductions in water availability due to climate change [Bates et al., 2008]. While afforestation is recognised as having a beneficial role in managing greenhouse gas emissions [UNFCC, 2009], it is also known to reduce catchment water yield [Brown et al., 2005]. The issue is compounded by sensitivities peculiar to water resources management. Access to water is regarded as a human right [WMO, 1992] requiring equitable management [Syme et al., 1999] and therefore robust assessments with high transparency. Unfortunately the availability of appropriate data to support such assessments is mixed. Most has been collected at small-scale sites (<100 km\textsuperscript{2}) and may not represent the forest treatments, environments [Lane et al., 2005] or flow regimes of interest [Brown et al., 2005].
A fundamental challenge to forest hydrology then, is to transparently validate the results of limited, small-scale assessments over regional-scales to inform planning and decision-making. Greenwood et al. [2011] demonstrated how this may be accomplished with non-linear quantile regression, well-known empirical rainfall-runoff curves and a cohesive, long-term Australian mean annual rainfall-runoff dataset. However, problems were encountered with the optimisation algorithms used in quantile regression. Convergence to multiple optimal parameter sets associated with extreme parameter interaction resulted in incoherent quantile curves. Such ill-posed behaviour was something of a surprise in simple two-parameter models.

Their ensuing analyses evaluated a single parameter tanh function and the well-known USDA curve number model [Greenwood et al., 2011]. The USDA function is parameterised to incorporate a fixed ratio of catchment losses which has been found to be suitable over a range of North American catchments [Figure 10-1, USDA-NRCS, 2004]. Lacking this flexibility, the tanh model gave a poorer description of the Australian data [Greenwood et al., 2011]. This study follows the procedure of Greenwood et al. [2011] and applies nonlinear quantile regression to data from regional Australian catchments to determine if a re-parameterised tanh function (Equation 3, Table 1 below) is superior to the USDA model in assessing the hydrological impacts of afforestation. The re-parameterisation was facilitated through the calibration of the two-parameter tanh model of McMahon et al. [1992] (Equation 2, Table 1 below) against regional Australian data using Bayesian techniques. Parameter identification was sufficiently robust to define the ratio of tanh loss parameters as a constant for regional Australia, analogous to the USDA initial abstraction ratio. Despite its improved flexibility the re-parameterised tanh model did not provide as good a description of regional data as the USDA model. A case study is presented which demonstrates how quantile curves can be used in validating site-scale hydrological afforestation data for application in regional assessment.

2 THOERY - REGIONAL AUSTRALIAN HYDRO-CLIMATOLOGY

Large regions of the Earth are known to exhibit distinctive hydrological characteristics. Budyko [1974] delineated large-scale geo-botanic zones according to their long-term mean annual heat and water balances. The dominant regional influences in controlling regional hydrological variability were climatic, while local discrepancies could be attributed to non-climatic, site-dependent influences. If it may be assumed that a large, mean annual rainfall-runoff dataset provides a comprehensive representation of hydrological variability for a given region, then regional runoff will be driven by the regional heat-water balance or potential evapotranspiration [Budyko, 1974]. The lower runoff quantiles will represent higher levels of potential evapotranspiration and higher quantiles lower potential evapotranspiration.

Australian runoff (streamflow) patterns have been recognised as distinctive among global data, arising from the broader variability of regional hydrology and climate [McMahon et al., 1992; Peel et al., 2004]. The El Niño-Southern Oscillation teleconnection across the equatorial Pacific Ocean; mid-latitude Southern Ocean systems and systems originating in the equatorial eastern Indian Ocean interact over south and south-eastern Australia, producing a distinctive regional hydro-climatic environment [Risbey et al., 2009]. Unlike continents of the northern hemisphere, Australia's modest orographic relief does not exert a major influence on its weather, but is subject to global climatic circulation patterns [Gentilli, 1971]. Consequently, regional hydrology should be heavily influenced by regional climatic drivers and a representative, regional-scale hydrological dataset should reflect regional climatic variability.
3 MATERIALS AND METHOD

Long-term, regional Australian data were taken from Peel et al. [2000]. The data were compiled to provide an extended time series of unimpaired streamflow information for use in both research and management of Australia's hydrological and ecological systems. They comprise mean annual rainfall-runoff data pairs and runoff coefficients from 331 catchments, their location (including administrative jurisdiction), catchment area, gauge number and station name (see Figure 1). Land-cover information was not recorded [Peel et al., 2000].

![Figure 1](image)

**Figure 1.** Regional data used in the study. Areas with potential for new plantations after Keenan et al. [2004] are shaded inset.

A subset of 313 sites was used in this assessment, to provide data that were more representative of the regions identified by Keenan et al. [2004] as having potential for the development of new plantations (Figure 1).

Hydrological grassland to plantation conversion data used here were compiled by Greenwood et al. [2011]. ‘Grassland’ runoff was taken as pre-treatment runoff, which ranged from true grassland to cleared forests. None of the data were contained within the Peel et al. [2000] temperate subset. The paucity of the conversion data prompted the use of a number of native vegetation sites to provide an indication of the validity of the selection of quantile ‘forest’ curves later in the study (see Figure 3 below).

The rainfall-runoff models used in this study are shown in Table 1 (below). Empirical rainfall-runoff relationships have a history of use in policy development and assessing the regional hydrological impacts of afforestation for water management, although legitimate concerns have been expressed around their predictive ability under conditions that vary from those of the original studies [see Greenwood et al., 2011 and citations therein]. They continue to find application in
Australia [Brown et al., 2007; Post et al., 2012], where the data available to support detailed modelling over large areas are limited.

A key assumption of the USDA equation (Equation 1, Table 1) is that an empirical, linear relationship may be formed between the initial abstraction (or initial loss, \( I_a \)) and the maximum potential retention (\( S \)) of a site, termed the initial abstraction ratio \( \lambda = I_a/S \). Moreover, \( \lambda \) may be equated to 0.2 for all sites across the United States to simplify field calculations [USDA-NRCS, 2004].

### Table 1: Rainfall-runoff functions.

<table>
<thead>
<tr>
<th>USDA-NRCS [2004]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ( RO = \frac{(P - \lambda S)^2}{(P + (1 - \lambda)S)} ); ( S = \frac{1000}{CN} - 10 )</td>
</tr>
<tr>
<td>Event-based runoff</td>
</tr>
<tr>
<td>( RO ) = runoff (inches/event)</td>
</tr>
<tr>
<td>( I_a ) = initial abstraction (inches/event)</td>
</tr>
<tr>
<td>( \lambda ) = initial abstraction ratio (( I_a/S ))</td>
</tr>
<tr>
<td>( CN ) = curve number (range: 1 to 100)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>McMahon et al. [1992]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 ( RO = P - IL - F \tanh\left(\frac{P - \gamma F}{F}\right) )</td>
</tr>
<tr>
<td>( RO ) = runoff (mm)</td>
</tr>
<tr>
<td>( P ) = precipitation (mm)</td>
</tr>
<tr>
<td>( IL ) = notional loss (mm)</td>
</tr>
<tr>
<td>( F ) = notional infiltration (mm)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Zhang et al. [2001]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 ( RO = P - \gamma F - F \tanh\left(\frac{P - \gamma F}{F}\right) )</td>
</tr>
<tr>
<td>loss ratio: ( \gamma = IL / F )</td>
</tr>
</tbody>
</table>

| | Runoff form, assuming no changes in storage. |
|-----------------|
| 4 \( RO = P - P \left[ \frac{1 + \omega} \right] \left( \frac{E_x}{P} \right)^{\frac{1}{3}} \) |
| \( RO \) = mean annual runoff (mm/year) |
| \( P \) = mean annual precipitation (mm/year) |
| \( \omega \) = plant available water coefficient, 0.5 for ‘grassland’ 2.0 for ‘forest’ |
| \( E_x \) = mean annual evapotranspiration optimised for a given land-cover, 1,100 mm/year for ‘grassland’, 1,410 mm/year for ‘forest’ |

Hyperbolic functions have been in use to describe runoff and other hydrological quantities since the early 20th Century [see Budyko, 1974 and work cited therein]. McMahon et al. [1992] used a two parameter tanh function to model annual runoff in a study of global runoff variability (Equation 2, Table 1). The structure of Equation 2 resembles the USDA function in that it contains terms depicting initial loss (\( IL \)) and infiltration (\( F \)). Equation 2 can be re-parameterised by defining a loss ratio \( \gamma = IL / F \) (Equation 3, Table 1), analogous to the USDA initial abstraction ratio \( \lambda = I_a/S \). If the ratio \( \gamma \) is known the tanh function reduces to a robust one-parameter model which retains some flexibility based on knowledge of average catchment losses.

Recent work on rainfall-runoff curves has used Markov chain Monte Carlo (MCMC) to calibrate Equation 2 against the temperate Peel et al. [2000] subset [Greenwood et al., in submission]. The

![Figure 2. Tanh joint posterior parameter distribution, after Greenwood et al. [in submission]](image-url)
multivariate normal form of the resulting joint posterior parameter distribution (Figure 2) resulted in high levels of confidence in parameter estimates for the tanh model ($L = 180$, $F = 910$ mm/year, rounded to nearest 5 mm/year). Interestingly, the value for $\gamma$ based on these data is $L/F = 180/910 = 0.2$, identical to the USDA estimate for regional application.

Quantile regression enables a more complete description of information than conventional regression by describing different proportions within a dataset. The principles underlying quantile regression may be described analogously to the more familiar least squares linear regression. Linear regression fits a straight line function to data. If the data are independent and the errors are normally distributed, the best fitting straight line arises from the parameters which minimise the variance between the data and modelled estimates. Given these assumptions, the estimate that minimises variance is the observed mean. However regression can be conducted around other estimates. Regression by absolute deviation minimises the sum of absolute errors, leading to an estimate of the median. It can also be shown that minimising the sum of asymmetrically weighted absolute residuals will yield the other quantiles [Koenker and Park, 1996].

This study applies the nonlinear quantile regression of Koenker [2009] to regional Australian data after Greenwood et al. [2011], using the re-parameterised tanh function (Equation 3). Regional ‘grassland’ and ‘forest’ quantile curves are selected as those with the highest agreement with the conversion data (mean annual rainfall <1000 mm/year) using the Kolmogorov-Smirnov (KS) test and root-mean-square-error (RMSE). Results are compared with the USDA model to determine which is better for assessing the regional impacts of afforestation. The following case study demonstrates the use of quantile curves in validating site-scale data for use in the regional assessment of afforestation.

4 RESULTS

The response of rainfall-runoff curves to parameter variation is shown in Figures 3(a) (Equation 1) and 3(b) (Equation 4) below. Both curve systems appear similar in their variability of form and potential to describe data. The theoretical linear limit of runoff equal to precipitation is shown in each panel for comparison.

Quantile curve arrays are shown in Figures 3(c) and (d). Log-log scales are used to facilitate examination behaviour in areas of low rainfall (note 1,000 mm/year \( \equiv \) 3.0 on log_{10} scale). Site-scale data used to identify regional quantile curves are shown in Figure 3(e) and (f). All sites were afforested with \textit{pinus radiata}. Only one comprised a classic paired catchment study - Redhill in NSW, data from which were collated for the cleared period (1a) and the control at Kiley’s Run (1b) (see Figure 3f). Balingup in WA (5) was planted with a mixture of pines and eucalypts. March Road (also WA, 6) was described as planted regrowth its runoff response may be closer to native vegetation. Rainfall was lower at March Road during the pre-treatment period and exceeded 1,000 mm/year following conversion (Figure 3b). The paucity and scatter of data from wetter sites (>1,000 mm/year rainfall) limited their value in identifying coherent curves, only data with <1,000 mm/year rainfall were used to select final quantile curves [see Greenwood et al., 2011].

Median curves of both models showed strong agreement with the regional data (KS p-values: 0.37 for tanh and 0.27 for USDA). Despite the lower KS p-value lower quantile USDA curves showed greater sensitivity to the entire dataset, providing a much better description of its variability at low rainfall (commencing around 800 mm/year, \( \equiv \)2.9 on log_{10} scale) (Figure 3d). Selected ‘grassland’ and ‘forest’ quantile curves are shown in Figures 3(e) and (f) along with standard errors of the parameter estimates and their significance [see Koenker, 2009].
Figure 3: Rainfall-runoff curves: (a) and (b) show functional responses for different parameter values; (c) and (d) show quantile curve arrays conditional on Equations 1 and 4 (log-scale); (e) and (f) show site-scale conversion data and selected regional ‘grassland’ and ‘forest’ quantile curves. Percentages of catchments treated are shown in brackets next to site names (Figures 3e and f). Dark grey conversion data >1,000 mm/year rainfall.

5 CASE STUDY

Relationships generated using empirical data from small-scale sites may be useful for estimating catchment responses to afforestation at larger scales [Bren and Hopmans, 2007].
Figure 4: (a) Conversion data from wet sites and an alternative empirical curve set. (b) Same data presented with quantile curves derived in this work, conversion data from dry areas, regional data and data from native bush sites.

Data from three wet plantation conversion sites (>1,000 mm/year) are shown in Figure 4(a). The data are high quality and have been used in a number of studies [for example Bren and Hopmans, 2007; Lane et al., 2005]. Also shown are Zhang [2001] curves (Equation 4 Table 1), an empirical curve system widely used in Australia. The conversion data accord with Zhang curves. Without any further statistical context the information shown in Figure 4(a) would encourage the extrapolation of both the data and the curves to other parts of Australia, which has indeed been the case [see Greenwood et al., 2011].

However, when wet conversion data are plotted with a coherent regional dataset (Figure 4b), the additional statistical context provides a transparent indication of their appropriateness for regionalisation. Levels of runoff from the pine plantation at Cropper’s Creek are particularly discordant with regional mean annual data (Figure 4b, forest site 8), much higher than its bush control sites (Figure 4b, sites 8a and b). However given similar environmental characteristics, native vegetation should yield higher runoff than pine plantations due to its open canopy [Putuhena and Cordery, 2000]. Clearly the site characteristics of Clem Creek, most obviously its steepness, play an important role in runoff generation [see also Greenwood et al., 2011], raising questions around its appropriateness for use in regional assessments.

Empirical curves can be used to construct rainfall-runoff reduction relationships relative to pre-treatment runoff. Figure 5 (below) shows such relationships constructed using the USDA and tanh quantile curves derived in this study and the Zhang curves. The four data points were taken from independent studies [see Greenwood et al. 2011 for details]. Reductions based on the quantile curves show the strongest agreement with the data (KS p values: 1.00 USDA; 0.77 tanh; 0.03 Zhang). In the 600 to 800 mm/year rainfall range, where most future planting is likely in Australia [Keenan et al., 2004], regional runoff reductions due to afforestation may be underestimated by 15-20%, using Zhang curves (and data from wet sites), with implications for confidence and equity in decision-making.
6 CONCLUSIONS

If a mean annual dataset provides a comprehensive representation of regional hydrological variability, quantile regression may be used to extract regional rainfall-runoff relationships that correspond to runoff responses from conceptual grassland and plantation land cover at a regional scale. The approach makes use of available site-scale data in a way which can simultaneously identify regional trends and scrutinise the appropriateness of site data for use in regionalisation. If significant discord is evident between regional- and site-scale data, it is likely that atypical, small-scale influences are obscuring regional relationships and the site-scale data concerned should not be used in regionalisation. The quality of the results depends more on the representativeness of the regional data and the ability of the chosen model to describe it, than the quality of the site data. As the number of available sites increases, the more robustly will the quantile curve depicting the particular regional land-cover runoff relationship be defined. A key feature of the approach is its transparency, which has the potential to engender confidence in the outcomes of regional assessments.

ACKNOWLEDGEMENTS

This study was funded by a University of Melbourne scholarship made available through the Cooperative Research Centre for Forestry. Additional Australian time series data were provided by Dr Lu Zhang and Dr Fangfang Zhao of CSIRO.

REFERENCES


Greenwood, A.J.B. et al., On the Bayesian evaluation of uncertainty and model choice with implications for water management in submission.


Putuhena, W.M., Cordery, I., Some hydrological effects of changing forest cover from eucalypts to *Pinus radiata*, *Agricultural and Forest Meteorology*, 100(1), 59-72, 2000.


