SMITH COLLEGE

Smith ScholarWorks

Engineering: Faculty Publications   Engineering

10-30-2017

Predicting Dry-Season Flows with a Monthly Rainfall–Runoff Model: Performance for Gauged and Ungauged Catchments

Perrine Hamel
Stanford University

Andrew John Guswa
Smith College, aguswa@smith.edu

Jake Sahl
Stanford University

Lu Zhang
CSIRO Land and Water

Follow this and additional works at: https://scholarworks.smith.edu/egr_facpubs

Part of the Engineering Commons

Recommended Citation

Predicting dry-season flows with a monthly rainfall-runoff model: performance for gauged and ungauged catchments

SCHOLARONE™
Manuscripts

1 Predicting dry-season flows with a monthly rainfall-runoff model:

2 performance for gauged and ungauged catchments

3

4 Perrine Hamel*[a], Andrew Guswa[b,a], Jake Sahl[a], Lu Zhang[c]

5

6 [a] Natural Capital Project, Woods Institute for the Environment, Stanford University.
7 371 Serra Mall, Stanford, CA 94305, USA. ph: +1 650-725-1783
8 perrine.hamel@stanford.edu; jakesahl@gmail.com

9 [b] Picker Engineering Program, Smith College, Northampton, MA 01063. aguswa@smith.edu

10 [c] CSIRO Land and Water, GPO Box 1666, Canberra ACT 2601, Australia. lu.zhang@csiro.au

11

12 * Corresponding author

1

13 **Abstract**

14 Hydrologic models are useful to understand the effects of climate and land-use changes on dry-

15 season flows. In practice, there is often a trade-off between simplicity and accuracy, especially

16 when resources for catchment management are scarce. Here, we evaluated the performance of

17 a monthly rainfall-runoff model (dynamic water balance model, DWBM) for dry-season flow

18 prediction under climate and land-use change. Using different methods with decreasing

19 amounts of catchment information to set the four model parameters, we predicted dry-season

20 flow for 89 Australian catchments, and verified model performance with an independent dataset

21 of 641 catchments in the United States. For the Australian catchments, model performance

22 without catchment information (other than climate forcing) was fair; it increased significantly as

23 the information to infer the four model parameters increased. Regressions to infer model

24 parameters from catchment characteristics did not hold for catchments in the United States,

25 meaning that a new calibration effort was needed to increase model performance there.

26 Recognizing the interest in relative change for practical applications, we also examined how

27 DWBM could be used to simulate a change in dry-season flow following land-use change. We

28 compared results with and without calibration data, and showed that predictions of changes in

29 dry-season flow were robust with respect to uncertainty in model parameters. Our analyses

30 confirm that climate is a strong driver of dry-season flow and that parsimonious models such as

31 DWBM have useful management applications: predicting seasonal flow under various climate

32 forcings when calibration data are available, and providing estimates of the relative effect of

33 land-use on seasonal flow for ungauged catchments.

34 **Keywords**: baseflow; land-use change; climate change; DWBM; prediction for ungauged basins

35

36 **1   Introduction**

37 With increasing pressure on water resources globally, managers of water resources need to

38 understand how streamflows – in particular, dry-season flows – respond to changes in land use

39 and climate. Applications vary broadly: at the global scale, hydrologists aim to better predict the

40 effect of agricultural expansion on water resources to avoid additional pressure in water-scarce

41 regions (Brauman et al., 2016). At the regional scale, water resources assessments are needed

42 to explore and implement efficient water-allocation plans (Kirby et al., 2014). For example, the

43 development of hydropower production facilities in Africa or South-East Asia requires the

44 prediction of annual and monthly flows (Vogl et al., 2016). In Latin America, the development of

45 investment in watershed services programs requires stakeholders to estimate the effect of land

46 management on hydrological services (Bremer et al., 2016; Guswa et al. 2014).

47 A number of knowledge gaps hinder the development of decision-aid tools for water resources

48 management. First, the effects of environmental changes on baseflow remain uncertain

49 (Andréassian, 2004; Brown et al., 2013, 2005; Price, 2011). Here, we define baseflow as

50 "streamflow fed from deep subsurface and delayed shallow subsurface storage between

51 precipitation and/or snowmelt events" (Price, 2011). Baseflow depends on many factors: climate

52 (magnitude and seasonality of precipitation and evapotranspiration), topography, geology, and

53 land use and land cover – with vegetation type and age as key subfactors (Brutsaert, 2008; Gao

54 et al., 2015; Zhang et al., 2014). In addition, the relative importance of these factors vary in time,

55 at the event and seasonal time scales (Devito et al., 2005; Jencso and McGlynn, 2011), making

56 it difficult to characterize in a given location. Second, relatedly, hydrologic models are limited in

57 their ability to estimate dry-season flow: lumped models tend to oversimplify the complexity of

58 hydrological processes, which casts doubt on their capacity to predict the effect of land use or

59 climate change. Complex models have high-data needs, require calibration, and often show

3

60 high uncertainty for predictions outside of the calibration conditions (in particular under land-use

61 change) (McIntyre et al., 2014; Smith et al., 2004).

62 Recognizing and seeking to fill these knowledge gaps is important, and taking stock of current

63 knowledge and its usefulness for practical applications is of equal priority for management. By

64 identifying questions that are of interest for water-resources management, hydrologists can

65 better understand where research gaps need to be filled. Typically, answering landscape

66 management questions requires an understanding of: i) the absolute magnitude of the *change*

67 in dry-season flow following land-use or climate change; ii) the relative difference in dry-season

68 flows among various land uses or management scenarios (e.g. afforestation, deforestation,

69 water abstraction for domestic or agricultural use); and iii) the spatial distribution of contributions

70 to baseflow (i.e. whether some part of the landscapes provide more baseflow than others)

71 (Guswa et al., 2014).

72 This paper explores the first two questions by analyzing how a simple monthly rainfall-runoff

73 model can capture major drivers of dry-season flow. Our aim is to quantify predictive uncertainty

74 in dry-season flow across a wide range of climate and catchment characteristics, and to assess

75 how this uncertainty changes as catchment information is introduced. In an era of increasingly

76 available data, in particular global daily precipitation data (Gehne et al., 2016), our work at the

77 monthly time step is justified by the parsimony of models operating at this time scale (Mouehli et

78 al., 2006). This characteristic facilitates regionalization and work in ungauged basins (Perrin et

79 al., 2001), as well as any analysis that does not necessitate short time-scale representation of

80 the flow regime: e.g., optimization approaches for reservoir operation or irrigation schemes

81 (Hughes, 2004; Kirby et al., 2014), or drought assessment (Smakhtin and Hughes, 2007). In

82 both circumstances, quantifying the uncertainty of uncalibrated models is important to produce

83 credible information for management, potentially overcoming the need for more sophisticated

84 models (Guswa et al., 2014).

4

85   Here, we used DWBM (dynamic water balance model) with a monthly time step (Zhang et al.

86   2008). The model has four parameters with physical interpretation and was shown to explain

87   flow variations for a large number of catchments in Australia (Zhang et al., 2016, 2008). After

88   describing the model and how climate influences its behavior, we examine the correlations

89   between catchment characteristics and calibrated model parameters. We examine how model

90   parameters are correlated with physical characteristics, and show that model performance for

91   dry-season flow prediction decreases sharply when catchment information is reduced. We also

92   examine predicted change in dry-season flow following a simulated land-use change, showing

93   that catchment information does not influence the general direction and magnitude of these

94   predictions. We discuss the implications of this work in Section 5, with a focus on the

95   importance of climate change relative to land-use change; we suggest that parsimonious

96   monthly models have practical utility when calibration data are available and when the main

97   objective of the study is to explore the *relative* effect of land use or climate change on seasonal

98   flow.

99

100   **2   A simple monthly water balance model for environmental change**

101   **2.1   Overview and comparison with other models**

102   The model used in this study, DWBM, is a four-parameter lumped catchment model that

103   partitions monthly precipitation into evapotranspiration and runoff (see full description in section

104   2.2). DWBM was developed by Zhang et al. (2008) with the aim to extend the Budyko theory, or

105   "limits" concept, to sub-annual timescales (Budyko, 1961; Hamel and Guswa, 2015). The model

106   also has a five-parameter version (Wang et al., 2011) but for the purpose of this study, we

107   employ the more parsimonious version, which has been verified on a subset of >200

108   catchments in Australia (Zhang et al. 2008).

5

109   DWBM is similar to a number of parsimonious lumped models, including *abcd* and G2M

110   (Mouelhi et al., 2006), which represent a catchment with one or two stores of water that

111   influence the basin-scale partitioning of precipitation into evapotranspiration and runoff. These

112   models continue to receive attention from the hydrologic community given the uncertainty

113   associated with complex models: for example, in their study of 429 catchments around the

114   world, Perrin et al. (2001) showed that models with a low number of parameters (<5) achieved a

115   performance comparable to more complex models, and recommended their use due to the ease

116   of assessing parameter uncertainty with such models. As described later, DWBM has the

117   advantage of using parameters with physical meaning, which facilitates interpretation of results

118   and inferring the effects of landscape modification. In general, we note that the selection of

119   DWBM does not impact the scope and ideas implemented in this study. Similar analyses could

120   be conducted with alternative models, and we suggest that a number of findings would hold: the

121   "equifinality of model structures", as defined by Perrin et al. (2001), suggest that most

122   parsimonious models would yield similar results.

123   **2.2   Model description**

124   • Model equations

125   The DWBM model operates with two stores of water for a catchment – the vadose zone and

126   groundwater. Monthly precipitation is partitioned among direct runoff, evapotranspiration,

127   storage in the vadose zone, and recharge to groundwater; monthly streamflow is a combination

128   of direct runoff and baseflow supplied by the groundwater store. The following section describes

129   the main equations but the reader is referred to the full description of model development for

130   additional details (Zhang et al. 2008).

131   For each month, the model first partitions precipitation into catchment wetting and direct runoff.

132   Catchment wetting, X, for a month, m, is bounded by both a supply limit (P_m, the precipitation

133 arriving in that month) and a demand limit, X_0. Mathematically, this "limit" concept is captured by

134 a bi-asymptotic function (Figure 1), and catchment wetting is computed as:

$$X(m) = P(m)F\left(\frac{X_0}{P(m)}, \alpha_1\right)$$

135   [1]

136 Where F is the bi-asymptotic function, defined as:

$$F(x, \alpha) = 1 + x - \left(1 + x^{\frac{1}{1-\alpha}}\right)^{1-\alpha},$$

138   [2]

139 α_1 is the retention efficiency, which determines how close X is to the supply and demand limits;

140 the "demand limit" X_0 is calculated as the sum of available storage capacity and

141 evapotranspiration demand, (here called potential evapotranspiration, PET):

$$X_0(m) = S_{max} - S(m-1) + PET(m)$$

142   [3]

143 where S_{max} is the maximum catchment storage capacity and S the catchment storage value at a

144 given time step.

145 [FIGURE 1]

146

147 For each month, X is used to compute an intermediate variable, available water, W:

$$W(m) = X(m) + S(m-1)$$

148   [4]

7

149   as well as the direct flow, Q_d, i.e. water not retained in the catchment that quickly becomes

150   streamflow:

$$Q_d(m) = P(m) - X(m)$$

151   [5]

152   The available water, W, is partitioned among evapotranspiration, storage and recharge. To do

153   so, the model computes the evapotranspiration opportunity, Y, i.e. the proportion of available

154   water that does not percolate below the root zone and become recharge. The supply limit for Y

155   is the available water, while the demand limit is the sum of potential evapotranspiration and

156   storage; therefore

$$Y(m) = W(m) \times F(\frac{PET(m) + S_{max}}{W(m)}, \alpha_2)$$

157   [6]

158   where α_2 is the evapotranspiration efficiency, which determines how close Y is to the supply and

159   demand limits (Figure 1).

160   Monthly evapotranspiration, ET, is bounded by the available water and energy demand (PET). It

161   is assumed that ET follows the same function as Y, i.e. that the evapotranspiration efficiency α_2

162   also determines how close ET is to the evapotranspiration demand:

$$ET(m) = W(m) \times F(\frac{PET(m)}{W(m)}, \alpha_2)$$

163   [7]

164   Recharge can then be calculated as the difference between available water and

165   evapotranspiration opportunity:

8

$$R(m) = W(m) - Y(m)$$

166 [8]

167 and storage is the difference between evapotranspiration opportunity and actual

168 evapotranspiration:

$$S(m) = Y(m) - ET(m)$$

169 [9]

170 Finally, monthly baseflow is calculated as:

$$Q_b(m) = d \cdot G(m-1)$$

171 [10]

172 where d is the groundwater store time constant, characterizing the groundwater drainage rate,

173 and G is groundwater storage, updated monthly as:

$$G(m) = G(m-1) - Q_b(m) + R(m)$$

174 [11]

175 Total streamflow is calculated as the sum of direct flow and baseflow.

176

177 • Interpretation in terms of environmental change

178 Given our focus on environmental change, we elaborate here on how climate and land-use

179 changes can be represented by the model. Table 1 summarizes the expected relation between

180 the four parameters and physical catchment characteristics. We suggest that changes in land-

181 use and land-cover will likely affect S_{max}, α_1 and α_2: changes to root depth and soil properties

9

182 may alter the partitioning between direct runoff and soil storage, along with the partitioning of

183 soil water between groundwater recharge and evapotranspiration. The parameter d affects only

184 the monthly timing of baseflow, and we suggest that d is primarily a function of geology and not

185 significantly influenced by land use or climate changes. (At the daily time scale, the dynamic

186 storage theory suggests that it also depends on antecedent conditions, i.e., on land use and

187 climate features, cf. Kirchner, 2009).

188 Seasonal changes in precipitation and potential evapotranspiration will be captured by the

189 climate forcing variables. Changes in the intensity of individual precipitation events, a

190 characteristic not described by the monthly total, will likely affect α_1, since higher intensity events

191 may result in more direct runoff. Indirect climate change effects may also affect soil and

192 vegetation properties, suggesting that α_2 and potentially S_{max} may be affected by climate change

193 (Table 1).

194 [TABLE 1]

195

196   **3**   **Methods**

197   **3.1**   **Overview**

198   Our aim is to quantify the uncertainty in minimum flow predictions across a wide range of

199   climate and catchment characteristics, and to understand how this uncertainty evolves as

200   catchment information is introduced. Our analyses rely on two metrics, minimum monthly flow

201   (Q_{min}) and total flow (Q_{tot}), computed as the minimum average monthly flow and average annual

202   flow, respectively, across the period of record. Here, minimum monthly flow is used to represent

203   dry-season flow, thereby using a flow-based definition of the dry season.

204   We first conduct a brief sensitivity analysis to illustrate the model response to climate forcing.

205   Building on previous work in Australia (Zhang et al. 2008), we compare observed minimum

206   monthly flows for 89 catchments to predictions from four versions of DWBM: one with

207   parameters obtained from calibration, two where parameters are determined via regression on

208   catchment characteristics, and one with no variation in model parameters among catchments

209   (i.e., the only variation in models among the basins is the climate forcing). We then use the

210   DWBM to predict low flows in 641 catchments in the United States. To assess the universality of

211   the regression models developed for the Australian catchments, we employ the same

212   regression models to determine model parameters for the US basins. We also evaluate the

213   performance of the DWBM with fixed parameters across the US catchments and with an

214   independent calibration. Finally, we explore the use of DWBM to assess the potential effect of

215   land-use change on dry-season flows for ungauged basins. In doing so, we evaluate whether

216   the model can predict land-use change effects in relative terms, even if the absolute magnitude

217   of minimum flows is not well predicted.

218

3.2 Sensitivity analysis: relative importance of catchment characteristics on annual and dry-season flow

221 To demonstrate model behavior, we present the sensitivity of our two variables of interest,

222 minimum monthly flow and total flow, to both climate forcing and model parameters (which are

223 proxies for catchment characteristics). We present three distinct climates, subtropical-dry

224 summer, tropical-dry winter, and humid continental. Details of the analyses and in-depth

225 discussion of the hydrological processes driving the results are presented in Appendix 1.

3.3 Parameter selection and model performance (Australian dataset)

227 Given the physical interpretation of DWBM parameters (Section 2.2), we expect their values to

228 be correlated with measurable characteristics of a catchment. We tested this hypothesis on a

229 dataset of 89 catchments in Australia for which the DWBM model was calibrated using four

230 objective functions related to low flows, high flows, time shift, and total mass balance (Zhang et

231 al. 2008). Catchment areas vary between 50 and 2000 km^2 and are located across a large

232 range of climate zones (Figure 2a). We examined twelve relevant and readily available

233 catchment characteristics as explanatory variables for the regression, including information on

234 climate, soil, topography, and land use (Table 2). Data sources for catchment streamflow time

235 series and characteristics are described by Shao et al. (2012). Each catchment had at least 10

236 years of climate and streamflow data, which we used to run the model and obtain a time series

237 of monthly modeled streamflow. For both observed and modeled time series, we computed the

238 average monthly flows and extracted the minimum and total annual flow to obtain the values of

239 Q_{min} and Q_{tot} for each catchment. After conducting a simple backward stepwise linear regression

240 model that had low predictive power (see Table 1), we developed two regression approaches

241 described below.

242

243   [FIGURE 2]

244   [TABLE 2]

245

246   • Regression with the full set of variables (regression trees)

247   We built regression trees to explore how much variability in parameter values could be

248   explained by the complete set of catchment characteristics given in Table 2. Regression trees

249   were selected for their high explanatory power, when compared with multiple linear regressions

250   and a multivariate adaptive regression spline (MARS) model (Shao et al., 2012). The analyses

251   were performed with the 'rpart'[1] package in the R environment. We tested simple and pruned

252   trees and finally selected a random forest method, using the 'randomForest'[2] package in R,

253   which gave the best performance. This method consists in creating thousands of unique

254   regression trees for the same dataset, using a random sampling of variables to create each tree

255   (Breiman, 2001). Each of these trees is used to predict the dependent variable, and the mean

256   prediction from the entire forest is the output. After "growing" a forest for each parameter, we

257   perform a 'leave-one-out' cross-validation, i.e. building a random forest using every observation

258   (the parameter values) except one, then using the model to predict the observation that was left

259   out. The process is repeated until the model has predicted every observation in the dataset,

260   after which the average prediction error is calculated.

261   • Multiple linear regression on a reduced set of variables

262   To assess the model performance in a situation with reduced data availability, we test a simple

263   linear regression model that relies on direct physical interpretation of parameters. Specifically,

264   we tested the correlation between each parameter and the catchment characteristics

---

[1] https://cran.r-project.org/web/packages/rpart/index.html
[2] https://cran.r-project.org/web/packages/randomForest/index.html

13

265 considered as the best proxies for the parameter. The following paragraphs explain the rationale

266 behind the selection of catchment characteristics for this simplified approach.

267 α_1 , the retention capacity, is closely related to the curve number (CN) used in the SCS-method

268 (NRCS-USDA, 2004). This empirical value captures the ability of a catchment to retain water in

269 the soil layer instead of producing direct runoff. Therefore, we tested the correlation between α_1

270 and CN values for each catchment. CN values were calculated as the weighted average of CN

271 for forest and grass land covers. Soil hydrologic groups were estimated from the HiHydroSoil

272 dataset (Boer, 2015)

273 α_2 is related to soil drainage and rain event frequency. Therefore, we used the subsoil hydraulic

274 conductivity and average storm depth as explanatory variables for α_2.

275 S_{max} is related to the product of soil depth and saturated water content. Because the soil dataset

276 we used did not show any variability in soil depths (all depths>2400mm), we only used

277 saturated water content in the regression.

278 d is related to hydraulic conductivity of deep layers. We used the subsoil hydraulic conductivity

279 as the only explanatory variable.

280 • Mean parameters

281 We also tested a case for which no catchment-specific information is used to estimate the

282 parameters. For this, we used the mean values of the calibrated parameters across all

283 Australian catchments. For these analyses, only climate forcing varies among the models from

284 one catchment to the next.

285 • Model performance

14

286   We ran the DWBM model three times for each Australian catchment with the parameter sets

287   described above, i.e. determined by the full regression model, the reduced regression model,

288   and the mean value. We compared the minimum flow and total flow predicted with each

289   parameterization, including the parameter set obtained by calibration, with the minimum flow

290   and total flow obtained from observed time series.

291   **3.4   Model verification (US dataset)**

292   We tested the performance of the modified DWBM, i.e. applied with the regressed set of

293   parameters, outside Australia. To compare the model performance when calibration data are

294   available, we also calibrated the model for the verification dataset. For this calibration, we used

295   a single objective function, the Nash-Sutcliffe efficiency for log-transformed flow, consistent with

296   our focus on low flows.

297   Our dataset of US catchments was developed by Newman et al. (2015), comprising 671

298   catchments (although we discarded 30 catchments for quality assurance reasons, see Appendix

299   B). Similar to the Australian dataset, the catchments range in size (1 to 25,800 km^2) and

300   hydroclimatic conditions (Figure 2b). To run DWBM on the US dataset, we summed

301   precipitation data at the monthly time step and computed monthly potential evapotranspiration

302   from monthly temperature data, using the modified Hargreaves method (Eq. 5 from Droogers et

303   al., 2002). Q_{min} and Q_{tot} and model performance metrics for the US dataset were calculated with

304   the method described above for the Australian dataset, i.e. we compared the Q_{min} and Q_{tot}

305   predictions based on the three alternative parameterizations with observations. To further

306   explore the variability in model performance, we grouped results by region, according to the

307   USGS HUC 02 classification.

308   **3.5   Variation of model performance with catchment characteristics**

309   We examined the correlation between model errors and catchment characteristics to identify the

310   conditions under which the model performs best. Specifically, we computed r^2 and p-values

311   between errors in Q_{min} and Q_{tot} obtained from each model parameterization, on one hand, and

312   all catchment characteristics listed in Table 2, on the other hand.

313   **3.6   Simulated effect of land-use change in ungauged basins**

314   The parameters for the DWBM each incorporate the effects of a host of climate, landscape, and

315   geologic factors, some of which are measurable and others which are not. Thus, detecting a

316   land-use signal in the parameters when moving from one catchment to another may be

317   challenging, as the effects of land-use alone may be lost amid the noise and other differences

318   between the catchments. Nonetheless, we were interested in assessing model predictions of

319   land-use change, in relative terms, within a particular catchment.

320   As noted in Table 1, land-use change presumably affects α_1 and α_2. Over the longer term, land-

321   use change may affect soil properties (i.e. S_{max}), but this effect is arguably weaker and ignored

322   in these analyses. It is possible that the flow response to a change in α_1 and α_2, representing

323   land-use change, may be a function of their original values. To test this hypothesis, we

324   investigated the effect of simultaneous 10% and 20% changes in α_1 and α_2 for each Australian

325   catchment, for both the calibrated dataset (for which the α_1 and α_2 parameters vary among the

326   basins) and the mean-value dataset (which all share the same parameter values). If the

327   changes in Q_{min} that result from changes in α_1 and α_2 are comparable between the two models

328   (calibrated and mean value), we can conclude that the effects of afforestation/deforestation on

329   minimum flows are independent of the original parameter values. Thus, in an ungauged basin

330   for which little information is available, the mean-value model could be used to predict the

331   effects of land-use change. The values of relative change (10 and 20%) were based on the

16

332 maximum change in parameter values predicted by the random forest model: α_1 and α_2

333 increased by a maximum of 6% and 13%, respectively, when forest cover was increased by

334 66% (for catchments with a cover <34%).

335

4   Results

4.1   Sensitivity analysis

338 In general, the model shows greater sensitivity to parameters for the subtropical and tropical

339 climates (Figure 3). In the humid climate, catchment properties have a lower impact on

340 minimum flows, since evapotranspiration is primarily energy-limited and changes in catchment

341 water storage have little effect on hydrologic partitioning. In subtropical dry-summer and dry-

342 winter climates, a small decrease in α_1 or α_2 may lead to a sharp relative increase in Q_{min}, due to

343 increases in the small amounts of surface runoff during dry months. Conversely, as α_1 or α_2

344 increase, Q_{min} generally decreases as water retained in the soil store is more likely to be

345 evapotranspired.

346 Based on the above analyses, predictions of minimum flows will be minimally impacted by

347 changes in parameter values when: climate is humid with low seasonality in precipitation, i.e.

348 variability in evaporative demand is the main driver of minimum flows; and when catchment

349 properties correspond to "insensitive" ranges for model parameters. For example, Figure 3

350 shows that minimum flows are not sensitive to low values of α_1 for the tropical dry-winter

351 climate. In such climate, minimum flows in catchments with low retention capacity (e.g. with

352 clayey or compacted soils) are unlikely to be affected by land use change.

353

354 [FIGURE 3]

17

355

356

4.2   Regression models for DWBM's parameters

358   The results from the random forest model are summarized in Table 3, showing that r^2 was high

359   for all parameters. The mean predictive errors obtained with the random forest method for α_1,

360   α_2, S_{max}, and d are reported in Table 3 and represent 45%, 39%, 40%, and 42% of their

361   respective mean value (Table 3). To gauge the impact of errors of this magnitude on model

362   outputs, we plotted these error ranges on the sensitivity analyses graphs (Figure 3): the effect of

363   parameter errors was relatively low for Q_{tot}, but for the semi-arid and tropical climates, errors in

364   α_2 and d may affect Q_{min} significantly (>50% error). In addition, the reduced regression model

365   showed much less explanatory power (Table 3): the reduced set of variables explaining less

366   than 20% of the variance in the calibrated parameter set.

367   Of the thirteen variables in Table 2, the curve number CN ranked as the most important variable

368   for α_1 and forest cover as the fourth most important variable. Here, importance is computed as

369   follows (see details in footnote 2 above): the mean square error is calculated on the out-of-bag

370   portion of the dataset, and again on the dataset with permuted variable; then, the average

371   difference in mean square error over all trees is computed, and normalized by the standard

372   deviation of the differences. For α_2, the four most important variables were all climate-related

373   (Peomonths, Aridity, Precipitation,CVP).

374

375   [TABLE 3]

376

377   **4.3   Model performance (Australian dataset)**

378   Figure 4 and Table 4 illustrate the model performance for Q_{min} for the Australian dataset.

379   Comparison of results from the calibrated models to the observed minimum flows yielded a root-

380   mean-square error (RMSE) of 2.37 mm/mo (i.e. about 50% of the average minimum flow, 4.63

381   mm/mo). The model with parameters obtained from the full regression ("full regression model"

382   hereafter) yielded good results for Q_{min}, with a RMSE of 2.32 mm/mo (Table 4) – similar to the

383   performance of the calibrated models. Model performance was lower when using the reduced

384   regression or the mean values for parameters, although these models still explained a large

385   proportion of the variance in Q_{min} (>53%).

386   The four models predicted total flows well, with r^2 values >0.87 (Table 4). The lowest RMSE for

387   total flows was obtained by the model with calibrated values (RMSE=42.4 mm/year), and the

388   highest was obtained by the model with mean values.

389

390   [FIGURE 4]

391   [TABLE 4]

392

393   **4.4   Model verification (US dataset)**

394   With the parameters obtained from the full regression, the performance of DWBM for minimum

395   flows was lower in the US (Table 4). The model explained 92% of the variance in Q_{tot}, but only

396   between 51 and 55% of the variance in Q_{min} (Table 4, Figure 5). As information was introduced

397   by the full regression and reduced regression, there was no improvement in model performance

398   over the mean-value model (RMSE in Q_{min} for the simplified regression was lower than the full

399   regression but the difference was not statistically significant, based on a Kolmogorov-Smirnov

19

400   test). However, calibration of the models based on log-transformed Nash-Sutcliffe efficiency

401   resulted in much higher performance –with 88% of the variance in Q_{min} explained, similar to the

402   Australian dataset. The calibrated value ranges were slightly broader than those of the

403   Australian dataset, [0.36;0.99], [0.16; 0.94], [0.10; 1], and [32; 500], respectively, for α_1, α_2,

404   Smax, and d (Australian ranges are reported in Table 3).

405

406   [FIGURE 5]

407

408   **4.5   Correlation between errors in Q_{min} and catchment characteristics**

409   When using the calibrated parameters for the Australian catchments, we found significant

410   correlations (p<0.01) between the relative error in minimum flow and three catchment

411   characteristics: aridity, precipitation, and PAWHC (all negative correlations). Errors in total flow

412   also showed strong correlations with catchment characteristics, in particular with climate

413   variables, and soil properties.

414   When using model predictions from the full regression model, errors in minimum flow showed

415   significant correlation only with the aridity index, and errors in total flow with precipitation and

416   the aridity index. No correlation was found for any catchment characteristics for predictions

417   obtained with the reduced regression or mean models.

418   We found no significant correlation between catchment characteristics and relative errors in

419   minimum flows for the US catchments, for any parameter set. However, relative errors in total

420   flows were correlated with a number of catchment characteristics (all variables in Table 2 except

421   CN and the relief ratio), and with two parameters (positive correlation, for both α_2, and d).

422   To explore the regional variation in model performance in the US, we separated the catchments

423   by region, using the USGS HUC 2 classification (Figure 2). Across these more homogenous

424   units, the calibrated model performance varied without significant pattern. However, the

425   improvement upon regression and mean-value models is more appreciable for HUCs with

426   higher values, which comprise more arid regions, a finding that seems consistent with the higher

427   performance of the model in arid catchments in Australia.

428

429   **4.6   Simulated effect of land-use change in ungauged basins**

430   Figure 6a represents the relative change in Q_{min} following 10% and 20% changes (both positive

431   and negative) in α_1 and α_2. All values in the bottom-left quadrant represent an increase in α_1 and

432   α_2, while all values in the top-right quadrant represent a decrease in the two parameters. The

433   direction of the change is consistent between the calibrated and mean datasets. The difference

434   in Q_{min} predicted by the models was small for the 10% change in parameter (RMSE of 0.36). For

435   the 20% change, the high RMSE (1.2) was largely driven by the negative change in parameter

436   values (i.e. circles in the top-right quadrant in Figure 6a). Of note, these high relative changes

437   correspond to low absolute changes: the RMSE for the absolute change in Q_{min} resulting from a

438   20% change (both positive and negative) in parameters is 1.5mm.

439   The results for total flow (Figure 6b) showed even smaller differences between the two models,

440   indicating that medium to high flows were less affected, in relative terms, by the change in

441   parameter values. RMSE were 0.15 and 0.37, respectively, for the 10% and 20% change in

442   parameter values.

443

444   [FIGURE 6]

21

445

5   Discussion and implications for predicting the effects of environmental change

448   The main objective of this paper is to assess the utility of a monthly, lumped hydrologic model

449   for predicting dry-season flows with varying degrees of information availability. As a rainfall-

450   runoff model governed by four parameters with physical meaning, DWBM has the potential to

451   be used for climate and land-use change scenarios analyses and inform landscape

452   management. The sensitivity analyses indicated that the importance of each parameter

453   depends on climate: for example, a larger storage capacity S_{max} will generally be needed in

454   highly seasonal climate to sustain baseflow during the dry season. The moderate sensitivity in a

455   number of environmental contexts (i.e. parameter sets) suggests that climate is the main driver

456   of seasonal flow, a fact that has been observed by many others (Devito et al., 2005; Jencso and

457   McGlynn, 2011). In practice, this means that a rough estimate of these parameters may be

458   sufficient to predict monthly flows with acceptable levels of certainty, as suggested by our

459   analyses on Australian and US catchments.

5.1   Model performance for absolute predictions of dry-season flow

461   The model performance, measured by r^2, in Australia was relatively high for both Q_{min} and Q_{tot}

462   predictions (Figure 4). RMSE for Q_{min} ranged from 2.4 mm/mo to 4.0 mm/mo, depending on the

463   model parameterization. Adding catchment information, i.e. moving from uniform parameters for

464   all catchments, to regressed parameter values, to streamflow time series for calibration,

465   generally improved model performance (measured by RMSE). The performance of the full

466   regression model was actually as good as the calibrated model, probably due to the large

467   number of explanatory variables and the high explanatory power of the regression (Table 3).

468   The poorer performance of the reduced regression could be due to two factors: poor selection of

22

469 model variables and over-fitting of the random forest full regression model. The stepwise

470 backward regression conducted in preliminary analyses confirmed that our selected variables

471 are among the best predictors, but that no single variable explained the variance significantly.

472 This suggests that the full regression model was probably over-fitted (12 variables for 89

473 observations), and therefore less likely to transfer outside the initial sample, for the US

474 catchments.

475 For the verification dataset with US catchments, the performance of DWBM with parameters

476 derived from the full regression was much lower: only 55% of the variance in Q_{min} was

477 explained. This suggests that although model parameters were strongly correlated with physical

478 characteristics in Australia, these relationships did not transfer to the US dataset. A number of

479 reasons could explain this negative result, in particular the consideration of snowmelt (see

480 below), and extrapolations of the regression outside the range of Australian values for a number

481 of physical characteristics (in particular soil variables and, to a lesser extent, climate variables).

482 However, after calibration, DWBM's performance was good for both total and minimum flows,

483 confirming the possibility to use the model with regional parameterization.

484 Because errors in Q_{min} were only weakly correlated with catchment characteristics, it is difficult

485 to predict where the model will perform best outside the set of catchments in Australia or the

486 US. However, the model seemed to perform better when the aridity index was lower (i.e. drier,

487 and thus minimum flows were lower), likely reflecting water-balance constraints in a water-

488 limited environment. Additionally, it is likely that snowmelt effects, ignored in this work,

489 contribute to errors in minimum flows. To test this hypothesis, we evaluated model performance

490 for the 97 US basins that were not influenced by snow precipitation (Guswa et al., 2017), and

491 found that r^2 increased to 0.70 (from 0.55) and 0.68 (from 0.51), respectively, for the "full

492 regression" and mean-value parameterizations (RMSE were 4.4 mm and 4.5 mm, respectively).

493 These results confirm that the relationships did not transfer to the US dataset. We also

23

494   hypothesized that performance would be higher where high values of α_1 and α_2 are predicted by

495   the regression, based on the sensitivity analyses, although the US dataset did not confirm this

496   hypothesis.

497   We conclude this section with methodological points that help interpret model performance, both

498   for absolute values or theoretical land-use change. First, we note that many catchments in our

499   datasets had low observed minimum flow (<3mm/mo), especially for the Australian dataset

500   dominated by the "humid temperate warm summer" climate zone. In absolute values, these

501   errors remain small as illustrated by Figure 4. In addition, the datasets included only "natural"

502   catchments, with the land use being mainly grassland or forest. This means that the effect of

503   different land uses is likely difficult to detect in these datasets, as suggested by the regressions

504   on catchment characteristics (forest cover was not significantly correlated with α_1 or α_2). This

505   could also explain the poor performance of the reduced regression model: variations in α_1 and

506   α_2 based on the simple regression models were small (for example, CN values only varied from

507   70 to 80, a narrower range compared to possible land use changes involving agricultural land).

508   We also note that further analyses could improve model performance in both regions. First, the

509   model calibration could be focused on low flows. The calibration for the Australian dataset was

510   performed using a combination of four objective functions, with only one focused on low flows

511   (Zhang et al., 2008). Second, the parameter values could be corrected for the bias in Q_{min} for

512   the US dataset. This bias may be related to the calibration function, but our analyses do not

513   provide evidence of this.

514   **5.2   Predicting the effect of environmental change for ungauged catchments**

515   • Climate change

516   Both the US and Australian dataset comprise catchments that range in climate, geology, and

517   land use. The fair performance of the model in both locations suggests that the model is able to

24

518   represent the variability of hydrological behavior induced by these factors. This gives confidence

519   that the effect of future climate forcing would be correctly represented: because the model uses

520   monthly climate time series as forcing variables, such analysis can be performed by substituting

521   current climate time series with future forecasts. Given the highest performance of the calibrated

522   model, climate change analyses are best performed with gauged catchments (calibrating the

523   model). However, they may be conducted on ungauged catchments too when information on a

524   relative change, rather than absolute, is sought. For example, Monte-Carlo-type analysis can be

525   performed by assuming parameter sets for the catchment of interest, and then running the

526   model for each set to provide upper and lower bounds of the expected change following climate

527   change.

528   • Land use change

529   The high performance of the full regression model in Australia was not found in the US.

530   Therefore, using model regression to infer parameter values is not a feasible option for

531   ungauged catchments globally. For the Australian dataset, CN and forest cover were found to

532   be important variables in the full regression on α_1 (Section 4.2), confirming the relationship

533   between this parameter and land-use variables. For both regions, calibrated parameter values

534   showed low or no correlation with land-use variables, which suggest that additional work is

535   needed to derive empirical relationships between the parameter values and land use

536   characteristics. Nonetheless, the land-use change analyses in Section 4.6 suggest that one can

537   use the baseline provided by the model to compute the relative change in Q_{min} following land-

538   use change. The motivation for this simple analysis was to understand, theoretically, the effect

539   of landscape interventions on dry-season flows. For example, such information can be used to

540   assess the potential for the "sponge effect" to occur in a given climate (i.e. that afforestation

541   would increase dry-season flow): specifically, the catchments with a relative change close to

25

542   zero in Figure 6a are unlikely to demonstrate an increase in dry-season flow with afforestation,

543   since the change in α parameter values only minimally affected Q_{min}.

544   We note that the absolute change in parameter values can be constrained by the calibrated

545   parameter set, if regional data are available (e.g. the US and Australian datasets used in this

546   study). As suggested above, Monte-Carlo runs can be performed to provide confidence intervals

547   around the change in Q_{min}. Additional work on the relationships between catchment

548   characteristics and parameter is in progress with catchments that have pre- and post-

549   afforestation streamflow data (Zhang et al., 2016). Preliminary results suggest that the

550   relationships hypothesized in Table 1 hold and that regional relationships can be used to predict

551   land-use change. The results also confirm that the land-use change signal (i.e. the increase in

552   forest cover, with all other variables held constant) may be confounded by other environmental

553   factors.

554

6   Conclusion

556   We have investigated how a simple rainfall-runoff model run at the monthly time step could

557   represent and predict the influence of climate and land-use change on dry-season flow. We

558   used the DWBM model, which assumes that streamflow, in particular during the dry season, is

559   driven by four main catchment characteristics: the retention efficiency of a catchment (ability to

560   store water for future release by discharge or evapotranspiration), evapotranspiration efficiency

561   (ability to use soil water for evapotranspiration rather than discharge), total soil storage, and

562   drainage rate. Our analyses confirmed that climate is a major driver of seasonal flows and that

563   the simple model DWBM, with default values obtained from the mean of calibrated catchments,

564   could provide a reasonable estimate of monthly flows. Model performance increases

565   significantly when calibration data are available, although in this work we found that regional

26

566   relationships to infer model parameters could not transfer to other regions (the regression on

567   catchment properties obtained in Australian did not result in high performance in the US). Our

568   analyses also suggest that DWBM can be used to estimate a change in annual and minimum

569   monthly flow following environmental change. Even without calibration data, the effects of land

570   use change (e.g. reduction in retention efficiency or in evapotranspiration efficiency) can be

571   explored and quantitatively estimated. The effect of climate change can also be assessed,

572   preferably with a calibrated model if absolute values are sought. The broad range of

573   environmental conditions used in that study confirmed that the simple structure is able to

574   capture the main hydrological processes driving runoff response. The model has low data

575   requirements, and all climate data and catchment information used in this study are available

576   globally.

577

581

582   **References**

583   Andréassian, V., 2004. Waters and forests: From historical controversy to scientific debate. J.

584   Hydrol. 291, 1–27. doi:10.1016/j.jhydrol.2003.12.015

585   Boer, F. de, 2015. HiHydroSoil: A High Resolution Soil Map of Hydraulic Properties. Report 134;

586   www.futurewater.nl.

587   Brauman, K.A., Richter, B.D., Postel, S., Malsy, M., Flörke, M., 2016. Water depletion: An

27

588   improved metric for incorporating seasonal and dry-year water scarcity into water risk

589   assessments. Elem. Sci. Anthr. 4, 83. doi:10.12952/journal.elementa.000083

590   Breiman, L., 2001. Random Forests, in: Machine Learning. pp. 5–32.

591   doi:10.1023/A:1010933404324

592   Brown, A.E., Western, A.W., McMahon, T.A., Zhang, L., 2013. Impact of forest cover changes

593   on annual streamflow and flow duration curves. J. Hydrol. 483, 39–50.

594   doi:10.1016/j.jhydrol.2012.12.031

595   Brown, A.E., Zhang, L., McMahon, T. a., Western, A.W., Vertessy, R. a., 2005. A review of

596   paired catchment studies for determining changes in water yield resulting from alterations

597   in vegetation. J. Hydrol. 310, 28–61. doi:10.1016/j.jhydrol.2004.12.010

598   Brutsaert, W., 2008. Long-term groundwater storage trends estimated from streamflow records:

599   Climatic perspective. Water Resour. Res. 44, n/a-n/a. doi:10.1029/2007WR006518

600   Budyko, M., 1961. The Heat Balance of the Earth's Surface. Washington D.C.

601   Devito, K., Creed, I., Gan, T., Mendoza, C., Petrone, R., Silins, U., Smerdon, B., 2005. A

602   framework for broad-scale classification of hydrologic response units on the Boreal Plain :

603   is topography the last thing to consider ? Hydrol. Process. 19, 1705–1714.

604   doi:10.1002/hyp.5881

605   Gao, Z., Zhang, L., Cheng, L., Zhang, X., Cowan, T., Cai, W., Brutsaert, W., 2015. Groundwater

606   storage trends in the Loess Plateau of China estimated from streamflow records. J. Hydrol.

607   530, 281–290. doi:10.1016/j.jhydrol.2015.09.063

608   Gehne, M., Hamill, T., Kiladis, G.N., Trenberth, K.E., 2016. Comparison of Global Precipitation

609   Estimates across a Range of Temporal and Spatial Scales. J. Clim.

28

610   Guswa, A., Brauman, K.A., Brown, C., Hamel, P., Keeler, B.L., Sayre, S.S., 2014. Ecosystem

611   Services: Challenges and Opportunities for Hydrologic Modeling to Support Decision

612   Making. Water Resour. Res. 50, 4535–4544. doi:10.1002/2014WR015497

613   Guswa, A.J., Hamel, P., Dennedy-Frank, P.J., 2017. Potential effects of landscape change on

614   water supplies in the presence of reservoir storage. Water Resour. Res.

615   Hamel, P., Guswa, A.J., 2015. Uncertainty analysis of a spatially-explicit annual water-balance

616   model: case study of the Cape Fear catchment, NC. Hydrol. Earth Syst. Sci. 19, 839–853.

617   doi:10.5194/hess-19-839-2015

618   Hughes, D.A., 2004. Incorporating groundwater recharge and discharge functions into an

619   existing monthly rainfall–runoff model/Incorporation de fonctions de recharge et de vidange

620   superficielle de nappes au sein d'un modèle pluie-débit mensuel existant. Hydrol. Sci. J.

621   49. doi:10.1623/hysj.49.2.297.34834

622   Jencso, K.G., McGlynn, B.L., 2011. Hierarchical controls on runoff generation: Topographically

623   driven hydrologic connectivity, geology, and vegetation. Water Resour. Res. 47, n/a-n/a.

624   doi:10.1029/2011WR010666

625   Kirby, J.M., Connor, J., Ahmad, M.D., Gao, L., Mainuddin, M., 2014. Climate change and

626   environmental water reallocation in the Murray–Darling Basin: Impacts on flows, diversions

627   and economic returns to irrigation. J. Hydrol. 518, 120–129.

628   doi:10.1016/j.jhydrol.2014.01.024

629   Kirchner, J.W., 2009. Catchments as simple dynamical systems: Catchment characterization,

630   rainfall-runoff modeling, and doing hydrology backward. Water Resour. Res. 45, W02429.

631   McIntyre, N., Ballard, C., Bruen, M., Bulygina, N., Buytaert, W., Cluckie, I., Dunn, S., Ehret, U.,

632   Ewen, J., Gelfan, A., Hess, T., Hughes, D., Jackson, B., Kjeldsen, T., Merz, R., Park, J.,

29

633   Connell, E.O., Donnell, G.O., Oudin, L., Todini, E., Wagener, T., Wheater, H., Mcintyre, N.,

634   2014. Modelling the hydrological impacts of rural land use change. Hydrol. Res. 45, 737–

635   754. doi:10.2166/nh.2013.145

636   Mouelhi, S., Michel, C., Perrin, C., Andréassian, V., 2006. Stepwise development of a two-

637   parameter monthly water balance model. J. Hydrol. 318, 200–214.

638   doi:10.1016/j.jhydrol.2005.06.014

639   Newman, A.J., Clark, M.P., Sampson, K., Wood, A., Hay, L.E., Bock, A., Viger, R.J., Blodgett,

640   D., Brekke, L., Arnold, J.R., Hopson, T., Duan, Q., 2015. Development of a large-sample

641   watershed-scale hydrometeorological data set for the contiguous USA: data set

642   characteristics and assessment of regional variability in hydrologic model performance.

643   Hydrol. Earth Syst. Sci. 19, 209–223. doi:10.5194/hess-19-209-2015

644   NRCS-USDA, 2004. Chapter 10. Estimation of Direct Runoff from Storm Rainfall, in: United

645   States Department of Agriculture (Ed.), Part 630 Hydrology. National Engineering

646   Handbook. United States Department of Agriculture,

647   http://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/water/?cid=stelprdb1043063.

648   Perrin, C., Michel, C., Andréassian, V., 2001. Does a large number of parameters enhance

649   model performance? Comparative assessment of common catchment model structures on

650   429 catchments. J. Hydrol. 242, 275–301. doi:http://doi.org/10.1016/S0022-

651   1694(00)00393-0

652   Price, K., 2011. Effects of watershed topography, soils, land use, and climate on baseflow

653   hydrology in humid regions: A review. Prog. Phys. Geogr. 35, 465–492. doi:DOI:

654   10.1177/0309133311402714

655   Shao, Q., Traylen, A., Zhang, L., 2012. Nonparametric method for estimating the effects of

656   climatic and catchment characteristics on mean annual evapotranspiration. Water Resour.

657   Res. 48, 1–13. doi:10.1029/2010WR009610

658   Smakhtin, V.., Hughes, D., 2007. Automated estimation and analyses of meteorological drought

659   characteristics from monthly rainfall data. Environ. Model. Softw. 22, 880–890.

660   doi:10.1016/j.envsoft.2006.05.013

661   Smith, M.B., Seo, D.-J., Koren, V.I., Reed, S.M., Zhang, Z., Duan, Q., Moreda, F., Cong, S.,

662   2004. The distributed model intercomparison project (DMIP): motivation and experiment

663   design. J. Hydrol. 298, 4–26. doi:10.1016/j.jhydrol.2004.03.040

664   Vogl, A.L., Dennedy-Frank, P.J., Wolny, S., Johnson, J.A., Hamel, P., Narain, U., Vaidya, A.,

665   2016. Managing forest ecosystem services for hydropower production. Environ. Sci. Policy

666   61, 221–229.

667   Wang, Q.J., Pagano, T.C., Zhou, S.L., Hapuarachchi, H. a P., Zhang, L., Robertson, D.E., 2011.

668   Monthly versus daily water balance models in simulating monthly runoff. J. Hydrol. 404,

669   166–175. doi:10.1016/j.jhydrol.2011.04.027

670   Zhang, L., Brutsaert, W., Crosbie, R., Potter, N., 2014. Long-term annual groundwater storage

671   trends in Australian catchments. Adv. Water Resour. 74, 156–165.

672   doi:10.1016/j.advwatres.2014.09.001

673   Zhang, L., Hickel, K., Shao, Q., 2016. Predicting afforestation impacts on monthly streamflow

674   using the DWBM model. Ecohydrology 10. doi:10.1002/eco.1821

675   Zhang, L., Potter, N., Hickel, K., Zhang, Y., Shao, Q., 2008. Water balance modeling over

676   variable time scales based on the Budyko framework – Model development and testing. J.

677   Hydrol. 360, 117–131. doi:10.1016/j.jhydrol.2008.07.021

31

678

679

680

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Appendix

681

A. Sensitivity analysis

682

683   We performed the sensitivity analyses with actual data from three climatically distinct locations:

684   (1) Nairobi, Kenya, (2) San Jose, Costa Rica, and (3) Cleveland, USA. Under the Köppen-

685   Geiger climate classification, Nairobi has a subtropical highland climate with dry summers and

686   an annual aridity index of 0.47; San Jose has a tropical climate, with dry winter and an aridity

687   index of 1.6; Cleveland has a humid continental climate with an aridity index of 2.4. Although we

688   could have used synthetic climate series to control climate variability, our objective here is to

689   illustrate model behaviors under different climate forcing, which is achieved by actual data from

690   different climate zones.

691   • Methods

692   Monthly precipitation and temperature data were acquired for each of these locations from the

693   National Oceanic and Atmospheric Administration's (NOAA) Global Historical Climatology

694   Network-Monthly (GHCN-M) dataset. From this dataset, we computed monthly averages. The

695   precipitation averages were used directly as model input, while the temperature averages were

696   used to calculate monthly potential evapotranspiration (PET) values using the modified

697   Hargreaves method (equation (5) from Droogers, et al., 2002). The precipitation and potential

698   evapotranspiration time series for each location are shown in Figure A1.

699   For each climate type, we first performed a one-at-a-time sensitivity analysis, using five levels at

700   equal intervals for each parameter. The range for each parameter was initially based on the

701   values obtained from the model calibration by Zhang et al. (2008), described in further details in

702   Section 3, and summarized in Table A1. Initial conditions affected flows for only the first few

33

703   years: to remove this 'warm-up' effect, the model was run for 10 years, repeating the same

704   climate forcing, and only the final year was used to compute the statistics.

705   Next, to quantify interaction effects among parameters, the model was run 24 additional times

706   for each climate type, varying every possible pair of parameters with all possible combinations

707   of upper and lower bounds for each parameter.

708   After computing the regression analyses (cf. Section 3.1), we also re-ran the one-at-a-time

709   sensitivity analyses varying mean values of each parameter by the average error in the random

710   forest model: the new range (twice the average parameter error around the mean value) gives a

711   more realistic assessment of sensitivity for the Australian dataset, and is plotted on Figure 3.

712   [TABLE A1 and FIGURE A1]

713   •   Results

714   In general, for the humid continental climate (Cleveland), total flow and minimum flow (Q_{min})

715   were not very sensitive to model parameters (Figure 3). The highest change in Q_{min} was 42%,

716   obtained for the minimum value of α_2. We note that in absolute value, effects of parameter

717   change were more significant than for other climates: for example, the 42% change in Q_{min}

718   represented 23 mm/mo. Larger variations in the relative sensitivity were observed for the two

719   other climate types.

720   **α_1** In the tropical dry winter (San Jose) and subtropical dry summer (Nairobi) climates, Q_{min} was

721   sensitive to increases in α_1 (with a maximum change of 53%) due to less direct flow during and

722   slightly after each precipitation event. In Nairobi, Q_{min} was more sensitive to low values of α_1:

723   decreasing α_1 lowers the baseflow contribution to streamflow significantly by reducing the

724   amount of water that is retained in soil storage, and thus in groundwater storage.

34

725 α_2. In the tropical and subtropical climates, Q_{min} decreased as α_2 increased (-83% and -96%,

726 respectively), due to a larger portion of water being evapotranspired. Q_{min} was sensitive to lower

727 values of α_2 in Nairobi (subtropical dry-summer), since evapotranspiration demand is high when

728 flows are low.

729 S_{max}. In all three climates, Q_{min} showed little sensitivity to S_{max}. Lower values tended to increase

730 Q_{min} in Nairobi, since they increased evapotranspiration opportunity (i.e. evapotranspiration and

731 recharge) in an arid environment. Conversely, lower values tended to decrease Q_{min} in San

732 Jose (tropical dry-winter) where water availability is higher, and low soil storage increased the

733 ratio of direct runoff over recharge.

734 **d**. As expected, Q_{min} was highly sensitive to d in seasonal climates (subtropical and tropical). In

735 particular, lower values of d resulted in sharp increases in Q_{min}, since the slow groundwater

736 release sustained a high baseflow throughout the year.

737 Interaction effects showed mostly subadditive effects. Only low values of d tended to exacerbate

738 sensitivity to S_{max} or α_2, while low values of S_{max} tended to exacerbate sensitivity to α_1 or α_2.

739

740 **B. Quality assurance of Newman's dataset (2015)**

741 [TABLE B1]

742   TABLES

743   *Table 1. Parameters and physical interpretation of DWBM. + and ++ indicate the strength of the likely*

744   *effect of land use change (LUC) and climate change (CC) on the parameters. Numbers in brackets are*

745   *coefficients obtained from a stepwise backward linear regression between calibrated parameter values*

746   *and forest cover (for LUC), or number of wet days (for CC) (* indicates significance at the 0.1 level, ***

747   *significance at the 0.01 level). See text and Table 2 for a definition of these variables.*

Parameter	Description	Affected by LUC	Affected by CC
S_{max} [5; 500]mm	Maximum catchment storage capacity. *Depends on*: soil depth and available water content (measurable soil characteristics)	+ [0.71]*	+ [-309]*
α_1 [0;1]	Catchment retention; affects the partitioning of precipitation into direct runoff and water that is available in the soil-moisture store (S) for evapotranspiration and groundwater recharge. *Depends on*: the soil infiltration capacity and the rainfall intensity (more intense, less frequent events means more direct runoff).	++ [1.8e^{-4}]	++ [-0.15]
α_2 [0;1]	Evapotranspiration efficiency; affects the partitioning of soil water into storage, recharge, and actual evapotranspiration. *Depends on*: the rainfall frequency, plant water-stress response, root depth, and soil properties such as hydraulic conductivity and critical moisture content (at which actual evapotranspiration is reduced for water	++ [4.1e^{-4}]	+ [-0.30]**

36

	stress). Note that α_2 does not depend on the crop coefficient, which is already included in PET	
d [0;1] month^{-1}	Groundwater store time constant; characterizes the groundwater drainage rate, ie the release of groundwater storage to baseflow. Note that d does not affect partitioning, only timing of baseflow. *Depends on*: aquifer characteristics (size, hydraulic conductivity, connectivity with the stream)	n.a. n.a.

748

749   *Table 2. Catchment characteristics assessed in this study*

	Abbreviation	Name	Description
Climate	**P**	Precipitation (mm/month)	Monthly rainfall
	Aridity	Aridity (-)	Precipitation divided by Potential Evapotranspiration
	Peomonths	Difference in peak potential evapotranspiration and precipitation (month)	The number of months that peak precipitation follows peak potential evapotranspiration
	ASD	Average Storm Depth (mm/day)	Depth of the average storm
	WetDays	Proportion of Wet Days (-)	Proportion of days per year with some precipitation
	CVP	Coefficient of Variation of Precipitation (-)	The standard deviation (of interannual precipitation) divided by the average

37

			annual rainfall
Vegetation	**forest**	Forest Cover (%)	Percent of catchment area covered by forest
	CN	Curve Number	Curve Number from the SCS method (also a function of soil group)
Topography	**Rr**	Relief Ratio (m/m)	Total catchment relief divided by the longest flow path
Soil	**soil_ksat_top**	Hydraulic Conductivity for topsoil (m/s)	Saturated hydraulic conductivity from 0-30cm depth
	soil_ksat_sub	Hydraulic Conductivity, subsoil (m/s)	Saturated hydraulic conductivity from 30-200cm depth
	soil_sat_wc	Saturated Water Content (m^3/m^3)	Maximum fraction of soil volume that can be occupied by water
	PAWHC	Plant Available Water Holding Capacity (mm)	Maximum depth of soil water that is available for removal by vegetation

750

751   *Table 3. Calibrated parameter values and strength of the correlation (r^2 and mean absolute error, MAE)*

752   *between calibrated parameters and their estimates from the full regression and reduced regression*

753   *methods, for the Australian dataset.*

	α_1	α_2	S_{max} (mm)	d (month^{-1})
Mean and range (calibrated	0.62	0.74	258	0.66

38

values)		[0.40; 0.74]	[0.42; 0.80]	[36.5; 500]	[0.10; 1]
Full regression	r^2	0.95	0.96	0.95	0.96
	MAE	0.017	0.022	34	0.10
Reduced regression	r^2	0.20	0.09	0.17	0.06
	MAE	0.038	0.057	86	0.24

754

755   *Table 4. Performance statistics, r^2 and root mean square error (RMSE), for the four model*

756   *parameterizations for the Australian and US catchments*

	Australia				US			
	Q_{tot}		Q_{min}		Q_{tot}		Q_{min}	
	r^2	RMSE	r^2	RMSE	r^2	RMSE	r^2	RMSE
Calibration	0.98	42.4	0.90	2.4	0.96	108.9	0.88	9.29
Full regression	0.96	47.8	0.84	2.3	0.92	167.5	0.55	9.92
Reduced regression	0.87	88.6	0.53	3.9	0.92	159.3	0.53	9.37
Mean	0.87	89.5	0.55	4.0	0.92	173.2	0.51	9.34

757

758   *Table A1. Parameter levels used in the sensitivity analyses, corresponding to the minimum, 25th, 50th, and*

759   *75th percentiles, and maximum parameter values.*

Parameter	**Min**	**25th**	**50th**	**75th**	**Max**

Smax (mm)	133	200	266	333	399
α_1	0.30	0.45	0.60	0.75	0.90
α_2	0.30	0.45	0.60	0.75	0.90
d	0.33	0.5	0.66	0.83	1.0

760

761 *Table B1. List of basins removed from Newman's dataset* (Newman et al., 2015)*, due to issues with the*

762 *time series. Other basins showed short gaps in the time series but the effect on the long-term average*

763 *was deemed minor. R is the reported average daily runoff in*

764 *basin_annual_hydrometeorology_characteristics_daymet.txt (from Newman et al.'s dataset); P is the*

765 *reported average daily precipitation in basin_annual_hydrometeorology_characteristics_daymet.txt; <q>*

766 *is the calculated average runoff from daily discharge and basin area; <p> is the calculated average*

767 *precipitation from daily precipitation*

Basin	Issue
03 02108000 NE Cape Fear, NC	Area and elevation in basin_characteristics file do not match USGS website or information in gage information file
03 02310947 Withlacoochee River near Cumpressco, FL	Multiple, long, discontinuous gaps in the streamflow record
03 02381600 Fausett Creek near Talking Rock, GA	Calculated average runoff from daily values, <q> is >1.5*reported average daily runoff, R
05 03357350 Plum Creek near Bainbridge, IN	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
09 05062500	Calculated average runoff from daily values, <q>

40

Wild Rice River at Twin Valley, MN	is >1.5*reported average daily runoff, R
09 05087500 Middle River at Argyle, MN	Calculated average runoff from daily values, <q> is >1.5*reported average daily runoff, R
09 05120500 Wintering River near Karlsruhe, ND	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
10 06468250 James River near Kensal, ND	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
10 06441500 Bad River near Fort Pierre, SD	Multiple long gaps in streamflow record
11 07067000 Current River at Van Buren, MO	Area and elevation in basin_characteristics file do not match USGS website or information in gage information file
12 08079600 Brazos River at Justiceburg, TX	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
15 09484000 Sabino Creek near Tucson, AZ	Multiple extended gaps in streamflow record throughout
15 09492400 East Fork White River near Apache, AZ	Calculated average runoff from daily values, <q> is >1.5*reported average daily runoff, R
16 10166430 West Canyon Creek near Cedar Fort, UT	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
16 10172700 Vernon Creek near Vernon, UT	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
16 10172800 South Willow Creek near Grantsville, UT	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
16 10242000	Calculated runoff from daily values, <q>, is less

41

Coal Creek near Cedar City, UT	than 50% of reported average daily runoff, R
16 10249300 South Twin River nr Round Mountain, NV	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
18 10259200 Deep Creek near Palm Desert, CA	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
18 10263500 Big Rock Creek near Valyermo, CA	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
18 11253310 Cantua Creek near Cantua Creek, CA	Calculated runoff from daily values, <q>, is less than 50% of reported average daily runoff, R
17 12040500 Queets River nr Clearwater, WA	Runoff ratio is greater than 1; <q> is greater than <p>
17 12041200 Hoh River nr Forks, WA	Runoff ratio is greater than 1; <q> is greater than <p>
17 12056500 NF Skokomish River near Hoodsport, WA	Runoff ratio is greater than 1; <q> is greater than <p>
17 12147500 NF Tolt River near Carnation, WA	Runoff ratio is greater than 1; <q> is greater than <p> and R is greater than P
17 12147600 SF Tolt River near Index, WA	Runoff ratio greater than 1; R is greater than P
17 12167000 NF Stillaguamish River near Arlington, WA	Runoff ratio greater than 1; <q> is greater than <p> and R is greater than P
17 12186000 Sauk River near Darrington, WA	Runoff ratio greater than 1; <q> is greater than <p> and R is greater than P
17 14158500 McKenzie River near Clear Lake, OR	Runoff ratio greater than 1; <q> is greater than <p>

17 14400000 Brookings, OR	Runoff ratio greater than 1; $\langle q \rangle$ is greater than $\langle p \rangle$ and R is greater than P

768

769

770

771 FIGURES

772 *Figure 1. "Limits" concept used for water partitioning in DWBM. The concept is used to partition both the*

773 *precipitation (P) between wetting (W) and direct runoff, and the wetting between evapotranspiration (ET)*

774 *and storage. The α parameters determine how close the variables are from their limits (dashed lines).*

775

776 *Figure 2. Distribution of the 89 Australian (left) and 641 US (right) catchments used in this study, with*

777 *associated aridity index values. Grey-scale background on the US map delineate the HUC2 regions*

778 *(darker colors represent higher HUC number).*

779

780 *Figure 3. Sensitivity of minimum flow (Q_{min}) to each model parameter. Grey lines represent the relative*

781 *error in parameter values from the regression model (Section 3.1)*

782

783 *Figure 4. Comparison of observed minimum flow Q_{min}, with predictions from the calibrated and mean-*

784 *value parameterizations for Australian catchments. Note that the plot window excludes between 12 and*

785 *16 catchments with high values of Q_{min}. rmse=root mean square error in mm/mo*

786

787 *Figure 5. Comparison of observed minimum flow Q_{min}, with predictions from the calibrated and mean-*

788 *value parameterizations for US catchments. Note that the plot window excludes between 12 and 16*

789 *catchments with high values of Q_{min}. rmse=root mean square error in mm/mo*

790

791

44

792 *Figure 6. Predictions for Q$_{min}$ (a) and Q$_{tot}$ (b) resulting from a hypothetical land use change – i.e. a change*

793 *in α$_1$ and α$_2$ values – for both the calibrated and the mean-value models. Each point represents one*

794 *catchment under either 10% (black) or 20% (grey) relative increase or decrease in α$_1$ and α$_2$. All values in*

795 *the bottom-left quadrant represent an increase in α$_1$ and α$_2$, while all values in the top-right quadrant*

796 *represent a decrease in the two parameters. Dashed lines represent a 50% difference between the*

797 *calibrate and mean-value predictions. RMSE for Q$_{min}$ is 0.36 for the "10% change" and 1.2 for "20%*

798 *change". For the increase in α$_1$ and α$_2$ (bottom-left quadrant), representing afforestation, RMSE for "20%*

799 *change" is 0.13.*

800

801 *Figure A1. Climate types used in this study: humid continental (Cleveland), subtropical with dry-summer*

802 *(Nairobi), tropical with dry-winter (San Jose)*

803

45

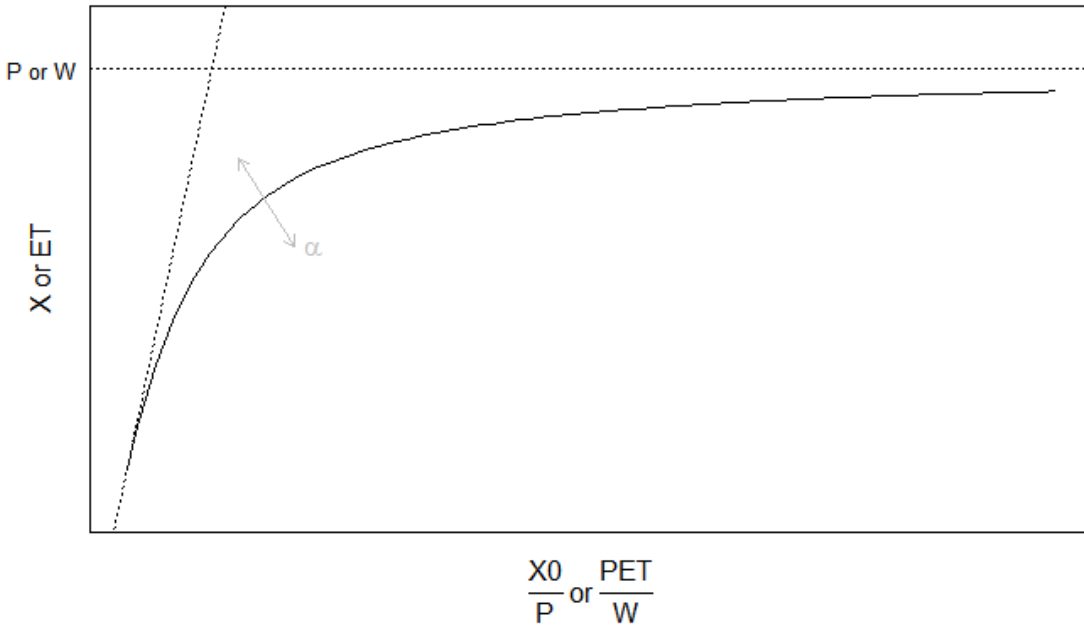*Figure 1. "Limits" concept used for water partitioning in DWBM. The concept is used to partition both the precipitation (P) between wetting (W) and direct runoff, and the wetting between evapotranspiration (ET) and storage. The α parameters determine how close the variables are from their limits (dashed lines).*
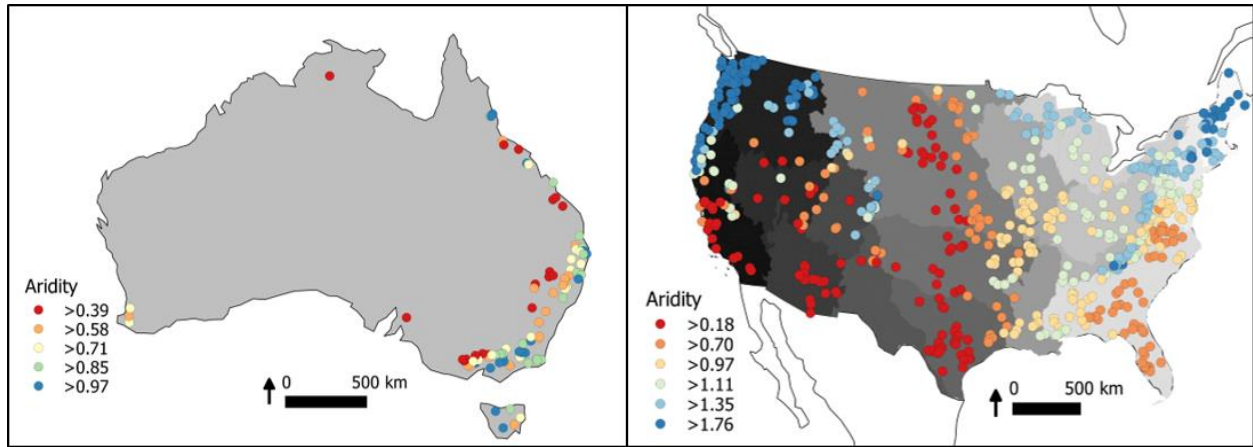
Figure 2. Distribution of the 89 Australian (left) and 641 US (right) catchments used in this study, with associated aridity index values. Grey-scale background on the US map delineate the HUC2 regions (darker colors represent higher HUC number).
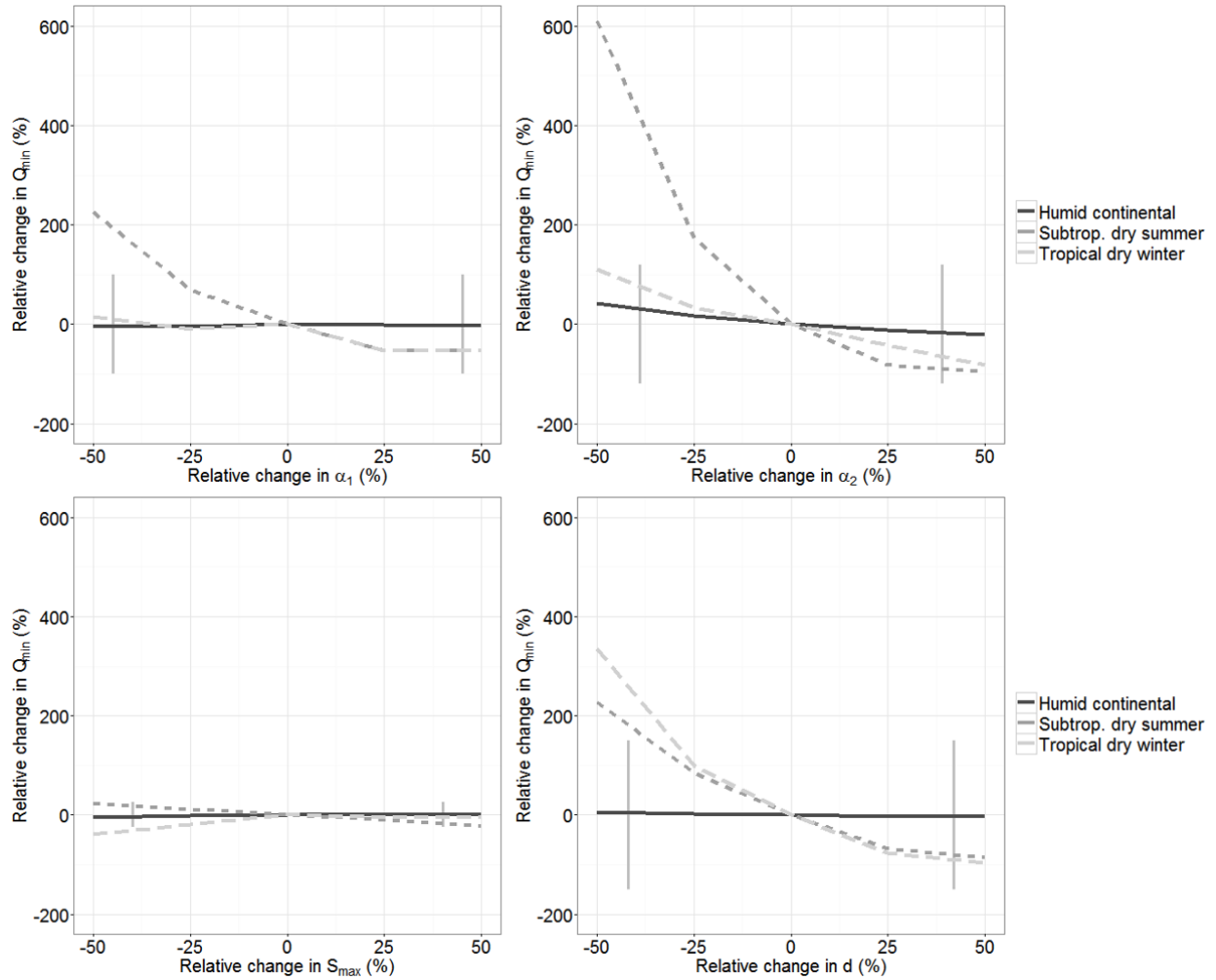
Figure 3. Sensitivity of minimum flow (Q_{min}) to each model parameter. Grey lines represent the relative error in parameter values from the regression model (Section 3.1)
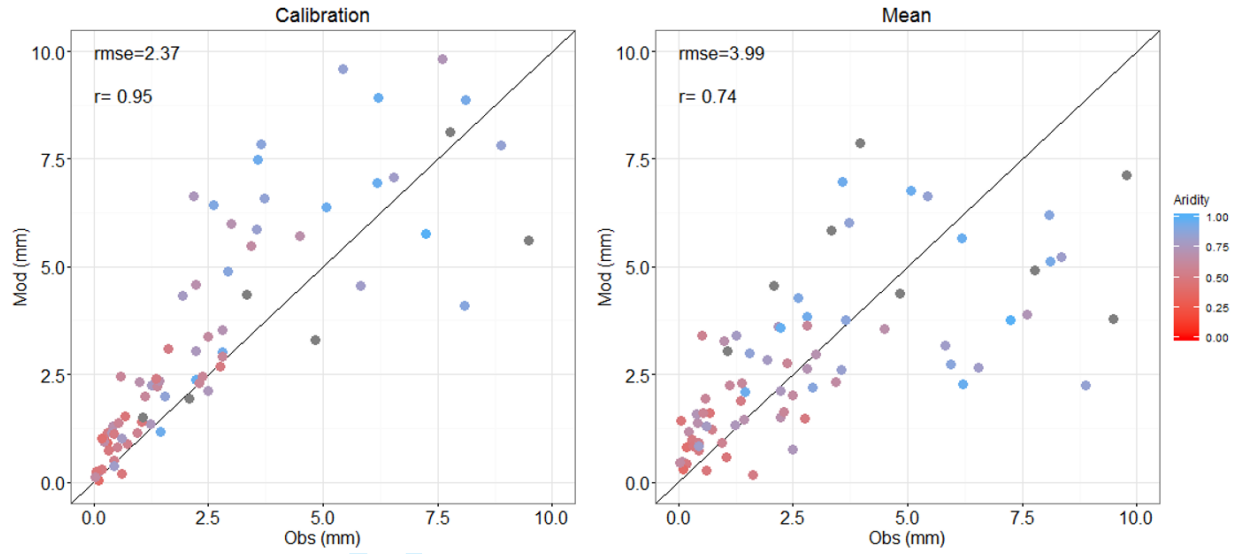
Figure 4. Comparison of observed minimum flow Q_{min}, with predictions from the calibrated and mean-value parameterizations for Australian catchments. Note that the plot window excludes between 12 and 16 catchments with high values of Q_{min}. rmse=root mean square error in mm/mo
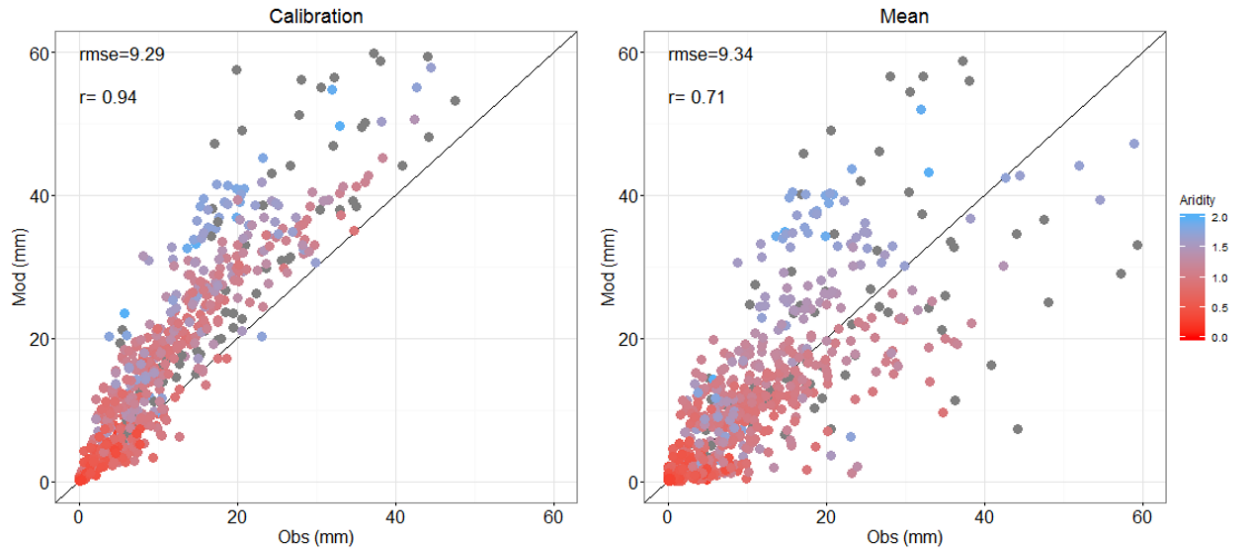
Figure 5. Comparison of observed minimum flow Q_{min}, with predictions from the calibrated and mean-value parameterizations for US catchments. Note that the plot window excludes between 12 and 16 catchments with high values of Q_{min}. rmse=root mean square error in mm/mo
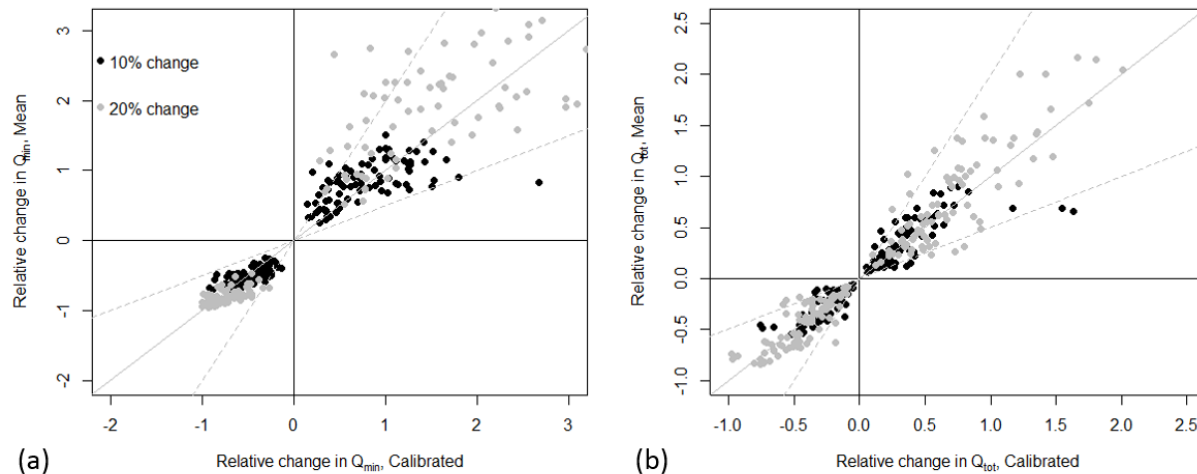
Figure 6. Predictions for Q$_{min}$ (a) and Q$_{tot}$ (b) resulting from a hypothetical land use change – i.e. a change in α$_1$ and α$_2$ values – for both the calibrated and the mean-value models. Each point represents one catchment under either 10% (black) or 20% (grey) relative increase or decrease in α$_1$ and α$_2$. All values in the bottom-left quadrant represent an increase in α$_1$ and α$_2$, while all values in the top-right quadrant represent a decrease in the two parameters. Dashed lines represent a 50% difference between the calibrate and mean-value predictions. RMSE for Q$_{min}$ is 0.36 for the "10% change" and 1.2 for "20% change". For the increase in α$_1$ and α$_2$ (bottom-left quadrant), representing afforestation, RMSE for "20% change" is 0.13.
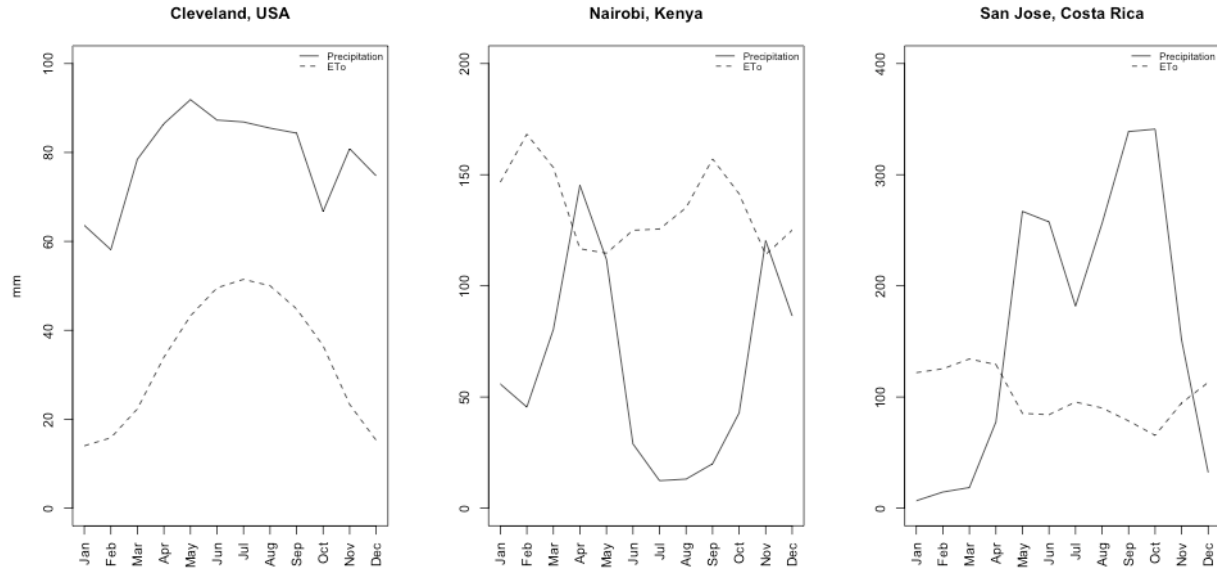
Figure A1. Climate types used in this study: humid continental (Cleveland), subtropical with dry-summer (Nairobi), tropical with dry-winter (San Jose)