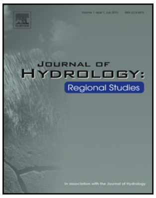
Journal of Hydrology: Regional Studies 2 (2014) 35–48

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh



Multivariate power-law models for streamflow prediction in the Mekong Basin

Guillaume Lacombe[a,*], Somphasith Douangsavanh[a], Richard M. Vogel[b], Matthew McCartney[a], Yann Chemin[c], Lisa-Maria Rebelo[a], Touleelor Sotoukee[a]

[a] International Water Management Institute, Southeast Asia Regional Office, PO Box 4199, Vientiane, Lao Democratic People's Republic
[b] Tuft University, Department of Civil and Environmental Engineering, Medford, MA 02155, USA
[c] International Water Management Institute, Headquarters, PO Box 2075, Colombo, Sri Lanka

ARTICLE INFO

Article history:
Received 7 March 2014
Received in revised form 8 August 2014
Accepted 13 August 2014
Available online 20 September 2014

Keywords:
Streamflow prediction
Ungauged catchment
Multivariate regression models
Mekong

ABSTRACT

Study region: Increasing demographic pressure and economic development in the Mekong Basin result in greater dependency on river water resources and increased vulnerability to streamflow variations.

Study focus: Improved knowledge of flow variability is therefore paramount, especially in remote catchments, rarely gauged, and inhabited by vulnerable populations. We present simple multivariate power-law relationships for estimating streamflow metrics in ungauged areas, from easily obtained catchment characteristics. The relations were derived from weighted least square regression applied to streamflow, climate, soil, geographic, geomorphologic and land-cover characteristics of 65 gauged catchments in the Lower Mekong Basin. Step-wise and best subset regressions were used concurrently to maximize the prediction *R*-squared computed by leave-one-out cross-validations, thus ensuring parsimonious, yet accurate relationships.

New hydrological insights for the region: A combination of 3–6 explanatory variables – chosen among annual rainfall,

---

* Corresponding author. Tel.: +856 21 77 14 38.
E-mail addresses: g.lacombe@cgiar.org (G. Lacombe), s.douangsavanh@cgiar.org (S. Douangsavanh), richard.vogel@tufts.edu (R.M. Vogel), m.mccartney@cgiar.org (M. McCartney), ychemin@gmail.com (Y. Chemin), l.rebelo@cgiar.org (L.-M. Rebelo), t.sotoukee@cgiar.org (T. Sotoukee).

http://dx.doi.org/10.1016/j.ejrh.2014.08.002

drainage area, perimeter, elevation, slope, drainage density and latitude – is sufficient to predict a range of flow metrics with a prediction *R*-squared ranging from 84 to 95%. The inclusion of forest or paddy percentage coverage as an additional explanatory variable led to slight improvements in the predictive power of some of the low-flow models (lowest prediction *R*-squared = 89%). A physical interpretation of the model structure was possible for most of the resulting relationships. Compared to regional regression models developed in other parts of the world, this new set of equations performs reasonably well.

1. Introduction

Population growth, economic development and climate change increase the vulnerability of people and ecosystems to variations in river flow. The Mekong Basin in Southeast Asia exemplifies these issues with growing irrigation water demand (Pech and Sunada, 2008), greater flood-risk exposure (Osti et al., 2011), and hydropower-induced changes in seasonal river flow and ecology (Arias et al., 2012; Ziv et al., 2012). Adaptation measures are hampered by uncertainties in projected streamflow changes (Kingston et al., 2011). A number of hydrological models have been developed for the Mekong Basin to predict streamflow variability, however their complexity and lack of transparency (Johnston and Kummu, 2012), often limit possible users to modeling experts, instead of the practitioners working closely with populations affected by flow extremes. Additionally, the majority of models have been developed to predict flow along the Mekong mainstem, precluding accurate assessments in headwater catchments where populations are repeatedly exposed to flash floods and/or water resource shortages.

Flow duration curves (FDCs) provide an integrated representation of flow variability that can be used for water resource planning, storage design and flood risk management (Castellarin et al., 2013). A period-of-record FDC indicates the percentage of time (duration) a particular value of streamflow is exceeded over a historical period. Similarly, a median annual FDC can reflect the percentage of time a particular value of streamflow is exceeded in a typical or median year (see Vogel and Fennessey, 1994). Various parametric and nonparametric statistical methods exist to predict an FDC in ungauged catchments and have been applied in many parts of the world (Castellarin et al., 2004).

We present a set of new multivariate power-law models to predict FDC percentiles as well as other flow metrics, at any location along the tributaries of the Lower Mekong River (Fig. 1) using easily determined catchment characteristics. Section 2 describes the main steps of the multiple regression analysis. Section 3 presents the data used to empirically develop the models. Section 4 presents the equations of the power-law models, discusses their significance and compares their performance with other case studies.

2. Multiple regression analysis

We used a multivariate power-law equation (Eq. (1)), already used in many parts of the world (Vogel et al., 1999; Castellarin et al., 2004), to estimate the river flow Q from m catchment characteristics X_i ($i = 1, \ldots, m$). A logarithmic transformation of Eq. (1) results in a log-linear model (Eq. (2)) whose coefficients β_i ($i = 1, \ldots, m$) can be determined by multiple linear regression.

$$Q = \exp^{\beta_0} \cdot X_1^{\beta_1} \cdot X_2^{\beta_2} \cdots X_m^{\beta_m} \cdot \nu \tag{1}$$

$$\ln(Q) = \beta_0 + \beta_1 \cdot \ln(X_1) + \beta_2 \cdot \ln(X_2) + \cdots + \beta_m \cdot \ln(X_m) + \varepsilon \tag{2}$$

β_0 is the intercept term of the model. ν (Eq. (1)) and ε (Eq. (2)) are the log-normally and normally distributed errors of the models, respectively. The natural logarithm (ln) being defined for strictly positive values only, catchment characteristics X_i and flow Q with possible zero values are incremented