Correlation and causation in tree-ring based reconstruction of paleohydrology in cold semi-arid regions

Amin Elshorbagy; Thorsten Wagener; Saman Razavi; David Sauchyn

Abstract

This paper discusses ways in which the tree-ring based reconstruction of paleohydrology can be better understood and better utilized to support water resource management, especially in cold semi-arid regions. The relationships between tree growth, as represented by tree-ring chronologies (*TRCs*), runoff (*Q*), precipitation (*P*), and evapotranspiration (*ET*) are discussed and analyzed within both statistical and hydrological contexts. Data from the Oldman River Basin (OMRB), Alberta, Canada, are used to demonstrate the relevant issues.. Instrumental records of *Q* and *P* data were available while actual *ET* was estimated using a lumped conceptual hydrological model developed in this study. Correlation analysis was conducted to explore the relationships between *TRCs* and each of *Q*, *P*, and *ET* over the entire historical record (globally) as well as locally in time within the wet and dry subperiods. Global and local correlation strengths and linear relationships appear to be substantially different. This outcome particularly affects tree-ring based inferences about the hydrology of wet and dry episodes when reconstructions are made using regression models. Important findings include: (i) reconstruction of paleo*Q* may not be as credible as paleo*P* and paleo*ET; (ii)* a moving average window of *P* and *ET* larger than one year might be necessary for reconstruction of these variables; and (iii) the long term mean of reconstructed *P*, *Q*, and *ET* leads us to conclude that there is uncertainty about the past climate. And finally, we suggest using the topographic index to pre-judge side suitability for dendrohydrological analysis.

Keywords: dendrohydrology; moving average; semi-arid regions; reconstructing evapotranspiration; reconstructing precipitation.

1. Introduction

Hydrology, as a natural science, has flourished and made significant advancements based on observations and empirical evidence derived from the relatively short instrumental record available. Hydrological processes exhibit different behaviors at different spatial and temporal scales, hence providing differences in information for scientific study (e.g. Sawicz et al., 2011; Singh et al., 2015). Such behavior, as expected, is yet to be fully understood given the high complexity and heterogeneity of hydrological systems, which has been a consistent challenge for hydrological modeling and prediction, as well as for water resources planning and management (McDonnell et al., 2007; Wagener et al., 2010). As a pragmatic solution, hydrologists rely on data for conceptualizing the functionality of hydrological systems, developing theory, and building both mechanistic and data driven models. However, records of hydrometeorological observations and measurements are usually too short to contain sufficient hydrological variability for long-term water management. In some parts of the globe the instrumental period extends to about a century, but in other regions it is considerably shorter. These short records not only limit the extractable knowledge but also affect our perception and definition of important hydrological principles, such as stationarity, change, and return periods of low frequency events (Hirsch, 2011; Kundzewicz, 2011; Salas et al., 2012; Sawicz et al., 2014; Razavi et al., 2015; Alam and Elshorbagy, 2015).

One of the solutions to the problem of limited hydrological records is extending such records using proxy data. Records extended in this manner can span over long periods, including episodes of wet and dry conditions of various lengths, thus providing water resources planners and managers with a means to develop more robust water policies. Such paleohydrological data series provide the potential for a valuable expansion of hydrological knowledge. Tree-rings (dendrohydrology) can be considered superior proxies of paleohydrology, compared to other natural proxy records, because of their ability to represent hydro-climatic behavior at a relatively fine temporal resolution – yearly and perhaps even sub-yearly (Crawford et al., 2015). Tree-ring proxy data are available in many regions in the world, including Britain (Jones et al., 1984), Chile (Urrutia et al., 2011), Canada (Boucher et al., 2011; Sauchyn et al., 2011), China (Gou et al., 2007), Morocco (Till and Guiot, 1990), and the United States (e.g., Cleaveland and Duvick, 1992; Gray and McCabe, 2010; Maxwell et al., 2011; Woodhouse and Lukas, 2006). Tree-ring data have been used to reconstruct the time series of various variables, including precipitation (Blasing et al., 1988; O'Donnell et al., 2015), temperature (Briffa et al., 1992; Fritts and Lough, 1985; Dorado Liñán et al., 2015), runoff (Meko et al., 2012; Axelson et al., 2009; Cleaveland and Stahle, 1989; Razavi et al., 2016), and drought index (Cook et al., 2004; Cleaveland and Duvick, 1992; Blasing et al., 1988; Tei et al., 2015).

Considerable success in reconstructing paleohydrology has been achieved; however, several unaddressed challenges, unresolved issues, and unexplored opportunities remain (Cook and Pederson, 2010); which limits their use by water resource managers. Investigating and understanding such challenges and opportunities, centered around the dynamics of tree growth and water uptake, and its relationship with other hydrological processes, is of fundamental importance for water resources planners and managers. The aim of this paper is to highlight some of the issues pertinent to the use of dendrohydrology for water resources planning and management purposes, particularly in semi-arid regions, and demonstrate them through a Canadian case study. In particular, differentiating between statistical correlation and physical and hydrological causation is the overarching objective. The next section of this paper provides a brief description of dendrohydrology and the main controls on tree growth, followed by discussions of some of the challenges of dendrohydrological reconstructions and their conceptual reliability. A description of the data and case study used in this paper for demonstrative purposes is given in Section 4. The methods and analysis approach is then explained in Section 5, followed by results and findings in Section 6. The paper ends with a brief section of conclusions.

2. A brief review of dendrohydrology

Thorough dendrohydrological investigations require understanding of both interactions between tree stands and surrounding hydroclimatic conditions, as well as ecological conditions, which include diseases and insect pests, soil fertility and erosion, plant competition, wildfire, and various sources of environmental pollution. These ecological conditions, affect the biological response of trees to climate and water availability (Loaiciga et al., 1993). However, they have received less attention due to the large uncertainty in quantifying their effects, owing to lack of data and the variability of response among individual trees. In the literature, the role of hydroclimatic conditions has been discerned from that of the ecological conditions by investigating tree populations and spatially-aggregated tree-ring chronologies rather than chronologies of individual trees (e.g., Breitenmoser et al., 2014; Ogle et al., 2015).

The major hydroclimatic controls (or limitations) on monthly tree growth, imprinted in tree-ring widths, are moisture, temperature, and solar radiation (Breitenmoser et al., 2014). For example, if temperature and radiation are not limiting, such as the case in low mid latitudes (e.g., southwestern USA), tree growth might be a function of moisture availability (Woodhouse et al., 2016). Therefore, (monthly) growth is assumed to be the minimum of the growth responses to temperature (g_T) and moisture (g_M), modulated by the response to insolation (g_E) (Breitenmoser et al., 2014). Annual ring width (*G*) is then defined as the sum of the monthly growth increments. The value of g_E depends on the mean monthly daytime length relative to that in the summer solstice month, which depends on the latitude of the site under consideration (Breitenmoser et al., 2014). The growth response values (g_M and g_T) can be scaled between zero and one to correspond to onset values of moisture and temperature (M_1 and T_1) and upper threshold values of M_2 and T_2, respectively. Tree growth is inhibited at values lower than M_1 and T_1, and insensitive to values higher than M_2 and T_2 (Figure 1). In a global dendrohydrological study (Breitenmoser et al., 2014), the values of T_1 (°C), T_2 (°C), M_1 (v/v), and M_2 (v/v) were found to be within the ranges of (1.74 – 8.41), (10.14 – 22.80), (0.019 – 0.025), and (0.11 – 0.64), respectively, where v/v refers to volume of soil moisture per unit volume of soil.

From a hydroclimatic point of view, moisture is the limiting factor; i.e., $g_M < g_T$, to tree growth in warm arid and semi-arid regions, such as the southwestern area of the United States. Such sites can be considered moisture-limiting sites. In cold regions such as the Canadian prairies and Rocky Mountains, temperature can be the limiting factor on the *annual* basis, causing a short temperature-permitting growing season (May – September). However, during the growing season, moisture becomes the limiting factor in the Canadian prairies and lower elevations of the Rocky Mountains. The insolation factor (g_E) plays a more important role at latitudes far away from the Equator in northern and southern hemispheres. In light of such understanding of the limiting factors for tree growth in a region, it is indeed sensible to focus reconstruction attempts on the limiting hydroclimatic variables (Breitenmoser et al., 2014).

A wealth of dendrohydrology studies, starting from the late 1960s (Fritts et al., 1971) and continuing to date (Gangopadhyay et al., 2015; Razavi et al., 2016), investigate long-term hydrological phenomena. Reconstruction of hydrological variables like runoff or precipitation is sought based on finding significant correlations between the variable and the tree-ring chronology, which is then used to develop data driven, mainly regression models (Loaiciga et al., 1993). However, maximizing correlation statistically, without much physical underpinning of the choices made, is not uncommon. Choices include (i) identifying the start and end of the water year that best correlates with the chronologies under consideration – July of the previous year to June of the current year (e.g., Gray and McCabe, 2010); August to July (e.g., Cleaveland and Duvick, 1992); or October to September (e.g., Till and Guiot, 1990); (ii) correlating chronologies with only particular month(s) of the year (e.g., Lutz et al., 2012); (iii) conducting principal component analysis and canonical correlation analysis (Fritts et al., 1971) to remove multicollinearity of inputs (chronology sites) and output (e.g., precipitation gauges) and to strengthen model predictive power (Axelson et al., 2009); and (iv) removing autocorrelation of tree-ring series (Starheim et al., 2013; Wise, 2010) and even the hydrological time series (Meko et al., 2011; Cleaveland and Stahle, 1989). Investigating the hydrological relationships that position dendrohydrology explicitly within the hydrological cycle can link statistical correlation with physical causation, and thus, maximize the utility of dendrohydrology for water resources planning and management.

3. Challenges and opportunities

3.1 Uncertainty and inaccuracy of dendrohydrological reconstructions

A common practice in dendrohydrology is the development of a statistical regression model that links a hydrological variable; e.g., runoff, to one or multiple tree-ring chronologies (*TRCs*). Using the available long record of *TRCs*, a corresponding long record of runoff can then be reconstructed. Some of the challenges facing dendrohydrology include the widely acknowledged uncertainty and inaccuracy with respect to these reconstructions (Cook and Pederson, 2010). It is common to reconstruct long records of runoff, or other variables, from proxy variables that account for less than 50% of their variability (Graumlich et al., 2003; Axelson et al., 2009). In fact, some published reconstructions account for as little as 35% (Gedalof et al., 2004), 37% (Axelson et al., 2009), and 38% (Watson et al., 2009) of hydrologic variability. Such limited explanatory power has been attributed to the limited representativeness of the sampled tree sites for the water balance of the larger catchment (Axelson et al., 2009), while the presence of other complex non-climate and biological factors that affect the response of trees as living objects might also play a role (Cook and Pederson, 2010).

3.2. Selection of hydrological variables for reconstruction

Another important challenge emphasized here is the selection of the hydrological variable to reconstruct in the first place. It is intuitive to hypothesize that evapotranspiration is the most relevant hydrological process, and thus, most correlated with tree growth, as represented by the tree-ring chronology (*TRC*). Through transpiration, trees fix CO_2 into organic compounds that form the plant's structural biomass. The reason for correlating other hydrological variables (e.g., precipitation, runoff) to *TRCs* is the availability of reasonably long records through direct measurement, while ET is generally an estimate itself. Both runoff and evapotranspiration are correlated with *TRCs* through a state variable, namely soil moisture. With respect to cause-effect relationships in the hydrological cycle, precipitation is the cause (independent variable) while runoff, evapotranspiration, and, thus, *TRC*, are effects (dependent variables) of a system experiencing a certain weather regime. One of the reasons for the nonlinearity of the relationships among the variables is soil moisture storage, which plays a central role in these relationships. The storage effect on the correlation coefficient can be substantial, and therefore, it can be further hypothesized that the storage medium leaves its signature on all variables that are dependent on it.

Ignoring the groundwater flow in a simplified annual water balance, where change in water storage in the basin is typically considered negligible if starting and ending at times with similar wetness stages, can be mathematically represented by

$$\text{Precipitation } (P) = \text{Evapotranspiration } (ET) + \text{Runoff } (Q) \quad (1)$$

Building a statistical relationship between runoff (*Q*) and *TRC*, the latter being related to *ET*, conceptually entails both statistical and physical problems. Physically, *ET* and *Q* are complementary variables that combine to close the water balance with *P*, assuming insignificant change in storage (which is a common practice for annual water balance). An increase in one does not necessarily mean an increase in the other, especially at certain time scales. This is particularly obvious in semi-arid regions where annual potential evapotranspiration (*PET*) exceeds annual values of *P*, leaving smaller amounts exiting the watershed in the form of *Q* after most of the *P* is depleted by *ET* during the growing season. For example, in Canadian semi-arid regions such as the Oldman River Basin in southern Alberta, runoff is snowmelt-dominated and peaks before *ET* climbs to its maximum rate in the summer. Runoff might peak when temperature is the limiting factor for tree growth, rendering a statistical relationship between tree growth and Q unreliable. This is typical in many northern watersheds.

Figure 2 presents the average monthly *P*, *Q*, and *PET* for the Oldman River Basin in western Canada. *PET* peaks in July concurrent with a decline in both *P* and *Q*. This time period also exhibits the highest values of actual evapotranspiration (*ET*) in the region, as indicated by a dense canopy and high leaf area index (tree growth) as well as measurements of *ET* using eddy covariance towers (Brown et al., 2014). The factual out-of-phase seasonality between *ET* and *Q* makes it difficult to argue in support of a causal relationship between *Q* and *ET* (effectively *Q* and *TRC*). The reason that predicted *ET* looks in phase with *P* and *Q* in Figure 2 will be discussed in Section 6.2. Indeed, striking isotopic evidence provided by Evaristo et al. (2015) and Brooks et al. (2010) show two almost separate "worlds or pools of water" contributing to runoff and evapotranspiration; specifically, runoff tends to be "new and mobile" water whereas tree water tends to come from older, tightly bound water in the soil. However, smoothing out the seasonal variability by considering the annual values, as typically done in dendrohydrology, can create a somewhat artificial correlation between *Q* and *ET*, especially in light of the fact that both water exits are modulated through the same filter, namely soil storage (Rinaldo et al., 2015). When causation is absent, the presence of statistical correlation should be treated with caution.

Statistically, constructing a predictive model for runoff based on a tree-ring chronology encompasses a risk of establishing a spurious relationship; that is, the correlation between the variables *Q* and *TRC* (or *ET*) is due, at least partially, to ignoring the confounding factor: *P*. A confounding factor is one that is driving or affecting each of the independent and dependent variables (Greenland et al., 2009). This can lead to epistemic and unquantifiable uncertainty as the established relationship may not work properly at times when moisture levels are insufficient to generate runoff, but sufficient for average tree growth. Not surprisingly, many dendrohydrologists have documented the under-predicting high runoff volume (e.g., Wise, 2010; Axelson et al., 2009).

In the previous discussion, a few open issues in dendrohydrology have been identified. In the remainder of this paper we will pursue the idea of establishing an association between *TRCs* and actual evapotranspiration *(ET),* thus, shedding some light on the association between *TRCs* and both *P* and *Q*. Only recently, dendrohydrologists started to attempt quantifying the relationships between TRCs and ET (Gangopadhyay et al., 2015) as well as soil water storage (Creutzfeldt et al., 2015).

4. Study region and data

The Oldman River Basin in southern Alberta, Canada, is presented as a case study. It is located in a suitable area (Figure 3) to discuss some of the issues highlighted above, as it is a semi-arid region with annual potential evapotranspiration exceeding annual precipitation (Figure 2). The tree-ring chronologies of the OMRB have been studied by Razavi et al. (2015, 2016), Sauchyn et al. (2015), and Axelson et al. (2009). The basin has a drainage area of approximately 26,700 km^2 (Alberta Environment, 2014) covering three natural regions: the Rocky Mountains, Foothills, and Grassland. The average annual precipitation in the OMRB is 488 mm (AMEC, 2009). In the warm months of May through August, precipitation is less than evapotranspiration and, hence, most agricultural areas rely on irrigation. The average annual natural flow of the Oldman River at the headwaters is 56 m^3/s, and peak runoff typically occurs in June (Figure 2). The headwaters include the Oldman, the Castle, and the Crowsnest Rivers, which join together in the Oldman River Reservoir. The St. Mary, Belly, and Waterton Rivers are the southern tributaries and originate from the state of Montana in the USA.

There are reasonably extensive records of *TRCs* of moisture-sensitive trees at various sites close to the headwaters of the Saskatchewan River Basin (SaskRB) in general, and the OMRB in particular. The longest records available date back to 1370 A.D., while the shortest record begins at 1750 A.D. Efforts have been made to reconstruct the annual runoff for different tributaries of the SaskRB (Axelson et al., 2009; Case and MacDonald, 2003; Sauchyn et al., 2011, 2015). Further, Fleming and Sauchyn (2013) analyzed previously reconstructed runoff time series of two major tributaries of the SaskRB to study the shifts and variance in water availability.

The *TRC* data of 16 sites in the OMRB were provided by the tree-ring lab at the University of Regina (http://www.parc.ca/urtreelab/) and are listed in Table 1 and shown in Figure 3. The length of the tree-ring records at different chronology sites varies considerably. Tree-rings were measured to within 0.001 mm from high-resolution images of polished wood samples using a semi-automated image analysis system (WinDendro Density). The measured tree-ring series were standardized using the program ARSTAN (Cook, 1985). Conservative detrending by a negative exponential curve was used to remove the juvenile biological growth trends in the tree-ring series. The standardized ring-width series were averaged for each site, using a mean value function that minimizes the effects of outliers.

A naturalized runoff, which is measure streamflow processed to remove the effects of flow regulation and water abstractions, time series for the period 1912-2001 at the watershed outlet (Figure 3) generated by Alberta Environment and Sustainable Resource Development was used to represent the basin runoff (Q). Daily meteorological variables (1953-2004), including air temperature (°C), relative humidity, wind speed (m/s), and vapor pressure (kPa), were obtained from Environment Canada and processed to produce monthly values. The *Second Generation of Daily Adjusted Precipitation for Canada* (Mekis and Vincent, 2011) product was used as the source of precipitation data. Precipitation records for the cities of Lethbridge and Medicine Hat and towns of Pincher Creek and Vauxhall (1936-2005) were considered and investigated for their completeness and correlation with the *TRCs* and runoff. The individual and spatial averages were considered, and the precipitation measured at Lethbridge was concluded to be the most suitable.

5. Methods and analysis approach

5.1 Topographical analysis to quantify site suitability

A common observation in dendrohydrology is that *TRCs* from different sites have different correlation strengths with hydrological variables, even for sites with the same tree species. Dendrohydrologists follow a purposeful approach to site selection, targeting sites with the oldest trees and confining sampling to sites where soil moisture is a limiting growth factor (i.e., the upper parts of slopes and those slopes that face south and west and are thus usually dry). However, a more systematic approach to select appropriate sites for dendrohydrological reconstruction is still needed. The intuitive approach is to attempt to investigate the role of topography in a systematic and explicit way, as topography is a major controller of hydrological conditions (Sørensen et al., 2006). In this study, the topographic index (*TI*), originally developed by Beven and Kirkby (1979) and defined as $\ln(a / \tan \beta)$, was calculated for the entire OMRB. The *TI* captures the two important elements that determine the wetness of an area: the upslope contributing area per contour length (a) and the local slope ($\tan \beta$) (Hornberger et al., 1998). *TI* was calculated for each cell of the basin using a digital elevation model (DEM), with 90m resolution, and ArcGIS. The objective of this step was to investigate the relationship between the *TI* of a site and the correlation between its *TRCs* and the hydrological variables.

5.2 Conceptual hydrological model

A lumped conceptual monthly hydrological model was developed for the OMRB as a way to predict the basin-scale actual evapotranspiration (*ET*). Linking *ET* to the *TRCs* is the main reason for developing this model and, therefore, a monthly temporal resolution was selected as an appropriate modeling time step. This step size is coarse enough to filter out processes occurring at finer temporal resolution, which may not be of direct relevance in this study, but also fine enough to represent the process seasonality. The monthly precipitation was considered to be rain or snow based on the average monthly temperature relative to 0.0 °C. *PET* values were estimated using the Penman-Monteith equation (ASCE, 2005) typically adopted by Alberta Agriculture for reference evapotranspiration:

$$PET = \frac{\left(0.408\times\Delta\times(R_n - G)\right) + \gamma\times\left(\frac{1600}{T_{Mean}+273}\right)\times u_2\times(e_s - e_a)}{\Delta + (\gamma\times(1+0.38\times u_2))} \quad (2)$$

where Δ is the slope of the saturation vapor pressure-temperature curve (kPa/°C), R_n is net radiation (MJ/m^2/day), G is soil heat flux (MJ/m^2/day), γ is a psychrometric constant (kPa/°C), T_{Mean} is mean daily temperature (°C), u_2 is wind speed at a height of 2 m (m/s), and e_s and e_a are the saturated and actual vapor pressure (kPa), respectively. On average, *G* is small and assumed to be zero, following the recommendation of Alberta Agriculture and Forestry (2015). This might lead to some overestimation of *PET* values, but is acceptable for our modeling purposes because *PET* was used as an upper limit for the predicted actual evapotranspiration, which is always much less than the *PET*.

In cold regions hydrology, various processes complicate hydrological modeling, such as snow accumulation, relocation and sublimation, snowmelt dependence on soil and air temperature and radiation, and infiltration into frozen soil (Pomeroy and Gray, 1995). However, for the monthly lumped model developed here, the question is more one of water balance (bucket approach) and translation of water from storage to exit (Rinaldo et al., 2015). It is not possible, nor necessary, to simulate local scale snow drifting, realistic instantaneous infiltration, and ponding processes. Therefore, the monthly runoff generation was developed using the simple concept of a runoff coefficient, combined with a simple form of the interesting concept of dynamic residence time or different residence time for different "waters" in the watershed (Rinaldo et al., 2015). This *monthly* model distinguishes winter months from spring/summer months in any given year using an average monthly air temperature threshold of zero degrees Celsius. The conceptual model developed considers three sources of runoff:

(i) Baseflow (Q_b): This was found through manual calibration to be around 1.7 mm/month;

(ii) Snowmelt (Q_s): If the average monthly air temperature (T_{Mean}, °C) < 0, then no runoff is generated, and all precipitation is assumed to be snow and accumulates, as cumulative snow water equivalent (*CSWE*) over the number of months for which T_{Mean} is less than 0.0 °C, in a virtual tank according to

$$CSWE = \sum_{i=1}^{w} SWE_i \qquad (3)$$

where *i* is a winter month and *w* is the total number of winter months. This approach makes w a dynamic variable that changes from one year to another. *CSWE* is contributing to runoff based on the ratio of the mean monthly temperature (T_j) relative to the summation of the mean temperature over the spring/summer season,

$$Q_{s\,j} = c_j \left(\frac{T_j}{T_1 + \cdots T_j + \cdots + T_k} \right) x\; CSWE \qquad (4)$$

where c_j is a coefficient that is distributing the snowpack as a snowmelt component of the runoff (Q_s) in month j over a k number of months of the spring and summer season (e.g., May, June, July) during which T_{Mean} is higher than 0.0 °C. Both c_j and k are calibration parameters.

(iii) Summer runoff (Q_r): During the spring and summer (e.g., May, June, July, August, September, October) when the precipitation is in the form of rain (R), a portion of the rain contributes to runoff based on calibration runoff monthly coefficients $d_1, .. d_j, ... d_k, ... d_m$: where m is the total number of spring and summer months, and j and k as defined ealier.

$$Q_{r\,j} = d_j\left(R_j\right) \quad (5)$$

Other than contributing to runoff, a portion of the melting snowpack contributes towards evapotranspiration and was estimated to close the water balance. On average, this was found to be 17% of the total snow. Actual evapotranspiration (*ET*) was calculated during the spring and summer months (*1, ..., m*):

$$ET_j = \left(1 - d_j\right)R_j \quad (6)$$

The total seasonal *ET* is calculated based on the cumulative ET_j from Equation (6) over *m* months plus the contribution of snowmelt to *ET*. The model parameters were calibrated manually using the maximization of Nash-Sutcliffe (NS) efficiency of the predicted and observed runoff as the objective function. The annual *ET* values (October – September) were investigated with respect to their relationship with the *TRCs*.

6. Results and analysis

6.1 Topographical analysis results

The topographic index (*TI*) values of the pixels within which the *TRCs* were sampled are provided in Table 1 and plotted in Figure 4, along with the correlation coefficients between *TRCs* and each of the yearly precipitation and runoff series. The expected trend is a decrease in correlation coefficients with increasing *TI* values, as higher *TI* relates to sites that for topographical reasons are more likely to be wet. Wetter sites obscure a direct (instantaneous) relationship between meteorological variables and the moisture available for trees. From a dendrohydrological point of view, higher *TI* adds uncertainty to the moisture response (g_M) at moisture-sensitive sites. Even though the trend presented in Figure 4 is not completely consistent with this hypothesis, 15 of the 16 sites (except TAB, which was removed from the graph because it shows substantial deviation from the other sites) support the conclusion with respect to both runoff and precipitation. Given that these results are based on a coarse resolution (90 m) DEM, this finding is important as it allows for formalized site selection for the purpose of identifying moisture-limited *TRCs* that more strongly correlate with hydrological variables.

6.2 Hydrological model results

The observed and simulated runoff time series at the outlet of the OMRB are shown in Figure 5. Two-thirds of the available record (1953-1984) was used for model calibration and one third (1985-2001) was used for evaluation. The model performance is quite similar over both periods with Nash-Sutcliffe coefficients of 0.59 and 0.60 for the calibration and evaluation periods, respectively. Except for the runoff flood event in June 1995, the model captured the runoff pattern reasonably well over the entire simulation period. Other error measures (correlation coefficient, R, and root mean squared error, RMSE) are also shown in Figure 5.

The calibration parameter *k* was found to be equal to 3, i.e., the snowmelt water is contributing to runoff over a three-month period in the spring, and the time-varying values of *c* and *d* are provided in Table 2. The model error (residuals) were found to be independent, with a very small autocorrelation coefficient of 0.07 and a mean value of -0.80 mm, indicating slight bias (underprediction). The simulated monthly actual evapotranspiration values (*ET*), shown in Figure 2, were aggregated into annual time series to investigate their correlation with various *TRCs*. As expected, *ET* starts to rise in the spring (May and June), benefiting from both rain and snowmelt. During the summer months (July, August, and September), *ET* and *Q* together are higher than incoming rainfall, suggesting that both processes are using water remaining from snowmelt. During summer, the rate of *ET* decline is much smaller than the rate of *Q* decline. The predicted ET shown in Figure 2 peaks in June in response to the moisture availability, as the developed model, like most available watershed models, may not capture the process of the trees taking up older water in the soil (Brooks et al., 2010; Ogle et al., 2015). This may not be accurate and other evidences from the region indicate that ET typically peaks in July (Brown et al., 2014).

6.3 Correlation analysis results

TRCs from four sites were selected for further analysis on the basis of their relatively high correlation coefficient values with the measured *P* and *Q*. The corresponding correlation values for all 16 sites are provided in Table S1 as a supplemental material. Interestingly, the concurrent (annual) correlation coefficient between the *ET* series and the *TRCs* were found to be as low as 0.44, 0.33, 0.35, and 0.33 for records from BVL, CAB, CAL, and OMR sites, respectively.

These are lower values than the correlations between the *TRCs* and both *Q* and *P* (Table 3; concurrent values are denoted with subscript *t*). Another interesting observation is the weaker correlation between *TRCs* and precipitation, compared to that with runoff, for three of the four sites. This issue can be attributed in part to storage (Creutzfeldt et al., 2015). This finding is supported by the observation that *TRC* signals (time series) and *P* are not consistently in phase, i.e., years with considerably low precipitation did not cause reduction in the tree growth; especially when they happen after wet years. The biological carryover phenomenon in tree growth (Brockway and Bradley, 1995) should be tracked in the correlation, without attempting to filter it out as it is entangled with the instantaneous climatic signal.

The correlation between *TRCs* at year *t* (October of year *t-1* to September of year *t*) and the running average of *ET* over year *t* and *t-1* (October of year *t-2* to September of year *t*) was investigated to account for such "carryover" effect, as well as the effect of old water. When the running average of ET is considered, the correlation coefficient notably improves, as shown in the last two columns of Table 3. Two observations are notable at this point, one hydrological and one dendrohydrological. First, what is typically modeled as an instantaneous process in catchment scale hydrology is based on the assumption of complete (Brooks et al., 2010) or partial mixing of water in a way similar to the conceptual model developed in this study. The instantaneous value of *ET*, which is water vapor leaving the watershed as a lithospheric sub-system, can include water that has been stored for a long time – perhaps longer than one year. This storage can be either within the soil (Brooks et al., 2010; Ogle et al., 2015) or within the tree itself. Indeed water can be stored within the tree for a short time period (Pfautsch et al., 2015), and there is no evidence to support the idea of much longer within-tree water storage. Therefore, if the long-term water storage concept is correct, then logically, precipitation should be correlated with *TRCs* in the same fashion, i.e., old water is available to trees over an extended period average. The first two columns of Table 3 support this hypothesis, as the correlation coefficients between *TRCs* at year *t* and *P*-average over two years are higher than for instantaneous correlations. Second, reconstructing a 2-year running average of *ET* and *P* from yearly *TRCs* is more reliable than reconstructing yearly values. This is not only due to statistically higher correlation values, but also because of the higher confidence in the physics of the relationships and the tree physiology. In a recent study, Ogle et al. (2015) investigated the issue of ecological memory and quantified the individual contributions of endogenous effects (e.g., antecedent tree growth) and exogenous effects (e.g., antecedent precipitation) to current tree growth. They found that the antecedent precipitation of the previous two years, independent of the antecedent growth, could affect the current year's tree growth; they attribute this to water stored in the deep soil for this period of time. Endogenous effects were found to be an inherent property of individual trees, varying significantly within the population of the same site, whereas exogenous effects are consistent across the population. Therefore, models operating at the population level may reliably use the antecedent climate driver (Ogle et al., 2015).

The two-year average might include two wet years, two dry years, or one wet and one dry year. It is argued here that this is what the tree-ring width reflects and what should be modeled. This is notably different from just considering year *t+1* of the *TRC* as an additional independent variable for predicting *P* or *ET*. Doing the latter, which is a common practice in dendrohydrology (Meko et al., 2011; Brockway and Bradley, 1995), might improve the correlation, but is less related to process understanding. For example, in the case of the BVL site, R is increased to 0.57, when 2-year average of *TRC* is considered, from the original value of 0.53, which is quite lower than 0.71 that results from our approach of running average, explained above. From a hydrological point of view, this argument points to the existence of change in storage term in Equation 1, even for annual water balance. Even though the n-year average window can be taken as a generalized concept, it is important to note that the two-year average window identified in this study is region-specific. Other regions might require different window lengths and/or different weight for each year included in the window.

The results provided in Table 3 suggest that the concept of a running average discussed above does not apply to runoff. In most cases, the instantaneous correlation between Q_t and the TRC_t is stronger than that between TRC_t and the running average of Q. This finding is important for this discussion, and should be understood in light of the water balance discussion of Section 3 and the concept of residence time and age of runoff (Rinaldo et al., 2015; McDonnell and Beven, 2014). The annual runoff signal reflects the integrated response of the entire watershed that is impacted by complex storage effects (e.g., delay and attenuation), which can create a similar effect to the old tree water uptake and, thus, no running average of Q may be required. In other words, the effect of the transfer function that relates P to Q can have some similarity with that of the transfer function that relates P to ET, although both functions are different in nature. This argument can be supported quantitatively by the autocorrelation coefficients of all variables under consideration. While the lag-1 autocorrelation coefficients (ρ) of the annual ET and P series are -0.12 and -0.03, the value is much higher for Q (0.23) and closer to that for the BVL TRC (0.40). The Q signal carries a memory already without calculating a running average, which does not exist in the P and predicted ET signals. However, the running average process elevates the lag-1 ρ values of ET and P to 0.50 and 0.46, respectively. It is important to note that the existence of memory in both TRC and Q signals can make them statistically correlated even though the cause of the signal's memory in each case is different.

From the discussion above, both causation and correlation are suggested in the case of the relationship between *P* and *TRCs* (or *ET*). More precipitation leads to more soil moisture, which can be instantaneously (within one year) taken up by the plants for use in photosynthesis. However, when *P* is less than the optimal *ET* requirements (during dry years), the plant can still resort to stored old water and resist any immediate dry-conditions effects. Meteorological drought must persist before resulting in significant soil moisture deficit that leads to ecological/agricultural drought (Mays, 2011), and it is agricultural drought that leads to slowing of tree growth. Therefore, instantaneous correlation can occur between *TRCs* and high values of *P*, but may not be the same between *TRCs* and low values of *P*. On the other hand, agricultural drought, when it persists, leads to hydrological drought. Therefore, low values of *Q* occur after the persistence of agricultural drought, which will have already affected tree growth (*TRCs*). However, even after severe droughts, flash floods or high flow periods can occur (due to sudden snowmelt or rain on snow); which will be reflected in runoff, but the trees might not have enough time to recover to full growth rate. So, in the case of runoff, there can be instantaneous correlation between *TRCs* and low values, but not necessarily high values of *Q*. Indeed, it is difficult for *TRCs* to instantaneously reflect high runoff signals unless high *Q* values persist for extended periods throughout the year. This analysis provides hydrological explanation for earlier empirical findings in dendrohydrology literature (Wise, 2010; Axelson et al., 2009).

Here, the correlation argument is supported by presentation of correlation coefficient values. Each of the available records of measured *P* (1939-2001) and *Q* (1913-2001) was divided based on the mean of the entire record into wet (above mean) and dry (below mean) subsets. The correlation with the corresponding *TRCs* was calculated based on annual instantaneous values (Table 4). The findings are important to dendrohydrology in semi-arid regions, as caution must be exercised when making inferences or interpretations based on reconstructed records, especially interpretation of dry years and wet years based on reconstructed precipitation and runoff, respectively. The *TRCs* have considerably better instantaneous correlations with *P* in wet years but very low instantaneous correlation values in dry years. This relationship is exactly the opposite of that found with *Q*. When developing a model assuming an overall correlation between two series based on the entire record, similar to the practice of developing regression-based models using the entire record, model reliability should be challenged when the assumed relationship is invalid over half of the series.

6.4 The dilemma of reconstruction of hydrological variables

Figure 6 shows 223 years of reconstructed runoff (*Q*) and precipitation (*P*) for four sites. These simple reconstructions were performed using linear regression with the *TRC* for each site as the independent variable. The regression models were developed using the available records from 1955-2001. The models were developed using 66% of the record and validated based on the remaining 34%. *ET* series were reconstructed in a similar way and follow the same pattern, but are not shown here. Both *P* and *ET* were reconstructed using the 2-year running average of years *t-1* and *t* of the dependent variable and the annual value of year *t* of the independent variable (*TRC*), following the logical argument provided earlier. Obtaining the annual values of reconstructed *P* and *ET* is straightforward, as the annual series can be easily obtained from the 2-year running average series simply by knowing the initial conditions (or the value of a single year), i.e., knowing the value of the first year of the reconstructed series. The R^2 of the calibration and the reduction of error (RE) of the validation for the developed models are provided in Table 5. In this study, we used only one independent variable in the reconstruction models, however, when multiple sites are used in the model, the regression R^2 value did not increase more than 2%.

The obvious and typical reconstruction uncertainty due to the choice of predictors, evident in Figure 6, is frequently discussed in the dendrohydrology literature (e.g., Gangopadhyay et al., 2009; Razavi et al. 2016). Even considering the two best and closest models in each case still reveals obvious uncertainty. Reconstructed 2-year running average precipitation (Figure 6b) using the BVL and CAL chronologies show differences in patterns during four periods in particular: 1738-1748, 1760-1772, 1848-1864, 1890-1896, and 1942-1950. Reconstructed annual runoff using the CAB and OMR chronologies show less obvious differences, however, notable differences are obvious during the 1760-1772 period. Furthermore other differences exist in comparison with the other two reconstructions. The differences are in particular important for inferences with respect to periods of possible droughts, as such paleo-reconstructions are commonly used for assessing the timing and magnitude of droughts in the period preceding the observational records. The uncertainty due to the choice of predictors is commonly masked by including all sites as independent variables (regressors) in the reconstruction model; such practice might result in improved regression fit (observational period) but less *reliable* reconstructions (pre-observational period).

The important reliability-related shortcomings that are emphasized in this research relate to the practice of regressing *TRCs* versus hydrological variables. The findings presented in Table 4 clearly indicate possible shortcomings of the conventional regression approach, where different behaviors are observed in wet and dry years in terms of the relationship of tree growth with hydrologic variables. Figure 7a, as an example, shows the relationship between annual precipitation and *TRC* of the OMR site. The overall trend, represented by the black line (slope

0.0012), is influenced by the trend (green dashed line, slope 0.0014) for the wet years (green dots). The dry years (red dots) show almost no trend (red dashed line, slope 0.0003).

The process of averaging the precipitation using a running window of two years helps homogenize the record and makes the slopes of the trend lines of both wet and dry subsets somewhat similar (at 0.0004 and 0.0010, respectively) and also improves the overall correlation, as shown earlier. However, it does not solve the problem of the difference between the overall correlation coefficient and those for the wet and dry subsets (Figure 7b). This finding is significant from both hydrological and modeling viewpoints. It certainly reiterates our earlier statement about the inconsistency of inferences, made using regression models, regarding wet and dry periods (Figure 7a) or even both (Figure 7b). Obviously, if more wet years (expanded green dots) or dry years (expanded red dots) occur, the slope of the overall trend lines in Figure 7 can change, thus casting more doubt and uncertainty regarding the constancy of the developed model. Records with longer runs of wet or dry periods than those in the instrumental period might have existed in the pre-instrumental period. This also helps explain the substantial change of the correlation coefficient over time, even during the instrumental period. Apparently what happens is that more wet or dry years are included within the window considered, and thus, the correlation coefficient changes. Therefore, the reliability of the developed model becomes conditioned on the observed climatology of the instrumental record. Even though this aspect is inherent in most hydrological models, it is more critical with data driven models, such as the regression-based reconstruction models. Indeed, the model structure and parameters are solely dependent on the data used for development. Other *TRCs* (not shown here) behave similarly, as indicated by the data in Table 4. However, a reverse pattern occurs with regard to Q, i.e., the slope is significant at low Q values (dry years) and insignificant at high Q values (wet years).

Regression models, frequently used to reconstruct paleohydrology, can still be of use to the water resources management community. One of the advantages of linear regression is the ability to produce an unbiased value of the mean of the calibration series. The means of the regression-based reconstructed records in this study were compared with the instrumental period's mean values to assess the relative change (Table 6). Two observations can be made with regard to the results. First, similar to typical projections about future climate by various general circulation models (GCMs), *TRC*-based reconstructions of paleohydrology also show a range of changes of the long-term mean, from a 5.3% decrease to 2.7% increase, from a 5.9% decrease to 5.0% increase, and from a 6.0% decrease to 3.4% increase for *P*, *Q*, and *ET*, respectively, in the Oldman River Basin. The differences in the estimation of the long-term mean of *P, Q*, and *ET* are due to the use of different TRC sites for the reconstruction of past records. Second, given that the regression models for *P*, *Q*, and *ET* were developed independently based on measured values of *P* and *Q* and model-based values of *ET* (using only the BVL *TRC*), one can argue that the values, from a water balance point of view, are reasonably consistent. This reflects well-estimated long-term water balance from each chronology site, as the relative change in the input (*P*) is consistent with the relative changes in the outputs (*Q* and *ET*).

As expected, the variance (or the standard deviation) of the regression-based values is considerably less than the variance of the observed records. A summary of the statistical properties of the observed and modeled records is provided in Table 7. The underestimation of variance of reconstructed paleohydrology is acknowledged in dendrohydrology, and was corrected in some cases (Cook et al., 2004). Due to the change in the long-term mean, we recommend that the coefficient of variation (CV), rather than the variance, be used to inflate the variance of the reconstructed records.

7. Conclusions

Precipitation, actual evapotranspiration, and runoff were investigated closely in this paper with regard to how their interrelations affect dendrohydrology and associated dendrohydrological reconstructions. Exercising caution is advised when attempting to reconstruct runoff, as increasing evidence points towards runoff and evapotranspiration (tree water) being drawn from different pools of water, which affects the reliability of their statistical linkage. Reconstructing precipitation and evapotranspiration is therefore more intuitive from hydrological and ecological points of view. Depending on the climatology of the region and how fast the water balance resets, a moving average of the precipitation and evapotranspiration with window length exceeding one year might be considered for reconstruction purposes. In the case of the Oldman River Basin in western Canada, a window length of two years was found to be necessary. This is important from a water resources management point of view because it increases the reliability of reconstructed hydrological variables. Furthermore, it is important from a general hydrological perspective as it indicates the importance of the change in storage even in closing the annual water budget.

Using a linear regression technique, which is common for reconstructing pre-instrumental hydrological time series, should be challenged and revisited as the correlation between tree-ring chronologies and the hydrological variables was found to be substantially different during wet versus dry periods. This difference affects inferences made about past dry or wet episodes using such regression-based models. However, the ability of the regression models to estimate the unbiased mean of the series is advantageous and can be used to estimate the relative change in long-term pre-instrumental record’s mean compared to the instrumental period. Assessment of such relative change is useful for water resources planning and management. Using the long-term mean of the reconstructed runoff, precipitation, and actual evapotranspiration in the Oldman River Basin leads us to conclude that there is uncertainty about the past climate. This uncertainty seems to be similar in nature to that typically produced by GCMs about future projections. Various reconstructions of past hydrological variables in the Oldman River Basin show a range of possibilities in the long-term mean of runoff, precipitation, and evapotranspiration, from a 6% decrease to a 5% increase. It will be interesting to conduct a similar study with regard to future projections to compare the values and assess if past occurrences already contain potential future variability. The doubts cast in this study about the use of regression models for reconstruction of paleohydrology suggest a need for and use of local modeling techniques; i.e., multiple local models parameterized over subsets of the data domain, to capture the pattern of the predictor-predictand relationships in various sub-regions of the space of the variable under consideration. Various forms of the *K*-nearest neighbors (*K*NN) technique (Gangopadhyay et al., 2009) are possibly suitable candidates for this task.

This study also points at the importance of the concept of mobile (new) and immobile (old) waters in the watershed. Even in cases where a coarse annual time scale is used, such as in this study, the concept is applicable and influential. Serious attempts should be made to include this concept in watershed models. Dendrohydrologists follow a purposeful approach for selecting the sites of sampling tree-ring chronologies, and we have shown in this study that using the topographic index is a good way to quantify the site suitability. Caution must be exercised regarding the generalization of the findings of this research. We recommend replicating this study using various sites from other regions of the world to validate or nullify the hypothesis made in this study.

Acknowledgements

Thanks to Toby Dunne and Nasim Hosseini for their help with producing the Topographic Index and the map of the Oldman River Basin, respectively. This work was conducted while the first author was at the University of Bristol, holding the Benjamin Meaker Visiting Professor Fellowship. For enquiries about the data used in this study, please contact Amin Elshorbagy at: amin.elshorbagy@usask.ca.

References

Alam, Md. S. and Elshorbagy, A. 2015. Quantification of the climate change-induced variations in Intensity-Duration-Frequency curves in the Canadian Prairies. *J. Hydrol.*, 527, 990-1005.

Alberta Agriculture and Forestry. 2015. http://www.agric.gov.ab.ca/acis/imcin/aimm.jsp. Last accessed on December 2, 2015.

Alberta Environment. 2014. Alberta's River Basins. http://www.environment.alberta.ca/apps/basins/default.aspx.

AMEC. 2009. South Saskatchewan River Basin in Alberta: Water Supply Study. Alberta Agriculture and Rural Development. Lethbridge, Alberta.

ASCE. 2005. Standardized Reference Evapotranspiration Equation. Environmental and Water Resources Institute (EWRI) of the American Society of Civil Engineers, 59 pp.

Axelson, J. N., D. J. Sauchyn, and J. Barichivich. 2009. New reconstructions of streamflow variability in the South Saskatchewan River Basin from a network of tree ring chronologies, Alberta, Canada. *Water Resour. Res*., 45, W09422, doi:10.1029/2008WR007639.

Beven, K. and M. Kirkby. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrolol. Sci. Bull.*, 24, 43–69.

Blasing, T. J., D. W. Stahle, and D. N. Duvick. 1988. Tree ring-based reconstruction of annual precipitation in the south-central United States from 1750 to 1980. *Water Resour. Res.*, 24(1), 163–171.

Boucher, É., T. B. M. J. Ouarda, Y. Bégin, and A. Nicault. 2011. Spring flood reconstruction from continuous and discrete tree ring series. *Water Resour. Res*, 47(7), W07516.

Breitenmoser, P., S. Brönnimann, and D. Frank, 2014. Forward modelling of tree-ring width and comparison with a global network of tree-ring chronologies. *Clim. Past*, 10, 437–449.

Briffa, K., P. Jones, and F. Schweingruber, 1992. Tree ring density reconstructions of summer temperature patterns across western North America since 1600. *Am. Meteorol. Soc.*, 5, 735-754.

Brockway, C. G. and A. A. Bradley. 1995. Errors in streamflow drought statistics reconstructed from tree ring data. *Water Resour. Res*, 31(9), 2279-2293.

Brooks, J., H. Barnard, R. Coulombe, and J. McDonnell. 2010. Ecohydrologic separation of water between trees and streams in a Mediterranean climate. *Nat. Geosci.*, 3, 100-104.

Brown, S., R. Petrone, L. Chasmer, C. Mendoza, M. Lazerjan, S. Landhäusser, U. Silins, J. Leach and K. Devito. 2014. Atmospheric and soil moisture controls on evapotranspiration from above and within a Western Boreal Plain aspen forest. *Hydrol. Process*., 28(15), 4449–4462.

Case, R. A. and G. M. MacDonald. 2003. Tree ring reconstructions of streamflow for three Canadian prairie rivers. *J. Am. Water Resour. Assoc.*, 39(3), 703-716.

Cleaveland, M. K. and D. N. Duvick. 1992. Iowa climate reconstructed from tree rings, 1640–1982. *Water Resour. Res*, 28(10), 2607-2615.

Cleaveland, M. K. and D. W. Stahle. 1989. Tree ring analysis of surplus and deficit runoff in the White River, Arkansas. *Water Resour. Res.*, 25(6), 1391–1401.

Cook, E. and N. Pederson. 2010. Chapter 4: Uncertainty, emergence, and statistics in dendrochronology. In: Dendroclimatology, Developments in Paleoenvironmental Research, M. Hughes et al. (Eds.), Springer Netherlands.

Cook, E. R., C. A. Woodhouse, M. Eakin, D. M. Meko, and D. W. Stahle. 2004. Long-term aridity changes in the western United States. *Science*, 306, 1015-1018.

Cook, E.R., 1985. A time series approach to tree-ring standardization, Ph.D. dissertation, Univ. of Ariz., Tucson.

Crawford, C. J., D. Griffin, and K. F. Kipfmueller. 2015, Capturing season-specific precipitation signals in the northern Rocky Mountains, USA, using earlywood and latewood tree rings. *J. Geophys. Res. Biogeosci*., 120, 428–440. doi:10.1002/2014JG002740.

Creutzfeldt, B., I. Heinrich and B. Merz. 2015. Total water storage dynamics derived from tree-ring records and terrestrial gravity observations. *J. Hydrol*., 529, 640-649.

Dorado Liñán, I., Zorita, E., González-Rouco, J.F. et al. 2015. Eight-hundred years of summer temperature variations in the southeast of the Iberian Peninsula reconstructed from tree rings. *Clim Dyn*. 44: 75. doi:10.1007/s00382-014-2348-5.

Evaristo, J., S. Jasechko, J. McDonnell. 2015. Global separation of plant transpiration from groundwater and streamflow. *Nature*, 525, 91-94.

Fleming, S. W. and D. J. Sauchyn. 2013. Availability, volatility, stability, and teleconnectivity changes in prairie water supply from Canadian Rocky Mountain sources over the last millennium. *Water Resour. Res*., 49, doi:10.1029/2012WR012831.

Fritts, H. C., and J. M. Lough. 1985. An estimate of average annual temperature variations for North America, 1602 to 1961. *Clim. Change*, 7(2), 203–224.

Fritts, H., Blasing, T., Hayden, B., and Kutzbach, J. 1971. Multivariate techniques for specifying tree-growth and climate relationships and for constructing anomalies in paleoclimate. *J. Appl. Meteorol.*, 10(5), 845–864.

Gangopadhyay, S., G. McCabe and C. Woodhouse, 2015. Beyond annual streamflow reconstructions for the Upper Colorado River Basin: A paleo water-balance approach. *Water Resour. Res*., doi: 10.1002/2015WR017283 (in press).

Gangopadhyay, S., B. L. Harding, B. Rajagopalan, J. J. Lukas, and T. J. Fulp. 2009, A nonparametric approach for paleohydrologic reconstruction of annual streamflow ensembles, *Water Resour. Res*., 45, W06417, doi:10.1029/2008WR007201

Gedalof, Z., D. L. Peterson, and N. J. Mantua. 2004. Columbia River flow and drought since 1750. *J. Am. Water Resour. Assoc*., 40, 1579–1592, doi:10.1111/j.1752-1688.2004.tb01607.x.

Gou, X., F. Chen, E. Cook, G. Jacobs, M. Yang, and J. Li. 2007. Streamflow variations of the Yellow River over the past 593 years in western China reconstructed from tree rings. *Water Resour. Res.*, 43(6): W06434.

Graumlich, L. J., M. F. J. Pisaric, L. A. Waggoner, J. S. Littell, and J. C. King. 2003. Upper Yellowstone River flow and teleconnections with Pacific Basin climate variability during the past three centuries. *Clim. Change*, 59, 245–262.

Gray, S. T. and G. J. McCabe. 2010. A combined water balance and tree ring approach to understanding the potential hydrologic effects of climate change in the central Rocky Mountain region. *Water Resour. Res*., 46, W05513, doi:10.1029/2008WR007650.

Greenland, S., J. M. Robins, and J. Pearl. 1999. Confounding and collapsibility in causal inference. *Statist. Sci.*, 14(1), 29–46.

Hirsch, R.M. 2011. A perspective on nonstationarity and water management. *J. Am. Water Resour. Assoc.*, 47(3), 436-446.

Hornberger, G., J. Raffensperger, P. Wiberg, and K. Eshleman. 1998. Elements of Physical Hydrology. John Hopkins University Press, Baltimore, Maryland.

Jones, P. D., K. R. Briffa, and J. R. Pilcher. 1984. Riverflow reconstruction from tree rings in southern Britain. *J. Climatol.*, 4(5): 461-472.

Kundzewicz, Z. 2011. Nonstationarity in water resources – Central European perspective. Journal of the American Water Resources Association (JAWRA), 47(3): 550-562. doi: 10.1111/j.1752-1688.2011.00549.x.

Loaiciga, H. A., L. Haston, and J. Michaelsen. 1993. Dendrohydrology and long-term hydrologic phenomena. *Rev. Geophys*., 31(2), 151–171, doi:10.1029/93RG00056.

Lutz, E. R., A. F. Hamlet, and J. S. Littell. 2012. Paleoreconstruction of cool season precipitation and warm season streamflow in the Pacific Northwest with applications to climate change assessments. *Water Resour. Res*., 48, W01525, doi:10.1029/2011WR010687.

Maxwell, R. S., A. E. Hessl, E. R. Cook, and N. Pederson. 2011. A multispecies tree ring reconstruction of Potomac River streamflow (1950–2001). *Water Resour. Res.*, 47(5).

Mays, L. W. 2011. Water resources engineering. John Wiley & Sons.

McDonnell, J. J. and K. Beven. 2014. Debates—The future of hydrological sciences: A (common) path forward? A call to action aimed at understanding velocities, celerities, and residence time distributions of the headwater hydrograph, *Water Resour. Res*., 50, 5342–5350, doi:10.1002/2013WR015141.

McDonnell, J. J., et al. 2007. Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, Water Resour. Res., 43, W07301, doi:10.1029/2006WR005467.

Mekis, É and L. A. Vincent. 2011. An overview of the second generation adjusted daily precipitation dataset for trend analysis in Canada. *Atmos-Ocean,* 49(2), 163-177.

Meko, D., C. Woodhouse, and K. Morino. 2012. Dendrochronology and links to streamflow. *J. Hydrol.*, 412-413, 200-209.

Meko, D., D. Stahle, D. Griffin, and T. Knight. 2011. Inferring precipitation-anomaly gradients from tree rings. *Quatern. Int*., 235, 89-100

O'Donnell AJ, Cook ER, Palmer JG, Turney CSM, Page GFM, et al. 2015. Tree Rings Show Recent High Summer-Autumn Precipitation in Northwest Australia Is Unprecedented within the Last Two Centuries. *PLoS ONE* 10(6): e0128533. doi: 10.1371/journal.pone.0128533

Ogle, K., J. Barber, G. Baron-Gafford, L. Bentley, J. Young, T. Huxman, M. Loik, and D. Tissue. 2015. Quantifying ecological memory in plant and ecosystem processes. *Ecol. Letters*, 18: 221-235.

Pfautsch, S., J. Renard, M. G. Tjoelker, and A. Salih. 2015. Phloem as capacitor: Radial transfer of water into xylem of tree stems occurs via symplastic transport in ray parenchyma. *Plant Physiol.*, 167, 3, 963–971.

Pomeroy, J. and D. Gray. 1995. Snowcover accumulation, relocation and management. National Hydrology Research Institute, Saskatoon, Canada, 144 pp.

Razavi, S., A. Elshorbagy, H. Wheater, and D. Sauchyn. 2015. Toward understanding nonstationarity in climate and hydrology through tree ring proxy records, *Water Resour. Res*., 51, doi:10.1002/2014WR015696.

Razavi, S., A. Elshorbagy, H. Wheater, and D. Sauchyn. 2016. Time scale effect and uncertainty in reconstruction of paleo-hydrology. *Hydrol. Process.,* doi: 10.1002/hyp.10754.

Rinaldo, A., P. Benettin, C. J. Harman, M. Hrachowitz, K. J. McGuire, Y. van der Velde, E. Bertuzzo, and G. Botter. 2015. Storage selection functions: A coherent framework for quantifying how catchments store and release water and solutes, *Water Resour. Res*., 51, 4840–4847, doi:10.1002/2015WR017273.

Salas, J., B. Rajagopalan, L. Saito and C. Brown, 2012. Special Section on Climate Change and Water Resources: Climate Nonstationarity and Water Resources Management. *J. Water Resour. Plann. Manage.*, 138(5): 385-388.

Sauchyn, D., J. Vanstone, and C. Perez-Valdivia. 2011. Modes and forcing of hydroclimatic variability in the Upper North Saskatchewan River Basin since 1063. *Can. Water Resour. J.*, 36(3): 205-217.

Sauchyn, D., J. Vanstone, J.-M. St. Jacques and R. Sauchyn. 2015. Dendrohydrology in Canada's western interior and applications to water resource management. *J. Hydrol*., 529, 548-558.

Sawicz, K., T. Wagener, M. Sivapalan, P. Troch and G. Carrillo. 2011. Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrol. Earth Syst. Sci*., 15, 2895–2911.

Sawicz, K., C. Kelleher, T. Wagener, P. Troch, M. Sivapalan and G. Carrillo. 2014. Characterizing hydrologic change through catchment classification. *Hydrol. Earth Syst. Sci*., 18, 273-285.

Singh, A., L. Reinhardt, and E. Foufoula-Georgiou. 2015. Landscape reorganization under changing climatic forcing: Results from an experimental landscape, *Water Resour. Res.*, 51, 4320–4337, doi:10.1002/2015WR017161.

Sørensen, R., U. Zinko, and J. Seibert. 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrol. Earth. Sys. Sci*., 10, 101–112.

Starheim, C. C. A., D. J. Smith, and T. D. Prowse. 2013. Dendrohydroclimate reconstructions of July–August runoff for two nival-regime rivers in west central British Columbia. *Hydrol. Process.*, 27, 405–420. doi:10.1002/hyp.9257.

Tei, S., H. Yonenobu, A. Sugimoto, T. Ohta, and T. C. Maximov. 2015. Reconstructed summer Palmer Drought Severity Index since 1850 AD based on $\delta^{13}C$ of larch tree rings in eastern Siberia. *J. Hydrol*., 529, 442-448.

Wagener, T., M. Sivapalan, P. A. Troch, B. L. McGlynn, C. J. Harman, H. V. Gupta, P. Kumar, P. S. C. Rao, N. B. Basu, and J. S. Wilson. 2010. The future of hydrology: An evolving science for a changing world, Water Resour. Res., 46, W05301, doi:10.1029/2009WR008906.

Till, C. and J. Guiot. 1990. Reconstruction of precipitation in Morocco since 1100 A.D. Based on *Cedrus atlantica* tree-ring widths. *Quatern. Res.*, 33(3), 337-351.

Urrutia, R. B., A. Lara, R. Villalba, D. D. Christie, C. Le Quesne, and A. Cuq. 2011. Multicentury tree ring reconstruction of annual streamflow for the Maule River watershed in south central Chile. *Water Resour. Res.*, 47(6): W06527.

Watson, T. A., F. A. Barnett, S. T. Gray, and G. A. Tootle. 2009. Reconstructed streamflows for the headwaters of the Wind River, Wyoming, United States, *J. Am. Water Resour. Assoc*., 45(1), 224–236.

Wise, E. K. 2010. Tree ring record of streamflow and drought in the upper Snake River, *Water Resour. Res*., 46, W11529, doi:10.1029/2010WR009282.

Woodhouse, C. A., G. T. Pederson, K. Morino, S. A. McAfee, and G. J. McCabe. 2016. Increasing influence of air temperature on upper Colorado River streamflow, *Geophys. Res. Lett*., 43, 2174–2181, doi:10.1002/2015GL067613.

Woodhouse, C. A. and J. J. Lukas. 2006. Multi-century tree-ring reconstructions of Colorado streamflow for water resource planning. *Clim. Change*, 78(2-4), 293–315.

Table 1. List of chronology sites (*TRCs*) in the Oldman River Basin, tree species, and data availability. The topographic index reflects the potential wetness of the site, the wetter the site, the higher the index value.

No	Chronology Site Code	Tree Species[a]	Period years	Elevation (m)	Topographic index
1	WCK	PM	1750-2005	1536	9.4
2	CAL	PM	1640-2004	1677	7.5
3	BMN	PF	1580-2007	1297	10.2
4	LBC	PM	1610-2004	1602	6.0
5	WSC	PM	1570-2004	1575	7.4
6	OMR	PM	1370-2007	1331	6.0
7	DCK	PM	1660-2004	1648	6.1
8	BDC	PM	1550-2004	1661	6.0
9	BCK	PM	1660-2006	1592	6.5

10	OMR	PF	1640-2003	1427	4.6
11	OCP	PC	1790-2003	1280	9.7
12	CAB	PM	1440-2004	1395	4.9
13	HEM	PF	1510-2007	1308	6.5
14	ELK	PF	1540-2004	1384	6.1
15	BVL	PM	1730-2004	1567	4.9
16	TAB	LL	1616-2010	1838	4.7

[a]PM, *Pseudotsuga menziesii*; PF, *Pinus flexilis*; PC, *Pinus contorta*; LL, *Larix lyallii*

Table 2. Values of the calibration parameters of the conceptual model developed.

c	Value	D	value
c_1^*	0.40	d_1^*	0.40
c_2	0.45	d_2	0.25
c_3	0.45	d_3	0.30
		d_4	0.20
		d_5	0.20
		d_6**	0.10

*The first month in the spring when the average monthly temperature is higher than zero (e.g., April).
**The coefficient value of the sixth month remains unchanged for any subsequent month with temperature higher than zero.

Table 3. Correlation between *TRCs* and both instantaneous (t) and 2-year average ("t" + "t+1") hydrological processes (1953-2001). The concept of 2-year average improves the correlation in case of P and ET.

Site	Precipitation (*P*)		Runoff (*Q*)		Evapotranspiration (*ET*)	
	P_t	$P_t + P_{t-1}$	Q_t	$Q_t + Q_{t-1}$	ET_t	$ET_t + ET_{t-1}$
CAB_t	0.45	0.62	0.61	0.51	0.33	0.58
CAL_t	0.45	0.60	0.58	0.60	0.35	0.51
OMR_t	0.47	0.55	0.61	0.51	0.33	0.49
BVL_t	0.53	0.71	0.47	0.34	0.44	0.71

Table 4. Instantaneous correlation between *TRCs* and *P* or *Q* for wet and dry years. Substantial correlation exists between TRC and wet precipitation and dry runoff years. Dry precipitation and wet runoff years do not show similar correlation with TRCs.

***TRC* site**	**Wet *P* (28 years)**	**Dry *P* (34 years)**	**Wet *Q* (41 years)**	**Dry *Q* (48 years)**
BVL	0.27	0.05	-0.10	0.30
CAB	0.55	0.01	0.21	0.43
CAL	0.42	0.13	0.18	0.40
OMR	0.42	0.05	0.16	0.34

Table 5. Performance measures of the regression-based reconstruction models. R^2 is the coefficient of determination and RE is the reduction of error statistic.

TRC site	Precipitation (*P*)		Runoff (*Q*)	
	R^2	RE	R^2	RE
BVL	**0.50**	**0.63**	0.22	0.20
CAB	0.39	0.52	**0.37**	**0.38**
CAL	**0.36**	**0.62**	0.34	0.34
OMR	0.30	0.33	**0.38**	**0.51**

* Bold numbers represent the best sites.

Table 6. Reconstructed relative change in the long-term mean of *P*, *Q*, and *ET* in the OMRB. Different TRCs lead to different conclusions with regard to the wetness and dryness of the pre-instrumental period.

TRC site	Percent change relative to instrumental period		
	P	*Q*	*ET*
BVL	-5.30%	-5.90%	-6.00%
CAB	0.80%	2.20%	1.80%
CAL	2.70%	5.00%	3.40%
OMR	-0.80%	0.00%	0.10%

Table 7. Statistical properties of observed and modeled hydrological variables during the instrumental period.

	Site	BVL		CAB		CAL		OMR	
		Obs.	Mod.	Obs.	Mod.	Obs.	Mod.	Obs.	Mod.
P	Mean	397	397	397	397	397	397	397	401
	Std.	72	47	72	45	72	43	72	39
	CV	0.18	0.12	0.18	0.11	0.18	0.11	0.18	0.10
	ρ	0.56	0.36	0.56	0.30	0.56	0.44	0.56	0.42
Q	Mean	126	126	126	126	126	126	126	126
	Std.	37	17	37	23	37	22	37	23
	CV	0.30	0.14	0.30	0.18	0.30	0.17	0.30	0.18
	ρ	0.23	0.37	0.23	0.29	0.23	0.44	0.23	0.41
ET	Mean	260	260	260	260	260	260	260	260
	Std.	51	37	51	30	51	26	51	25
	CV	0.19	0.14	0.19	0.11	0.19	0.10	0.19	0.09
	ρ	0.52	0.39	0.52	0.30	0.52	0.44	0.52	0.42

Obs.: observed values; Mod.: Based on regression models; Std.: standard deviation; CV: coefficient of variation; ρ: autocorrelation coefficient.

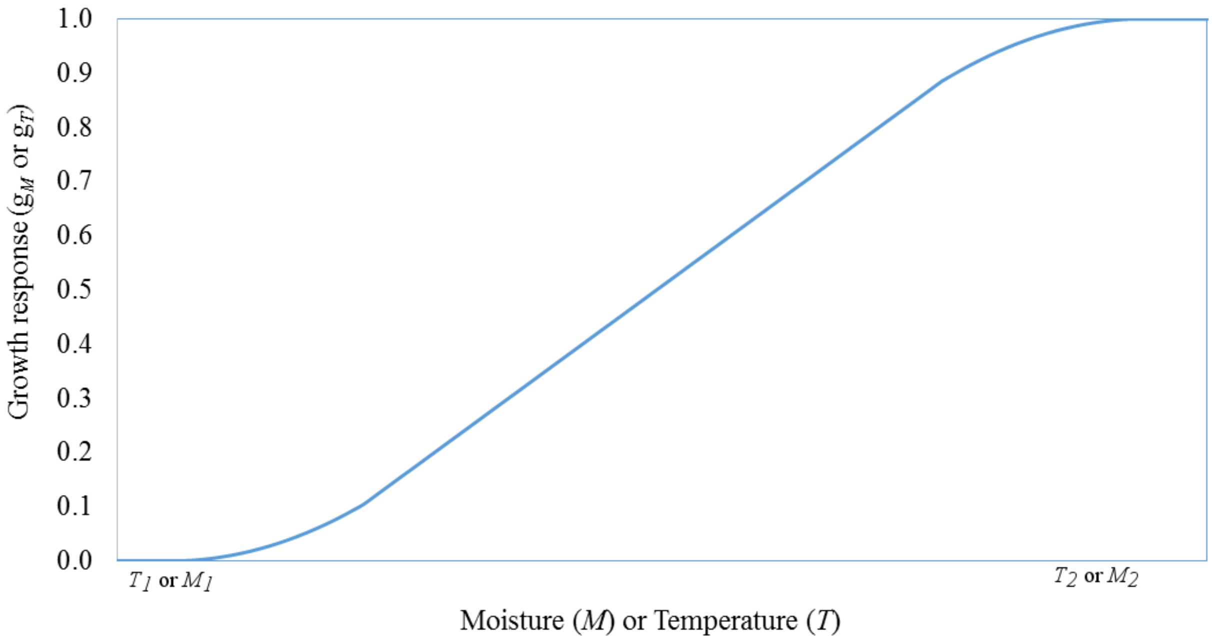


Figure 1. Generic representation of tree growth as a function of moisture (M) and temperature (T) thresholds – g_M and g_T are growth response to moisture and temperature, respectively, and T_1 (M_1) and T_2 (M_2) are the lower and upper temperature (moisture) thresholds, respectively.

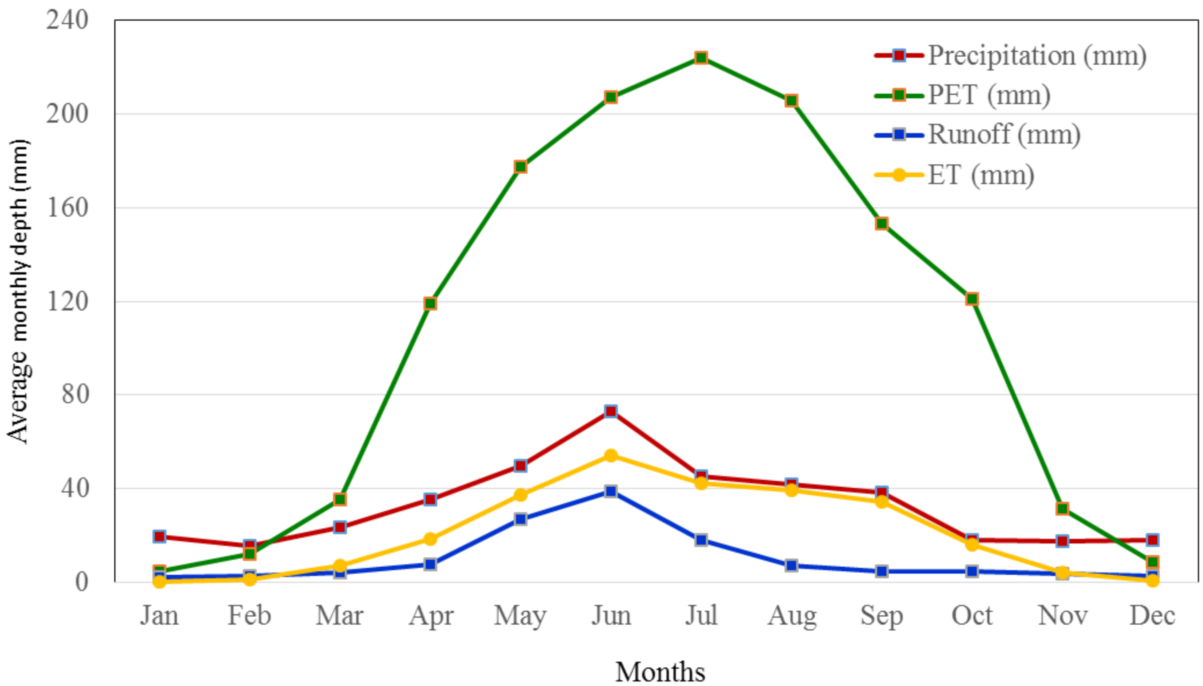


Figure 2. Average monthly values of water balance components in the Oldman River Basin (1953-2001). Precipitation and runoff are based on measured values; PET is estimated using Penman-Monteith method, and actual evapotranspiration (ET) is predicted using the conceptual model developed in this study (introduced later in Section 5.2).

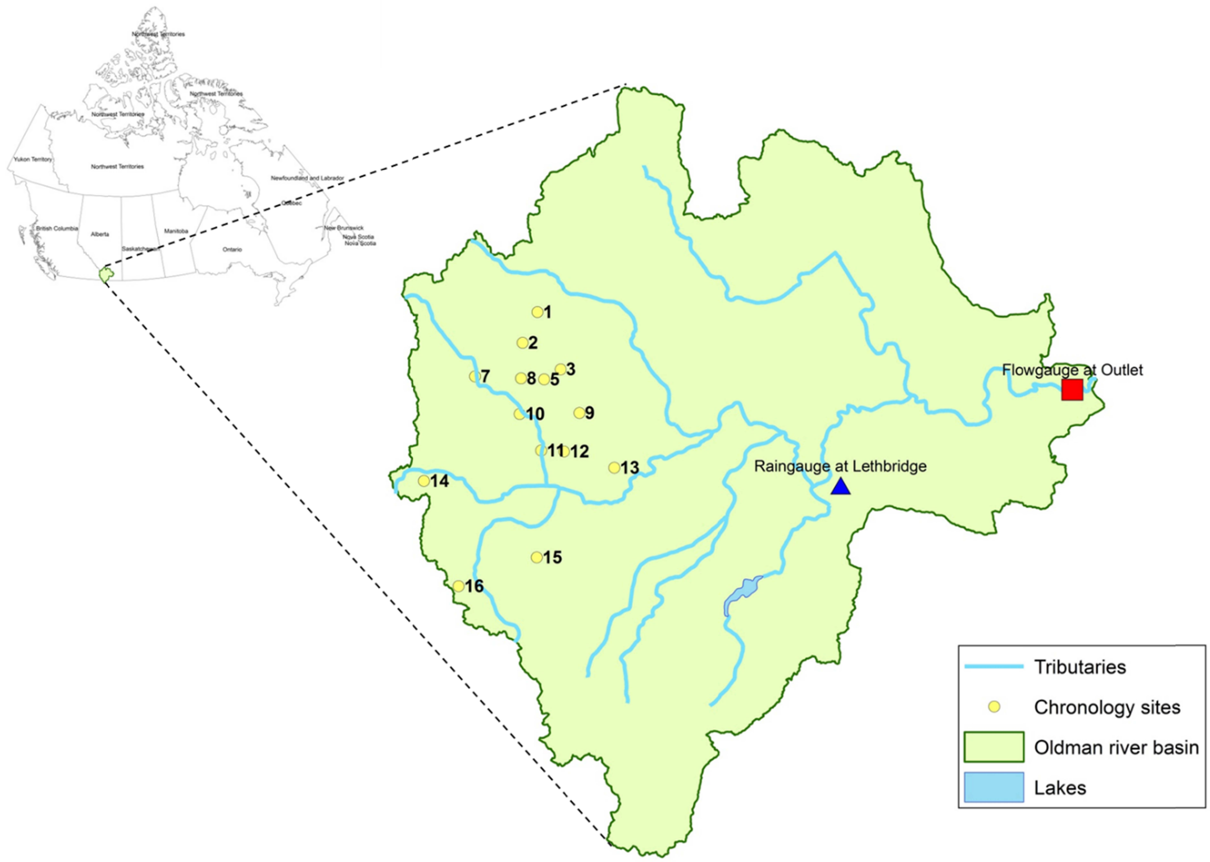


Figure 3. The Oldman River Basin in southwestern Alberta, Canada. 16 TRC sites (listed in Table 1) are shown along with the Precipitation gauge and the runoff sites. Sites 4, 6, and 8 are in the same area but at different elevations.

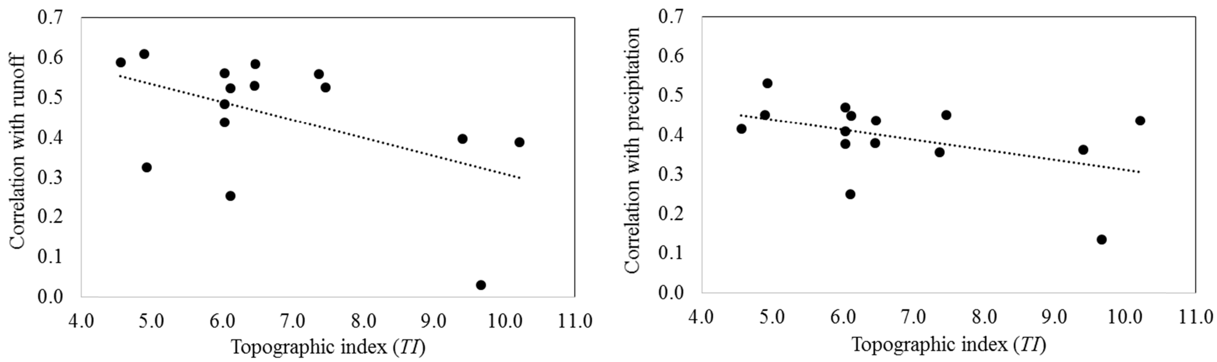


Figure 4. Topographic index and correlation between *TRCs* and hydrological variables at various sites in the OMRB. There is obvious trends showing that the higher the topographic index values (potentially wetter sites), the lower the correlation coefficient values between TRCs and hydrological variables.

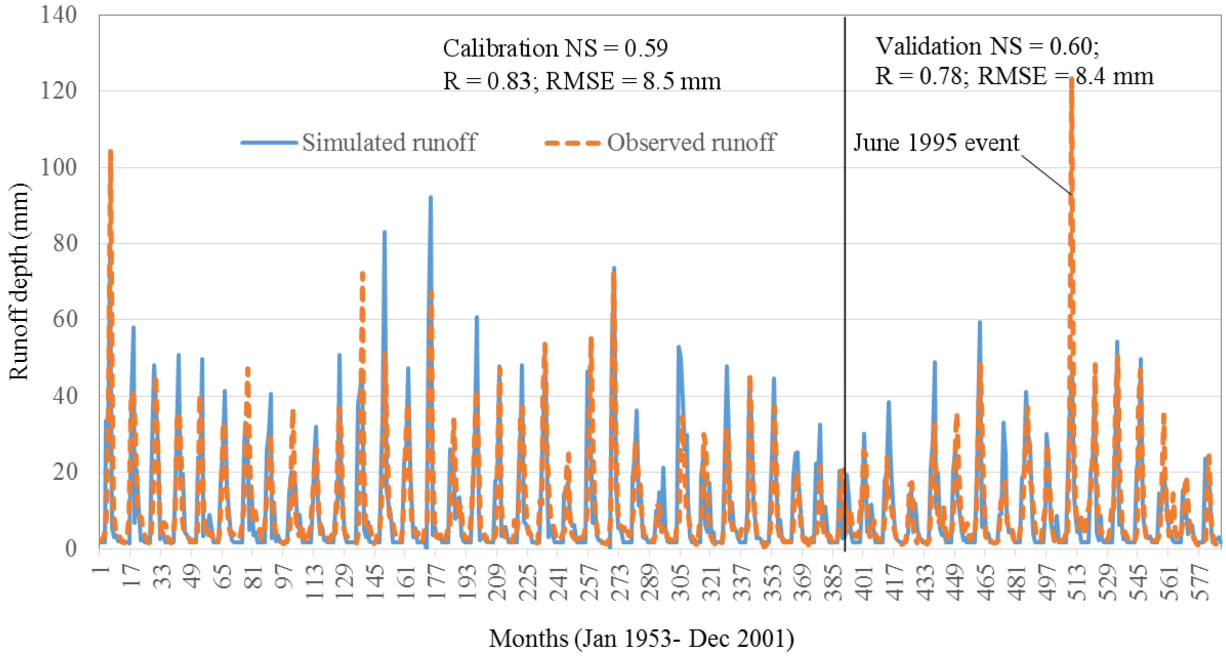


Figure 5. Observed and simulated monthly runoff at the outlet of the OMRB, Alberta, Canada. The developed conceptual model performs well in predicting the monthly runoff values with Nash-Sutcliffe value of 0.60 and correlation coefficient of 0.78.

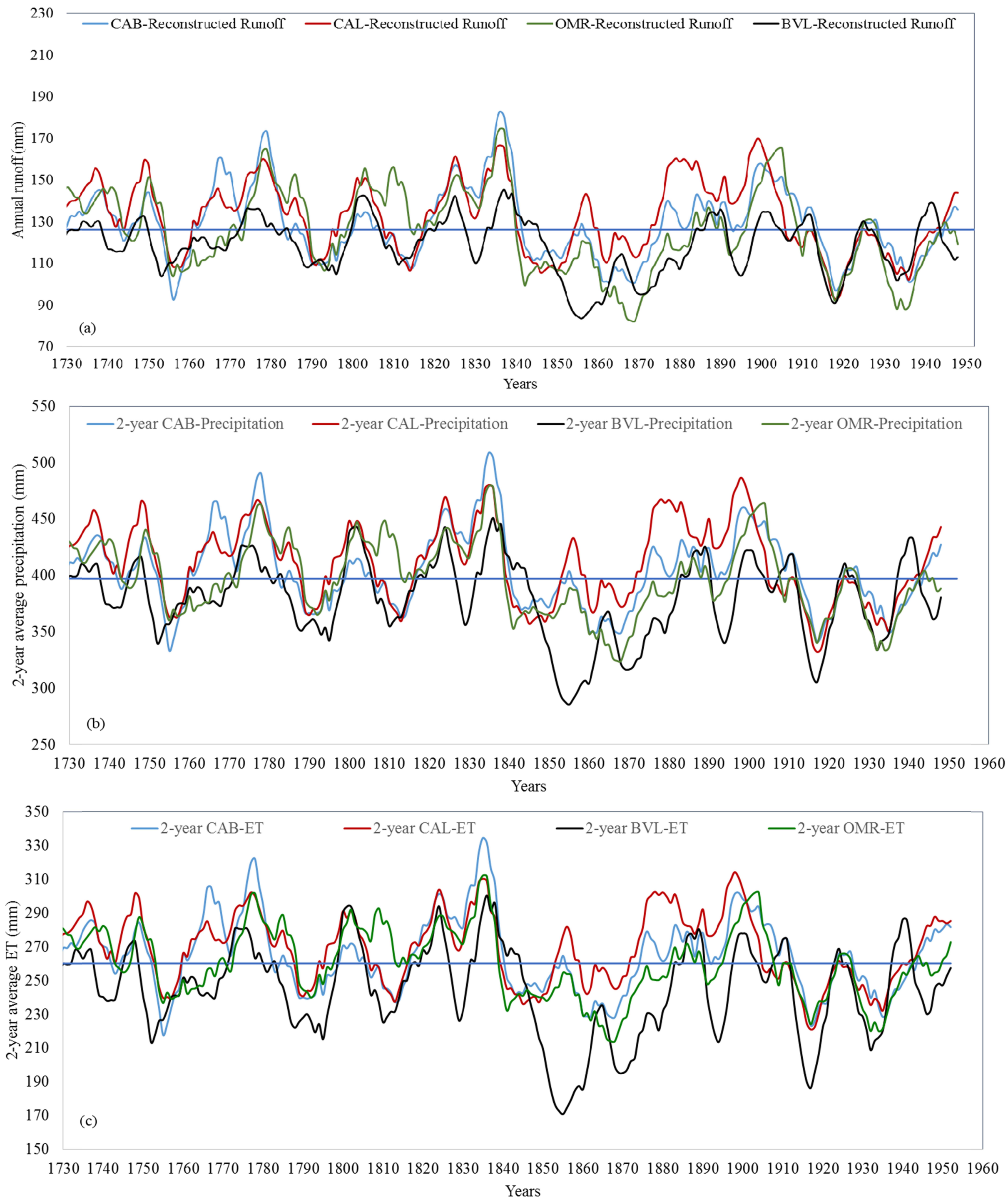


Figure 6. 5-year running average of (a) reconstructed annual runoff,(b) 2-year average precipitation, and (c) 2-year average evapotranspiration for the OMRB. The horizontal line shows the mean value based on the instrumental period (1955-2001).

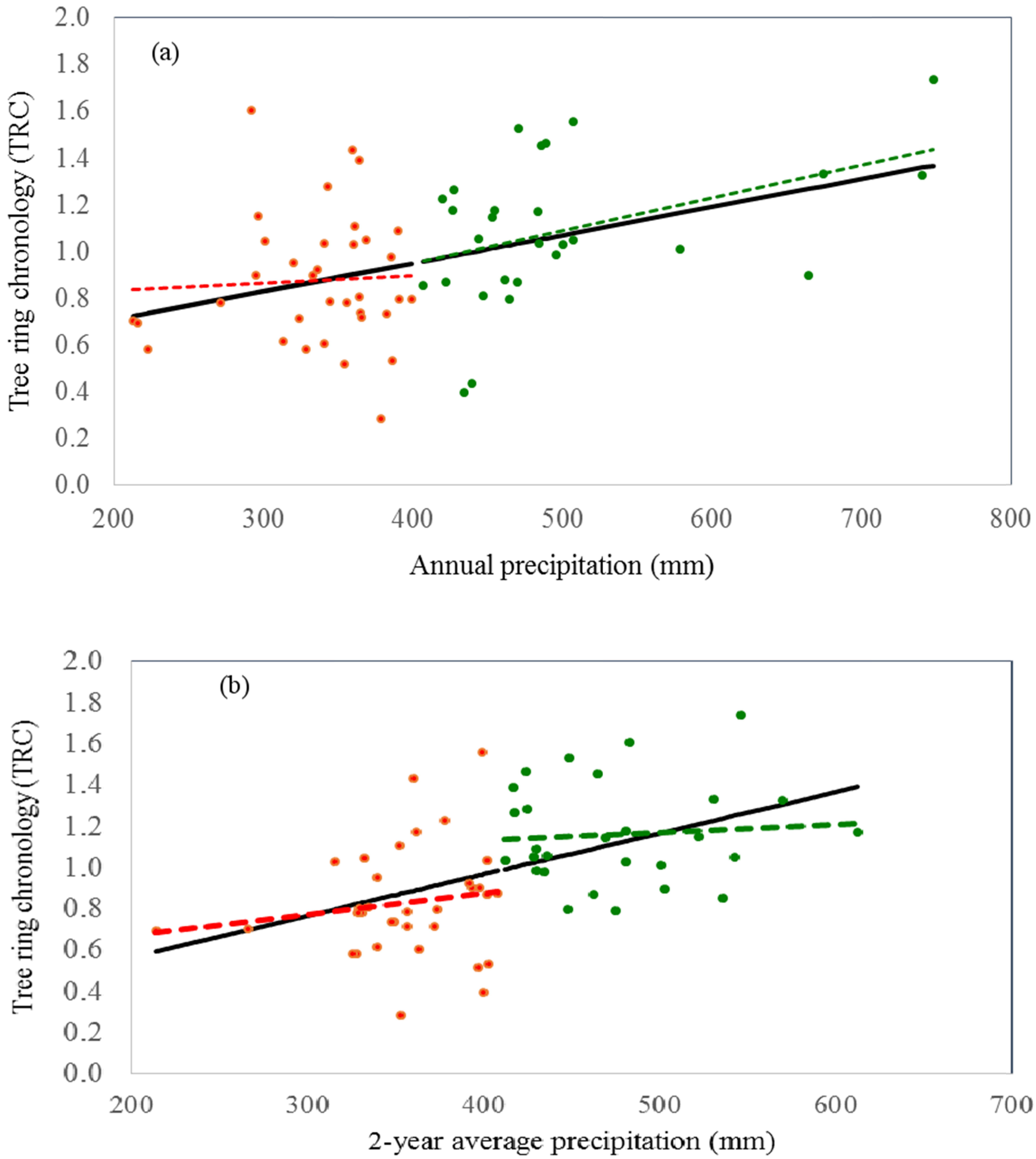
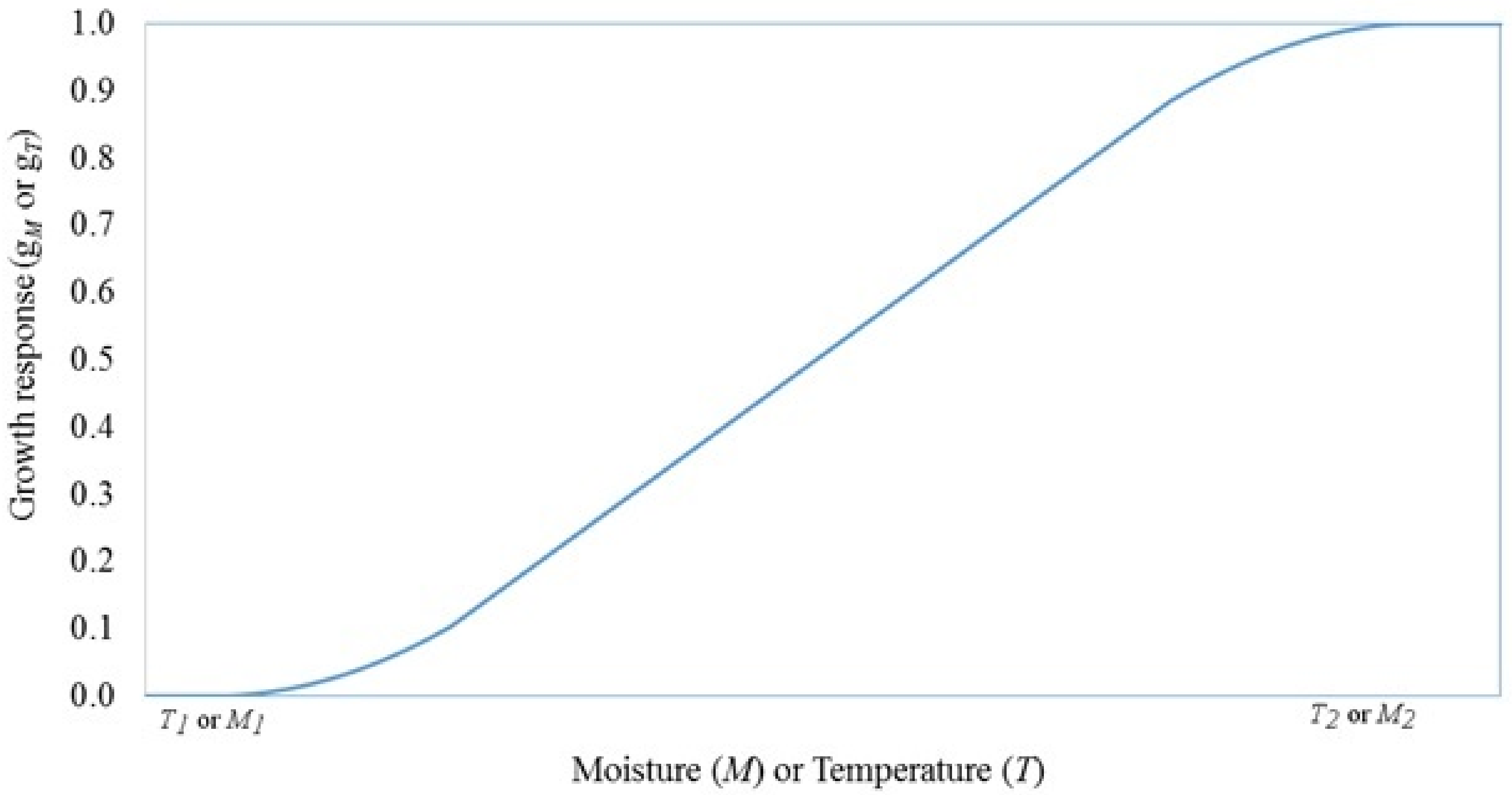


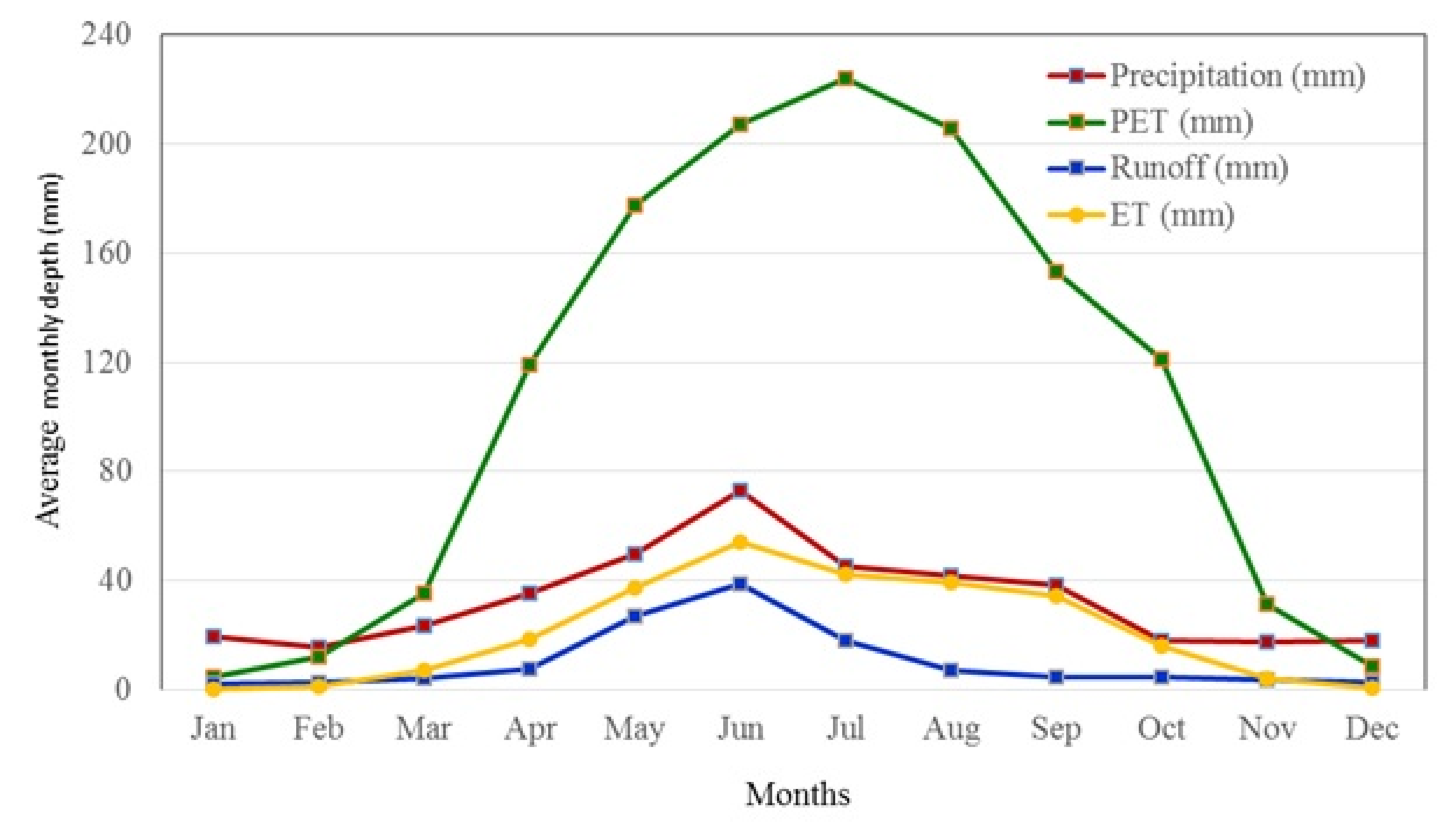
Figure 7. Overall trend as well as the trend lines over sub-regions of wet and dry years of (a) annual precipitation and (b) 2-year average precipitation with respect to *TRCs* from the OMR site. There are obvious differences between the trend based on the entire record and zonal trends based on wet and dry years separately.

Figure 1. Figure



1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Growth response (g_M or g_T)
T_1 or M_1
T_2 or M_2
Moisture (M) or Temperature (T)

Figure 2. Figure



240
200
160
120
80
40
0
Average monthly depth (mm)
Precipitation (mm)
PET (mm)
Runoff (mm)
ET (mm)
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Months

Figure 3. Figure

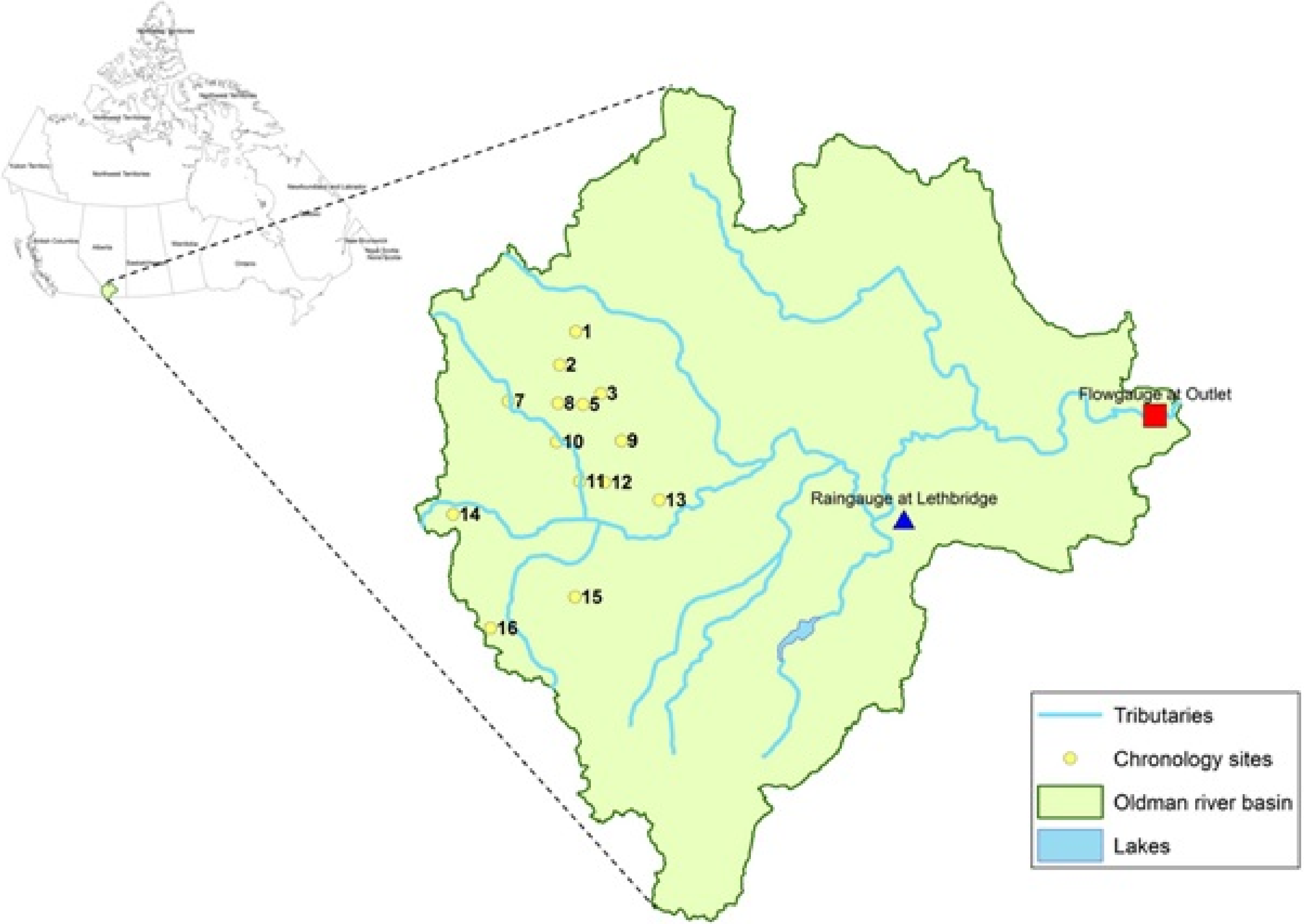
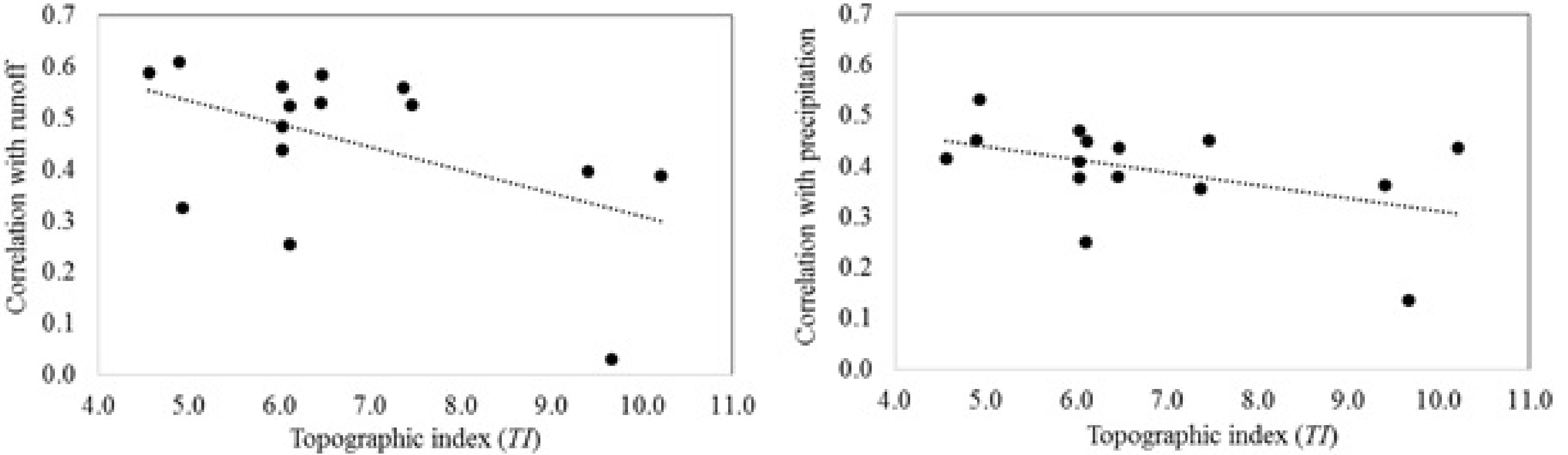

1
2
3
5
7
8
9
10
11
12
13
14
15
16
Flowgauge at Outlet
Raingauge at Lethbridge
Tributaries
Chronology sites
Oldman river basin
Lakes

Figure 4. Figure



Correlation with runoff
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
Topographic index (TI)
Correlation with precipitation
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
Topographic index (TI)

Figure 5. Figure

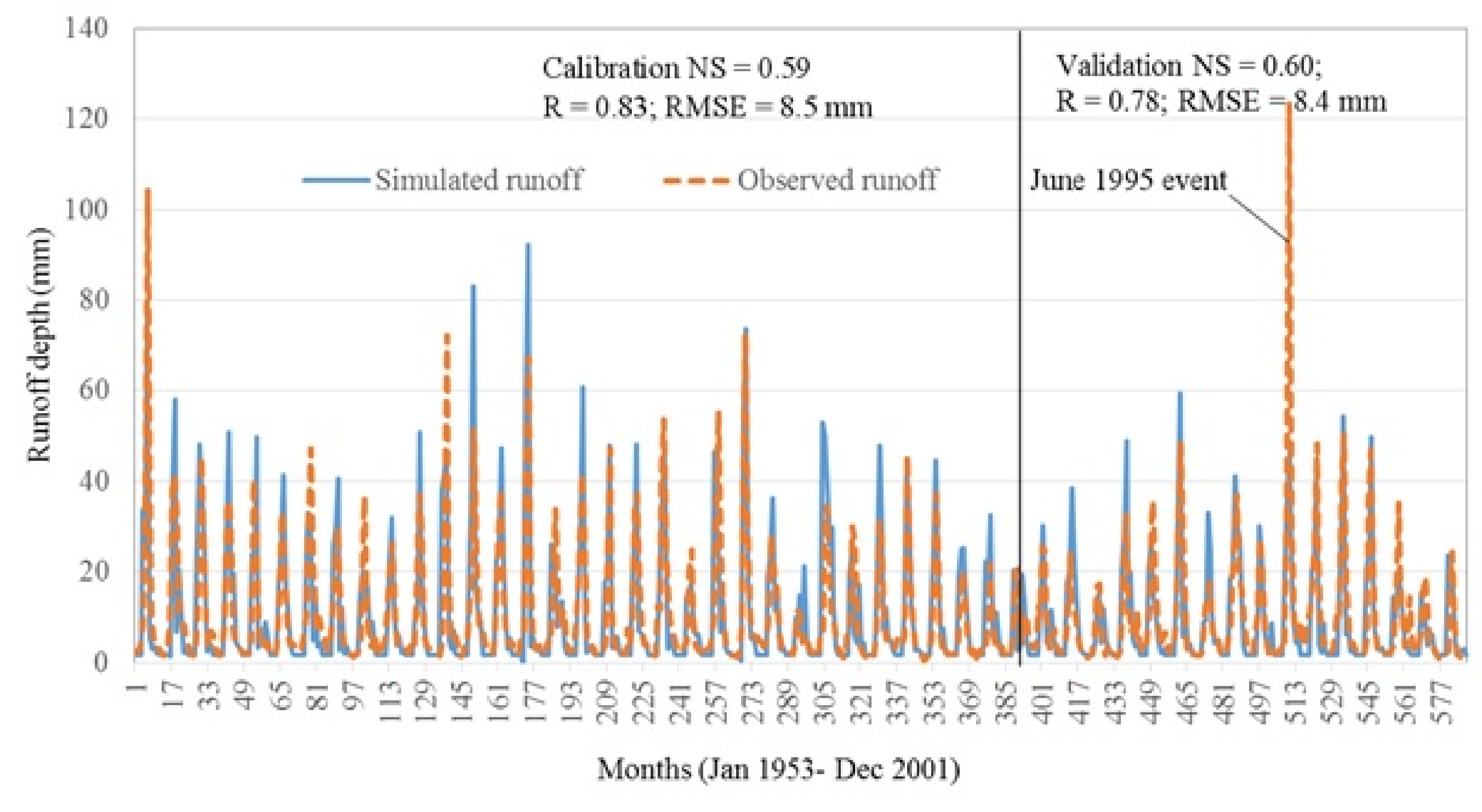

Calibration NS = 0.59
R = 0.83; RMSE = 8.5 mm
Validation NS = 0.60;
R = 0.78; RMSE = 8.4 mm
Simulated runoff
Observed runoff
June 1995 event
Runoff depth (mm)
Months (Jan 1953- Dec 2001)
0
20
40
60
80
100
120
140
1
17
33
49
65
81
97
113
129
145
161
177
193
209
225
241
257
273
289
305
321
337
353
369
385
401
417
433
449
465
481
497
513
529
545
561
577

Figure 6. Figure

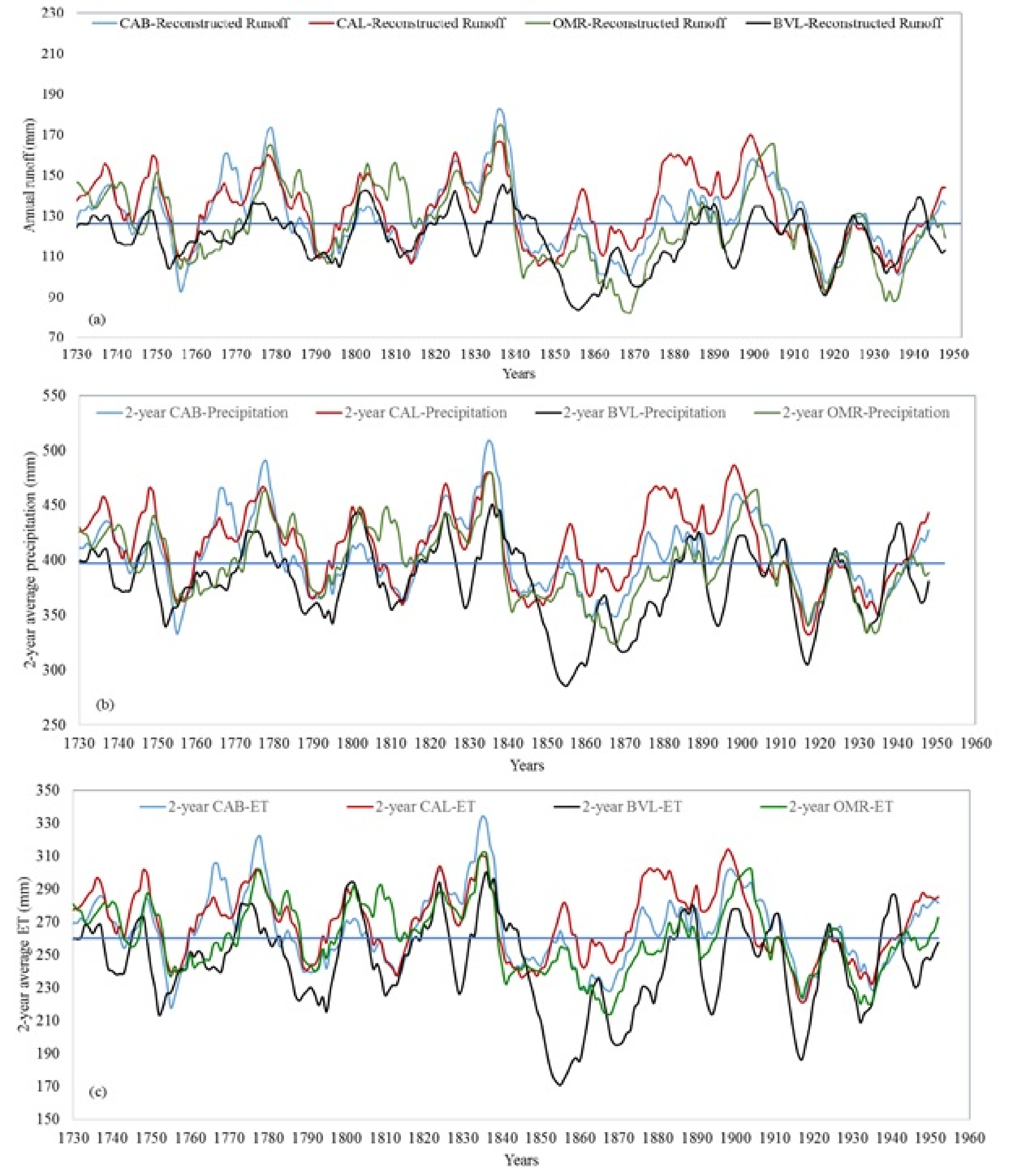
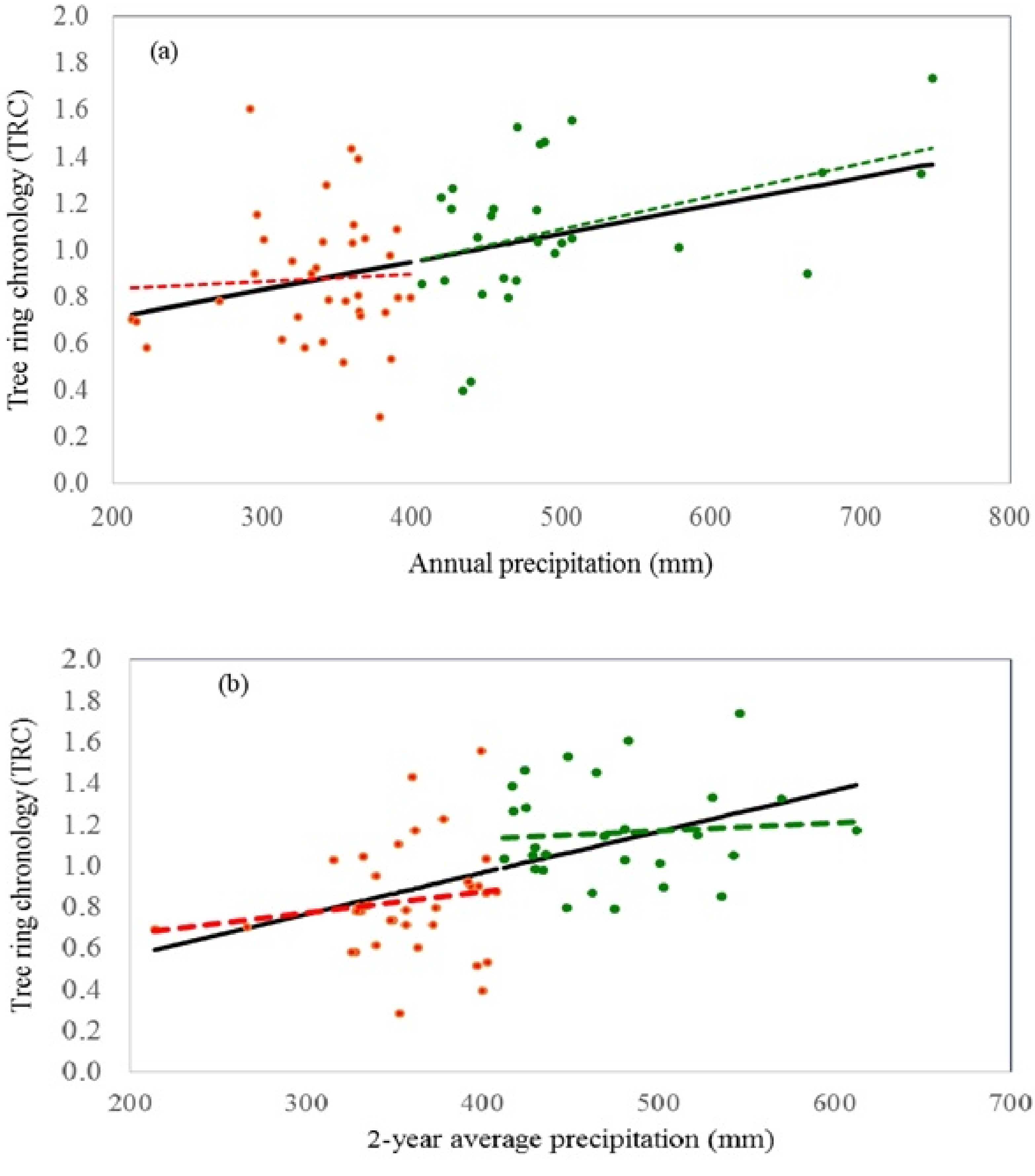

CAB-Reconstructed Runoff
CAL-Reconstructed Runoff
OMR-Reconstructed Runoff
BVL-Reconstructed Runoff
Annual runoff (mm)
230
210
190
170
150
130
110
90
70
(a)
1730 1740 1750 1760 1770 1780 1790 1800 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950
Years
2-year CAB-Precipitation
2-year CAL-Precipitation
2-year BVL-Precipitation
2-year OMR-Precipitation
2-year average precipitation (mm)
550
500
450
400
350
300
250
(b)
1730 1740 1750 1760 1770 1780 1790 1800 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960
Years
2-year CAB-ET
2-year CAL-ET
2-year BVL-ET
2-year OMR-ET
2-year average ET (mm)
350
330
310
290
270
250
230
210
190
170
150
(c)
1730 1740 1750 1760 1770 1780 1790 1800 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950 1960
Years

Figure 7. Figure



(a)
Tree ring chronology (TRC)
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
200
300
400
500
600
700
800
Annual precipitation (mm)
(b)
Tree ring chronology (TRC)
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
200
300
400
500
600
700
2-year average precipitation (mm)