Journal of Hydrology 380 (2010) 305–317

Contents lists available at ScienceDirect

Journal of Hydrology

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

Regional estimation of extreme suspended sediment concentrations using watershed characteristics

Yves Tramblay *, Taha B.M.J. Ouarda, André St-Hilaire, Jimmy Poulin

Chair in Statistical Hydrology, INRS-ETE, 490 Rue de la Couronne, Québec, Canada G1K 9A9

ARTICLE INFO

Article history:
Received 25 June 2008
Received in revised form 2 October 2009
Accepted 10 November 2009

This manuscript was handled by L. Charlet, Editor-in-Chief, with the assistance of Jose D. Salas, Associate Editor

Keywords:
Frequency analysis
Regionalisation
Suspended sediment
Concentrations
Water quality

SUMMARY

The number of stations monitoring daily suspended sediment concentration (SSC) has been decreasing since the 1980s in North America while suspended sediment is considered as a key variable for water quality. The objective of this study is to test the feasibility of regionalising extreme SSC, i.e. estimating SSC extremes values for ungauged basins. Annual maximum SSC for 72 rivers in Canada and USA were modelled with probability distributions in order to estimate quantiles corresponding to different return periods. Regionalisation techniques, originally developed for flood prediction in ungauged basins, were tested using the climatic, topographic, land cover and soils attributes of the watersheds. Two approaches were compared, using either physiographic characteristics or seasonality of extreme SSC to delineate the regions. Multiple regression models to estimate SSC quantiles as a function of watershed characteristics were built in each region, and compared to a global model including all sites. Regional estimates of SSC quantiles were compared with the local values. Results show that regional estimation of extreme SSC is more efficient than a global regression model including all sites. Groups/regions of stations have been identified, using either the watershed characteristics or the seasonality of occurrence for extreme SSC values providing a method to better describe the extreme events of SSC. The most important variables for predicting extreme SSC are the percentage of clay in the soils, precipitation intensity and forest cover.

Introduction and literature review

High suspended sediment concentrations (SSC) are harmful to certain species of fish and aquatic organisms, increase the cost of drinking water treatment and possibly carry large amounts of pollutants (Waters, 1995). High SSC affect the biota by reducing the density, productivity and abundance of primary producers and macro invertebrates (Wood and Armitage, 1997). Concentrations above 80 mg/L may affect some fish populations, and concentrations above 200 mg/L are assumed to be harmful to most North American fish (Waters, 1995; Newcombe and Jensen, 1996).

From 1982 to 1998, the number of sediment monitoring stations in the USA has decreased by 65% (Gray and Glysson, 2002) and the same trend is observable for stations in Canada (Day, 1992). Meanwhile, the need for reliable, cost-effective, spatially and temporally consistent sediment data has never been greater for engineering, environmental and regulatory considerations since suspended sediment is increasingly considered to be a key water quality variable (Gray and Glysson, 2002; Meybeck et al., 2003). To compensate for this lack of measurements, it is necessary to develop reliable estimation methods for suspended sediment transport.

Estimation of SSC is often done with rating curves that establish an empirical relation between sediment concentration and discharge. Several studies have shown that rating curve models tend to underestimate high SSC values (Fergusson, 1986; Asselman, 2000) even when a bias correction method is applied (Walling and Webb, 1988). Improvements in estimating SSC have been noted when data sets were subdivided into seasonal or hydrological groupings (Asselman, 2000; Horowitz, 2003) or with the use of truncated rating curves relating only the highest quantiles of SSC and river flow (Meybeck et al., 2003; Simon et al., 2004). Inaccurate prediction of extreme SSC can be related to the scattered relation that could exist between sediment and discharge caused by hysteresis effect (Smith and Croke, 2005). Weak correlations between SSC and discharge have been reported in several studies. Tramblay et al. (2008) have shown that correlation between annual maximum SSC and corresponding discharge (taking into account the possible lag time caused by hysteresis) was significant in only 92 out of 208 rivers in North America. These results suggest that for many rivers, suspended sediment transport is rather supply limited than discharge limited. Discharge alone is therefore not sufficient to estimate magnitude of extreme SSC in numerous rivers.

Aside from discharge, several studies have established significant correlations between some watershed characteristics (including topography, geology, climate, and land use) and suspended sediment flux (Slaymaker, 1982; Bray and Xie, 1993; Ludwig and

* Corresponding author. Tel.: +1 418 654 2524; fax: +1 418 654 2600.
E-mail address: ytramblay@gmail.com (Y. Tramblay).

0022-1694/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2009.11.006

Probst, 1998; Restrepo et al., 2006), mean SSC values (Jarvie et al., 2002; Robertson et al., 2006; Siakeu et al., 2004; Dodds and Whiles, 2004) or extreme SSC (Tramblay et al., 2007). Robertson et al. (2006) used a regression tree analysis to create groups of rivers in the Great Lakes region in order to estimate several water quality parameters, including suspended sediments. Jarvie et al. (2002) constructed multiple regression models to explore the linkages between land characteristics and SSC in central England. To our knowledge no studies used this approach for extreme SSC.

An alternative way of estimating extreme SSC can be considered, based on a probabilistic approach. Frequency analysis is a statistical approach commonly used in hydrology to relate the magnitude of extreme events (e.g. floods or low flows) to a probability of occurrence (Rao and Hamed, 2001). Only a few studies used a probabilistic modelling approach for SSC. Van Sickle (1982) developed a peak-over-threshold Poisson model for annual sediment flux for two small Oregon streams. Watts et al. (2003) computed exceedance probabilities of SSC for 1–6-day durations using peaks-over-threshold techniques for reaches of the Lower Swale, UK. Simon et al. (2004) estimated suspended sediment transport conditions at the 1.5-year recurrence interval for rivers in the USA. Galéa et al. (2004) proposed the transposition of the discharge–duration–frequency analysis concept (or QdF) to the wash load in the Bega sub-basin in Romania. Soler et al. (2007) fitted the log-normal distribution to annual maximums and partial duration series of SSC for the Vallcebre basin in Spain. Application of the frequency analysis approach and results for annual maxima of SSC in 179 rivers of North America are detailed in Tramblay et al. (2008).

In this study, the feasibility of estimating extreme SSC at a site in the absence of measurements is tested using some approaches developed originally for the regionalisation of floods. The goal of flood regionalisation is to estimate flood quantiles at ungauged catchments using the watershed characteristics (GREHYS, 1996). The two main steps are the identification of groups of hydrologically homogeneous basins and the regional estimation within each delineated region to estimate flood characteristics at the site of interest. In this study, two approaches are compared to identify hydrological homogeneous regions. The first approach is based on grouping catchments according to their physiographic similarity. This approach is commonly used in the regionalisation of floods (Nathan and McMahon, 1990; GREHYS, 1996). The second approach tested is based on the similarity in seasonality of occurrence of extreme SSC. Regionalisation methods based on seasonality have been recently gaining increased popularity among hydrologists (Ouarda et al., 2006). In these seasonal regional models, the delineation of homogeneous regions is based on the seasonal behaviour of flood flows in the various stations (Burn, 1997; Ouarda et al., 2006). The two main objectives of this study are:

(1) To identify the most significant physiographic variables affecting extreme SSC. Subsequently they may be used in regionalisation methods.
(2) To compare the estimation performance of two regionalisation models versus a unique (i.e. encompassing all study sites) for extreme SSC. The first approach considers the physiographic characteristics of the watershed and the second uses seasonality of extreme SSC to delineate regions.

Study area and data collection

Suspended sediment data

Daily SSC data from 140 gauging stations in North America constitute the basic data base of the project (2505 station-years). In North America, the annual maximum SSC occurs more often in spring, except for rivers located in California and Rio Grande or Colorado systems (Tramblay et al., 2008). Stations where annual maxima of SSC occur in the spring were chosen because they represent the largest population of stations having long records (10 years or more) of SSC in North America (Tramblay et al., 2008). Peaks of SSC in spring are usually associated with snowmelt but could also be produced by rainfall events on soils not yet protected with a fully grown vegetation cover (Lecce et al., 2006). High concentrations in summer months are generally caused by thunderstorms with high rainfall intensities (Meade et al., 1990; Lecce et al., 2006).

In Canada, data were retrieved from the Environment Canada HYDAT Database. In the USA, data were provided by the US Geological Survey (USGS) Sediment Database. The uniform data collection methods developed by the USGS in the USA, also used in Canada, provide comparable data. The database has been screened for errors and is available on the Internet (http://co.water.usg.gov/sediment for USA, http://www.wsc.ec.gc.ca/products/hydat/main_e.cfm?cname=hydat_e.cfm for Canada). The daily mean concentration is a time-weighted mean value; the sampling frequency is changing from one stream to another depending on its size, location and climatic zone. The frequency is increased during high flow periods, but hydrologic sampling can be complicated and hazardous during extreme events. Detailed information about sampling protocols and data collection can be found in Edwards and Glysson (1988). Some data on bed material and suspended-sediment size is also available, but these data are periodic in nature and not available for all stations. Fig. 1 shows the number of stations with SSC records available for each year. Most of the data are available for the 1970s and 1980s. Subsequently, a major decline of the number of available stations is observed in the USA and Canada (see Fig. 2).

Selected stations

The selection criteria for stations were record length, number of missing data, hypotheses test results, watershed size and regulation level. All stations have 10 years or more of daily SSC data. Drainage area ranges between 20 km^2 and 200 000 km^2. These stations cover a wide range of climates and landscapes. Average record length is 17 years and median record length is 15 years.

Data series were screened to ensure they include no more than 20 missing days during the "spring season" (February–July). Stations on basins with major dams or reservoirs were excluded since sediment transport can be greatly affected by regulation in rivers (Walling and Fang, 2003; Walling, 2006). Metadata from Environment Canada HYDAT Database contain information about regulated rivers in Canada. In the USA, the database of the National Inventory of Dams (NID) from the US Army Corps of Engineers was used to detect the presence of large dams or reservoirs on the main stem in the selected rivers. Time series of flow discharge were also analyzed to detect major shifts in the data that could be caused by dam construction, using the Wilcoxon and Kendall non-parametric tests. Thus, 38 stations were excluded because of evidence of strong flow regulation.

Annual maxima of SSC occurring in spring (February–July) were extracted from the data, and the hypotheses of homogeneity, stationarity and independence were verified since they are a prerequisite for frequency analysis. The non-parametric tests of Wilcoxon for homogeneity, Wald–Wolfowitz for autocorrelation and Kendall for stationarity were used at the 1% significance level. Fifteen stations had a significant trend detected by the Kendall test and eight series showed autocorrelation detected by the Wald–Wolfowitz test. Homogeneity was tested to detect shifts in the chronological distribution as well as on a seasonal basis to test whether events of early spring (February–April), had the same median amplitude than late spring events (May–July). Only 10 sta-