Available online at www.sciencedirect.com
SCIENCE DIRECT®
Advances in Water Resources 27 (2004) 1061–1073
Advances in Water Resources
www.elsevier.com/locate/advwatres

Longitudinal interpolation of parameters characterizing channel geometry by piece-wise polynomial and universal kriging methods: effect on flow modeling

Brett F. Sanders *, Constantios V. Chrysikopoulos

Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697, USA

Received 23 February 2004; received in revised form 26 August 2004; accepted 26 August 2004

Abstract

Channel geometry often is described by a set of longitudinally varying parameters measured at a set of survey stations. To support flow modeling at arbitrary resolution, three methods of parameter interpolation are described including piece-wise linear interpolation, monotone piece-wise-cubic Hermitian interpolation, and universal kriging. The latter gives parameter estimates that minimize the mean square error of the interpolator, and therefore can be used as a standard against which the accuracy of polynomial methods can be assessed. Based on the application of these methods to a dataset describing cross-sectional properties at 283 stations, piece-wise linear interpolation gives parameter estimates that closely track universal kriging estimates and therefore this method is recommended for routine modeling purposes. Piece-wise-cubic interpolation gives parameter estimates that do not track as well. Differences between cubic and kriging estimates were found to be 2–10 times larger than differences between linear and kriging parameter estimates. In the context of one-dimensional flow modeling, the sensitivity of steady state water level predictions to the channel bed interpolator is comparable to a 5% change in the Manning coefficient.

1. Introduction

Topographic data describing the geometry of channels are required for all types of channel flow modeling including one-dimensional (cross-sectionally integrated), two-dimensional (vertically or laterally integrated), and three-dimensional approaches. These data are typically obtained by surveying transverse profiles, or cross-sections, at a series of stations distributed in the longitudinal direction [7], although relatively new vessel mounted, geo-referenced profiling technologies have allowed bed elevation to be recorded along a more general two-dimensional trajectory. Widely used river modeling software such as HEC-RAS [2] and UNET [18], developed by the US Army Corps of Engineers, directly use cross-sectional survey data to characterize geometric properties of the channel such as the top width, cross-sectional area, and conveyance. Data from digital elevation maps (DEMs) available from the Earth Resource Observation Systems (EROS) Data Center of the US Geologic Survey, which resolve topography on a cartesian grid at resolutions as fine as 30 m, are useful for mapping floodplain topography and making overland flow predictions, but not for channel topography because EROS DEMs do not resolve flooded topography and are too coarse to resolve channel topography at the resolution necessary for flow predictions.

When survey data consist of transverse profiles, measurements can be written as $z_o(x_i, y_j)$ where x_i represents the position of the ith survey station in the longitudinal direction (where $i = 1, \ldots, n$ and n is the

* Corresponding author. Tel.: +1 949 824 4327; fax: +1 949 824 3672.
E-mail addresses: bsanders@uci.edu (B.F. Sanders), costas@eng.uci.edu (C.V. Chrysikopoulos).

0309-1708/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2004.08.010

Nomenclature

A	wetted cross-sectional area, defined in (32), L^2
$\mathbf{A}$	matrix $(n+p \times n+p)$ of the kriging system, defined in (15)
$\mathbf{b}$	vector $(n+p)$ of the kriging system, defined in (17)
C_P	autocovariance function
$E[\,]$	expectation operator
f	known trial or base spatially dependent functions
Fr	Froude number
g	gravitational acceleration constant, L/t^2
L	Lagrangian, defined in (11)
MSE	mean square error of kriging estimator, defined in (18)
m	number of survey points in transects
n	number of available observations
n_m	Manning coefficient
n_p	number of parameters used to characterize cross-sectional geometry
p	number of drift coefficients
P	spatially dependent stochastic variable
P_w	wetted perimeter, L
Q	volumetric discharge, L^3/t
r	separation distance between measurements, L
S_f	friction slope, defined in (33)
T	bank to bank width of the channel at the free surface, L
V	cross-sectionally averaged fluid velocity, L/t
w_b	bottom width of the channel, L
w_t	bank width, L
x	spatial coordinate in the direction of flow, L
z_b	bottom elevation of the channel, L
z_t	bank elevation, L
$\mathbf{x}$	vector $(n+p)$ of unknowns of the kriging system, defined in (16)

Greek Letters

β	deterministic but unknown drift coefficients
γ	semi-variogram
ε	zero-mean stochastic process or residual
ζ	exponential semi-variogram model parameter, L
η	free surface elevation, L
$\lambda_1,\ldots,\lambda_n$	deterministic weight coefficients
μ_ε	mean or expected value of $\varepsilon(t)$
$\nu_1,\ldots,\nu_p$	1/2 Lagrange multipliers
ξ	Gaussian semi-variogram model parameter, L
σ^2	variance of the exponential semi-variogram model

number of survey stations), y_j represents position of the jth survey point in the transverse direction (where $j = 1,\ldots,m(i)$ and $m(i)$ is the number of survey points at each survey station). The spacing of survey stations is usually not consistent with the needs of computer models which may require many computational elements between neighboring survey stations [2]. Therefore, interpolation is needed to estimate geometric properties at intermediate stations. Fig. 1a provides a conceptual illustration of how channel geometry is described by a combination of surveyed and interpolated transects. To describe the longitudinal variability of cross-sectional geometry, it is assumed that section geometry can be characterized by a set of n_p longitudinally varying parameters, $P_k(x_i)$ where $k = 1,\ldots,n_p$ and $i = 1,\ldots,n$. These parameters could simply represent bed elevation at a regular set of points across the channel, although this alternative would be useful primarily for multi-dimensional river modeling. To support one-dimensional modeling, cross-sections can fitted by a particular shape such as a rectangle, trapezoid, or parabola, then coefficients that scale the particular geometry can be interpolated. For example, Chow [4] notes that many small and medium-sized natural channels can be approximated by a parabola, and more recent studies have adopted power-law expressions to characterize cross-sectional channel geometries [3,20]. For models that use "look-up tables" which characterize hydraulic properties of the channel, i.e., top width, cross-sectional area, and conveyance versus water level, another alternative is to fit polynomial, exponential, or power-law equations to these tables and interpolate the equation coefficients.

The present paper describes several longitudinal interpolators that can be adopted to estimate cross-sectional geometry, describes differences in cross-sectional properties estimated by different interpolators, and assesses the impact of these differences in the context of steady state flow predictions made by a deterministic flow model. Two types of interpolators are considered, namely piece-wise polynomial interpolation (linear and cubic) and universal kriging. Compared to universal kriging, piece-wise polynomial interpolation represents a fast, efficient, easily implemented and widely used approach of parameter estimation. However, universal kriging gives parameter estimates that minimize the mean square error of the interpolator, so in theory, it gives the most accurate estimate possible of bed elevation

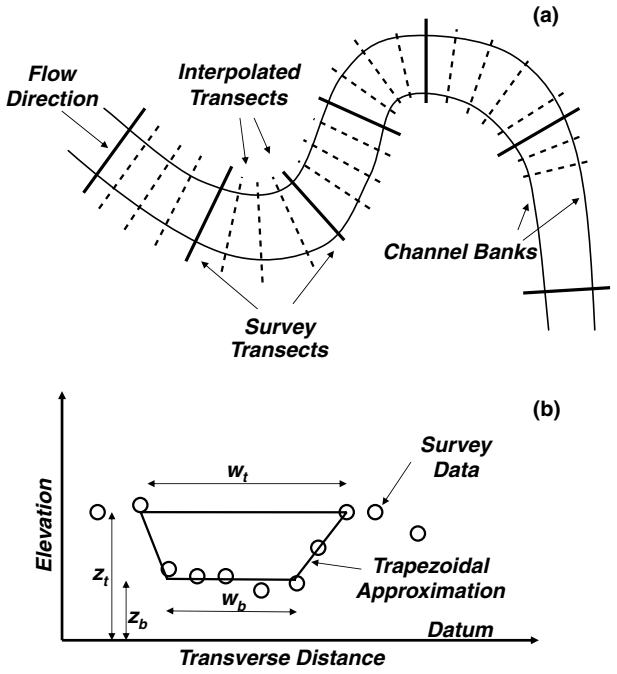

Fig. 1. Conceptual illustration of channel geometry. Plan view is presented in (a) including survey transects, where channel geometry is measured, and interpolated transects. Cross-sectional view is presented in (b) including hypothetical survey measurements and a trapezoidal approximation of this geometry. Cross-section view illustrates the four geometrical parameters estimated in this study: z_t, z_b, w_t, and w_b.

data at points where it is not measured based on data at points where it is measured. By comparing linear and cubic polynomial interpolation estimates of bed elevation data against universal kriging estimates, insight into the more accurate polynomial interpolator is obtained.

2. Field data

The data used in this study were obtained from the Santa Clara Valley Water District and consist of transects at 283 stations ($n = 283$) along San Francisquito Creek, which drains from the Santa Cruz Mountains east to San Francisco Bay in northern California as shown in Fig. 2. These transects span a 12.2 km portion of San Francisquito Creek below USGS gaging station 11164500, where the channel bed is earthen and characterized by a cross-sectional shape that closely resembles a trapezoid. Due to this shape and for consistency with the hydraulic model used in this study, a four parameter description ($n_p = 4$) was adopted involving the bottom width w_b, bottom elevation z_b, bank width w_t, and bank elevation z_t, as shown in Fig. 1b. These parameters were estimated by a relatively subjective graphical fitting procedure and are presented in Fig. 3. Bank elevation and bottom elevation are presented in Fig. 3a, while bank width and bottom width are presented in Fig. 3b. The elevation data illustrate that the channel is characterized

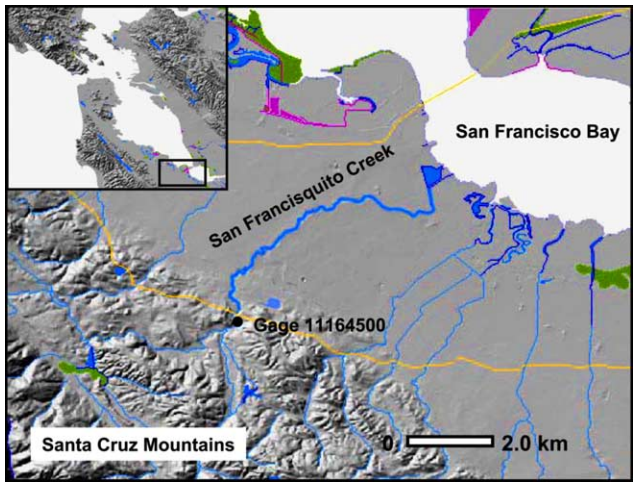

Fig. 2. San Francisquito Creek and surrounding watershed showing the study reach between the stream gage (x) and San Francisco Bay.

by a concave-up longitudinal profile [13], with a bed slope of roughly 0.5% over the first 2 km, a slope of roughly 0.25% between kilometer 2 and 10, and a nearly horizontal bed slope between kilometer 10 and 12. The bottom width and bank width average roughly 10 and 30 m, respectively, over the first 10 km though these parameters vary locally by as much as 10 m. Bridges cross the creek at several points denoted in Fig. 3 by solid vertical lines. Bridge abutments, which restrict the flow, are modeled as a rectangular cross-section so the bottom and bank widths are identical. Between kilometer 10 and 12, the channel is considerably wider.

3. Methods

Piece-wise polynomial and universal kriging interpolation methods applied to this dataset are described in this section, along with a cross-sectionally integrated flow model used to assess the sensitivity of flow predictions to the channel bed description. To simplify the presentation of interpolation methods, the parameter notation $P_k(x_i)$ is written as $P(x_i)$.

3.1. Piece-wise polynomial interpolation

For a set of n available observations of the spatially variable parameter to be interpolated, $P(x_1), \ldots, P(x_n)$, where $x_1 \ldots x_n$ are monotonically increasing, piece-wise linear interpolation provides an estimate of P at x_0 based only on neighboring data. That is, for $x_i \leqslant x_0 \leqslant x_{i+1}$, $P(x_0)$ is estimated in terms of $P(x_i)$ and $P(x_{i+1})$. For linear polynomials, the estimate is given as

$$\widehat{P}(x_0) = \frac{x_0 - x_{i+1}}{x_{i+1} - x_i} P(x_i) + \frac{x_0 - x_i}{x_{i+1} - x_i} P(x_{i+1}) \quad x_i \leqslant x \leqslant x_{i+1}, \tag{1}$$

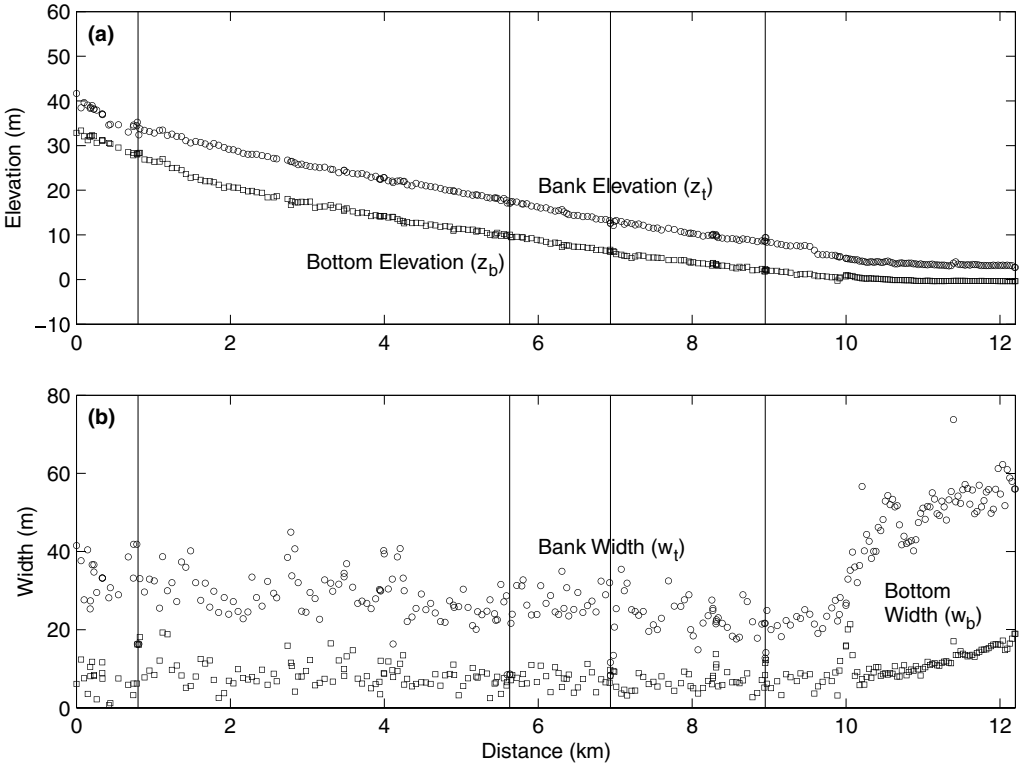

Fig. 3. Parameter values at survey stations. Bank elevation z_t (circles) and bottom elevation z_b (squares) are presented in (a), and bank width w_t (circles) and bottom width w_b (squares) are presented in (b). Solid vertical lines denote the position of bridge abutments modeled as rectangular cross-sections.

where the hat signifies an estimate. For cubic polynomials, the estimates is given as,

$$\widehat{P}(x_0) = a_i + b_i(x_0 - x_i) + c_i(x_0 - x_i)^2 + d_i(x_0 - x_i)^3 \quad x_i \leqslant x \leqslant x_{i+1}, \tag{2}$$

where a_i, b_i, c_i, and d_i are weighting coefficients of the ith segment. There are several different ways to determine these coefficients. Each enforces the property that $\widehat{P}(x_i) = P(x_i)$ and $\widehat{P}(x_{i+1}) = P(x_{i+1})$, but there are differences in how the remaining two constraints are obtained. For example, use of $\widehat{P}'(x_i) = P'(x_i)$ and $\widehat{P}'(x_{i+1}) = P'(x_{i+1})$ corresponds to piece-wise-cubic Hermite interpolation, where the prime notation signifies the derivative, and therefore this method requires that both the unknown parameter and its derivative be given as input data. The interpolator is a cubic spline when the cubic polynomial is constrained by continuity of the second derivative, $\widehat{P}''(x)$, at both endpoints, x_i and x_{i+1}. In this study, the PCHIP routine (piece-wise-cubic Hermite interpolating polynomial) implemented in the commercial software MATLAB (The MathWorks, Natick, Mass.) is adopted [8]. Using this approach, the slopes $P'(x_i)$ are computed from parameter measurements $P(x_i)$ in such a way that $P(x)$ is monotonic. Cubic splines were initially tested but not used because the resulting parameter estimates were not at all realistic, i.e., the interpolator introduced new maxima and minima that greatly exceeded measurements.

3.2. Interpolation using universal kriging

Universal kriging is a relatively flexible geostatistical method of parameter identification. The kriging estimator is a weighted linear combination of all available experimental data, in contrast to piece-wise polynomial interpolation which only combines data from neighboring points

$$\widehat{P}(x_0) = \sum_{i=1}^{n} \lambda_i P(x_i), \tag{3}$$

where the hat signifies an estimate; and $\lambda_1, \ldots, \lambda_n$ are deterministic but unknown weight coefficients. These are determined so that on the average the estimator error is zero (unbiasedness property) and the square estimation error is as small as possible (minimum variance property). The weights are obtained from the solution of a system of linear equations (kriging system).

The parameter P is assumed to be described by the following model

$$P(x) = \sum_{k=1}^{p} \beta_k f_k(x) + \varepsilon(x), \tag{4}$$

where $\beta_1,\ldots,\beta_p$ are deterministic but unknown coefficients often referred to as drift coefficients; $f_1(x),\ldots,f_p(x)$ are known space-dependent functions which are called trial or base functions; and $\varepsilon(x)$ is a zero-mean stochastic process (residual). Consequently, the mean of the parameter P is given by

$$\overline{P}(x) = \sum_{k=1}^{p} \beta_k f_k(x), \tag{5}$$

where $\overline{P}(x) = E[P(x)]$ is the mean of P (where $E[\,]$ is the expectation operator). Note that the functions for the space-dependent parameter P and its mean defined in (4) and (5), respectively, are both linear in the drift coefficients.

In order to satisfy the unbiasedness requirement the weight coefficients $\lambda_1,\ldots,\lambda_n$ should be selected so that the average estimation error is zero

$$E\left[\widehat{P}(x_0) - P(x_0)\right] = 0 \tag{6}$$

for any of the unknown drift coefficients $\beta_1,\ldots,\beta_p$. In view of (3)–(5) the preceding expression can be written as

$$\sum_{k=1}^{p}\left\{\sum_{i=1}^{n}\lambda_i f_k(x_i) - f_k(x_0)\right\}\beta_k = 0. \tag{7}$$

For this condition to be valid for any drift coefficient $\beta_1,\ldots,\beta_p$ it is evident that

$$\sum_{i=1}^{n}\lambda_i f_k(x_i) = f_k(x_0). \tag{8}$$

In order to satisfy the minimum variance requirement, the variance or mean square error $\mathrm{MSE} = E\left[(\widehat{P}(x_0) - P(x_0))^2\right]$ should be as small as possible. In view of (3)–(5) the variance of the linear estimator $\widehat{P}(x_0)$ can be expressed as

$$\mathrm{MSE} = E\left[\left(\widehat{P}(x_0) - P(x_0)\right)^2\right] = \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j \mathrm{C_P}(x_i,x_j) - 2\sum_{i=1}^{n}\lambda_i \mathrm{C_P}(x_i,x_0) + \mathrm{C_P}(x_0,x_0), \tag{9}$$

where

$$\mathrm{C_P}(x_i,x_j) = E\left[\left(P(x_i) - \overline{P}(x_i)\right)\left(P(x_j) - \overline{P}(x_j)\right)\right] = E\left[\varepsilon(x_i)\varepsilon(x_j)\right] \tag{10}$$

is the autocovariance function representing the mutual variability between $P(x_i)$ and $P(x_j)$, or equivalently the mutual variability between $\varepsilon(x_i)$ and $\varepsilon(x_j)$. Minimization of the objective function (9) subject to the constraint (8) can be obtained by the Lagrange multipliers method [12]. This method requires formation of the Lagrangian

$$L(\lambda_1,\ldots,\lambda_n,\nu_1,\ldots,\nu_p) = \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j \mathrm{C_P}(x_i,x_j) - 2\sum_{i=1}^{n}\lambda_i \mathrm{C_P}(x_i,x_0) + \mathrm{C_P}(x_0,x_0) + 2\sum_{k=1}^{p}\nu_k\left\{\sum_{i=1}^{n}\lambda_i f_k(t_i) - f_k(x_0)\right\}, \tag{11}$$

where $2\nu_1,\ldots,2\nu_p$ are the Lagrange multipliers (the 2 is used only for mathematical convenience). A system of $n+p$ linear equations is formed by taking the derivatives of $L(\lambda_1,\ldots,\lambda_n,\nu_1,\ldots,\nu_p)$ with respect to $\lambda_1,\ldots,\lambda_n,\ \nu_1,\ldots,\nu_p$ and setting them equal to zero, as follows:

$$\sum_{j=1}^{n}\lambda_j \mathrm{C_P}(x_i,x_j) + \sum_{k=1}^{p}\nu_k f_k(x_i) - \mathrm{C_P}(x_i,x_0) = 0, \quad i = 1,2,\ldots,n, \tag{12}$$

$$\sum_{i=1}^{n}\lambda_i f_k(x_i) - f_k(x_0) = 0, \quad k = 1,2,\ldots,p. \tag{13}$$

Eqs. (12) and (13) define a system of $n+p$ linear equations with $n+p$ unknowns that can be solved for the unknown coefficients $\lambda_1,\ldots,\lambda_n,\ \nu_1,\ldots,\nu_p$. This system of equations is the kriging system that can be expressed in matrix notation as

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{14}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathrm{C_P}(x_1,x_1) & \cdots & \mathrm{C_P}(x_1,x_n) & f_1(x_1) & \cdots & f_p(x_1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{C_P}(x_n,x_1) & \cdots & \mathrm{C_P}(x_n,x_n) & f_1(x_n) & \cdots & f_p(x_n) \\ f_1(x_1) & \cdots & f_1(x_n) & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_p(x_1) & \cdots & f_p(x_n) & 0 & \cdots & 0 \end{bmatrix}, \tag{15}$$

$$\mathbf{x} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \nu_1 \\ \vdots \\ \nu_p \end{bmatrix}, \tag{16}$$

$$\mathbf{b} = \begin{bmatrix} C_P(x_1, x_0) \\ \vdots \\ C_P(x_n, x_0) \\ f_1(x_0) \\ \vdots \\ f_p(x_0) \end{bmatrix}. \tag{17}$$

Assuming that $C_P(x_i, x_j)$ is known or it can be determined from available data, the coefficients $\lambda_1, \ldots, \lambda_n$, $v_1, \ldots, v_p$ are easily determined by solving the kriging system (14). Subsequently, the estimate $\widehat{P}(x_0)$ is evaluated from (3). Note that neither $P(x_0)$ nor $P(x_0 + r)$ need to be measured for the estimation of $\widehat{P}(x_0)$. Furthermore, in view of (12) and (13) the MSE expression (9) can be simplified as

$$\mathrm{MSE} = -\sum_{i=1}^{n} \lambda_i C_P(x_i, x_0) - \sum_{k=1}^{p} v_k f_k(x_0) + C_P(x_0, x_0). \tag{18}$$

3.2.1. Autocovariance function determination

A random field is partially described by second-order characteristics, e.g. its mean function or expected value, and autocovariance function [5]. Consequently, the stationary stochastic random field for the parameter $\varepsilon(t)$ is characterized by the mean function:

$$\mu_\varepsilon(x) = E[\varepsilon(x)] = 0, \tag{19}$$

indicating a zero-mean stochastic process, and by the autocovariance function:

$$C_P(x_i, x_j) = C_P(r) = E[\varepsilon(x)\varepsilon(x + r)], \tag{20}$$

representing the mutual variability between $\varepsilon(x_i)$ and $\varepsilon(x_j)$ or equivalently the mutual variability between $\varepsilon(x)$ and $\varepsilon(x + r)$, where $r = |x_i - x_j|$. The autocovariance function of a stationary random field is related to the semi-variogram by [9,11]

$$C_P(r) = C_P(0) - \gamma(r), \tag{21}$$

where $C_P(0) = E[\varepsilon^2(x)]$ is the variance of $\varepsilon(x)$, representing the mean square deviation of $\varepsilon(t)$ from its mean value, $\mu_\varepsilon(x) = 0$; and $\gamma(r)$ is the semi-variogram defined as

$$\gamma(r) = \frac{1}{2} E[(\varepsilon(x) - \varepsilon(x + r))^2]. \tag{22}$$

To determine the autocovariance function $C_P(x_i, x_j)$ the detrended data sets, shown in Fig. 4e–h , were used for the construction of the appropriate raw semi-variograms or scatter plots. The raw semi-variogram is essentially a plot of the square difference $1/2[\varepsilon(x_i) - \varepsilon(x_i + r)]^2$ as a function of separation distance between measurements, r.

For n experimental data points, there are $n(n-1)/2$ such pairs that comprise the raw variogram. For the $z_{t(u)}$ detrended data set (see Fig. 4e$_u$), there are 202 measurements that yield a cloud of 20,301 pairs, indicated by the solid circles in Fig. 5a$_u$. Dividing the axis of r into nine consecutive intervals and by averaging the pairs of measurement in each interval, the experimental semi-variogram of the detrended data is constructed and is illustrated by the solid squares in Fig. 5a$_u$. The experimental semi-variogram is fitted with the following exponential, theoretical semi-variogram model:

$$\gamma(r) = \sigma^2 \left[1 - \exp\left(-\frac{r}{\zeta}\right)\right], \tag{23}$$

$$C_P(r) = \sigma^2 \exp\left(-\frac{r}{\zeta}\right), \tag{24}$$

where $\sigma^2 = C_P(0) > 0$ is the variance or sill of the exponential semi-variogram, and $\zeta > 0$ is a model parameter that determines how fast the semi-variogram increases to its sill value. For a value of $r = 3\zeta$ the exponential semi-variogram is approximately equal to 95% of σ^2; this distance is known as the range of the exponential semi-variogram model. The exponential model is frequently used in various hydrologic applications [6,14,17,19]. The unknown model parameters of the semi-variogram are determined by a relatively subjective graphical fitting procedure that leads to estimates $\sigma^2 = 31.98$ and $\zeta = 193.80$ meters. In Fig. 5a$_u$, the fitted model is indicated by the solid curve and is in agreement with the experimental semi-variogram. In a similar fashion we constructed the raw variograms (scatter plots) and obtained the appropriate theoretical semi-variograms for the $z_{t(l)}$, $z_{b(u)}$, $z_{b(l)}$, $w_{t(u)}$, $w_{t(l)}$, $w_{m(u)}$, and $w_{t(l)}$ data sets shown in Fig. 4.

3.2.2. Interpolation at desired grid locations

In view of equation (5) and the estimated drift coefficients for $z_{t(u)}$ presented in Table 1, it is evident that

$$\begin{aligned} \bar{z}_{t(u)}(x) &= E[z_{t(u)}(x)] \\ &= 33.33681 - 1.28497 \times 10^{-3} x. \end{aligned} \tag{25}$$

Consequently, the appropriate trial or base functions for $z_{t(u)}$ are:

$$f_1(x) = 1, \tag{26}$$

$$f_2(x) = x. \tag{27}$$

Furthermore, in view of (21), (23) and the fitted theoretical semi-variogram for $z_{t(u)}$ listed in Table 2, the corresponding autocovariance function is:

$$C_{x_{t(u)}}(r) = 31.98 \exp\left[-\frac{r}{193.8}\right]. \tag{28}$$

For the $z_{t(u)}$ data set, $n = 202$, and $p = 2$. Consequently, in view of (26)–(28) and (14) the corresponding kriging system can be written as

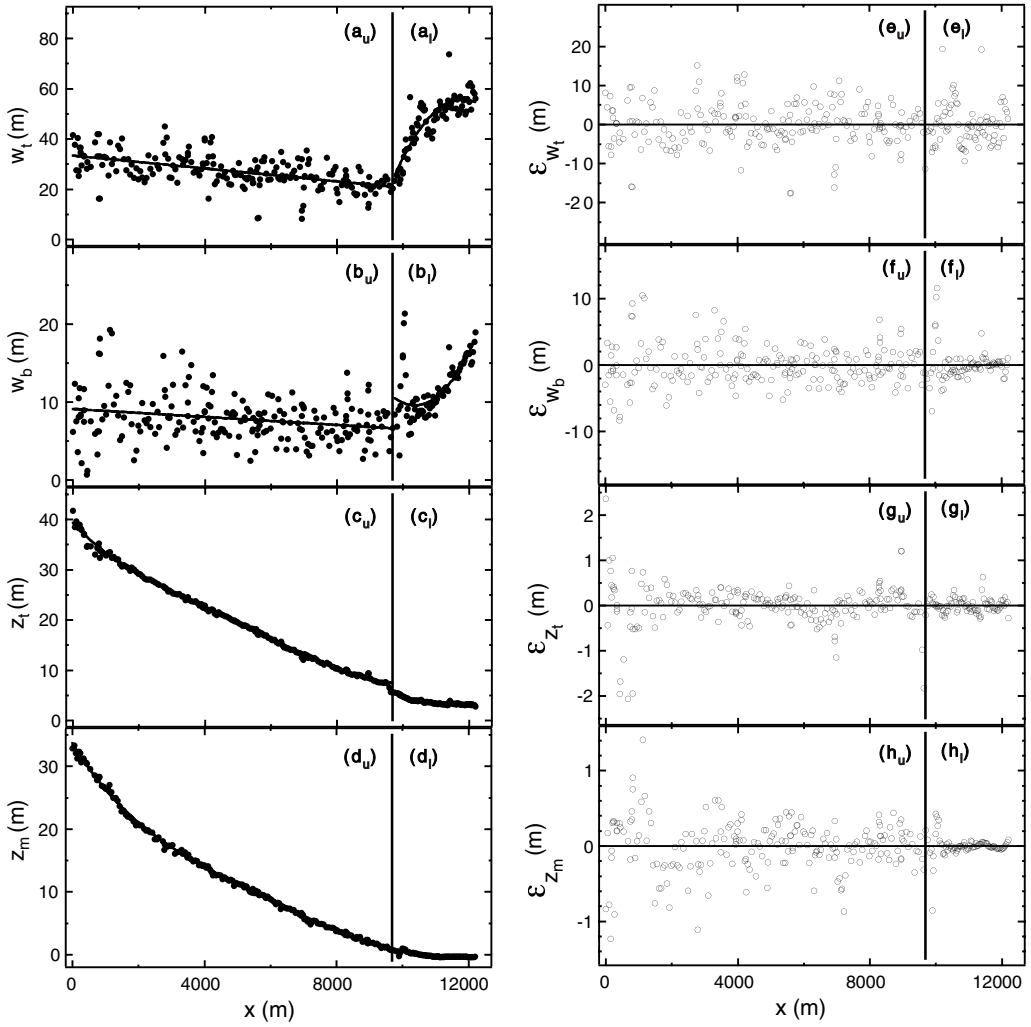

Fig. 4. Field measurements at various downstream locations (solid circles) and fitted polynomials (solid curves) for: (a_u) $w_{t(u)}$, (a_l) $w_{t(l)}$, (b_u) $w_{b(u)}$, (b_l) $w_{b(l)}$, (c_u) $z_{t(u)}$, (c_l) $z_{t(l)}$, (d_u) $z_{m(u)}$, and (d_l) $z_{t(l)}$. The residuals or detrended data (open circles) are presented in (e) through (h). The subscript 'u' refers to the upper portion of the reach (0 to 9700 m) and 'l' refers to the lower portion of the reach (9700 to 12,200 m).

$$\begin{bmatrix} 31.98 & \cdots & 31.98\exp\left[-\frac{|x_1-x_n|}{193.8}\right] & 1 & x_1 \\ 31.98\exp\left[-\frac{|x_2-x_1|}{193.8}\right] & \cdots & 31.98\exp\left[-\frac{|x_2-x_n|}{193.8}\right] & 1 & x_2 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 31.98\exp\left[-\frac{|x_{202}-x_1|}{193.8}\right] & \cdots & 31.98 & 1 & x_{202} \\ 1 & \cdots & 1 & 0 & 0 \\ x_1 & \cdots & x_{202} & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{202} \\ v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 31.98\exp\left[-\frac{|x_1-x_0|}{193.8}\right] \\ 31.98\exp\left[-\frac{|x_2-x_0|}{193.8}\right] \\ \vdots \\ 31.98\exp\left[-\frac{|x_{202}-x_0|}{193.8}\right] \\ 1 \\ x_0 \end{bmatrix}. \tag{29}$$

The preceding set of linear equations is solved for the unknown coefficients $\lambda_1,\ldots,\lambda_{202}$, v_1, and v_2, for 9700 different x_0 values (grid points) ranging from 0 to 9700 m. For each x_0 considered the corresponding $\widehat{z}_{t(u)}(x_0)$ is determined by (3) using the determined coefficients $\lambda_1,\ldots,\lambda_{202}$ and the data presented in Fig. 4a_u. All other desired parameter values are interpolated in a similar fashion.

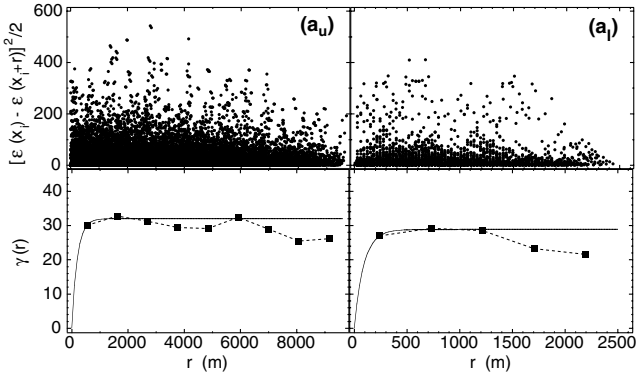

Fig. 5. Scatter plots of the detrended data sets (dots), experimental semi-variograms of detrended data sets (solid squares) and fitted theoretical semi-variogram models (solid curve) for: (a_u) $z_{t(l)}$ and (a_l) $z_{t(l)}$.

3.3. Cross-sectionally integrated flow model

Cross-sectionally integrated flow models (i.e., one-dimensional) are widely used for watershed scale modeling involving floods, sediment transport, and water quality predictions and are the basis of government sponsored modeling software such as the river analysis system HEC-RAS [2], the channel network model UNET [18], and the hydrologic and water quality modeling system HSPF [1]. Cross-sectionally integrated flow models are derived from macroscopic mass and momentum balance equations where the control volume, shown in Fig. 6, is a tube of water bounded laterally by the channel bed, banks, and free-surface and longitudinally by control sections separated by a finite length, Δx [7]. In the limit that $\Delta x \to 0$, these equations are converted to a differential form known as the St. Venant equations,

$$\frac{\partial \eta(x,t)}{\partial t} + \frac{1}{T(x,t)} \frac{\partial Q(x,t)}{\partial x} = 0, \tag{30}$$

$$\frac{\partial Q(x,t)}{\partial t} + \frac{\partial}{\partial x}[V(x,t)Q(x,t)] + gA(x,t)\frac{\partial \eta(x,t)}{\partial x} = -gA(x,t)S_f(x,t), \tag{31}$$

where η is the free surface elevation, Q is the volumetric flow rate, V is the cross-sectionally averaged fluid velocity, T is the bank to bank width of the channel at the free surface (top width), A is the wetted cross-sectional area, g is the gravitational acceleration constant, S_f is the friction slope, x measures distance along the channel centerline (longitudinal distance), and t is time. The free-surface elevation is related to the cross-sectional area by the geometry of the channel as follows:

$$A(x,t) = \int_{z_b(x)}^{\eta(x,t)} w(x,z)\mathrm{d}z, \tag{32}$$

where $z_b(x)$ is the bottom elevation (lowest point) of the channel cross-section, $w(x,z)$ represents the bank-to-bank width, and z measures vertical position as shown in Fig. 6. Note that $w(x,\eta(x,t)) = T(x,t)$. The volumetric discharge, cross-sectional area, and cross-sectionally averaged velocity are related by $Q(x,t) = V(x,t)A(x,t)$, and the friction slope is determined by an empirical

Table 2
Fitted theoretical semi-variograms and variances

Parameter	$\gamma(r)$	$C_p(0) = \gamma(\infty)$
$w_{t(u)}$	$31.98\left[1-\exp(-\frac{r}{193.80})\right]$	31.98
$w_{t(l)}$	$28.82\left[1-\exp(-\frac{r}{88.91})\right]$	28.82
$w_{b(u)}$	$9.25\left[1-\exp(-\frac{r}{161.21})\right]$	9.25
$w_{b(l)}$	$6.11\left[1-\exp(-\frac{r}{32.96})\right]$	6.11
$z_{t(u)}$	$0.34\left[1-\exp(-\frac{r}{2988})\right]$	0.34
$z_{t(l)}$	$0.024\left[1-\exp(-\frac{r}{25.20})\right]$	0.024
$z_{b(u)}$	$0.11\left[1-\exp(-\frac{r}{74.04})\right]$	0.11
$z_{b(l)}$	$0.018\left[1-\exp(-\frac{r}{79.65})\right]$	0.018

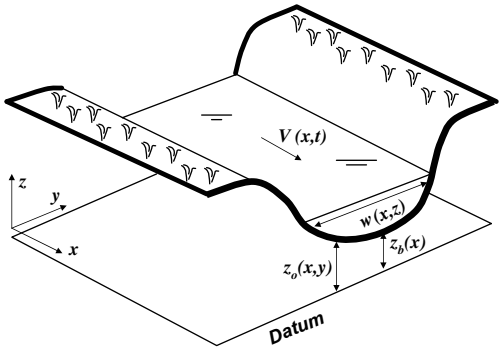

Fig. 6. Typical channel geometry illustrating geometrical shape parameters.

Table 1
Drift coefficients

Parameter	β_1	β_2	β_3	β_4	β_5
$w_{t(u)}$	33.33681	-1.28497×10^{-3}			
$w_{t(l)}$	−952.7980	1.709383×10^{-1}	-7.243566×10^{-6}		
$w_{b(u)}$	9.08768	-2.567335×10^{-4}			
$w_{b(l)}$	281.4035	-5.245153×10^{-2}	2.530053×10^{-6}		
$z_{t(u)}$	39.33364	-7.084518×10^{-3}	1.275390×10^{-6}	-1.730340×10^{-10}	8.446015×10^{-15}
$z_{t(l)}$	927.0335	-2.745761×10^{-1}	2.967751×10^{-5}	-1.360626×10^{-9}	2.163467×10^{-14}
$z_{b(u)}$	33.69423	-8.671870×10^{-3}	1.456767×10^{-6}	-1.566120×10^{-10}	6.470396×10^{-15}
$z_{b(l)}$	−4827.1370	1.7740	-2.436105×10^{-4}	1.481862×10^{-8}	-3.369931×10^{-13}

resistance law such as the Manning equation, which gives the friction slope as

$$S_f(x,t) = \frac{n_m(x)V(x,t)^2 P_w(x,t)^{4/3}}{A(x,t)^{4/3}}, \quad (33)$$

where n_m is the Manning coefficient and $P_w(x,t)$ is the wetted perimeter.

The St. Venant equations describe unsteady, gradually varied channel flow under the following set of conditions [7]: (i) the velocity and water level are laterally uniform, (ii) streamline curvature is small and vertical accelerations are negligible, (iii) the effects of boundary friction and turbulence can be accounted for by resistance laws analogous to those used for steady flow, and (iv) the average channel bed slope is small, so that the cosine of the angle it makes with the horizontal may be replaced by unity. Because of these restrictions, the St. Venant equations given by Eqs. (30) and (31) are applicable only to channels with relatively simple geometry such as that shown in Fig. 6, or channels that are rectangular, trapezoidal, or triangular in cross-section. Modification to these equations are needed to model flow in channels with irregular cross-sections, when overbank flow occurs, or when any of the above conditions are not applicable. For water quality modeling, flow models which solve Eq. (30) and a simplified form of Eq. (31) are often used (e.g., [1]).

In the context of modeling flow in San Francisquito creek (where the cross-section is roughly trapezoidal), the St. Venant equations are applicable to flood flows that do not crest above the banks of the channel. It has been estimated this occurs when the discharge exceeds 170 m^3/s, while the 100 year return period flood has been estimated to be 250 m^3/s [10]. To characterize the sensitivity of flood stage predictions to the topographic interpolation, but not exceed the limitations of the flow model, the St. Venant equations were solved to predict water levels in San Francisquito Creek based on a volumetric flow rate of 100 m^3/s, which corresponds roughly to a 5-year flood event [10]. A total variation diminishing (TVD) finite volume numerical method was used to numerically integrate the St. Venant equations [15], or more specifically macroscopic mass and momentum balance equations applied to the control volume shown in Fig. 1. That is, this numerical method does not discretize the differential form given by Eqs. (30) and (31) but rather integral equations. A more detailed description of these differences is presented in [7]. This numerical method is ideal for resolving channel flow over a wide range of Froude numbers because it can accommodate discontinuities arising from hydraulic jumps. Source terms related to spatial variations in the channel shape are modeled in a robust way [16]. Finally, energy losses are accurately predicted even though the

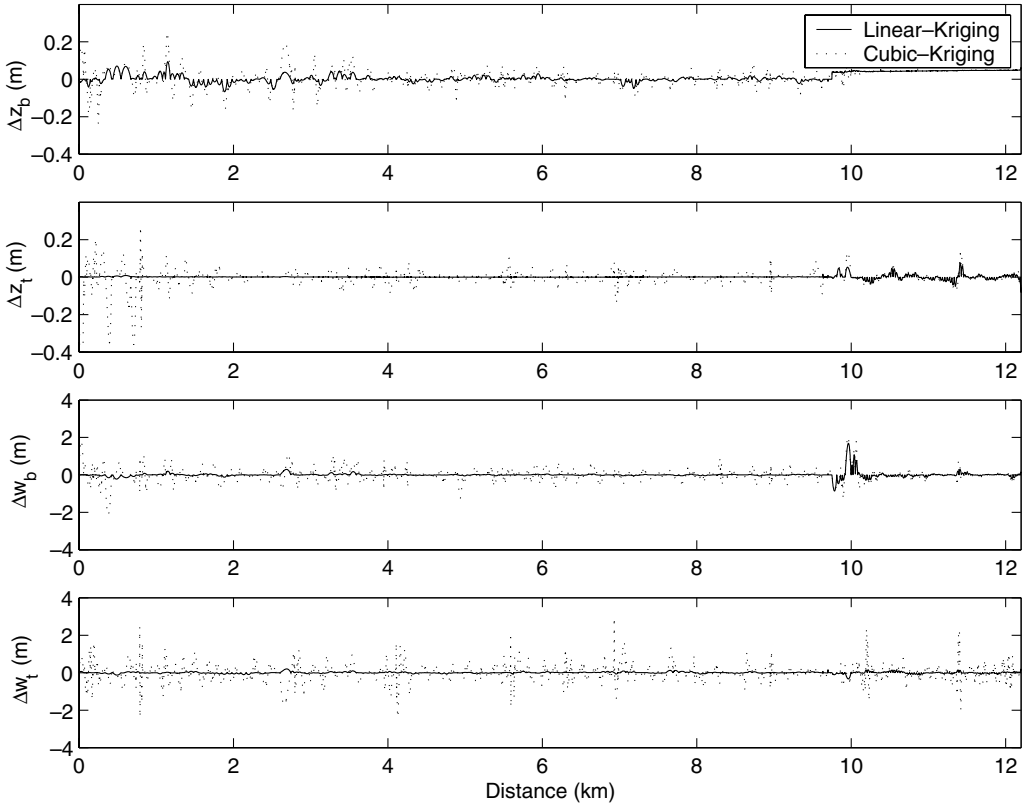

Fig. 7. Differences between piece-wise linear and kriging based estimates of channel geometry parameters (solid line) and between piece-wise-cubic and kriging based estimates (dots).

model is a momentum-only formulation, i.e., the model does not include an energy balance equation [16].

4. Results and discussion

Based on the data from 283 stations and the methods described above, interpolation was performed by piece-wise linear polynomials (linear), monotone piece-wise-cubic Hermitian interpolating polynomials (cubic), and universal kriging (kriging) on a grid of 12,201 points which discretize the 12.2 km reach of San Francisquito Creek at 1 m resolution. For all of the parameters, kriging estimates were tracked considerably better by linear estimates than cubic estimates. Fig. 7 presents differences between polynomial and universal kriging estimates for z_b, z_t, w_b, and w_t. The delta notation is used to denote a difference, either between the linear and kriging estimates or the cubic and kriging estimates. The root-mean-square difference between linear and kriging estimates of z_b, z_t, w_b, and w_t was computed to be 2.5, 0.8, 13 and 4.3 cm, respectively; while the root-mean-square difference between cubic and kriging estimates for z_b, z_t, w_b, and w_t was computed to be 4.6, 5.8, 35 and 44 cm, respectively. Hence, the linear estimates are in all cases better predictors than cubic estimates, given that kriging estimates are most accurate. The implication of this comparison is that, when interpolating channel topography between survey stations, piece-wise linear interpolation is preferable to high order methods such as monotone piece-wise-cubic Hermitian interpolating polynomials or cubic splines.

Deterministic flow predictions were made on three separate 1 km reaches of San Francisquito Creek to assess sensitivities to channel geometry. These reaches correspond to km 0–1, 6–7, and 10–11 and were chosen to sample a steep, intermediate and nearly horizontal portion of the channel, respectively. For analysis purposes, a uniform value of the Manning coefficient, $n_m = 0.025$ was used and a supercritical upstream boundary condition based upon a depth of 1 m was enforced. These conditions give rise to regions of both subcritical and supercritical flow allowing sensitivities to be assessed over a range of flow conditions. In reality, the Manning coefficient for San Francisquito Creek is likely to be higher since it is vegetated, and vary longitudinally. Results of model predictions are shown in Figs. 8–10 in terms of the free surface elevation η, the difference

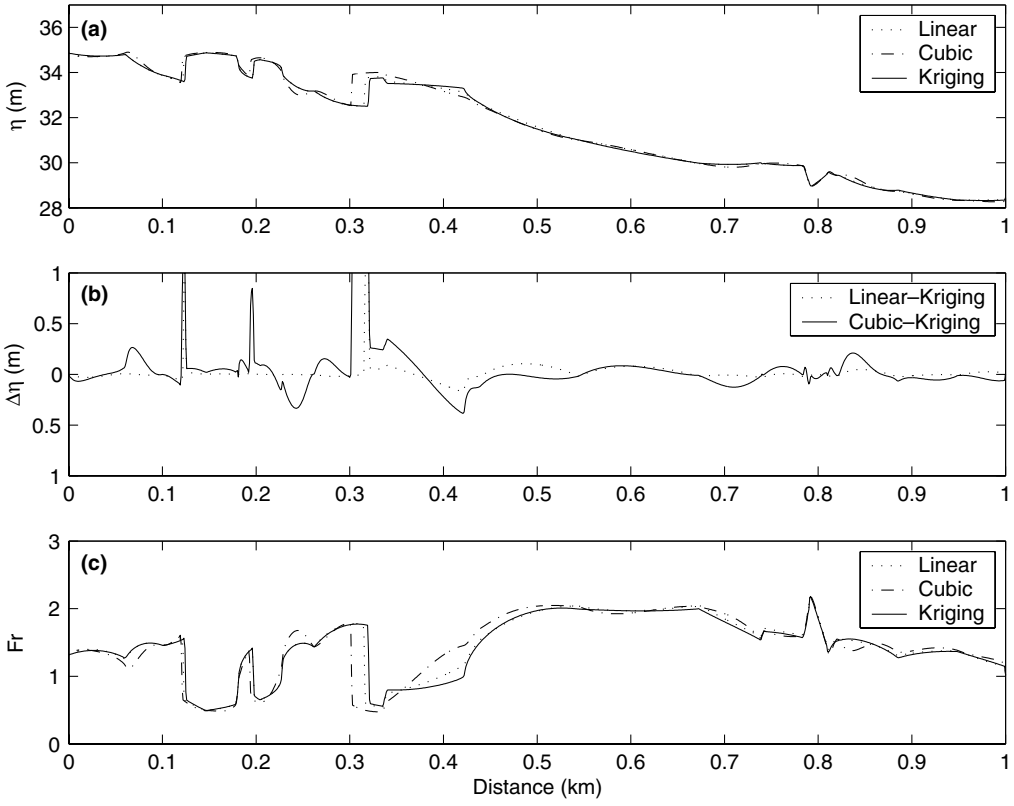

Fig. 8. Flow model predictions for km 0–1. Water elevation η is presented in (a) based on linear, cubic, and kriging based estimates of channel geometry. Differences between linear and kriging based estimates of water elevation, and between cubic and kriging based estimates, are presented in (b). The Froude number, Fr based on each prediction is presented in (c) to indicate regions of supercritical ($Fr > 1$) and subcritical ($Fr < 1$) flow, as well as hydraulic jumps.

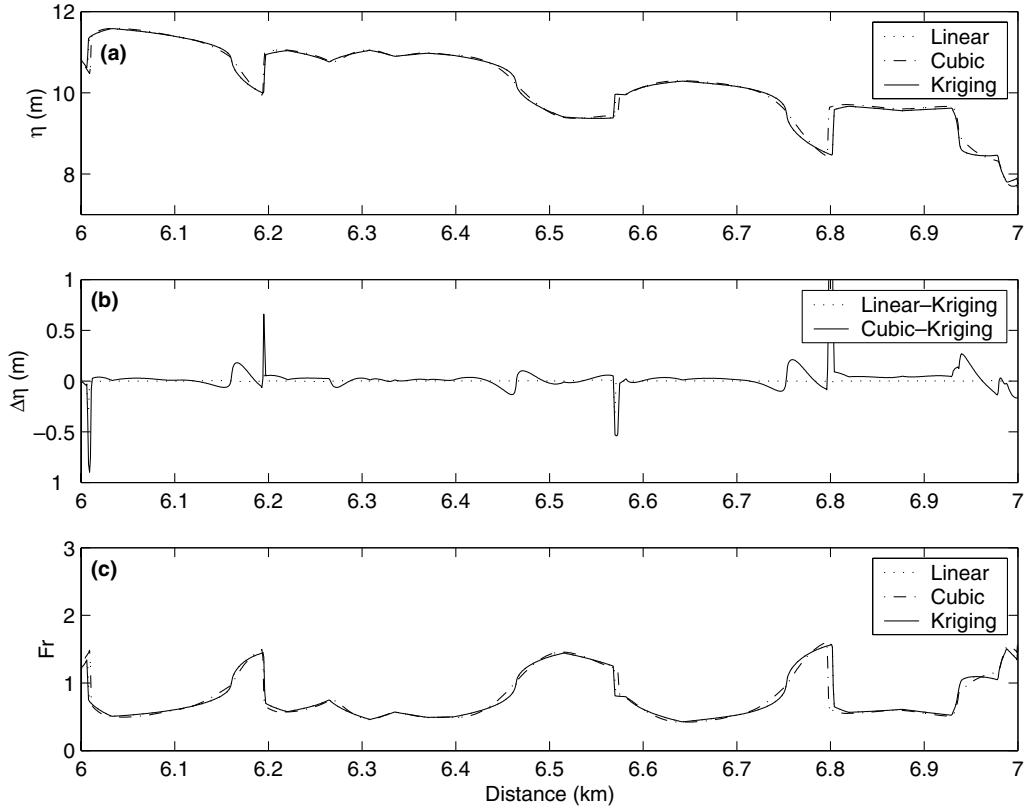

Fig. 9. Flow model predictions for km 6–7. Water elevation η is presented in (a) based on linear, cubic, and kriging based estimates of channel geometry. Differences between linear and kriging based estimates of water elevation, and between cubic and kriging based estimates, are presented in (b). The Froude number, Fr based on each prediction is presented in (c) to indicate regions of supercritical ($Fr > 1$) and subcritical ($Fr < 1$) flow, as well as hydraulic jumps.

between model predictions of η using polynomial and kriging based topographic parameters, and the Froude number Fr which corresponds to the ratio of the cross-sectionally averaged velocity to the wave speed and defines regions of subcritical ($Fr < 1$) and supercritical ($Fr > 1$) flow. Note that over the first reach flow is largely supercritical but includes several hydraulic jumps, over the second reach flow is supercritical and subcritical in equal proportion, and over the third reach flow is largely subcritical.

Differences between linear and kriging based predictions of η are smaller than differences between cubic and kriging based predictions, which is not surprising considering that linear based topographic parameters are a better predictor of the kriging based topographic parameters than are the cubic based parameters. The root-mean-square difference between estimates of η, based on linear and kriging based parameter estimates, was computed to be 2.9, 0.5 and 1.3 cm for km 0–1, 6–7, and 10–11, respectively; while based on cubic and kriging parameters the root-mean-difference was computed to be 6.8, 3.4 and 2.1 cm, respectively. On the whole, these differences are small compared to the depth in each of these reaches, the average of which was computed to be 1.9, 2.8, and 2.0 m for km 0–1, 6–7, and 10–11, respectively. At most, this represents less than a 5% error in the depth prediction. Locally, however, these differences can be significant. For example, roughly at $x = 0.3$ km a hydraulic jump is predicted to occur but its position, based on cubic prediction of channel geometry, is roughly 20 m upstream of where it is predicted to be based on linear and kriging based channel geometry.

Another source of uncertainty in channel flow predictions arises from the selection of resistance parameters, in this case the Manning coefficient, n_m. To quantify the effect of channel interpolation methods on water level predictions, in comparison to the effect of uncertainty in resistance parameters, steady state flow predictions were repeated for each reach (km 0–1, 6–7, and 10–11) using a 5% larger Manning coefficient. Kriging based estimates of the channel geometry were used for these predictions. The root-mean-square difference between estimates of η, based on a 5% increase in n_m, was computed to be 2.7, 3.3 and 3.0 cm for km 0–1, 6–7, and 10–11, respectively. These differences in η are slightly larger than differences in predictions based on linear and kriging based channel geometry

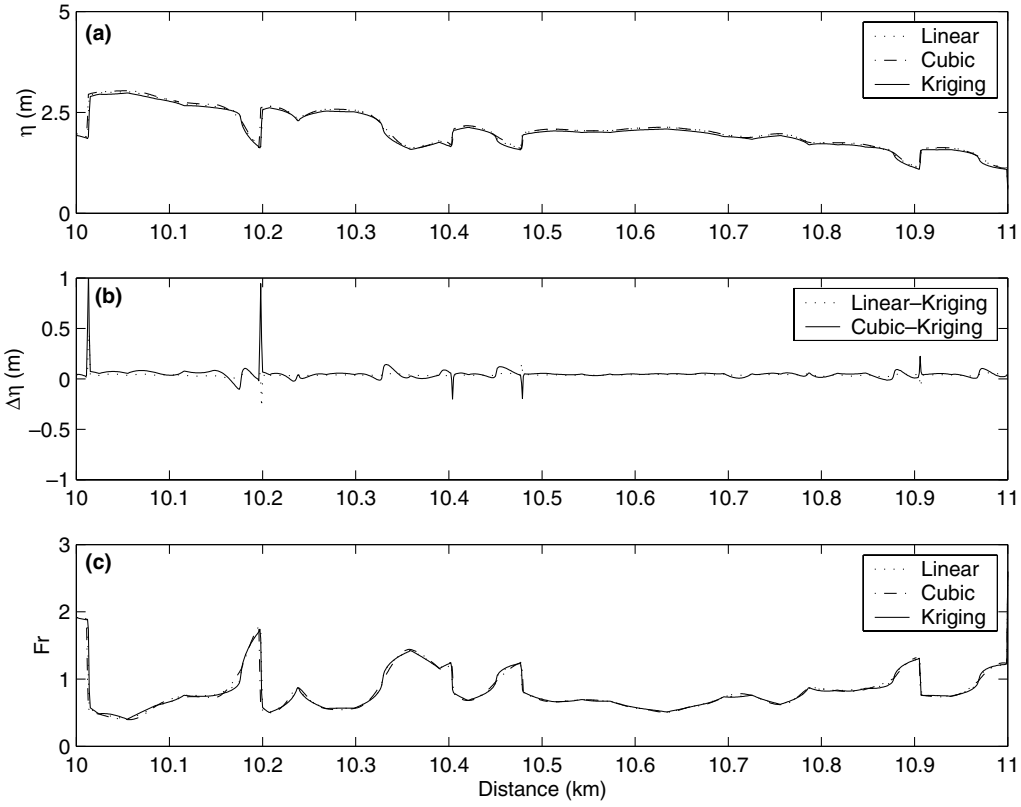

Fig. 10. Flow model predictions for km 10–11. Water elevation η is presented in (a) based on linear, cubic, and kriging based estimates of channel geometry. Differences between linear and kriging based estimates of water elevation, and between cubic and kriging based estimates, are presented in (b). The Froude number, Fr based on each prediction is presented in (c) to indicate regions of supercritical ($Fr > 1$) and subcritical ($Fr < 1$) flow, as well as hydraulic jumps.

estimates, but similar or slightly smaller than differences based on cubic and kriging based channel geometry estimates.

Many streams feature geometrical properties considerably more complex than those of San Francisquito Creek involving features such as multiple thalwegs, pronounced flood plains, and terraces. Methods of longitudinal interpolation presented in this study can be applied without modification to such systems by using larger numbers of parameters. For example, a channel characterized by a thalweg bordered by active floodplain would require at least an eight parameter description (elevations and widths for the thalweg, left flood plain, right flood plain, and bank elevations). For more complex channels, it may become advantageous to target hydraulic properties for interpolation instead of geometrical properties, assuming the number of parameters needed to fit each hydraulic property remains constant while the number of parameters needed to parameterize channel geometry increases. Additional study is warranted to carefully examine these tradeoffs, to characterize the performance of interpolators in more complex channels, and to characterize the performance of interpolators when data are sparse and/or data gaps exist.

5. Summary

Interpolation of parameters characterizing the geometry of channel beds is needed to support flow modeling at varying grid resolutions. Three methods of interpolating geometric parameters between survey stations were described including piece-wise linear interpolation, monotone piece-wise-cubic Hermitian interpolation, and universal kriging. The latter gives parameter estimates that minimize the mean square error of the interpolator and therefore is considered the most accurate method. Based on the application of these methods to a dataset describing cross-sectional properties at 283 stations, piece-wise linear interpolation gave parameter estimates that very closely track universal kriging estimates. Piece-wise-cubic interpolation, including monotone piece-wise-cubic Hermitian interpolation and cubic spline interpolation, gave parameter estimates that did not track as well. In fact, parameter estimation based on cubic splines is not a viable method because it predicts unphysical, oscillatory parameter values. Given that the implementation of universal kriging for interpolation purposes is considerably more involved than piece-wise polynomial interpolation, the results of this study support use of piece-wise linear interpolation

over higher order polynomial interpolants for deterministic flow modeling.

In the context of one-dimensional flow predictions, root-mean-square differences in water level predictions arising from different interpolators were at most 5% of the depth. The largest differences were observed in the vicinity of hydraulic jumps, because the interpolation method was observe to alter the predicted longitudinal position of the jump by as much as 20 m. Differences in water level predictions based on linear and kriging based channel geometry estimates were slightly smaller than differences based on cubic and kriging based channel geometry estimates. The sensitivity of flow predictions to the channel bed interpolation method was found to be similar to 5% uncertainty in the Manning coefficient.

Acknowledgments

The authors wish to thank D. Jaffe for organizing the dataset used in this study, and R. Argall for assisting with illustrations. This work was supported in part by grants from the University of California Water Resources Center (#W 942) and the National Science Foundation (# CMS-9984579).

References

[1] Bicknell BR, Imhoff JC, Kittle JL, Jr, Donigian AS, Jr, Johanson RC. Hydrological Simulation Program–Fortran: User's manual for version 11: US Environmental Protection Agency, National Exposure Research Laboratory, Athens, GA., EPA/600/R-97/080, 1997. 755p.

[2] Brunner GW. HEC-RAS, River Analysis System User's Manual, US Army Corps of Engineers, Hydrologic Engineering Center, Davis, CA, 2002.

[3] Buhman DL, Gates TK, Watson CC. Stochastic variability of fluvial hydraulic geometry: Mississippi and Red rivers. J Hydraul Eng 2002;128(1):426–37.

[4] Chow VT. Open-channel hydraulics. McGraw Hill; 1959. 680p.

[5] Christakos G. Random field models in earth sciences. San Diego, CA: Academic Press; 1992.

[6] Chrysikopoulos CV, Kitanidis PK, Roberts PV. Analysis of one-dimensional solute transport through porous media with spatially variable retardation factor. Water Resour Res 1990;26(3):437–46.

[7] Cunge JA, Holly Jr FM, Verwey A. Practical aspects of computational river hydraulics. Pitman; 1980.

[8] Fritsch FN, Carlson RE. Monotone piecewise cubic interpolation. SIAM J Numer Anal 1980;17:238–46.

[9] Isaaks EH, Strivastava RM. An introduction to applied geostatistics. New York, NY: Oxford University Press; 1989.

[10] Jaffe DA. Levee breaches for flood reduction. PhD Dissertation. Environmental Engineering Graduate Program, University of California, Irvine, 2002.

[11] Journel AG, Huijbregts CJ. Mining Geostatistics. London, UK: Academic Press; 1978.

[12] Kitanidis PK. Introduction to geostatistics: applications in hydrogeology. Cambridge, UK: Cambridge University Press; 1997.

[13] Leopold LB, Wolman MG, Miller JP. Fluvial processes in geomorphology. San Francisco, Calif: W.H. Freeman; 1964.

[14] Russo D, Zaidel J, Laufer A. Stochatic analysis of solute transport in partially saturated heterogeneous soils, 1. Numerical experiments. Water Resour Res 1994;30(3):769–79.

[15] Sanders BF. High-resolution and non-oscillatory solution of the St. Venant equations in non-rectangular and non-prismatic channels. J Hydraul Res 2001;39(3):321–30. [see also Discussion and Reply by Authors, J Hydraul Res 41(6):668–72].

[16] Sanders BF, Jaffe DA, Chu AK. Discretization of integral equations describing flow in nonprismatic channels with uneven beds. J Hydraul Eng 2003;129(3):235–44.

[17] Sudicky EA. A natural gradient experiment on solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role in the disperion process. Water Resour Res 1986;22(13): 2069–82.

[18] US Army Corps of Engineers. UNET, One-dimensional unsteady flow through a full network of open channels. User's manual, Hydrologic Engineering Center, Institute for Water Resources, Davis, CA, 2001.

[19] Vogler ET, Chrysikopoulos CV. Dissolution of nonaqueous phase liquid pools in anisotropic aquifers. Stoch Environ Res Risk Assess 2001;15:33–46.

[20] Western AW, Finlayson BL, McMahon TA, O'Neill IC. A method for characterising longitudinal irregularity in river channels. Geomorphology 1997;21:39–51.