On polynomial predictions for river flows*

E. G. Birgin[†] J. M. Martínez[‡]

December 18, 2023

Abstract

This paper deals with the prediction of river levels by means of polynomial regression models using only elevation data and inflow forecasts. Different models for this purpose are examined and a new approach based on the concept of virtual stations is presented. Detailed numerical experiments show that this proposal may be useful as a tool for making predictions when the physical characteristics of the river are uncertain.

Key words: Flow predictions, natural rivers, Saint-Venant equations, parameter estimation.

1 Introduction

River flow modelling is an important tool for analysing and predicting dam failures and their consequences. The main mathematical procedure for this task is based on the solution of partial differential equations (PDE). The equations of Saint Venant [20] are the best known equations for this purpose. Their numerical solution requires initial and boundary conditions in terms of river wetted cross-sections and flow-rates. In addition, geometric descriptions of the cross sections and bed elevations are required. Finally, Manning roughness coefficients, which may be spatially and temporally dependent, must be determined. See [1, 2, 3, 5, 4, 6, 7, 8, 11, 12, 13, 14, 17, 18, 19, 20, 21].

Typically, partial observations of river surface elevations at different spatial and temporal coordinates are available. These observations make it possible the estimation of the unknown characteristics of the river, which are necessary for the numerical integration of the partial differential equations. The resulting PDE-constrained parameter estimation problem can be difficult to solve, requires integration of the PDE's for different instances, and is subject to instability and lack of reliability of results. However, this problem has been the subject of valuable research over many years. See [1, 2, 3, 5, 6, 7, 8, 11, 14, 15, 17, 18].

The PDE approach obtains predictions by means of the estimation of unknown physical characteristics and associated PDE integration. Moreover, the estimation of unknown physical characteristics is based on fitting the direct solution of the PDE's to available observations. This suggests the possibility of obtaining river predictions directly from available data without the need to estimate the physical characteristics of the river. The obvious drawback of this approach relies on the fact that

*This work was supported by FAPESP (grants 2013/07375-0, 2018/24293-0, and 2022/05803-3) and CNPq (grants 302538/2019-4 and 302682/2019-8).

[†]Department of Computer Science, Institute of Mathematics and Statistics, University of São Paulo, Rua do Matão, 1010, Cidade Universitária, 05508-090, São Paulo, SP, Brazil. e-mail: egbirgin@ime.usp.br

[‡]Department of Applied Mathematics, Institute of Mathematics, Statistics, and Scientific Computing (IMECC), State University of Campinas, 13083-859 Campinas SP, Brazil. e-mail: martinez@ime.unicamp.br

we do not have reliable physical models that directly link observations to predictions. For this reason we believe that data-based predictions should generally be considered in conjunction with PDE predictions, although the specific form of this relation is highly problem-dependent [6].

Reliable data-based approaches should start with a reliable identification of cause-effect relationships. For example, in the case of river flow phenomena, a high correlation may be found between upstream discharge and downstream elevations. Obviously, upstream discharges are the cause of downstream elevations and not the other way round. If a cause-effect relationship is established, the next step could be to propose an appropriate form of dependence relationship, the specific form of which should be based on previous data analysis.

Let us consider an example that is well suited to introduce and motivate the rest of this paper. It has been widely observed that water elevation at an arbitrary fixed station of a natural river is a smooth function of the upstream (inlet) flow-rate. See [12] and [2, Fig.12b]. In Figure 1, we consider data for the Fork River published in [9]. Figure 1a shows observations of the elevation z corresponding to the section $x = 751$ m, together with linear, quadratic and cubic polynomials representing elevation as a function of the inflow rate $Q_{\min}$ (in m^3/s). The polynomials were fitted using simple least squares. Figure 1b shows the same information but related to the section $x = 3256$ m. The observations are taken every 12 hours starting at zero hours on day 3. The polynomial coefficients and the corresponding root mean square deviation (RMSD) are given in Table 1.

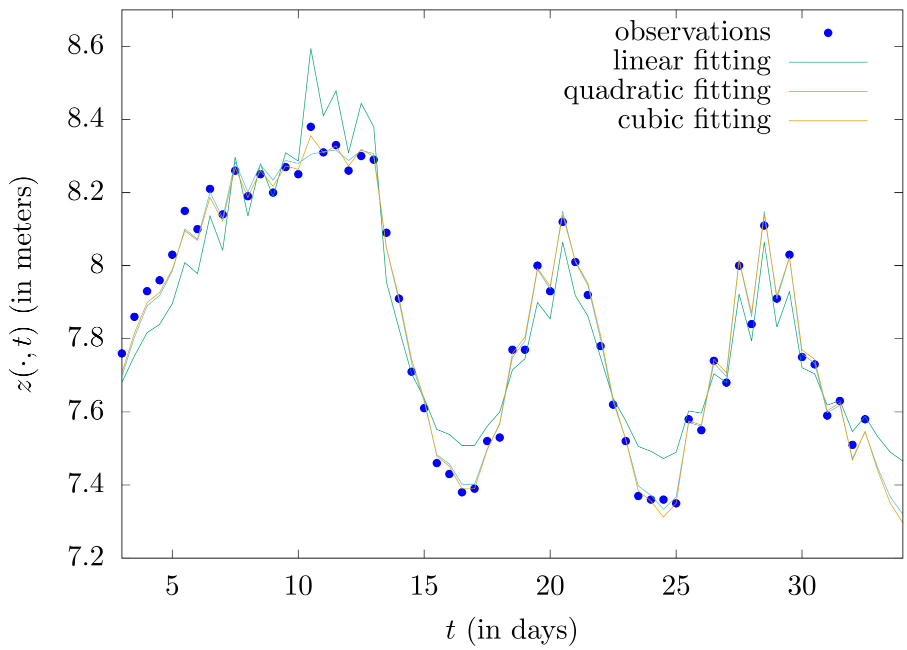
Station	Polynomial	RMSD	c_0	c_1	c_2	c_3
751 m	linear	8.69579603E-02	7.35113673	3.75568519E-02	−−	−−
	quadratic	2.69668513E-02	7.08338033	8.19336547E-02	-1.36086954E-03	−−
	cubic	2.42162234E-02	7.01642805	9.97412870E-02	-2.60953038E-03	2.47226020E-05
3256 m	linear	6.02123240E-02	5.44084782	3.91356263E-02	−−	−−
	quadratic	3.13462816E-02	5.28175904	6.62052107E-02	-8.39381211E-04	−−
	cubic	3.07813747E-02	5.24970397	7.49278445E-02	-1.45802975E-03	1.23271897E-05

Table 1: Fork river: fitting polynomials, their coefficients, and the corresponding RMSD (in meters).

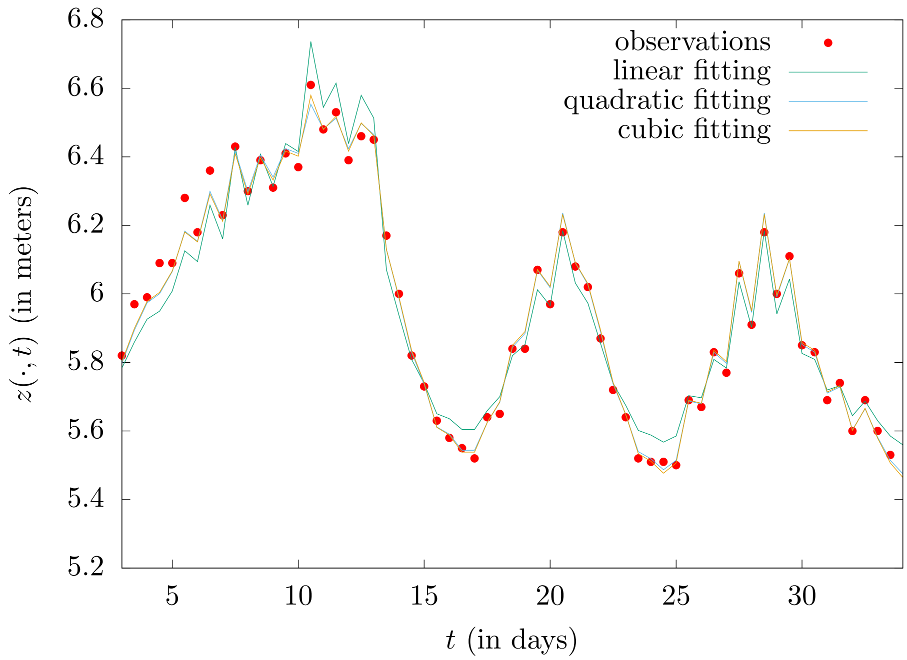
It is interesting to fit the data of, say, the first 10 days and observe if the approximating curves fit well the data for the remaining days. Figure 2 and Table 2 show the results. Throughout this paper surface elevations and the corresponding RMSD errors are expressed in meters. So, for example, the testing error of the cubic polynomial for $x = 751$ m meters is 4.80 cm according to Table 2. This error is quite small for practical prediction purposes regarding a real river.

Station	Polynomial	RMSD		c_0	c_1	c_2	c_3
		training	testing				
751 m	linear	4.36362260E-02	2.24297179E-01	7.66210301	2.34403945E-02	−−	−−
	quadratic	1.97586581E-02	1.11828660E-01	7.33545084	5.90878962E-02	-8.67373515E-04	−−
	cubic	1.06148710E-02	4.79787043E-02	6.94435497	1.24958948E-01	-4.23241309E-03	5.33788086E-05
3256 m	linear	4.41380828E-02	1.67298488E-01	5.67319985	2.89537322E-02	−−	−−
	quadratic	3.43615061E-02	8.69890046E-02	5.44061014	5.43362118E-02	-6.17605427E-04	−−
	cubic	2.73302025E-02	7.05794611E-02	4.95184090	1.36658084E-01	-4.82303941E-03	6.67097817E-05

Table 2: Fork river: Fitting polynomials, their coefficients, and the corresponding RMSD. In this case, observations of the first 10 days were used as training data to fit the polynomials. The remaining observations (20 or 21 days for Sections $x = 751$ m and Section $x = 3256$ m, respectively) were not used in the fitting (training) phase and, then, were used to test the predictions provided by the fitted polynomials.



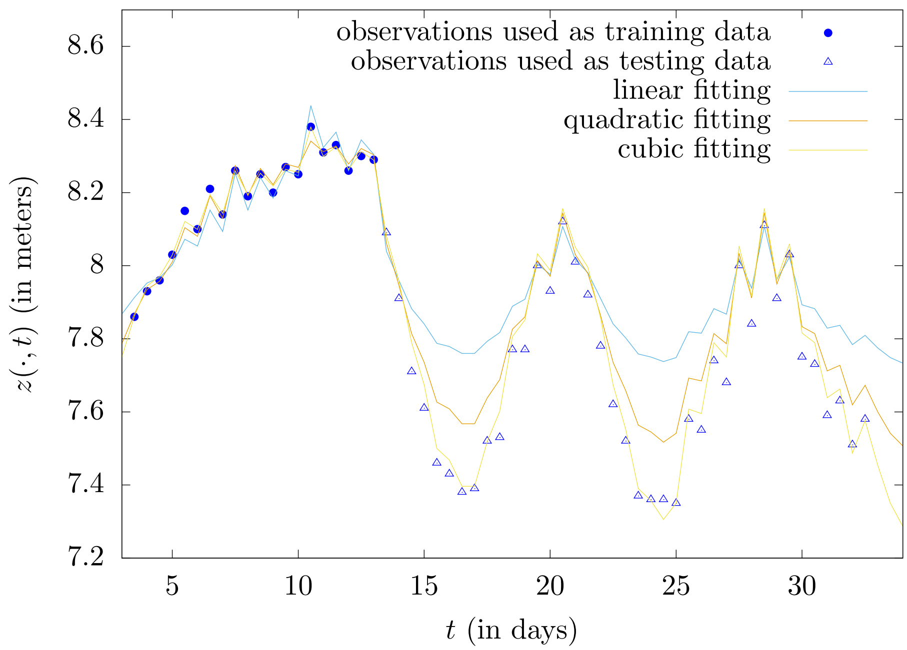
(a) Section $x = 751\,\mathrm{m}$.



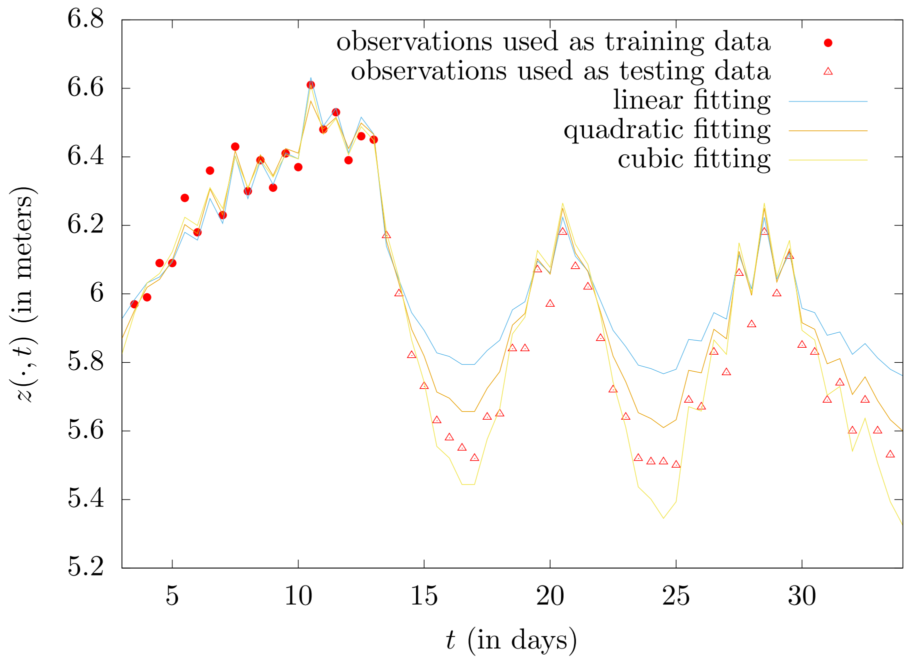
(b) Section $x = 3256\,\mathrm{m}$.

Figure 1: Fork river: Observed elevations at a given station and their approximation as a (linear, quadratic and cubic) fitting polynomial of the inlet discharge.

These results suggest that, for predicting elevations at a fixed station x in "future days" under suitable forecast on the inlet discharge, it is enough to fit the curve of the surface elevation $z(x, t)$ versus $Q_{\min}(t)$ using available data at station x, with the reasonable belief that, in the next days, this



(a) Section $x = 751\,\mathrm{m}$.



(b) Section $x = 3256\,\mathrm{m}$.

Figure 2: Fork river: Observed elevations at a given station and their approximation as a (linear, quadratic and cubic) fitting polynomial of the inlet discharge. In this case, observations of the first 10 days were used as training data to fit the polynomials. The remaining observations (20 or 21 days for Sections $x = 751\,\mathrm{m}$ and Section $x = 3256\,\mathrm{m}$, respectively) were not used in the fitting (training) phase and, then, were used to test the predictions provided by the fitted polynomials.

curve will provide reasonable elevation estimates, provided that inlet discharge forecasts are reliable. In fact, this should be the case if one has data for a suitable number of days before "today" and for all the relevant stations along the river. Unfortunately both situations are unlike to occur. Usually, one needs previsions for the future employing a possibly moderate number of past data at a possibly moderate number of stations x.

For example, according to Table 1, for $x = 751\,\text{m}$, the best third-order polynomial that represents $z(x,t)$ as a function of $Q_{\min}(t)$ is given by

$$z(751,t) \approx 7.02 + 9.97 \times 10^{-2} Q_{min}(t) - 2.61 \times 10^{-3} Q_{min}(t)^2 + 2.47 \times 10^{-5} Q_{min}(t)^3, \qquad (1)$$

while, for $x = 3256\,\text{m}$, the best third-order polynomial that represents $z(x,t)$ as a function of $Q_{\min}(t)$ is given by

$$z(3256,t) \approx 5.25 + 7.49 \times 10^{-2} Q_{min}(t) - 1.46 \times 10^{-3} Q_{min}(t)^2 + 1.23 \times 10^{-5} Q_{min}(t)^3. \qquad (2)$$

However, if $x \notin \{751, 3256\}$, we do not know, for example, which is the best third-order polynomial that fits the elevations $z(x,t)$ at Section $x = 555\,\text{m}$ as a function of $Q_{\min}(t)$. This question is addressed in the present paper.

We will start from the empirical observation that, in real rivers, inlet discharge is the dominant cause of river elevations at different stations. This fact supports the idea that, given a spatial position x, the elevation $z(x,t)$ can be well approximated by a low-order polynomial $P(Q_{\min}(t))$. We will see that third-order polynomials are the more appropriate for this purpose. In order to recover elevations at stations x that are not represented in the data we analyse the employment of two-dimensional polynomials in the variables x and $Q_{\min}(t)$. However, the need to preserve the accuracy of the one-dimensional fits leads us to propose a different strategy based on the concept of "virtual stations". This paper proposes an algorithm for selecting suitable virtual stations and demonstrates its reliability through detailed numerical experiments.

This research is conducted within CRIAB, a Latin-American academic group that involves collaborators of several countries. The group is dedicated to analyzing, comprehending and mitigating dam-breaking and related accidents. River modelling is one of the techniques that must be mastered in the broader landscape of modelling embankments and basins. Optimization regression techniques are among the tools used for this purpose.

This paper is organized as follows. Section 2 analyses the compatibility of one-dimensional regression with two-variable polynomial fitting. Section 3 introduces the method of virtual stations and describes the algorithm that will be used in the experiments. Section 4 describe the generation of synthetic data. Numerical experiments are reported in Section 5, while conclusions and future research directions are presented in Section 6.

Notation. $\#A$ will denote the number of elements of the set A. If A and B are sets, $A \setminus B$ denotes the set of elements of A that do not belong to B.

2 Two-variable polynomial fitting

Consider an arbitrary one-dimensional flow where the spatial (length) coordinate x goes from $x_{\min}$ to $x_{\max}$. The surface elevation for space coordinate x and time coordinate t will be denoted $z(x,t)$. Assume that at p different stations $x_1, \ldots, x_p \in [x_{\min}, x_{\max}]$ we have observations of surface elevations at different times. The inlet discharge (flow-rate at $x = x_{\min}$) at time $t \in [t_{\min}, t_{\max}]$ is denoted $Q_{\min}(t)$. For simplicity, if confusion is not possible, we omit the dependence of t in this notation (denoting $Q_{\min} = Q_{\min}(t)$). Assume that, at each station x_j, we fit a polynomial $P_j(Q_{\min})$ with degree q, in the least-squares sense, in order to minimize the deviations with respect to measured elevations.

We may consider the model

$$z(x,t) \approx W_1(x) P_1(Q_{\min}(t)) + \cdots + W_p(x) P_p(Q_{\min}(t)), \tag{3}$$

where, for all $j = 1, \ldots, p$, $W_j(x)$ is a polynomial with degree $p-1$ such that $W_j(x_j) = 1$ and $W_j(x_\ell) = 0$ if $\ell \neq j$. Namely,

$$W_j(x) = \frac{\prod_{i \neq j}(x - x_j)}{\prod_{i \neq j}(x_i - x_j)}. \tag{4}$$

The right-hand side of (3) is a sum of $p(q+1)$ monomials of the form $\gamma_{i,j} x^i Q_{\min}^j$ for $i = 0, 1, \ldots, p-1$ and $j = 0, 1, \ldots, q$.

This suggests the model

$$z(x,t) \approx \sum_{i=0}^{s} \sum_{j=0}^{q} \gamma_{i,j} x^i Q_{\min}(t)^j. \tag{5}$$

In (5), we postulate that the elevation at each point (x,t) is a two-variable polynomial with variables x and $Q_{\min}(t)$, with degree s in the variable x and degree q in the variable $Q_{\min}$. Note that in (3) we have that $s = p - 1$.

The model (5) induces a linear least-squares problem, in which the coefficients $\gamma_{i,j}$ are the unknowns and observations are available at different stations and times. We wonder whether, if observations are given at a finite number of stations $x_1, \ldots x_p$, the solution of the least-squares problem comes from addressing p separate least squares problems, one corresponding to each station. In this case, we could compute the best polynomial of degree q with respect to measurements at the considered station and the predicted values at arbitrary points (x,t) would come from interpolation according to (3) and (4).

The following theorem gives an answer to this question.

Theorem 2.1 *Assume that elevations $z_{k,\ell}$ are given at p stations x_k, $k = 1, \ldots, p$, and time instants t_ℓ, $\ell = 1, \ldots, r_k$. Assume, moreover, that for each observed $z_{k,\ell}$ the inlet flow $Q_{\min}(t_\ell)$ (in short Q_ℓ) is known. Consider the linear least-squares problems*

$$\text{Minimize} \sum_{k=1}^{p} \sum_{\ell=1}^{r_k} \left[\sum_{j=0}^{q} \sum_{i=0}^{s} \gamma_{i,j} x_k^i Q_\ell^j - z_{k,\ell} \right]^2 \tag{6}$$

and

$$\text{Minimize} \sum_{k=1}^{p} \sum_{\ell=1}^{r_k} \left[\sum_{j=0}^{q} \beta_{k,j} Q_\ell^j - z_{k,\ell} \right]^2. \tag{7}$$

Then, the objective function value at the solution of (7) is less than or equal to the objective function value at the solution of (6). Moreover, if $s \geq p-1$ both objective functions are identical at respective solutions.

Proof: Problem (6) is equivalent to

$$\text{Minimize} \sum_{k=1}^{p} \sum_{\ell=1}^{r_k} \left[\sum_{j=0}^{q} \beta_{k,j} Q_\ell^j - z_{k,\ell} \right]^2 \tag{8}$$

subject to

$$\beta_{k,j} = \sum_{i=0}^{s} \gamma_{i,j} x_k^i \text{ for all } k = 1, \ldots, p, j = 0, 1, \ldots, q. \tag{9}$$

Therefore, problem (6) is equivalent to problem (7) with the additional constraints (9). So, the feasible region of (7) contains the feasible region of (8,9). This implies that the objective function of (7) at its solution is smaller than or equal to the objective function of (8,9) at its solution. Both objective function values are identical if the feasible region of (7) is the same as the feasible region of (8,9), that is, if for all $\beta_{k,j} \in \mathbb{R}$ there exist $\gamma_{i,j}$ such that the identity (9) holds. This would mean that the linear system (9) (with unknowns $\gamma_{i,j}$) and independent term given by $\beta_{k,j}$) is compatible.

By (9), for $j = 0, 1, \ldots, q$, we have

$$\beta_{1,j} = \gamma_{0,j} x_1^0 + \gamma_{1,j} x_1^1 + \cdots + \gamma_{s,j} x_1^s, \tag{10}$$

$$\beta_{2,j} = \gamma_{0,j} x_2^0 + \gamma_{1,j} x_2^1 + \cdots + \gamma_{s,j} x_2^s, \tag{11}$$

$$\cdots$$

$$\beta_{p,j} = \gamma_{0,j} x_p^0 + \gamma_{1,j} x_2^1 + \cdots + \gamma_{s,j} x_p^s. \tag{12}$$

If $s < p-1$ the systems (10)–(12) are overdetermined and the solution set may be empty. In that case, the objective function value at the solution of (6) could be bigger than the objective function value at the solution of (7). If $s = p-1$, for each $j = 0, 1, \ldots, q$, the equations (10)–(12) define a $p \times p$ Vandermonde system. See [10, pp.203-207]. So, the $q+1$ systems (10–(12) are compatible and the unknowns $\gamma_{0,j}, \ldots, \gamma_{p-1,j}$ are (uniquely) determined by the constraints (9). If $s > p-1$ the systems (10)–(12 are underdetermined and particular solutions come from completing the solutions of the case $s = p-1$ with $\gamma_{p,j} = \ldots, \gamma_s = 0$. Therefore, when $s \geq p-1$, the constraints (9) do not impose any constraint at all to the solution of (8). Thus, the problems (6) and (7) are equivalent when $s \geq p-1$. This completes the proof. □

However, if observations $z_{\text{obs}}(x_k, t_k)$ are available at different times and stations (x_k, t_k), $k \in K_{\text{obs}}$, we must rely directly on the least squares problems induced by (5). Namely,

$$\text{Minimize} \sum_{k \in K_{\text{obs}}} \left[\sum_{i=0}^{s} \sum_{j=0}^{q} \gamma_{i,j} x_k^i Q_{\min}(t_k)^j - z_{\text{obs}}(x_k, t_k) \right]^2. \tag{13}$$

Note that problems of the form (6) are of the form (13) but the reciprocal is not true. Observe, moreover, that the number of parameters γ_{ij} that are estimated when we use (13) is $(s+1)(q+1)$, where s is the degree of the polynomial with respect to the variable x and q is the degree of the polynomial with respect to the variable $Q_{\min}$.

3 Method of virtual stations

Assume that we have p observation stations with spatial coordinates $x_1, \ldots, x_p$ and that, for all $i = 1, \ldots, p$, N_i elevation observations are available for N_i different temporal coordinates. It is plausible that, as suggested in Section 1, and as will be confirmed by forthcoming experiments, the best model for the predicted elevations at any given station should come from a least-squares fitting of a suitable polynomial using the observed associated elevations. If the degree of each polynomial is q, the number of coefficients of this model is $p(q+1)$. It is disappointing that this number is, in general, bigger than $(s+1)(q+1)$, which is the number of coefficients associated with the two-variable polynomial model discussed in Section 2. Therefore, solving (13) does not lead to the likely optimal elevation prediction, given the data availability mentioned in this paragraph.

On the other hand, the procedure based on (13) seems to be suitable for the case where one has observations at different space-time positions, not necessarily concentrated at fixed stations. In this section we will assume that available elevation data $z_{\rm obs}(x_k, t_k)$ are given at $n_{\rm dat}$ space-time points (x_k, t_k) for $k = 1, \dots, n_{\rm dat}$. We also assume that inlet discharge $Q_{\rm min}(t)$ is available whenever necessary.

We consider that $x_{\rm min} \leq \bar{x}_1 < \bar{x}_2 < \cdots < \bar{x}_{n_{\rm stat}} \leq x_{\rm max}$. Each spatial position $\bar{x}_j$ will be called "virtual station". The unknowns of our problem will be the coefficients $c_{0,j}, c_{1,j}, c_{2,j}, c_{3,j}$ for all $j = 1, \dots n_{\rm stat}$. Note that our fitting problem has $4n_{\rm stat}$ unknowns. The objective function f will be a sum of squared errors, each error corresponding to an elevation observation. Namely,

$$f(c) = \sum_{k=1}^{n_{\rm dat}} \left[z_{\rm cal}(x_k, t_k, c) - z_{\rm obs}(x_k, t_k) \right]^2, \tag{14}$$

where c is the vector of estimated coefficients c_{ij} stored columnwise and $z_{\rm cal}(x_k, t_k, c)$ is the elevation computed by the model at the point (x_k, t_k) when the model coefficients are given by the vector c.

Let us describe how $z_{\rm cal}(x_k, t_k, c)$ is computed. Given $k \in \{1, \dots, n_{\rm dat}\}$ we define $x_{{\rm left}(k)}$ as the biggest $\bar{x}_j$ such that $\bar{x}_j \leq x_k$ and we define $x_{{\rm right}(k)}$ as the smallest $\bar{x}_j$ such that $x_k < \bar{x}_j$, except in the cases that $x_k < \bar{x}_1$ or $x_k > \bar{x}_{n_{\rm stat}}$. If $x_k < \bar{x}_1$ we define $x_{{\rm left}(k)} = \bar{x}_1$ and $x_{{\rm right}(k)} = \bar{x}_2$. If $x_k > \bar{x}_{n_{\rm stat}}$ we define $x_{{\rm left}(k)} = \bar{x}_{n_{\rm stat}-1}$ and $\bar{x}_{{\rm right}(k)} = x_{n_{\rm stat}}$. The coefficients $c_{0,{\rm left}(k)}, c_{1,{\rm left}(k)}, c_{2,{\rm left}(k)}, c_{3,{\rm left}(k)}$ and $c_{0,{\rm right}(k)}, c_{1,{\rm right}(k)}, c_{2,{\rm right}(k)}, c_{3,{\rm right}(k)}$ will be the only coefficients that appear in the definition of $z_{\rm cal}(x_k, t_k, c)$.

We define

$$w_{{\rm left}(k)} = c_{0,{\rm left}(k)} + c_{1,{\rm left}(k)} Q_{\rm min}(t_k) + c_{2,{\rm left}(k)} Q_{\rm min}(t_k)^2 + c_{3,{\rm left}(k)} Q_{\rm min}(t_k)^3 \tag{15}$$

and

$$w_{{\rm right}(k)} = c_{0,{\rm right}(k)} + c_{1,{\rm right}(k)} Q_{\rm min}(t_k) + c_{2,{\rm right}(k)} Q_{\rm min}(t_k)^2 + c_{3,{\rm right}(k)} Q_{\rm min}(t_k)^3. \tag{16}$$

Finally,

$$z_{\rm cal}(x_k, t_k, c) = \frac{x_k - x_{{\rm right}(k)}}{x_{{\rm left}(k)} - x_{{\rm right}(k)}} \, w_{{\rm left}(k)} + \frac{x_k - x_{{\rm left}(k)}}{x_{{\rm right}(k)} - x_{{\rm left}(k)}} \, w_{{\rm right}(k)}. \tag{17}$$

According to (15), (16), and (17), $z_{\rm cal}(x_k, t_k, c)$ depends linearly on the unknown coefficients c. Therefore, the minimization of (14) is a linear least-squares problem. This problem has $n_{\rm dat}$ equations and $4n_{\rm stat}$ unknowns. Note that the number of virtual stations and their positions are arbitrary and should be chosen taken into account the coordinates of the available data.

3.1 Choosing virtual stations

The positions of the virtual stations $\bar{x}_1, \dots, \bar{x}_{n_{\rm stat}} \in [x_{\rm min}, x_{\rm max}]$ are "hyper-parameters" of the model presented in Section 3. The objective function in the model "with variable virtual stations" is given by (14) and each $z_{\rm cal}(x_k, t_k, c)$ is defined by (17), but $x_{{\rm right}(k)}$ and $x_{{\rm left}(k)}$ are now variables of the problem that may change in order to obtain better values of the objective function. Therefore, a more precise definition of the objective function is

$$f(c, \bar{x}) = \sum_{k=1}^{n_{\rm dat}} \left[z_{\rm cal}(x_k, t_k, c) - z_{\rm obs}(x_k, t_k) \right]^2, \tag{18}$$

where the coordinates of $\bar{x}$ are $\bar{x}_1, \ldots, \bar{x}_{n_{\text{stat}}}$ and, for all $k = 1, \ldots, n_{\text{dat}}$,

$$z_{\text{cal}}(x_k, t_k, \bar{x}, c) = \frac{x_k - x_{\text{right}(k)}}{x_{\text{left}(k)} - x_{\text{right}(k)}} w_{\text{left}(k)} + \frac{x_k - x_{\text{left}(k)}}{x_{\text{right}(k)} - x_{\text{left}(k)}} w_{\text{right}(k)}. \tag{19}$$

Let us define now an algorithm that we effectively use for choosing the coordinates of stations $\bar{x}_1, \ldots, \bar{x}_{n_{\text{stat}}}$. Let us initialize the set $\mathcal{O}$ in the following way:

$$\mathcal{O} = \{x \in [x_{\min}, x_{\max}] \text{ such that there exists } k \in \{1, \ldots, n_{\text{dat}}\} \text{ with } x = x_k\}. \tag{20}$$

Note that we could define

$$\mathcal{O} = \{x_1, \ldots, x_{n_{\text{dat}}}\},$$

but this definition should be ambiguous, inducing that the number of elements of $\mathcal{O}$ is n_{dat}. This is not the case, because x-coordinates may be repeated in the set of observations. In fact, the number of elements of $\mathcal{O}$ is less than or equal to n_{dat}. From now on, we will assume that the cardinality of $\mathcal{O}$ is not smaller than 2. Therefore, one has at least two values of spatial coordinates x for which we have at least one observation. Note that the number of elements of $\mathcal{O}$ is between 2 and n_{dat} and that this number may be strictly smaller than n_{dat}. The set of positions of the virtual stations will be called $\mathcal{S}$. It will be defined recursively in the following way:

Algorithm 3.1.1. Initialize $\mathcal{S} \leftarrow \emptyset$.

Step 1. If $\#\mathcal{S} \geq n_{\text{stat}}$ or $\mathcal{O} = \emptyset$, stop.

Step 2. Compute a solution $\widehat{x}$ of the problem

$$\underset{x \in \mathcal{O}}{\text{Maximize}} \min \{\#\mathcal{L}(x), \#\mathcal{R}(x)\} \tag{21}$$

where

$$\mathcal{L}(x) := \{k \in \{1, \ldots, n_{\text{dat}}\} \mid x_{\text{left}(k)} = x\} \text{ and } \mathcal{R}(x) := \{k \in \{1, \ldots, n_{\text{dat}}\} \mid x_{\text{right}(k)} = x\}.$$

Step 3. Update $\mathcal{S} \leftarrow \mathcal{S} \cup \{\widehat{x}\}$ and $\mathcal{O} \leftarrow \mathcal{O} \setminus \{\widehat{x}\}$ and go to Step 1.

At each iteration, the algorithm chooses the virtual station that maximizes the minimum number of available observations to determine each of the n_{stat} station cubic polynomial by means of least-square calculations. It is clear that, after a finite number of steps we have that the number of elements of $\mathcal{S}$ is n_{stat} or that $\mathcal{O}$ is empty and the algorithm stops.

4 Generation of synthetic data

In order to evaluate the effectiveness of different regression models for river predictions, we need to rely on synthetic experiments. In our present research we decided to generate synthetic data by means of integration of the Saint-Venant equations [20], which are given by

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \tag{22}$$

and

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{Q^2}{A}\right) + gA\frac{\partial z}{\partial x} + \frac{n_g^2 Q|Q|}{AR^{4/3}} = 0 \tag{23}$$

for $x \in [x_{\min}, x_{\max}]$ and $t \in [t_{\min}, t_{\max}]$, where $h(x,t) = z(x,t) - z_b(x)$ is the depth of the river at (x,t), $A(x,t) = h(x,t)\,w(x)$ is the cross wetted area at (x,t), $P(x,t) = w(x) + 2h(x,t)$ is the wetted perimeter at (x,t), $R(x,t) = A(x,t)/P(x,t)$ is the hydraulics radius at (x,t), $V(x,t) = Q(x,t)/A(x,t)$ is the speed of the fluid at (x,t), and g is the acceleration of gravity taken as $9.81 m/s^2$. Equation (22) describes mass conservation and equation (23) represents conservation of the linear momentum. The coefficient n_g is known as Manning roughness coefficient. It is unclear in which way this coefficient depends on x or t. On the one hand, the roughness coefficient depends on x due to the morphological differences of the river along its course. On the other hand, sediment deposition can also affect the roughness coefficients over time. In (23), n_g has units $m^{1/6}$.

The Saint-Venant equations were solved approximately by means of an explicit diffusive finite-difference method [13, 19] with the following specifications:

- $x_{\min} = 0$ and $x_{\max} = 3000$ (meters).
- $t_{\min} = 0$ and $t_{\max} = 29 + \frac{23}{24}$ (days) or, equivalently, 719 hours or 2,588,400 seconds.
- Initial conditions $z(x, t_{\min})$ given in Figure 3 and $Q(x, t_{\min}) = 3.9\,\mathrm{m}^3/\mathrm{s}$ for all $x \in [x_{\min}, x_{\max}]$.
- Boundary condition $Q(x_{\min}, t)$ given in Figure 4.
- Manning coefficient $n_g(x) = 0.078$ for all $x \in [x_{\min}, x_{\max}]$.
- Time step $\Delta t = 1$ second, spatial step $\Delta x = 30$, and diffusion coefficient 0.99.

Note that, according to the considered discretization, the finite difference method computes the values of $z(x,t)$ and $Q(x,t)$ at 101×2588401 points. We store only the values of $z(x,t)$ and $Q(x,t)$ for $x = 0, 30, 60, \dots, 3000$ meters and for $t = 0, 1, 2, \dots, 719$ hours. In other words, the "observed" elevations are given by a matrix of 101×720 positions. The level sets defined by this matrix is given in Figure 5.

5 Numerical Experiments

The data used in the numerical experiments are generated as described in Section 4. The employment of synthetic data allows us to test regression models in situations in which real data are not available.

5.1 Single-station one-dimensional models

In this short subsection, using synthetic data, we perform the same one-dimensional models experiment described in Section 1. In this case we use the stations defined by $x = 720\,\mathrm{m}$ and $x = 3000\,\mathrm{m}$. We wish to verify whether the performance of the polynomial one-dimensional models for reproducing synthetic data is similar to the performance reported for real data in Section 1. Figures 6 and 7 and Tables 3 and 4 show the results. Clearly, in terms of quality of fitting and predictions, the performance of the polynomial models using synthetic data is similar to the one that has been reported in Section 1 for data of the real Fork River.

5.2 Experiments using observations on a mesh

In this subsection we consider as observations data between days $t = 3$ and $t = 1010$, every 12 hours, at 26 equally spaced stations between $x_{\min} = 0$ and $x_{\max} = 3000$ meters. The objective is, with these meshed data, to predict the elevation $z(x,t)$ at 26 the equally spaced stations between $x_{\min} = 0$ and $x_{\max} = 3000$ meters and $t \in \{11, 12, \dots 29\}$. We consider six different ways of prediction by combining

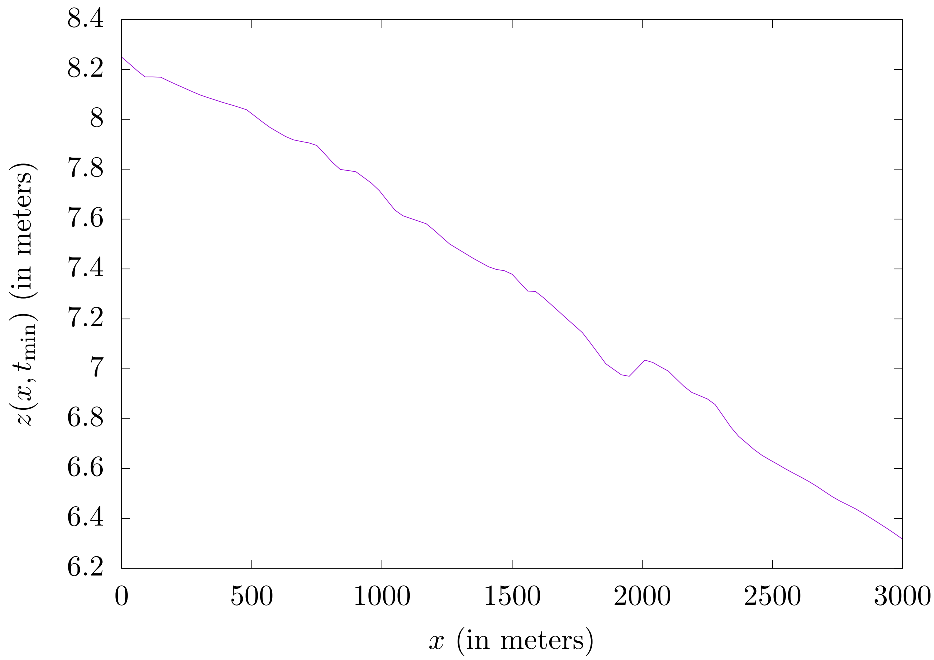


Figure 3: Initial condition for z used in the generation of synthetic data.

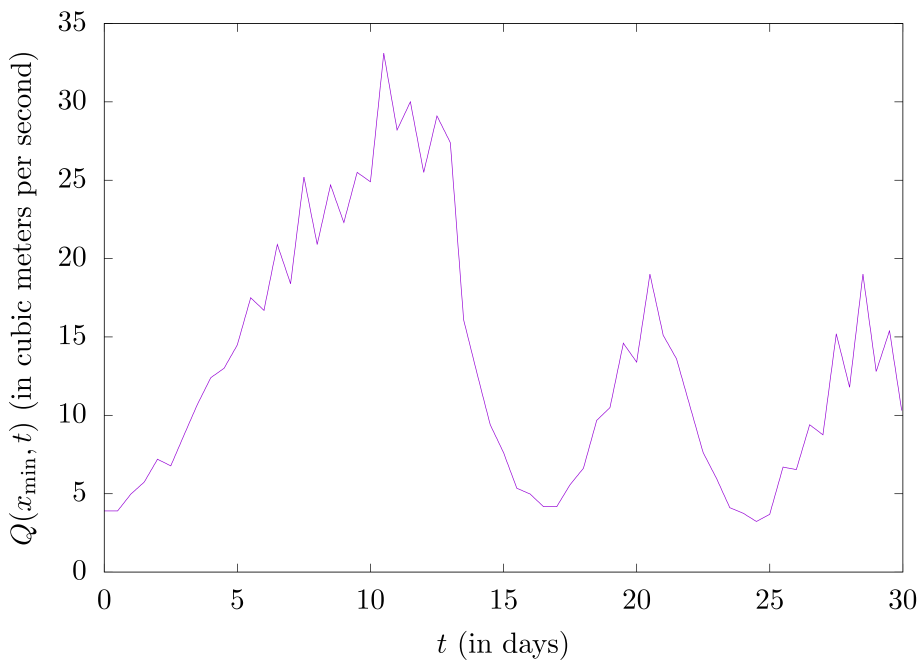


Figure 4: Q boundary condition used in the generation of synthetic data.

two types of polynomials (interpolating and least squares) and three possible degrees (linear, quadratic and cubic). Specifically, each of the six experiments consists of:

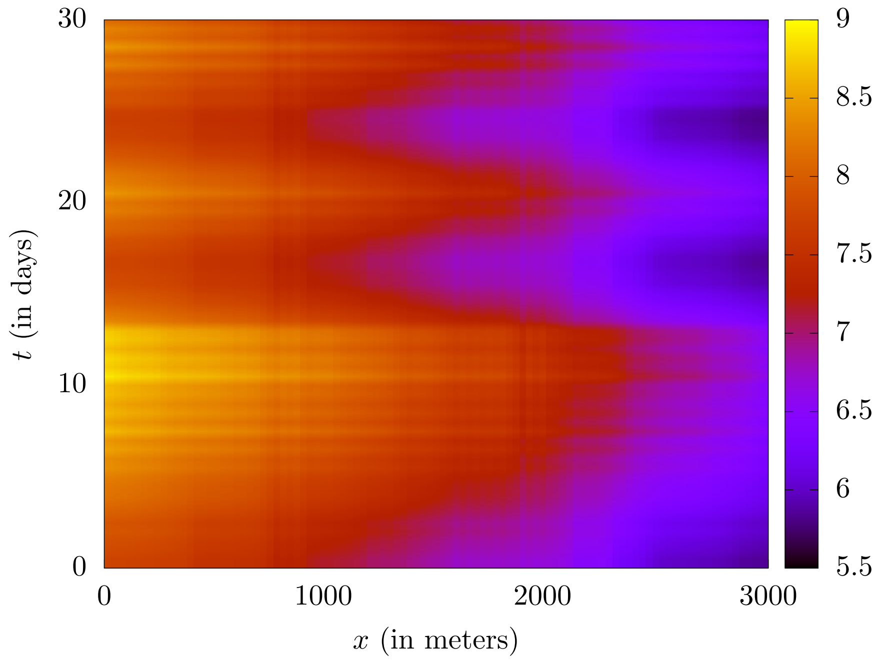


Figure 5: Synthetic Elevations.

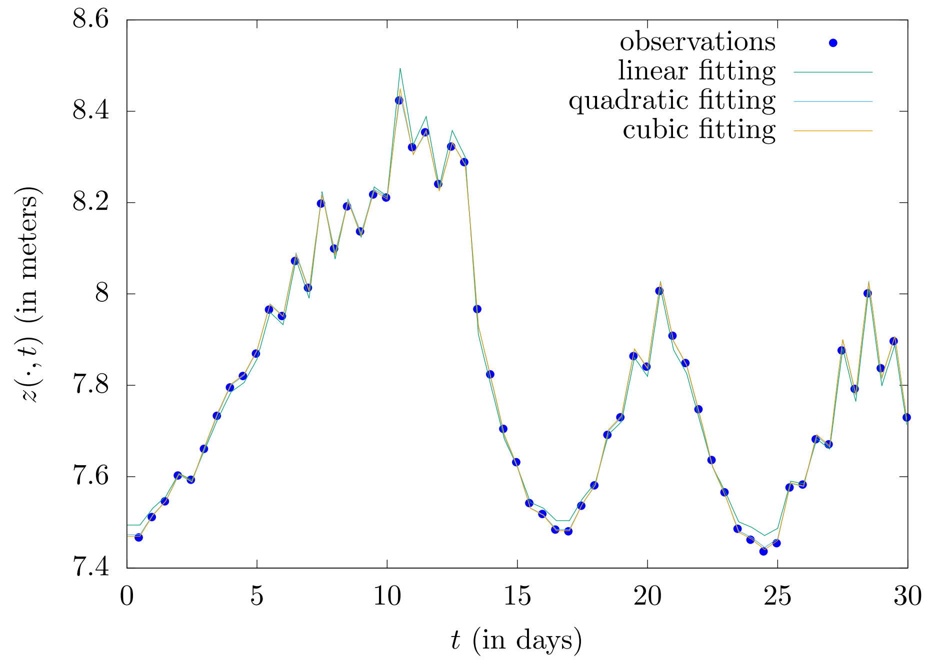
Station	Polynomial	RMSD	c_0	c_1	c_2	c_3
720 m	linear	1.68217006E-02	7.36053069	3.42611821E-02	−−	−−
	quadratic	3.84570654E-03	7.30939110	4.29580343E-02	-2.70630819E-04	−−
	cubic	2.59511479E-03	7.29329906	4.73743343E-02	-5.86479099E-04	6.35214794E-06
3000 m	linear	4.19519692E-02	5.81624209	2.58508659E-02	−−	−−
	quadratic	1.30136586E-02	5.69169716	4.70311095E-02	-6.59092116E-04	−−
	cubic	5.98388456E-03	5.62617288	6.50135886E-02	-1.94517664E-03	2.58649476E-05

Table 3: Section 5.1. Fitted polynomials, their coefficients and the corresponding RMSD using synthetic data. Observations up to 30 days.

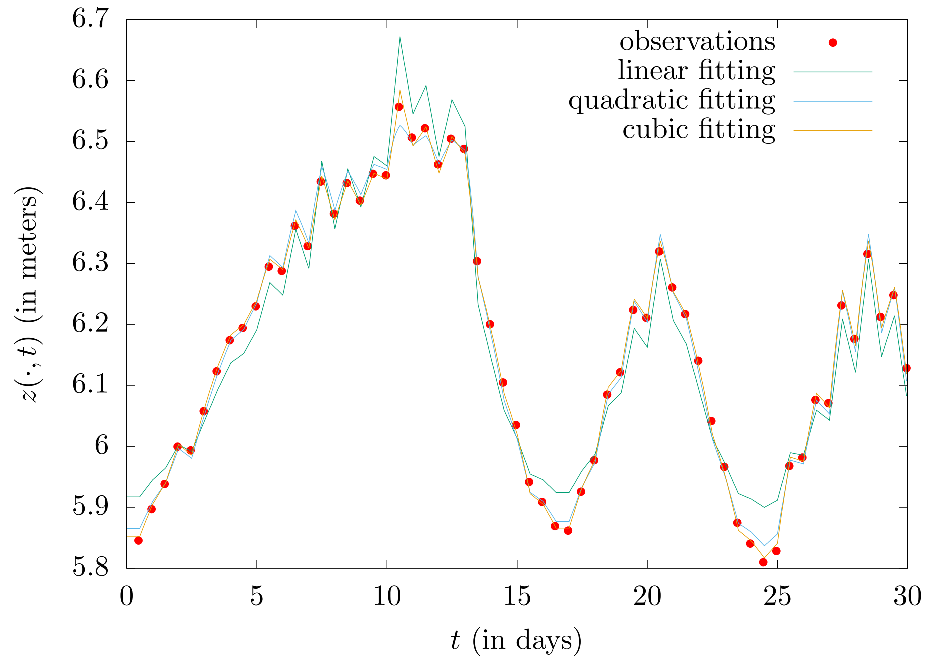
Station	Polynomial	RMSD		c_0	c_1	c_2	c_3
		training	testing				
720 m	linear	1.16720184E-02	1.94601125E-02	7.35751723	3.46560681E-02	−−	−−
	quadratic	2.20001211E-03	5.41537380E-03	7.30850076	4.32306214E-02	- 2.86981705E-04	−−
	cubic	1.67445669E-03	3.33740105E-03	7.29464777	4.71442613E-02	- 5.87655387E-04	6.73839556E-06
3000 m	linear	3.07334766E-02	4.72819963E-02	5.81094708	2.65408393E-02	−−	−−
	quadratic	7.34536455E-03	2.00261102E-02	5.68333537	4.88642187E-02	- 7.47141135E-04	−−
	cubic	3.07346962E-03	1.08396889E-02	5.61856923	6.71614492E-02	- 2.15286467E-03	3.15036593E-05

Table 4: Section 5.1. Fitting polynomials, their coefficients, and the corresponding RMSD using synthetic data. In this case, observations of the first 10 days were used as training data to fit the polynomials. Observations of the remaining 20 days were considered unknown in the fitting phase and then they were used to test predictions given by the fitted polynomials.

Experiment 1: We assume that observed elevations correspond to instants $t_1 = 9.5$ and $t_2 = 10$ days



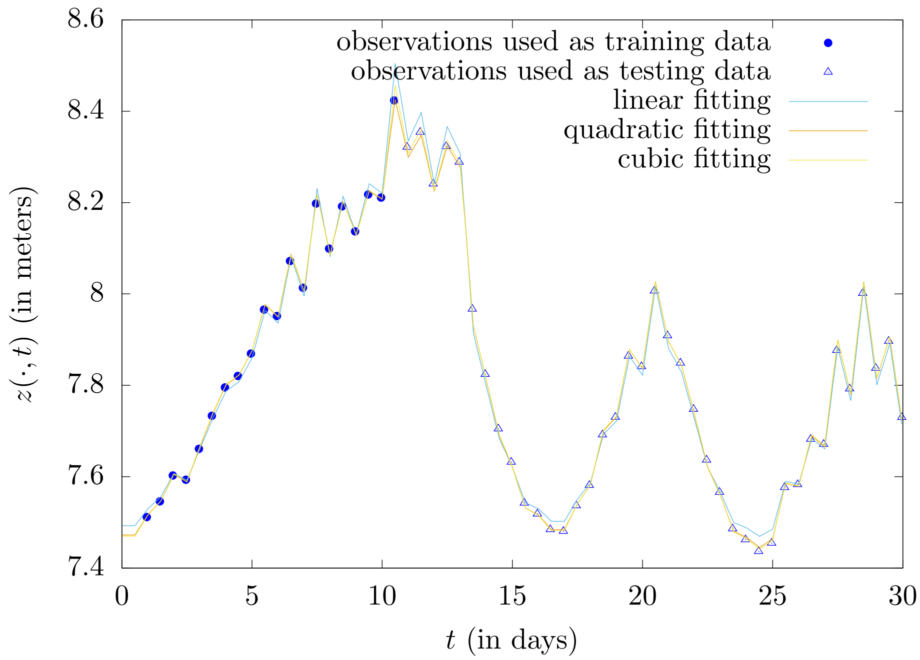
(a) Section $x = 720\,\text{m}$.



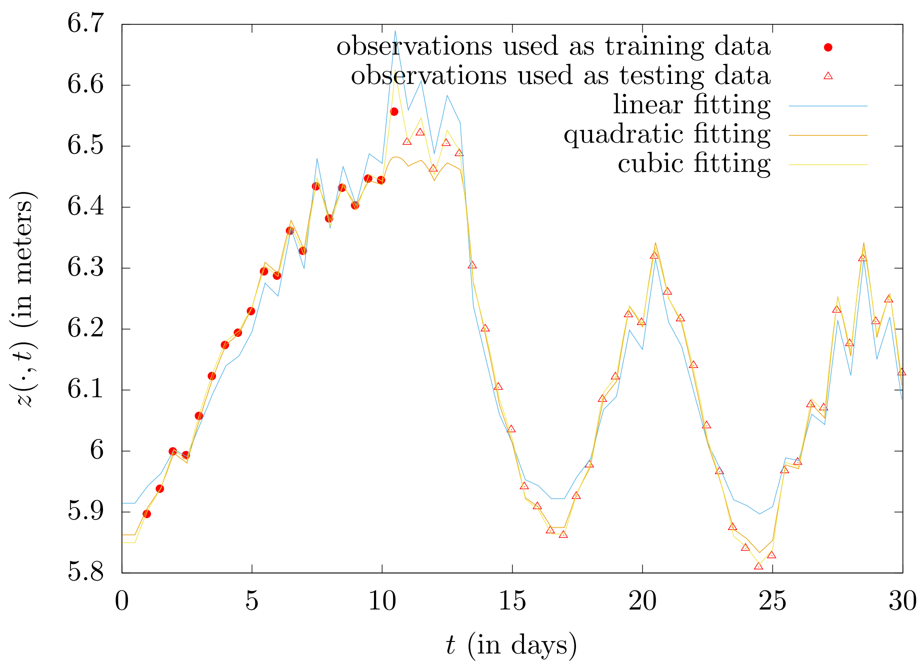
(b) Section $x = 3000\,\text{m}$.

Figure 6: Section 5.1. Synthetic observed elevations at a given station and their approximations as (linear, quadratic, and cubic) polynomials of the inlet discharge. Observations up to 30 days were used to fit the polynomials

and 26 equally spaced stations between $x_{\min} = 0$ and $x_{\max} = 3000$ meters. We consider that the inlet discharge $Q_{\min}(t)$ at times t_1 and t_2 are also observed. We employ the model (5) with



(a) Section $x = 720\,\text{m}$.



(b) Section $x = 3000\,\text{m}$.

Figure 7: Section 5.1. Synthetic observed elevations at a given station and their approximation as a (linear, quadratic, and cubic) polynomial of the inlet discharge. In this case, observations of the first 10 days were used as training data to fit the polynomials. Observations of the remaining 20 days were considered unknown in the fitting phase and, then, they were used to test predictions produced by the fitted polynomials.

$q = 1$ and $s = p - 1 = 25$. Note that, due to Theorem 2.1, it is not necessary to fit explicitly a polynomial with degree 25 in order to obtain predictions for the future at the given stations. Using this fitting, and considering suitable forecasts for the inlet discharges, we can predict elevations for days $11, 12, 13, \ldots, 29$ for 101 values of x equally spaced between $x_{\min}$ and $x_{\max}$ and we can compare these predictions with the observed elevations. Note that, in this case, the RMSD-error corresponding to the training set is necessarily equal to 0. The result of this experiment is given in Table 9.

Experiment 2: Observed elevations correspond to instants $t_1 = 9, t_2 = 9.5$, and $t_3 = 10$ days. Elevation data correspond to these instances and the model (5) uses $q = 2$ and $p - 1 = 25$. So, the elevation at each station is modelled by a quadratic interpolating polynomial. The result of this experiment is given in Table 10.

Experiment 3: Observed elevations correspond to instants $t_1 = 8.5, t_2 = 9, t_3 = 9.5$, and $t_4 = 10$ days. Elevation data correspond to these instances and the model (5) uses $q = 3$ and $p-1 = 25$. So, the elevation at each station is modelled by a cubic interpolating polynomial. The result of this experiment is given in Table 11.

Experiment 4: Observed elevations correspond to instants $t \in \{3.5, 4, 4.5, \ldots, 10\}$ days. Elevation data correspond to these instances and the model (5) is a line that fits the observed elevations at those instants in the least-squares sense. The result of this experiment is given in Table 12.

Experiment 5: Observed elevations correspond to instants $t \in \{3.5, 4, 4.5, \ldots, 10\}$ days. Elevation data correspond to these instances and the model (5) is a quadratic polynomial that fits the observed elevations at those instants in the least-squares sense. The result of this experiment is given in Table 13.

Experiment 6: Observed elevations correspond to instants $t \in \{3.5, 4, 4.5, \ldots, 10\}$ days. Elevation data correspond to these instances and the model (5) is a cubic polynomial that fits the observed elevations at those instants in the least-squares sense. The result of this experiment is given in Table 14.

Tables 9–14 in the Appendix show the results. Figures 8 and 9 give a graphical representation of the predictions' RMSD as a function of $t \in \{11, 12, \ldots, 29\}$. For each t, the RMSD of the 26 equally spaced $x \in [0, 3000]$ meters is shown. The experiment shows that polynomial interpolators of past data are bad at extrapolating to predict the future. One reason may be that they are based on little data and focus on capturing local behavior. Thus, the linear and quadratic options are less bad than the cubic, which quickly goes to infinity under the influence of local behavior. On the other hand, least squares polynomials computed with more data better capture the trend implicit in the data and thus better predict the future. Of the three options (linear, quadratic, and cubic), the cubic provides the best predictions.

5.3 Next-day predictions using observations on a mesh

In this experiment, we evaluate the six approaches considered in the previous subsection to predict the "elevation of the next day". We consider $t_{\text{today}} \in \{2, 3, \ldots, 28\}$ days and $t_{\text{tomorrow}} = t_{\text{today}} + 1$. Available data of $z(x, t)$ with t multiple of half day and $t \leq t_{\text{today}}$ was used as training data. For the interpolating polinomials, only the most recent information was considered, while for least squares, all available data was considered. For each of the 26 equally spaced stations x between $x_{\min} = 0$ and $x_{\max} = 3000$ meters, the six approaches were used to predict the elevation $z(x, t_{\text{tomorrow}})$. Table 5 shows the details. As seen in previous experiments, least squares polynomials gave very reasonable

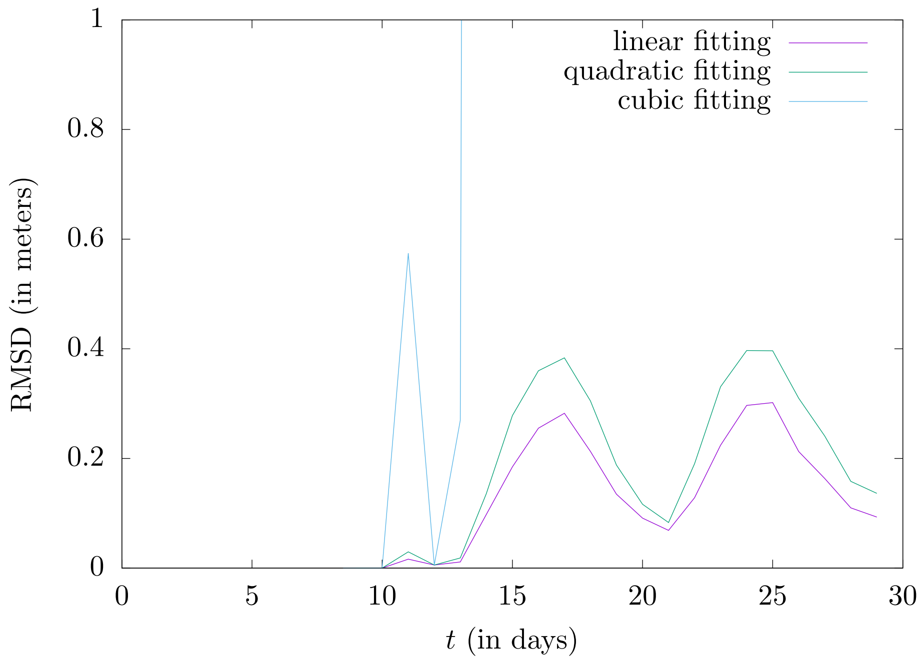


Figure 8: Section 5.2. RMSD of predictions of $z(x,t)$ for $t \in \{10, 11, \dots, 29\}$ when predictions are given by interpolating polinomials (linear, quadratic, and cubic) computed using training data with $t < 10$. For each t, the RMSD of the 26 equidistant $x \in [0, 3000]$ meters is being displayed.

predictions (with an average error of 1 centimeter in the case of the cubic polynomial) and performed better than interpolating polynomials. As expected, the cubic was better than the quadratic, which was better than the linear. Unlike previous experiments, interpolating polynomials were also useful in many cases, because in the present experiments we are dealing with next-day predictions, i.e. interpolating polynomials are used to extrapolate only a little outside the interpolating range.

5.4 Next-day predictions using irregularly distributed data

In the experiments of the previous subsection, we considered observations every 12 hours between day $t = 3$ and day $t = 10$ (15 time instants) at 26 stations equidistant between 0 and 3000 meters, totalizing 390 observations. However, considering our synthetic data, in that same domain of space (x, t) we have available data from hour to hour and at 101 equidistant stations, amounting to $101 \times 169 = 17069$ available data. With the intuition of using random subsets of data with uniform distribution, in the next experiment we draw the observations among the available data with probability $\frac{390/17069}{\nu} \approx \frac{0.0288}{\nu}$, with $\nu \in \{1, 2, 4\}$. With this way of determining the observations, we constituted training data sets with 394, 195, and 95 elevation observations.

The experiment consists of (a) positioning n_{stat} stations using Algorithm 4.1, (b) with the stations already positioned solve the linear least squares problem (18) that computes the cubic polynomial of each station, and (c) use those polynomials to predict the elevation of the "next day", that is, the day $t_{\text{tomorrow}} = 11$ at 101 equidistant points between 0 and 3000 meters. We wish to understand how the predictions behave for different values of n_{stat}.

Table 6 shows the results when 394 observations are available with the number of virtual stations varying from 2 to 100. The first column shows the number of stations. The second column reports

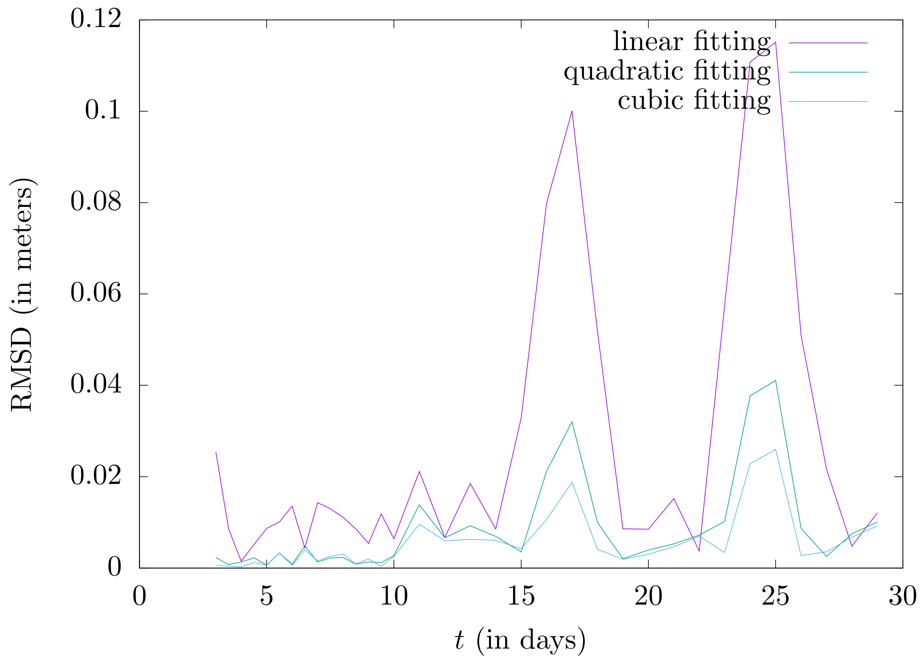


Figure 9: Section 5.2. RMSD of predictions of $z(x,t)$ for $t \in \{10, 11, \ldots, 29\}$ when predictions are given by best fitting polinomials (linear, quadratic, and cubic) computed by solving a linear least squares problem using training data with $t \in \{3.5, 4, 4.5, \ldots, 9.5\}$. For each t, the RMSD of the 26 equidistant $x \in [0, 3000]$ meters is being displayed.

"minobs", the minimum number of observations that were used, given the positions of the virtual stations, to determine each of the n_{stat} cubic polynomials by means of least-square calculations. The third column shows the RMSD of the training data. The last column shows the RMSD of the next-day prediction at the 101 points equidistant between 0 and 3000 meters. It is clear from the figures in the table that the RMSD of the training data decreases monotonically as the number of stations increases. On the other hand, the RMSD of the next-day prediction remains more or less constant (between 3 and 6 centimeters) when the number of virtual stations is between 2 and 49 and deteriorates rapidly when this number is 50 or more. In fact, the optimal number of virtual stations is, in this case, 19. For completeness, in this case we report the results up to 100 virtual stations.

In Table 7 and Table 8 we report the same type of results when the number of available observations is 185 and 95, respectively. In the first case, the number of virtual stations goes from 2 to 37 and in the second case it goes from 2 to 19 because larger numbers of virtual stations yield prediction errors that are bigger than 1 meter. Again, the error in the training set decreases with the number of virtual stations, as the number of free parameters is increased.

6 Conclusions

This paper discusses the potential of methods based on surface elevation data alone for predicting river levels, provided that reliable inlet discharge forecasts $Q(x_{\min}, t)$ are available. We have focused on low-degree polynomial models because they are simple and economical in terms of the number of unknown parameters. The various alternatives presented in this paper can be considered successful in the sense that they provide results that are accurate enough for predicting the levels of real rivers.

t_{today}	Interpolating polynomial of degree:			Fitting polynomial of degree:		
	1	2	3	1	2	3
2	0.0062	0.0069	0.0199	0.0141	0.0021	0.0199
3	0.0130	0.0679	0.4752	0.0421	0.0219	0.0426
4	0.0051	0.0011	0.0033	0.0302	0.0108	0.0057
5	0.0014	0.0154	0.0348	0.0329	0.0142	0.0043
6	0.0086	0.0120	0.0150	0.0248	0.0094	0.0022
7	0.0071	0.0071	0.0071	0.0266	0.0105	0.0052
8	0.0021	0.0005	0.0141	0.0176	0.0028	0.0047
9	0.0035	0.0038	0.0041	0.0303	0.0075	0.0030
10	0.0163	0.0295	0.5743	0.0461	0.0225	0.0055
11	0.0083	0.0018	0.0056	0.0035	0.0024	0.0056
12	0.0008	0.0030	0.0034	0.0133	0.0013	0.0037
13	0.0905	0.0246	1.1340	0.0357	0.0131	0.0053
14	0.0199	0.0089	0.0180	0.0043	0.0113	0.0081
15	0.0162	0.0038	0.0069	0.0484	0.0079	0.0010
16	0.0063	0.0122	0.0163	0.0600	0.0165	0.0059
17	3.7608	1.0794	0.2602	0.0075	0.0045	0.0021
18	0.0290	0.0174	3.9772	0.0297	0.0090	0.0008
19	0.0168	0.0198	0.0503	0.0391	0.0057	0.0012
20	0.0104	0.0121	0.0124	0.0405	0.0046	0.0021
21	0.0281	0.0142	0.3077	0.0299	0.0129	0.0055
22	0.0353	0.0331	0.0751	0.0174	0.0052	0.0060
23	0.0173	0.0048	0.0022	0.0656	0.0194	0.0076
24	0.0018	0.0015	0.0014	0.0635	0.0199	0.0094
25	0.0126	0.7339	3.3581	0.0019	0.0079	0.0043
26	0.0806	0.1488	0.2467	0.0230	0.0113	0.0035
27	0.0224	0.0517	0.3592	0.0396	0.0099	0.0030
28	0.0075	0.0044	0.0029	0.0430	0.0107	0.0054
	0.7243	0.2537	1.0417	0.0353	0.0118	0.0102

Table 5: Section 5.3. For a given day t_{today} and a given experiment (interpolating or fitting polynomial of degree 1, 2, or 3) the table shows the RMSD of the next-day predicted elevation of all 26 stations. The last row shows the overall RMSD of each approach.

In particular, the strategy of virtual stations presented in this paper seems to be useful in the case where observations are irregularly distributed. Moreover, this strategy preserves the best third-order polynomial approximations in the regularly distributed case.

It is interesting to consider the problem of predicting flow-rates $Q(x,t)$ from elevation observations $z(x,t)$ only. From the mass-conservation equation we have that

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0.$$

Therefore,

$$Q(x,t) = Q(x_{\min},t) - \int_{x_{\min}}^{x} \frac{\partial A}{\partial t}(\xi,t)d\xi.$$

Virtual stations	minobs	RMSD training	RMSD testing	Virtual stations	minobs	RMSD training	RMSD testing
2	394	5.3283820057185995E-002	6.1583164478249554E-002	52	10	8.0329354317997658E-003	0.12968356044883630
3	194	3.2648300820944318E-002	4.3726318071762964E-002	53	10	8.0306312766861461E-003	0.12972255677477529
4	101	3.2573813718738569E-002	4.5187290605368642E-002	54	10	8.0187525830073773E-003	0.12981398947867417
5	96	3.2349694130166182E-002	4.6027937601361361E-002	55	9	8.0179333149921674E-003	0.12982155115549487
6	96	3.1639597470529801E-002	4.8685500605615238E-002	56	9	8.0079086981791198E-003	0.12997349516966566
7	96	2.2853081231014000E-002	3.5410636986479435E-002	57	9	7.9116879215693353E-003	0.13080885606219347
8	47	2.2333106995790799E-002	3.4395169668978959E-002	58	9	7.1120982718531822E-003	0.12975686606923589
9	47	2.2005150036587609E-002	3.5766673494248759E-002	59	9	7.0527831979785103E-003	0.12960764964637364
10	47	1.9266463631366117E-002	3.4100377670761683E-002	60	9	6.8200629596160949E-003	0.12970197134584049
11	47	1.8330036730951786E-002	3.8099252360178991E-002	61	9	6.8099002730974117E-003	0.13443085010705041
12	47	1.8247387283753448E-002	3.9287689605826834E-002	62	9	6.7569652637862266E-003	0.13432463882639850
13	47	1.6544563006014871E-002	3.6201727385641792E-002	63	8	6.7551245197326791E-003	0.13450127671524303
14	46	1.6109150476189795E-002	3.4048905093485481E-002	64	8	6.7356313704111668E-003	0.14227538180090210
15	46	1.5513927451101283E-002	3.2651624306286674E-002	65	8	6.7297367192108533E-003	0.15228129547423630
16	26	1.5337223488082084E-002	3.2449894571859761E-002	66	8	6.7266867935987414E-003	0.71829333520509586
17	26	1.5158717471376091E-002	3.2345999462907726E-002	67	3	6.7264323627948203E-003	0.71833153306163988
18	13	1.5156034949535248E-002	3.2305120343670203E-002	68	3	4.7223533876023351E-003	0.71808112648009770
19	13	1.5034755623023519E-002	3.1751814157051458E-002	69	3	4.6359280777656803E-003	0.80170172105254001
20	13	1.5018746780348795E-002	3.2280097922278560E-002	70	3	4.6330920332480260E-003	0.84002240218473490
21	13	1.4923539265815706E-002	3.3525846097928733E-002	71	3	4.6122345451842994E-003	0.83973278380279726
22	13	1.4859547396657539E-002	3.4453557117947238E-002	72	3	4.6102751068605808E-003	1.5559016307267453
23	13	1.3691065318011980E-002	5.3843426618723260E-002	73	3	4.5875787632858409E-003	1.5560072582416900
24	13	1.3659881827357711E-002	5.4143100409803191E-002	74	3	4.4777490065165777E-003	1.5562018595829066
25	13	1.3263279056685832E-002	5.3164207679156618E-002	75	3	4.4755600778543063E-003	1.6023004097916536
26	13	1.3175509007241182E-002	5.2752211202618811E-002	76	3	4.3307476396778153E-003	1.6026236880641358
27	13	1.2893886036635917E-002	5.3294152115026458E-002	77	3	3.9620225228191950E-003	1.6027498818868846
28	13	1.2272901759381408E-002	5.3934282047854318E-002	78	3	2.8886317493116574E-003	1.6104032246579405
29	13	1.2090872110758055E-002	5.5372017107663166E-002	79	3	2.8869280012161587E-003	1.6211332816669319
30	13	1.2043880504159150E-002	5.6348350230236106E-002	80	3	2.8599432980772987E-003	1.8984411394933036
31	13	1.1923895103591405E-002	5.6647214439744825E-002	81	3	2.7391034801038162E-003	1.8988767221527962
32	13	1.1864076693219016E-002	5.6304646204128742E-002	82	3	2.5652606094846665E-003	1.9160614771120590
33	13	1.1682392020077523E-002	5.7208920168502021E-002	83	3	2.5267865337923233E-003	1.9168010751012954
34	13	1.1670097881090850E-002	5.6984529825676034E-002	84	3	2.3775396132823608E-003	2.3251907404973591
35	13	1.1628659265365073E-002	5.5947727821659403E-002	85	3	2.3001696336729283E-003	2.3255810361822524
36	13	1.0416666600669637E-002	3.7676481978826011E-002	86	3	2.2706294451196921E-003	2.3784129971256784
37	13	1.0399791281676873E-002	3.7889247768345305E-002	87	3	2.1418982774388468E-003	2.3898547222289541
38	13	1.0383501483021772E-002	3.7653477041362522E-002	88	3	2.1403856079314533E-003	2.3907039892197419
39	13	9.9227025361913641E-003	3.8432530450767118E-002	89	3	2.1403856079314563E-003	2.4025492211247181
40	13	9.8817703068327638E-003	3.7475166481987364E-002	90	3	2.1403856079314503E-003	2.4257624451134867
41	12	9.5545957721222922E-003	3.8662642310044168E-002	91	3	1.2241141853634287E-003	2.7471502163021850
42	12	9.5259047134240975E-003	3.8864212195983724E-002	92	3	1.1935575705138712E-003	2.7739273204290273
43	12	9.5068693663715384E-003	3.9334825174204911E-002	93	3	1.1935575705138105E-003	2.7914574146435789
44	12	9.4031730453848442E-003	4.0138579708408451E-002	94	3	1.1935575705138755E-003	10.769455041417071
45	11	9.3878986230084352E-003	4.0381481232975115E-002	95	3	1.1935575705138961E-003	13.194184220825461
46	11	9.2827157053270003E-003	4.1040039457329369E-002	96	3	1.1919087363459152E-003	15.209373170340974
47	11	9.1301445568780712E-003	4.3543218230213378E-002	97	3	1.0625333177912046E-003	16.484479467068468
48	11	8.4370965844998008E-003	4.3316691550977185E-002	98	3	1.0625333177911834E-003	18.255510661328792
49	11	8.4166324951854554E-003	4.3061145849278733E-002	99	3	1.0625333177911502E-003	18.414545717188812
50	11	8.2245582976612011E-003	0.12923190100341689	100	3	1.0625333177912118E-003	18.510711323780274
51	10	8.2013285286650847E-003	0.12971323080775976				

Table 6: Section 5.4. 394 random observations in the first 10 days. Effect of increasing the number of virtual stations. Reporting training and test.

So, if we have a good approximation for $A(x,t)$, then we can obtain, in principle, a good approximation of $Q(x,t)$ [6]. Moreover, according to the results of the present paper, a good approximation of $z(x,t)$ can be obtained using elevation observations and $Q_{\min}$ forecasts. Unfortunately, the cross wetted area $A(x,t)$ can be obtained from $z(x,t)$ only if we already know the bed elevation $z_b(x)$ and the geometric characteristics of the channel. This is the information that we have considered uncertain, and whose use we have tried to avoid above under the "only elevation" approach of the present paper. Therefore, predicting flow rates from data alone is more problematic than predicting surface elevations, and this issue deserves future study.

Virtual stations	minobs	RMSD training	RMSD testing
2	185	5.1248811463696108E-002	7.4089510567510924E-002
3	91	3.2529197002708163E-002	3.5519815707586833E-002
4	47	3.1284750173733118E-002	4.3991827708239235E-002
5	46	2.9878458008594868E-002	8.3143696581750387E-002
6	46	2.2450370975549642E-002	6.8068533438989373E-002
7	22	2.1764814785424559E-002	6.5415859677875707E-002
8	22	2.0766831805240742E-002	6.7534463202462633E-002
9	22	2.0573231136337876E-002	6.6790025825457275E-002
10	22	2.0104363353167974E-002	6.3158699486610168E-002
11	22	1.8826320026367610E-002	7.3860255266057703E-002
12	22	1.8270195922934992E-002	9.1669436715929684E-002
13	22	1.5006593710845378E-002	7.7134955384989309E-002
14	10	1.4856901876581540E-002	7.7603153454404841E-002
15	10	1.4558586705511977E-002	7.0383131196731591E-002
16	10	1.4367127212092557E-002	7.2430701225312186E-002
17	10	1.4056078954309907E-002	7.6150540705157130E-002
18	10	1.3944065662380992E-002	7.3380393469394803E-002
19	10	1.2663706268095641E-002	9.2371846936037255E-002
20	10	1.2527967989703639E-002	9.3799256133081668E-002
21	10	1.2452824633423691E-002	9.2679586134847808E-002
22	10	1.2430540420894570E-002	9.2397760084283950E-002
23	8	1.2308000711774588E-002	0.10149722107077552
24	8	1.2170734928201659E-002	0.10745970660633239
25	8	1.0773903953212724E-002	0.11569597875671561
26	8	1.0703768079820394E-002	0.12342432166878507
27	8	8.0880895546961377E-003	0.13473581683984887
28	5	8.0856468697171491E-003	0.13471641725588770
29	5	8.0624368937640255E-003	0.13730815995101689
30	5	7.5576773867298457E-003	0.14055458227230450
31	5	7.4451793362722103E-003	0.14283338653559183
32	5	7.3306658105015731E-003	0.15803100201596343
33	5	6.9433436122286933E-003	0.20812431672514720
34	5	6.7419124067725706E-003	0.84362496442554258
35	5	6.7171350118954724E-003	0.84673106438894241
36	5	5.1430668400310421E-003	0.89984696571610800
37	5	2.7526193304603067E-003	1.5797211827586006

Table 7: Section 5.4. 185 random observations in the first 10 days. Effect of increasing the number of virtual stations. Reporting training and test RMSD.

References

[1] A. Agresta, M. Baioletti, C. Biscarini, F. Caraffini, A. Milani, and V. Santucci, Using optimisation meta-heuristics for the roughness estimation problem in river flow analysis, *Applied Sciences* 11, 10575, 2021.

Virtual stations	minobs	RMSD training	RMSD testing
2	95	5.1699965039000095E-002	6.3173540327799205E-002
3	47	3.1098552914075955E-002	3.7681240571308741E-002
4	23	3.0008346683687421E-002	5.4680180663268317E-002
5	23	2.7343428686985052E-002	0.10084121692806179
6	13	2.6263510371117418E-002	0.12687890844372421
7	13	2.5469694685872728E-002	0.14131992067387011
8	11	2.5009773875898412E-002	0.14894525241456291
9	11	1.8148205210940294E-002	9.3893818673443930E-002
10	11	1.7516836515265203E-002	0.13792608105500823
11	11	1.4868146929722386E-002	0.14102668906719881
12	7	1.4669480794577859E-002	0.14028858123690985
13	7	1.4344105609466579E-002	0.15561454960887050
14	7	1.3939622987893766E-002	0.17476910376467428
15	5	1.3899191317637099E-002	0.17136513895286537
16	5	1.3641822085803633E-002	0.17664288679815060
17	5	8.7311056846887617E-003	0.55468488663305626
18	5	8.4229878615030840E-003	0.55920223834725657
19	5	3.6141721204698040E-003	2.0837223835016192

Table 8: Section 5.4. 95 random observations in the first 10 days. Effect of increasing the number of virtual stations. Reporting training and test RMSD.

[2] M. T. Ayvaz, A linked simulation-optimization model for simultaneously estimating the Manning's surface roughness values and their parameter structures in shallow water flows, *Journal of Hydrology* 500, pp. 183–199, 2013.

[3] M. Kh. Askar and K. K. Al-jumaily, A nonlinear optimization model for estimating Manning's roughness coefficient, in *Proceedings of the Twelfth International Water Technology Conference, IWTC12*, Alexandria, Egypt, 2008, pp. 1299–1306.

[4] E. G. Birgin, M. R. Correa, V. A. González-López, J. M. Martínez, and D. S. Rodrigues, Randomly supported models for the prediction of flows in channels, Technical Report of CRIAB-Mathematical Engineering, *submitted*.

[5] E. G. Birgin and J. M. Martínez, Accelerated derivative-free nonlinear least-squares applied to the estimation of Manning coefficients, *Computational Optimization and Applications* pp. 689–715, 2022.

[6] E. G. Birgin and J. M. Martínez, A PDE-informed optimization algorithm for river flow predictions, *Numerical Algorithms*, to appear (DOI: 10.1007/s11075-023-01647-1).

[7] Y. Ding, Y. Jia, and S. S. Y. Wang, Identification of Manning's roughness coefficients in shallow water flows, *Journal of Hydraulic Engineering*, pp. 501–510, 2004.

[8] Y. Ding and S. S. Y. Wang, Identification of Manning's roughness coefficients in channel network using adjoint analysis, *International Journal of Computational Fluid Dynamics* 19, pp. 3–13, 2005.

[9] W. W. Emmett, W. W. Myrick, and R. H. Meade, Field data describing the movement and storage of sediment in the East Fork River, Wyoming, Part 1, *River Hydraulics and Sediment Transport*, Report No. 1, 1979.

[10] G. H. Golub and Ch. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 2013.

[11] K. Guta and K. S. H. Prasad, Estimation of open channel flow parameters by using optimization techniques, *International Journal of Science and Research* 6, pp. 1295–1304, 2018.

[12] Y. Jia and S. S. y. Wang, 2001, CCHE2D: Two-dimensional Hydrodynamic and Sediment Transport Model for Unsteady Open Channel Flows Over Loose Bed, National Center for Computational Hydroscience and Engineering Technical Report No.: NCCHE-TR-2001-1, 2001.

[13] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Lectures in Mathematics, ETH Zürich, Birkäuser, 1992.

[14] W. A. Marcus, K. Roberts, L. Harvey, and G. Tackman, An evaluation of methods for estimating Manning's n in small mountain streams, *Mountain Research and Development* 12, pp. 227–239, 1992.

[15] J. M. Martínez and L. T. Santos, Inexact-restoration modelling with monotone interpolation and parameter estimation, *Optimization and Engineering*, to appear (DOI: 10.1007/s11081-023-09861-5).

[16] R. H. Meade, W. W. Myrick, and W. W. Emmett, Field data describing the movement and storage of sediment in the East Fork River, Wyoming, Part 2, *River Hydraulics and Sediment Transport*, Report No. 2, 1979.

[17] F. Pappenberger, K. Beven, M. Horrit, and S.Blazkova, Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations, *Journal of Hydrology* 302, pp. 46–69, 2005.

[18] A. P. Piotrowski and J. J. Napiorkowski, Optimizing neural networks for river flow forecasting - Evolutionary Computation methods versus the Levenberg-Marquardt approach, *Journal of Hydrology* 407, pp. 12–27, 2011,

[19] R. M. Porto, *Hidráulica Básica*, EESC-USP, São Paulo, 2000.

[20] A. J. C. Saint-Venant, Théorie du mouvement non-permanent des eaux, avec application aux crues des rivière at à l'introduction des marées dans leur lit, *Comptes Rendus des Séances de Académie des Sciences* 73, pp. 147–154, 1871.

[21] A. L. Simões, H. E. Schulz, and R. M. Porto, *Métodos Computacionais em Hidráulica*, Edufba, Bahia, 2017.

Appendix

	t in days	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
	$Q_{\min}(t)$	28.20	25.50	27.40	12.70	7.62	4.98	4.18	6.62	10.50	13.40	15.10	10.60	5.98	3.76	3.69	6.54	8.76	11.80	12.80
x in meters	0	0.0041	0.0019	0.0030	0.0666	0.1419	0.2061	0.2316	0.1641	0.0948	0.0600	0.0438	0.0923	0.1785	0.2464	0.2494	0.1655	0.1213	0.0769	0.0648
	120	0.0051	0.0021	0.0037	0.0665	0.1397	0.2020	0.2268	0.1613	0.0941	0.0601	0.0442	0.0916	0.1752	0.2412	0.2442	0.1627	0.1197	0.0765	0.0647
	240	0.0065	0.0025	0.0046	0.0635	0.1271	0.1793	0.1999	0.1456	0.0882	0.0579	0.0432	0.0856	0.1568	0.2117	0.2142	0.1466	0.1102	0.0723	0.0615
	360	0.0082	0.0030	0.0057	0.0615	0.1132	0.1509	0.1650	0.1272	0.0830	0.0568	0.0432	0.0803	0.1349	0.1730	0.1749	0.1278	0.1004	0.0691	0.0595
	480	0.0090	0.0033	0.0063	0.0738	0.1444	0.2019	0.2246	0.1654	0.1021	0.0679	0.0509	0.0985	0.1770	0.2375	0.2407	0.1662	0.1261	0.0837	0.0712
	600	0.0096	0.0034	0.0067	0.0696	0.1331	0.1851	0.2060	0.1521	0.0953	0.0643	0.0486	0.0919	0.1624	0.2179	0.2209	0.1527	0.1168	0.0786	0.0671
	720	0.0110	0.0038	0.0077	0.0669	0.1182	0.1562	0.1712	0.1327	0.0891	0.0626	0.0482	0.0858	0.1397	0.1796	0.1819	0.1329	0.1059	0.0746	0.0644
	840	0.0123	0.0044	0.0086	0.0884	0.1660	0.2219	0.2419	0.1881	0.1213	0.0819	0.0618	0.1165	0.1984	0.2526	0.2556	0.1884	0.1472	0.0997	0.0850
	960	0.0133	0.0048	0.0093	0.0970	0.1895	0.2640	0.2920	0.2181	0.1350	0.0897	0.0672	0.1293	0.2317	0.3070	0.3113	0.2185	0.1661	0.1099	0.0930
	1080	0.0137	0.0048	0.0095	0.0965	0.1869	0.2589	0.2862	0.2147	0.1339	0.0894	0.0670	0.1281	0.2276	0.3010	0.3053	0.2150	0.1642	0.1091	0.0925
	1200	0.0149	0.0053	0.0104	0.1033	0.2010	0.2802	0.3105	0.2317	0.1438	0.0958	0.0717	0.1373	0.2456	0.3267	0.3317	0.2319	0.1764	0.1168	0.0988
	1320	0.0150	0.0052	0.0104	0.0983	0.1891	0.2625	0.2908	0.2176	0.1362	0.0914	0.0687	0.1300	0.2303	0.3060	0.3108	0.2178	0.1664	0.1109	0.0941
	1440	0.0167	0.0057	0.0116	0.1018	0.1923	0.2636	0.2903	0.2208	0.1404	0.0951	0.0718	0.1337	0.2325	0.3042	0.3089	0.2206	0.1702	0.1146	0.0973
	1560	0.0174	0.0060	0.0121	0.1044	0.1981	0.2749	0.3048	0.2286	0.1442	0.0975	0.0736	0.1371	0.2408	0.3205	0.3261	0.2283	0.1750	0.1174	0.0996
	1680	0.0181	0.0062	0.0125	0.1059	0.2004	0.2789	0.3101	0.2315	0.1461	0.0991	0.0749	0.1388	0.2438	0.3267	0.3327	0.2312	0.1771	0.1190	0.1009
	1800	0.0183	0.0062	0.0127	0.1049	0.1972	0.2734	0.3036	0.2275	0.1444	0.0983	0.0744	0.1372	0.2393	0.3196	0.3255	0.2271	0.1746	0.1178	0.1000
	1920	0.0186	0.0062	0.0128	0.0946	0.1722	0.2358	0.2610	0.1979	0.1286	0.0892	0.0683	0.1220	0.2072	0.2744	0.2795	0.1973	0.1536	0.1056	0.0901
	2040	0.0187	0.0061	0.0128	0.0831	0.1445	0.1942	0.2143	0.1649	0.1109	0.0790	0.0613	0.1051	0.1717	0.2248	0.2291	0.1643	0.1302	0.0919	0.0791
	2160	0.0210	0.0070	0.0145	0.1012	0.1707	0.2183	0.2361	0.1916	0.1343	0.0961	0.0743	0.1274	0.1973	0.2452	0.2492	0.1907	0.1559	0.1118	0.0963
	2280	0.0222	0.0074	0.0153	0.1104	0.1805	0.2162	0.2267	0.1984	0.1458	0.1048	0.0810	0.1384	0.2017	0.2313	0.2340	0.1973	0.1672	0.1219	0.1050
	2400	0.0209	0.0070	0.0144	0.1152	0.2119	0.2840	0.3099	0.2426	0.1583	0.1084	0.0821	0.1497	0.2525	0.3227	0.3279	0.2417	0.1894	0.1290	0.1095
	2520	0.0218	0.0074	0.0151	0.1256	0.2409	0.3369	0.3741	0.2804	0.1755	0.1179	0.0885	0.1655	0.2938	0.3930	0.4008	0.2793	0.2129	0.1414	0.1192
	2640	0.0218	0.0073	0.0151	0.1231	0.2358	0.3300	0.3665	0.2747	0.1719	0.1156	0.0869	0.1620	0.2876	0.3849	0.3926	0.2735	0.2085	0.1385	0.1167
	2760	0.0196	0.0066	0.0136	0.1132	0.2189	0.3077	0.3419	0.2555	0.1587	0.1062	0.0796	0.1496	0.2678	0.3591	0.3663	0.2544	0.1931	0.1276	0.1073
	2880	0.0200	0.0067	0.0138	0.1112	0.2165	0.3089	0.3455	0.2542	0.1563	0.1044	0.0783	0.1470	0.2668	0.3640	0.3718	0.2530	0.1906	0.1254	0.1053
	3000	0.0139	0.0049	0.0097	0.0944	0.1926	0.2810	0.3160	0.2281	0.1353	0.0879	0.0648	0.1273	0.2408	0.3337	0.3409	0.2272	0.1677	0.1073	0.0894
		0.0163	0.0056	0.0113	0.0976	0.1848	0.2552	0.2824	0.2128	0.1348	0.0912	0.0689	0.1282	0.2241	0.2968	0.3018	0.2126	0.1634	0.1098	0.0932

Table 9: Section 5.2. For each value of x in the first column, predictions for $t > 10$ days use data observed at $t_1 = 9.5$ and $t_2 = 10$ days. Observed data correspond to $z(x, t_1)$ and $z(x, t_2)$ and the prediction is given by the *linear polynomial in* $Q_{\min}(t)$ *that interpolates the data*. Each cell of the table shows $|z_{\text{pred}}(x, t) - z(x, t)|$, where values of $z(x, t)$ correspond to synthetic data that is not used in the prediction. The last line in the table shows the RMSD for each t.

	t in days	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
	$Q_{\min}(t)$	28.20	25.50	27.40	12.70	7.62	4.98	4.18	6.62	10.50	13.40	15.10	10.60	5.98	3.76	3.69	6.54	8.76	11.80	12.80
x in meters	0	0.0101	0.0019	0.0062	0.0383	0.0656	0.0685	0.0651	0.0677	0.0503	0.0335	0.0246	0.0508	0.0696	0.0622	0.0613	0.0683	0.0602	0.0437	0.0384
	120	0.0116	0.0021	0.0071	0.0461	0.0832	0.0930	0.0920	0.0877	0.0618	0.0403	0.0293	0.0622	0.0913	0.0904	0.0896	0.0885	0.0753	0.0530	0.0462
	240	0.0134	0.0025	0.0083	0.0578	0.1130	0.1383	0.1433	0.1226	0.0796	0.0502	0.0360	0.0799	0.1301	0.1453	0.1451	0.1239	0.0997	0.0671	0.0579
	360	0.0159	0.0030	0.0098	0.0732	0.1534	0.2017	0.2160	0.1705	0.1033	0.0632	0.0447	0.1035	0.1836	0.2234	0.2241	0.1725	0.1327	0.0857	0.0731
	480	0.0176	0.0033	0.0109	0.0781	0.1562	0.1958	0.2052	0.1704	0.1081	0.0675	0.0483	0.1088	0.1824	0.2097	0.2095	0.1726	0.1368	0.0909	0.0783
	600	0.0183	0.0034	0.0114	0.0833	0.1695	0.2152	0.2266	0.1859	0.1162	0.0719	0.0512	0.1167	0.1993	0.2322	0.2321	0.1882	0.1478	0.0972	0.0834
	720	0.0204	0.0038	0.0126	0.0968	0.2057	0.2723	0.2920	0.2291	0.1373	0.0833	0.0587	0.1376	0.2475	0.3022	0.3031	0.2320	0.1773	0.1135	0.0967
	840	0.0233	0.0044	0.0145	0.1055	0.2177	0.2857	0.3067	0.2406	0.1470	0.0909	0.0648	0.1482	0.2603	0.3182	0.3190	0.2439	0.1884	0.1232	0.1059
	960	0.0253	0.0048	0.0157	0.1124	0.2248	0.2843	0.3005	0.2448	0.1547	0.0969	0.0695	0.1565	0.2636	0.3094	0.3092	0.2483	0.1963	0.1308	0.1131
	1080	0.0257	0.0048	0.0160	0.1151	0.2317	0.2949	0.3123	0.2529	0.1587	0.0992	0.0711	0.1605	0.2727	0.3217	0.3214	0.2566	0.2019	0.1341	0.1157
	1200	0.0279	0.0053	0.0173	0.1248	0.2501	0.3167	0.3345	0.2722	0.1716	0.1074	0.0771	0.1738	0.2937	0.3444	0.3438	0.2764	0.2181	0.1452	0.1255
	1320	0.0278	0.0052	0.0172	0.1254	0.2535	0.3230	0.3419	0.2767	0.1732	0.1079	0.0773	0.1752	0.2987	0.3522	0.3518	0.2808	0.2206	0.1461	0.1261
	1440	0.0305	0.0057	0.0189	0.1393	0.2847	0.3675	0.3918	0.3121	0.1931	0.1198	0.0855	0.1953	0.3378	0.4054	0.4054	0.3168	0.2470	0.1625	0.1400
	1560	0.0317	0.0060	0.0197	0.1447	0.2946	0.3770	0.3997	0.3218	0.2003	0.1244	0.0889	0.2027	0.3481	0.4124	0.4116	0.3268	0.2559	0.1688	0.1455
	1680	0.0327	0.0062	0.0203	0.1499	0.3056	0.3905	0.4134	0.3337	0.2076	0.1289	0.0921	0.2102	0.3611	0.4260	0.4249	0.3390	0.2654	0.1750	0.1508
	1800	0.0330	0.0062	0.0205	0.1520	0.3111	0.3990	0.4231	0.3402	0.2109	0.1306	0.0932	0.2133	0.3682	0.4364	0.4355	0.3455	0.2699	0.1775	0.1528
	1920	0.0328	0.0062	0.0204	0.1538	0.3193	0.4145	0.4417	0.3511	0.2150	0.1322	0.0939	0.2169	0.3803	0.4567	0.4564	0.3564	0.2762	0.1799	0.1543
	2040	0.0322	0.0061	0.0200	0.1545	0.3257	0.4278	0.4580	0.3603	0.2178	0.1328	0.0938	0.2192	0.3903	0.4746	0.4749	0.3655	0.2809	0.1812	0.1548
	2160	0.0369	0.0070	0.0229	0.1767	0.3791	0.5091	0.5500	0.4226	0.2500	0.1516	0.1071	0.2518	0.4599	0.5727	0.5741	0.4288	0.3249	0.2075	0.1772
	2280	0.0393	0.0074	0.0244	0.1879	0.4096	0.5645	0.6169	0.4607	0.2667	0.1609	0.1137	0.2685	0.5037	0.6464	0.6495	0.4675	0.3488	0.2208	0.1884
	2400	0.0373	0.0070	0.0232	0.1723	0.3569	0.4685	0.5034	0.3927	0.2393	0.1478	0.1055	0.2425	0.4274	0.5234	0.5237	0.3991	0.3080	0.2014	0.1734
	2520	0.0391	0.0074	0.0243	0.1776	0.3591	0.4569	0.4838	0.3899	0.2440	0.1524	0.1094	0.2483	0.4235	0.4995	0.4976	0.3968	0.3118	0.2071	0.1792
	2640	0.0389	0.0073	0.0242	0.1771	0.3582	0.4558	0.4828	0.3888	0.2434	0.1520	0.1091	0.2476	0.4225	0.4987	0.4968	0.3958	0.3110	0.2066	0.1787
	2760	0.0351	0.0066	0.0219	0.1591	0.3198	0.4050	0.4283	0.3462	0.2179	0.1364	0.0981	0.2219	0.3762	0.4422	0.4403	0.3525	0.2780	0.1853	0.1606
	2880	0.0355	0.0067	0.0221	0.1607	0.3215	0.4030	0.4238	0.3468	0.2199	0.1379	0.0992	0.2240	0.3764	0.4364	0.4338	0.3532	0.2800	0.1872	0.1623
	3000	0.0257	0.0049	0.0160	0.1123	0.2162	0.2599	0.2685	0.2286	0.1505	0.0963	0.0700	0.1546	0.2479	0.2744	0.2712	0.2334	0.1898	0.1302	0.1139
		0.0295	0.0056	0.0184	0.1355	0.2783	0.3599	0.3835	0.3051	0.1881	0.1165	0.0832	0.1903	0.3307	0.3967	0.3964	0.3099	0.2410	0.1583	0.1363

Table 10: Section 5.2. For each value of x in the first column, predictions for $t > 10$ days use data observed at $t_1 = 9$, $t_2 = 9.5$, and $t_3 = 10$ days. Observed data correspond to $z(x, t_1)$, $z(x, t_2)$, and $z(x, t_3)$ and the prediction is given by the *quadratic polynomial in* $Q_{\min}(t)$ *that interpolates the data*. Each cell of the table shows $|z_{\text{pred}}(x, t) - z(x, t)|$, where values of $z(x, t)$ correspond to synthetic data that is not used in the prediction. The last line in the table shows the RMSD for each t.

	t in days	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
	$Q_{\min}(t)$	28.20	25.50	27.40	12.70	7.62	4.98	4.18	6.62	10.50	13.40	15.10	10.60	5.98	3.76	3.69	6.54	8.76	11.80	12.80
x in meters	0	0.1884	0.0019	0.0884	5.0472	15.3205	23.9477	27.0883	18.2898	8.5959	4.1676	2.4647	8.4059	20.3765	28.8421	29.1420	18.5422	12.3497	6.3488	4.9138
	120	0.2171	0.0021	0.1018	5.8152	17.6504	27.5875	31.2045	21.0707	9.9036	4.8018	2.8398	9.6847	23.4744	33.2241	33.5694	21.3614	14.2281	7.3148	5.6616
	240	0.2542	0.0025	0.1192	6.8082	20.6602	32.2868	36.5174	24.6625	11.5939	5.6218	3.3249	11.3377	27.4749	38.8794	39.2832	25.0027	16.6552	8.5636	6.6283
	360	0.3027	0.0030	0.1420	8.1078	24.5982	38.4332	43.4658	29.3613	13.8058	6.6951	3.9599	13.5008	32.7081	46.2752	46.7554	29.7662	19.8311	10.1980	7.8936
	480	0.3372	0.0033	0.1582	9.0355	27.4170	42.8435	48.4563	32.7280	15.3867	7.4612	4.4128	15.0463	36.4590	51.5897	52.1257	33.1791	22.1029	11.3650	8.7966
	600	0.3512	0.0034	0.1648	9.4102	28.5532	44.6183	50.4634	34.0839	16.0245	7.7707	4.5959	15.6701	37.9695	53.7266	54.2847	34.5538	23.0189	11.8363	9.1615
	720	0.3917	0.0038	0.1838	10.4938	31.8360	49.7420	56.2557	38.0008	17.8686	8.6656	5.1254	17.4736	42.3320	59.8918	60.5135	38.5247	25.6665	13.1990	10.2165
	840	0.4503	0.0044	0.2112	12.0706	36.6212	57.2164	64.7070	43.7130	20.5547	9.9678	5.8954	20.0996	48.6940	68.8878	69.6028	44.3152	29.5247	15.1824	11.7513
	960	0.4888	0.0048	0.2293	13.1047	39.7637	62.1341	70.2716	47.4666	22.3169	10.8218	6.4002	21.8224	52.8761	74.8134	75.5906	48.1204	32.0571	16.4833	12.7579
	1080	0.4978	0.0048	0.2335	13.3473	40.4988	63.2816	71.5693	48.3438	22.7298	11.0221	6.5188	22.2262	53.8529	76.1951	76.9867	49.0096	32.6500	16.7884	12.9940
	1200	0.5408	0.0053	0.2537	14.5013	44.0008	68.7550	77.7600	52.5248	24.6952	11.9751	7.0823	24.1479	58.5103	82.7860	83.6463	53.2481	35.4732	18.2399	14.1174
	1320	0.5380	0.0052	0.2524	14.4250	43.7685	68.3912	77.3484	52.2471	24.5650	11.9121	7.0451	24.0206	58.2010	82.3479	83.2037	52.9667	35.2860	18.1439	14.0431
	1440	0.5912	0.0057	0.2773	15.8525	48.0982	75.1538	84.9949	57.4149	26.9957	13.0910	7.7424	26.3974	63.9570	90.4873	91.4274	58.2056	38.7771	19.9393	15.4328
	1560	0.6148	0.0060	0.2884	16.4865	50.0230	78.1641	88.4012	59.7135	28.0757	13.6146	8.0520	27.4534	66.5176	94.1148	95.0931	60.5356	40.3286	20.7368	16.0500
	1680	0.6353	0.0062	0.2980	17.0343	51.6852	80.7623	91.3405	61.6980	29.0087	14.0670	8.3196	28.3656	68.7282	97.2445	98.2557	62.5474	41.6687	21.4259	16.5833
	1800	0.6413	0.0062	0.3008	17.1946	52.1710	81.5206	92.1977	62.2776	29.2815	14.1994	8.3979	28.6324	69.3738	98.1570	99.1776	63.1350	42.0604	21.6275	16.7394
	1920	0.6363	0.0062	0.2985	17.0565	51.7504	80.8615	91.4514	61.7746	29.0456	14.0853	8.3305	28.4021	68.8137	97.3620	98.3739	62.6253	41.7215	21.4537	16.6051
	2040	0.6249	0.0061	0.2931	16.7475	50.8114	79.3924	89.7892	60.6527	28.5187	13.8300	8.1797	27.8873	67.5644	95.5921	96.5853	61.4882	40.9645	21.0649	16.3045
	2160	0.7188	0.0070	0.3372	19.2708	58.4593	91.3322	103.2885	69.7791	32.8142	15.9140	9.4124	32.0876	77.7289	109.9618	111.1036	70.7401	47.1322	24.2383	18.7609
	2280	0.7662	0.0074	0.3594	20.5416	62.3073	97.3308	110.0664	74.3685	34.9771	16.9637	10.0333	34.2026	82.8392	117.1742	118.3897	75.3927	50.2365	25.8364	19.9980
	2400	0.7280	0.0070	0.3415	19.5250	59.2371	92.5522	104.6687	70.7105	33.2496	16.1241	9.5362	32.5122	78.7655	111.4305	112.5881	71.6838	47.7585	24.5584	19.0078
	2520	0.7642	0.0074	0.3585	20.5020	62.2071	97.2032	109.9337	74.2596	34.9151	16.9311	10.0132	34.1400	82.7191	117.0381	118.2556	75.2812	50.1520	25.7875	19.9587
	2640	0.7609	0.0073	0.3569	20.4136	61.9390	96.7844	109.4599	73.9397	34.7646	16.8581	9.9700	33.9928	82.3627	116.5336	117.7459	74.9569	49.9359	25.6763	19.8726
	2760	0.6876	0.0066	0.3225	18.4484	55.9770	87.4691	98.9249	66.8231	31.4182	15.2352	9.0102	30.7205	74.4352	105.3176	106.4133	67.7422	45.1292	23.2045	17.9594
	2880	0.6942	0.0067	0.3256	18.6246	56.5133	88.3109	99.8788	67.4644	31.7185	15.3806	9.0962	31.0141	75.1499	106.3341	107.4409	68.3923	45.5611	23.4262	18.1309
	3000	0.5055	0.0049	0.2371	13.5685	41.1748	64.3475	72.7786	49.1561	23.1091	11.2054	6.6266	22.5951	54.7554	77.4831	78.2905	49.8317	33.1948	17.0667	13.2084
		0.5743	0.0056	0.2694	15.3999	46.7243	73.0076	82.5683	55.7752	26.2250	12.7173	7.5214	25.6436	62.1302	87.9043	88.8180	56.5431	37.6696	19.3700	14.9921

Table 11: Section 5.2. For each value of x in the first column, predictions for $t > 10$ days use data observed at $t_1 = 8.5$, $t_2 = 9$, $t_3 = 9.5$, and $t_4 = 10$ days. Observed data correspond to $z(x, t_1)$, $z(x, t_2)$, $z(x, t_3)$, and $z(x, t_4)$ and the prediction is given by the *cubic polynomial in $Q_{\min}(t)$ that interpolates the data*. Each cell of the table shows $|z_{\text{pred}}(x, t) - z(x, t)|$, where values of $z(x, t)$ correspond to synthetic data that is not used in the prediction. The last line in the table shows the RMSD for each t.

x in meters	RMSD of observations – t in days														
	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
0	0.0253	0.0082	0.0019	0.0048	0.0092	0.0117	0.0127	0.0063	0.0125	0.0108	0.0084	0.0077	0.0036	0.0115	0.0076
120	0.0242	0.0078	0.0018	0.0046	0.0088	0.0110	0.0121	0.0058	0.0120	0.0105	0.0082	0.0074	0.0036	0.0109	0.0071
240	0.0199	0.0066	0.0012	0.0037	0.0071	0.0090	0.0102	0.0046	0.0104	0.0093	0.0075	0.0064	0.0035	0.0093	0.0057
360	0.0144	0.0054	0.0004	0.0024	0.0049	0.0065	0.0080	0.0033	0.0086	0.0079	0.0068	0.0052	0.0034	0.0075	0.0041
480	0.0214	0.0072	0.0012	0.0040	0.0075	0.0093	0.0112	0.0046	0.0116	0.0104	0.0086	0.0070	0.0041	0.0101	0.0059
600	0.0184	0.0063	0.0010	0.0034	0.0064	0.0079	0.0097	0.0038	0.0102	0.0093	0.0078	0.0062	0.0038	0.0088	0.0049
720	0.0129	0.0050	0.0001	0.0021	0.0042	0.0054	0.0074	0.0025	0.0083	0.0079	0.0071	0.0050	0.0038	0.0069	0.0033
840	0.0225	0.0082	0.0008	0.0041	0.0078	0.0098	0.0124	0.0047	0.0131	0.0118	0.0101	0.0079	0.0050	0.0112	0.0062
960	0.0280	0.0094	0.0017	0.0055	0.0098	0.0117	0.0147	0.0054	0.0152	0.0136	0.0113	0.0092	0.0053	0.0130	0.0075
1080	0.0270	0.0092	0.0015	0.0053	0.0094	0.0113	0.0143	0.0052	0.0149	0.0134	0.0112	0.0090	0.0053	0.0126	0.0072
1200	0.0293	0.0098	0.0018	0.0059	0.0102	0.0121	0.0154	0.0054	0.0161	0.0144	0.0120	0.0097	0.0057	0.0136	0.0077
1320	0.0267	0.0090	0.0015	0.0053	0.0092	0.0108	0.0141	0.0048	0.0148	0.0134	0.0113	0.0089	0.0054	0.0124	0.0069
1440	0.0259	0.0089	0.0013	0.0052	0.0088	0.0102	0.0139	0.0043	0.0148	0.0135	0.0115	0.0089	0.0056	0.0122	0.0064
1560	0.0269	0.0091	0.0015	0.0055	0.0092	0.0105	0.0143	0.0043	0.0152	0.0139	0.0118	0.0091	0.0057	0.0125	0.0066
1680	0.0269	0.0090	0.0015	0.0056	0.0091	0.0103	0.0144	0.0042	0.0153	0.0140	0.0120	0.0092	0.0058	0.0125	0.0065
1800	0.0260	0.0087	0.0014	0.0054	0.0087	0.0099	0.0139	0.0039	0.0149	0.0137	0.0118	0.0089	0.0058	0.0121	0.0062
1920	0.0203	0.0069	0.0009	0.0042	0.0066	0.0072	0.0111	0.0025	0.0122	0.0115	0.0102	0.0073	0.0052	0.0097	0.0044
2040	0.0141	0.0050	0.0004	0.0029	0.0043	0.0044	0.0080	0.0010	0.0093	0.0091	0.0084	0.0055	0.0046	0.0070	0.0025
2160	0.0155	0.0065	0.0004	0.0027	0.0047	0.0055	0.0097	0.0018	0.0115	0.0111	0.0105	0.0067	0.0058	0.0088	0.0033
2280	0.0149	0.0074	0.0012	0.0022	0.0044	0.0059	0.0103	0.0022	0.0125	0.0119	0.0116	0.0071	0.0064	0.0095	0.0035
2400	0.0266	0.0095	0.0010	0.0054	0.0089	0.0101	0.0148	0.0039	0.0161	0.0147	0.0130	0.0095	0.0064	0.0128	0.0064
2520	0.0337	0.0112	0.0021	0.0073	0.0115	0.0127	0.0179	0.0049	0.0189	0.0171	0.0146	0.0113	0.0069	0.0153	0.0081
2640	0.0329	0.0108	0.0021	0.0071	0.0112	0.0122	0.0175	0.0046	0.0184	0.0167	0.0142	0.0110	0.0068	0.0149	0.0078
2760	0.0313	0.0102	0.0021	0.0068	0.0107	0.0117	0.0165	0.0045	0.0173	0.0156	0.0132	0.0104	0.0062	0.0140	0.0075
2880	0.0313	0.0099	0.0023	0.0070	0.0107	0.0114	0.0163	0.0042	0.0170	0.0154	0.0129	0.0102	0.0061	0.0137	0.0073
3000	0.0313	0.0098	0.0025	0.0069	0.0109	0.0121	0.0160	0.0050	0.0162	0.0142	0.0116	0.0098	0.0052	0.0135	0.0079
	0.0254	0.0086	0.0015	0.0052	0.0087	0.0101	0.0135	0.0044	0.0143	0.0130	0.0111	0.0086	0.0054	0.0119	0.0064

x in meters	RMSD of predictions – t in days																		
	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
0	0.0248	0.0096	0.0207	0.0046	0.0379	0.0850	0.1053	0.0536	0.0093	0.0068	0.0119	0.0076	0.0639	0.1175	0.1200	0.0545	0.0246	0.0002	0.0058
120	0.0231	0.0088	0.0194	0.0046	0.0360	0.0814	0.1010	0.0512	0.0088	0.0065	0.0115	0.0070	0.0609	0.1127	0.1153	0.0520	0.0233	0.0004	0.0059
240	0.0190	0.0068	0.0161	0.0041	0.0290	0.0653	0.0811	0.0414	0.0074	0.0055	0.0100	0.0054	0.0489	0.0904	0.0925	0.0420	0.0189	0.0007	0.0055
360	0.0145	0.0045	0.0125	0.0034	0.0196	0.0424	0.0520	0.0280	0.0057	0.0041	0.0081	0.0036	0.0321	0.0577	0.0592	0.0281	0.0133	0.0008	0.0048
480	0.0199	0.0068	0.0170	0.0051	0.0301	0.0693	0.0864	0.0441	0.0078	0.0062	0.0114	0.0050	0.0513	0.0963	0.0990	0.0444	0.0197	0.0015	0.0070
600	0.0167	0.0054	0.0145	0.0047	0.0257	0.0606	0.0763	0.0382	0.0067	0.0055	0.0101	0.0039	0.0444	0.0854	0.0880	0.0383	0.0168	0.0016	0.0066
720	0.0123	0.0031	0.0108	0.0038	0.0165	0.0386	0.0486	0.0250	0.0050	0.0040	0.0080	0.0023	0.0281	0.0545	0.0564	0.0247	0.0112	0.0016	0.0057
840	0.0212	0.0068	0.0183	0.0062	0.0295	0.0636	0.0769	0.0432	0.0085	0.0069	0.0130	0.0045	0.0483	0.0841	0.0865	0.0429	0.0200	0.0023	0.0088
960	0.0245	0.0082	0.0212	0.0079	0.0378	0.0879	0.1086	0.0572	0.0098	0.0088	0.0157	0.0050	0.0649	0.1197	0.1233	0.0569	0.0249	0.0034	0.0110
1080	0.0236	0.0078	0.0205	0.0078	0.0362	0.0840	0.1041	0.0548	0.0095	0.0085	0.0154	0.0047	0.0619	0.1150	0.1187	0.0544	0.0239	0.0034	0.0109
1200	0.0252	0.0083	0.0218	0.0088	0.0391	0.0924	0.1148	0.0600	0.0101	0.0094	0.0167	0.0047	0.0676	0.1269	0.1312	0.0594	0.0257	0.0041	0.0122
1320	0.0225	0.0072	0.0197	0.0083	0.0353	0.0841	0.1050	0.0545	0.0092	0.0087	0.0155	0.0039	0.0613	0.1163	0.1205	0.0539	0.0232	0.0040	0.0116
1440	0.0213	0.0064	0.0188	0.0086	0.0332	0.0792	0.0982	0.0521	0.0088	0.0087	0.0156	0.0031	0.0576	0.1081	0.1121	0.0512	0.0219	0.0045	0.0122
1560	0.0217	0.0065	0.0191	0.0092	0.0345	0.0853	0.1073	0.0551	0.0090	0.0092	0.0163	0.0028	0.0611	0.1189	0.1238	0.0541	0.0226	0.0050	0.0130
1680	0.0214	0.0063	0.0189	0.0094	0.0344	0.0865	0.1097	0.0555	0.0089	0.0093	0.0165	0.0025	0.0614	0.1221	0.1274	0.0544	0.0225	0.0053	0.0134
1800	0.0204	0.0059	0.0182	0.0093	0.0328	0.0830	0.1053	0.0533	0.0085	0.0090	0.0161	0.0022	0.0588	0.1172	0.1224	0.0521	0.0215	0.0053	0.0132
1920	0.0148	0.0035	0.0136	0.0081	0.0248	0.0652	0.0834	0.0417	0.0065	0.0074	0.0134	0.0008	0.0454	0.0931	0.0976	0.0405	0.0163	0.0051	0.0118
2040	0.0088	0.0009	0.0086	0.0068	0.0162	0.0459	0.0599	0.0290	0.0043	0.0056	0.0104	0.0007	0.0309	0.0673	0.0710	0.0278	0.0105	0.0047	0.0100
2160	0.0125	0.0018	0.0117	0.0074	0.0157	0.0390	0.0495	0.0273	0.0057	0.0061	0.0123	0.0004	0.0272	0.0547	0.0581	0.0257	0.0113	0.0049	0.0114
2280	0.0140	0.0020	0.0130	0.0072	0.0125	0.0219	0.0245	0.0204	0.0064	0.0058	0.0128	0.0000	0.0174	0.0250	0.0269	0.0185	0.0105	0.0046	0.0116
2400	0.0212	0.0058	0.0190	0.0104	0.0313	0.0749	0.0921	0.0513	0.0090	0.0096	0.0175	0.0014	0.0543	0.1004	0.1049	0.0495	0.0212	0.0063	0.0150
2520	0.0260	0.0079	0.0230	0.0127	0.0416	0.1059	0.1335	0.0691	0.0107	0.0121	0.0210	0.0020	0.0748	0.1474	0.1543	0.0670	0.0273	0.0077	0.0179
2640	0.0248	0.0075	0.0221	0.0126	0.0404	0.1036	0.1306	0.0675	0.0103	0.0119	0.0206	0.0016	0.0729	0.1440	0.1509	0.0654	0.0264	0.0078	0.0178
2760	0.0237	0.0074	0.0211	0.0118	0.0387	0.0988	0.1244	0.0645	0.0098	0.0112	0.0194	0.0017	0.0697	0.1369	0.1434	0.0625	0.0253	0.0072	0.0166
2880	0.0226	0.0070	0.0202	0.0121	0.0389	0.1030	0.1310	0.0659	0.0095	0.0114	0.0193	0.0013	0.0716	0.1450	0.1521	0.0638	0.0251	0.0076	0.0169
3000	0.0246	0.0086	0.0214	0.0108	0.0404	0.1042	0.1318	0.0665	0.0097	0.0108	0.0181	0.0026	0.0733	0.1456	0.1522	0.0649	0.0260	0.0062	0.0148
	0.0211	0.0067	0.0185	0.0085	0.0329	0.0798	0.1001	0.0518	0.0086	0.0085	0.0152	0.0037	0.0576	0.1108	0.1151	0.0509	0.0216	0.0048	0.0121

Table 12: Section 5.2. For each value of x in the first column, predictions for $t > 10$ days use data observed at $t \in \{3.5, 4, 4.5, \dots, 10\}$ days. Observed data correspond to $z(x,t)$ and the prediction is given by the *best fitting linear polynomial (solution of a linear least squares problem)*. Each cell of the table shows $|z_{\text{pred}}(x,t) - z(x,t)|$. In the left-hand part of the table, values of $z(x,t)$ correspond to synthetic trainind data (used in the fitting), while in the right-hand part of the table $z(x,t)$ correspond to synthetic (testing) data that is not being used in the fitting. The last line in the table shows the RMSD for each t.

	RMSD of observations – t in days														
x in meters	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
0	0.0027	0.0008	0.0017	0.0019	0.0011	0.0012	0.0003	0.0026	0.0003	0.0002	0.0005	0.0001	0.0006	0.0012	0.0012
120	0.0026	0.0008	0.0016	0.0018	0.0010	0.0013	0.0003	0.0027	0.0002	0.0003	0.0002	0.0002	0.0005	0.0011	0.0013
240	0.0017	0.0006	0.0011	0.0014	0.0005	0.0014	0.0003	0.0025	0.0002	0.0007	0.0004	0.0003	0.0001	0.0008	0.0014
360	0.0004	0.0002	0.0002	0.0006	0.0001	0.0015	0.0003	0.0021	0.0007	0.0013	0.0013	0.0005	0.0008	0.0004	0.0014
480	0.0017	0.0006	0.0010	0.0015	0.0004	0.0019	0.0004	0.0030	0.0005	0.0011	0.0009	0.0005	0.0004	0.0009	0.0017
600	0.0014	0.0005	0.0008	0.0013	0.0002	0.0019	0.0004	0.0029	0.0006	0.0013	0.0012	0.0005	0.0006	0.0007	0.0017
720	0.0000	0.0001	0.0000	0.0005	0.0005	0.0019	0.0004	0.0025	0.0011	0.0018	0.0021	0.0007	0.0014	0.0002	0.0017
840	0.0011	0.0003	0.0006	0.0014	0.0001	0.0024	0.0006	0.0036	0.0010	0.0017	0.0018	0.0007	0.0009	0.0008	0.0021
960	0.0024	0.0008	0.0014	0.0023	0.0005	0.0029	0.0007	0.0046	0.0008	0.0016	0.0014	0.0007	0.0005	0.0013	0.0025
1080	0.0022	0.0007	0.0013	0.0021	0.0004	0.0029	0.0007	0.0045	0.0009	0.0017	0.0015	0.0007	0.0007	0.0012	0.0025
1200	0.0026	0.0008	0.0015	0.0025	0.0005	0.0032	0.0008	0.0050	0.0009	0.0018	0.0016	0.0007	0.0007	0.0014	0.0027
1320	0.0023	0.0007	0.0013	0.0022	0.0004	0.0031	0.0007	0.0048	0.0010	0.0019	0.0018	0.0008	0.0008	0.0012	0.0026
1440	0.0020	0.0007	0.0011	0.0021	0.0002	0.0034	0.0008	0.0050	0.0013	0.0022	0.0022	0.0009	0.0011	0.0011	0.0028
1560	0.0023	0.0008	0.0013	0.0024	0.0003	0.0036	0.0008	0.0053	0.0013	0.0022	0.0022	0.0009	0.0011	0.0012	0.0030
1680	0.0023	0.0008	0.0013	0.0024	0.0002	0.0037	0.0008	0.0055	0.0013	0.0023	0.0023	0.0009	0.0012	0.0012	0.0031
1800	0.0021	0.0008	0.0012	0.0023	0.0001	0.0037	0.0008	0.0054	0.0014	0.0024	0.0025	0.0009	0.0013	0.0012	0.0031
1920	0.0014	0.0006	0.0008	0.0018	0.0002	0.0036	0.0007	0.0049	0.0015	0.0026	0.0028	0.0010	0.0016	0.0009	0.0029
2040	0.0006	0.0004	0.0003	0.0012	0.0006	0.0033	0.0006	0.0043	0.0016	0.0028	0.0031	0.0010	0.0020	0.0005	0.0028
2160	0.0006	0.0001	0.0005	0.0007	0.0011	0.0037	0.0009	0.0045	0.0024	0.0035	0.0042	0.0013	0.0027	0.0002	0.0030
2280	0.0019	0.0007	0.0014	0.0000	0.0016	0.0037	0.0011	0.0044	0.0030	0.0039	0.0050	0.0015	0.0033	0.0001	0.0030
2400	0.0016	0.0005	0.0008	0.0022	0.0002	0.0042	0.0011	0.0059	0.0019	0.0029	0.0032	0.0012	0.0018	0.0011	0.0034
2520	0.0032	0.0010	0.0018	0.0034	0.0004	0.0047	0.0012	0.0070	0.0016	0.0027	0.0026	0.0011	0.0012	0.0017	0.0037
2640	0.0032	0.0010	0.0018	0.0034	0.0004	0.0047	0.0012	0.0070	0.0016	0.0027	0.0026	0.0011	0.0012	0.0017	0.0037
2760	0.0032	0.0010	0.0018	0.0032	0.0005	0.0043	0.0011	0.0065	0.0014	0.0023	0.0022	0.0010	0.0010	0.0017	0.0034
2880	0.0035	0.0011	0.0020	0.0035	0.0006	0.0044	0.0011	0.0067	0.0013	0.0023	0.0021	0.0009	0.0009	0.0018	0.0035
3000	0.0039	0.0012	0.0023	0.0034	0.0010	0.0035	0.0009	0.0057	0.0007	0.0013	0.0009	0.0006	0.0000	0.0018	0.0027
	0.0023	0.0007	0.0013	0.0022	0.0006	0.0033	0.0008	0.0049	0.0014	0.0022	0.0023	0.0009	0.0013	0.0012	0.0027

	RMSD of predictions – t in days																		
x in meters	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
0	0.0090	0.0030	0.0061	0.0030	0.0057	0.0264	0.0376	0.0121	0.0009	0.0023	0.0021	0.0021	0.0160	0.0446	0.0463	0.0123	0.0019	0.0029	0.0037
120	0.0091	0.0033	0.0062	0.0031	0.0053	0.0255	0.0364	0.0116	0.0010	0.0023	0.0022	0.0023	0.0153	0.0433	0.0450	0.0117	0.0017	0.0031	0.0039
240	0.0081	0.0033	0.0054	0.0029	0.0032	0.0184	0.0268	0.0082	0.0009	0.0019	0.0022	0.0023	0.0105	0.0321	0.0335	0.0081	0.0008	0.0029	0.0038
360	0.0063	0.0033	0.0041	0.0024	0.0003	0.0062	0.0102	0.0024	0.0007	0.0013	0.0020	0.0023	0.0025	0.0127	0.0136	0.0020	0.0007	0.0025	0.0036
480	0.0094	0.0042	0.0063	0.0038	0.0022	0.0184	0.0276	0.0081	0.0011	0.0023	0.0029	0.0034	0.0097	0.0331	0.0351	0.0077	0.0001	0.0039	0.0053
600	0.0087	0.0041	0.0058	0.0035	0.0015	0.0164	0.0252	0.0070	0.0011	0.0021	0.0028	0.0033	0.0083	0.0306	0.0325	0.0065	0.0003	0.0037	0.0050
720	0.0068	0.0041	0.0044	0.0029	0.0017	0.0054	0.0102	0.0015	0.0009	0.0014	0.0025	0.0032	0.0010	0.0133	0.0147	0.0008	0.0016	0.0032	0.0046
840	0.0107	0.0052	0.0071	0.0047	0.0009	0.0083	0.0130	0.0041	0.0012	0.0026	0.0038	0.0046	0.0031	0.0154	0.0170	0.0030	0.0014	0.0049	0.0069
960	0.0136	0.0060	0.0091	0.0061	0.0016	0.0219	0.0322	0.0104	0.0018	0.0037	0.0047	0.0059	0.0109	0.0376	0.0403	0.0092	0.0007	0.0065	0.0087
1080	0.0134	0.0060	0.0090	0.0061	0.0009	0.0199	0.0299	0.0094	0.0018	0.0036	0.0047	0.0059	0.0095	0.0353	0.0380	0.0082	0.0009	0.0064	0.0087
1200	0.0148	0.0066	0.0099	0.0069	0.0011	0.0232	0.0348	0.0110	0.0020	0.0041	0.0052	0.0067	0.0111	0.0409	0.0442	0.0095	0.0011	0.0073	0.0098
1320	0.0140	0.0064	0.0093	0.0066	0.0005	0.0209	0.0319	0.0098	0.0019	0.0038	0.0050	0.0065	0.0096	0.0377	0.0409	0.0083	0.0013	0.0070	0.0094
1440	0.0143	0.0069	0.0095	0.0069	0.0008	0.0174	0.0267	0.0083	0.0020	0.0039	0.0054	0.0071	0.0071	0.0313	0.0344	0.0066	0.0019	0.0074	0.0101
1560	0.0151	0.0072	0.0101	0.0075	0.0005	0.0216	0.0336	0.0100	0.0022	0.0043	0.0057	0.0077	0.0090	0.0397	0.0436	0.0081	0.0020	0.0080	0.0108
1680	0.0154	0.0074	0.0103	0.0077	0.0006	0.0228	0.0360	0.0104	0.0023	0.0044	0.0059	0.0080	0.0093	0.0429	0.0473	0.0084	0.0022	0.0083	0.0112
1800	0.0151	0.0074	0.0101	0.0076	0.0010	0.0213	0.0340	0.0097	0.0023	0.0043	0.0058	0.0079	0.0084	0.0405	0.0448	0.0077	0.0023	0.0082	0.0111
1920	0.0134	0.0071	0.0089	0.0068	0.0020	0.0163	0.0268	0.0071	0.0021	0.0037	0.0053	0.0073	0.0054	0.0323	0.0360	0.0052	0.0027	0.0074	0.0101
2040	0.0113	0.0066	0.0074	0.0058	0.0030	0.0111	0.0195	0.0043	0.0018	0.0030	0.0046	0.0064	0.0024	0.0239	0.0271	0.0026	0.0030	0.0064	0.0088
2160	0.0116	0.0072	0.0074	0.0062	0.0072	0.0026	0.0013	0.0022	0.0017	0.0029	0.0053	0.0072	0.0068	0.0029	0.0057	0.0044	0.0048	0.0069	0.0099
2280	0.0111	0.0074	0.0070	0.0061	0.0114	0.0215	0.0258	0.0103	0.0012	0.0025	0.0056	0.0072	0.0182	0.0291	0.0278	0.0128	0.0064	0.0067	0.0101
2400	0.0161	0.0082	0.0107	0.0086	0.0042	0.0102	0.0173	0.0055	0.0024	0.0047	0.0068	0.0092	0.0014	0.0200	0.0236	0.0029	0.0038	0.0094	0.0128
2520	0.0196	0.0091	0.0132	0.0105	0.0018	0.0269	0.0421	0.0131	0.0031	0.0060	0.0079	0.0110	0.0102	0.0492	0.0550	0.0100	0.0032	0.0114	0.0152
2640	0.0195	0.0091	0.0131	0.0105	0.0018	0.0268	0.0418	0.0131	0.0032	0.0060	0.0078	0.0110	0.0101	0.0486	0.0543	0.0100	0.0033	0.0114	0.0151
2760	0.0182	0.0083	0.0123	0.0098	0.0013	0.0261	0.0402	0.0130	0.0030	0.0057	0.0073	0.0103	0.0103	0.0465	0.0519	0.0100	0.0028	0.0107	0.0141
2880	0.0188	0.0085	0.0127	0.0101	0.0005	0.0312	0.0481	0.0151	0.0031	0.0059	0.0074	0.0106	0.0130	0.0558	0.0619	0.0121	0.0026	0.0110	0.0144
3000	0.0164	0.0067	0.0112	0.0089	0.0014	0.0332	0.0497	0.0163	0.0028	0.0054	0.0063	0.0091	0.0153	0.0574	0.0629	0.0137	0.0014	0.0096	0.0124
	0.0139	0.0066	0.0093	0.0070	0.0035	0.0212	0.0320	0.0100	0.0021	0.0040	0.0053	0.0072	0.0102	0.0377	0.0411	0.0087	0.0026	0.0075	0.0101

Table 13: Section 5.2. For each value of x in the first column, predictions for $t > 10$ days use data observed at $t \in \{3.5, 4, 4.5, \ldots, 10\}$ days. Observed data correspond to $z(x,t)$ and the prediction is given by the *best fitting quadratic polynomial (solution of a linear least squares problem)*. Each cell of the table shows $|z_{\text{pred}}(x,t) - z(x,t)|$. In the left-hand part of the table, values of $z(x,t)$ correspond to synthetic trainind data (used in the fitting), while in the right-hand part of the table $z(x,t)$ correspond to synthetic (testing) data that is not being used in the fitting. The last line in the table shows the RMSD for each t.

x in meters	RMSD of observations – t in days														
	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
0	0.0004	0.0004	0.0003	0.0004	0.0003	0.0011	0.0001	0.0011	0.0003	0.0009	0.0010	0.0002	0.0008	0.0000	0.0009
120	0.0004	0.0004	0.0003	0.0004	0.0003	0.0012	0.0001	0.0013	0.0003	0.0010	0.0012	0.0003	0.0009	0.0000	0.0010
240	0.0004	0.0003	0.0002	0.0005	0.0003	0.0013	0.0001	0.0016	0.0005	0.0012	0.0013	0.0004	0.0010	0.0001	0.0011
360	0.0002	0.0002	0.0001	0.0005	0.0002	0.0015	0.0003	0.0020	0.0007	0.0014	0.0015	0.0005	0.0010	0.0002	0.0013
480	0.0004	0.0004	0.0002	0.0006	0.0004	0.0018	0.0002	0.0022	0.0008	0.0015	0.0017	0.0005	0.0012	0.0002	0.0015
600	0.0004	0.0004	0.0002	0.0006	0.0004	0.0019	0.0002	0.0023	0.0008	0.0016	0.0018	0.0005	0.0012	0.0002	0.0016
720	0.0003	0.0002	0.0001	0.0006	0.0003	0.0019	0.0004	0.0026	0.0010	0.0017	0.0019	0.0007	0.0012	0.0003	0.0017
840	0.0003	0.0002	0.0001	0.0008	0.0004	0.0024	0.0005	0.0031	0.0012	0.0020	0.0023	0.0007	0.0014	0.0004	0.0020
960	0.0005	0.0004	0.0003	0.0010	0.0006	0.0028	0.0004	0.0033	0.0013	0.0022	0.0026	0.0008	0.0018	0.0003	0.0022
1080	0.0005	0.0004	0.0003	0.0010	0.0005	0.0028	0.0004	0.0034	0.0013	0.0023	0.0026	0.0008	0.0018	0.0003	0.0022
1200	0.0006	0.0005	0.0003	0.0012	0.0006	0.0031	0.0005	0.0037	0.0014	0.0025	0.0029	0.0008	0.0020	0.0003	0.0024
1320	0.0006	0.0004	0.0003	0.0011	0.0006	0.0031	0.0005	0.0037	0.0014	0.0024	0.0029	0.0008	0.0019	0.0003	0.0024
1440	0.0006	0.0004	0.0003	0.0012	0.0006	0.0033	0.0006	0.0041	0.0016	0.0027	0.0031	0.0009	0.0021	0.0004	0.0026
1560	0.0006	0.0005	0.0003	0.0013	0.0007	0.0035	0.0006	0.0042	0.0017	0.0028	0.0033	0.0010	0.0022	0.0004	0.0027
1680	0.0007	0.0005	0.0003	0.0014	0.0007	0.0037	0.0006	0.0044	0.0017	0.0029	0.0034	0.0010	0.0023	0.0004	0.0028
1800	0.0007	0.0005	0.0003	0.0014	0.0007	0.0037	0.0006	0.0044	0.0017	0.0029	0.0034	0.0010	0.0023	0.0004	0.0028
1920	0.0006	0.0004	0.0003	0.0013	0.0007	0.0035	0.0006	0.0044	0.0017	0.0029	0.0033	0.0010	0.0022	0.0004	0.0028
2040	0.0006	0.0004	0.0002	0.0012	0.0006	0.0033	0.0006	0.0042	0.0016	0.0028	0.0032	0.0010	0.0021	0.0005	0.0028
2160	0.0003	0.0001	0.0000	0.0013	0.0006	0.0037	0.0010	0.0051	0.0022	0.0032	0.0036	0.0013	0.0021	0.0007	0.0031
2280	0.0000	0.0004	0.0002	0.0013	0.0005	0.0038	0.0014	0.0057	0.0025	0.0033	0.0037	0.0014	0.0020	0.0010	0.0033
2400	0.0005	0.0003	0.0002	0.0016	0.0008	0.0042	0.0009	0.0052	0.0022	0.0033	0.0039	0.0012	0.0024	0.0006	0.0032
2520	0.0008	0.0006	0.0003	0.0018	0.0010	0.0047	0.0008	0.0054	0.0022	0.0035	0.0042	0.0012	0.0028	0.0005	0.0034
2640	0.0008	0.0006	0.0003	0.0018	0.0010	0.0047	0.0008	0.0054	0.0022	0.0035	0.0042	0.0012	0.0028	0.0005	0.0034
2760	0.0008	0.0006	0.0003	0.0017	0.0009	0.0043	0.0007	0.0049	0.0020	0.0031	0.0038	0.0011	0.0025	0.0004	0.0030
2880	0.0008	0.0007	0.0004	0.0017	0.0009	0.0043	0.0006	0.0049	0.0019	0.0032	0.0039	0.0011	0.0026	0.0004	0.0031
3000	0.0007	0.0006	0.0003	0.0014	0.0008	0.0035	0.0004	0.0036	0.0014	0.0023	0.0030	0.0007	0.0021	0.0002	0.0022
	0.0006	0.0004	0.0003	0.0012	0.0006	0.0033	0.0006	0.0040	0.0016	0.0026	0.0031	0.0009	0.0020	0.0004	0.0025

x in meters	RMSD of predictions – t in days																		
	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
0	0.0014	0.0018	0.0009	0.0016	0.0009	0.0124	0.0197	0.0044	0.0007	0.0008	0.0010	0.0018	0.0060	0.0245	0.0258	0.0043	0.0004	0.0018	0.0023
120	0.0020	0.0021	0.0013	0.0017	0.0008	0.0124	0.0197	0.0044	0.0008	0.0009	0.0012	0.0020	0.0060	0.0245	0.0258	0.0043	0.0004	0.0020	0.0025
240	0.0035	0.0026	0.0022	0.0020	0.0002	0.0098	0.0159	0.0035	0.0007	0.0010	0.0015	0.0021	0.0045	0.0198	0.0210	0.0032	0.0006	0.0022	0.0029
360	0.0057	0.0032	0.0037	0.0023	0.0007	0.0050	0.0086	0.0017	0.0006	0.0012	0.0019	0.0023	0.0016	0.0109	0.0118	0.0013	0.0009	0.0024	0.0035
480	0.0051	0.0035	0.0033	0.0029	0.0006	0.0104	0.0173	0.0037	0.0010	0.0015	0.0022	0.0032	0.0040	0.0216	0.0233	0.0031	0.0012	0.0032	0.0044
600	0.0056	0.0036	0.0036	0.0029	0.0006	0.0106	0.0177	0.0037	0.0010	0.0015	0.0023	0.0032	0.0041	0.0221	0.0239	0.0031	0.0012	0.0032	0.0044
720	0.0076	0.0042	0.0049	0.0031	0.0012	0.0068	0.0121	0.0023	0.0009	0.0016	0.0026	0.0032	0.0020	0.0153	0.0167	0.0016	0.0014	0.0033	0.0047
840	0.0081	0.0048	0.0053	0.0042	0.0026	0.0034	0.0068	0.0014	0.0011	0.0021	0.0034	0.0045	0.0004	0.0084	0.0099	0.0003	0.0022	0.0045	0.0064
960	0.0073	0.0050	0.0048	0.0049	0.0024	0.0103	0.0174	0.0041	0.0016	0.0025	0.0037	0.0056	0.0027	0.0210	0.0233	0.0027	0.0026	0.0055	0.0075
1080	0.0077	0.0051	0.0051	0.0050	0.0027	0.0094	0.0165	0.0037	0.0016	0.0025	0.0038	0.0057	0.0021	0.0202	0.0227	0.0022	0.0026	0.0056	0.0076
1200	0.0082	0.0056	0.0054	0.0056	0.0031	0.0110	0.0192	0.0043	0.0018	0.0028	0.0042	0.0065	0.0024	0.0234	0.0263	0.0026	0.0030	0.0063	0.0085
1320	0.0084	0.0056	0.0055	0.0055	0.0030	0.0105	0.0186	0.0041	0.0018	0.0027	0.0042	0.0063	0.0022	0.0228	0.0257	0.0024	0.0030	0.0061	0.0083
1440	0.0096	0.0061	0.0063	0.0060	0.0037	0.0087	0.0157	0.0036	0.0019	0.0030	0.0047	0.0069	0.0010	0.0189	0.0217	0.0017	0.0033	0.0067	0.0092
1560	0.0097	0.0064	0.0064	0.0064	0.0039	0.0115	0.0207	0.0045	0.0021	0.0032	0.0049	0.0074	0.0019	0.0253	0.0289	0.0024	0.0036	0.0072	0.0098
1680	0.0100	0.0066	0.0066	0.0067	0.0041	0.0127	0.0232	0.0049	0.0022	0.0033	0.0051	0.0077	0.0022	0.0285	0.0326	0.0027	0.0038	0.0075	0.0101
1800	0.0103	0.0066	0.0067	0.0067	0.0041	0.0123	0.0224	0.0047	0.0022	0.0033	0.0051	0.0077	0.0020	0.0276	0.0316	0.0025	0.0038	0.0075	0.0101
1920	0.0107	0.0066	0.0070	0.0063	0.0038	0.0112	0.0204	0.0043	0.0020	0.0031	0.0049	0.0072	0.0018	0.0250	0.0286	0.0023	0.0035	0.0070	0.0095
2040	0.0110	0.0066	0.0072	0.0058	0.0032	0.0105	0.0188	0.0040	0.0018	0.0029	0.0046	0.0064	0.0020	0.0231	0.0263	0.0023	0.0030	0.0063	0.0088
2160	0.0147	0.0077	0.0096	0.0068	0.0053	0.0031	0.0087	0.0010	0.0017	0.0035	0.0058	0.0074	0.0028	0.0112	0.0141	0.0011	0.0039	0.0074	0.0105
2280	0.0176	0.0084	0.0114	0.0073	0.0073	0.0095	0.0104	0.0037	0.0014	0.0038	0.0066	0.0075	0.0096	0.0118	0.0101	0.0060	0.0044	0.0077	0.0113
2400	0.0128	0.0076	0.0084	0.0080	0.0064	0.0039	0.0093	0.0020	0.0023	0.0040	0.0062	0.0091	0.0030	0.0109	0.0143	0.0007	0.0048	0.0088	0.0121
2520	0.0116	0.0078	0.0077	0.0090	0.0069	0.0121	0.0233	0.0050	0.0029	0.0044	0.0067	0.0107	0.0003	0.0280	0.0333	0.0016	0.0056	0.0102	0.0136
2640	0.0115	0.0078	0.0076	0.0090	0.0069	0.0120	0.0230	0.0050	0.0029	0.0044	0.0066	0.0107	0.0003	0.0274	0.0328	0.0016	0.0056	0.0102	0.0136
2760	0.0102	0.0070	0.0067	0.0083	0.0064	0.0111	0.0211	0.0047	0.0027	0.0041	0.0061	0.0099	0.0003	0.0250	0.0300	0.0015	0.0053	0.0094	0.0125
2880	0.0098	0.0071	0.0065	0.0084	0.0063	0.0146	0.0268	0.0059	0.0029	0.0041	0.0060	0.0102	0.0011	0.0319	0.0375	0.0026	0.0053	0.0096	0.0127
3000	0.0059	0.0050	0.0040	0.0069	0.0053	0.0138	0.0250	0.0056	0.0025	0.0033	0.0047	0.0086	0.0015	0.0296	0.0345	0.0027	0.0046	0.0080	0.0104
	0.0096	0.0060	0.0063	0.0061	0.0043	0.0106	0.0188	0.0041	0.0019	0.0030	0.0047	0.0070	0.0034	0.0228	0.0260	0.0027	0.0036	0.0068	0.0092

Table 14: Section 5.2. For each value of x in the first column, predictions for $t > 10$ days use data observed at $t \in \{3.5, 4, 4.5, \ldots, 10\}$ days. Observed data correspond to $z(x,t)$ and the prediction is given by the *best fitting cubic polynomial (solution of a linear least squares problem)*. Each cell of the table shows $|z_{\text{pred}}(x,t) - z(x,t)|$. In the left-hand part of the table, values of $z(x,t)$ correspond to synthetic trainind data (used in the fitting), while in the right-hand part of the table $z(x,t)$ correspond to synthetic (testing) data that is not being used in the fitting. The last line in the table shows the RMSD for each t.