uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Shohreh Khosravan Najafabadi
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Civil Engineering)
GRADE / DEGREE

Department of Civil Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Optimal Vector Interpolation of Asynoptic Spatial Survey of Vector Quantities for Interpolating ADCP Water Velocity Measurements

TITRE DE LA THÈSE / TITLE OF THESIS

Dr. C. Rennie
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Dr. B. Daneshfar

Dr. I. Nistor

Dr. P. Simms

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Optimal Vector Interpolation of Asynoptic Spatial Survey of Vector Quantities

for

Interpolating ADCP Water Velocity Measurements

Shohreh Khosravan Najafabadi

A Thesis
submitted under the supervision of Dr. Colin D. Rennie
in partial fulfillment of the requirements for the degree of
Master of Applied Science in Civil Engineering

The Ottawa-Carleton Institute for Civil Engineering
Department of Civil Engineering
Faculty of Engineering
University of Ottawa
Ottawa, Ontario

Your file Votre référence
ISBN: 978-0-494-25794-4
Our file Notre référence
ISBN: 978-0-494-25794-4

To

the Bahai youth in Iran

who are deprived of higher education

ACKNOWLEDGEMENTS

I would like to express my deep appreciation to my supervisor Dr. Colin Rennie for his guidance and encouragement and also his patience throughout this study. His insights and guidance were always valuable and beneficial. I would like to thank Francois Rainville from Environment Canada for providing the data for this work.

I wish to acknowledge the Bahai community of Canada and many people in Bahai community of Ottawa who kindly supported me throughout my study. In particular I thank Mrs. Sherri Yazdani, Mr. Ted Slavin, Ms. Barbara Pope, and Mr. Wes Maneval.

My special love goes to my beloved parents; my mother, Jannat and my father Mehdigholi for inspiring and encouraging the pursuit of my academic life. I also thank my brothers Shahriar and Shahram, my sister in law, Saghar, and my nephew Adib, who spiritually supported me during the time of my study.

ABSTRACT

In fields of study such as geophysics and hydraulics, many random variables are vector quantities, not scalars. Vector quantities require statistical techniques that are independent of choice of coordinate system. In this research a new optimal vector interpolation method, suitable for interpolation of asynoptically measured spatial vector fields, was developed and tested. The new method was compared to scalar interpolation by kriging. The test data were spatial Acoustic Doppler Current Profiler (ADCP) surveys of depth average fluvial water velocity in reaches upstream and downstream of a bridge. The interpolation procedures were evaluated by interpolating the fields with various amounts of data removal, and comparing to the actual measured field using a vector correlation coefficient previously developed by Crosby et al. (1993). The new optimal vector interpolation method was superior to kriging when all data were utilized (upstream reach) and for data removal rates of up to 30% (downstream reach).

CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY

The following symbols are used in this thesis:

A= coefficient matrix;

a_i and a_j= a pair of measured value;

C = sound speed (m/s);

$C(\ )$ = covariance function;

D_{ij} = Variogram;

$d(\underline{X},\underline{Y})$= the temporal and spatial distance between $\underline{X}$ and $\underline{Y}$;

$E(\ )$ = mean function;

E= expectation value;

F_D = Doppler shift frequency (Hz);

F_s = transmitted frequency of the sound from a stationary source (Hz);

$G(H)$ = variance of difference between measured values as a function of the distance between them (H);

H = spatial or temporal distance between points $\underline{X}$ and $\underline{X}+H$;

h= size of the variable;

M = mean;

m= number of observations in actual field (Y);

N= number of measured values at N points;

$N(H)$ = the number of $\underline{X}$ and $\underline{Y}$ pairs separated by H;

n=number of observations in *a priori* data field and also number of interpolated data;

n_1=size of sample one;

n_2=size of sample two;

P_i= individual sample values of variable P

$\overline{P}$ =mean of the variable P;

Q_i= individual sample values of variable Q;

$\overline{Q}$ = mean of the variable Q;

R = distance between measured points and interpolated point;

r_1=Correlation coefficient of sample one;

r_2=Correlation coefficient of sample two;

S = covariance;

s_u^2 = covariance of u;

s_v^2 = covariance of v;

s_{uv}^2 =cross covariance of u and v;

t= time;

$\hat{U}(\underline{X}_0)$ = linear estimator in kriging method;

$\langle U \rangle(\underline{X})$ = true point value;

u = easting component of a vector;

ui_{Xj} = easting component of the ith vector in jth block of a priory data field;

ui_{Yj} = easting component of the ith vector in jth block of actual data field;

V= relative velocity of the sound source and the sound receiver (m/s);

V_{water} = absolute water velocity;

V_{measured} = measured velocity by ADCP;

$V_{boatbybottomtracking}$ = relative velocity of the boat to the bottom of the river measure by bottom tracking method;

$V(\underline{X})$ = measured vector at location $\underline{X} = (x,y)$;

$V_{streamwise}$ = streamwise component of the vector in streamwise-cross-stream coordinate system;

$V_{cross-stream}$ = cross-stream component of the vector in streamwise-cross-stream coordinate system;

$V_{easting}$ = easting component of the vector in easting-northing coordinate system;

$V_{northing}$ = northing component of the vector in easting-northing coordinate system;

Vi_{Xj} = ith vector in jth block in a priori data field;

Vi_{Yj} = ith vector in jth block in actual data field;

v = northing component of a vector;

vi_{Xj} = northing component of the ith vector in jth block of a priory data field;

vi_{Yj} = northing component of the ith vector in jth block of actual data field;

W = two dimensional vector;

X= random vector (a set of n values, a priori data);

X' = *a priori* vector field;

X_i = ith block in a priori data field;

$\hat{X}$ =estimated vector (a set of n value);

Y=random vector (a set of m values, actual field);

Y' = spatially measured vector field;

Y_i = ith block in actual data field;

Z =Correlation coefficient testing statistics;

z_1=Conversion of r_1 to the value which estimated a population parameter;

z_2=Conversion of r_2 to the value which estimated a population parameter;

$\Gamma(H)$ = model fitting variogram ;

$\sum_E$ = error covariance matrix;

$\sum_{XX}$ = covariance matrix of the a priori data;

$\sum_{XY}$ = covariance matrix between a priori data and actual data;

$\sum_{YY}$ =covariance matrix of the actual data;

$\sum_{11}$ = covariance matrix of two dimensional vector W_1;

$\sum_{12}=\sum_{21}$ = cross covariance matrix of two dimensional vectors W_1 and W_2;

$\sum_{22}$ = covariance matrix of two dimensional vector W_2;

α=Type I error probability for significant testing;

χ^2=non-central chi-square distribution;

$\delta_{\underline{XY}}$ = Kronecker function;

$\varepsilon(\underline{X})$ = random error;

$\gamma(H)$ = observed variogram;

λ_i = coefficient determined by accurate statistical description of measured data;

θ = the average of the vector angles in a vector field;

ρ=cross correlation;

ρ_v^2 = Crosby vector correlation coefficient;

$\sigma_{z_1-z_2}$ = Z statistics normalization factor;

σ_R = standard deviation;

σ^2 = error correlation;

τ = Lagrange multiplier;

1 INTRODUCTION

1.1 Problem Statement

In various fields of study measured random variables can not be used directly for further studies such as analytical studies or numerical models. This is because the measured data are noisy due to measurement error. Further, measurement of the whole field is impractical, because not all the points or stations in a field of study are accessible to be measured, and sampling the whole field is time consuming, expensive, and difficult. Therefore data are only available at a limited number of points which are not uniformly distributed in the field of measurement. Despite of all these deficiencies, having a uniform map is possible using interpolation methods. Interpolation methods permit estimation of data in locations that are not sampled using data that are available in sampled locations.

Furthermore, for advanced studies, computing the association between fields is necessary to compare fields which are separated from each other in space and time, obtained with different methods, or describe different parameters. This comparison is achieved through calculating the correlation coefficient between the fields .

On the other hand, in various fields of engineering, random variables are vectors rather than scalars. Orientation of rock joints in geological formation, characteristic

hydraulics parameters such as velocity, surface wind velocity in oceanographic and meteorological studies, geophysical data, and characteristic parameters of oil reservoirs are examples of vector variables. Accordingly, vector statistics have received attention in fields of study such as meteorology, oceanography, and recently in hydraulics surveys (Young 1987).

According to Crosby (1993), vector quantities require special statistical consideration because a vector includes both magnitude and direction and it is further complicated by the fact that direction is a circular function. Most importantly, the choice of coordinate system produces a different result for variance and all higher statistical moments. Thus, statistical procedures do not yield unique results. For instance, if a vector is decomposed to its scalar components, NW and SE components will be different from N and E components and consequently the statistical information of second order or higher moments are not constant in different coordinate systems.

Second moments are used to calculate the correlation coefficient or to interpolate an irregular vector field by kriging (see below) and since they differ in different coordinate systems, there will not be a unique result for correlation and interpolation in a vector field. Therefore utilization of vector statistics is necessary for calculating valid statistical factors for both components of a vector (i.e. N and E components of a vector).

In the present research the goal is to create evenly distributed velocity vector fields in

rivers. Such velocity vector maps can be utilized to calibrate two-dimensional and three-dimensional hydraulic models.

1.2 Outline of Chapters

Relevant literature is reviewed in Chapter 2. The necessity of the interpolation strategies and specifically vector statistics are discussed. Two interpolation methods; a) kriging interpolation method and b) optimal vector interpolation method along with a vector correlation coefficient method are introduced. A brief history of the previous studies on each strategy is presented and the application of each strategy in hydrometric surveys is reviewed.

Chapter 3 explains the method of collecting the survey data along with the description and the formula of each kriging method, optimal vector interpolation method and Crosby vector correlation coefficient method. Moreover, the procedure carried out for applying each method in this study is presented.

Chapter 4 commences with the results of interpolating the vector fields estimated by kriging interpolation method, followed by the results achieved by applying optimal vector interpolation method. The Crosby vector correlation coefficient computed from correlating interpolated fields using each interpolation method and measured field are compared, and the effect of data removal on interpolation procedure with each method is studied. The dependency of each interpolation method on the choice

of the coordinate system is also scrutinized by comparing the estimated field in different coordinate systems. The achieved results then are discussed to conclude this chapter. Chapter 5 discusses the validity of each interpolation method and difficulties and advantages of each method for estimating vector maps on uniform grids.

2 LITRATURE REVIEW

2.1 Vector Interpolation

2.1.1 Introduction

Generally speaking, interpolation strategies smooth a noisy measured field, and produce data on a uniform grid. The estimation method which is used to interpolate a noisy measured field needs to meet the condition of unbiasedness. On the other hand, it needs to be optimum in the sense that it minimizes the estimation variance.

Spatiotemporal phenomena display random aspects with respect to time and space. Thus, estimating the variables at unsampled spaces and unsampled times is required to fill gaps in the sampled field. (see Rohani et al. 1990). Interpolation techniques for spatiotemporally distributed data were mainly developed for mining purposes but their principals are used in a number of scientific fields of study.

Bentamy (1996) utilized interpolation to make a regular field out of irregular measurements of spatiotemporally distributed velocity vectors in wind fields. In such spatiotemporal phenomena, data are collected irregularly due to sampling equipment and techniques, high cost, conditions in the field, and limited accessibility to the

phenomena in question. Therefore, measured data are irregular with gaps and cannot be used directly for further procedures.

Using an interpolation technique, a modeled surface can be created and mapped. Depending on the field of study, the modeled surface may represent a velocity vector field (Rennie and Millar 2004), a wind field map (Feliks et al. 1996; Inggs et al. 1995), mesoscale surface velocity stream function (Wilkin et al. 2002), etc. The resulting maps can be employed to give information for consequent studies such as bed material transport or description of the wind field in an area based on the location and number of the observations.

One field of study in which interpolation is employed is oceanographic surveys. Bretherton et al. (1976) utilized interpolation to evaluate the spatial structure of time-dependent velocity and density fields in deep oceans. In oceanographic surveys, surface wind information is also necessary. Given a measured wind field, it is necessary to develop vector interpolation methods to make a gridded field. The uniform wind field is required in the assessment of the air quality, dispersion of contaminants, initialization of diagnostic wind models such as mass consistent models (Sherman 1978; Ludwig et al, 1991), and development of prediction models with the available data (Daley 1991). These models are very sensitive to the interpolated wind vector field, particularly in complex topographies.

In river hydraulics, uniform maps created by a proper interpolation approach have

been used to develop the knowledge of spatial distribution of bed material transport which is necessary in river management (Rennie and Millar 2004). Such maps may be used to predict channel changes, analysis of the stability of hydraulic structures, and for the study of aquatic habitat.

2.1.2 Vector Interpolation Methods

A common and simple interpolation procedure is the inverse square distance ($1/R^2$) interpolation approach. In this formula R is the distance that separates the measured points from the interpolated point. The estimated value is a weighted sum of nearby measured points, where the weights are equal to $1/R^2$ (Porch and Rodriguez 1987).

A common approach for vector interpolation is to decompose the vectors into their magnitude and direction scalar components and interpolate the scalars using standard correlation and regression techniques. Porch and Rodriguez (1987) interpolated a wind field by interpolating the wind speed and direction in a complex terrain. Their interpolation procedure utilized measured data at the time of interpolation and statistical relationships from historical data between different stations. The interpolated field showed a better interpolated wind vector magnitude comparing to the $1/R^2$ interpolation method but not much improvement in the interpolated direction.

However, decomposing the two dimensional vectors into two independent scalars such as speed and direction reduces the physical information due to the physical

vector property (Crosby 1993). Furthermore, in the $1/R^2$ method, only the separation distance between the measurement stations is taken into account, thus the change of the wind direction and magnitude due to the terrain is not considered (Feliks et al. 1996).

The aforementioned inconveniences of common methods of interpolation for vector fields make studies on methods of interpolation, which are able to deal with vectors, necessary. In this study two interpolation techniques, which previously have been applied on wind fields, are evaluated and adapted for velocity vector fields in rivers.

2.1.2.1 Kriging

Kriging is a geostatistical method which is based on the Gauss-Markov theorem. (Matheron, 1971, Servain et al., 1993). Kriging determines interpolated values based on a linear combination of weighted neighbouring points, with the weights assigned based on the spatial variance of the field. In this section kriging is described and then studies are reviewed which used the kriging method to interpolate noisy and non-uniform measurements of vector fields to make maps of vector quantities on uniform grids.

Kriging is based on objective analysis in a least-squares sense, and yields the Best Linear Unbiased Estimator (BLUE) of a spatial distribution (Isaaks and Srivastava 1989). It is the best estimator because it is based on Gauss-Markov linear estimator which optimizes the variance of the residuals, it is linear since it is a linear

combination of available observations, and it is unbiased, since the mean of prediction errors (residuals) is zero (Issak and Srivastava 1989).

The objective of kriging is estimation of a value in each gridded point, based on the scattered measurements. The procedure assigns a calculated weight to neighbouring measured points. The technique which is used to calculate the weight of each point employs a modeled variogram. The variogram is simply a form of the spatial autocorrelation function which is defined as the inverse of covariance function (See Bentamy et al. 1996). Kriging thus depends on the variance of the field in question.

Stationarity of the measured data is one of the most important assumptions that are made in deriving the equations in kriging method. Stationarity indicates that the field in question has uniform mean and variance. In other words, data stationarity means that the joint distribution between two points depends on the distance that separates them from each other rather than their locations (Issak and Srivastava 1989). One advantage of the kriging method is that stationarity can be tested. Unfortunately, the data stationarity assumption is often violated; consequently, the kriging estimations might sometimes be non-BLUE due to the influence of non-stationarity on the variograms. (Rennie and Millar 2004).

The stationarity assumption is made during three steps of deriving the kriging equations. First, stationary is required to prove that unbiasedness requires that the sum of the kriging weights in a search area of a particular location be equal to one.

Second, the errors are minimized utilizing the data stationary assumption. Finally, this assumption is made for the third time when the variance error is minimized by estimating the variogram (Issak and Strivastava 1989).

If stationarity is maintained, in each grid there is an estimated vector which is optimal. Although there can be more than one estimated vector which is linearly dependent to the available measurements, only the vector calculated with this method is called optimum. That is because this one, on the average, has the smallest error. Besides, in this method the accuracy of the calculated vector in grids can be assessed by an expression for variance error which is called kriging variance (see Bentamy et al. 1996).

Minimizing the error variance is another advantage of the kriging method. To minimize the error variance, the standard deviation of the estimated value (σ_R)is calculated. It should be recognized that σ_R is not the standard deviation of error between modeled and actual field. Rather, σ_R is the square root of the difference between the modeled variance of the data field (i.e. the variogram sill) and the modeled variance between neighbouring points and the estimate location. In effect, the parameter σ_R shows how much the model has been capable to gain information from the data in adjacent points.

The disadvantage of the kriging method for vector fields is the effect of changing the coordinate system on the kriging results. Changing the coordinate system changes

the level of correlation between data and consequently changes the variogram observed from available data. Typically, the modeled variogram used in the kriging equations is fitted to the observed variogram. Since the modeled variogram is the main factor that defines the kriging weights of neighbouring points, interpolated results are different in different choice of the coordinate system.

Hayashi 1980, Legler et al. 1985, Bentamy 1996, etc. carried out different studies to develop methods for providing regularly spaced wind fields, particularly from scatterometer data. Wind vectors over the oceans may be measured by scatterometry, which is a microwave remote sensing technology utilizing a polar orbiting satellite. The wind vectors are asynoptic in that they measured irregularly in space and time because the satellite scans only a portion of the Earth at any instant. This problem is confounded by the fact that weather systems move, thus the measured field is unsteady.

Inggs and Lord (1995) presented a kriging method for vector interpolation. The method was a generalization of kriging of scalar fields, and employed the estimation variance and vector semi-variogram of the magnitude of difference vectors in the wind field.

Bentamy et al.(1996) utilized ERS-1 scatterometer-derived wind speeds and directions measured over a region of tropical Atlantic and tested the application of a kriging method. Their kriging method interpolated the vector magnitudes and vector

directions separately. Kriging of the direction was carried out by kriging the east and north components. Loss of some vector properties was inevitable during the vector decomposition.

The main assumption of the method was that the wind speed and the vector components, (u) and (v), were stationary and the area of study was appropriate for applying this assumption. The method used spatial and temporal structure functions for eastward and northward components of wind vector. Based on the Gauss-Markov theorem, the method objectively analyzed the vector field and made a gridded wind field. The estimated field was tested for accuracy by comparing with block averaged data.

In this study we will apply kriging method on a velocity vector field in order to estimate a uniform vector field. The interpolation procedure is carried out by decomposing the vectors to their u (streamwise) and v (cross-stream) components. Thus, the magnitudes of vectors components are utilized as the kriging variables to estimate two individual variograms. The streamwise and cross-stream components are utilized to minimize correlation between components, which should minimize influence of choice of coordinate system. However, this is not an interpolation method that strictly considers vector quantities.

2.1.2.2 Optimal vector interpolation

The optimal vector interpolation method (Bretherton et al. 1976), which is also called

the objective method, is also based on the Gauss-Markov theorem, thus it provides the least square error estimate that is a linear combination of neighbouring values. By means of this method, the statistics of the field and the measurement error along with an expression for the root mean square (rms) error are estimated using the observations from a limited number of stations. (Bretherton et al.1976). The statistical assumptions, upon which the method is based, are explicitly stated and stable statements. Applying these assumptions, an estimated value is calculated in each point which is optimal. The estimated value is optimal, since out of all estimated values which are linearly dependent on the measured data, it has relatively the smallest amount of error. (see Bretherton et al. 1976). The method has been used on open ocean data to provide two-dimensional maps of oceanographic variables in homogeneous, isotropic, and stationary fields. (see Freeland and Gould 1976; McWilliams 1976 for details).

Feliks et al. (1996) applied the optimal interpolation method to vector wind fields. The method explicitly considers the physical vector property along with the changes in the wind field due to topography, thus stationarity is not required. The method is unbiased and unique. Feliks et al. (1996) proved that changing or rotating the coordinate system of the field to be interpolated does not change the result. Likewise, the given error of estimate is independent of the choice of coordinate system.

However, this method requires *a priori* knowledge of the measured field and instruments to estimate the vector field. In particular, an initial estimate of the

measured field is required. It also needs an assessment about the optimal number of measured vectors and stations which are required to have a uniform vector field. Bretherton et al. (1976) found that the level of the accuracy of the *a priori* statistical information about the variable to be mapped and the density of the measured points determines the accuracy of the interpolated map. Previous studies suggest that with good statistical information about the flow field and a sufficient density of measured points, this approach yields an estimated map with a reasonable accuracy.

Feliks et al. (1996) employed the optimal vector interpolation method to interpolate a vector field of surface wind. For this study, eleven stations were selected in a complex terrain. The wind magnitude and direction were measured every second for 96 hours at these eleven stations. Data were then filtered by averaging the measured wind vectors in each fifteen minutes and decomposed to the easting and northing (u and v) components. By choosing two stations as observation stations, the time series were estimated at the nine remaining stations, using available historical data for the *a priori* estimates. Since the mean of the data was removed before interpolation, the mean was added to estimated wind vectors as the last step of interpolation. Comparison of the interpolated wind vector field from the optimal vector interpolation method with the inverse square distance method showed significantly better results in interpolating the vector fields with regards to both magnitude and direction.

Bretherton et al. (1976) used this technique to design sampling arrays for Mid Ocean Dynamics Experiment (MODE-I). Since the method works by minimizing the

estimated interpolation errors, the resultant error is based on the statistics of the field rather than the measured field itself. However as Bretherton et al. (1976) state, one of the important factors in this approach is a good approximation of the statistics of the field. Statistical interface, derived from a small amount of data, provide relatively poor interpolated fields.

2.1.3 Comparison of Kriging and Optimal Vector Interpolation Methods

In this section kriging and optimal vector interpolation methods are compared briefly. One of the important differences between these two methods which makes the optimal method advantageous over kriging method is its independency to the choice of coordinate system, whereas in kriging the coordinate system must be rotated such that the influence of the coordinate system is minimized.

Furthermore, in the kriging method, the density of the sampling points and the value to be mapped determine the accuracy of the resultant field (Bentamy et al.1996). In the optimal method, a realistic density of the measurement stations and their locations along with the appropriate *a priori* data reproduce an interpolated field with a high level of accuracy.

Kriging and optimal vector interpolation methods are both linear combinations of randomly measured variables. However, in kriging the estimated field is a linear combination of weighted neighbouring points, and the weight is computed based on the variance structure as a function of distance between the points. In the optimal

method, on the other hand, the weights are calculated through a covariance matrix of all pairs of data. In other words, while in the kriging method the distance between the points directly affects the computed weight, in the optimal method, only the similarity between points influences the covariance matrix.

Furthermore, the kriging method is based on the data stationarity assumption thus, when this assumption is violated, kriging does not offer the best linear unbiased estimate (Rennie and Millar, 2004). Similarly, Bretherton et al. (1976) state that the optimal method provides a powerful result, given the statistics of a measured field only if the measurement error is unbiased.

Kriging and the optimal interpolation method both appear to be capable of providing gridded wind vector data, sampled by the methods and equipment for wind fields. In particular, wind fields tend to be measured as time series at fixed weather stations scattered throughout the field of interest. The present study was performed to assess the accuracy of these two techniques in velocity vector fields which are measured asynoptically using an ADCP from a boat moving throughout the field. No at-a-point time series data are available to calculate variances.

2.1.4 Application of Interpolation Strategy to Hydrometric Surveys

According to Rennie and Millar (2004), knowledge of the spatial distribution of bed is required for prediction of channel changes, analysis of stability of hydraulic structures such as bridges, and evaluation of aquatic habitat. Maps of spatial

distribution of bedload velocity and bedload transport for calibration of numerical morphodynamic river models (e.g. Wu et al. 2000, Ruther and Olsen 2005) can be provided using the bottom-tracking feature of Acoustic Doppler Current Profiler (ADCP) (Rennie and Millar 2004). This technique measures apparent bed load velocity asynoptically throughout a reach of a river using a moving boat. The resulting data are non-uniformly distributed and noisy. An appropriate interpolation procedure is required to yield a smoothed vector map on a uniform grid. Rennie and Millar (2004) employed scalar kriging of streamwise and cross-stream velocity components.

Dinehart and Burau recently studied the influence of secondary circulation in channel bends (Dinehart and Burau 2005a) and river difluence junctions (Dinehart and Burau 2005b) on juvenile salmon smolt migration in the lower Sacremento River. In the river bend study, repeat cross-section surveys were conducted by ADCP at several cross-sections (Dinehart and Burau 2005a). They provide a view of secondary flow and sediment movement by interpolating the ADCP measurements. An inverse distance interpolation algorithm was used to yield straight cross-sections average velocity slices of the flow field from the repeat cross-section surveys that were measured along irregular paths. The interpolation strategy was used to provide uniform maps of cross-sectional velocity and backscatter which show the cross-stream velocity and suspended sediment distributions. Consequently, the cross-stream velocity profiles were utilized to determine the streamwise velocity profile for further hydraulic studies.

In the channel junction study velocity vector maps were created by inverse-distance interpolation of the measured velocity vectors obtained asynoptically by ADCP survey throughout the junction (Dinehart and Burau 2005b). Further, the ADCP surveys were repeated, thus three dimensional animations were created using time-series of interpolated velocity vector and backscatter fields. The animations were utilized to scrutinize the flow dynamic of the reach and the spatial movement of suspended sediment through the reach of the study. ADCP measurements and interpolated fields showed the potential to simplify the analysis of the flow and sediment discharge.

One of the other applications of interpolation techniques in hydrometric surveys, is in quantitative flow analysis from particle image velocimetry (PIV) data. PIV involves measurement of two-dimensional (or three-dimensional) velocities by illumination of particles in a plane using a pulsed laser light sheet. Continuous video images, obtained perpendicular to a light sheet parallel to the main flow, are converted to digital data and processed in order to record the velocity. Interpolation techniques are used to estimate missing values. Values may be missed due to incomplete particle information in some parts of the images, or refusal of some incorrect vectors after a vector validation process. (see Stamhuis and Videler 1995).

The goal of the present study is to compare kriging and optimal vector interpolation for smoothing the noisy maps of water velocity vector data, measured asynoptically by an ADCP. Subsequently the interpolated maps are evaluated by calculating the

correlation coefficient between the interpolated maps and the actual field to check the validity of the interpolation.

2.2 Vector Correlation Coefficient

2.2.1 Introduction

As discussed in the previous section, the validity of the interpolation methods needs to be evaluated by comparing the resultant interpolated field and the actual field. This comparison is carried out using a valid correlation coefficient technique. A larger correlation coefficient value demonstrates that two fields are more similar to each other. Thus, the correlation coefficient provides a test of the success of the interpolation method.

2.2.2 Background

Vector correlation coefficient difficulties, similar to vector interpolation problems, arise because a vector is composed of both magnitude and direction. This problem was recognized at least 75 years ago when Detzius (1916), Sverdrup (1917), Charles (1959), Buell (1971), Breckling (1989) tried to apply correlation coefficient techniques in meteorology and oceanography.

The correlation between parameters must be considered in some surveys because measurement of different parameters needs to be carried out by different instruments or in different time or space. For instance, in order to determine the correctness of

remotely sensed geophysical factors, one has to perform comparison of these parameters with simultaneous co-located in-situ measurements.

In coastal engineering, calculation and validation of wave hindcast information is very important. The common procedure to validate the hindcast is to compare the hindcast information with measured data at the coincident location. Since this particular study deals with circular data for wave direction, the validation of wave direction is more complex, but wave height and period can be validated easily using linear statistics. (Tracy 2002).

Common correlation coefficient methods such as Pearson product-moment can only be applied to vector fields if the vector is decomposed into its scalar components; magnitude and direction. The problem arises when correlation tests are done on the scalar components of the vectors in different choices of coordinate system. Different results of correlation coefficient are obtained in different choice of coordinate system. For instance, the results of correlating vector components decomposed in spherical, earth-oriented coordinate system is different from the ones from correlating the same vectors decomposed in a natural coordinate system (Crosby 1993).

Alternatively, correlating the vectors directly without decomposition offers a value as an assessment of correlation of the vectors, without missing the effect of vector direction and magnitude. Crosby (1990) redefined a vector correlation originally proposed by Hooper (1959) and Jupp and Mardia (1980). Other correlation

coefficients that consider both magnitude and direction include Durst (1957), Court (1958), and Breckling (1989). Mardia and Puri (1978) and Stephens (1979) developed correlation coefficients for direction based on the unit circle. Although there is not a general definition that is used currently for vector correlation coefficient, the technique proposed by Crosby (1990) has relatively complete series of useful statistical properties for both population parameter and sampled data statistics. Besides, it is a generalization of the standard definition used for scalar fields and it accounts for both the direction and the magnitude, thus it is a desirable method in vector studies.

2.2.3 Method

The Crosby (1993) correlation coefficient is:

$$\rho_v^2 = \mathrm{Tr}[(\sum_{11})^{-1}\sum_{12}(\sum_{22})^{-1}\sum_{21}]. \qquad [2.1]$$

In two dimensional problems ρ_v^2 is between 0.0 and 2.0, whereas the conventional correlation coefficient for scalar variables ranges between -1.0 and 1.0.

The formula offered by Crosby et al. (1993) is applied simultaneously to u and v components of the vectors, without decomposing them into their scalar components. The Crosby vector correlation coefficient is symmetric and does not change by changing the coordinate system. Conversion in coordinate system includes rotation of the axis as well as changing the scale of the coordinate system. The proof of this property of the method is prepared by Crosby (1993) for two dimensional problems, but it can be generalized to three and four dimensional fields. If the vectors are

linearly dependent, the correlation coefficient is maximum whereas it is zero when the vectors are independent. This method has been tested in the fields of meteorology and oceanography.

Crosby (1993) summarized the ρ_v^2 correlation coefficient properties as follows; it is the generalization of the square of the standard scalar correlation coefficient and its properties give an instinctive meaning to the definition. Also its properties relate the definition to the multivariate coefficient of determination.

Crosby (1993) emphasized that a high correlation coefficient does not mean that two compared fields are equal, rather, it suggests two fields differ in a similar way. For example in two velocity vector fields, where all velocities of one field are rotated by 90° and scaled by a constant factor compared to the other field, these two fields are perfectly correlated.

2.2.4 Application to Hydrometric Survey

Vector correlation coefficients are mostly utilized in hydraulic surveys for comparing the outputs of a two dimensional hydrodynamic model with the actual field. The model outputs are normally two dimensional maps of depth averaged velocity vectors. This comparison procedure is carried out for calibration purposes. In this case, we need to use statistics to quantitatively evaluate the relationship of observed data and data obtained from modeling the fields.

Rennie and Millar (2004) employed the correlation coefficient method developed by Crosby et al. (1993) to compare the distributions of bedload velocity and the near-bed and depth averaged water velocity. The method was chosen because it is independent of the choice of coordinate system.

In the present study, the Crosby vector correlation coefficient is utilized to evaluate kriging and optimal vector interpolation by comparing the interpolated velocity vector field and the actual velocity vector field. Data for this study are collected using an Acoustic Doppler Current Profiler (ADCP) in a river. In chapter three, the method used for collecting the data is explained. Also, the ADCP is briefly introduced. Chapter three also contains detailed descriptions of the two interpolation methods. Further, the proposed correlation coefficient method for vectors, proposed by Crosby (1993) is scrutinized Chapter 4 provides the results of the comparison between the two interpolation methods.

3 METHODS

3.1 Introduction

In this chapter two interpolation methods are employed to make uniform velocity maps by interpolating irregular velocity vector fields measured by ADCP. The two vector interpolation methods to be assessed are the kriging method (Bentamy 1996) and a modified version of the optimal vector interpolation method (Feliks et al. 1996). In order to evaluate these two techniques, the proposed vector correlation coefficient by Crosby (1993) is utilized.

Available data for this study come from about eight hours of velocity measurement in Gatineau River (45.46°N, 75.70°W). Measuring the velocity vector field was carried out using an acoustic Doppler current profiler (ADCP) mounted on a boat and connected to the Global Positioning Systems (GPS) receivers. The survey was completed in a 400 meter reach of the river on November 3rd 2004. The survey was designed by the author, but was conducted by Environment Canada personnel. The area to be measured was in a location where the river intersected a highway bridge with elliptical piers of 2.4 m width. Therefore, the measured area was divided into two separate portions (upstream and downstream of the bridge) by the highway bridge. These two separate portions provide the possibility of interpolating the downstream and upstream flow independently. This experimental design allows for

assessment of possible influence on the interpolation procedures of greater flow field spatial variability due to the bridge pier wake.

Acquired data by ADCP were output from WinRiver with separate files for each path that the boat took during the measurement process. Data were then imported to Matlab 7.1 programming software and processed for further analysis. Matlab is a programming software for numerical computations, providing 2D and 3D graphical displays of information and solving problems in various engineering and science fields. Surfer8 software was employed to interpolate the field with kriging method. This software is a powerful tool for contouring and 3D surface mapping . In order to interpolate the field using optimal vector interpolation method, program algorithms were written in Matlab (a Matlab tutorial is available at http://www.cs.wvu.edu/~trapp/wvumatlab.htm). The vector correlation coefficient (Crosby 1993) was applied to compare the interpolated fields and spatially block averaged measured data.

3.2 Data Collection

Velocity vectors for this study where sampled by an acoustic Doppler current profiler (ADCP) mounted on a vessel. The vessel-mounted ADCP measured velocity profiles along a defined path in the reach. In this section, ADCPs are briefly introduced. Also the data collection procedure for this study using an ADCP is explained.

3.2.1 Acoustic Doppler Current Profiler (ADCP)

ADCPs are the only feasible systems for measuring the discharge in rivers and estuaries with tidal effects or rivers and canals with unsteady flow. (Simpson and Oltman 1993). They are also capable to be substituted with conventional discharge measurement systems and practical to use in upland rivers where wading measurement is not possible due to the depth. (Simpson 2001). ADCPs were designed to work based on acoustic velocity principle and Doppler shift effect.

3.2.2 Doppler Principle

According to Simpson (2001), the ADCP employs sound in ultrasonic range to record the velocity information in vertical column of water. The measurement is based on the Doppler principle. Doppler principle states that the change in frequency of a source is related to the relative velocities of the source and the observer. Doppler shift frequency is calculated from the following formulae:

$$F_D = F_S(\frac{V}{C}) \qquad [3.1]$$

where F_D is the Doppler shift frequency, F_s is the transmitted frequency of the sound, V is the relative velocity of the sound source and the sound receiver and C is the speed of the sound.

With an ADCP, an ultrasonic sound pulse is bounced off small particles of sediment and other material which are assumed to move at the same velocity as the fluid in a water column. The ADCP then receives the return echo of the acoustic backscatters in

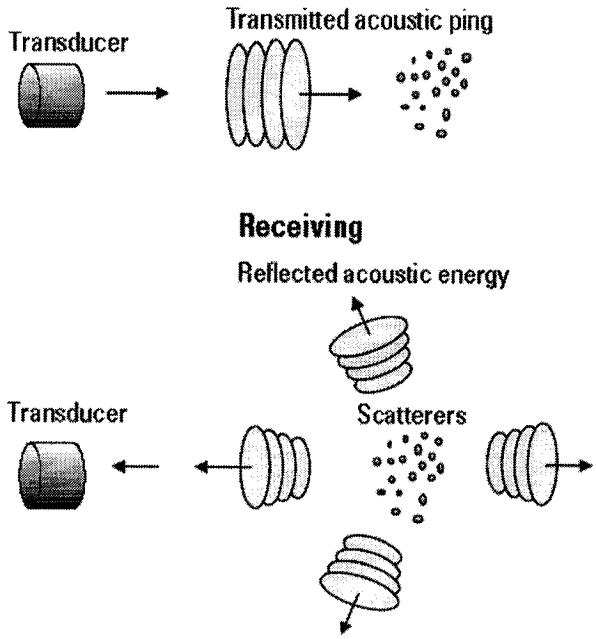

Figure 3. 1 An acoustic pulse being scattered . (Water survey of Canada. ADCP Principles, 2004)

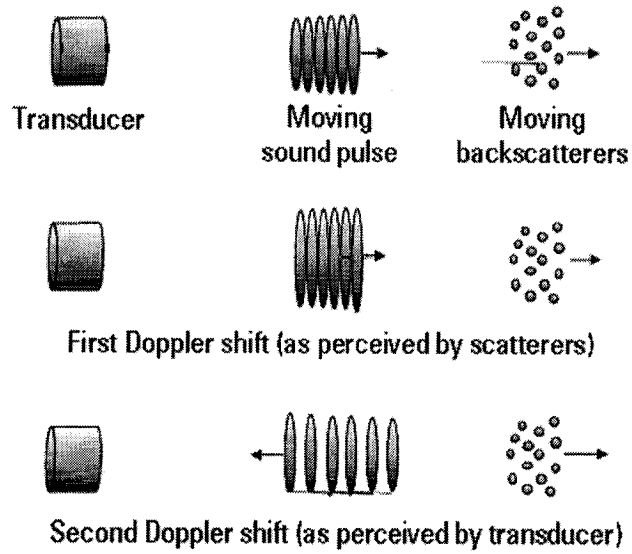

Figure 3.2 Reflected pulse showing two Doppler shift. (Water survey of Canada. ADCP Principles, 2004)

the water column. Doppler shift is calculated using an onboard signal processor in ADCP upon receiving the return echo. In Figure 3.1 a diagram of transmitted acoustic pulse and the returned acoustic energy can be seen. If the scatters move away from the ADCP, the sound frequency decreases. The change in the frequency is proportional to the relative velocity of the ADCP and the scatters.

Part of the sound which has been affected by Doppler shift is backscattered by the particles towards the ADCP. Therefore the sound is shifted two times, one time when it is received by scatters and the second time when it is received by the ADCP. Consequently, the Equation 3.1 changes to:

$$F_D = 2F_S(\frac{V}{C}) \qquad [3.2]$$

3.2.3 Bottom Tracking and Real Time Kinematics

Acoustic Doppler current profiler technique measures the velocity of the water relative to the ADCP. The absolute water velocity can be calculated by knowing the measured velocity by ADCP and the relative velocity of the ADCP to the river bottom which is equal to the relative velocity of the boat to the river bottom. ADCPs that are used to measure the discharge are able to calculate the vessel velocity using bottom track pulses. It also measures the profiled depth range for each beam, and the average depth which is used as the water depth for the measured velocity profile.

The relative velocity of the ADCP to the river bottom also can be measured using

differential Global Positioning System (DGPS). Boat velocity data which are obtained by bottom tracking method are more accurate than data which are given by DGPS, but when the bed is mobile, the bottom tracking is not reliable. (See Simpson 2001, for more details).

3.2.4 GPS

Global Positioning System (GPS) is an alternative for measuring the relative velocity between ADCP and river bottom. GPS positions are determined by triangulation of distances to known satellite positions, with distances determined based on propagation time of radio signals from the GPS satellites. Required positioning accuracy is obtained when the GPS signal is differential, in which errors are determined at a known base station and corrections are sent by radio to the GPS receiver. The errors result from propagation delays of the GPS satellite radio signals, and can be determined from phase shift of the GPS Pseudo Random Code (code phase) or more precisely from phase shift of the radio wave itself (carrier phase). Code phase differential corrections are obtained from a Coast Guard beacon if it is available in the area, or carrier phase Real Time Kinematic (RTK) corrections can be obtained by using two GPSs which work as rover and base GPS. RTK is the most reliable and accurate correction system (See Simpson 2001).

Figure 3.3 shows the typical arrangement of rover and base GPS. Rover GPS is located above the ADCP. RTK correction is done by the base GPS which is positioned above a control point on shore. The compass calibration is necessary to be

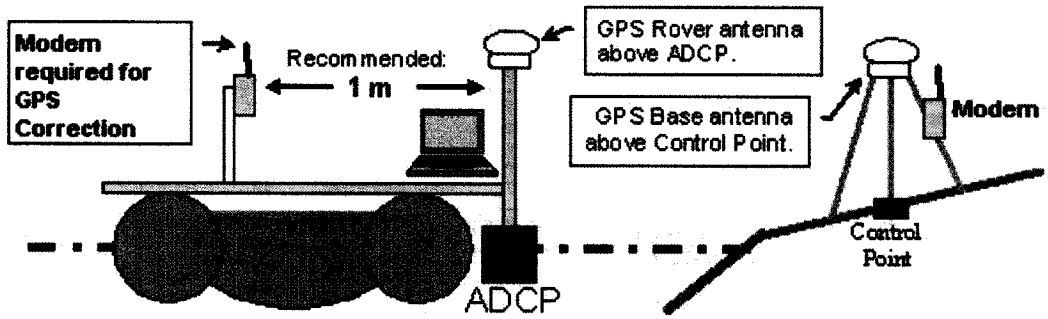

Figure 3.3 Typical deployment of GPS base and rover along with ADCP (Simpson 2001)

performed when GPS is used for correction.

3.2.5 Velocity Field Measurement Using an ADCP

The measurement process in this technique is carried out by a vessel mounted ADCP moving on a defined path in the survey field. ADCPs have either three or four transducers which are orthogonal to each other and spaced in 120° or 90°. The transducers are mounted near the water surface by a downward direction with the angles of 15° to 30° from vertical axis (Figure 3.4). Each transducer sends the sound and receives the backscatter itself. Given measured velocity components with at least three beams, three dimensional velocity vectors in Cartesian coordinate system are computed using trigonometry. (Figure 3.5)

Since three velocity components measured by three different beams are not in the same point in space, the assumption of horizontally homogeneous velocity field is essential to make the calculated velocity correct. By homogeneous velocity field, it means that all measured velocities in the field have the same magnitude and direction. In four beam ADCPs, the fourth beam is used to compute the error velocity and test the assumption of homogeneous velocity field. (Figure 3.6). See Simpson 2001 for more details.

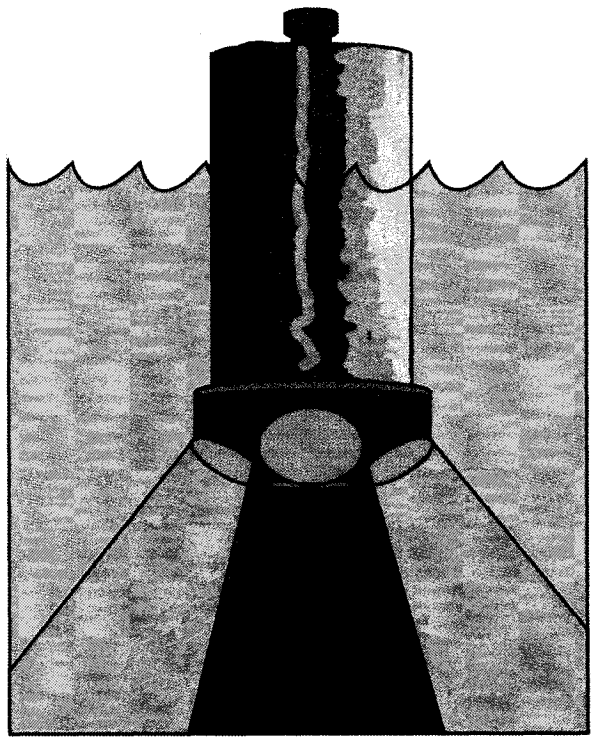

Figure 3.4 Downward looking ADCP. Water survey of Canada. ADCP Principles, 2004)

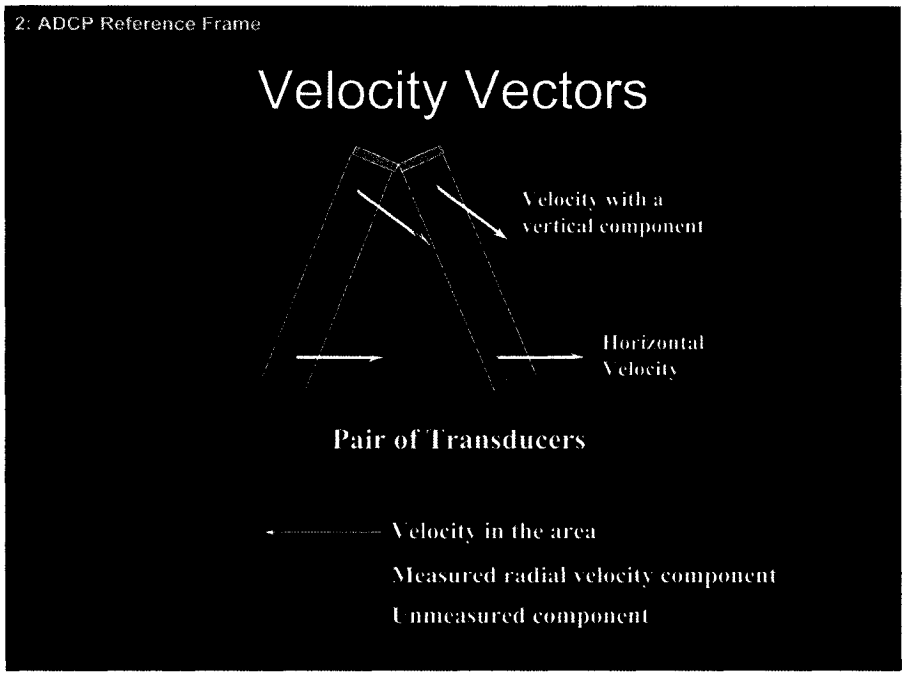

Figure 3.5 Velocity vectors. (Water survey of Canada. ADCP Principles, 2004)

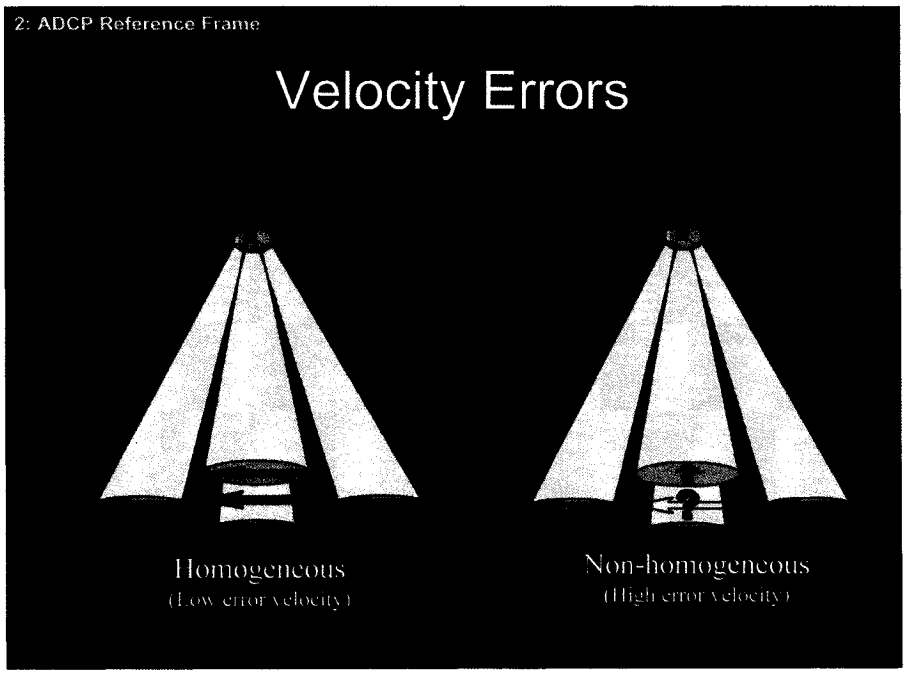

Figure 3.6 Homogeneous versus non homogeneous. (Water survey of Canada. ADCP Principles, 2004)

Using an internal compass, the three dimensional velocity vector in Cartesian coordinate system is converted to the Earth coordinate system (East, North, Up). The measured data influenced by magnetic field required to be corrected for magnetic declination. The magnetic declination is the difference between the true North and the magnetic North.

3.2.6 Data Collection

In this study, the velocity vectors were measured using an ADCP in a 400m reach of the Gatineau River (45.46°N, 75.70°W). The width of the river in the area of survey was 200 m and the maximum depth was about 8 m. the ADCP survey was completed in about 8 hours. River stage was relatively low on the day of measurement, with depth average velocities of about 0.6 m/s. The area of measurement was intersected by a highway bridge. The bridge piers were elliptical with a width of 2.4 m. Weak alternating vortices were generated from the bridge piers in a typical Karman street. The relevant Reynolds number was 1.4E6.

Data from ADCP were processed in WinRiver software and exported to Matlab software for further computations. Before exporting the data to Matlab, magnetic declination of the site which was estimated to be -14.3° (ref NRCan website) by the ADCP was set to its true value in the location of study which is -13.65° (Rennie and Rainville 2006). The vector plot of the bedload velocity data for the reach is presented in Figure 3.7. The continuous line represents the path of the boat which was

started from downstream region of the bridge. Figure 3.7 shows a stationary bed except the areas around the bridge piers.

This observation is encouraging that bottom tracking data can be used as the velocity reference. However the average bed speeds of -0.0393 m/s for easting direction and -0.0307 m/s for northing direction with bottom tracking reference suggest the idea of non stationary bed. Therefore, GPS reference was checked as the velocity reference. Still it showed the bed speed of -0.0330 easting and northing. Alternatively the plotted data demonstrated biased bed speed vectors toward upstream concentrated in the measured longitude around the piers. In other words, the bed speed direction is opposite of the flow direction.

These observations together suggested that the bed was not moving. Instead, the calculated bed speed in the area close to the piers is because of the inaccuracy of the operation of the ADCP compass affected by ambient magnetic field created by the steel used in the construction of the bridge piers and deck. Consequently the compass returns from north toward the piers and influences the process of measurement around the pier. The solution to this problem is excluding the measured paths around the piers in both downstream and upstream of the bridge. Therefore bottom tracking can be employed as the reference for calculating the absolute water velocity through Equation 3.3:

$$V_{water} = V_{measured} - V_{boatbybottomtraching} \qquad [3.3]$$

where V_{water} is the absolute water velocity, $V_{measured}$ is the measured velocity by ADCP

and $V_{\text{boatbybottomtracking}}$ is the relative velocity of the boat to the bottom of the river measure by bottom tracking method.

3.2.6.1 Separating Downstream and Upstream

Figure 3.8 shows the velocity vector field in the 400 m reach of Gatineau River measured by means of an ADCP. Since the measurement is performed in depth cells, measured velocity vectors in each water column were depth averaged to use for further analysis. Data measurement was carried with boat speed of approximately 0.5 m/s and ADCP sent signals every 0.5 s (2 Hz). Because of the high density of four measured vectors per 1 meter, in Figure 3.8 ten ensembles of resultant depth averaged velocity vectors were averaged.

The unmeasured area under the bridge, which is seen in the middle of the velocity vector map in Figure 3.8, along with the corrupted data near the bridge piers, were excluded from the whole data field by separating portions downstream and upstream of the bridge (Figure 3.9a and b). In this way the results will not be influenced by the vectors affected by bridge piers. Furthermore, the influence of the vortices caused by bridge piers in the downstream portion of the river can be studied. (Figures3.9b). In the following sections the procedure of interpolating the upstream and downstream portion of the measured velocity vector field along with the formulae of each method will be presented .

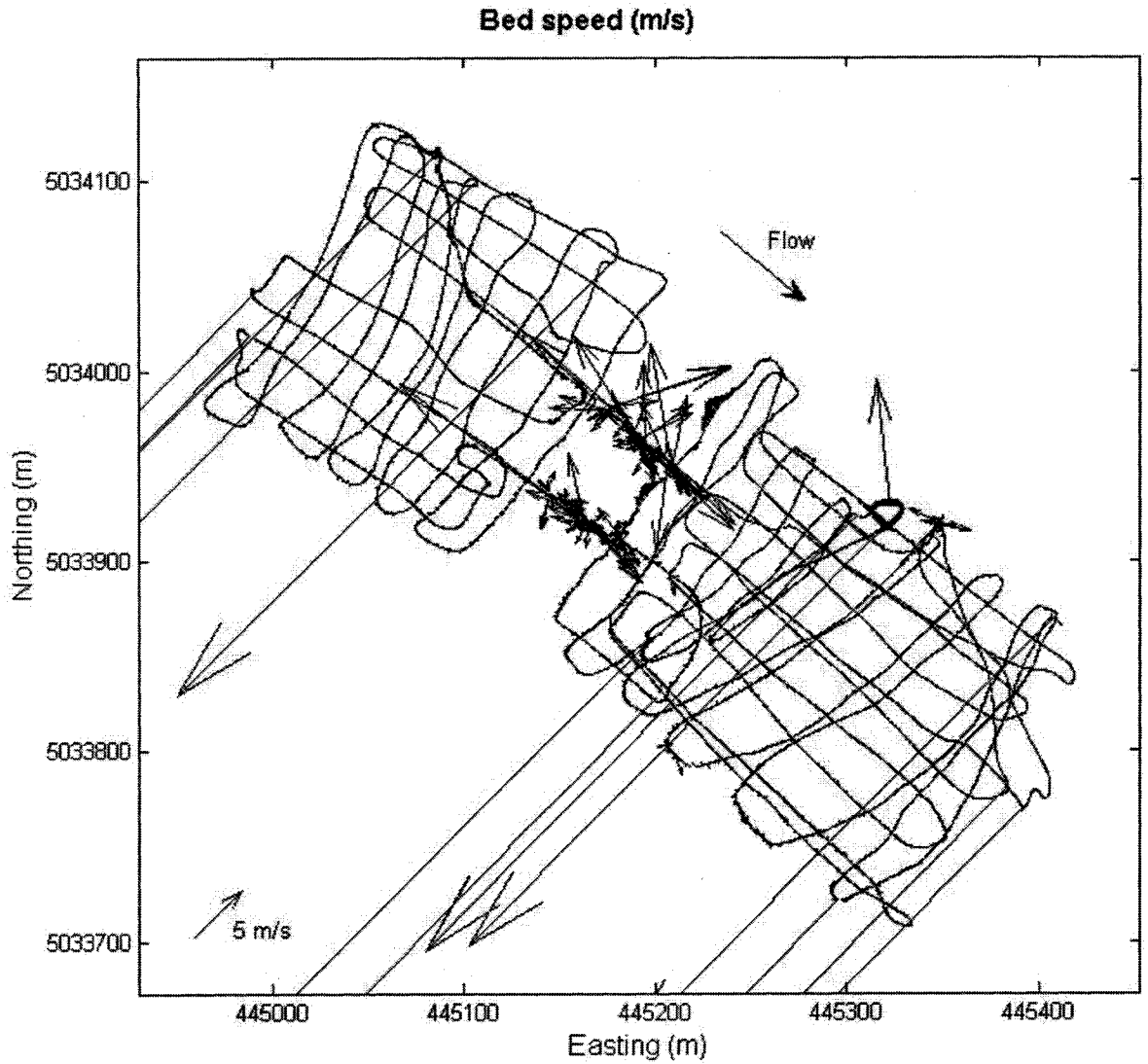

Figure 3.7 Bed speed map

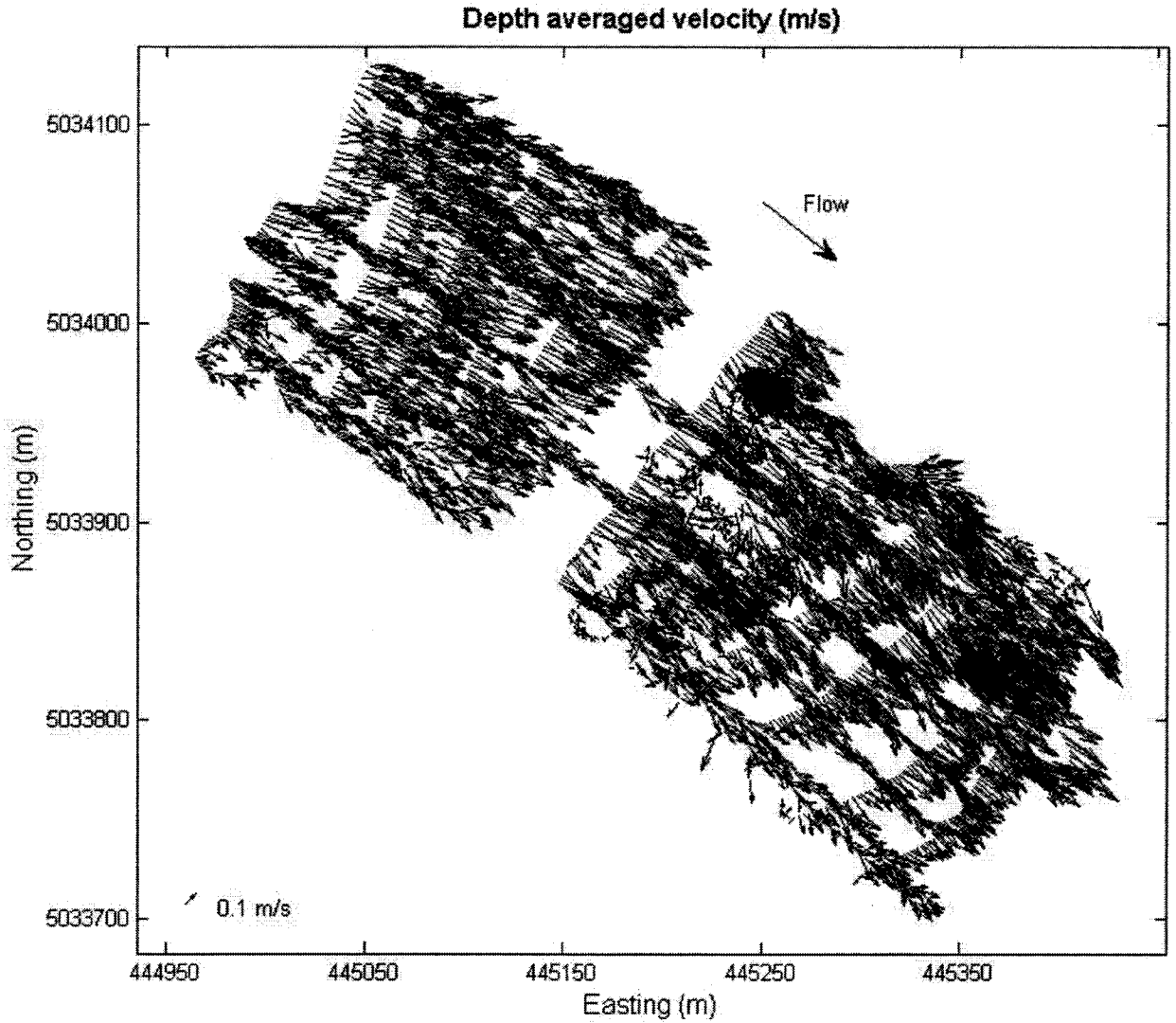

Figure 3.8 Depth average velocity vector for every 10 ensembles

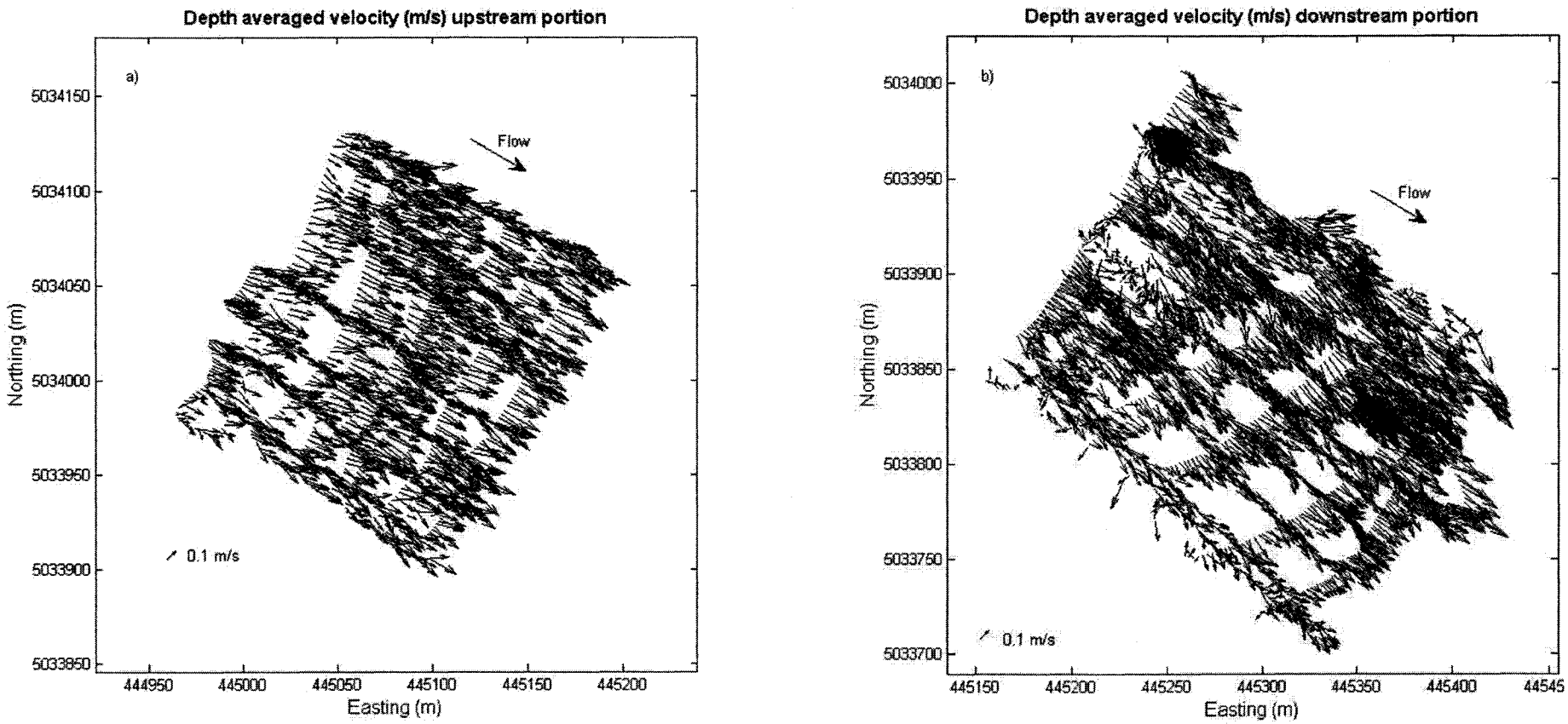

Figure 3.9 Depth average velocity of every 10 ensembles in a) upstream and b) downstream portion of the measured field. (vectors are scaled)

3.3 Kriging Method

3.3.1 Introduction

Formulae for kriging developed by Bentamy et al. (1996) are presented here. By means of this technique, one dimensional unmeasured values are estimated and a uniform map of the quantity is made from irregular measurements. Bentamy et al. (1996) used the method to estimate wind vectors which are two dimensional quantities. To make the application of the method possible, they resolved the vectors to their magnitude and direction and interpolated the scalars separately.

3.3.2 Formulae

In the kriging method, a linear combination of N measured vectors $V(\underline{X})$ in location $X=(x, y)$ and time t, estimates $\hat{U}(\underline{X}_0)$ with zero bias in each grid point:

$$\hat{U}(\underline{X}_0) = \sum_{i=1}^{N} \lambda_i . V(\underline{X}_i) \qquad [3.4]$$

If there is either geophysical or instrumental error $\varepsilon(\underline{X})$, then the measured vector

$$V(\underline{X}) = \langle U \rangle(\underline{X}) + \varepsilon(\underline{X}) \qquad [3.5]$$

where $\langle U \rangle(\underline{X})$ is the true point value.

The parameter λ_i in Equation 3.4 is the coefficient determined by accurate statistical description of the data. Calculating λ_i requires assumptions of the spatial and temporal variability of the data in the area of measurement. The basic assumption is

that the measured field is statistically homogeneous and the mean and covariance functions of the field are known:

$$M = E(\langle U\rangle(\underline{X})) \qquad [3.6]$$

$$C(H) = E((\langle U\rangle(\underline{X}+H) - M)(\langle U\rangle(\underline{X}) - M)) \qquad [3.7]$$

E ( ) and C ( ) are mean and covariance function respectively. Kriging requires that the mean of the variable $\langle U\rangle(\underline{X})$ is not dependent on the location $\underline{X}$, and the covariance of $\langle U\rangle(\underline{X})$ and $\langle U\rangle(\underline{X}+H)$ is only dependent on H which is the spatial and temporal distance between $\underline{X}$ and $\underline{X}+H$. In other words, the mean is constant over the spatiotemporal domain and data are stationary.

Another assumption required for kriging is zero correlation between the errors and between the errors and measured data ‹U›.

$$E(\varepsilon(\underline{X}).\varepsilon(\underline{Y})) = \delta_{\underline{XY}}.\sigma^2 \quad \text{and} \quad E(\varepsilon(\underline{X}).\langle U\rangle(\underline{Y})) = 0 \qquad [3.8]$$

where $\underline{X}$ and $\underline{Y}$ are a pair that are separated from each other by H in space or time and $\delta_{\underline{XY}}$ is the Kronecker function. The simplest interpretation of Kronecker function says that it is 0 when $\underline{X} \neq \underline{Y}$ and is 1 when $\underline{X} = \underline{Y}$)

Equations 3.6 and 3.7 can be simplified:

$$E(\langle U\rangle(\underline{X}+H) - \langle U\rangle(\underline{X})) = 0 \qquad \text{for a fixed } \underline{X} \text{ and } H \qquad [3.9]$$

$$E((\langle U\rangle(\underline{X}+H) - \langle U\rangle(\underline{X}))^2) = G(H) \qquad [3.10]$$

These two equations satisfy the assumption of stationarity; Equation 3.9 shows no expected difference between $\langle U\rangle$ values over the location (x, y) and time (t) and

Equation 3.10 defines $G(H)$ which is the variance of differences between measured $\langle U \rangle$ values as a function of H.

However, $G(H)$ is related to the variance function $C(H)$ in Equation 3.7

$$G(H) = E((\langle U \rangle(\underline{X} + H) - M) - (\langle U \rangle(\underline{X}) - M))^2) = 2.(C(0) - C(H)) \quad [3.11]$$

In practice we consider the quantity:

$$\Gamma(H) = \frac{1}{2}.G(H) = C(0) - C(H) \quad [3.12]$$

$\Gamma(H)$ is the variogram which behaves similar to the inverse of covariance function.

With the assumptions discussed above, the Gauss-Markov theorem offers the best estimator $\hat{U}(\underline{X})$ as the least-square optimum linear estimator of $\langle U \rangle(\underline{X})$. The estimation of the linear estimator $\hat{U}$ is equal to the estimation of coefficients λ_i which produce the minimum of the function L:

$$L = E((\hat{U} - \langle U \rangle)^2) - 2.\tau(1 - \sum_{i=1}^{N} \lambda_i) \quad [3.13]$$

where τ is the Lagrange multiplier. Satisfying the following constraint ensures zero bias on the estimator (3.4):

$$\sum_{i=1}^{N} \lambda_i = 1 \quad [3.14]$$

Using Equations (3.4) and (3.8) the function L is defined as:

$$L = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i((\lambda_j.C(i,j)) - 2C(0,i)) + C(0,0) + 2.\tau(\sum_{i=1}^{N} \lambda_i^2 - 1) + \sum_{j=1}^{N} \lambda_i^2 \sigma^2 \quad [3.15]$$

Where $C(i,j) = E(\langle U \rangle(\underline{X}_i).\langle U \rangle(\underline{X}_j).\langle U \rangle(\underline{X}_j))$ [3.16]

When the first derivative of L is set to zero with respect to coefficient $\lambda_1, \lambda_2, \ldots \lambda_N$, it means that those coefficients are the answer for the N+1 linear system:

$$\sum_{i=1}^{N} \lambda_i . C(i,j) - C(j,0) - \tau + \lambda_j \sigma^2 = 0, \qquad \sum_{i=1}^{N} \lambda_i = 1 \qquad [3.17]$$

Here j ranges from 1 to N.

The following linear system is gained, when the basic assumptions are applied:

$$-\sum_{i=1}^{N} \lambda_i . \Gamma(i,j) + \Gamma(j,0) - \tau + \lambda_j \sigma^2 = 0, \qquad \sum_{i=1}^{N} \lambda_i = 1 \qquad [3.18]$$

The above system is a linear system which is called the kriging system. Its solution shows that the variance between the estimated field and the actual field is:

$$E((\hat{U} - \langle U \rangle)^2) = C(0,0) + \tau - \sum_{i=1}^{N} \lambda_i . C(i,0) \qquad [3.19]$$

The variance value shows the quality of the estimator and how good the estimator is for interpolating data in unsampled time and place. Better knowledge of the variogram given by Equation 3.12 gives a better estimated field by the kriging method.

Field stationarity grows with the convergence of the observed variogram:

$$\gamma(H) = \frac{1}{N(H)} \Sigma_{d(\underline{X},\underline{Y})=H} (V(\underline{X}) - V(\underline{Y}))^2 \qquad [3.20]$$

and $G = E((\langle U \rangle(\underline{X} + H) - \langle U \rangle(\underline{X}))^2)$ in Equation 3.10.

$d(\underline{X},\underline{Y})$ is the temporal and spatial distance between $\underline{X}$ and $\underline{Y}$, and $N(H)$ is the number of $\underline{X}$ and $\underline{Y}$ pairs separated by H. In Equation 3.12, $\Gamma(H)$ is the model fitting of observed variogram ($\gamma(H)$) in the least-squares sense.

3.3.2.1 Variograms

The goal of using kriging interpolation method is achieving the best linear unbiased

estimation of an irregularly measured field. Since the estimated values are the result of the linear combination of measured points, each point needs to be assigned with a weight value. The weight is calculated using spatial continuity of the points. In other words, the weight is calculated based on how far the points are from each other. This satisfies the stationarity assumption which tells us that the joint distribution of data only depends on their locations. The tool which is used to calculate the weight of each point in kriging method is the variogram.

In a measured field, there is a distance in which pairs of data start to become more different from each other as the distance between the pairs of observations become larger. Variograms are employed to calculate this distance. The variogram of a pair of measurements (a_i and a_j) is the mean value of D_{ij}

$$D_{ij} = 0.5(a_i - a_j)^2 \qquad [3.21]$$

Variograms are either experimental variograms or model variograms. The experimental variograms smooth the data to highlight the trend. A model variogram is the most suitable fitted standard mathematical function to the data. See Figure 3.10. A typical model variogram is shown in Figure 3.11. Two points that are very close to each other might have different values. This difference produces an expected value of $0.5(a_i - a_j)^2$ greater than zero which is called "nugger effect". The maximum value of the expected value of $0.5(a_i - a_j)^2$ is the maximum height of the curve and is

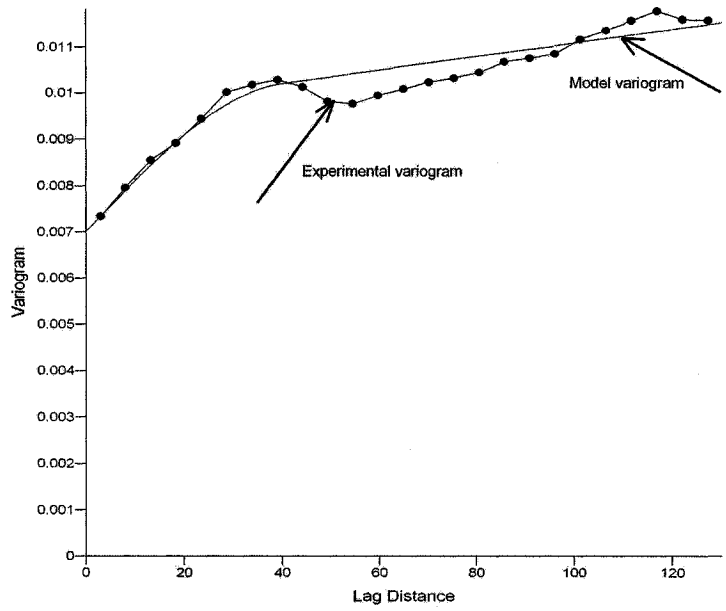

Figure 3.10 Example experimental and modelled variograms.

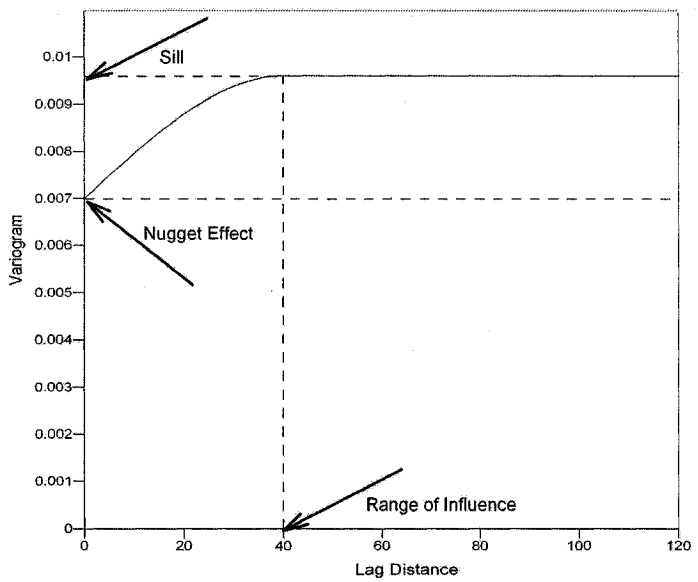

Figure 3.11 A typical model variogram with the nugget effect, the sill and the range of influence

called the "sill". This happens in the "range of influence" which is a distance in which two points have effectively independent values. The standard mathematical models for variograms are Gaussian model, spherical model, exponential model, and power model. (B.F. Manly 2001)

3.3.3 Kriging Procedure

In this study, the kriging method has been used to interpolate two dimensional vector velocities. The measured field was assumed to be stationary. As discussed previously, the kriging method is derived based on data stationarity assumption. In other words, if the measured field is not stationary, the resultant estimated field is not BLUE. Calculated mean and variance of 0.20 m/s and 0.01 m^2/s^2 for downstream portion of the bridge comparing with relevant mean and variance of 0.21 m/s and 0.01 m^2/s^2 for upstream portion respectively suggest that the stationary assumption is not violated in this study.

Since the kriging method is an interpolation method which is able to estimate scalars or one dimensional quantities, an approach is needed to adapt the method for two dimensional vectors. Bentamy et al. (1996) kriged wind vectors by interpolating the velocity magnitudes and separately interpolating the easting and northing components to estimate the directions. In this study the velocity vectors were decomposed to their easting and northing components and the vector components were interpolated separately. The same approach was employed by Rennie and Millar (2004). However, in both approaches there is a possibility of losing vector interpolation.

Furthermore, According to Rennie and Millar (2004), by utilizing the kriging method, interpolated vector fields that are decomposed in different coordinate systems are different. However, by resolving the velocity vectors into streamwise and cross-stream components, the cross-correlation between the components are minimized and the estimated result improves.

The data produced by ADCP are given in easting and northing coordinate system. The data were resolved to streamwise and cross-stream components by means of Equation 3.22a and 3.22b

$$V_{\text{streamwise}} = V_{\text{easting}} \times \cos\theta + V_{\text{northing}} \times \sin\theta \qquad [3.22a]$$

$$V_{\text{cross-stream}} = V_{\text{northing}} \times \cos\theta - V_{\text{easting}} \times \sin\theta \qquad [3.22b]$$

The required angle θ in above equations is the average of the vector angles which is toward the downstream direction of the river. This angle was calculated separately for the downstream and upstream portions of the measured river reach.

Two separate sets of scalar data ($V_{\text{streamwise}}$ and $V_{\text{cross-stream}}$) were provided for each portion of the field. Variograms were computed by Surfer 8 for each set of scalar data to be interpolated. The produced variograms by Surfer 8 are accessible to be modified and modeled by the best fitted standard mathematical functions. The variogram produced for the streamwise velocity component in the downstream portion was modified to a spherical variogram since the original variogram did not perfectly present the spatial variability of the data. The lag width of 5 m set by the software

was accepted for the computation. For the upstream portion the same variograms with lag distance of 4 m calculated by the software were used. The variograms are shown in Figures 4.1 and 4.2 in the next chapter.

Basically, the aim of interpolation is to bring the irregular measured data in regular grids. It is apparent that in a velocity vector field, higher resolution grids produce more intensive and smooth maps which can clearly show the behaviour of the flow. In the present study, the measured field was interpolated onto 5 m by 5m grids.

The interpolated quantities, which were in streamwise and cross-stream coordinate system, were returned to their original coordinate system by opposite rotation angle used in the first transformation:

$$V_{\text{easting}} = V_{\text{streamwise}} \times \cos(-\theta) + V_{\text{cross-stream}} \times \sin(-\theta) \quad (3.23a)$$

$$V_{\text{northing}} = V_{\text{cross-stream}} \times \cos(-\theta) - V_{\text{streamwise}} \times \sin(-\theta) \quad (3.23b)$$

3.4 Optimal Vector Interpolation

3.4.1 Introduction

Optimal vector interpolation method proposed by Feliks et al. (1996) is presented here. This method also utilizes the Gauss-Markov (GM) theorem. The method was proved by Feliks et al.(1996) to be independent of the choice of the coordinate system. Furthermore, the technique takes into account the cross-covariance between multiple variables. Although vectors are resolved to their components and

interpolated separately, due to utilizing cross-covariance between vector components in the calculation, the method may improve the estimation of the components. In this section, the Gauss-Markov theorem for a scalar field along with the formulae developed with Feliks et al.(1996) for two dimensional vector fields are reviewed.

3.4.1.1 Covariance

Covariance is a sample statistic that indicates the relation between random variables.

$$S = \frac{1}{h-1}\sum_{i=1}^{h}(P_i - \overline{P})(Q_i - \overline{Q}) \qquad [3.24]$$

P_i and Q_i are individual sample values of variables P and Q, which are of size h. $\overline{P}$ and $\overline{Q}$ are the means of P and Q. If large values of P occur with large values of Q the calculated S is positive. On the contrary, S is negative if the large values of P occur simultaneously with small values of Q *(*or *vice versa)*.

3.4.2 Formulae

The GM theorem was first employed in geophysical scalar studies by Gandin (1965) for atmospheric studies and by Bretheron et al. (1976) for oceanic researches. In order to explain the GM theorem in a scalar field, we assume that X and Y are two random vectors of length n and m with zero mean and random measurement error. Here, the term "vector" indicates a set of l one dimensional values rather than vector in the physical sense such as two dimensional velocity vectors. The following equations are the covariance matrices and E is the expectation operator;

$$\sum_{XX} = E[XX^{T}] \qquad [3.25]$$

$$\sum_{XY} = E[XY^{\mathrm{T}}] \qquad [3.26]$$

$$\sum_{YY} = E[YY^{\mathrm{T}}] \qquad [3.27]$$

where T indicates matrix transpose, and (XY^{T}) is the outer product defined as:

$$(XY^{\mathrm{T}})_{ij} = x_i y_j \qquad [3.28]$$

Given m observations of Y which is the actual measured data set, the aim of the method is estimating the interpolated field $\hat{X}$, based on a first guess of X, by the following formula:

$$\hat{X} = A\,Y \qquad [3.29]$$

$\hat{X}$ contains n values and is assumed to be the best linear estimate of X by minimizing the error $[\sum_{i=1}^{N} E(\hat{X}_i - X_i)^2]$, and obtaining the optimal $n \times m$ matrix A. The error matrix is minimized by minimizing the trace of the error covariance matrix

$$\sum_{E} = E[(\hat{X} - X)(\hat{X} - X)^T] \qquad [3.30]$$

If $\hat{X}$ is unbiased, according to the GM theorem $E[\hat{X}] = E[X]$, then A is uniquely defined by:

$$A = \sum_{XY} \sum_{YY}^{-1} \qquad [3.31]$$

When the error covariance matrix is

$$\sum_{E} = \sum_{XX} - \sum_{XY} \sum_{YY}^{-1} \sum_{XY}^{T} \qquad [3.32]$$

3.4.2.1 Application of Gauss-Markov Theorem to Physical Vector Fields

In order to adapt the method in two dimensional vector fields Feliks et al. (1996)

supposed to have m measured two dimensional vectors at m locations for each time step of the time series.

$$Y^{\mathrm{T}} = [(u_1, v_1), (u_2, v_2), ..., (u_{\mathrm{m}}, v_{\mathrm{m}})] \qquad [3.33]$$

where u_1, v_1 are the two components of the vectors in measured field Y

Importantly, Feliks et al. (1996) measured time series of wind vectors at each location, thus Y^{T} was available for each time step in the time series. The purpose of the method is estimating the vectors at n grid locations by means of Equation 3.29 in the two dimensional vector fields at each time step of the time series :

Here:

$$\hat{X}^{\mathrm{T}} = [(\hat{u}_1, \hat{v}_1), (\hat{u}_2, \hat{v}_2), ..., (\hat{u}_{\mathrm{n}}, \hat{v}_{\mathrm{n}})], \qquad [3.34]$$

where $\hat{u}, \hat{v}$ are the two components of the vectors in the estimated field $\hat{X}$.

Matrix A which is a $2n \times 2m$ matrix in two dimensional vector fields instead of $n \times m$ matrix in scalar fields, is defined as best linear estimator by GM theorem in Equation 3.31.

The advantage of this estimator is independency to the choice of the coordinate system. To prove this, Feliks et al. (1996) assumed to have two coordinate systems with two estimators $\hat{X}_1$ and $\hat{X}_2$ in each. Both $\hat{X}_1$ and $\hat{X}_2$ contain physical two dimensional vectors. Based on GM theorem both $X - \hat{X}_1$ and $X - \hat{X}_2$ must yield the minimum error defined by Equation 3.32. Since X is same in both, $\hat{X}_1$ has to be equal to $\hat{X}_2$ in order to yield the unique error.

3.4.3 Optimal Vector Interpolation Procedure

The same downstream and upstream portions of the surveyed field which were interpolated by the kriging method in the previous section were utilized to evaluate the optimal vector interpolation method. Each field had to be detrended by demeaning the measured vectors according to Feliks et al. (1996). Further, an *a priori* data field needed to be estimated and zero-meaned as well.

In the Feliks et al. (1996) study, wind vectors were measured in eleven stations in time series. Data were stationary in each station, and covariance matrices could be calculated from these time series. The optimal vector interpolation method was then used to estimate time series in unmeasured stations. In the present study, data were not measured as time series at fixed stations. Rather, they were asynoptic spatial measurements. A novel approach had to be developed to employ the optimal vector interpolation equations for interpolation of aysnoptic spatial data. The approach which was used to adapt the Feliks et al. (1996) method in spatially measured data fields was to divide the fields (X' and Y') into equal blocks($X' = [X1, X2, ..., Xn]$ and $Y' = [Y1, Y2, ..., Ym]$) and to consider these blocks as measurement stations. In other words, it was assumed that all data in each spatial block were stationary, thus data measured within a block at different times represented samples of a stationary process. Covariance matrices could be calculated from the data sets in each spatial block.

The estimated fields in the downstream and upstream portions by means of kriging

method with grid spacing of 1 m were considered as *a priori* data. The *a priori* field along with the actual field were grided in 5 m by 5m blocks in order to interpolate the field with the same resolution as the interpolated field by kriging. Demeaning both actual and *a priori* data fields was carried out by demeaning the easting and northing velocity components of the vectors over the entire fields X' and Y'.

The process of calculating the covariance matrix of $\sum_{XY} = E[XY^{T}]$ is explained here in detail as an example. Gridded fields into 5 m by 5 m blocks, provided n stations in $X' = [X1, X2, ..., Xn]$ and m stations in $Y' = [Y1, Y2, ..., Ym]$. m and n are equal when both fields are divided into an equal grid spacing. Since the actual field contained unsampled areas and consequently there were no data in some of the grids, the empty grids were excluded from the actual fields and m became different from n. Each 5m by 5m block of data, contained different number of vectors. According to Equation 3.24, Pi and Qi should have the same number of vectors, therefore, only k first vectors in each block were taken into account to calculate the covariance, when k is the minimum available vector in the blocks.

$$X_i^{T} = [V1_{Xi}, V2_{Xi}, ... Vk_{Xi}] \qquad [3.35]$$

$$Y_i^{T} = [V1_{Yi}, V2_{Yi}, ... Vk_{Yi}] \qquad [3.36]$$

where $V1_{Xi}, V2_{Xi}, ... Vk_{Xi}$ are the k first vectors in the i_{th} station (i=1 to n) which are decomposed to their u and v components as below:

$$X_i^{T} = [(u1_{Xi}, v1_{Xi}), (u2_{Xi}, v2_{Xi}), ...(uk_{Xi}, vk_{Xi})] \qquad [3.37]$$

Similar to Equation 3.35, $V1_{Yi}, V2_{Yi}, ... Vk_{Yi}$ in Equation 3.36 are the first k vectors in

the i_{th} station in measured field. The resolved vectors to their u and v components are shown below:

$$Y_i^{T} = [(u1_{Yi}, v1_{Yi}),(u2_{Yi}, v2_{Yi}),....(uk_{Yi}, vk_{Yi})] \quad [3.38]$$

and:

$$\sum_{XY} = \begin{bmatrix} cov(X_1Y_1) & cov(X_1Y_2) & . & . & cov(X_1Y_m) \\ cov(X_2Y_1) & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ cov(X_nY_1) & . & . & . & cov(X_nY_m) \end{bmatrix} \quad [3.39]$$

which is an n by m matrix.

Each component in matrix 3.39 is a 2 by 2 matrix, achieved by Equation 3.40 based on Equation 3.24. Thus, $\sum_{XY}$ becomes a $2n$ by $2m$ matrix. For instance $cov(X_pY_q)$ is calculated as following:

$$S = \frac{1}{k-1}\sum_{i=1}^{k}(Vi_{Xp})(Vi_{Yq})^{T} = \begin{pmatrix} s_u^2 & s_{uv} \\ s_{vu} & s_v^2 \end{pmatrix} \quad [3.40]$$

where:

$$s_u^2 = \frac{1}{k-1}\sum_{i=1}^{k}(ui_{Xp})(ui_{Yq}) \quad [3.41]$$

$$s_{uv} = s_{vu} = \frac{1}{k-1}\sum_{i=1}^{k}(ui_{Xp})(vi_{Yq}) \quad [3.42]$$

$$s_v^2 = \frac{1}{k-1}\sum_{i=1}^{k}(vi_{Xp})(vi_{Yq}) \quad [3.43]$$

It should be noted that demeaning occurred on the whole field, but not within the

spatial blocks, and that within block means were not subtracted when calculating the covariance matrix. It was assumed that by demeaning the whole field, the means within the blocks were almost zero.

The $\sum_{YY}$ matrix in Equation 3.31 needed to be inverted. After computing the matrix $\sum_{YY}$, the determinant of the matrix was equal to zero for both downstream and upstream portions. When the determinant of a matrix is equal to zero, the matrix is called singular and cannot be inverted. Thus the calculation of matrix A and the rest of interpolation procedure is impossible. The singularity of a matrix is due to the big size of the matrix as well as its components. In order to continue the process of interpolation, some of the grids needed to be left out to generate a non singular matrix.

The filtering procedure to achieve a non-singular $\sum_{YY}$ was carried out through an optimization procedure. A number of vectors was chosen as the minimum required vector in each grid. Based on this number, only the grids which contain this amount of vectors were taken into account for generating the covariance $\sum_{YY}$. The determinant of the produced matrix was calculated in the next step and if it was equal to zero, the number of minimum required vector increased. Consequently, the number of remaining grids decreased up to the point that the covariance $\sum_{YY}$ was not singular any longer. This process was done for both the downstream and upstream portions. Note that the process of filtering was done on the original field which was

not zero mean, and demeaning was then done on the selected remaining field. The same chosen grids were utilized to calculate $\sum_{XY}$. It is apparent that a large number of the measured vectors are missing for further calculations.

In a time series situation, the unsampled value in each time is estimated as a combination of *Y*, when *Y* contains the vectors in sampled stations in the specific time, using the coefficient matrix *A* (Equation 3.31). In this study, the average vector in each block in actual field (*Y*) is considered as the representative of each block. One vector is estimated in each block and located in the middle point of the 5m by 5m blocks to make a uniform map.

3.5 Capability of the Kriging and Optimal Vector Interpolation methods in Less Intensive Measured Fields

The ability of two above mentioned interpolation methods in the data fields with smaller number of measured vectors were assessed by removing the vectors gradually from the original measured fields and applying the interpolation methods on the new fields. The evaluation was carried out for kriging method by interpolating 90%, 80%, 70% 60%, 50%, 40%, 30%, 20%, 10%, and 1% of the whole field. For optimal vector interpolation method, data reduction was done by reducing 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 99% of the vectors within the grids which were selected for producing the non singular $\sum_{YY}$ matrix.

It is predicted that by interpolating the fields after reducing the number of measured vectors, less similarity occurs between the original field and the estimated field. Thus, if the calculated correlation coefficient between the original and estimated field by gradually decreasing the number of measured vectors employed in the interpolation procedure diminishes, this will serve as validation that the employed correlation coefficient method is functioning properly.

3.6 Correlation Coefficient Method

In order to asses the validity of the vector interpolation methods, the estimated fields with each of the methods are correlated with the actual data field. The correlation coefficient method offered by Crosby et al (1993) was utilized in this study. The method compares the components of a vector by taking into account their cross-correlation. Thus, the vectors are not decomposed to their magnitude and direction scalars and a single value is calculated as the measure of the level of correlation between the vectors. Besides, unlike the other methods that deal with vectors and are dependent to the choice of the coordinate system, the results of correlation coefficient, calculated by this method do not change by changing the coordinate system.

In following paragraphs the Crosby (1993) vector correlation coefficient method is reviewed. This definition is a generalization of the standard correlation coefficients and deals with both direction and magnitude.

The Crosby vector correlation coefficient is based on the definition given by Hooper (1956) which was developed by Jupp and Mardia (1980):

$$\rho_v^2 = \mathrm{T_r}[(\Sigma_{11})^{-1}\Sigma_{12}(\Sigma_{22})^{-1}\Sigma_{21}] \qquad [3.44]$$

if $W_1 = (u_1, v_1)$ and $W_2 = (u_2, v_2)$ is a pair of vectors, the definition is given by Equations 3.45, 3.46 and 3.47, in term of u an v:

$$\rho_v^2 = f/g \qquad [3.45]$$

where:

$$\begin{aligned} f = {} & \sigma(u_1,u_1)\{\sigma(u_2,u_2)[\sigma(v_1,v_2)]^2 + \sigma(v_2,v_2)[\sigma(v_1,u_2)]^2\} + \sigma(v_1,v_1)\{\sigma(u_2,u_2)[\sigma(u_1,v_2)]^2 \\ & + \sigma(v_2,v_2)[\sigma(u_1,u_2)]^2\} + 2[\sigma(u_1,v_1)\sigma(u_1,v_2)\sigma(v_1,u_2)\sigma(u_2,v_2)] \\ & + 2[\sigma(u_1,v_1)\sigma(u_1,u_2)\sigma(v_1,v_2)\sigma(u_2,v_2)] - 2[\sigma(u_1,u_1)\sigma(v_1,u_2)\sigma(v_1,v_2)\sigma(u_2,v_2)] \\ & - 2[\sigma(v_1,v_1)\sigma(u_1,u_2)\sigma(u_1,v_2)\sigma(u_2,v_2)] - 2[\sigma(u_2,u_2)\sigma(u_1,v_1)\sigma(u_1,v_2)\sigma(v_1,v_2)] \\ & - 2[\sigma(v_2,v_2)\sigma(u_1,v_1)\sigma(u_1,u_2)\sigma(v_1,u_2)] \end{aligned} \qquad [3.46]$$

and

$$g = \{\sigma(u_1,u_1)\sigma(v_1,v_1) - [\sigma(u_1,v_1)]^2\}\{\sigma(u_2,u_2)\sigma(v_2,v_2) - [\sigma(u_2,v_2)]^2\} \qquad [3.47]$$

The conventional scalar correlation coefficient is normalized to have a value between 0.0 and 1.0. Jupp and Mardia did not divide the right-hand side of Equation 3.44 by a constant to make sure that ρ_v^2 is between 0.0 and 1.0. In the two dimensional case ρ_v^2 varies from 0.0 to 2.0. Note that both vector fields must be sampled at the same locations. Significance testing for sufficiently large samples is based on the χ^2 distribution, with a 95% confidence. Significant correlation is equal to $\rho_v^2 > 9.488/n$, when n is the number of vectors in a vector field. In general, large vector fields require very small correlation to be significantly correlated.

3.6.1 Correlation Coefficient Procedure

Estimated downstream and upstream vector fields which were interpolated by means of kriging and optimal vector interpolation methods were correlated with the related actual data fields to validate the strategies of the interpolation. Raw data fields are vector fields measured irregularly. Since in the correlation coefficient ρ_v^2 method, each vector in one field is compared to the related vector in other field, irregularly measured raw data fields had to be block averaged in 5 m by 5 m blocks. In addition, due to the measurement method, some of the blocks do not contain data and required to be excluded from calculation in both measured and interpolated fields. Interpolated fields after data removal are correlated with the same spatially block averaged raw data field in order to study of the influence of gradually reducing actual data field on the interpolation methods.

4 RESULTS AND DISCUSSION

Kriging and optimal vector interpolation methods were employed to estimate the uniform velocity vector maps. In this chapter, the validation of each method is assessed by comparing the resultant fields and actual data using Crosby vector correlation coefficient method. In addition, the capability of the methods in estimating less intensive measured field is reviewed. Also the influence of rotating the coordinate system on the estimated fields for each method is evaluated.

4.1 Kriging Interpolation Method

The velocity vector field (Figure 3.8) measured by means of ADCP was separated into upstream and downstream portion of the bridge (Figure 3.9a and 3.9b). The kriging technique which is an interpolation method for scalar quantities was employed on velocity vector field by interpolating each velocity vector component separately. In order to eliminate the effect of the cross-correlation between the decomposed quantities, they were rotated to streamwise and cross-stream coordinate system. The variograms which were modeled based on experimental variograms for each component of the portions of the field are shown in Figures 4.1 and 4.2.

4.1.1 Cross Correlation

Since the kriging method is only able to interpolate the scalars, the velocity vectors needed to be decomposed to their components before interpolation. In this way, each component is interpolated independently and any association between them is ignored. Due to this reason, vector properties are missing in this method. In order to prevent this problem to occur, the vectors were rotated to a coordinate system in which the cross correlation between the vector components is zero or very small.

The cross correlation between the vector components to be interpolated was diminished by rotating the easting-northing coordinate system to streamwise-cross-stream coordinates. In the downstream portion, the cross correlation of the components in original coordinate decreased from ρ= -0.24 to ρ= 0.03 in the rotated coordinates. Likewise, in upstream portion, the cross correlation decreased from ρ= -0.08 in easting-northing decomposition to ρ=-0.03 in streamwise-cross-stream decomposition. The results show that the cross correlation is eliminated successfully to an insignificant amount for both downstream and upstream portions.

4.1.2 Variograms

Figures 4.1a and 4.1b do not display spatial structure in correlation for the upstream portion in the original coordinate system, and consequently resolving the velocities to streamwise and cross-stream components does not change the spatial correlation very much. (Figures 4.1c and 4.1d).

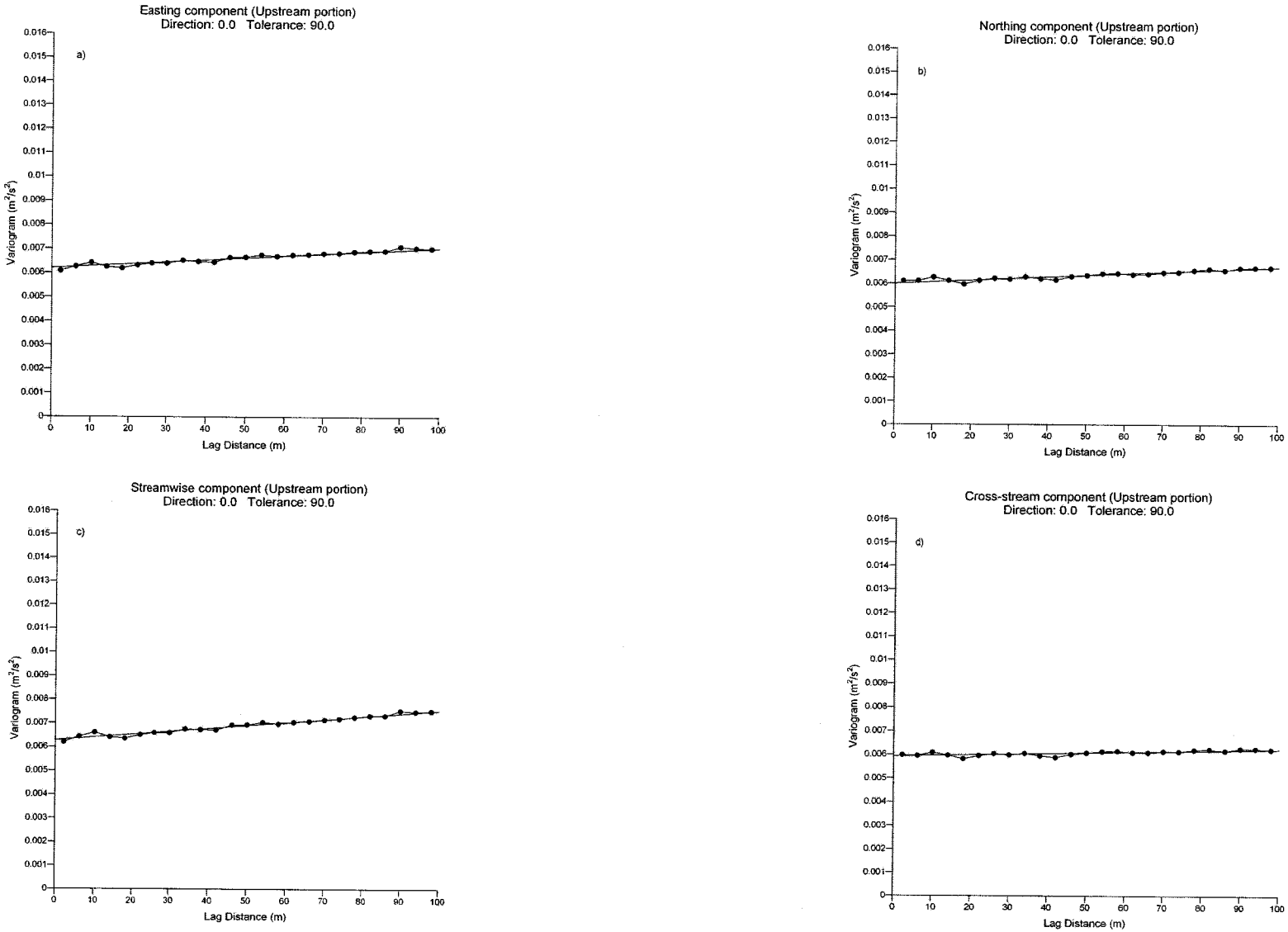

Figure 4. 1 Vector components variograms for upstream portion

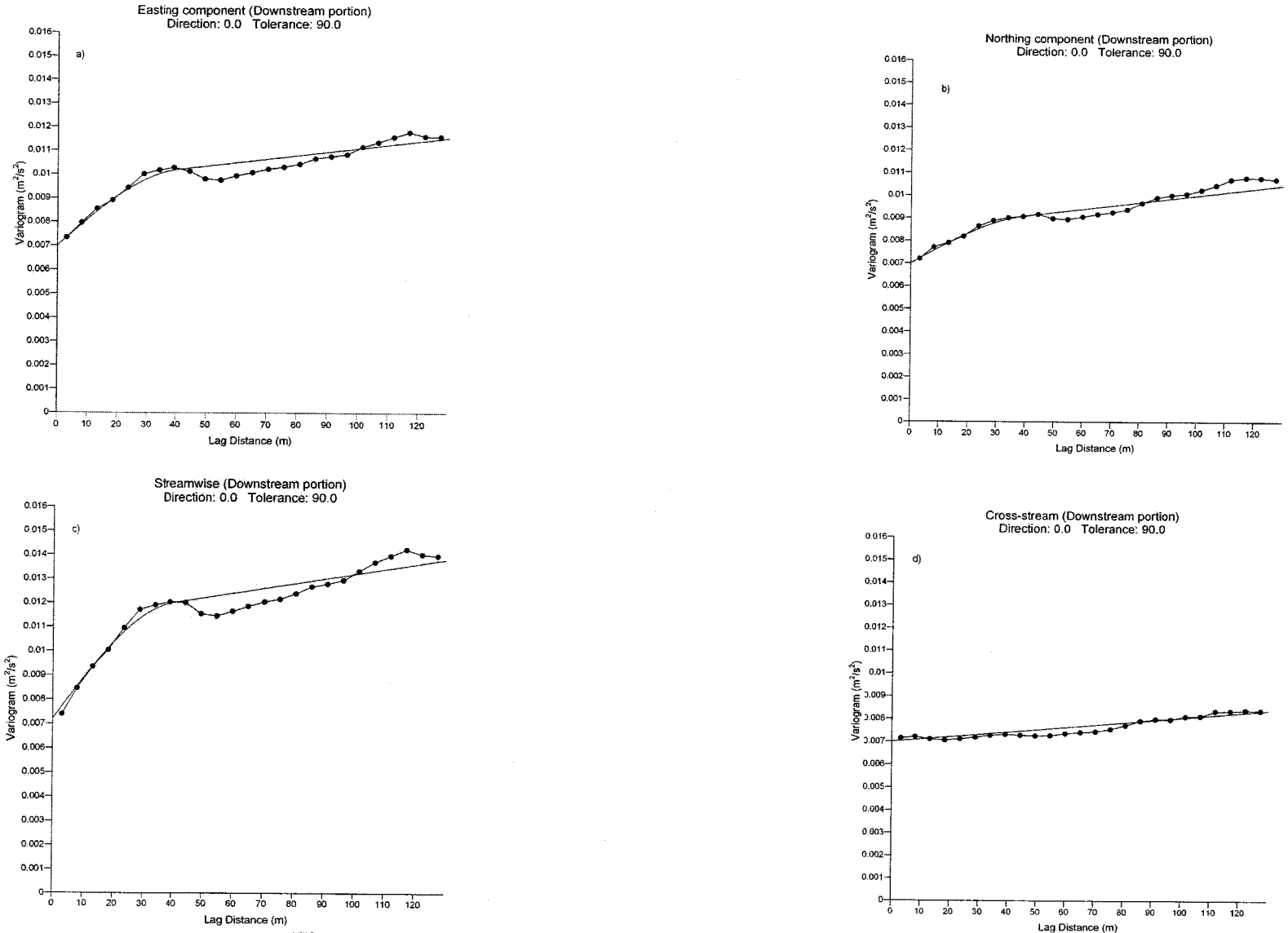

Figure4.2 Vector components variograms for downstream portion

In the downstream portion, variograms indicate a spatial correlation up to the spatial lag of 40 m for data in easting and northing coordinate system. (Figures 4.2a and 4.2b). After decomposing the vectors to streamwise and cross-stream coordinates the spatial correlation of the data decreased for the streamwise component (Figure 4.2c). However, as it can be seen in Figure 4.2d, the spatial correlation increased for the cross-stream component.

The streamwise component variograms for upstream and downstream portions have ordinate intercepts ("nuggets") of 0.006 m^2/s^2 and 0.007 m^2/s^2 respectively. The nugget indicates the amount of variance due to sampling error and due to true variability at scales smaller than the smallest lag distance. Assuming similar sampling errors in the upstream and downstream portions, it appears that closely situated vectors were slightly more similar in the upstream portion than the downstream portion, as expected for a more uniform flow field. More importantly, the variogram for the downstream portion has a range of influence of about 40 m with a sill of approximately 0.012 m^2/s^2, whereas no sill is apparent in the upstream portion and the variogram is still 0.007 at large lags. In other words, in the upstream portion the flow is similar throughout the field, and depth averaged vectors are similar even if separated by 100 m. In the downstream portion, on the other hand, vectors spaced far apart are less similar. This is a result of the influence of the bridge piers, with vortex streets creating a non-uniform flow field.

4.1.3 Interpolation

The velocity components were interpolated in 5m by 5m spacing grids in streamwise and cross-stream coordinate and rotated back to easting and northing components. The vector maps of the interpolated raw data fields for upstream and downstream portions of the field are shown in Figures 4.3 and 4.4. with contours showing the bathymetry of the river Estimated vectors vary from 0.102 m/s to 0.295 in the upstream and 0.014 m/s to 0.415 m/s in the downstream of the bridge.

The vector plot of the upstream portion presents a more uniform flow (Figure 4.3). Alternatively, as it was expected, resultant flow field for the downstream area of the field shows a more variable flow (Figure 4.4). The pattern of the vortices due to the bridge piers is observed around the locations of the piers downstream of the bridge. These vortices are the consequence of flow separation at the sides of the piers, and are described as a von Karman street (Mellville and Coleman 1999). The vortices are the origins of the significant changes to the flow pattern. These changes show themselves by generating a more variable flow in downstream portion of the bridge compared to the flow in the upstream portion.

In order to evaluate the capability of the kriging interpolation method in less intensive measured fields, the procedure was repeated for vector fields by reducing the number of velocity vectors to 90%, 80%, 70%,...10% and finally only 1% of the measured vectors and interpolating the reduced fields using the same variograms modeled for

the whole data fields. Maps of Interpolated raw data fields are compared with interpolated fields after 10% and 90% data reduction in Figures 4.5 and 4.6 for upstream portion and Figures 4.7 and 4.8 for downstream portion.

The vectors are interpolated vectors in every 5m by 5m grid. Superimposed maps of interpolated whole actual field and interpolated field after 10 percent data reduction demonstrate a very high similarity between interpolated data both in downstream and upstream portions (Figure 4.5 and 4.7). In contrast, comparing the same entire data interpolated maps with related interpolated fields after removing 90 percent of the measured data shows relatively smaller similarity between pairs of vectors both in upstream and downstream parts. (Figures 4.6 and 4.8).

After reducing 10% and 90% of the whole data in upstream portion, the velocity ranges from 0.102m/s to 0.299m/s and 0.147m/s to 0.294m/s respectively. In downstream portion the velocity vector values show higher variability. Interpolation of 90% and 10% of the measured data in the downstream portion estimated velocity vectors from 0.011 to 0.413 m/s and 0.006 and 0.442 m/s respectively.

4.2 Optimal Vector Interpolation Method

Optimal vector interpolation method was applied to the same measured fields in the downstream and upstream portions. Contrary to kriging method which interpolates

the scalars separately, the optimal method is capable of interpolating the two dimensional vector quantities.

Data were interpolated every 5m in both easting and northing directions. Interpolated downstream and upstream portions within 1m by 1m grids by kriging method were utilized as *a priori* data. The same set of plots as in the previous section is presented in this section. Figures 4.9 and 4.10 present the whole interpolated fields in upstream and downstream portions. Similar to the resultant interpolated maps in the kriging method, the flow shows more variability in the downstream portion as compared to the upstream portion due to the existence of the bridge piers. The velocity varies from 0.010 m/s to 0.443 m/s for downstream and 0.113 m/s to 0.282 m/s in upstream portion for the interpolated raw data fields. After 10% data reduction within the selected blocks of measured fields, the estimated velocity ranges from 0.139 m/s to 0.296 m/s and 0.009 m/s to 0.455 m/s in the upstream and downstream portions, respectively (Figures 4.11 and 4.13).

The maximum amount of data reduction in which the covariance matrix of $\sum_{YY}$ was not singular was 30 percent and 50 percent for upstream and downstream portions, respectively. Accordingly, the optimal interpolation procedure was not attempted for greater amounts of data reduction. Figures 4.12 and 4.14 present the estimated velocity vectors which range from 0.036 m/s to 1.701 m/s in upstream and 0.010 m/s to 2.351 m/s in downstream part.

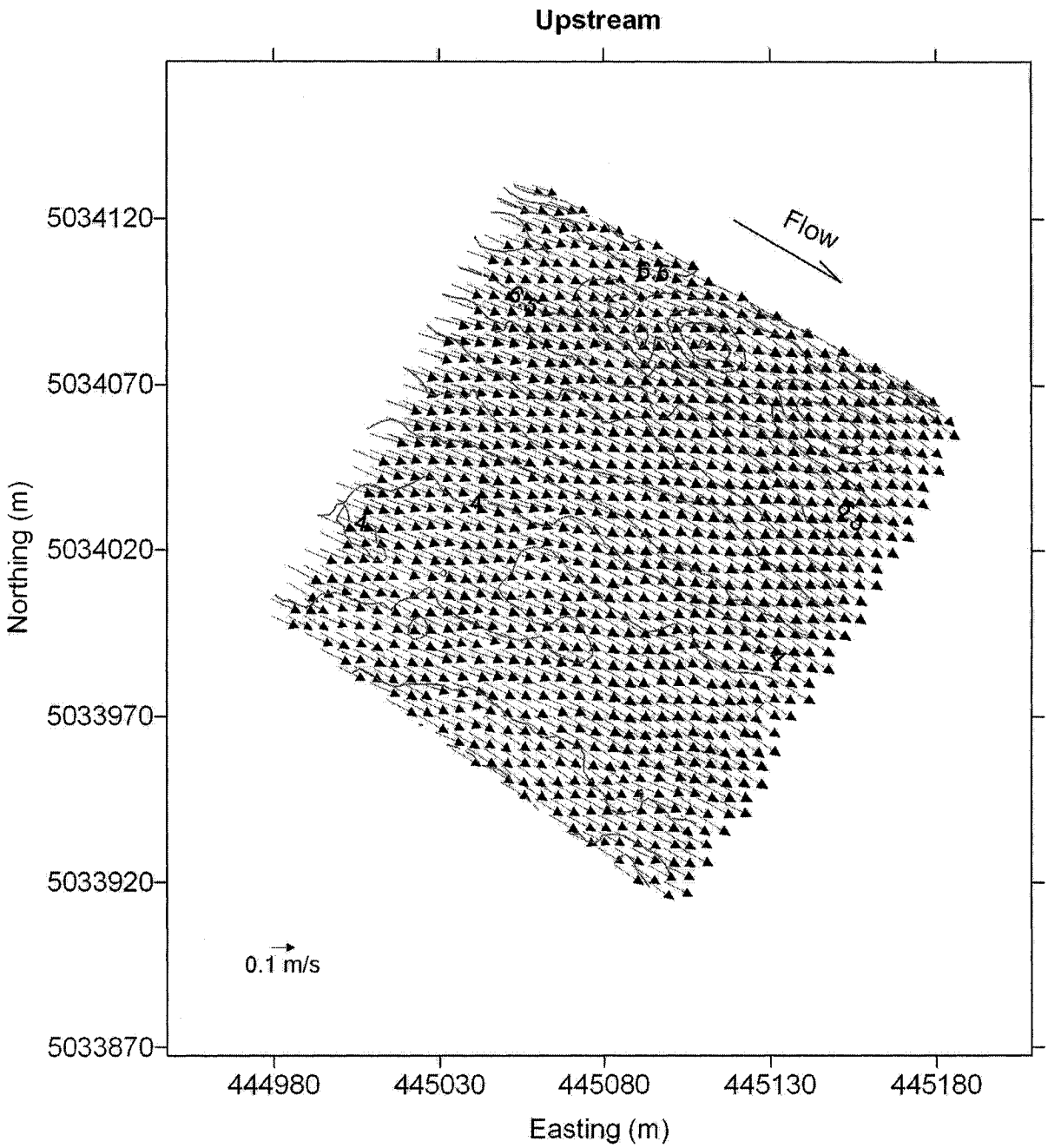

Figure 4. 3 Vector plot of the upstream portion interpolated by kriging method. Vector magnitudes range from 0.102 m/s -0.295 m/s. Bathymetry contours in m based on depths on day of survey.

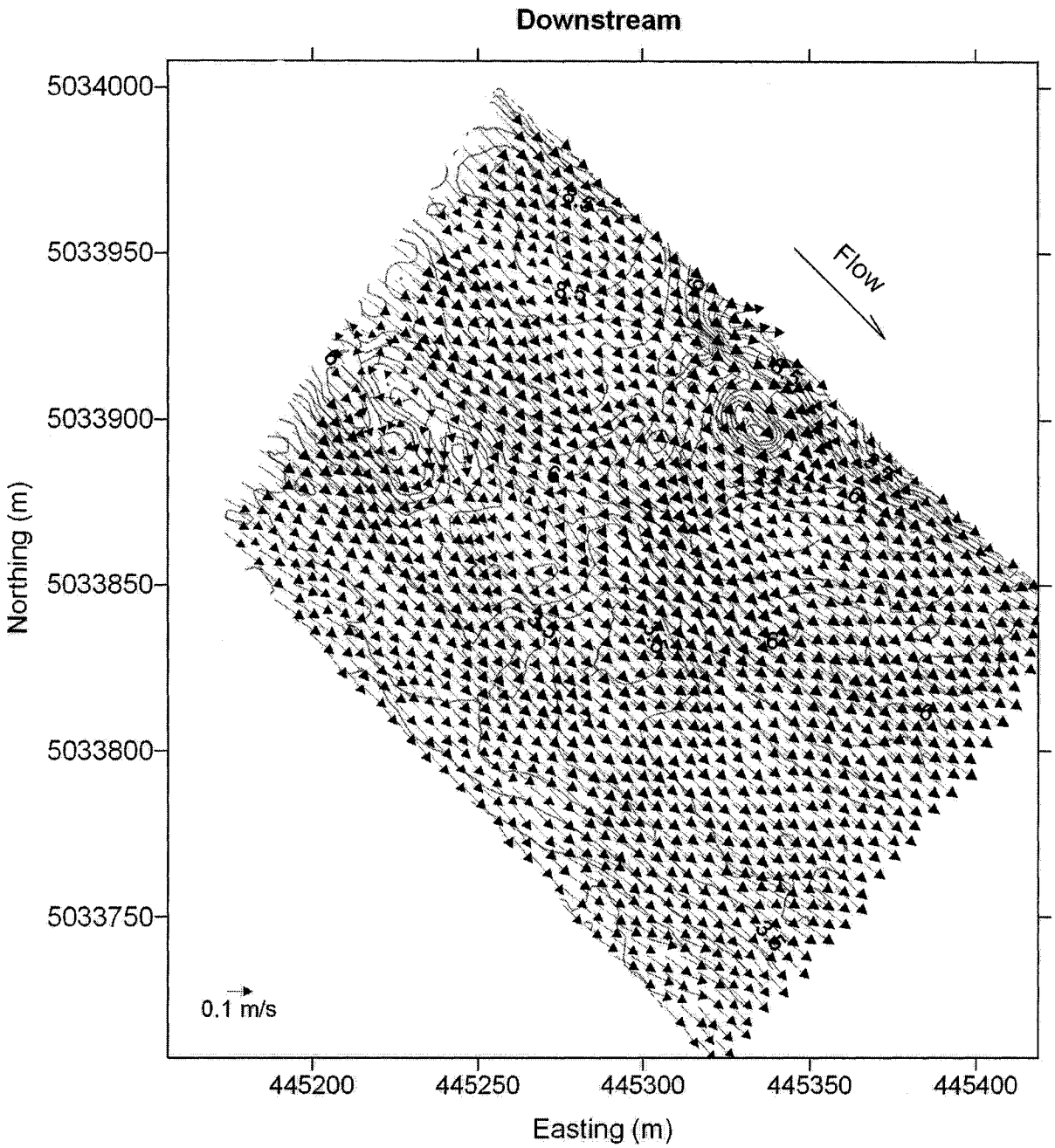

Figure 4. 4 Vector plot of the downstream portion interpolated by kriging method. Vector and spatial scaling equal to Figure 4.3. Vector magnitudes range from 0.014 m/s -0.415 m/s. Bathymetry contours in m based on ADCP measured depths on day of survey.

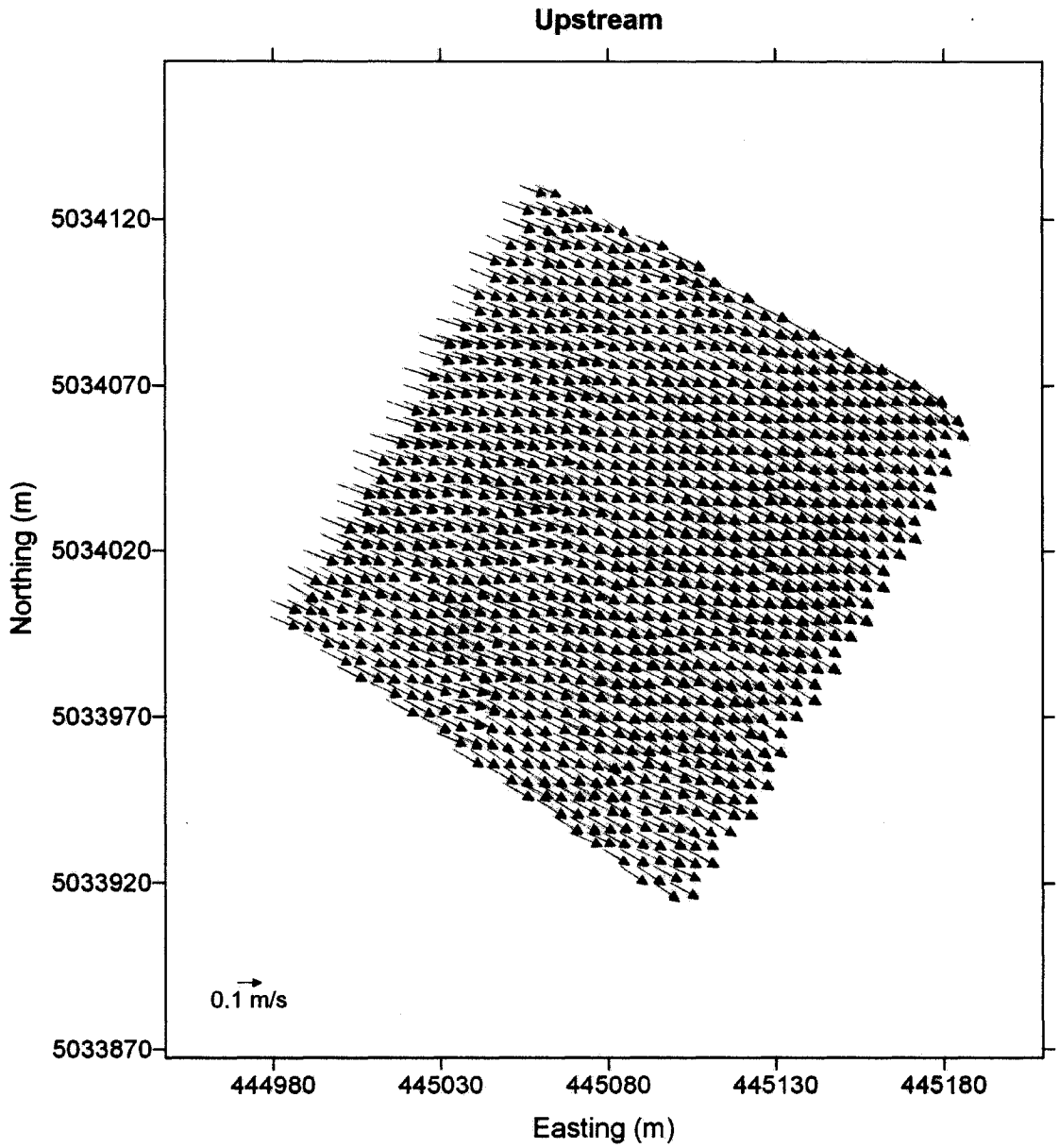

Figure 4. 5 Vector plot of the upstream of the bridge estimated by kriging method; interpolated actual data field (red, 0.102 m/s- 0.295 m/s) and interpolated field with 10 percent of data removal (black, 0.102 m/s- 0.299 m/s). Vector and spatial scaling equal to Figure 4.3.

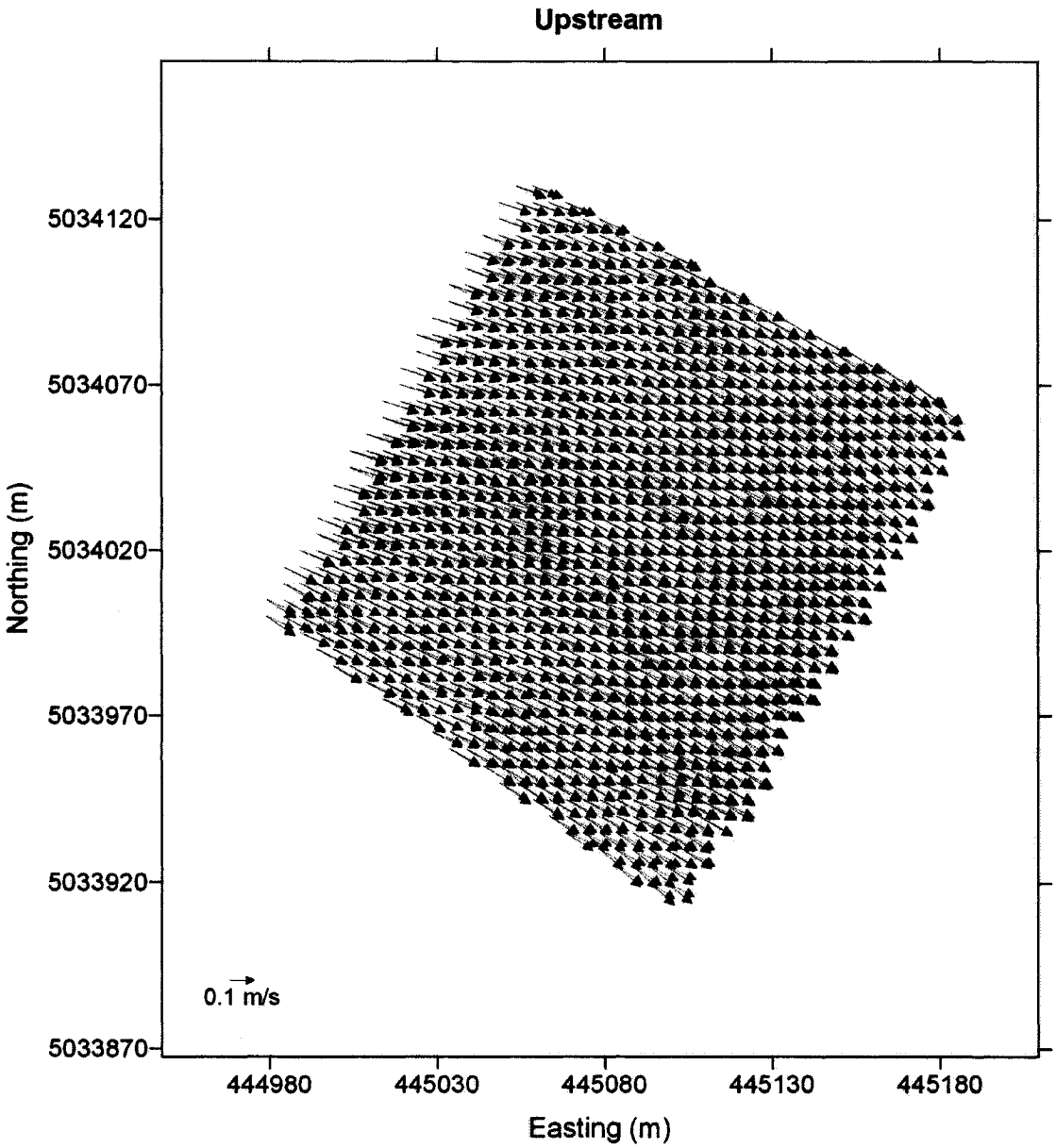

Figure 4. 6 Vector plot of the upstream of the bridge estimated by kriging method; interpolated actual data field (red, 0.102 m/s- 0.295 m/s) and interpolated field with 90 percent of data removal (black, 0.147 m/s- 0.294). Vector and spatial scaling equal to Figure 4.3.

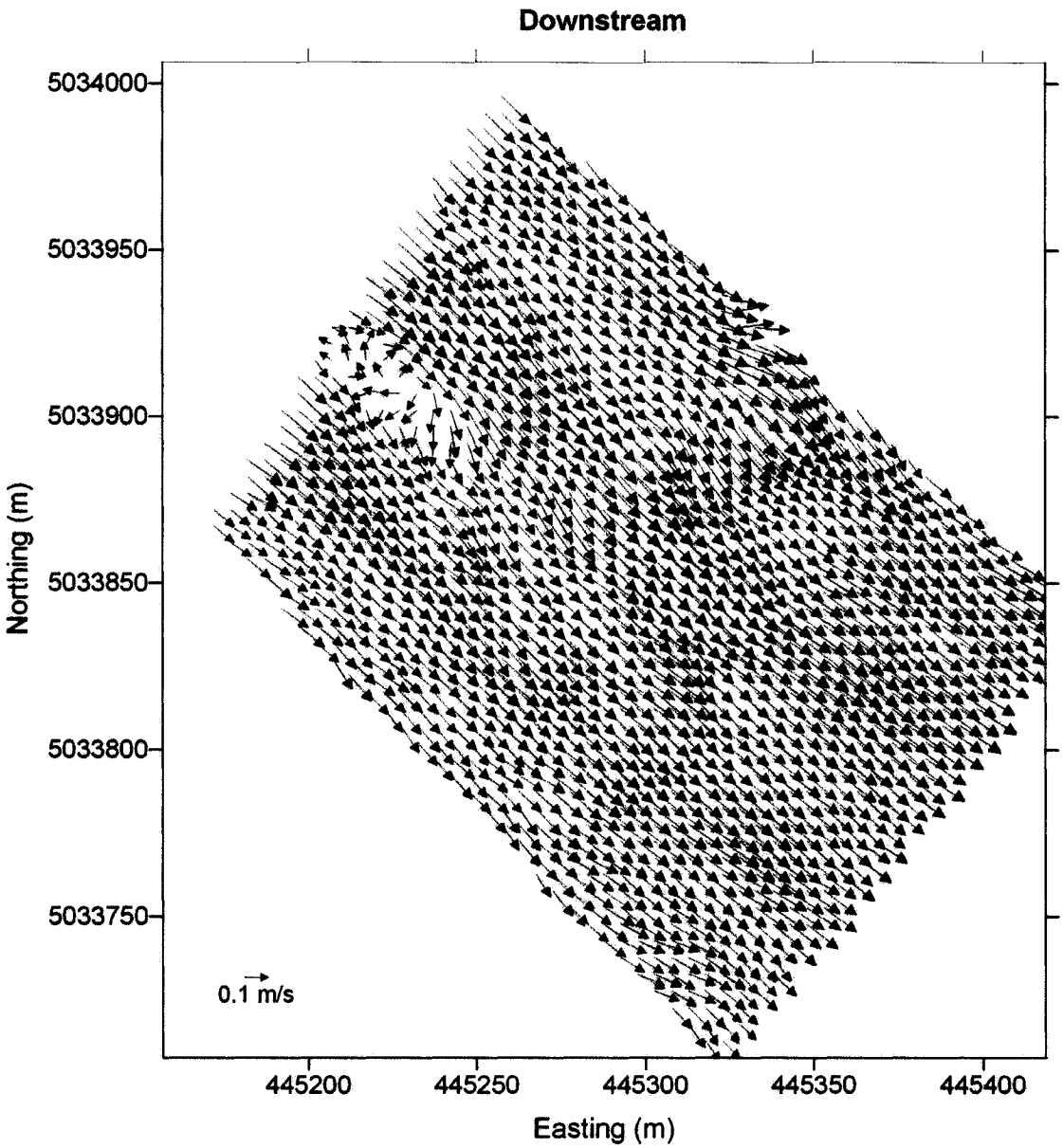

Figure 4. 7 Vector plot of the downstream of the bridge estimated by kriging method; interpolated actual data field (red, 0.014 m/s- 0.415 m/s) and interpolated field with 10 percent of data removal (black, 0.011 m/s- 0.413 m/s). Vector and spatial scaling equal to Figure 4.3.

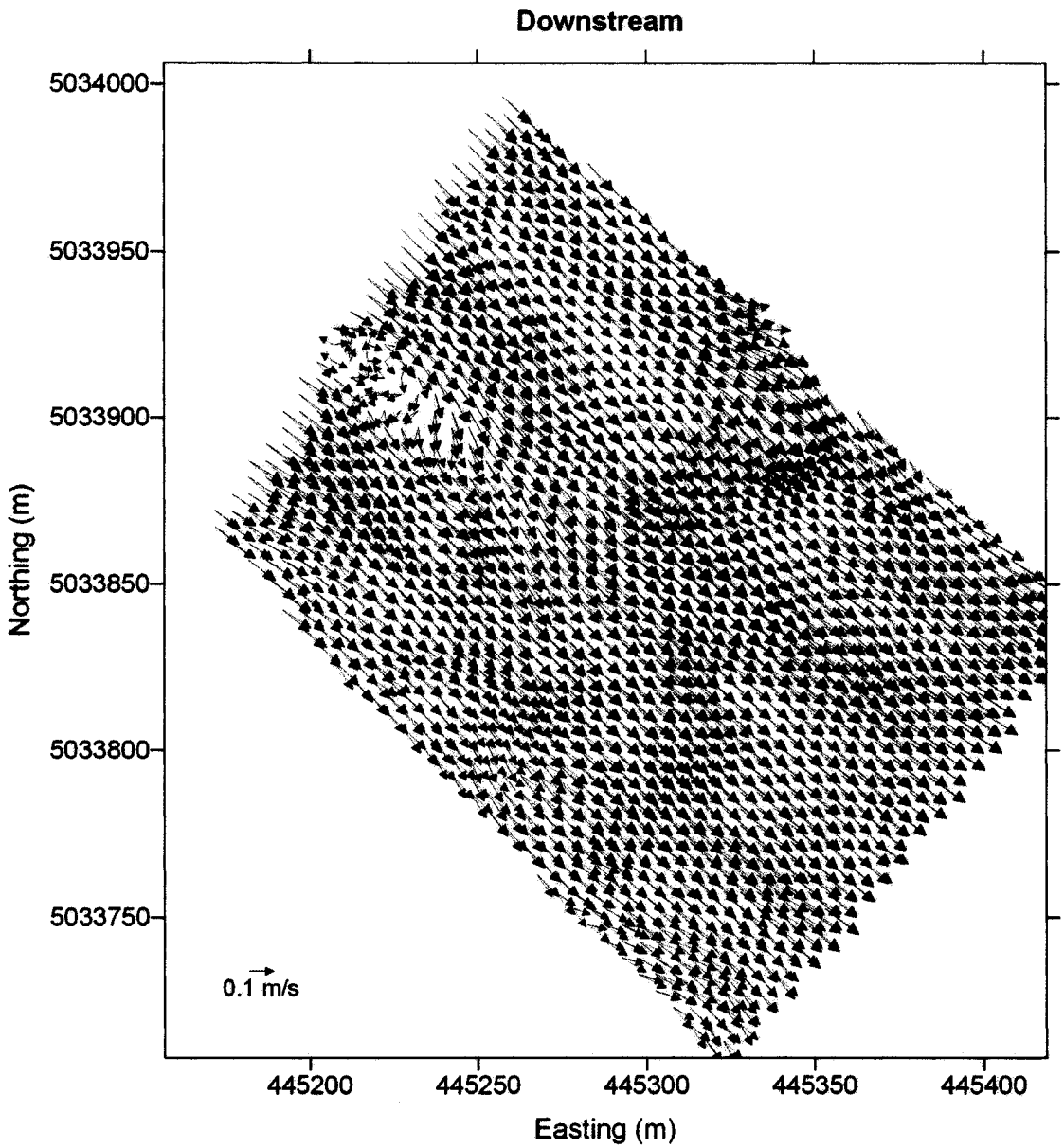

Figure 4. 8 Vector plot of the downstream of the bridge estimated by kriging method; interpolated actual data field (red, 0.014 m/s- 0.415 m/s) and interpolated field with 90 percent of data removal (black, 0.006 m/s- 0.442 m/s). Vector and spatial scaling equal to Figure 4.3.

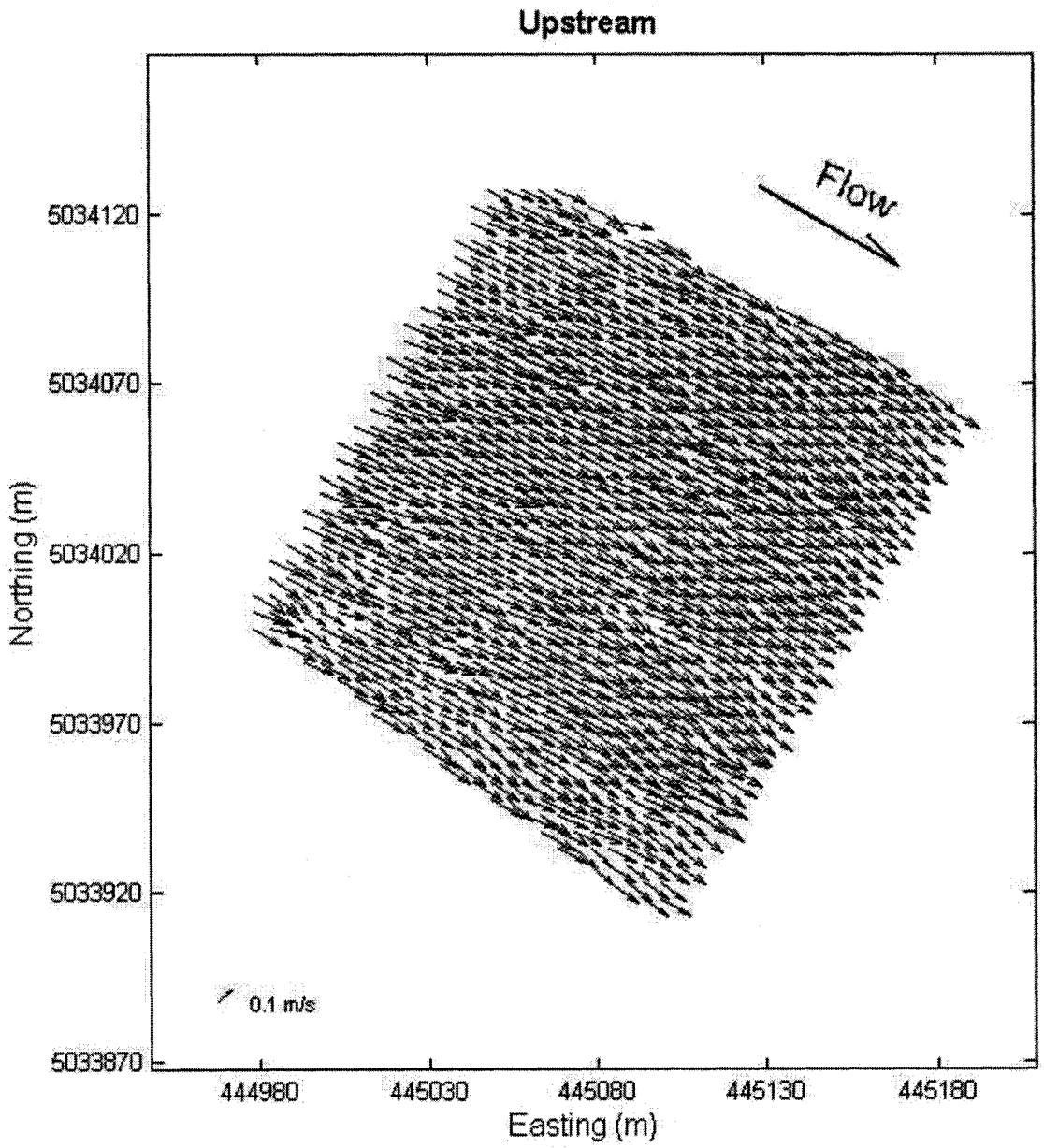

Figure 4. 9 Vector plot of the upstream portion interpolated by optimal vector interpolation method. Vector magnitudes range from 0.113 m/s to 0.282 m/s.

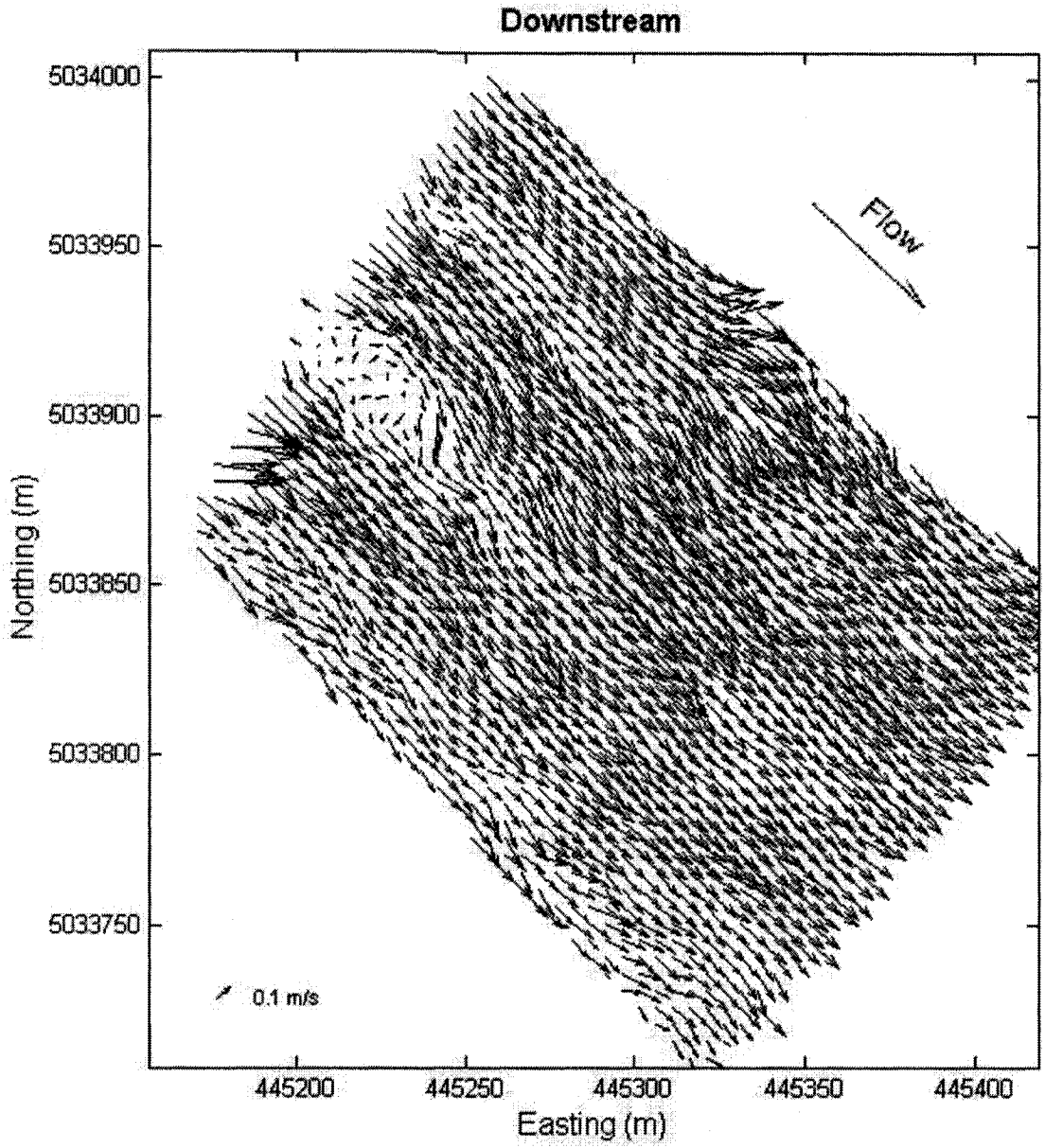

Figure 4. 10 Vector plot of the downstream portion interpolated by optimal vector interpolation method. Vector magnitudes range from 0.010 m/s to 0.443 m/s.

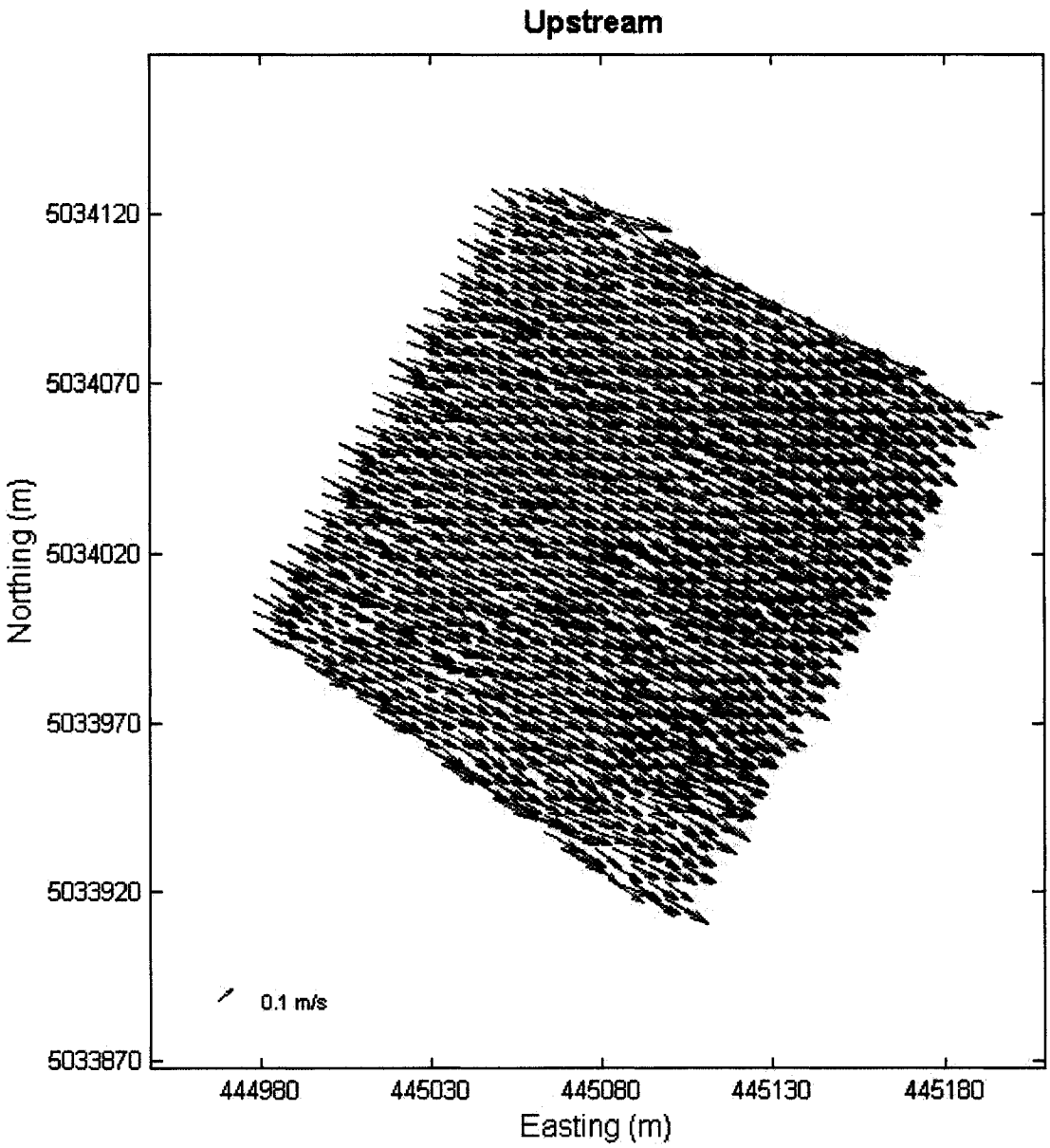

Figure 4. 11 Vector plot of the upstream of the bridge estimated by optimal vector interpolation method; interpolated actual data field (red, 0.113 m/s- 0.282 m/s) and interpolated field with 10 percent of data removal (black, 0.139 m/s- 0.296 m/s). Vector fields are equally scaled.

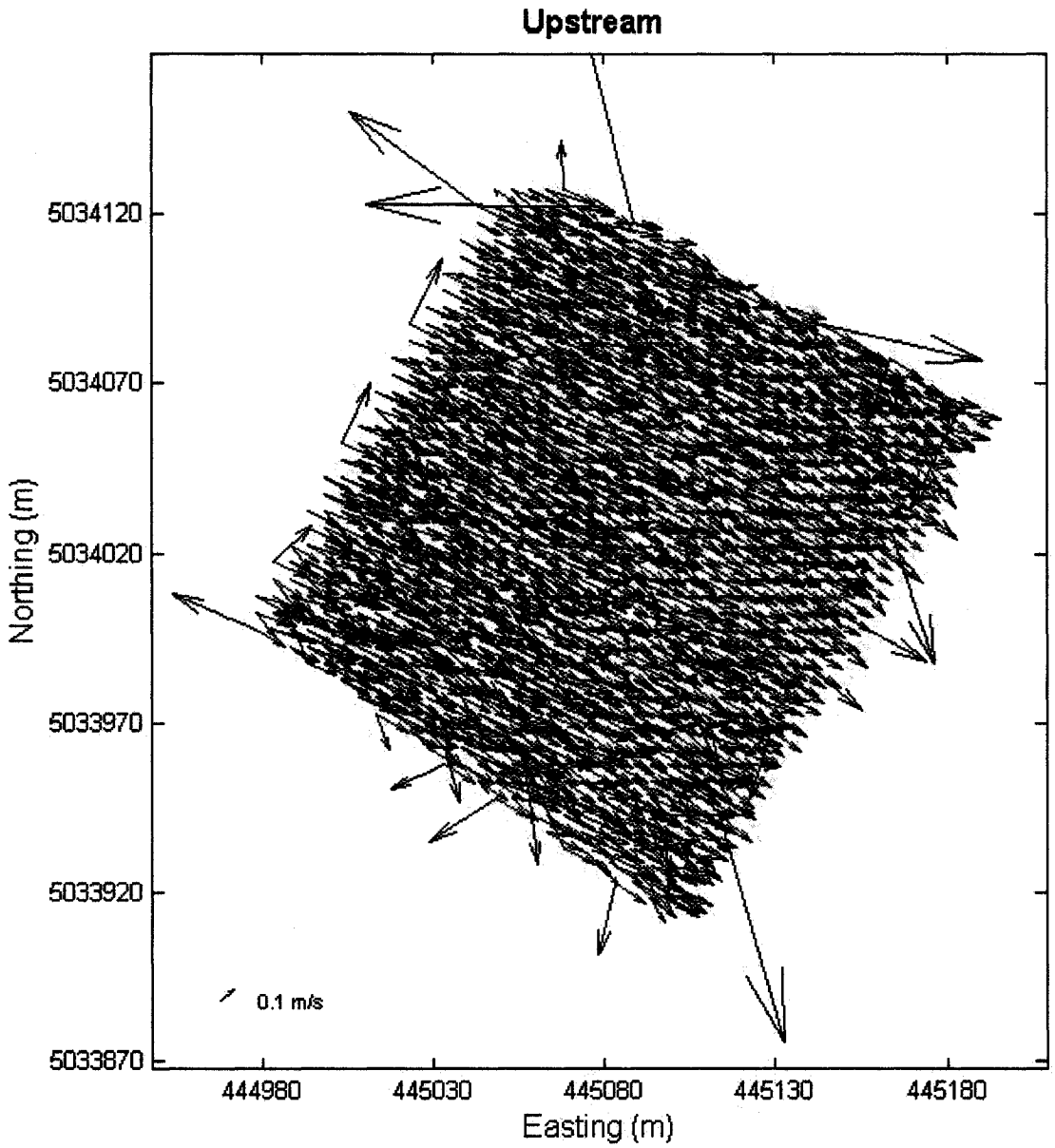

Figure 4. 12 Vector plot of the upstream of the bridge estimated by optimal vector interpolation method; interpolated actual data field (red, 0.113 m/s- 0.282 m/s) and interpolated field with 30 percent of data removal (black, 0.036 m/s- 1.701 m/s). Vector fields are equally scaled.

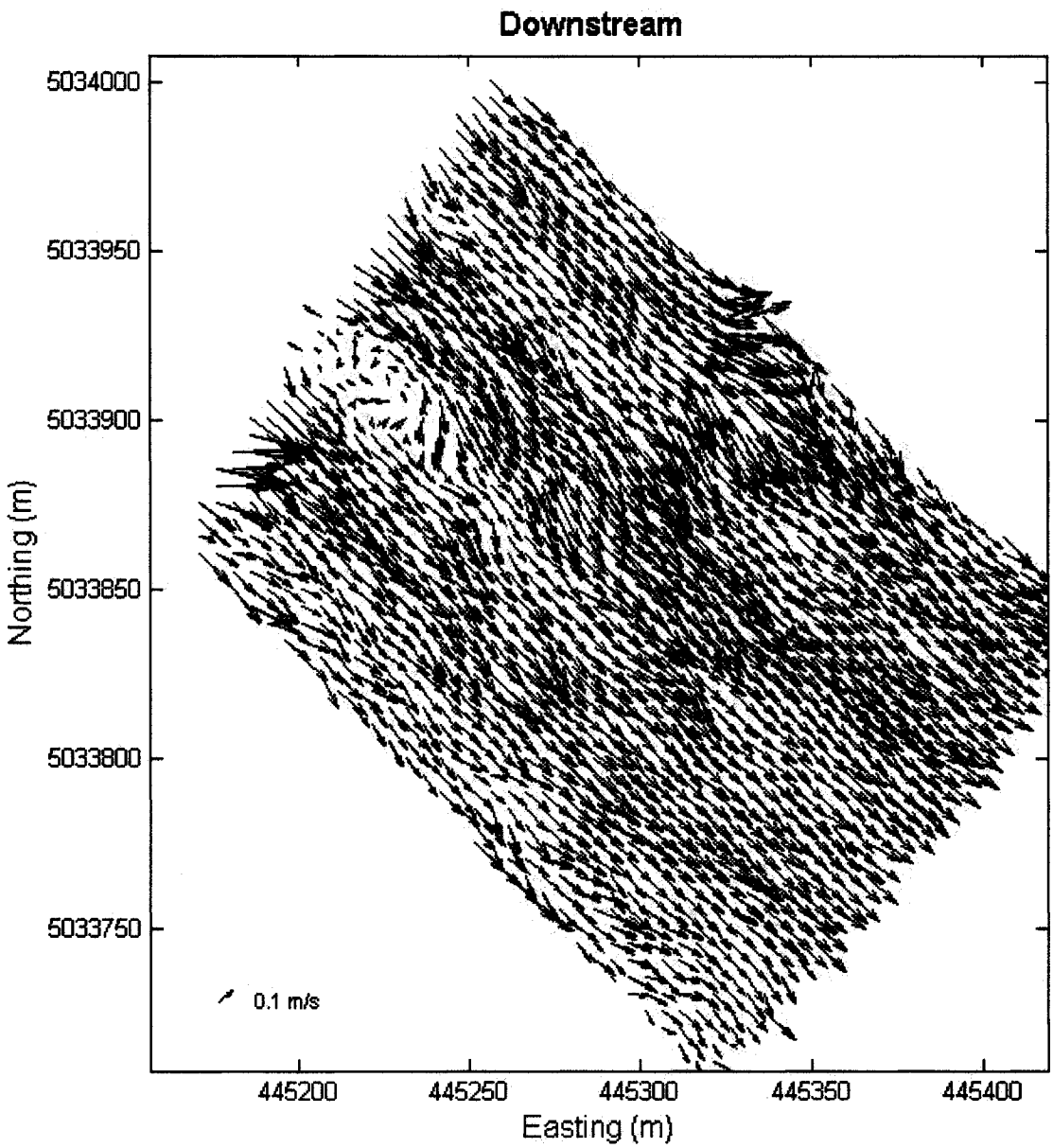

Figure 4. 13 Vector plot of the downstream of the bridge estimated by optimal vector interpolation method; interpolated actual data field (red, 0.010 m/s- 0.443 m/s) and interpolated field with 10 percent of data removal (black, 0.009 m/s- 0.455 m/s). Vector fields are equally scaled.

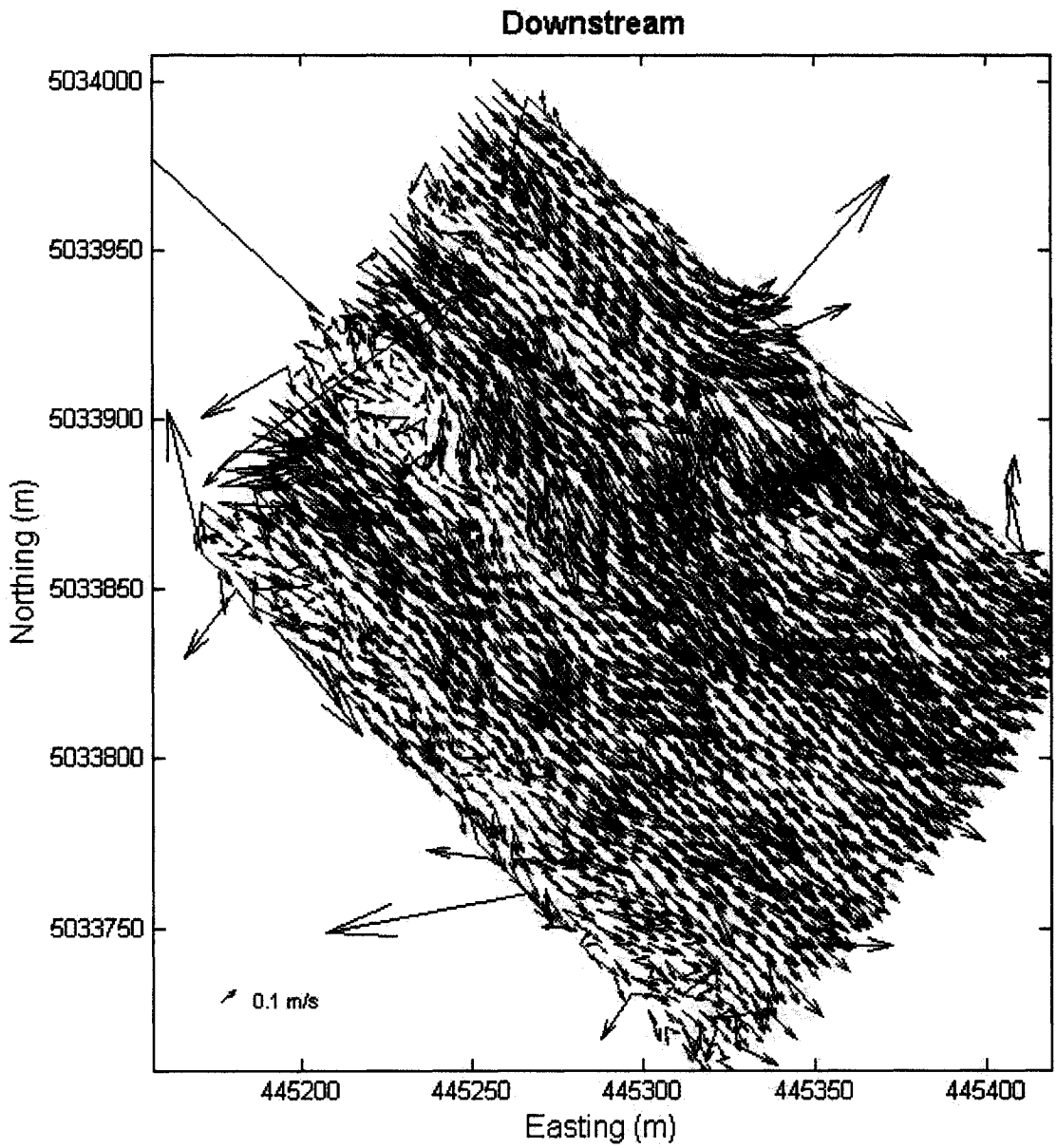

Figure 4. 14 Vector plot of the downstream of the bridge estimated by optimal vector interpolation method; interpolated actual data field (red, 0.010 m/s- 0.443 m/s) and interpolated field with 50 percent of data removal (black, 0.010m/s- 2.351m/s). Vector fields are equally scaled.

4.3 Vector Correlation Coefficient

4.3.1 Interpretation of the Correlation Coefficient

The correlation between the 5 m by 5 m interpolated field (by kriging or optimal vector interpolation method) with spatially block averaged measured fields in 5 m spacing blocks was estimated using the Crosby vector correlation coefficient method.

Before assessing the results of applying the Crosby correlation coefficient method in this study, two important points should be considered about the Crosby correlation coefficient. First, the Crosby correlation coefficient asymptotically follows a Chi-square distribution with four degrees of freedom, from which significance tests can be derived (Crosby et al. 1993). Second, the Crosby correlation coefficient is only a measure of the linear interdependence between two variables rather than a precise correlation coefficient value. It is due to the complexity of the joint variability of the variables to be correlated that cannot be summarized in one parameter. This comment should be considered even more strongly for the association between the vectors. To make the above comment more clear, it is worth considering situations that lead to a perfect vector correlation (ρ_v^2=2). There are four cases that result in a perfect correlation coefficient: when the pairs are the same, when the magnitude of the vectors in the first vector set are multiplied by a constant and used as the second set, when the directions of the vectors in the first set are rotated by a constant angle, and when the second set is obtained by combining a constant change in both magnitude and direction. These cases include all the situations in which one vector set is a linear combination of the other.

It is worth noting that the test of interpolation was conducted by correlating the interpolated field to the observed actual field. Thus, the interpolation was only tested where observed vectors were available. Unavoidably, interpolation in the unmeasured portion of the field was not verified.

For the downstream portion of the field, kriging and optimal vector interpolation methods compared to the related block averaged measured field and respectively resulted in correlation coefficients of ρ_v^2=1.37 and ρ_v^2==1.51. The superimposed maps for this comparison are presented in Figures 4.15 and 4.16 respectively. The same comparison was carried out for the portion upstream of the bridge with less variable flow and surprisingly the calculated ρ_v^2 was less than for downstream with more variable flow. The correlation coefficient was ρ_v^2=0.88 for the kriging method and ρ_v^2=1.06 for the optimal method (Figures 4.17 and 4.18 respectively). All calculated Crosby correlation coefficients are significant correlations when compared to the critical correlation coefficient of 0.017. Figures 4.19 to 4.22 compare the raw data field with interpolated field after 30 percent data reduction. The interpolated fields with 30% data removal differ more from the measured field than when all the survey data are utilized for interpolation (compare Figures 4.19-4.22 to Figures 4.15-4.18).

Crosby correlation coefficients between spatially block averaged fields and interpolated fields by the two different interpolation methods were tested to assess if

the results were significantly different. The testing procedure (Zar 1996) employs the statistic:

$$Z = \frac{z_1 - z_2}{\sigma_{z_1 - z_2}}, \qquad [4.1]$$

where z_1 and z_2 are calculated from Equation 4.2 for r_1 and r_2 correlation coefficients :

$$z = 0.5 \ln\left(\frac{1+r}{1-r}\right) \qquad [4.2]$$

where z was proposed by Fisher to normalize the distribution of r, and

$$\sigma_{z_1 - z_2} = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}} \qquad [4.3]$$

Where n_1 and n_2 are the size of each sample. Calculated Z is compared with the critical value of Z which is $Z_{0.05(2)} = 1.960$. In order to utilize this test, which is based on the conventional scalar correlation coefficient (r), all values of ρ_v^2 were normalized by dividing by two.

Computed Z for downstream and upstream portions were Z=2.448 and Z=1.967 respectively which were compared with $Z_{0.05(2)} = 1.960$ and verified that optimal vector interpolation method is significantly (α=0.05) more successful in interpolating the vector field than kriging method, especially in the downstream portion.

The correlation coefficient method was also applied on interpolated fields after randomly removing 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% , and 99% of

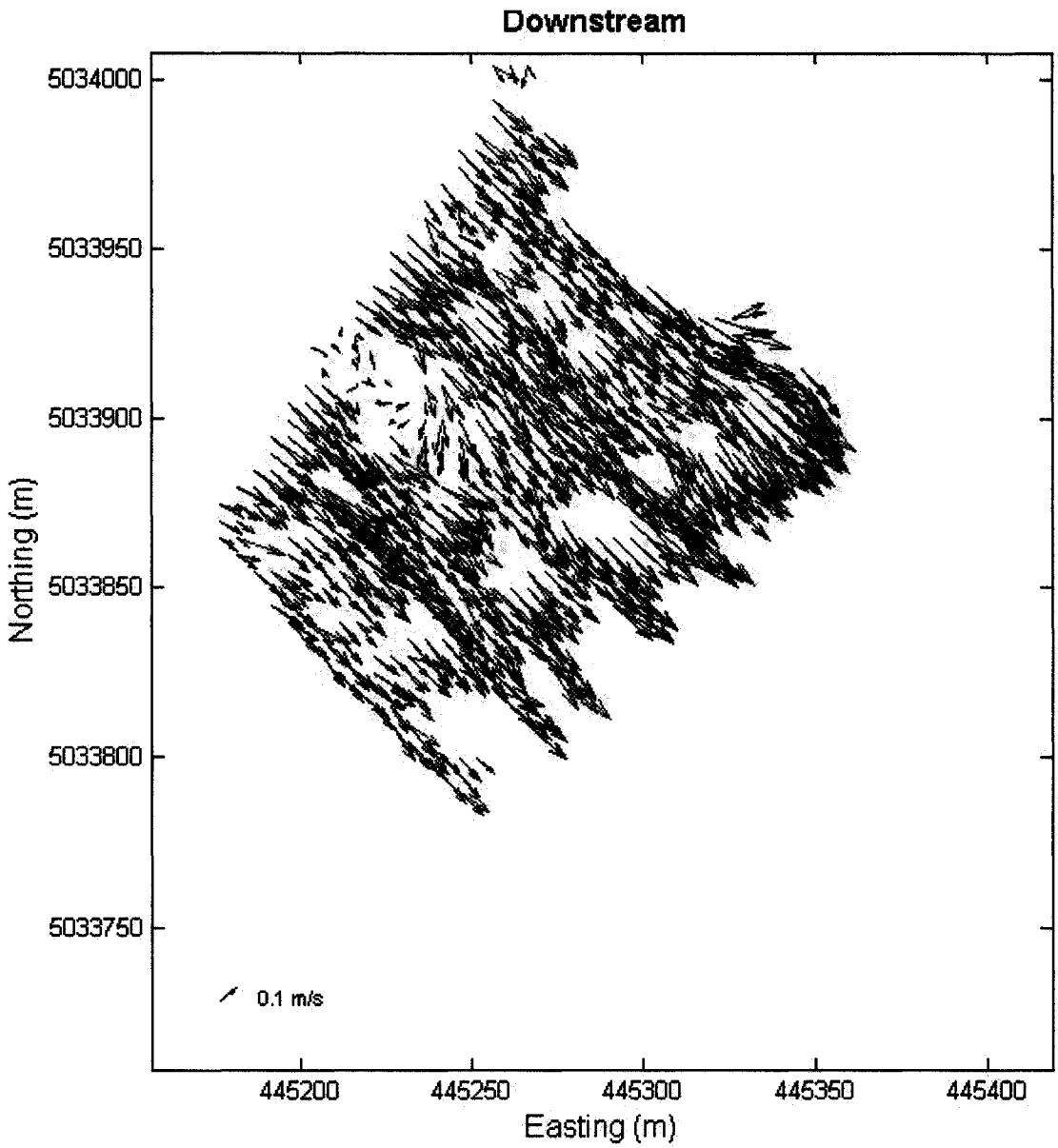

Figure 4. 15 Vector plot of the downstream of the bridge; spatially block averaged (red) and interpolated raw data field using kriging method (black).

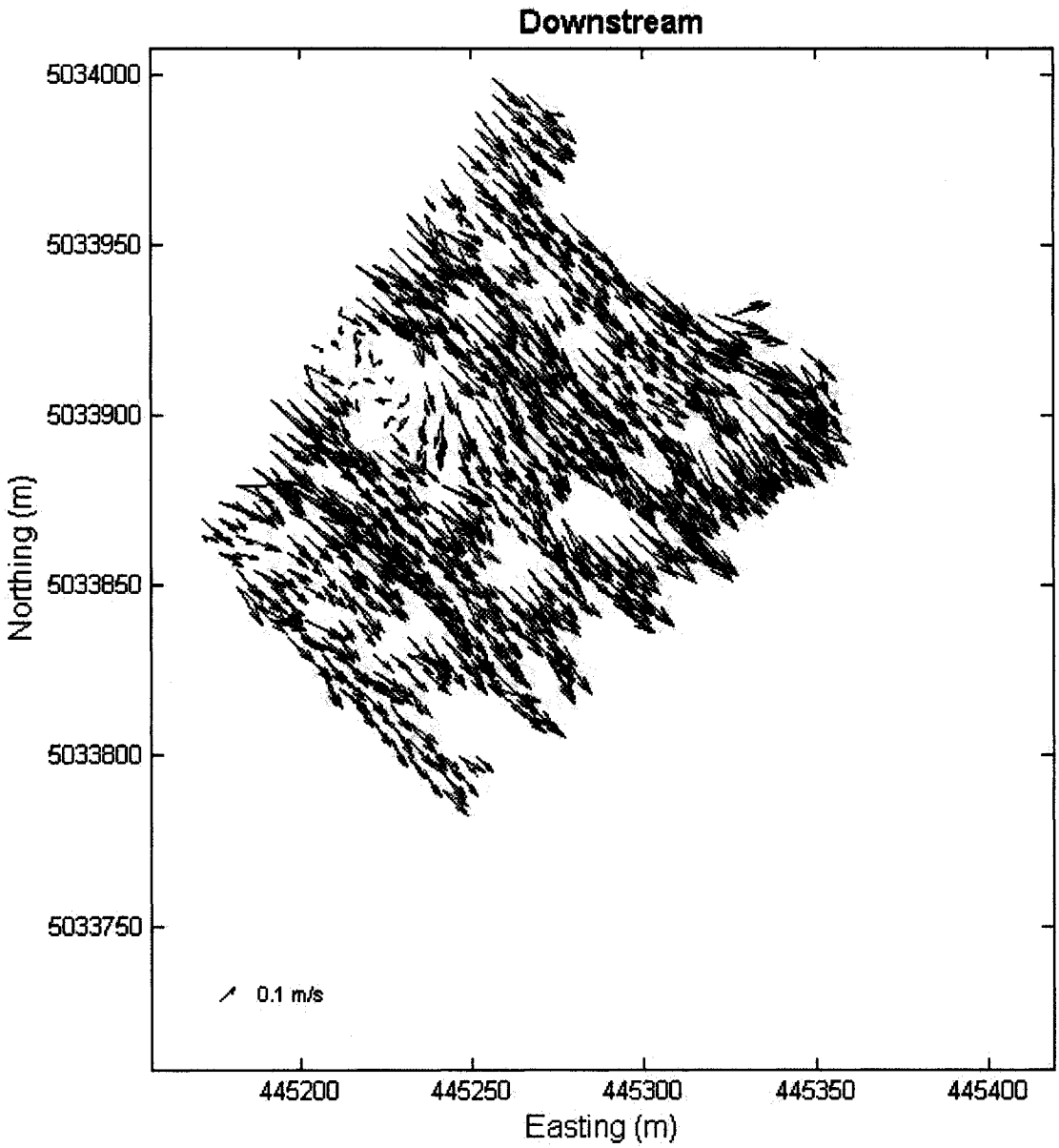

Figure 4. 16 Vector plot of the downstream of the bridge; spatially block averaged (red) and interpolated raw data field using optimal vector interpolation method (black). Vector fields are equally scaled.

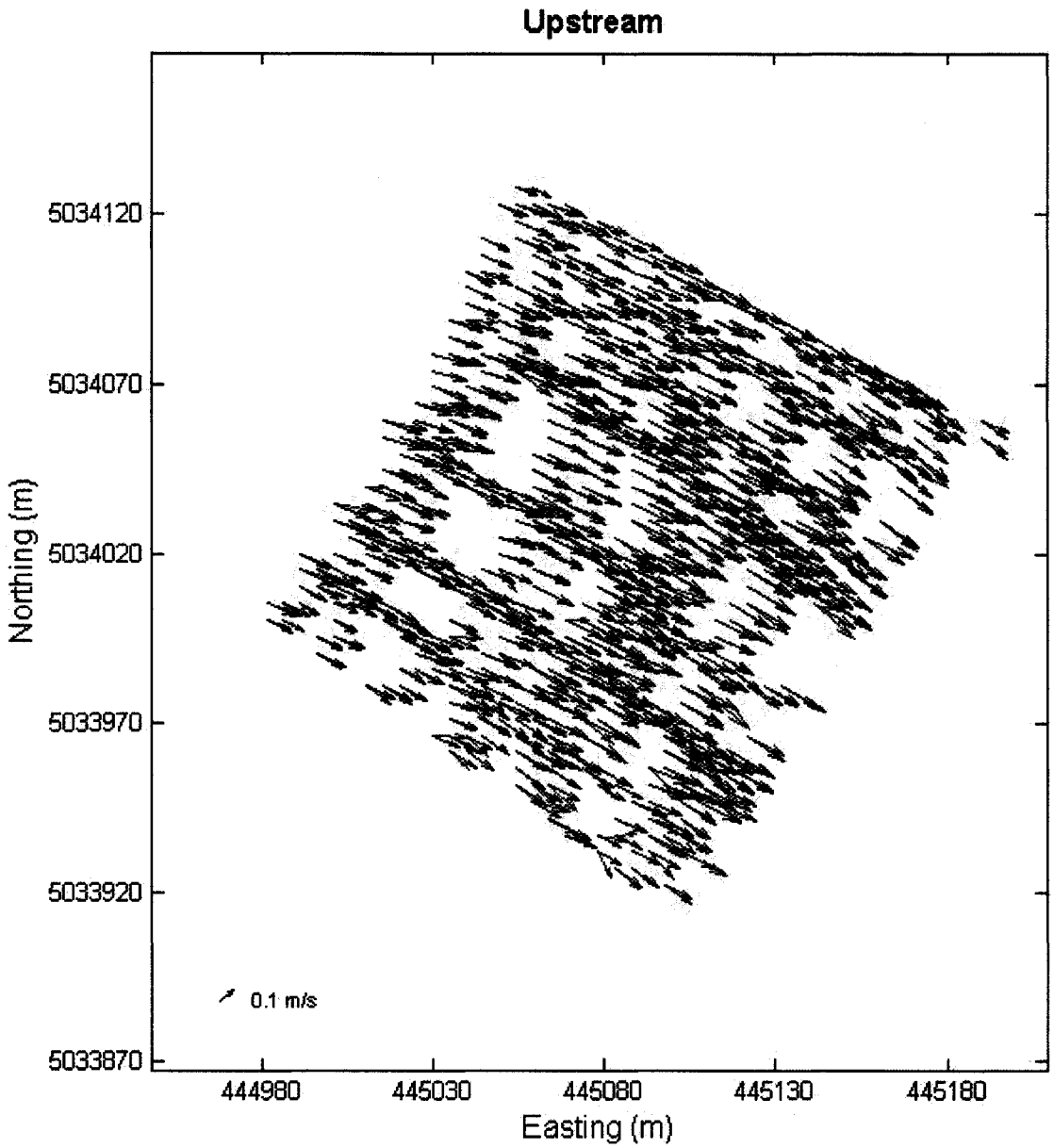

Figure 4. 17 Vector plot of the upstream of the bridge; spatially block averaged (red) and interpolated raw data field using kriging method (black). Vector fields are equally scaled.

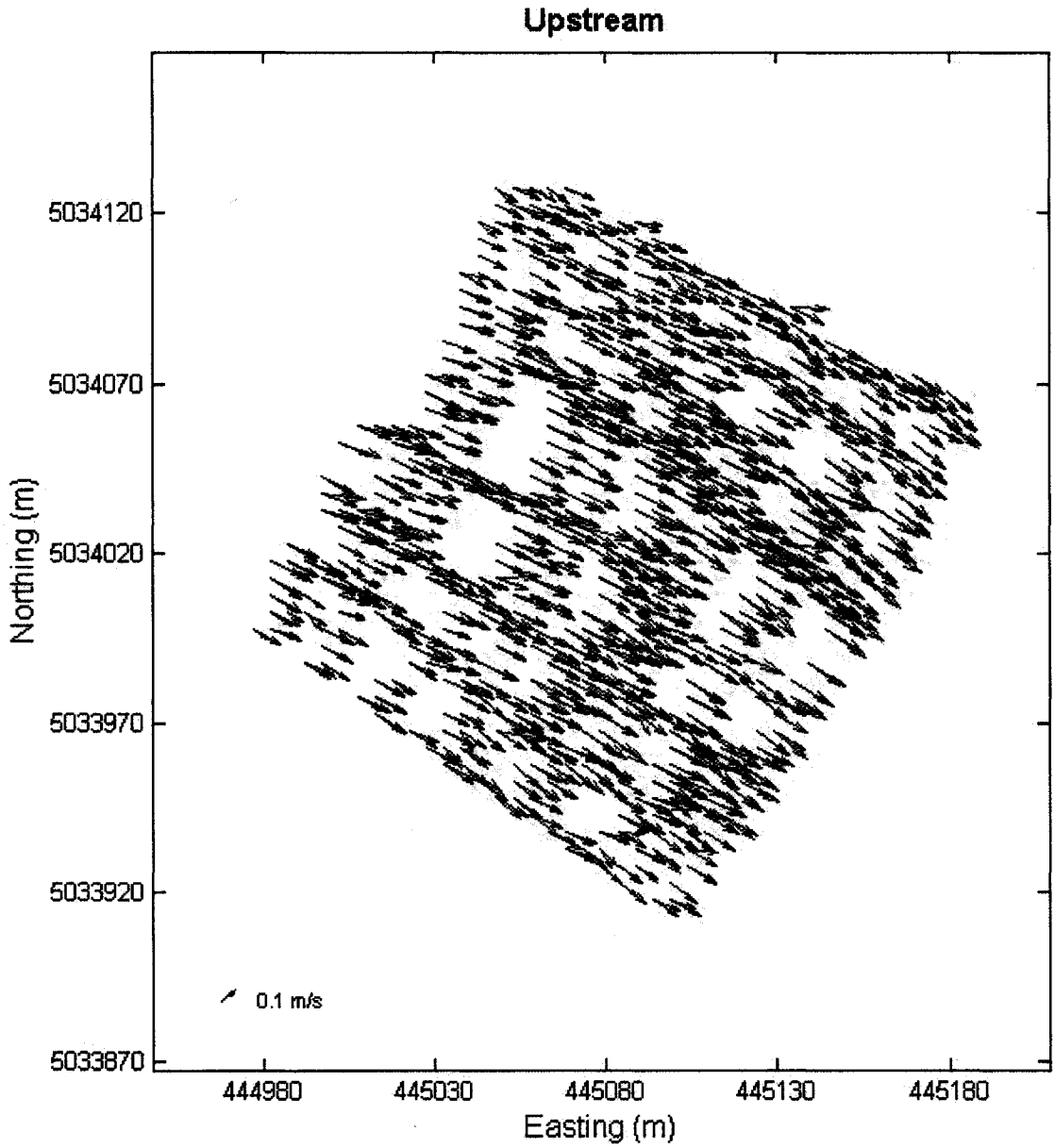

Figure 4. 18 Vector plot of the upstream of the bridge; spatially block averaged (red) and interpolated raw data field using optimal vector interpolation method (black). Vector fields are equally scaled.

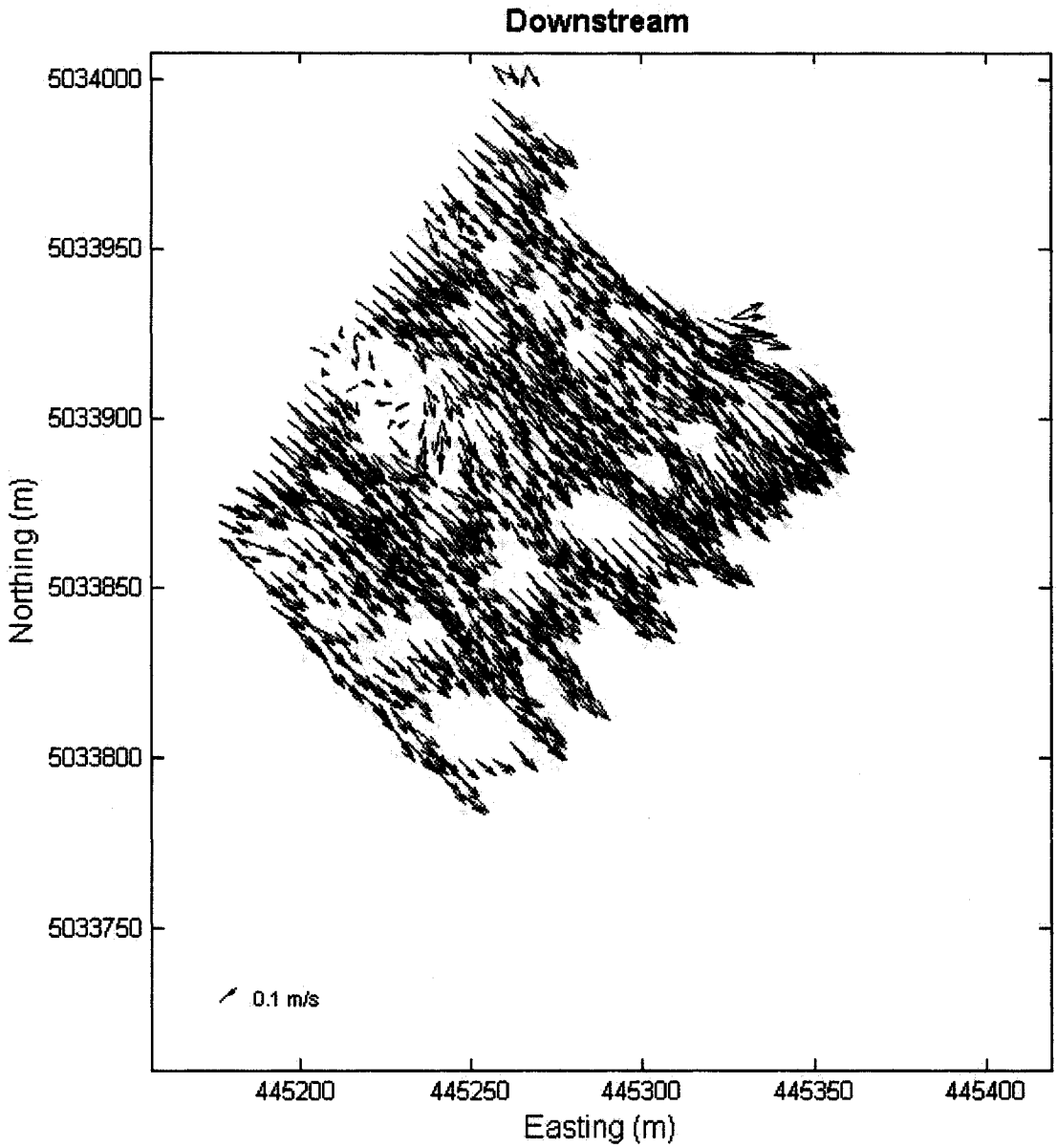

Figure 4. 19 Vector plot of the downstream of the bridge; spatially block averaged (red) and interpolated data field after 30 percent data reduction using kriging method (black). Vector fields are equally scaled.

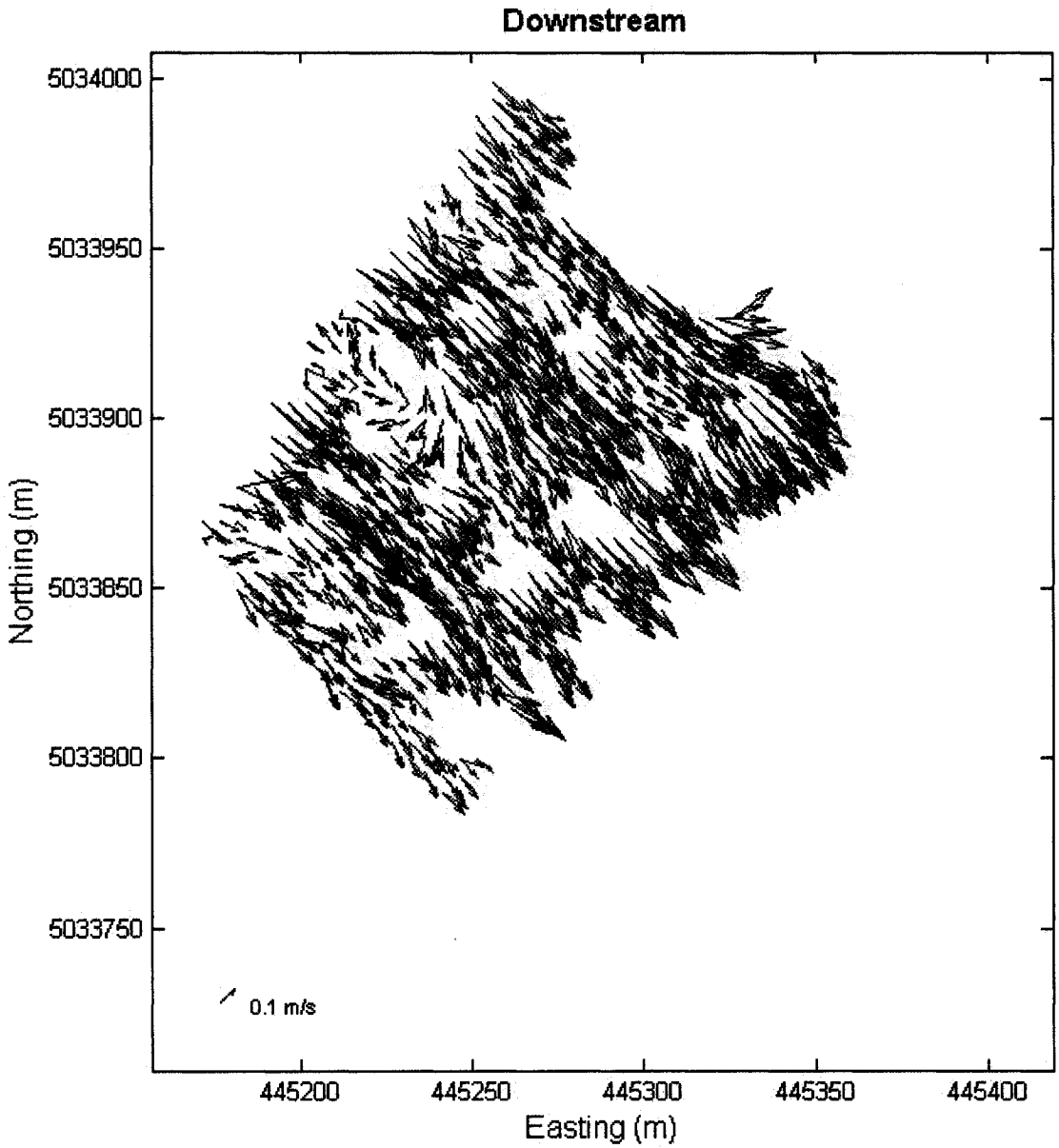

Figure 4. 20 Vector plot of the downstream of the bridge; spatially block averaged (red) and interpolated data field after 30 percent data reduction using optimal vector interpolation method (black). Vector fields are equally scaled.

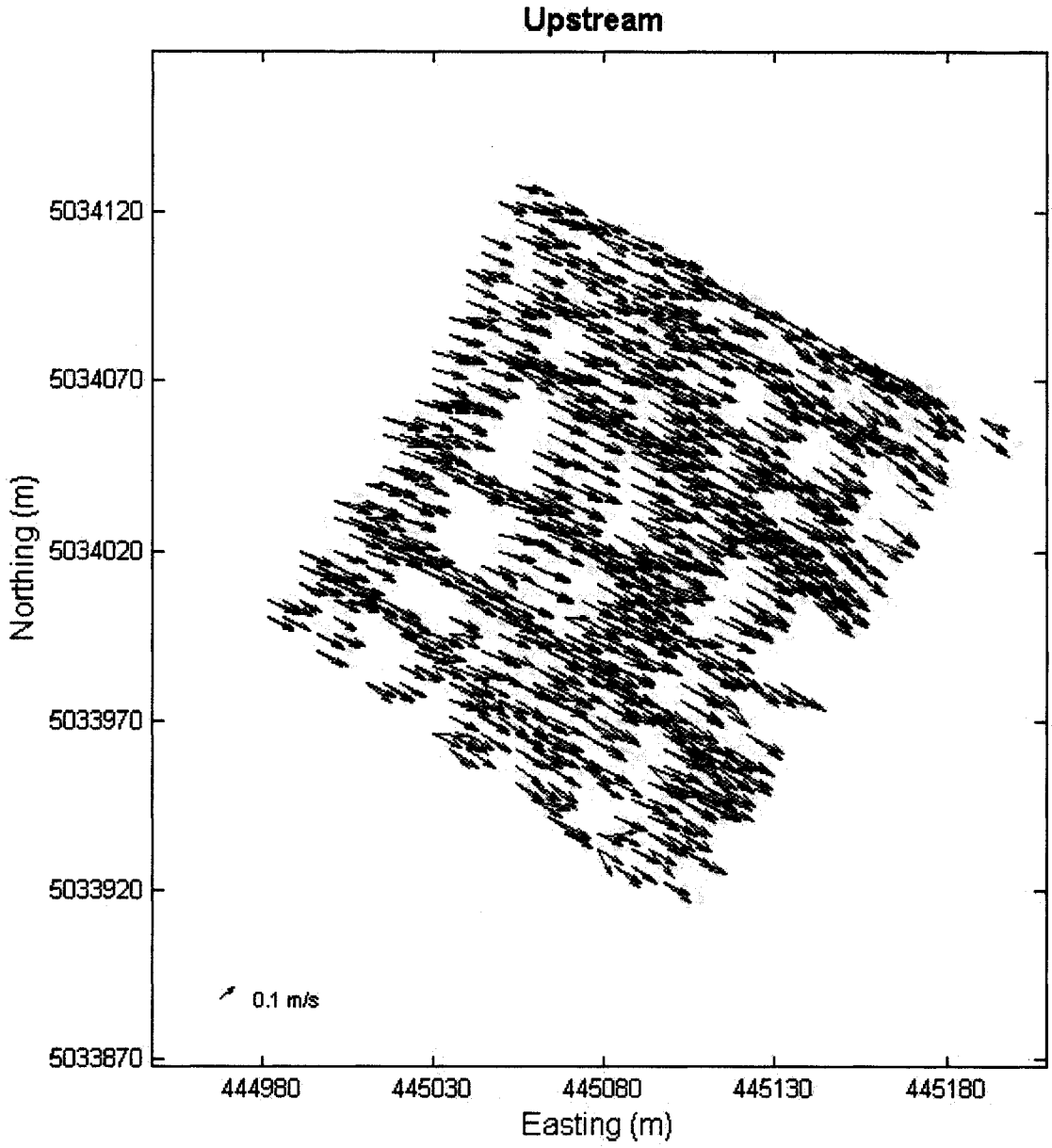

Figure 4. 21 Vector plot of the upstream of the bridge; spatially block averaged (red) and interpolated data field after 30 percent data reduction using kriging method (black). Vector fields are equally scaled.

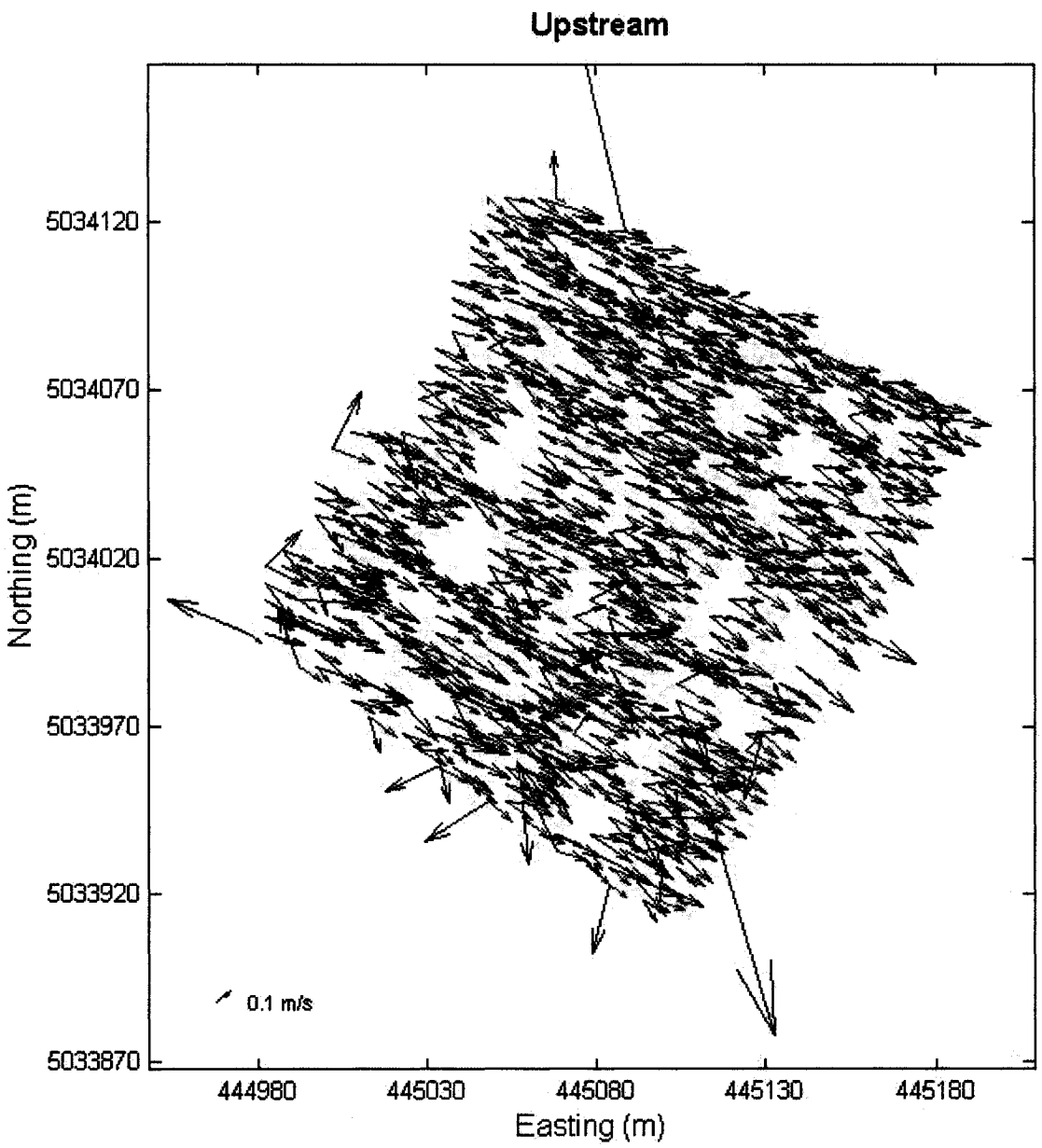

Figure 4. 22 Vector plot of the upstream of the bridge; spatially block averaged (red) and interpolated data field after 30 percent data reduction using optimal vector interpolation method (black). Vector fields are equally scaled.

Table 4. 1 Crosby correlation coefficient between interpolated vector field and spatially block averaged raw field

	Kriging Interpolation Method		Optimal Vector Interpolation Method	
Percentage of Removed Data	Upstream Portion	Downstream Portion	Upstream Portion	Downstream Portion
0	0.8846	1.3674	1.0615	1.5153
10	0.8773	1.3575	0.6822	1.4913
20	0.8433	1.3427	0.6193	1.4424
30	0.8453	1.3389	0.0605	1.4683
40	0.7854	1.3172	N/A	1.3379
50	0.7704	1.304	N/A	1.0819
60	0.7561	1.2726	N/A	N/A
70	0.6676	1.2245	N/A	N/A
80	0.6663	1.1319	N/A	N/A
90	0.6389	1.0146	N/A	N/A
99	0.4439	0.3631	N/A	N/A

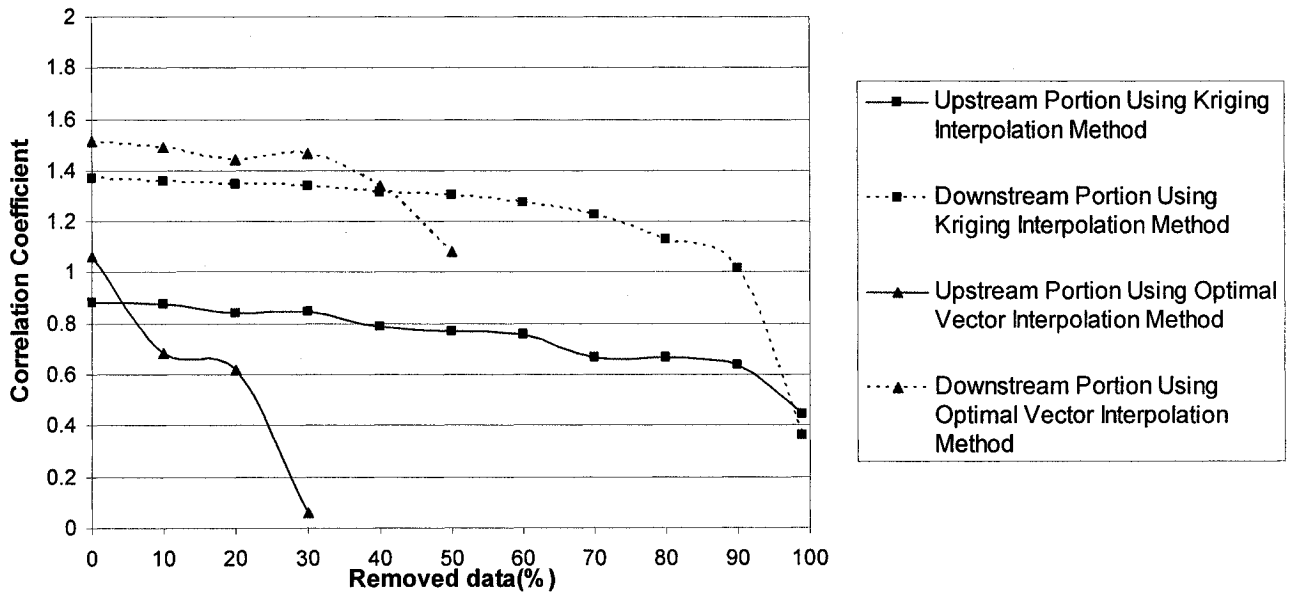

Figure 4. 23 Correlation coefficients values against percentage of data removal.

the data for both employed interpolation methods to evaluate the capability of the methods in successfully interpolating less intensive data fields (Table 4.1, Figure 4.23). As noted above, the data reduction in optimal vector interpolation method was possible only up to 30 percent (upstream) and 50 percent (downstream) removal of data due to the singularity of the covariance matrix $\sum_{YY}$.

The results of the vector correlation assessment are shown in Table 4.1 and Figure 4.23 for both kriging and optimal vector interpolation method. The number of vectors to be compared in downstream and upstream portions was almost the same and the critical value for significant correlation was calculated to be 0.017, suggesting that all correlations were significant. However, in both methods, correlation between interpolated field and raw data field decreased by decreasing the raw data. (Figure 4.23).

As it can be seen in Table 4.1, by increasing the amount of data removal from 0% to 99%, the vector correlation coefficient for kriging decreased from ρ_v^2=0.88 to ρ_v^2=0.44 in the upstream portion and from ρ_v^2=1.37 to ρ_v^2=0.36 in the downstream portion. The reduction of correlation with increasing data removal suggests that the kriging interpolation strategy is less capable of estimating the unsampled data in vector fields with a large degree of missing vectors, yet interpolation is still feasible. The same comparison for optimal method in Table 4.1, shows that decreasing the data to be interpolated from 100% to the minimum possible data (70% for upstream and 50% for downstream portion), diminished the correlation coefficient from ρ_v^2=1.06 to

ρ_v^2=0.06 for upstream and ρ_v^2=1.51 to ρ_v^2=1.08 for downstream part. Unfortunately, the procedure was not feasible with a high degree of missing data, due the inability to calculate covariance matrices from sparse data for each spatial block.

4.4 Evaluation of the Dependency of the Interpolation Methods to the Choice of the Coordinate System

According to Rennie and Millar (2004), the kriging method is dependent on the choice of the coordinate system due to the influence of individual scalar variograms (i.e. spatial correlation) on the results. In other words, interpolating the same vector field in different coordinate systems causes differences between resultant fields. Changing the coordinate system is necessary in kriging method in order to eliminate the cross correlation between vector components, and should limit the influence of choice of coordinate system. Alternatively, Feliks et al. (1996) declared that the optimal vector interpolation method is independent to the rotation of the coordinate system. In following paragraphs, the effect of changing the coordinate system is reviewed for kriging and optimal vector interpolation methods respectively. Vector maps of the resultant interpolated fields decomposed in easting-northing and streamwise-cross-stream systems are superimposed to demonstrate the differences between two interpolated methods in estimating the vectors in two different coordinate systems.

4.4.1 The Effect of Rotating the Coordinate System on Kriging Interpolation Method

Easting and Northing components of the vector field were interpolated by kriging method using the modeled variograms related to each downstream and upstream portion of the field. The modeled variograms of easting and northing components of data in downstream and upstream portion are presented in Figures 4.1 and 4.2. Using Equation 3.22a and 3.22b, the easting-northing velocity vectors components were rotated by the average of the vector angle which was θ=-41.95° for downstream portion and θ=-27.59° for upstream portion. Data were interpolated in the new coordinate system and rotated back to their original, using the opposite rotation angles. High correlation coefficients of ρ_v^2=1.97 for upstream and ρ_v^2=1.95 for downstream were calculated when comparing the interpolated fields in two different coordinate systems.

Figure 4.24 shows the superimposed map of rotated and unrotated data for upstream. Although the range of interpolated velocity vectors differs from 0.113-0.293 m/s for unrotated to 0.102-0.295 m/s for rotated data, the two fields are very similar. The comparative variograms for data in rotated and unrotated coordinate systems suggest the same idea since they are similar in both coordinate systems, largely due to a general lack of spatial trend in the variograms. (Figure 4.1).

However, comparing the related variograms of downstream part (Figure 4.2) demonstrates reasonable differences between the spatial correlation of easting-

northing versus streamwise-cross-stream data. In Figure 4.25 the interpolated field in easting and northing coordinate system is superimposed with the interpolated field in streamwise and cross-stream coordinate. The dependency of the interpolation method to the choice of the coordinate system is apparent in this figure. This occurs because there is spatial trend in the variograms for the downstream portion (Figure 4.2).

4.4.2 The Effect of the Rotating the Coordinate System on Optimal Vector Interpolation Method

The same procedure was carried out by interpolating data with east-north decomposition and streamwise-cross-stream decomposition by means of the optimal vector interpolation method. The superimposed map for the downstream portion is shown in Figure 4.26. Clearly, the underneath map can not be seen and the vector maps interpolated in two different coordinate systems are precisely the same. Furthermore, the perfect correlation coefficient of ρ_v^2=1.99 for both upstream and downstream portion confirms the independency of the method to the choice of the coordinate system.

4.5 Discussion

4.5.1 Kriging Method

Kriging is an interpolation method which is dependent on the choice of the coordinate system, thus the method can only reliably interpolate scalars and not vectors. To apply the method on the velocity vector fields, two dimensional quantities have to be

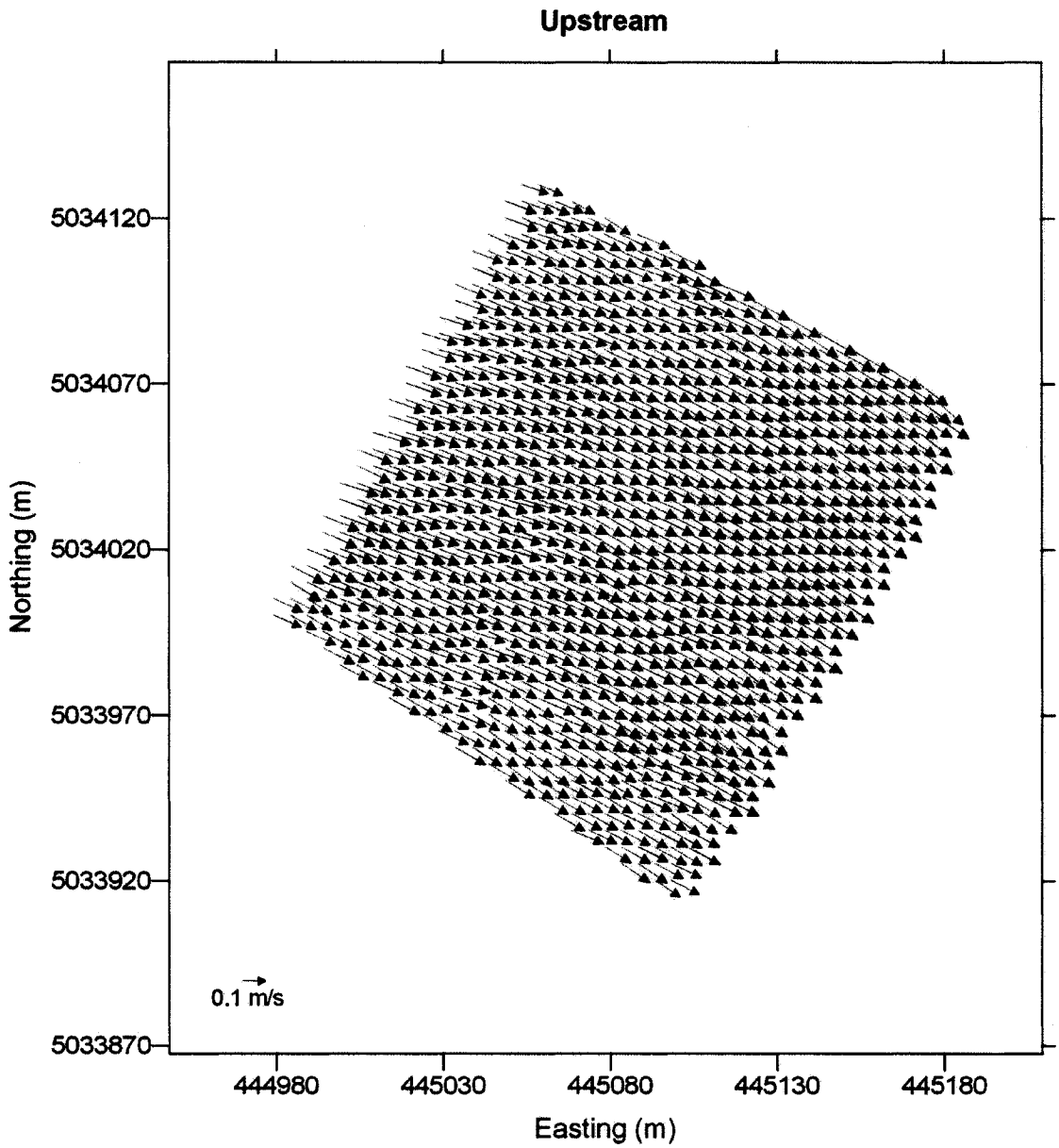

Figure 4. 24 Superimposed maps of interpolated upstream portion using kriging in two different coordinate system. Vector fields are equally scaled.

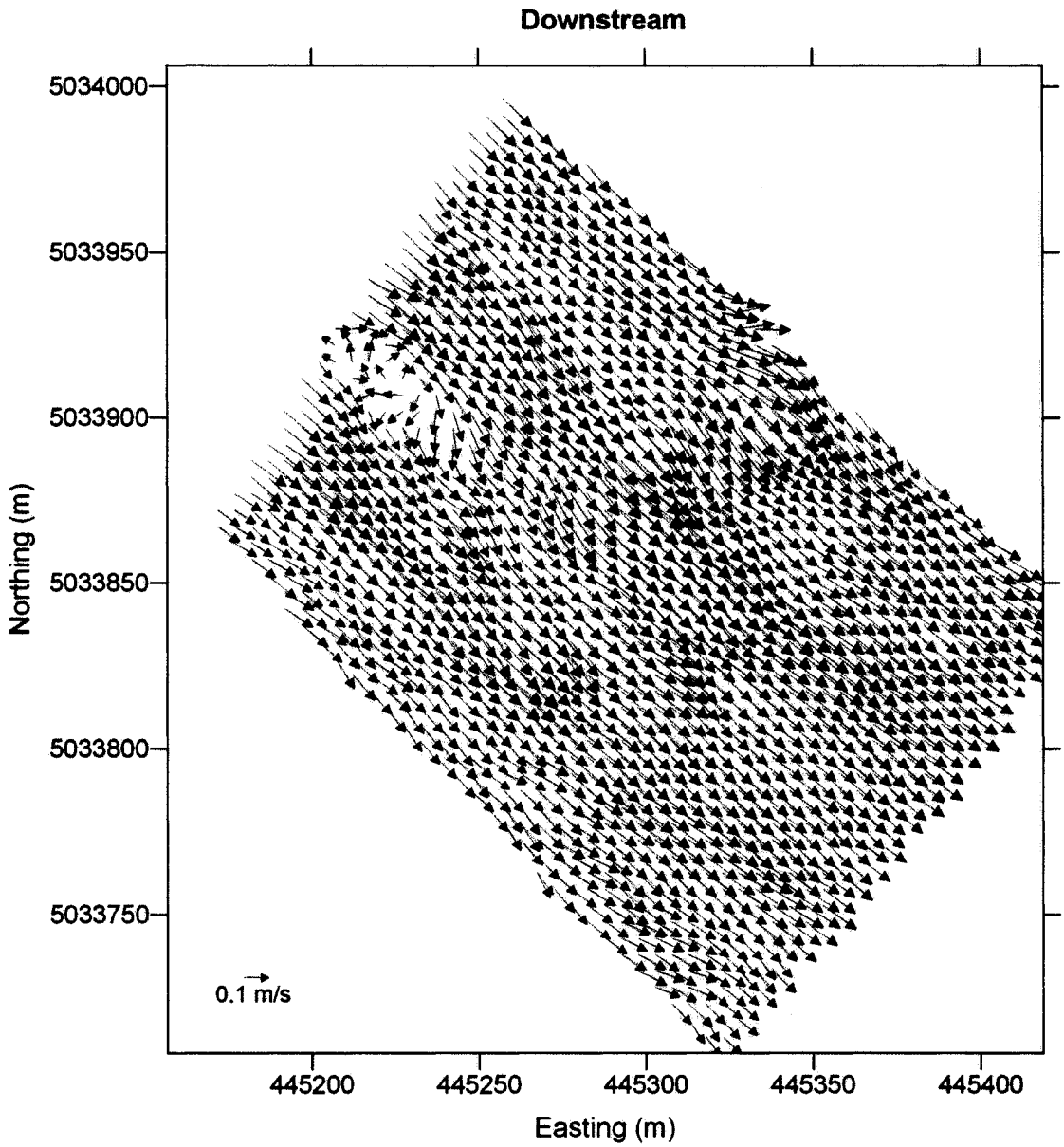

Figure 4. 25 Superimposed maps of the interpolated downstream portion using kriging in two different coordinate system. Vector fields are equally scaled.

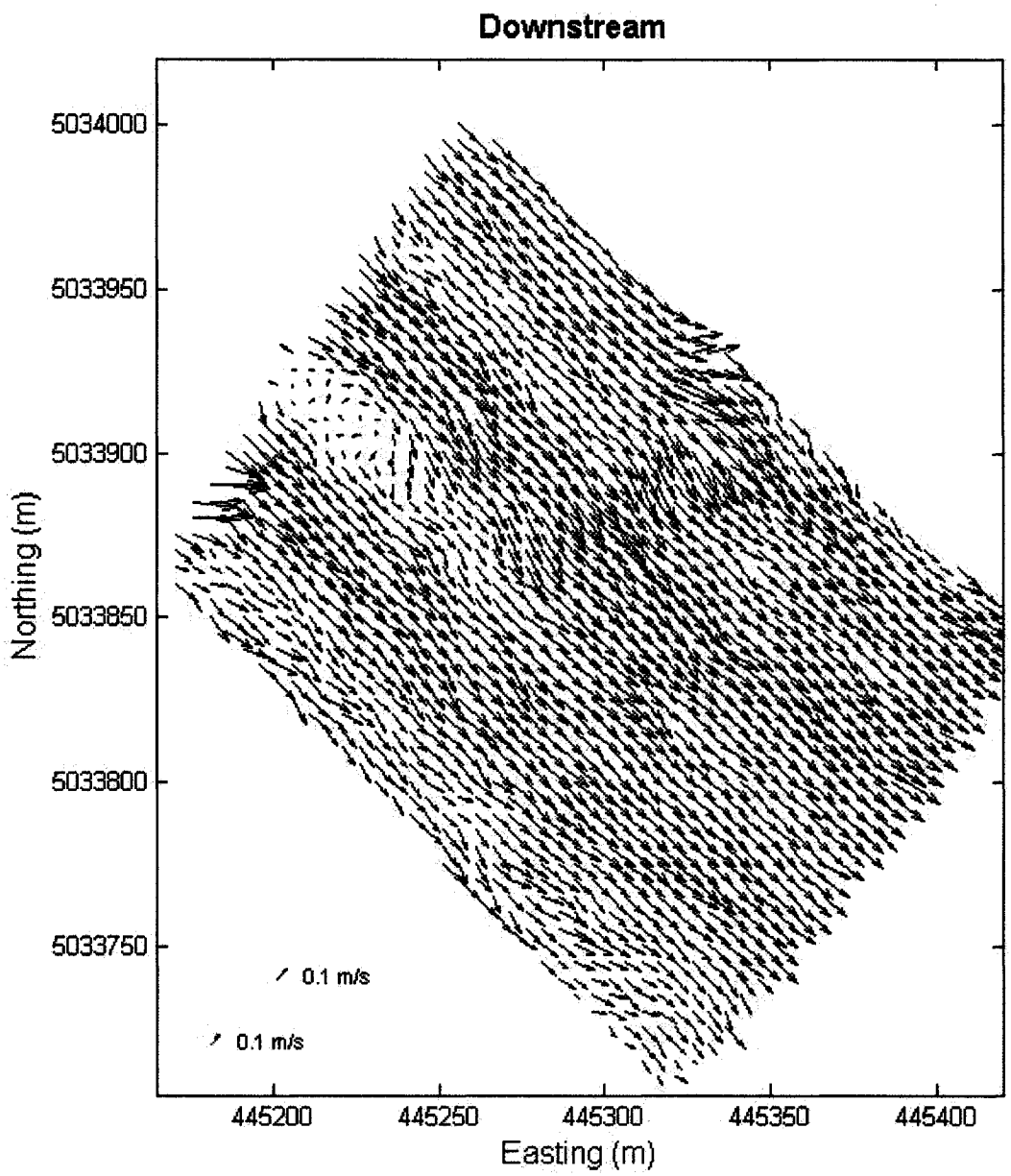

Figure 4. 26 Superimposed maps of interpolated downstream portion using optimal method in two different coordinate system. Vector fields are equally scaled.

decomposed to their scalars and interpolated independently. Consequently the association between vector components is not taken into account and vector properties are ignored during the interpolation procedure. In order to avoid this problem, the fields to be interpolated are rotated into a coordinate system in which the cross correlation between the components is zero. This solution, however, effects the spatial correlation of the velocity components and consequently influences the interpolation procedure. In this study, rotating the coordinate system to eliminate the cross correlation decreased the spatial correlation of the streamwise velocity components in downstream portion which can be seen in Figure 4.2c.

The correlation coefficient amount of ρ_v^2=0.88 for upstream portion and ρ_v^2=1.37 for downstream portion which exceeded the significant correlation criterion of 0.017 suggests that the method is successful in interpolating the velocity vector fields both in variable flow (downstream portion) and uniform flow (upstream portion).

Although the smaller amount of vector correlation coefficient after data reduction in both downstream and upstream portions demonstrates that the technique is less capable to interpolate fields with a high degree of missing data, still the minimum correlation coefficient in each field (ρ_v^2= 0.44 for upstream and ρ_v^2=0.35 for downstream) exceeded the significant correlation of 0.017. Even with 99% data removal, the estimated vectors were significantly correlated to the raw data. Kriging appears to be robust in situations with sparse data.

4.5.2 Optimal Vector Interpolation

Interpolating the velocity vector fields using the optimal vector interpolation method and correlating the resultant uniform vector maps with spatially block averaged fields in downstream and upstream portions returned correlation coefficient of ρ_v^2=1.06 for the upstream portion and ρ_v^2=1.51 for the downstream portion.

Reducing the data within the selected grids which are considered as measurement stations resulted in singular covariance matrices for $\sum_{YY}$ .The maximum possible data reductions were 30% and 50% for upstream and downstream portions respectively. By removing more than these amounts of data for each portion, the $\sum_{YY}$ became a singular matrix which is not invertible. Therefore the interpolation procedure was not possible to be performed further than the abovementioned data removal.

The criteria for using a block as measurement station were a minimum of 47 and 60 vectors in each block in the upstream and downstream portions respectively. Therefore, the number of vectors in upstream blocks was less than in downstream blocks. Due to the smaller number of vector in upstream blocks, the covariance matrix in upstream portion becomes singular after only 30 % of data reduction as compared to 50% data reduction in the downstream portion. Still, the optimal vector interpolation method yielded significant minimum correlation coefficients for both upstream and downstream portions (ρ_v^2=0.06 and ρ_v^2=1.08>0.017). These results state that the interpolated fields are correlated with the measured fields.

In general, the singularity of the covariance matrix $\sum_{YY}$, which according to Equation 3.31 needs to be inverted, is a major disadvantage of the method. As it was mentioned before, using the entire measured field produces a singular $\sum_{YY}$ matrix. A huge number of measured vectors were rejected to gain an invertible covariance matrix $\sum_{YY}$ .In addition, assessing the capability of the optimal vector interpolation throughout a complete procedure of the 10% to 99% data removal was not achievable due to the same problem.

4.5.3 Comparison of the Interpolation Methods

A clear advantage of the optimal vector interpolation method developed herein for interpolation of asynoptic spatial vector surveys is that it is independent of the choice of coordinate system. Significantly greater correlation (α=0.05) between the estimated and actual fields were achieved using the optimal vector interpolation method (ρ_v^2=1.51 in downstream portion and ρ_v^2=1.06 for upstream portion) than with kriging (ρ_v^2=1.37 for downstream portion and ρ_v^2=0.88 for upstream).

However, the optimal vector interpolation method is very dependent on the *a priori* first guess of the interpolated field. Obviously, more successful interpolation is achieved when the *a priori* data better represents the actual field. Further, although the optimal vector interpolation method gives the impression of being more competent to interpolate irregular vector fields as compared to the kriging method, it should be emphasized this is only possible in more intensively measured fields, in which some measured data can be

removed, in order for the structure of $\sum_{YY}$ matrix to be non-singular. In other words, the optimal method requires very intensive measured data within the measurement stations or blocks to be able to estimate unsampled vectors.

4.5.4 Interpolation of Upstream versus Downstream Portions

An unexpected result was that correlation values between interpolated and measured fields were greater in the downstream portion than the upstream portion. For example, the correlation coefficient between the raw data field and interpolated actual field for downstream by kriging was 1.37 which was significantly greater than the corresponding correlation coefficient of 0.88 for upstream (the calculated Z for comparison of correlation coefficients was 5.918). The upstream portion had more uniform flow, thus it was expected that interpolation would be less successful in the downstream portion.

However, this counter-intuitive result may be an artefact of the relatively flat variogram observed and modeled in the upstream portion (Figure 4.1c), versus a spatial autocorrelation structure in the downstream portion indicating less similarity for points separated by distances greater than 40 m (Figure 4.2c). In the upstream portion, the flat variogram produces a highly smoothed interpolated surface, because all interpolated values are determined by linear combination of the same large set of distant neighbouring points. The estimated interpolated vectors are all very similar (Figures 4.3 and 4.9), despite actual small local variances in vectors (Figure 3.9a), which lowers the correlation between interpolated and actual fields. In the downstream portion, on the other hand, the lack of spatial autocorrelation at greater lag distance ensures that only nearby vectors are

utilized to estimate an interpolated vector, which explains the increased correlation between interpolated and actual fields in the downstream portion. It may be possible to increase the correlation for the upstream portion by kriging if the variogram were modeled with a sill at a small range of influence. However, this can not be achieved for the optimal vector interpolation method, because lag distances are not explicitly considered in covariance matrices that are computed from the whole field.

One consideration about the Crosby (1993) vector correlation coefficient is that vectors in each vector field should be independent. The measured variograms suggest that vectors are correlated when separated by distances of up to 40 m for the downstream portion, and over even larger distances for the upstream portion. Thus, it may be prudent to calculate the Crosby correlation coefficient on a subsample of each vector field, in which the vectors are spaced by 40 m. This was performed in order to ensure that the achieved correlation coefficients were not due to the similarity between the adjacent vectors in 5m by 5m blocks. Spaces of 10m, 20m, 30m, and 40m between the vectors which were correlated were chosen. The results are shown in table 4.2. As can be seen, the correlation coefficient between observed spatial block average fields and interpolated fields remain greater for the downstream portion even when the vectors are 40m apart.

Table 4. 2 Crosby correlation coefficient between interpolated vector field and spatially block averaged raw field in subsamples greater than 5m.

	Subsample Size			
Portion	10m	20m	30m	40m
Upstream Portion	0.94	0.96	1.24	1.16
Downstream Portion	1.31	1.32	1.56	1.70

5 CONCLUSION

Two vector interpolation methods were applied on velocity vector data measured asynoptically in a 400m reach of Gatineau River. The reach was divided into upstream and downstream portions with a highway bridge which intersected the river. Required data were measured by means of an ADCP. Raw data vectors in upstream and downstream portions were interpolated separately. Data were first interpolated using the kriging interpolation method (Bentamy 1996) and in the next step, the same data sets were interpolated by means of a novel optimal vector interpolation method modified from that proposed by Feliks et al. (1996). The method proposed by Feliks et al. (1996) was applicable to time series collected at a few individual stations, whereas the method presented in this thesis can interpolate a spatial field from data collected in asynoptic spatial surveys wherein data are collected as single point measurements distributed quasi-randomly throughout the spatial domain. The interpolation methods were also applied on less intensive fields by gradually removing 10, 20, 30, 40, 50, 60, 70, 80, 90, and 99 percent of the entire data. This procedure was performed to assess the capability of the methods to interpolate the field with sparse data. Furthermore, dependency or independency of the interpolation methods to the choice of the coordinate system were assessed by comparing the resultant interpolated data in different coordinate systems. The comparison between the interpolated and raw data fields was performed by employing the Crosby (1993) vector correlation coefficient.

Velocity vectors were decomposed to their scalars and interpolated independently by means of the kriging interpolation method. In order to diminish the effect of cross correlation between the scalars, which is ignored during the interpolation procedure, the velocity vectors are decomposed in streamwise-cross-stream coordinates. The interpolated velocity vector fields using kriging method correlated well with spatially block averaged actual fields in both upstream and downstream portion of the field. Significant Crosby vector correlation coefficients were ρ_v^2=0.88 for the upstream portion and ρ_v^2=1.37 for the downstream portion.

The same measured fields were interpolated using optimal vector interpolation method. Significant (α=0.05) correlation coefficients of ρ_v^2=1.06 for upstream portion and ρ_v^2=1.51 for downstream portion were obtained when the interpolated data were correlated with spatially block averaged measured fields..

Interpolation of the removed data fields still showed that all interpolated vector fields were significantly (α=0.05) correlated with the actual data field in both methods. However, diminishing the correlation coefficient by increasing the percentage of data removal (see Table 4.1) suggests that both methods of interpolation which are used in this study are less capable in estimating the velocity vector fields with large amount of the unmeasured vectors. This incapability is more apparent for the optimal vector interpolation method. In this method interpolating the data was possible only up to 30% of the data removal in the upstream portion and 50% in the downstream portion (See Table 4.1).

Interpolating the measured fields in two different coordinate systems using kriging method and superimposing the resultant maps demonstrated a slight difference between interpolated maps which is due to the dependency of the method to the choice of the coordinate system. Conversely, the same test for the optimal vector interpolation method did not cause any change between interpolated maps in two different coordinate systems and confirmed the independency of the method to the choice of the coordinate system.

Calculated correlation coefficients for optimal vector interpolation method were significantly (α=0.05) greater than correlation coefficients for the kriging method. The significant difference between the Crosby vector correlation coefficient achieved from correlating the measured field and interpolated fields using two different correlation coefficient methods verified that the new optimal vector interpolation method is more effective. Furthermore, contrary to the kriging method, the optimal vector interpolation method deals with vectors by considering the cross correlation between the vector components. It also is independent to the choice of the coordinate system.

Therefore, the method should be considered as a viable alternative to the kriging method. However, the optimal vector interpolation method is very dependent to the *a priori* data. Moreover, the number of the measured vectors in the grids and also the number of the grids need to be optimized in order for the interpolation to be soluble. Furthermore, kriging was more capable of interpolating measured fields with less number of sampled vectors.

The downstream portion of the bridge had more variable flow due to the effect of the bridge piers, thus it was expected that the interpolation would be more successful in the upstream portion with more uniform flow. The greater Crosby correlation coefficient achieved by both methods in downstream portion was the unexpected result in this study, and may have been due to the relatively flat variogram in the portion that resulted in an oversmoothed interpolated vector field.

In order to ensure vector independence when calculating the correlation coefficient, the space between the vectors was increased gradually. The Crosby vector correlation coefficient for upstream portion was still smaller than for the downstream portion even with 40m spacing between the vectors.

Based on the above observations it is concluded that the new optimal vector interpolation method is superior to kriging when sufficient measured data are available. Kriging, on the other hand, is more reliable when interpolating field with sparse vectors. Since the interpolated field using optimal vector interpolation method is very dependent to the *a priori* field, it is expected that the method is superior when reasonable *a priori* data are available. Further, it is necessary to test the $\sum_{YY}$ covariance matrix for singularity when using the new optimal vector interpolation method. However, independency of the optimal vector interpolation method to the choice of the coordinate system and the capability of the method for interpolating the vectors by taking the cross correlation of the components into account makes the method advantageous compared to the kriging interpolation method.

REFERENCES

Bentamy A, Quilfen Y, Gohin F, Grima N, Lenaour M, Servain J. 1996. Determination and validation of average wind fields from ERS-1 scatterometer measurements. *The Global Atmosphere and Ocean System* **4:** 1-29.

Breckling J. 1989. *The analysis of directional time series: application to wind speed and direction.* Springer-Verlag.

Bretherton Francis P, Davis Russ E, Fandry C.B.1976. A technique for objective analysis and design of oceanographic experiments applied to MODE-73. *Deep-Sea Research* **23:** 559-582.

Buell C.E. 1971. Two-point wind correlations on an isobaric surface in a non homogeneous non-isotropic atmosphere. *Journal of Applied Meteorology* **10:** 1266-1274.

Charles B.N. 1959. Empirical models of interlevel correlation of winds. *Journal of Meteorology* **16:** 581-585.

Court A. 1958. Wind correlation and regression. AFCRC TN-58-230. Contract AF19(604)-2060. Cooperative Research Foundation, 16 pp.

Crosby D. S, Breaker L. C, Gremmill W. H. 1990. A definition for vector correlation and its application to marine surface winds. National Meteorological Center Office Note No 365, NOAA/ National Weather Service, Washington D.C., 50 pp. [Available from the National Meteorological Center, 2500 Auth Rd., Washington D.C. 20233.]

Crosby D. S, Breaker L. C, Gremmill W. H. 1993. A proposed definition for vector correlation in geophysics: theory and application. *Journal of Atmospheric and Oceanic Technology* **10:** 355-367

Daley R. 1991. *Atmospheric Data Assimilation.* Cambridge University Press.

Detzius, R., 1916. Extension of cirrelaton methods and the method of least square to vectors. *Sitzungsber. Akas. Wiss. Wien.* **125(Ha):** 3-20.

Durst B.A. 1957. A statistical study of the variation of wind with height. Professional Notes. Vol 8. No.121. Air Ministry. Meteorological Office, 10 pp. [Available from NOAA Asheville Library, NCDC Library, Federal Building MC16, Asheville, NC 28801-2696.]

Dinehart R. L, Burau J. R. 2005 (a). Averaged indicators of secondary flow in repeated acoustic Doppler current profiler crossing of bends. *Water Resources Research* **41:** 1-18.

Dinehart R. L, Burau J. R. 2005 (b). Repeated surveys by acoustic Doppler current profiler for flow and sediment dynamics in a tidal river. *Journal of Hydrology* **314:** 1-21.

Feliks Yizhak, Gavez Ehud, Givati Reuven. 1996. Optimal vector interpolation of wind fields. *Journal of Applied Meteorology* **35:** 1153-1165.

Freeland H. J, Gould W. J. 1976. Objective analysis of meso-scale ocean circulation features. *Deep-Sea research.* **23:** 915-923.

Gandin L.S. 1965. Objective analysis of meteorological fields. Israel Program for Scientific Translation. 242 pp.

Hayashi Y. 1980. A method of estimating space-time spectra from polar orbiting satellite data. *Journal of Atmospheric science* **37:**1385-1392.

Hooper J. W. 1959. Simultaneous equations and canonical correlation theory. *Econometrica* **27:** 245-256.

Inggs Micheal R, Lord Richard T. 1995. Interpolating satellite derived wind field data using ordinary kriging, with application to the nadir gap. *International Geoscience and remote sensing symposium* **1:** 141-143.

Isaaks E.H, Srivastava R.M. 1989. *An Introduction to Applied Geostatistics.* Oxford University Press, New York.

Jupp P.E, Mardia K.V. 1980. A general correlation coefficient for directed data and related regression problems. *Biometrika* **67:** 163-173.

Legler D, O'Brien J. 1985. Development and testing assimilation technique to derive average wind fields from simulated scatterometer data. *Mon Weather Rev* **113:** 1791-1800

Ludwig F. L, Livingstone J.M, Endlich R. M. 1991. Use of ,as conservation and critical diving streamline concepts for efficient objective analysis of winds in complex terrain. *Journal of Applied Meteorology* **30:** 1490-1499.

Mc Williams J.C. 1976. Maps from the mid-ocean dynamics experiment: Part I. Geostrophic streamfunction. *Journal of physical Oceanography* **6:** 810-827.

Manly Bryan F.J. 2001. *Statistics for Environmental Science and Management.* Chapman & Hall/CRC.

Mardia K. V, Puri M. L. 1978. A spherical correlation coefficient robust against scale. *Biometrika* **65:** 391-395.

Matheron G. The theory of regionalized variables and its applications. Techn. Rep. Fascicule5, Les cahiers du centre dee Morphologie Mathemathique de Fontainbleau. Ecole Superieure des Mines. Paris 1971.

Mellville Bruce W, Coleman Stephen E. 1999. *Bridge Scour.* Water Resources Publications, LLC.

Porch W, Rodriguez D. 1987. Spatial interpolation of meteorological data in complex terrain using temporal statistics. *Journal of Climate and Applied Meteorology.* **26** 1696-1708.

Rennie Colin D, Rainville F. 2006. Case study of precision of GPS differential correction strategies: Influence on aDcp velocity and discharge estimates. *Journal of Hydraulic Engineering* (ASCE) **132(3):** 225-234.

Rennie Colin D, Millar Robert G. 2004. Measurement of the spatial distribution of fluvial bedload transport velocity in both sand and gravel. *Earth Surface Processes and Landforms* **29:** 1173-1193.

Rouhani Shahroukh, Myers Donald E. 1990. Problems in space-Time kriging of geohydrological data. *Mathematical Geology* **22:** 611-623.

Ruther N, Olsen N. 2005. Three-dimensional modeling of sediment transport in a narrow 90° channel bend. *Journal of Hydraulic Engineering* **131:** 917-920.

Servain J, Gohin F, Muzellec A. 1993. Wind fields at the surface determined from combined ship and satellite altimeter data. *Journal of Atmospheric and Oceanic Technology* **10:** 880-886.

Sherman C. A. 1978. Mathew: A mass-consistent model for wind fields over complex terrain. *Journal of Applied Meteorology* **17:** 312-319.

Simpson, M.R, Oltman R.N. 1993, Discharge measurement system using an acoustic Doppler current profiler with application to large rivers and estuaries. *U.S. Geological Survey Water-Supply Paper* 2395.

Simpson M. R. 2001. Discharge measurements using a broad-band acoustic Doppler current profile.: *U.S. Geological Survey Open-File Report* 01-1, at http://pubs.water.usgs.gov/ofr0101.

Stamhuise Eize J, Videler John J. 1995. Quantitative flow analysis around aquatic animals using laser sheet particle image velocimetry. *The Journal of Experimental Biology* **198:** 283-294.

Stephens M. A. 1979. Vector correlation. *Biomatrika* **66:** 41-48.

Sverdrup H. U. 1917. On correlation between vectors with applications to meteorological problems. *Meteor.* Z. **34:** 285-291.

Tracy Barbara A. 2002. Directional characteristics of the 1990-1999 wave information studies Gulf of Mexico. *Proceeding 7th International Workshop on Wave Hindcasting and Forcasting, Invironment Canada.*

Wilkin John L, Bowen Melissa M, Emery William J. 2002. Mapping mesoscale currents by optimal interpolation of satellite radiometer and altimeter data. *Ocean Dynamics* **52:** 95-103.

Wu W, Rodi W, Wnka Th. 2000. 3D numerical modeling of flow and sediment transport in open channels. *ASCE Journal of Hydraulic Engineering* **126:** 4-15

Young Dae S. 1987. Random vectors and spatial analysis by geostatistics for geotechnical applications. *Mathematical Geology* **19:** 467-479.

Zar Jerold H. 1996. *Biostatistical Analysis.* Prentice-Hall, Inc, New Jersey.

Matlab tutorial. http://www.cs.wvu.edu/~trapp/wvumatlab.htm

NRCan Magnetic Declination Calculator.
http://gsc.nrcan.gc.ca/geomag/field/mdcalc_e.php

APPENDIX

Matlab codes used for interpolation by optimal vector interpolation method

```
%Load measured field (Y)
load depthavgvelmatrixds.mat

eastingst=depthavgvelmatrixds(:,1);
northingst=depthavgvelmatrixds(:,2);
ust=depthavgvelmatrixds(:,3)*1000;
vst=depthavgvelmatrixds(:,4)*1000;


mineastingst=445168.813;
maxeastingst=445418.813;
minnorthingst=5033707.943;
maxnorthingst=5034007.943;



k=1
for x=1:length(ust)
   if eastingst(x)>mineastingst & eastingst(x)<maxeastingst&
northingst(x)>minnorthingst& northingst(x)<maxnorthingst
      u(k)=ust(x);
      v(k)=vst(x);
      easting(k)=eastingst(x);
      northing(k)=northingst(x);
      k=k+1;
   end
end




%Load a priori data field (X)
load estimatedfield_1mspace.mat
estimatedfieldmatrixds=[utmegood',utmngood',field1eastgood',field1northgood'];
eastingestst=estimatedfieldmatrixds(:,1);
northingestst=estimatedfieldmatrixds(:,2);
ueststst=estimatedfieldmatrixds(:,3)*1000;
vestst=estimatedfieldmatrixds(:,4)*1000;
xestst=[ueststst,vestst];
n=length(ueststst);
```

```
k=1;
for x=1:length(ueststst)
   if eastingeststst(x)>mineastingst & eastingeststst(x)<maxeastingst&
northingeststst(x)>minnorthingst& northingeststst(x)<maxnorthingst
      uest(k)=ueststst(x);
      vest(k)=veststst(x);
      eastingest(k)=eastingeststst(x);
      northingest(k)=northingeststst(x);
      k=k+1;
   end
end



uestmean=mean(uest);
vestmean=mean(vest);
uestnomean=uest-uestmean;
vestnomean=vest-vestmean;
xnomean=[uestnomean',vestnomean'];



mineastingeststst=445168.813;
maxeastingeststst=445418.813;
minnorthingeststst=5033707.943;
maxnorthingeststst=5034007.943;







step=5



numberofblocksinY=0;

countY=0;

h=1;
```

```
for  minnorthing=minnorthingst:step:maxnorthingst-step
   for  mineasting=mineastingst:step:maxeastingst-step
      for k=1:length(u)
         if easting(k)>mineasting &easting(k)<mineasting+step &
northing(k)>minnorthing & northing(k)<minnorthing+step
            countY=countY+1;

         end
      end

         if countY>=60
         nvy(h)=countY;
         h=h+1;
      numberofblocksinY=numberofblocksinY+1;
      countY=0;
         else
            countY=0;
      end
      end
end



numberofblocksinX=0;

countX=0;

h1=1;
for  minnorthing=minnorthingst:step:maxnorthingst-step
   for  mineasting=mineastingst:step:maxeastingst-step
      for k=1:length(uestnomean)
         if eastingest(k)>mineasting &eastingest(k)<mineasting+step &
northingest(k)>minnorthing & northingest(k)<minnorthing+step

            countX=countX+1;
         end
      end

     if countX~=0

         nvx(h1)=countX;
         h1=h1+1;
         numberofblocksinX=numberofblocksinX+1;
          countX=0;
      else
         numberofblocksinX=numberofblocksinX+1;
```

```
        countX=0;
      end

      end
end



%sigmaYY matrix befor demeaning the field Y
C1(1:numberofblocksinY*2,1:numberofblocksinY*2)=zeros(2*numberofblocksinY,2*nu
mberofblocksinY);
num=1;
 n=0;
m=0;
unomeangood1=0;
vnomeangood1=0;
unomeangood2=0;
vnomeangood2=0;


for  minnorthing1=minnorthingst:step:maxnorthingst-step
   for  mineasting1=mineastingst:step:maxeastingst-step

    ugood1=0;
    vgood1=0;

      e=1;
  count1=0;
      for k=1:length(u)
         if easting(k)>mineasting1 &easting(k)<mineasting1+step &
northing(k)>minnorthing1 & northing(k)<minnorthing1+step


         ugood1(e)=u(k);
         vgood1(e)=v(k);
         eastingood1(e)=easting(k);
          northingood1(e)=northing(k);

         e=e+1;
        count1=count1+1 ;
         end
      end



      if count1>=60
```

```
        n=n+1;

m=0;
  if n<=numberofblocksinY
  for k=1:count1
            uplot(num)=ugood1(k);
            vplot(num)=vgood1(k);
            eastingplot(num)=eastingood1(k);
            northingplot(num)=northingood1(k);
            num=num+1;
           end

    for minnorthing2=minnorthingst:step:maxnorthingst-step;
      for mineasting2=mineastingst:step:maxeastingst-step;

          ugood2=0;
          vgood2=0;

          f=1;
          count2=0;
          for l=1:length(u)
          if easting(l)>mineasting2 &easting(l)<mineasting2+step &
northing(l)>minnorthing2 & northing(l)<minnorthing2+step


          ugood2(f)=u(l);
          vgood2(f)=v(l);


          f=f+1;
         count2=count2+1;
          end
          end




          if count2>=60

             m=m+1;
  if m<=numberofblocksinY



for k=1:60
```

```
    C1(n*2-1:n*2,m*2-1:m*2)=C1(n*2-1:n*2,m*2-1:m*2)+[ugood1(k)*ugood2(k)
ugood1(k)*vgood2(k) ;vgood1(k)*ugood2(k) vgood1(k)*vgood2(k)];
end

if length(ugood1)==1
   C1(n*2-1:n*2,m*2-1:m*2)=C1(n*2-1:n*2,m*2-1:m*2)./length(uestnomeangood) ;

else
  C1(n*2-1:n*2,m*2-1:m*2)=C1(n*2-1:n*2,m*2-1:m*2)./(60-1) ;

end

   end
    end
  end
 end
  end
end
   end
end



umean=mean(uplot);
vmean=mean(vplot);
unomean=uplot-umean;
vnomean=vplot-vmean;
ynomean=[unomean',vnomean', eastingplot', northingplot'];

umean1=0;
umean2=0;
vmean1=0;
vmean2=0;

%sigmaYY matrix after demeaning field Y
C(1:numberofblocksinY*2,1:numberofblocksinY*2)=zeros(2*numberofblocksinY,2*nu
mberofblocksinY);

n=0;
m=0;
unomeangood1=0;
vnomeangood1=0;
unomeangood2=0;
vnomeangood2=0;
```

```
for  minnorthing1=minnorthingst:step:maxnorthingst-step
   for  mineasting1=mineastingst:step:maxeastingst-step

    unomeangood1=0;
    vnomeangood1=0;

      e=1;
  count1=0;
      for k=1:length(unomean)
         if eastingplot(k)>mineasting1 &eastingplot(k)<mineasting1+step &
northingplot(k)>minnorthing1 & northingplot(k)<minnorthing1+step


         unomeangood1(e)=unomean(k);
         vnomeangood1(e)=vnomean(k);

         e=e+1;
        count1=count1+1 ;
         end
      end



     if count1>=60
        n=n+1;

m=0;
   if n<=numberofblocksinY


    for minnorthing2=minnorthingst:step:maxnorthingst-step;
      for mineasting2=mineastingst:step:maxeastingst-step;

         unomeangood2=0;
         vnomeangood2=0;

         f=1;
         count2=0;
         for l=1:length(unomean)
         if eastingplot(l)>mineasting2 &eastingplot(l)<mineasting2+step &
northingplot(l)>minnorthing2 & northingplot(l)<minnorthing2+step


         unomeangood2(f)=unomean(l);
         vnomeangood2(f)=vnomean(l);
```

```
          f=f+1;
         count2=count2+1;
          end
          end



          if count2>=60
             m=m+1;
   if m<=numberofblocksinY
for k=1:60
    C(n*2-1:n*2,m*2-1:m*2)=C(n*2-1:n*2,m*2-
1:m*2)+[unomeangood1(k)*unomeangood2(k)  unomeangood1(k)*vnomeangood2(k)
;vnomeangood1(k)*unomeangood2(k) vnomeangood1(k)*vnomeangood2(k)];
end
if length(unomeangood1)==1
   C(n*2-1:n*2,m*2-1:m*2)=C(n*2-1:n*2,m*2-1:m*2)./length(unomeangood1) ;
else
  C(n*2-1:n*2,m*2-1:m*2)=C(n*2-1:n*2,m*2-1:m*2)./(60-1) ;
end
   end
    end
  end
 end
  end
end
   end
end


%sigmaXY matrix
B(1:numberofblocksinX*2,1:numberofblocksinY*2)=zeros(2*numberofblocksinX,2*nu
mberofblocksinY);
num=1;
 n=0;
m=0;
uestnomeangood=0;
vestnomeangood=0;
unomeangood=0;
vnomeangood=0;

umean1=0;
umean2=0;
```

```
vmean1=0;
vmean2=0;

for  minnorthingest=minnorthingeststst:step:maxnorthingeststst-step
   for  mineastingest=mineastingeststst:step:maxeastingeststst-step
    n=n+1;
    uestnomeangood=0;
    vestnomeangood=0;

      e=1;
  count1=0;
      for k=1:length(uestnomean)
         if eastingest(k)>mineastingest &eastingest(k)<mineastingest+step &
northingest(k)>minnorthingest & northingest(k)<minnorthingest+step

         uestnomeangood(e)=uestnomean(k);
         vestnomeangood(e)=vestnomean(k);

         e=e+1;
        count1=count1+1 ;
         end
      end



     if count1~=0

   if n<=numberofblocksinX

   m=1;
    for minnorthing=minnorthingst:step:maxnorthingst-step;
       for mineasting=mineastingst:step:maxeastingst-step;

          unomeangood=0;
          vnomeangood=0;

          f=1;
          count2=0;
          for l=1:length(unomean)
          if eastingplot(l)>mineasting &eastingplot(l)<mineasting+step &
northingplot(l)>minnorthing & northingplot(l)<minnorthing+step
          unomeangood(f)=unomean(l);
          vnomeangood(f)=vnomean(l);
          f=f+1;
         count2=count2+1;
          end
```

```
        end



        if count2>=60

  if m<=numberofblocksinY


for k=1:length(uestnomeangood)
    B(n*2-1:n*2,m*2-1:m*2)=B(n*2-1:n*2,m*2-
1:m*2)+[uestnomeangood(k)*unomeangood(k)  uestnomeangood(k)*vnomeangood(k)
;vestnomeangood(k)*unomeangood(k) vestnomeangood(k)*vnomeangood(k)];
end

if length(uestnomeangood)==1
  B(n*2-1:n*2,m*2-1:m*2)=B(n*2-1:n*2,m*2-1:m*2)./length(uestnomeangood) ;
else
 B(n*2-1:n*2,m*2-1:m*2)=B(n*2-1:n*2,m*2-1:m*2)./(length(uestnomeangood)-1) ;
end
  m=m+1;
  end
    end
 end
  end
end
  end
end



p=0;
for minnorthing=minnorthingst:step:maxnorthingst-step
       for mineasting=mineastingst:step:maxeastingst-step
h=0;
usum=0;
vsum=0;
for k=1:length(unomean)
    if eastingplot(k)>mineasting &eastingplot(k)<mineasting+step &
northingplot(k)>minnorthing & northingplot(k)<minnorthing+step
       usum=usum+unomean(k);
       vsum=vsum+vnomean(k) ;
h=h+1;
    end
```

```
end

if h>=60
   p=p+1;
umeaninblock(p)=usum/h;
vmeaninblock(p)=vsum/h ;
        end
end
end


Y=[umeaninblock' vmeaninblock'];


count=1;
for k=1:length(umeaninblock)
   Yinline(count)=Y(k,1);
      count=count+1;
   Yinline(count)=Y(k,2);
      count=count+1;
end


Yinline=Yinline'


A=B*inv(C)
Xhatinline=A*Yinline;


k=1;
for count=1:2:length(Xhatinline)
  Xuhat(k)=Xhatinline(count);
  Xvhat(k)=Xhatinline(count+1);
  k=k+1;
end


Xhat=[Xuhat' Xvhat'];


k=1;
for x=1:length(Xuhat)
   if Xuhat(x)==0
      Xuhatwithmean(k)=2;
      Xvhatwithmean(k)=2;
```

```
      k=k+1;
   else
Xuhatwithmean(k)=Xuhat(x)+umean;
Xvhatwithmean(k)=Xvhat(x)+vmean;
k=k+1;
   end
end

Xhatwithmean=[Xuhatwithmean' Xvhatwithmean']
meanXuhatwithmean=mean(Xuhatwithmean)
meanXvhatwithmean=mean(Xvhatwithmean)

p=1;
for minnorthingest=minnorthingeststst:step:maxnorthingeststst-step;
        for mineastingest=mineastingeststst:step:maxeastingeststst-step;
eastinghat(p)=mineastingest+(mineastingest+step-mineastingest)/2;
northinghat(p)=minnorthingest+(minnorthingest+step-minnorthingest)/2;
p=p+1;
end
end


OI_wholefield5mspace_innerproduct_1mest_ds_demeanlater=[eastinghat; northinghat;
Xuhatwithmean; Xvhatwithmean]';
save OI_wholefield5mspace_innerproduct_1mest_ds_demeanlater
```

Matlab codes for calculating the vector correlation coefficient using Crosby method

```
file1east=input('input filename of first vector field east component in single quotes
without extension');
file1eastwithext=[file1east '.dat']
file1north=input('input filename of first vector field north component in single quotes
without extension');
file1northwithext=[file1north '.dat']
file2east=input('input filename of second vector field east component in single quotes
without extension');
file2eastwithext=[file2east '.dat']
file2north=input('input filename of second vector field north component in single quotes
without extension');
file2northwithext=[file2north '.dat']
filetemp=['load -ascii ' file1eastwithext];
eval(filetemp)
templine=['utme=' file1east '(:,1);'];
eval(templine)
templine=['utmn=' file1east '(:,2);'];
eval(templine)
filetemp=['load -ascii ' file2eastwithext];
eval(filetemp)
templine=['field2east=' file2east '(:,3);'];
eval(templine)
filetemp=['load -ascii ' file2northwithext];
eval(filetemp)
templine=['field2north=' file2north '(:,3);'];
eval(templine)

%load the 5mx5m block average raw data
load outputintenseus.mat
field1eastingmidpointtemp=outputintenseus(:,1);
field1northingmidpointtemp=outputintenseus(:,2);
field1easttemp=outputintenseus(:,3);
field1northtemp=outputintenseus(:,4);

field1eastingmidpoint=[];
field1northingmidpoint=[];

%  rearrenge matlab results
for m=1:115
   for k=1:114
      field1eastingmidpoint((k-1)*115+m)=field1eastingmidpointtemp(114*(m-1)+k);
      field1northingmidpoint((k-1)*115+m)=field1northingmidpointtemp(114*(m-1)+k);
      field1east((k-1)*115+m)=field1easttemp(114*(m-1)+k);
      field1north((k-1)*115+m)=field1northtemp(114*(m-1)+k);
```

```
  end
end

badinput=(isfinite(field1east));
for k=1:length(badinput)
  if badinput(k)==0
    field1east(k)=100;
    field1north(k)=100;
  end
end

%delete bad data -convert the large surfer blank number to NaN
k=1;
for x=1:length(field1east)
  if field1east(x)>10 | field1north(x)>10 | field2east(x)>10 | field2north(x)>10
    field1east(x)=NaN;
    field1north(x)=NaN;
    field2east(x)=NaN;
    field2north(x)=NaN;
  else
    field1eastgood(k)=field1east(x);
    field1northgood(k)=field1north(x);
    field2eastgood(k)=field2east(x);
    field2northgood(k)=field2north(x);
    field1eastingmidpointgood(k)=field1eastingmidpoint(x);
    field1northingmidpointgood(k)=field1northingmidpoint(x);
    k=k+1;
  end
end



%run vectorcorrcoeff
pv2block=vectorcorrcoef(field1eastgood,field1northgood,field2eastgood,field2northgood
)
nblock=length(field1eastgood)-length(find(isnan(field1eastgood)));
significantpv2block=9.488/nblock

numvectors=length(find(isfinite(field1east)))
```