Principal Component Analysis on Horizontal Acoustic Doppler Current Profilers Measurement

Tuy N.M. Phan, John C. Wells
Dept. of Civil Engineering
Ritsumeikan University
Kusatsu, Japan
tuy_phan@yahoo.com

Yusuke Uchiyama
Dept. of Civil Engineering
Kobe University
Kobe, Japan
uchiyama@harbor.kobe-u.ac.jp

James S. Bonner, Mohammad S. Islam, William D. Kirkey
Dept. of Civil and Environmental Engineering
Clarkson University
Potsdam, NY, USA
jbonner@clarkson.edu

Abstract— **Principal Component Analysis (PCA) is proposed on long-beam velocities of Horizontal Acoustic Doppler Current Profiler (HADCP) to reveal flow dynamics in near-bank regions and check simulation results, particularly Regional Ocean Modeling System (ROMS) performed for the Hudson River estuary (New York, USA). We here analyzed data measured by Horizontal Acoustic Doppler Current Profiler (HADCP) at West Point, on the inner bank, downstream of a bend. Results show an agreement between HADCP and ROMS in the first mode. Low-order PCA temporal coefficients have fluctuation in ebb tide which may be associated with topographically-generated eddies. In addition, it appears an asymmetry between ebb and flood in the second temporal coefficient. Tidal effect is found in lower order temporal coefficients and their spectra.**

Keywords—ADCP, beam velocity, Principal Component Analysis, ROMS

I. INTRODUCTION

Principal Component Analysis (PCA) is a common tool making predictive models and statistically analyzing data. Using the PCA, a data set of possibly correlated variables is linearly transformed into a data set of uncorrelated (orthogonal) variables called principal components. The first principal component has the largest variance in the data, followed by the second principal component and so on. The PCA thus best explains the variance in the data. It is also called the discrete Karhunen–Loève transform in signal analysis and pattern recognition, empirical orthogonal functions in meteorology flow and oceanography, or proper orthogonal decomposition in turbulent flow.

Acoustic Doppler Current Profiler (ADCP) is becoming a dominant instrument to monitor water flows in rivers, estuaries and coastal oceans. Boat-mounted ADCPs are commonly used to measure velocity field in cross sections and flow discharge. Fixed ADCPs, notably bottom-mounted ADCPs and Horizontal ADCPs which measures the horizontal velocity profile at a fixed height, are used to monitor discharge[1], and to measure wave and currents near the shore [2], [3]. Some applications of ADCPs are also found in studies of sediment transport [4], [5] and validation of numerical simulations [6], [7].

ADCPs measure a real-time 2D or 3D velocity profile using two or more acoustic beams emitted from a sensor head in different directions. ADCP transducers transmit sound at a known frequency and measure Doppler shift in frequency of echoes from scatterers in water. They, therefore, can measure only velocity parallel to the beam in depth cells or bins. If the flow is assumed to be homogenous in a particular bin of opposing beams, the velocities in an orthogonal coordinate system of the instrument are obtained as a weighted sum of beam velocities in that bin. However, the flow across the beam is not homogenous when the distance between beams is comparable to the scale of turbulent eddies. Therefore, raw beam velocities are used, instead of velocities in orthogonal coordinates, to study turbulence.

Turbulence measurement with ADCPs is relatively new. Some works can be found in estuaries [8]–[10] and laboratory scales [11]. They applied the variance technique[12] to measure Reynolds stress and turbulence kinetic energy from the variance of beam velocities of four-beam ADCPs with extra assumptions besides the assumption of homogeneity. ADCPs can measure turbulence statistics, such as turbulence intensities [13], which are determined by large-scale structures, or low-frequency phenomenon such as tides [12]. However, how to use Horizontal ADCP (HADCP) to investigate large-scale structures in near-bank regions?

We here apply the PCA technique on beam velocities of HADCP measurement to extract the most important spatial and temporal information on near-bank turbulence. Besides, we also aim to check simulation results using PCA analysis.

978-0-933957-40-4 ©2013 MTS

II. HADCP DEPLOYMENT

A 600 KHz RDI HADCP is installed on a dock near the USGS gauging station at West Point, nearly 20 m from the western bank of the Hudson River Estuary (New York State, USA) and 80 km north of the Battery, the mouth of the river (Fig. 1). It has three beams with the beam angle θ of 30 degrees, measuring the velocity on a horizontal transect of the river at a depth approximately 3m above the river bed. The heading of the HADCP is manually set to normal to the bank with the variance of 5 degrees. The velocity is recorded in 120 bins with bin size set at 1 m and the first bin of 2.12 m. The HADCP is set up to generate data in earth coordinates. Because such velocity is estimated based on the assumption of homogeneous flow on the plane parallel to the ADCP head, we reconvert the data into the raw beam data.

In our HADCP measurement, each beam velocity is considered from bins 1 to $m = 79$ because bad records appeared frequently in farther bins. One snapshot (or sample at one point in time) of 3 beam velocities can be expressed as a $M = 237$ dimensional row vector $\vec{V}$.

$$\vec{V} = [b_1(\vec{x}_1^1),...,b_1(\vec{x}_m^1),b_2(\vec{x}_1^2),...,b_2(\vec{x}_m^2),b_3(\vec{x}_1^3),...,b_3(\vec{x}_m^3)] \quad (1)$$

III. SIMULATION MODEL

ROMS (Regional Ocean Modeling System) is a hydrostatic, primitive equation ocean model that solves the Reynolds averaged Navier-Stokes equations on a horizontal orthogonal curvilinear Arakawa "C" grid and uses stretch terrain following coordinates in the vertical [14]. Previous ROMS models for the Hudson River estuary mainly focus on influences of salinity stratification on vertical mixing and residual estuarine circulation [15], [16]. We performed high-resolution ROMS, UCLA version 2008, for a reach of this river between the USGS monitoring stations at Poughkeepsie (north) and Piermont (south) to resolve thermal plume from the Indian Point power plant by adding an empirical near-field buoyant source submodel [17]. We run 30 day simulation (Oct 2, 2011 – Oct 31, 2011) with a grid configuration of 160 × 800 × 20 in the zonal, meridional and vertical respectively. The upstream boundary conditions are clamped to the in-situ discharge, temperature, salinity, and water level. The downstream boundary conditions are clamped to the in-situ temperature, salinity, and water level. All in-situ data was taken from USGS monitoring stations. The model output was compared with the USGS gauges at West Point with a reasonable agreement. In this paper, to validate the model with HADCP data, the grid around HADCP location at West Point is refined. Grid spacing is 30-40 m (Fig. 2), which is about three times smaller than the grid spacing presented in Uchiyama et al [17].

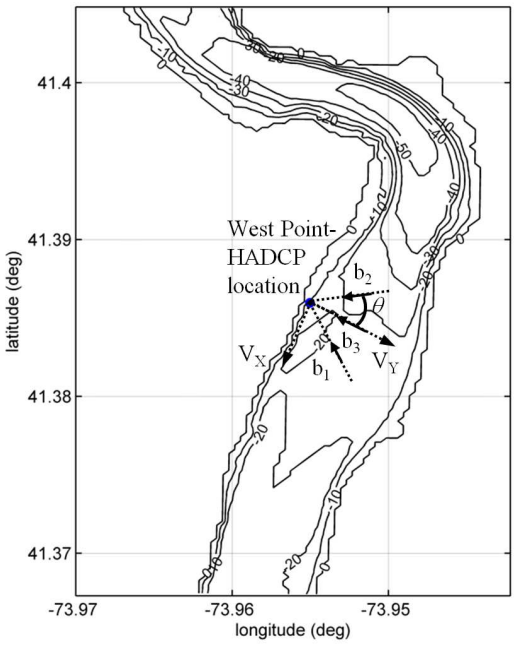

Fig. 1. A reach of the Hudson River around a three-beam HADCP at West Point. Depth contour interval is 10 m

IV. PRINCIPAL COMPONENT ANALYSIS

PCA is applied to HADCP beam velocities and "virtual" beam velocities extracted from ROMS simulation results. The PCA is a technique that decomposes a spatial-temporal data set into a set of orthogonal spatial basis functions and a set of corresponding temporal coefficients. Let $\mathbf{V}(\mathbf{x},t)$ be the matrix in which each row is a single snapshot of three beam velocities $\vec{V}$. The PCA approximates $\mathbf{V}(\mathbf{x},t)$ as below

$$\mathbf{V}(\mathbf{x},t) = \sum_{k=1}^{N} \zeta_k(t)\psi_k(\mathbf{x}) \quad (2)$$

where N is the number of snapshots of three beam velocities

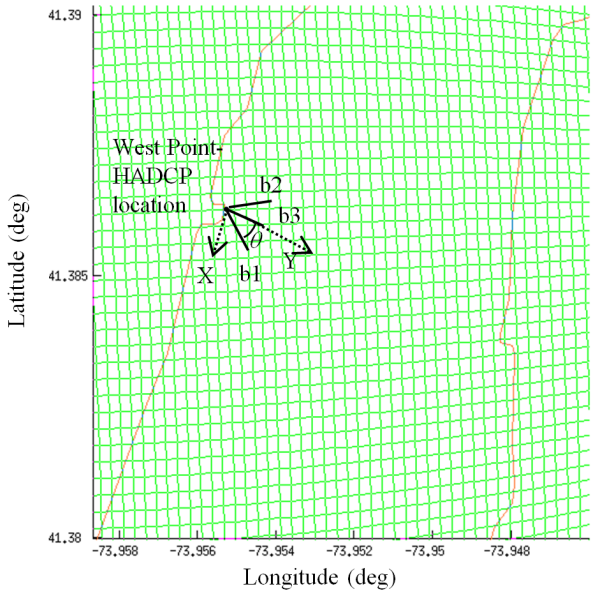

Fig. 2. ROMS grid around a three-beam HADCP at West Point

which must be larger than the measurement points ($N > M$), $\zeta_k(t)$ are called the temporal coefficients, and $\boldsymbol{\psi}_k(\mathbf{x})$ are called spatial basis functions. The PCA basis functions $\boldsymbol{\psi}_k(\mathbf{x})$ are obtained by solving an eigenvalue-eigenvector problem of the correlation matrix given as

$$\mathbf{C}_{ij} = \frac{1}{N}\int \mathbf{V}(\mathbf{x},t_i)\mathbf{V}(\mathbf{x},t_j)d\mathbf{x} \tag{3}$$

The eigenvalue λ represents the distribution of the energy, or variance, of all three beam velocities in each eigenvector. The eigenvectors and eigenvalues are sorted in the order of decreasing eigenvalues. The PCA is thus a spectral decomposition. Energy spectra are calculated as

$$St_i = \frac{\lambda_i}{\sum_{i=1}^{N}\lambda_i} \tag{4}$$

The PCA eigenvectors and temporal coefficients are calculated as

$$\boldsymbol{\psi}_k(\mathbf{x}) = \sum_{i=1}^{N}\alpha_{ki}\mathbf{V}(\mathbf{x},t_i) \tag{5a}$$

$$\zeta_k(t_j) = \int \mathbf{V}(\mathbf{x},t_j)\boldsymbol{\psi}_k(\mathbf{x})d\mathbf{x} = N\sum_{i=1}^{N}\alpha_{ki}\mathbf{C}_{ij} \tag{5b}$$

with

$$\alpha_{ki} = \frac{\upsilon_i^k}{\sqrt{N\sum_{m=1}^{N}\sum_{r=1}^{N}\upsilon_m^k\upsilon_r^k\mathbf{C}_{mr}}} \tag{6}$$

where υ_i^k is the i-th element of eigenvector $\boldsymbol{\upsilon}^k$ corresponding to the eigenvalue λ_k of the correlation matrix $\mathbf{C}_{ij}$

V. Results and Discussion

Time series of beam velocities measured by HADCP and spatially interpolated from ROMS at bin 79, corresponding to Y = 80.12 m from the HADCP location, are shown in Fig. 3. The datasets, mainly fluctuating with tides, were obtained during 30 days (Oct 2, 2011 – Oct 31, 2011), approximately 2 spring-neap cycles. HADCP dataset is 5-minute time-averaged, and ROMS snapshots are instantaneous with sample interval of 5 minutes. Note that the beam velocity is positive when beam velocity vector is directed to the HADCP. The velocity along beam 1 thus becomes positive during flood tides. Root-mean

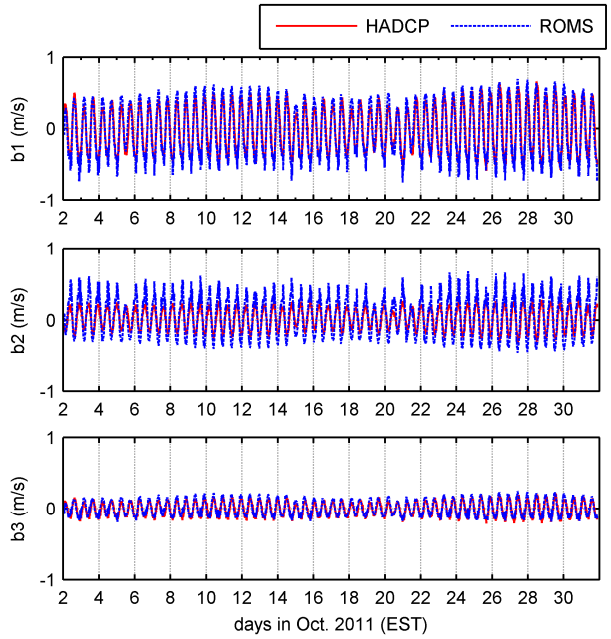

Fig. 4. Beam velocities measured by HADCP (red) and extracted from ROMS (blue) at bin 79, Y = 80.12 m. Sample interval is 5 minutes.

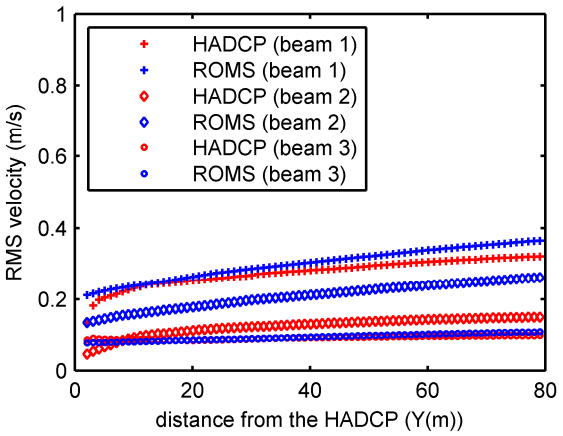

Fig. 3. Root-mean square velocity measured by HADCP (red) and extracted from ROMS (blue) along beam 1 (cross), beam 2 (diamond) and beam 3 (circle) during 30 days (Oct 2-31, 2011).

square velocities along beam 2 extracted from the ROMS are found to be twice as high as those measured by the HADCP (Fig. 4). Meanwhile, root-mean square velocities along beam 3, nearly normal to the river bank, appear a good agreement between two datasets.

Cumulative energy of PCA on two datasets of beam velocities measured by HADCP and extracted from ROMS shown in Fig. 3 are displayed in Fig. 5. The first PCA mode contains more than 97% energy of the flow along three beams. The first two PCA modes capture 99% of energy. It means that, using only these modes, the mean-squared error of reconstruction is about 1%. They thus can reveal statistically

essential dynamics of the flow measured along three HADCP beams.

Fig. 6 shows the first two PCA spatial basis functions of three beam velocities measured by HADCP and spatially interpolated from ROMS. Both have a reasonable agreement in the first PCA spatial structure, which is closer to the tidal flow. Discrepancy in higher-order spatial basis functions between two datasets clearly appears near the bank. Spatial functions of beam velocities interpolated from ROMS look nearly straight because grid spacing is coarse compared with that of the real HADCP. The second PCA spatial functions of three HADCP beam velocities change sign and have a zero-crossing around the location of 40 m from the HADCP location. Note that velocity along a beam becomes zero when the flow is normal to the beam. Because the beams are not parallel, if beam velocities are zero at the same location, vortices might occur there.

The first two PCA coefficients of both HADCP and ROMS have fluctuation in ebb tide which may be associated with topographically-generated eddies (Fig. 7). However, the ROMS yields more fluctuating coefficients than the HADCP does. It may be that the ROMS data is instantaneous and HADCP data is time-averaged. Besides, the second PCA temporal evolution has an asymmetry between ebb and flood. It might be due to the north-south asymmetry of the river bathymetry (Fig. 1). In ebb tide, high-speed water comes from the deep main channel and spreading spanwise to the bank. In flood tide, the upstream flow (to the south) comes from the

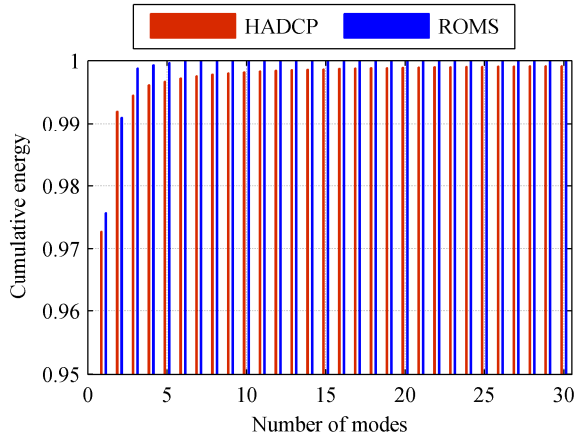

Fig. 5. Cumulative energy of PCA on beam velocities in measured by HADCP (red) and extracted from ROMS (blue) during 30 days (Oct 2-31, 2011).

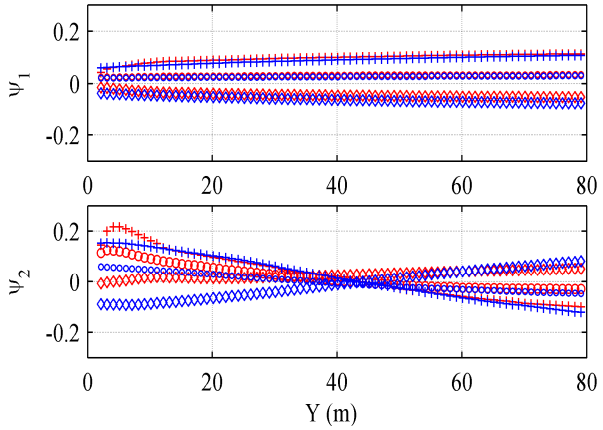

Fig. 6. The first two PCA basis functions of velocities measured by HADCP (red) and extracted from ROMS (blue) along beam 1 (cross), beam 2 (diamond) and beam 3 (circle) during 30 days (Oct 2-31, 2011).

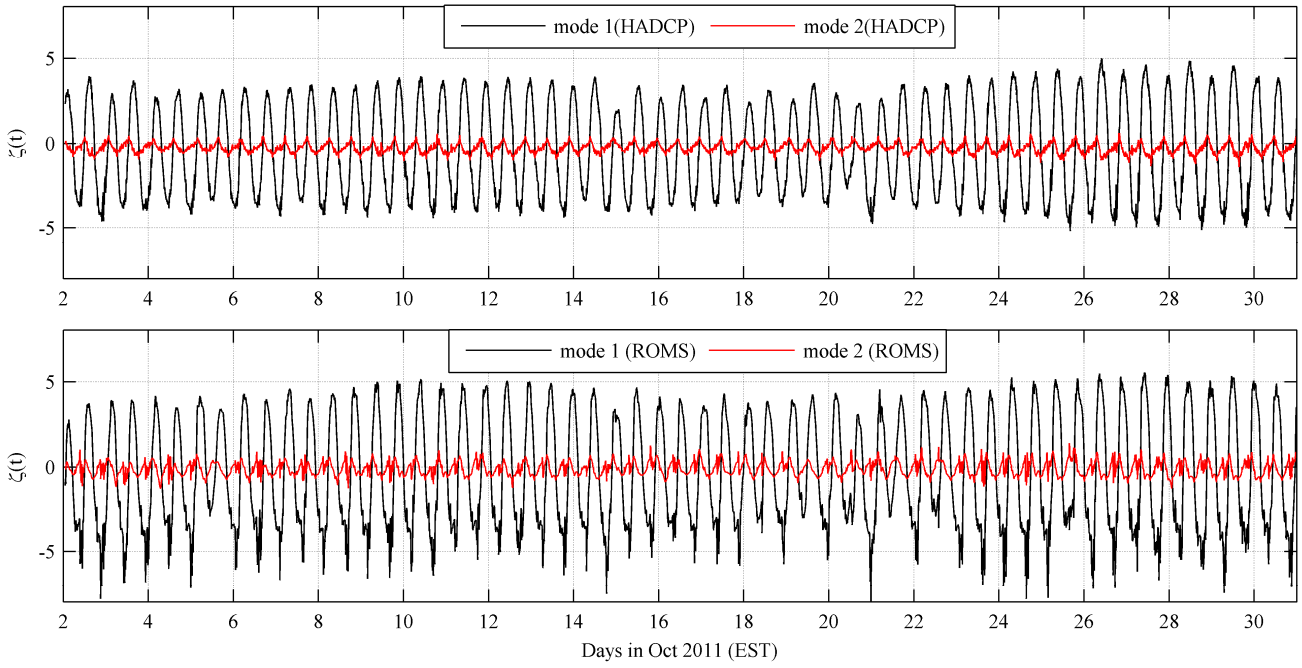

Fig. 7. The first three PCA temporal coefficients of beam velocities measured by HADCP (upper) and extracted from ROMS (lower).

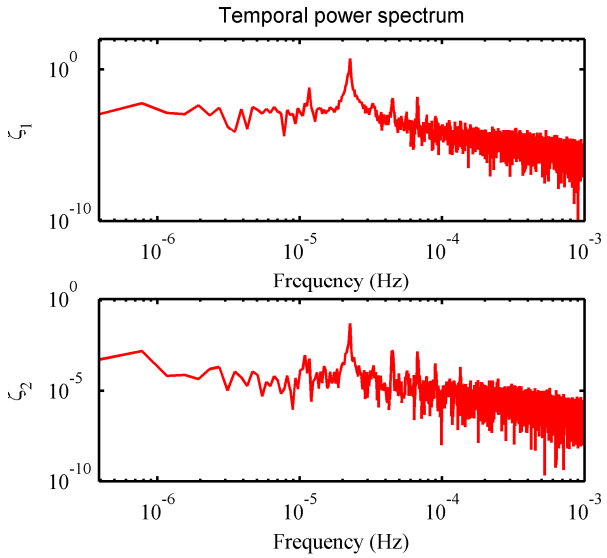

Fig. 8. Temporal power spectrum of the first two PCA temporal coefficients of HADCP beam velocities

shallow region. A peak is found in temporal power spectrum of the low-order PCA temporal coefficients of HADCP beam velocities at frequency corresponding to the semidiurnal tidal frequency of 2.23×10^{-5} Hz (Fig.8).

VI. Conclusion

PCA has been considered as a tool to compare most statistical characteristics of the flow from HADCP measurement and ROMS model. Comparison results shows that both HADCP and ROMS have a best match in the first PCA mode which captures more than 97% energy of the flow along three beam velocities. This mode approximates to the tidal flow. The second PCA temporal coefficient shows an asymmetry between ebb and flood that might be due to the north-south asymmetry in bathymetry around the HADCP location. Low-order PCA temporal coefficients of both HADCP and ROMS have fluctuation in ebb tide which may be associated with topographically-generated eddies. Moreover, the PCA coefficients of ROMS more fluctuate than those of HADCP. The first reason may be that HADCP dataset is time-averaged, and ROMS is instantaneous. The second may be that ROMS uses constant horizontal viscosity coefficients which can be scaled by grid size. Currently, they are set to be zero. In future, we will perform ROMS with some different horizontal viscosity coefficients and compare with HADCP data.

References

[1] Y. Nihei and A. Kimizu, "A new monitoring system for river discharge with horizontal acoustic Doppler current profiler measurements and river flow simulation," Water Resour. Res., vol. 44, p. 15 PP., Dec. 2008.

[2] P. A. Work, "Nearshore directional wave measurements by surface-following buoy and acoustic Doppler current profiler," Ocean Eng., vol. 35, no. 8–9, pp. 727–737, Jun. 2008.

[3] M. A. de Schipper, R. C. de Zeeuw, S. de Vries, M. J. F. Stive, and J. Terwindt, "Horizontal ADCP measurements of waves and currents in the very nearshore," in *Current, Waves and Turbulence Measurements (CWTM), 2011 IEEE/OES 10th*, 2011, pp. 159 –166.

[4] G. Wall, E. Nystrom, and S. Litten, "Suspended Sediment Transport in the Freshwater Reach of the Hudson River Estuary in Eastern New York," Estuaries Coasts, vol. 31, no. 3, pp. 542–553, 2008.

[5] S. A. Moore, J. Le Coz, D. Hurther, and A. Paquier, "On the application of horizontal ADCPs to suspended sediment transport surveys in rivers," Cont. Shelf Res., vol. 46, pp. 50–63, Sep. 2012.

[6] S. Baranya and J. Jozsa, "Flow analysis in river Danube by field measurement and 3D CFD turbulence modelling," Period. Polytech. Civ. Eng., vol. 50, no. 1, pp. 57–68, 2006.

[7] B. Castelle, P. Bretel, S. Morisset, P. Bonneton, N. Bonneton, M. Tissier, C. Sotin, A. Nahon, N. Bruneau, J. P. Parisot, and others, "Rip current system over strong alongshore non-uniformities: on the use of HADCP for model validation," J. Coast. Res. SI, vol. 56, pp. 1746–1750, 2009.

[8] A. E. Gargett, "Observing Turbulence with a Modified Acoustic Doppler Current Profiler," J. Atmospheric Ocean. Technol., vol. 11, no. 6, pp. 1592–1610, Dec. 1994.

[9] M. T. Stacey, S. G. Monismith, and J. R. Burau, "Observations of Turbulence in a Partially Stratified Estuary," J. Phys. Ocean., vol. 29, no. 8, pp. 1950–1970, Aug. 1999.

[10] E. Osalusi, J. Side, and R. Harris, "Reynolds stress and turbulence estimates in bottom boundary layer of Fall of Warness," Int. Commun. Heat Mass Transf., vol. 36, no. 5, pp. 412–421, May 2009.

[11] E. A. Nystrom, C. R. Rehmann, and K. A. Oberg, "Evaluation of mean velocity and turbulence measurements with ADCPs," J. Hydraul. Eng., vol. 133, no. 12, pp. 1310–1318, 2007.

[12] A. Lohrmann, B. Hackett, and L. P. Røed, "High Resolution Measurements of Turbulence, Velocity and Stress Using a Pulse-to-Pulse Coherent Sonar," J. Atmospheric Ocean. Technol., vol. 7, no. 1, pp. 19–37, Feb. 1990.

[13] D. K. Barua and K. H. Rahman, "Some aspects of turbulent flow structure in large alluvial rivers," J. Hydraul. Res., vol. 36, no. 2, pp. 235–252, 1998.

[14] A. F. Shchepetkin and J. C. McWilliams, "The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model," Ocean Model., vol. 9, no. 4, pp. 347–404, 2005.

[15] J. C. Warner, W. R. Geyer, and J. A. Lerczak, "Numerical modeling of an estuary: A comprehensive skill assessment," J. Geophys. Res., vol. 110, p. 13 PP., May 2005.

[16] M. E. Scully, W. R. Geyer, and J. A. Lerczak, "The influence of lateral advection on the residual estuarine circulation: A numerical modeling study of the Hudson River estuary," J. Phys. Ocean., vol. 39, no. 1, pp. 107–124, 2009.

[17] Y. Uchiyama, S. Ishii, P. Tuy, J. C. Wells, W. D. Kirkey, M. S. Islam, and J. S. Bonner, "Transient thermal plume dispersal in the Hudson River Estuary," presented at the 2012 ASLO Aquatic Sciences Meeting, Shiga, Japan, 2012.