

Spatial-scale partitioning of in situ turbulent flow data over a pebble cluster in a gravel-bed river

R. W. Jay Lacey,¹ Pierre Legendre,² and André G. Roy¹

Received 15 March 2006; revised 12 August 2006; accepted 6 November 2006; published 14 March 2007.

[1] Spatial structures within turbulent flow data were investigated through the use of a new multivariate variation partitioning analysis technique involving principal coordinates of neighbor matrices (PCNM), which is a form of distance-based eigenvector maps (DBEM). The analysis revealed a significant ($\alpha = 0.01$) spatial dependence, 58%, for the mean and turbulent flow variables. The flow variables were obtained from instantaneous two-dimensional velocities collected in situ along a streamwise section that crosses over a pebble cluster in a gravel-bed river. Using the orthogonal property of the PCNM variables, the explained variation was partitioned over four significant ($\alpha = 0.01$) spatial scales: very large (VL, 17%), large (L, 24%), medium (M, 6%) and fine (F, 2%). Nearly 75% of the variance of the main turbulent flow indicators, such as the root-mean-square of the streamwise and vertical velocity components and the mean uv component Reynolds shear stress, was explained by the VL- and L-scale PCNM submodels, which have streamwise and vertical length scales of the order of $\Delta x = 5.3H - 2.6H$ and $\Delta y = 1.0H - 0.5H$ (where H is the flow depth), respectively. Through a multivariate mapping procedure, clear spatial patterns within the explained flow variables emerge around the cluster, where the flow separation zone seems to play a significant role at a range of scales. As well, intervariable correlations at each spatial scale, obtained through eigenvector scatterplots, show intricate relationships between the flow variables. The interdependence of the Reynolds shear stress and the u component turbulent energy is much stronger at the VL scale than at the L and M scales. The application of PCNM analysis on the turbulent flow field shows the power of the technique to resolve the relevant spatial scales and patterns, and demonstrates its potential use in a variety of water resources studies.

Citation: Lacey, R. W. J., P. Legendre, and A. G. Roy (2007), Spatial-scale partitioning of in situ turbulent flow data over a pebble cluster in a gravel-bed river, *Water Resour. Res.*, 43, W03416, doi:10.1029/2006WR005044.

1. Introduction

[2] A recent trend in turbulent flow research is to attempt to identify and characterize the temporal and spatial scales of coherent flow structures over smooth as well as rough beds [Adrian *et al.*, 2000; Liu *et al.*, 2001; Shvidchenko and Pender, 2001; Chen and Hu, 2003; Roy *et al.*, 2004]. Space-time correlations of high-frequency velocity time series have shown the existence of large depth-scaled coherent flow structures consisting of high- and low-speed wedges [Nakagawa and Nezu, 1981; Buffin-Bélanger *et al.*, 2000; Roy *et al.*, 2004]. An alternative method of investigating the spatial scales of turbulence is through the analysis of the spatial distribution of the mean turbulent parameters of interest. Following this approach, a few studies have investigated the spatial distribution of turbulence properties over and around large roughness elements such as isolated clasts

or pebble clusters [Brayshaw *et al.*, 1983; Paola *et al.*, 1986; Buffin-Bélanger and Roy, 1998; Lawless and Robert, 2001; Shamloo *et al.*, 2001; Tritico and Hotchkiss, 2005]. In natural gravel-bed rivers, the bed is made up of poorly sorted clasts ranging in size from very coarse to fine grains, which results in abrupt roughness transitions [Robert *et al.*, 1992]. These rapid changes in roughness have a direct effect on the turbulent length scales and on the spatial patterns of the mean and turbulent flow properties [Buffin-Bélanger and Roy, 1998]. The mechanism responsible for these changes is the shedding of vortices in the lee of protuberant particles. Such vortices have a range of sizes and frequencies [Kirkbride, 1993]. This suggests a high dependence of the spatial pattern of turbulent flow structures on the distribution of large clasts and of bed forms such as clusters on the heterogeneous bed.

[3] The relationships between spatial patterns of turbulent structure and large roughness elements on the bed were investigated by Buffin-Bélanger and Roy [1998]. Through an intense measurement scheme around a pebble cluster, they were able to delineate six characteristic regions of the flow field (acceleration, recirculation, shedding, reattachment, upwelling, and recovering flow), and showed the relationships between these regions and the protuberant

¹Département de Géographie, Université de Montréal, Montréal, Québec, Canada.

²Département de Sciences Biologiques, Université de Montréal, Montréal, Québec, Canada.

clasts. While their study provided qualitative descriptions of numerous flow variables and of their spatial patterns, it did not attempt to quantitatively explain the dependence of the flow variables on the spatial structure, and did not estimate the proportion of the variation within the flow variables explained by the spatial structures.

[4] In ecological studies, quantification of spatial structure is often obtained through trend surface analysis. This is a standard method used to explain the variance associated with spatial trends in variables measured at points in space through polynomial regressions [Legendre and Legendre, 1998]. The higher the polynomial order, the finer the spatial structures which can be explained. Yet the terms within the polynomial are often highly correlated with one another, which prevents the modeling of linearly independent structures at different scales [Borcard and Legendre, 2002]. Furthermore, trend surface analysis is devised to model large-scale spatial structures with simple shapes and cannot adequately model finer structures [Borcard and Legendre, 2002; Borcard et al., 2004].

[5] Borcard and Legendre [2002] have recently developed a new statistical spatial modeling method: principal coordinates of neighbor matrices (PCNM). The method, the theory for which has been further explored by Dray et al. [2006], is a form of distance-based eigenvector maps (DBEM), and has been successfully applied in aquatic ecological studies to describe the dominant spatial scales at which species variation occurs [Borcard et al., 2004; Brind'Amour et al., 2005]. PCNM analysis resembles Fourier analysis and harmonic regression but has the advantage of providing a broader range of signals and can also be used with irregularly spaced data [Borcard and Legendre, 2002]. PCNM analysis is based on the orthogonal spectral decomposition of the relationships among the spatial coordinates of a sampling design [Borcard et al., 2004]. The orthogonal property of the PCNM technique allows an exact partitioning (no intercorrelation) of the explained variance over different spatial scales. PCNM analysis is used in conjunction with multiple regression to study the spatial structure of a single variable, or with canonical redundancy analysis (RDA) when studying the spatial structure of multiple variables. RDA is an extension of multiple regression used to model multivariate data. It is based on the eigenvalue decomposition of the table of regressed fitted values, which reduces the large number of associated (linearly correlated) fitted vectors to a smaller composite of linearly independent variables [Legendre and Legendre, 1998]. With eigenvalue decomposition, most of the variability is often summarized in the first few dimensions, which facilitates interpretation. Eigenvalue analysis has been used to study turbulent coherent structures through proper orthogonal decomposition (POD) [Liu et al., 2001], and is used extensively in ecology with data sets which include large numbers (hundreds, thousands) of interrelated variables.

[6] PCNM analysis bears some similarity to POD. For POD, the eigenvalue decomposition is performed directly on a two-point correlation coefficient matrix of the flow variable under investigation using Fourier modes which are sinusoidal (quasi-trigonometric) eigenvectors termed eigenfunctions [Moin and Moser, 1989; Berkooz et al., 1993]. As POD is a direct eigenvalue decomposition of the flow variable correlation matrix, the sum of the eigenvalues

is equal to the total variance of the flow variable matrix. PCNM analysis is a regression technique which identifies only the fraction of the total variance in a response variable that is spatially dependent. An advantage of the PCNM technique is that the PCNM variables represent the eigenvalue decomposition of the relationships of a specific sampling grid, and can be used to analyze irregularly spaced data with nonrectangular boundaries. PCNM analysis is, as well, a multivariate regression technique; it offers the added advantage over POD (which can only analyze a single variable at a time) of allowing the analysis of all response variables at once.

[7] Buffin-Bélanger and Roy [1998] investigated each flow parameter, or ratio of flow parameters individually, an approach common in studies investigating the turbulent structure of flows [Bennett and Best, 1995; Lawless and Robert, 2001]. This approach used for investigating spatial patterns of flow structure could be greatly improved using PCNM and RDA, which can identify and quantify the spatial dependence of all flow parameters at once, thus providing an efficient means of summarizing and interpreting the spatial patterns. This paper examines the potential use of PCNM analysis as a statistical tool for investigating the spatial-scale dependence of turbulent flow processes as a complement to traditional analyses. The paper revisits the turbulent flow data reported by Buffin-Bélanger and Roy [1998] adjacent to and overtop of a pebble cluster in a gravel-bed river. Our study furthers the previous work by explaining and partitioning the variance of the flow parameters over four spatial scales, providing a quantification of the spatially explained variance, and indicating the intercorrelations among the turbulence variables at each scale. This leads to new insights into the turbulent flow field around clusters and protuberant clasts in rivers by suggesting the appropriate scale dependence of the turbulent flow variables, and demonstrates the potential use of PCNM analysis for a wider field of application in water resources studies.

2. Field Measurements and Turbulent Statistics

[8] The collection and processing of the instantaneous velocity measurements used in our study was described in detail by Buffin-Bélanger and Roy [1998] and are briefly summarized here. Velocity measurements were collected from the Eaton North River, Quebec, Canada, on a streamwise x , – vertical y , transect plane with a mean height (H) above the bed of 0.38 m and a streamwise length of 4.0 m. The x – y plane crosses through the center of a naturally formed pebble cluster. The crest of the cluster is located at $x = 0.77$ m (Figure 1) and has a height (h_s) of 0.20 m. Electromagnetic current meters (ECMs) were used to collect instantaneous streamwise u and vertical v velocity measurements at a sampling frequency of 20 Hz. The present study consists of 29 vertical profiles (the two most upstream profiles of the original 31 were omitted due to their inconsistent separation distances). Each profile contains six to 13 vertical measurement locations (Figure 1). In total, the data set consists of 340 velocity time series of 60 s duration. Each time series corresponds to a point on the sampling grid of Figure 1. The spacing between measurements along the vertical profiles is 0.02 m (with the exception of two offset grid points at $x = 1.1$ m), while

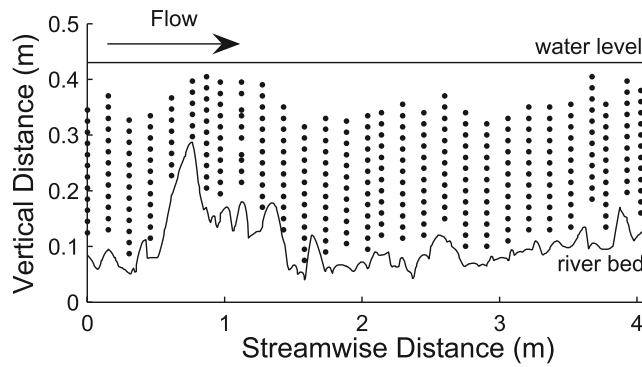


Figure 1. Sampling x - y transect plane with measurement points (dots). Modified from Buffin-Bélanger and Roy [1998].

spacing in the streamwise direction varied from 0.1 to 0.15 m.

[9] The mean downstream velocity ($\bar{U}$), where the over bar represents averaging over time, for the x - y plane is 0.28 m s^{-1} resulting in a Reynolds number Re of 8.0×10^4 , which indicates a fully turbulent regime. Buffin-Bélanger and Roy [1998] investigated and presented a set of 22 flow variables. These consisted of mean and turbulent statistics, and ratios between some of the flow variables. For the current study, we selected a subset of 15 variables to be used as response variables in the PCNM analysis. This subset was selected on the basis that they covered a range of spatial patterns and scales in order to properly test the PCNM method without introducing excessive redundancy between the variables. As such, the flow variables included mean flow statistics ($\bar{U}$, $\bar{V}$); second-order moment statistics (root-mean-square values, u' and v') which are a measure of the turbulent intensity and had been shown by Buffin-Bélanger and Roy [1998] to exhibit broad-scale spatial patterns; and the third-order moment (skewness, Sk_u and Sk_v). Skewness is a measure of the asymmetry of the velocity distribution, and it reveals the presence of high-magnitude events within the velocity signal [Buffin-Bélanger and Roy, 1998]. For example, a positive Sk_v indicates intermittent, infrequent events of vertical velocity directed toward the surface. Previous boundary layer studies have observed near-bed velocity distributions to be positively skewed [Grass, 1971].

[10] We have also included the average statistics of turbulent events such as the percent of time (T) and the frequency (f) of low-speed flow ejections ($Q2$) and high-speed sweeps ($Q4$). Event statistics are estimated by conditionally sampling the fluctuating component of the velocity signals following Lu and Willmarth [1973]. Ejections (quadrant 2) are defined by negative u and positive v excursions from the mean, while sweeps (quadrant 4) are defined by positive u and negative v excursions. Following Lu and Willmarth [1973], a threshold hole size, $T_h = |uv|/u'v'$, was used to distinguish between stronger, more energetic, events ($T_h = 2$) and all events ($T_h = 0$). The terminology for “hole” corresponds to the more quiescent contributions which are obtained by subtracting events estimated with $T_h = 2$ from those estimated with $T_h = 0$. Ejections and sweeps are a common feature of turbulent flows over smooth and rough

boundaries [Grass, 1971], and contribute to the bulk of the positive Reynolds shear stress [Williams *et al.*, 1989]. The event statistics displayed a more localized spatial pattern than the mean turbulent statistics previously mentioned. Mean Reynolds shear stress ($-\rho\bar{uv}$) and integral timescale (ITS_u , ITS_v) obtained from the autocorrelation functions of the streamwise and vertical velocity time series were also included. The mean shear stress is a measure of the mean turbulent momentum exchange at a sampling location, and ITS is a rough measure of the interval over which velocities are autocorrelated, giving an estimation of the size of turbulent coherent structures.

[11] The spatial mean values for the time-averaged turbulent statistics, event statistics, and integral timescales are presented in Table 1. Further details on the methods for estimating these variables are presented by Buffin-Bélanger and Roy [1998]. The spatial distributions of the standardized values (z-scores) of the 15 flow variables are presented in Figure 2. By investigating each subfigure as in the work by Buffin-Bélanger and Roy [1998], clear spatial patterns emerge. For instance, large patches of high u' and v' , $-\rho\bar{uv}$ are observed in the wake of the pebble cluster. The spatial patterns of $T_{Q4T_{h:0}}$ and $T_{Q2T_{h:2}}$ are patchier, with some better defined trends showing higher values in the wake of the cluster advecting toward the water surface with distance downstream from the cluster increases. While these and other general spatial patterns were described by Buffin-Bélanger and Roy [1998], they remain qualitative observations, and investigating each variable sequentially is cumbersome and does not lead to a global view where the interactions between the variables are fully exploited. To do so requires a multivariate approach that can deal simultaneously with the interrelations between the flow variables and with the spatial components of the data.

3. PCNM Statistical Analysis

[12] The PCNM technique is used to explain the spatial dependence and patterns of distributed variables over a sampling grid. The advantage of the PCNM technique is that the explained variance can be explicitly estimated for each spatial scale. The PCNM variables (PCNMs) are obtained by principal coordinate analysis (PCoA) of a

Table 1. Spatial Means and Standard Deviations of the Flow Variables

Variable	Mean	SD
$\bar{U}$, m s^{-1}	0.28	0.11
$\bar{V}$, m s^{-1}	0.00	0.01
u' , m s^{-1}	0.05	0.01
v' , m s^{-1}	0.04	0.01
$-\rho\bar{uv}$, N m^{-2}	0.60	0.66
Sk_u , $\text{m}^3 \text{s}^{-3}$	0.15	0.38
Sk_v , $\text{m}^3 \text{s}^{-3}$	0.02	0.27
$T_{Q2T_{h:0}}$, %	29.62	3.44
$T_{Q4T_{h:0}}$, %	28.58	3.82
$T_{Q2T_{h:2}}$, %	2.62	1.42
$T_{Q4T_{h:2}}$, %	2.95	1.27
$f_{Q2T_{h:2}}$, s^{-1}	0.34	0.13
$f_{Q4T_{h:2}}$, s^{-1}	0.40	0.13
ITS_u , s	0.79	0.46
ITS_v , s	0.31	0.20

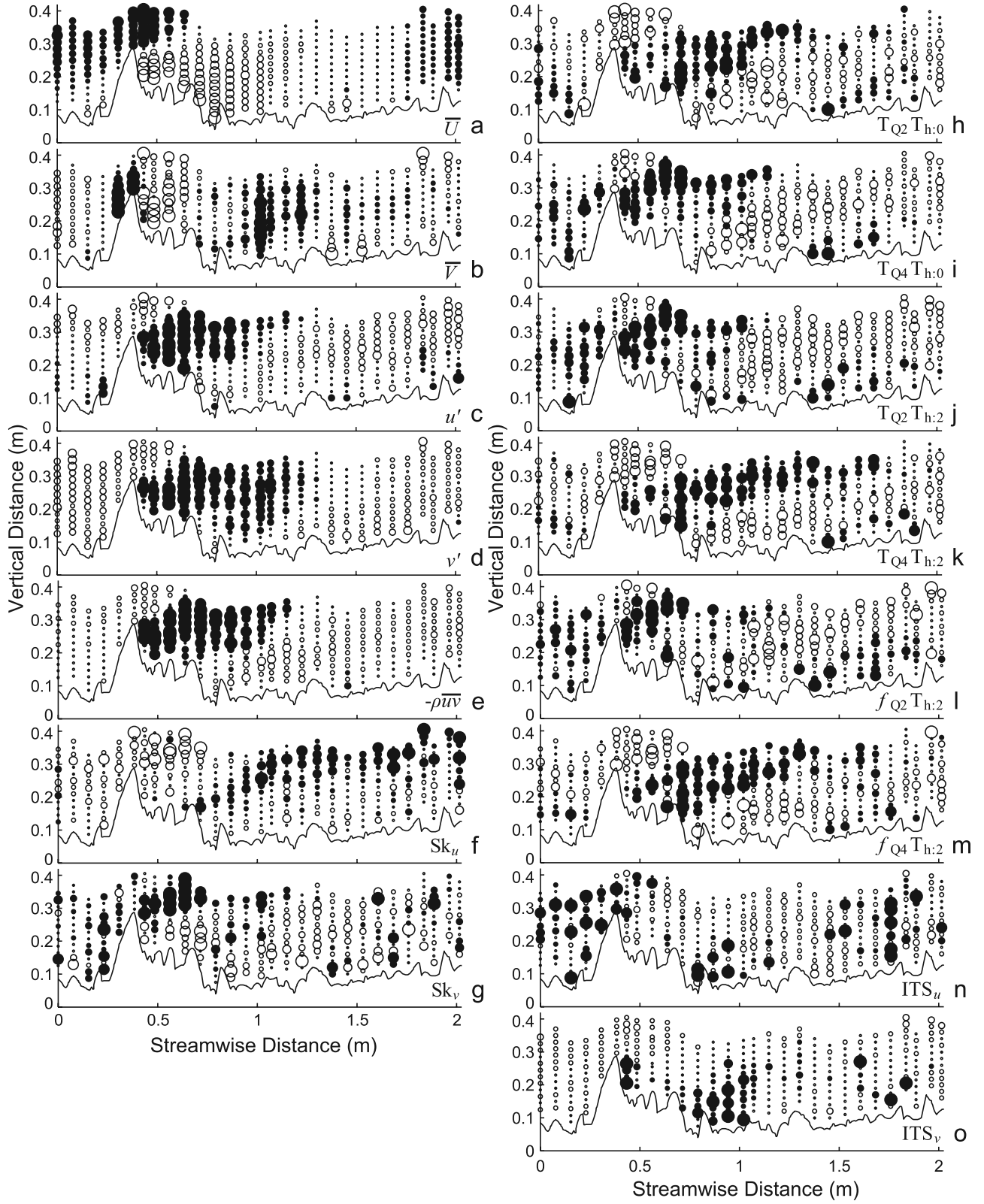


Figure 2. The x - y transect plot of standardized values (z-scores) of the mean and turbulent flow statistics: (a) $\bar{U}$, (b) $\bar{V}$, (c) u' , (d) v' , (e) $-\rho \bar{u}v$, (f) Sk_u , (g) Sk_v , (h) $T_{Q2}T_{h:0}$, (i) $T_{Q4}T_{h:0}$, (j) $T_{Q2}T_{h:2}$, (k) $T_{Q4}T_{h:2}$, (l) $f_{Q2}T_{h:2}$, (m) $f_{Q4}T_{h:2}$, (n) ITS_u , and (o) ITS_v . Positive values are represented by solid circles, and negative values are represented by open circles. Circle size indicates relative magnitude. Flow is from left to right. Modified from *Buffin-Bélanger and Roy [1998]*.

truncated pair-wise geographical Euclidian distance matrix among the sampling points. PCNM variables are thus orthogonal to one another (null scalar product). PCNMs are sinusoidal and of decreasing periods, and as such they can be grouped into submodels corresponding to different scales. Selecting the number of submodels to use and the scale associated with each one (i.e., large, medium, and fine) is a subjective process based on the objectives of the analysis (the level of detail desired) and the similarity between the significant PCNM periods. Once the submodels are constructed, they are used as explanatory variables in RDA. The sinusoidal property of the PCNM variables bears some resemblance to spectral analysis using a Fourier transformation of the autocorrelation function and wavelet analysis. Wavelet analysis is often used to decompose time series into time-frequency space to determine both the dominant modes of variability and how those modes vary in time. Similarly, PCNM analysis fits the grouped sinusoidal PCNM variables to response variables. Here spatial data are used, obtaining a decomposition describing the dominant modes of variability, as well as their spatial variation. Furthermore, PCNM analysis quantifies the fit by an R^2 statistic within each scale. RDA is more interesting than multiple regression because it can analyze several response variables in a single analysis, and display graphically their regressed interrelationships. In RDA, each canonical ordination axis corresponds to a direction in the response variable space that is maximally correlated to a linear combination of the explanatory variables. The orthogonal nature of the PCNM variables means that the variance explained by each PCNM submodel is unique and additive. The total explained variance can therefore be partitioned among the different PCNM submodels, or spatial scales.

[13] The main constraint for applying a PCNM analysis is that best results are obtained when using a uniform sampling grid with equally spaced x and y . Small irregularities in the sampling grid result in an inability to explain the finer-scale spatial structures. As the variation between sampling point distances increases, the ability of the PCNM technique to resolve the finer-scale spatial structures is compromised (i.e., increases in the sample grid irregularities result in an inability to explain larger and larger fine-scale spatial structures). In the present analysis, the vertical heights were multiplied by a factor of 7.65 to achieve a roughly regular grid pattern between the streamwise and vertical sampling points. This adjustment allowed the finer-scale variation measured in the vertical dimension to be retained and analyzed. Irregularities at the sampling grid boundaries due to the nonhomogeneous bed and the pebble cluster could not be avoided, and resulted in a loss of fine-scale spatial explanation.

[14] Figure 3 illustrates the various steps involved in producing the PCNM variables from the x - y sampling points following PCNM analysis theory [Borcard and Legendre, 2002]. A Euclidean distance (D_1) matrix was calculated for all possible distances between sampling locations (Figures 3a and 3b) using the modified coordinates. The D_1 matrix was truncated at a threshold distance (d_t) which was equal to or larger than the minimum between-site connection distance corresponding to the distance that kept all sampling locations connected together in a network. Using hierarchical, single-linkage, agglomerative

cluster analysis, the appropriate d_t was estimated to be 0.17 m. Unfortunately, because of the inherent irregularities in the x - y sampling grid that were imposed from the irregular bed topography, truncating the D_1 matrix at 0.17 m resulted in highly disrupted and distorted PCNM variables. This distortion influenced the amplitude, phase, and period of the sinusoids, thereby complicating their interpretation as the PCNMs bear structures at several scales [Borcard and Legendre, 2002]. The distortion decreased as d_t was increased, but by increasing d_t , the explanation of fine-scale variability was compromised, because inherently PCNM variables are unable to explain any spatial variance at scales less than d_t [Legendre and Borcard, 2006]. The minimum value which produced the fewest singularities was $d_t = 0.35$ m, a value approximately twice as large as the d_t calculated with the cluster analysis. Consequently, any spatial structure occurring in the flow variables at scales below this threshold ($\Delta x < 0.35$ m and $\Delta y = 0.046$ m) could not be explained by our analysis, where Δx and Δy represent the physical length scales in the streamwise and vertical directions, respectively. This constraint should be kept in mind when designing new studies.

[15] Distances between sampling points above d_t were set to a value ($d_m = 1.4$ m), which is 4 times higher than d_t (Figure 3c), in order to retain only the distances smaller than d_t between neighboring sites (sampling points) within the D_1 matrix [Borcard et al., 2004]. The eigenvalues and principal coordinates (eigenvectors) of the truncated D_1 matrix were obtained using PCoA (Figure 3d). Of the 340 eigenvalues obtained, 180 were positive. A forward selection permutation procedure from the CANOCO program [ter Braak and Smilauer, 2002] was used to determine which PCNMs explained a significant ($\alpha = 0.01$) level of variation in the flow variables. Twenty-nine significant PCNMs were identified, and they were subjectively classified into four submodels according to the scales of their respective periods: very large scale (VL), large scale (L), medium scale (M), and fine scale (F). One PCNM from each class is presented in Figure 3e. The PCNMs can be seen as a series of two-dimensional (2-D) sinusoidal curves of decreasing periods.

[16] The largest detectable scale, which is linked to the period of the first PCNM, is dictated by the spatial extent of the farthest sampling locations. For instance, when computed from a distance matrix corresponding to n equidistant objects arranged as a straight line, the largest period is equal to $n + 1$ [Borcard and Legendre, 2002]. From Figure 3e, the streamwise period of PCNM 1 can be estimated as $x \approx 4$ m, while the period of PCNM 13 is $x \approx 0.6$ m. The physical length scales, Δx and Δy , associated with each PCNM class were estimated from the mean half-period of the grouped PCNMs. Δx and Δy are presented in Table 2 along with the range and number of PCNMs included per class. As indicated in Table 2, the maximum streamwise scale is around $5.3H$, and the minimum streamwise scale is equivalent to the flow depth ($1.0H$); the mean flow depth is $H = 0.38$ m. Any streamwise spatial variation in the data occurring at scales above and below these thresholds could not be resolved by our analysis.

[17] The PCNM submodels were used as explanatory (independent) variables in multiple regressions and canonical RDA (Figure 3f) for the turbulent flow data obtained

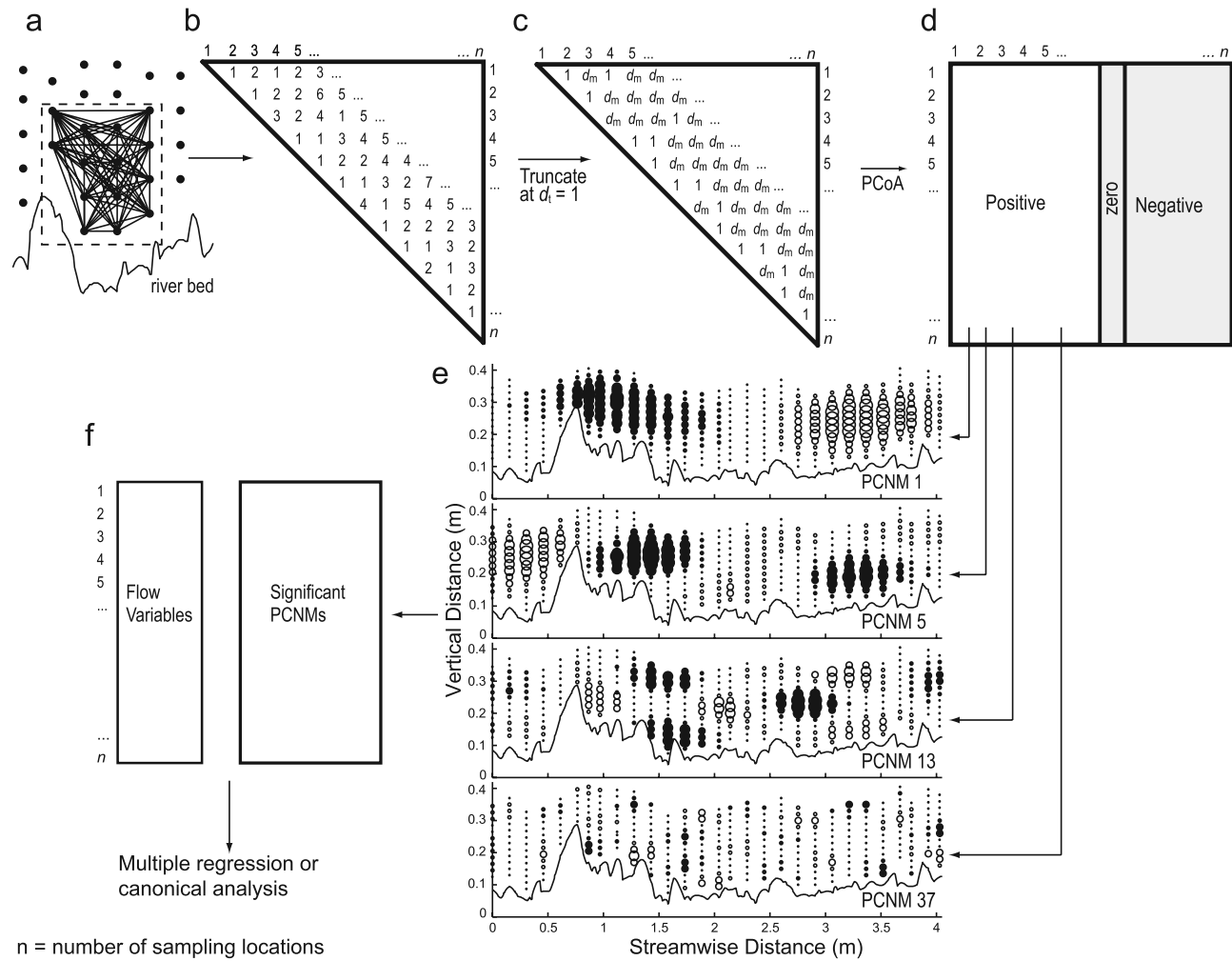


Figure 3. Methodology for developing principal coordinates of neighbor matrices (PCNM) variables. (a) Example of sampling locations with Euclidian links between coordinates. (b) Euclidian distance (D_1) matrix. (c) Truncated D_1 matrix (neighboring matrix) truncated at $d_t = 1.0$ by d_m . (d) Principal coordinate (eigenvector) matrix, indicating positive, zero, and negative eigenvalues. (e) Examples of PCNMs (positive eigenvalued principal coordinates) representing very large scale (PCNM 1), large scale (PCNM 5), medium scale (PCNM 13), and fine scale (PCNM 37). The size of the circles is proportional to the magnitude of the PCNM variable values (positive values solid). Flow is from left to right. (f) Response (flow) matrix and explanatory (PCNM) matrix used in multiple regression and canonical analysis. Modified from *Borcard and Legendre* [2002].

from *Buffin-Bélanger and Roy* [1998] and listed in Table 1. The flow variables were standardized, and both the multiple regressions and the RDA were computed using the CANOCO program [*ter Braak and Smilauer*, 2002]. The global model (containing 180 PCNMs) and all canonical axes of the submodels were tested for significance ($\alpha = 0.01$) using 999 Monte Carlo unrestricted permutations. The significant “fitted site scores” from each submodel were plotted on the sampling point coordinates, thus providing a means to interpret the spatial patterns of the results at each scale. The “fitted site scores” are the values obtained from the RDA. The term “fitted site scores” is commonly used in canonical analysis; it designates the principal components of the table of fitted values of the multiple regressions. “Fitted site scores” are calculated by multiplying the canonical eigenvectors by the fitted response variables. The spatial patterns and relative magnitude of each correlated flow

variable can be directly interpreted from these plots. Scatterplots of the RDA eigenvectors focusing on the correlations between the fitted response variables are also presented to give information on the correlations between the turbulent statistics at each submodel scale.

Table 2. Scale Classification of the Significant PCNM Variables^a

Scale	PCNM Range	Number of PCNMs	Physical Scale, m	
			$\Delta x/H$	$\Delta y/H$
Very large (VL)	1–2	2	5.3	1.0
Large (L)	3–10	8	2.6	0.5
Medium (M)	11–21	10	1.6	0.3
Fine (F)	22–58	9	1.0	0.14

^a $H = 0.38$ m is the mean flow depth.

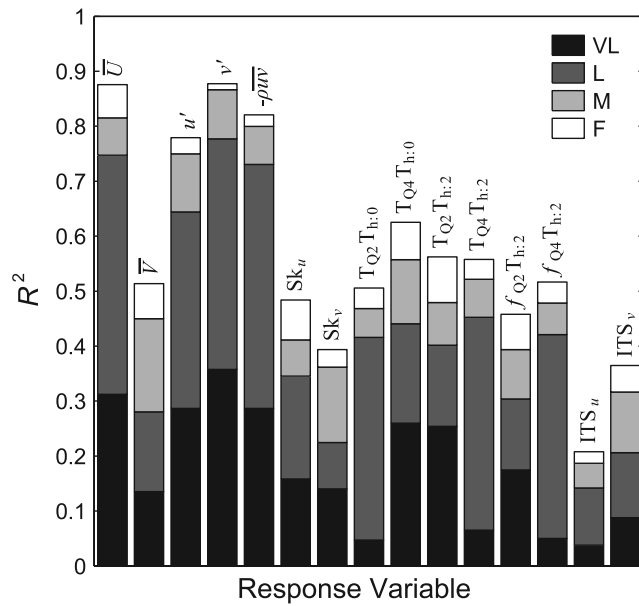


Figure 4. Fraction of explained variance (unadjusted coefficient of multiple determination, R^2) for individual mean and turbulent flow variables, for each significant PCNM submodel ($p < 0.05$): very large scale (VL), large scale (L), medium scale (M), and fine scale (F).

[18] The contribution of the independent variables (global PCNM model or PCNM submodels, in the present study) to the explanation of the response variables is given by the bivariate redundancy statistic (or canonical R^2) and its adjusted form, the adjusted bivariate redundancy statistic (or adjusted canonical R^2 , R_a^2). The adjusted form is corrected for the explanation that would be provided by the same number of random explanatory variables measured over the same number of observation points. The correction formula is the same as for the adjusted coefficient of multiple determination in multiple regression [Ezekiel, 1930]. The canonical R^2 can also be computed as the sum of the RDA canonical eigenvalues divided by the total variance in the array of standardized response variables. Because of the adjustment, the sum of the submodels R_a^2 does not equal the R_a^2 of all PCNM variables.

4. Results

[19] In the following, results of the RDA performed on the global model using all 180 PCNMs (positive eigenvalues) are discussed and compared with the explained variance obtained using a more traditional trend surface analysis. The RDA and the multiple regression results performed on the four submodels are presented. The spatial decomposition of the explained variance is presented by plotting the “fitted site scores” of each significant RDA canonical axes. The correlations between flow variables for each submodel are discussed using eigenvector scatterplots.

4.1. Global PCNM Model RDA

[20] The global RDA based on 180 PCNMs explains a significant portion of the variance of the mean and turbulent flow statistics. The adjusted bivariate redundancy

statistic, R_a^2 , is $R_a^2 = 0.58$ with an F statistic of 3.64 and an associated $p < 0.001$. The significantly large portion of the explained global variance clearly indicates the spatial dependence of the flow field parameters. For comparison purposes, an RDA was performed on the flow variables using a third-order polynomial created from the x - y sampling location coordinates. While the explained variance of the trend surface analysis was much lower ($R_a^2 = 0.34$, $F = 20.5$, $p < 0.01$) than for the PCNM analysis, this technique was still able to demonstrate the presence of a large-scale spatial pattern. Yet further interpretation is limited due to the highly correlated terms which prevent the modeling of independent structures at different scales [Borcard and Legendre, 2002].

4.2. PCNM Submodels RDA

[21] The results of the RDA performed using the four PCNM submodels indicate an unequal partitioning of the global variance between scales: VL, $R_a^2 = 0.17$; L, $R_a^2 = 0.24$; M, $R_a^2 = 0.06$; and F, $R_a^2 = 0.02$. All submodels were significant at $\alpha = 0.01$. The partitioned R_a^2 values indicate that a substantial portion of the variation of the mean and turbulent flow statistics is explained by the models at very large and large spatial scales; these scales are of the order of $\Delta x = 5.3H - 2.6H$ and $\Delta y = 1.0H - 0.5H$. The medium-scale and fine-scale submodels explain much smaller portions of the variation. These results are perhaps related to previous observations indicating that turbulent flow in gravel-bed rivers organizes itself into large depth-scaled coherent structures [Shvidchenko and Pender, 2001; Roy et al., 2004] which are surrounded by small-scale isotropic random eddies [Townsend, 1976]. Our results suggest that the turbulent flow variables contain a high spatial dependence, while at smaller scales the turbulent flow variables are more randomly distributed. The lack of spatial structure in the finer-scale submodels may also be caused by the poor resolution of finer scales due to the irregular sampling grid and the d_t used. While depth is a more commonly used variable for scaling turbulent coherent structures, the VL- and L-scale spatial scales could also be scaled by roughness element height (h_s) $\Delta x = 10h_s - 5h_s$ and $\Delta y = 2h_s - 1h_s$. The simple relationship between the VL and L scales and roughness element height supports previous work by Kirkbride [1993] suggesting the dependence of the shedding spatial patterns on bed roughness elements.

4.3. PCNM Submodel Multiple Regressions

[22] A multiple regression was conducted on each flow variable in order to isolate the response of individual flow variables by submodel. The unadjusted coefficient of multiple determination (R^2) represents the explained variation of the response variables by the PCNM submodel and provides a global account of the fit of each model. Details about the variance explained by each PCNM submodel, for each individual flow variable, are presented in Figure 4. The PCNM analysis summed over all scales explains nearly 80% of the spatial variation in $\bar{u}$, u' , v' , and $-\rho\bar{u}v$, with 75% being explained by the VL- and L-scale submodels. In other words, large-scale flow patterns are responsible for 75% of the variance in $\bar{u}$, u' , v' , and $-\rho\bar{u}v$. This finding is comparable to that of Liu et al. [2001], who found, through POD, that large-scale motions with length scales, $\Delta x > 1.6H$

and $\Delta y = 0.3H - 2H$, contained 50% of the total turbulent kinetic energy and two thirds to three quarters of the Reynolds shear stress at $Re = 5300 - 30\,000$. Their laboratory experiments were conducted on a smoothed wall rectangular channel and demonstrated the similarities of large-scale motions over smooth- and rough-walled flows. The R^2 for $\bar{U}$ is more than twice as large as that of $\bar{V}$. This is perhaps due to the influence of the pebble cluster on the v component velocity which generally has weaker spatial correlations than the u component [Nakagawa and Nezu, 1981]. The variance in $\bar{V}$ is explained in approximately equal portions by the VL, L, and M scales, indicating an equal superposition of scales within the x - y plane. Slightly more variance is explained for v' than is for u' and similarly more variance is explained for ITS_v than for ITS_u . This is seen in the z-score plots, as well (Figure 2), where the spatial patterns of v' and ITS_v are more regular than those of u' and ITS_u , respectively. While this result is counterintuitive, given the v component's weaker spatial correlations, it indicates that the v component turbulent statistics are influenced in a more spatially uniform manner by the pebble cluster than their u component counterparts.

[23] We expected the medium- and fine-scale submodels to explain a greater percentage of their variability, given that we specifically included turbulent statistics of higher moments (i.e., Sk_u) and of more localized variability (ITS_u) in our flow parameter data set. The R^2 values for the M- and F-scale submodels are inconsistently distributed between these turbulent variables and do not explain as much of the variability as the larger-scale submodels. It is possible that the resolution of the F-scale PCNMs was too coarse to pick up the fine-scale details where structures had periods too short to be represented. The higher minimum truncation distance and irregular sampling grid may have resulted in the distortion of the finer-scale PCNM variables resulting in a loss of fine-scale resolution.

4.4. Spatial Decomposition and Intercorrelation of PCNM Submodel Flow Variables

[24] Using the multivariate analysis of all flow variables, we mapped the “fitted site scores” of each significant canonical axis for the four submodels investigated (Figure 5). All canonical axes presented are significant at $\alpha = 0.01$ with the exception of the M-scale canonical axis 2 which is significant at $\alpha = 0.05$. These maps provide a spatial decomposition of the explained variance for each axis and allow for the perception of spatial patterns within the data. The R_i^2 bar graphs included on the right-hand side of Figure 5 indicate the unadjusted fraction of variance explained for each response variable. Since the canonical axes are orthogonal, the fractions of variation they express (R_i^2 , where i is the canonical axis index) are linearly independent of one another. Table 3 provides a summary of the unadjusted R_i^2 values for each flow variable expressed by the significant canonical axes of each submodel. The unadjusted coefficients of multiple determination, R^2 , obtained from all 29 significant PCNMs are, as well, included in the far right hand column of Table 3.

4.4.1. Very Large Scale Submodel

[25] The “fitted site score” plots of the VL-scale submodel, Figures 5a and 5b, display a depth-scale spatial pattern of $\Delta x = 5.3H$ and $\Delta y = 1.0H$ within the turbulent

flow variables. These scales are strikingly similar to the large-scale flow structures found in previous studies [Shvidchenko and Pender, 2001; Roy et al., 2004]. Values of u' , $-\rho\bar{u}v$, Sk_u , Sk_v , $T_{Q4T_{h:0}}$, $T_{Q2T_{h:2}}$, and $f_{Q2T_{h:2}}$ are strongly expressed by the first canonical axis (Figure 5a) while $\bar{U}$ and v' are strongly expressed by the second canonical axis (Figure 5b). The first canonical axis indicates that high values of u' , $-\rho\bar{u}v$, Sk_v , $T_{Q4T_{h:0}}$, $T_{Q2T_{h:2}}$, and $f_{Q2T_{h:2}}$ and low-magnitude Sk_u occur in a large region above and adjacent to the cluster between $x = 0.35$ and 1.5 m; the inverse trend occurs farther downstream between $x = 2.1$ and 3.7 m. The second canonical axis reveals a zone of low $\bar{U}$ and high v' between $x = 1.3$ and 2.6 m. Farther downstream between $x = 2.6$ and 4.0 m, this trend is reversed. Differences between u' and v' indicate that u' is of greater magnitude near the pebble cluster, while v' is larger farther downstream. The elevated $f_{Q2T_{h:2}}$, positive Sk_v , and negative Sk_u in the near-wake are an indication of ejecting structures, while downstream, the reverse skewness trend indicates high-speed sweeps in the $x = 2.1$ to 3.7 m range.

[26] The eigenvectors for the VL scale (Figure 6a) reveal high correlation between u' and $-\rho\bar{u}v$, and $T_{Q4T_{h:0}}$ and $T_{Q2T_{h:2}}$. $\bar{U}$ is negatively correlated with $\bar{V}$ and displays a near-zero correlation with u' and $-\rho\bar{u}v$. The inhomogeneous bed and the turbulence generated in the recirculation and shedding zones in the near-wake of the pebble cluster [Buffin-Bélanger and Roy, 1998] are likely responsible for disrupting the large-scale spatial correlation between $\bar{U}$ and u' .

4.4.2. Large-Scale Submodel

[27] The large-scale PCNM submodel explains the greatest fraction of the variance in the flow data. The physical scale of this submodel $\Delta x = 2.6H$ and $\Delta y = 0.5H$ is still within the range of sizes described as large-scale structures in previous studies [Liu et al., 2001; Nakagawa and Nezu, 1981]. Most of the explained variance is expressed by the first canonical axis ($R_1^2 = 0.16$). The fraction of R_1^2 expressed by individual flow variables (Figures 5c–5f) differs slightly from the VL-scale submodel. $\bar{U}$, u' , v' , $-\rho\bar{u}v$, $T_{Q2T_{h:0}}$, $T_{Q4T_{h:2}}$, and $f_{Q4T_{h:2}}$ are strongly expressed by the first axis. Sk_u and $\bar{V}$ are strongly expressed by the second ($R_2^2 = 0.13$) and the third ($R_3^2 = 0.1$) canonical axes, respectively. The PCNMs of the L-scale submodel vertically discriminate scales of $0.5H$ and, as such, are better able to distinguish spatial patterns which bisect the water column. Conversely, the VL-scale submodel was restricted to investigating depth-scale structures ($\Delta y = 1.0H$).

[28] The maps of the “fitted site scores” of the first canonical axis (Figure 5c) show zones of increased magnitude $\bar{U}$ and positive $\bar{V}$ overtop (stoss side) of the pebble cluster and in the far-wake ($x = 0.7$ – 2.8 m). In the upstream zone, the flow constriction induced by the pebble cluster causes a suppression of the turbulence statistics as flow is forced overtop of the cluster [Buffin-Bélanger and Roy, 1998]. The far-wake is characterized by fluid upwelling [Buffin-Bélanger and Roy, 1998]. These characteristic zones were discussed by Buffin-Bélanger and Roy [1998], and through the current analysis are related to large-scale turbulent flow structures. Between these two zones lies the near-wake region ($x = 0.8$ m to 2.2 m) which consists of a recirculation and eddy shedding zone [Buffin-Bélanger and Roy, 1998]. The recirculation zone is found here to be characterized by large-scale patterns of low magnitude $\bar{U}$,

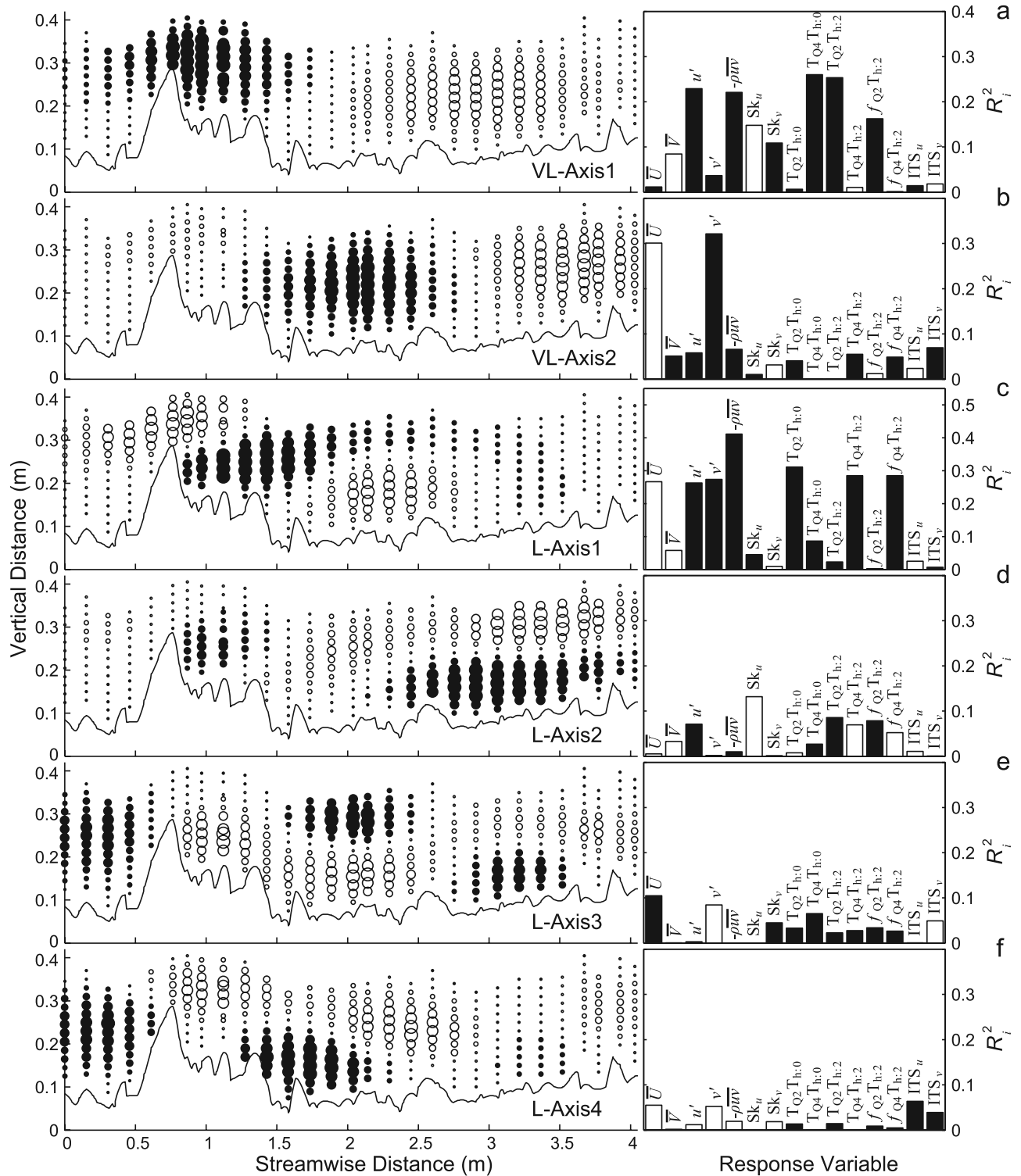


Figure 5. Significant canonical axes of the “fitted site scores” ($p < 0.05$) plotted on the sampling location coordinates: (a) VL-axis1, (b) VL-axis2, (c) L-axis1, (d) L-axis2, (e) L-axis3, (f) L-axis4, (g) M-axis1, (h) M-axis2, and (i) F-axis1. Positive values are represented by solid circles. The right-hand bar graphs present the fraction of variance for each flow variable explained by the canonical axes (open = negative correlations, solid = positive). Flow is from left to right.

high magnitude u' , v' , $-\rho\bar{u}v$, and a dominance of smaller magnitude ejections ($T_{Q2}T_{h:0}$) and high magnitude sweeps ($T_{Q4}T_{h:2}$). The PCNM technique was able to clearly show the division (shearing) caused by the shedding/recirculation

zone and the overlying fluid. The shear layer is initiated at the crest of the pebble cluster and is inclined toward the water surface. The second axis (Figure 5d) explains much of the variance in the distal downstream portion of the x - y

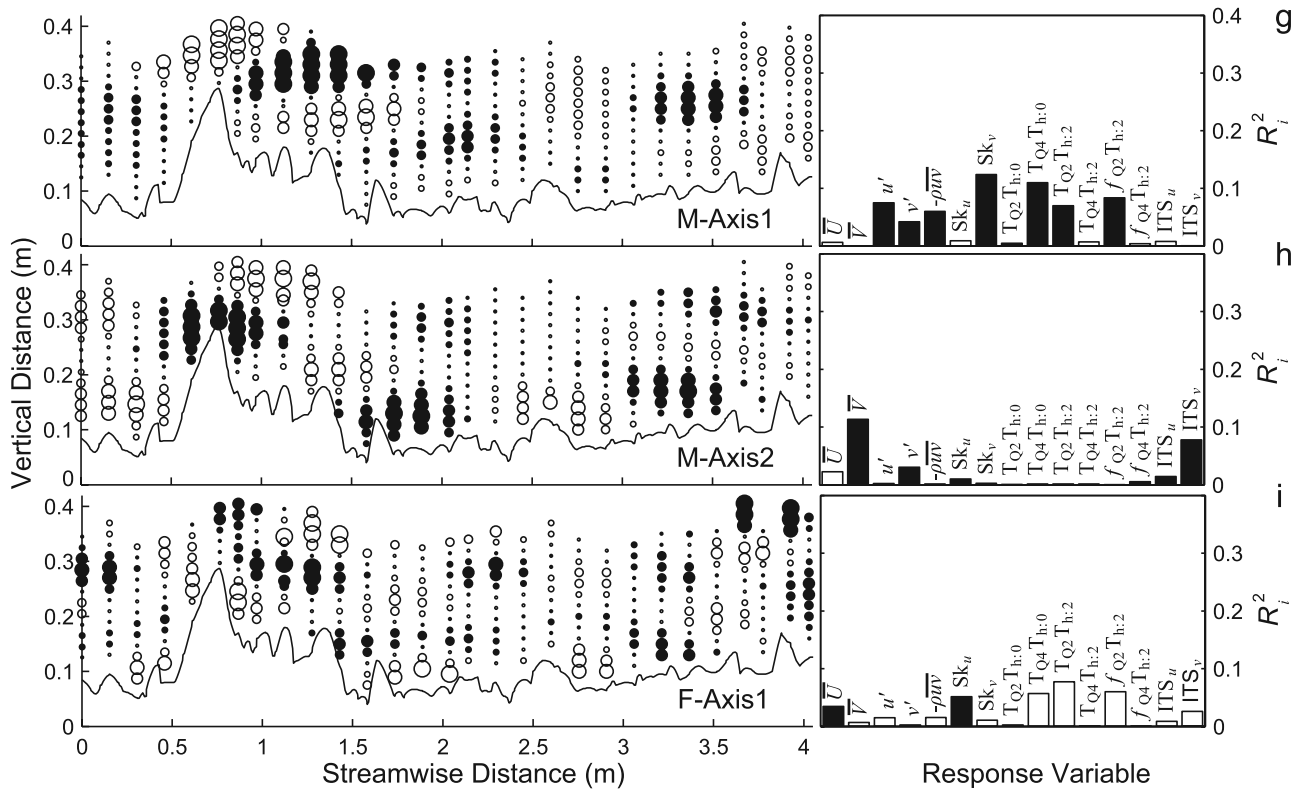


Figure 5. (continued)

transect, where high-magnitude Sk_u occurs in the upper water column, indicating infrequent events of accelerated fluid which coincide with the high-magnitude sweeps (Q4 events) also indicated by the analysis. The trend is reversed closer to the bed in the distal zone. Canonical axis 3 indicates elevated $\bar{U}$ upstream of the cluster and higher in the water column, and all variables are only weakly explained by canonical axis 4. The between variable corre-

lations (Figure 6b) are similar to those discussed for the VL scale. While the VL-scale plot indicates a much higher correlation between u' and $-\rho\bar{u}v$ than between v and $-\rho\bar{u}v$, the L-scale plot shows a near equal correlation between u'/v' and $-\rho\bar{u}v$.

4.4.3. Medium-Scale Submodel

[29] The PCNMs of the medium-scale submodel have multiple periods over the x - y transect plane in the stream-

Table 3. Estimated Fractions of Unadjusted Variance (R_i^2) of Flow Variables Expressed by Significant Canonical Axes^a

	Very Large Scale		Large Scale				Medium Scale		Fine Scale	All Scales
	R_1^2	R_2^2	R_1^2	R_2^2	R_3^2	R_4^2	R_1^2	R_2^2	R_1^2	R^2
$\bar{U}$	0.01	0.30	0.27	0.01	0.10	0.05	0.01	0.02	0.03	0.88
$\bar{V}$	0.08	0.05	0.06	0.03	0.00	0.00	0.00	0.11	0.01	0.51
u'	0.23	0.06	0.26	0.07	0.00	0.01	0.08	0.00	0.02	0.78
v'	0.04	0.32	0.27	0.00	0.08	0.05	0.04	0.03	0.00	0.88
$-\rho\bar{u}v$	0.22	0.07	0.41	0.01	0.00	0.02	0.06	0.00	0.02	0.82
Sk_u	0.15	0.01	0.04	0.13	0.00	0.00	0.01	0.01	0.05	0.48
Sk_v	0.11	0.03	0.01	0.00	0.04	0.02	0.12	0.00	0.01	0.39
$T_{Q2}T_{h:0}$	0.01	0.04	0.31	0.01	0.03	0.01	0.00	0.00	0.00	0.51
$T_{Q4}T_{h:0}$	0.26	0.00	0.09	0.03	0.07	0.00	0.11	0.00	0.06	0.63
$T_{Q2}T_{h:2}$	0.25	0.00	0.02	0.09	0.02	0.01	0.07	0.00	0.08	0.56
$T_{Q4}T_{h:2}$	0.01	0.05	0.29	0.07	0.03	0.00	0.01	0.00	0.00	0.56
$f_{Q2}T_{h:2}$	0.16	0.01	0.00	0.08	0.03	0.01	0.08	0.00	0.06	0.46
$f_{Q4}T_{h:2}$	0.00	0.05	0.28	0.05	0.03	0.01	0.00	0.01	0.00	0.52
ITS_u	0.01	0.02	0.03	0.01	0.00	0.06	0.01	0.01	0.01	0.21
ITS_v	0.02	0.07	0.01	0.00	0.05	0.04	0.00	0.08	0.03	0.36
Canonical eigenvalue	0.10 ^b	0.07 ^b	0.16 ^b	0.04 ^b	0.03 ^b	0.02 ^b	0.04 ^b	0.02 ^c	0.02 ^b	

^aOnly significant canonical axes of each submodel are presented. The bottom row presents the unadjusted canonical eigenvalues (λ_i of each significant axis (λ_i are estimated as the mean R_i^2 value for all 15 response variables); and the far right column presents the unadjusted coefficients of multiple determination (R^2) using all 29 significant PCNMs. Monte Carlo significance test (999 unrestricted permutations).

^bHere $p \leq 0.01$.

^cHere $p \leq 0.05$.

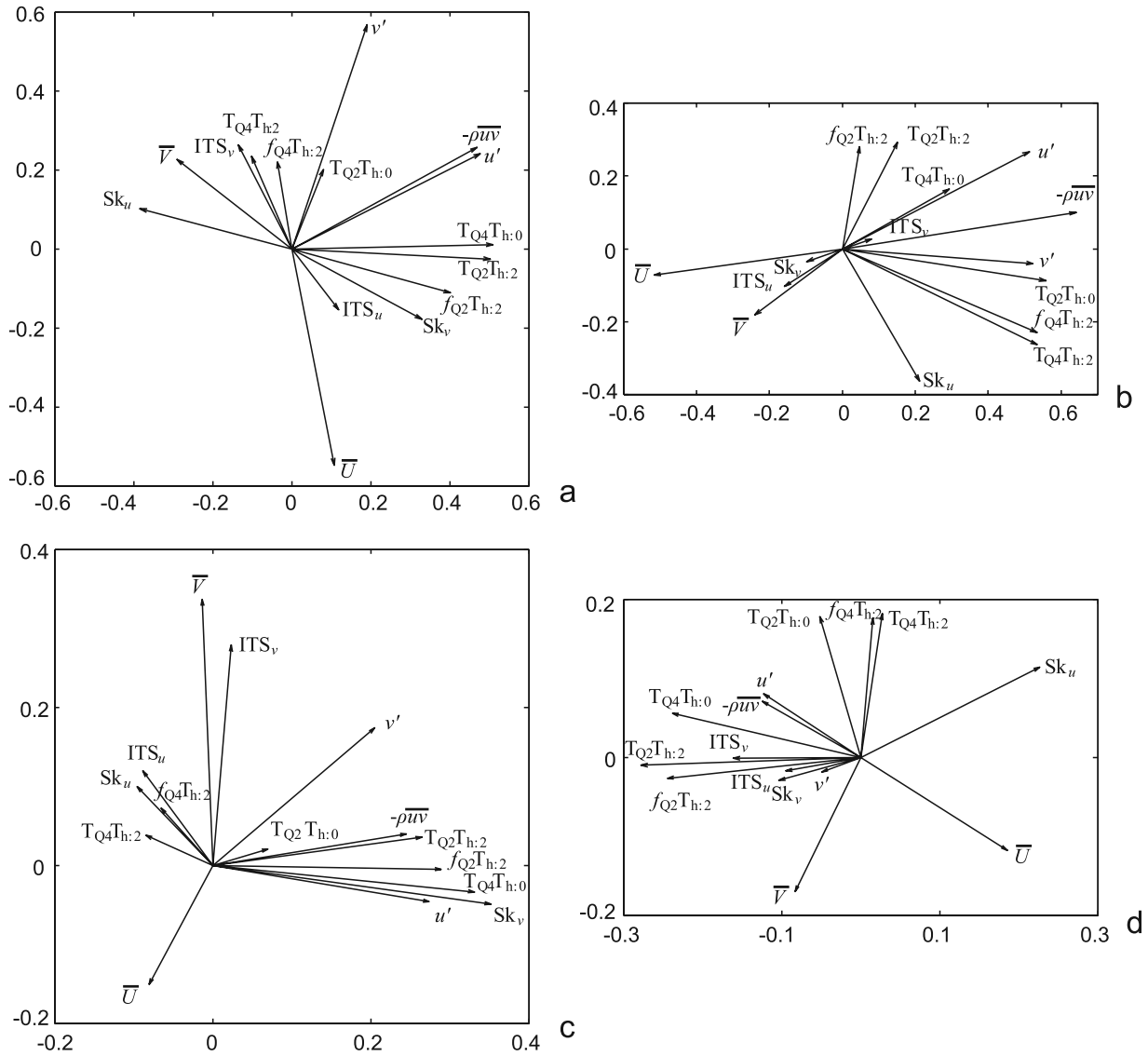


Figure 6. Eigenvector scatterplots. Canonical axis 1 is the abscissa, while canonical axis 2 is the ordinate: (a) very large scale, (b) large scale, (c) medium scale, and (d) fine scale.

wise and vertical directions, producing a “checkerboard” pattern. The variance explained by this scale is between $\Delta x = 1.6H$ and $\Delta y = 0.3H$. The R_1^2 expressed by individual flow variables is much smaller than for the two larger scales; only two variables, Sk_v and $T_{Q4}T_{h:0}$, are moderately expressed (Figure 5g). The shedding/recirculation zone identified at the larger scales is further subdivided into two distinct zones. The overlying shedding zone with elevated medium-scale turbulent statistics (positive Sk_v and elevated $T_{Q4}T_{h:0}$) initiates at the tip of the pebble cluster, while the recirculation zone below is characterized by decreased turbulent statistics at the medium scale. The second canonical axis explains pockets of elevated $\bar{V}$ and ITS_v directly above the pebble cluster (Figure 5h). These two variables are also highly correlated at the medium scale, as shown by the eigenvector plot (Figure 6c). The plot also indicates a tighter grouping of turbulent statistics than for

the two larger scales, where u' , Sk_v , $-\rho_{uv}$, $T_{Q2}T_{h:2}$, $T_{Q4}T_{h:0}$, and $f_{Q2}T_{h:2}$ are closely correlated.

4.4.4. Fine-Scale Submodel

[30] All flow response variables are weakly expressed at the F scale (Figure 5i), perhaps indicating that the resolution of the F-scale PCNMs (constrained by the selected d_t) is too coarse to pick up the fine-scale details. The larger residual variance in the higher moment and more localized turbulent statistics supports this point. The pattern of scatter in the plotted “fitted site scores” for the F-scale submodel (Figure 5i), which we include for completeness, is difficult to interpret. The effects of the irregular sampling grid preferentially distort the smaller-scale PCNMs and could be a factor for the apparent random patchiness. On the other hand, the random patchiness observed may have a physical basis in terms of turbulence theory, where at small scales, turbulent motion tends toward local isotropy [Townsend,

1976]. The apparent randomly scattered values of Figure 5i may represent fine-scale random autocorrelation between neighboring sampling locations.

5. Discussion

[31] Herein the PCNM method provides a powerful means of describing multivariate spatial patterns within the turbulent flow data. The turbulent flow field over very rough boundaries appears to be well organized at a range of scales identified and quantified by the PCNM analysis. The technique was able to decompose and summarize the complex interrelations of the flow variables over a range of scales. The explanation of 75% of the variation in the standard turbulent parameters u' , v' , and $-\rho\bar{u}v$ by spatial descriptors of $\Delta x = 5.3H - 2.6H$ and $\Delta y = 1.0H - 0.5H$ indicates that the turbulent energy and shear stress form consistent, large-scale spatial patterns under the influence of a pebble cluster. The relative importance of the large-scale modes is consistent with previous smooth-walled POD studies [Liu *et al.*, 2001] and highlights the similarities between smooth- and rough-walled flows. Previous studies comparing smooth- and rough-wall boundary layers have shown that in the outer boundary layer, $y/H > 0.2$, turbulent intensities and macrolength scales are weakly effected by roughness size [Grass, 1971; Nezu and Nakagawa, 1993]. Both smooth and rough boundary layers produce sweep and ejection events, irrespective of the surface roughness, even though their generation mechanisms close to the bed are different (as no viscous sublayer is present in rough boundary layers), and it has been found that both types of boundary layers are scaled by roughness size [Grass and Mansour-Tehrani, 1996; Smith, 1996]. Large-scale coherent structures are a common feature of turbulent flows [Shvidchenko and Pender, 2001; Roy *et al.*, 2004], and have been found to be little affected by the shedding vortices of large roughness elements [Lacey and Roy, 2007].

[32] The large-scale patterns observed in our study are determined from the spatial distribution of time-averaged flow parameters and, as such, are not equivalent to the depth-scale, time-dependent, coherent flow structures reported in previous studies based on the analysis of turbulent events [Shvidchenko and Pender, 2001; Roy *et al.*, 2004]. Yet the similar scaling suggests that flow depth is the limiting scale for both the spatial patterns of turbulent properties and the large-scale coherent turbulent structures. The dependence of the spatial patterns of the flow parameters on the heterogeneous bed is implied by the simple relationship of the VL- and L-scale PCNM submodels with pebble cluster height ($\Delta x = 10h_s - 5h_s$ and $\Delta y = 2h_s - 1.0h_s$). The streamwise spatial extent of the L-scale submodel is equal to the distance estimated in experiments by Best and Brayshaw [1985] from their roughness element to the reattachment point. The high explanation of flow variables at the L scale is perhaps due to the clear distinction between patterns occurring within and outside of the pebble cluster wake zone. The direct cluster height scaling supports the view that the general flow structure can be linked to the spatial distribution of roughness elements on the river bed [Clifford *et al.*, 1992].

[33] The observations made from the “fitted site scores” regarding the spatial differences between u' and v' , and the inverse relationship between Sk_v and Sk_u , were similarly

interpreted by Buffin-Bélanger and Roy [1998] and can be seen to some degree in the z-score turbulent flow plots (Figures 2c, 2d, 2f, and 2g). Buffin-Bélanger and Roy [1998] gave general descriptions of the turbulent flow variable spatial patterns but were unable to give scale-dependent, quantitative results. Through PCNM analysis, we are able to estimate the proportion of variances (R_1^2) associated with each flow variable at each spatial scale. The L-scale PCNM submodel clearly shows the division caused by the shedding/recirculation zone and the overlying fluid, and the M-scale PCNM submodel provides further detail on the shedding/shear layer. While similar spatial patterns were observed by Buffin-Bélanger and Roy [1998], PCNM analysis quantifies and efficiently summarizes the dependent turbulent parameters associated with these spatial patterns (i.e., high u' , v' , $-\rho\bar{u}v$, $T_{Q2}T_{h:0}$, $T_{Q4}T_{h:2}$, and $f_{Q4}T_{h:2}$ at the L scale (Figure 5c), and high Sk_v and $T_{Q4}T_{h:0}$ at the M scale (Figure 5g)).

[34] Plotting the “fitted site scores” is an advantage of this multivariate analysis because it provides an effective means of summarizing the numerous flow variables into a single plot (at each scale), as well as provides a means of determining which variables are strongly or weakly expressed. Large-scale spatial patterns of the strongly expressed variables can be readily identified from the “fitted site scores” plots. These large-scale patterns are, in some instances, difficult to distinguish on the z-score plots (Figure 2) due to the superposition of smaller-scale turbulent patterns. An additional advantage of the PCNM technique is that correlations between flow variables at each scale can be investigated through eigenvector scatterplots (Figure 6). These correlations are related to the entire study area and do not differentiate between different flow regions (i.e., upstream and downstream of the cluster), yet still give valuable information on the relationships between flow variables over different scales. For example, the VL-scale submodel revealed high correlation between u' and $-\rho\bar{u}v$, while the L-scale and M-scale submodels estimated similar correlations between u' and $-\rho\bar{u}v$, and v' and $-\rho\bar{u}v$. While correlations do not indicate a causal link, these results do indicate that at the largest scales, the Reynolds shear stress and the u component turbulent energy show a much stronger interdependence than for the v component.

[35] The main drawback of the PCNM method is the dependence of the PCNM variables on the uniformity of the sampling grid (i.e., the greater the inconsistency within the sampling grid, the greater the loss of fine-scale structure explanation). When large inconsistencies are present in the sampling grid, the PCNMs are still orthogonal and properly describe the sampling space; yet instead of containing regular sinusoidal waves, individual PCNMs are composed of sinusoidal waves of multiple scales [Borcard *et al.*, 2004]. This makes it difficult to partition the PCNMs into different scales. In our study, a larger truncation distance was selected to avoid PCNM inconsistencies, and as a result, the explanation of fine-scale variability was limited. The more localized, higher moment and event statistics were much less explained (higher residual variation) than the larger-scale core turbulent statistics. Understanding this constraint will help us design future studies.

[36] The PCNM technique illustrated here, using the characterization of the turbulent flow field, is a powerful

statistical tool with a much broader application in fluvial geomorphology. Considering the recent interest in reach-scale spatial attributes and patterns [Emery *et al.*, 2003; Clifford *et al.*, 2005], PCNM analysis could provide a new way of looking at the ecological and physical environmental variables on a scale basis. For instance, research investigating the spatial scales and patterns over which physical habitat units and ecological (biotic) units interact, could profit from the use of PCNM analysis. Proposed hierarchical models group physical habitat units such as “river styles” [Thomson *et al.*, 2001] and “physical biotopes” [Newson and Newson, 2000] at different scales. These classifications, such as the “river styles” geomorphic channel classification system [Thomson *et al.*, 2001], can be quite cumbersome due to the number of parameters required to place rivers into the correct unit at each scale. Physical units or flow parameters, which are relevant at one scale, are not necessarily important to biota at other scales. Hierarchical cluster analysis has been used as a spatial statistical tool to define physical units, yet lacks the differentiation of patterns between scales [Emery *et al.*, 2003]. PCNM analysis is a much more powerful technique to provide spatially explicit and scale-dependent relationships between flow variables and aquatic species data. These scale-dependent (hierarchical) spatial patterns could be used to determine the habitat units important for geomorphologists and ecologists. Similarly to Clifford *et al.* [2005], conducting PCNM analysis on the high-resolution velocity and depth data estimated by a hydrodynamic model would provide a powerful appraisal tool for river rehabilitation projects.

6. Conclusions

[37] PCNM analysis successively partitioned the variance (according to characteristic spatial scales) associated with the mean and turbulent flow variables calculated from instantaneous velocities measured over an in situ turbulent flow field in the presence of a pebble cluster in a gravel-bed river. The variation was partitioned over four spatial scales: very large, large, medium, and fine. The full model significantly explained 58% of the variance in the data (R_a^2), while the submodels VL, L, M, and F explained significant ($\alpha = 0.01$) portions of the variance (17%, 24%, 6%, and 2%, respectively). The PCNM analysis was able to demonstrate a high spatial dependence within the flow variables and to quantify the dominant spatial scales and patterns of the core turbulent variables in the mean and turbulent flow data. The VL and L scales explained 75% of the variance of the main turbulent flow indicators u' , v' , and $-\rho\bar{u}v$. The L-scale submodel explained the largest percentage of the variance throughout most of the flow variables. Our results suggest that flow depth and roughness element height are appropriate scales for the time-averaged spatial patterns which were observed.

[38] The usefulness of the PCNM statistical technique is that it quantifies the spatial dependence of individual response variables over a user-defined range of spatial scales. The method provides a spectral decomposition for spatially irregular sampling locations which is much more powerful than currently used methods of quantifying spatial structure (i.e., trend surface analysis). In our study, the mapped “fitted site scores” of the PCNM submodels not only provided similar information to the z-score flow

parameter plots, but also allowed for the quantification of the explained variance and for the identification of scale differences within the flow variables. PCNM analysis by canonical RDA was able to summarize the core spatially dependent variables in only a few plots, allowing for a rapid analysis and quantification of spatial patterns. The intercorrelations among individual response variables at each spatial scale are illustrated through eigenvector scatterplots. The technique allows researchers to refine their analysis and to examine in detail the structure of the data. This is a considerable advantage over the practice often followed in analyses of spatial structures in hydrodynamics, where plots of each flow variable are investigated in turn. The PCNM technique is a powerful tool to understand the spatial relations among complicated sets of variables. The present application to turbulent flow dynamics around a pebble cluster, a problem of great complexity, illustrates the power of the method and its potential for a broad range of applications in water resources and the Earth sciences.

[39] **Acknowledgments.** This research was conducted as part of the program of the Canada Research Chair in Fluvial Dynamics. Funding for this research was provided by the National Sciences and Engineering Research Council and by the Canadian Foundation for Innovation. We would like to thank Tom Buffin-Bélanger for allowing us to use his data and for his review and thorough comments on a preliminary version of this manuscript. We would like to thank Pascale Biron for her help with the collection of the original data in the field and Daniel Borcard for fielding technical questions on PCNM analysis. We are also grateful to the anonymous reviewers whose insightful comments helped us improve the manuscript.

References

- Adrian, R. J., C. D. Meinhart, and C. D. Tomkins (2000), Vortex organization in the outer region of the turbulent boundary layer, *J. Fluid Mech.*, 422, 1–54.
- Bennett, S. J., and J. L. Best (1995), Mean flow and turbulence structure over fixed, 2-dimensional dunes—Implications for sediment transport and bedform stability, *Sedimentology*, 42(3), 491–513.
- Berkooz, G., P. Holmes, and J. L. Lumley (1993), The proper orthogonal decomposition in the analysis of turbulent flows, *Annu. Rev. Fluid Mech.*, 25, 539–575.
- Best, J. L., and A. C. Brayshaw (1985), Flow separation—A physical process for the concentration of heavy minerals within alluvial channels, *J. Geol. Soc. London*, 142(5), 747–755.
- Borcard, D., and P. Legendre (2002), All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices, *Ecol. Modell.*, 153, 51–68.
- Borcard, D., P. Legendre, C. Avois-Jacquet, and H. Tuomisto (2004), Dissecting the spatial structure of ecological data at multiple scales, *Ecology*, 85, 1826–1832.
- Brayshaw, A. C., L. F. Frostick, and I. Reid (1983), The hydrodynamics of particle clusters and sediment entrainments in coarse alluvial channels, *Sedimentology*, 30, 137–143.
- Brind'Amour, A., D. Boisclair, P. Legendre, and D. Borcard (2005), Multi-scale spatial distribution of a littoral fish community in relation to environmental variables, *Limnol. Oceanogr.*, 50, 465–479.
- Buffin-Bélanger, T., and A. G. Roy (1998), Effects of a pebble cluster on the turbulent structure of a depth-limited flow in a gravel-bed river, *Geomorphology*, 25, 249–267.
- Buffin-Bélanger, T., A. G. Roy, and A. D. Kirkbride (2000), On large-scale flow structures in a gravel-bed river, *Geomorphology*, 32, 417–435.
- Chen, J., and F. Hu (2003), Coherent structures detected in atmospheric boundary-layer turbulence using wavelet transforms at Huaihe River basin, China, *Boundary Layer Meteorol.*, 107, 429–444.
- Clifford, N. J., K. S. Richards, and A. Robert (1992), The influence of micro-form bed roughness elements on flow and sediment transport in gravel bed rivers—Comment, *Earth Surf. Processes Landforms*, 17(5), 529–534.
- Clifford, N. J., P. J. Soar, O. P. Hammar, A. M. Gurnell, G. E. Petts, and J. C. Emery (2005), Assessment of hydrodynamic simulation results for eco-hydraulic and eco-hydrological applications: A spatial semivariance approach, *Hydrol. Processes*, 19(18), 3631–3648.

- Dray, S., P. Legendre, and P. R. Peres-Neto (2006), Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM), *Ecol. Modell.*, 196, 483–493.
- Emery, J. C., A. M. Gurnell, N. J. Clifford, G. E. Petts, I. P. Morrissey, and P. J. Soar (2003), Classifying the hydraulic performance of riffle-pool bedforms for habitat assessment and river rehabilitation design, *River Res. Appl.*, 19(5–6), 533–549.
- Ezekiel, M. (1930), *Methods of Correlation Analysis*, John Wiley, Hoboken, N. J.
- Grass, A. J. (1971), Structural features of turbulent flow over smooth and rough boundaries, *J. Fluid Mech.*, 50, 233–255.
- Grass, A. J., and M. Mansour-Tehrani (1996), Generalized scaling of coherent bursting structures in the near-wall region of turbulent flow over smooth and rough boundaries, in *Coherent Flow Structures in Open Channels*, edited by P. J. Ashworth et al., pp. 41–61, John Wiley, Hoboken, N. J.
- Kirkbride, A. D. (1993), Observations of the influence of bed roughness on turbulence structure in depth limited flows over gravel beds, in *Turbulence: Perspectives on Flow and Sediment Transport*, edited by N. J. Clifford et al., pp. 185–196, John Wiley, Hoboken, N. J.
- Lacey, R. W. J., and A. G. Roy (2007), A comparative study of the turbulent flow field with and without a pebble cluster in a gravel-bed river, *Water Resour. Res.*, doi:10.1029/2006WR005027, in press.
- Lawless, M., and A. Robert (2001), Three-dimensional flow structure around small-scale bedforms in a simulated gravel-bed environment, *Earth Surf. Processes Landforms*, 26(5), 507–522.
- Legendre, P., and D. Borcard (2006), Quelles sont les échelles spatiales importantes dans un écosystème?, in *Analyse Statistique de Données Spatiales*, edited by J.-J. Droesbeke et al., pp. 425–442, Technip, Paris.
- Legendre, P., and L. Legendre (1998), *Numerical Ecology*, Elsevier, New York.
- Liu, Z., R. J. Adrian, and T. J. Hanratty (2001), Large-scale modes of turbulent channel flow: Transport and structure, *J. Fluid Mech.*, 448, 53–80.
- Lu, S. S., and W. W. Willmarth (1973), Measurements of the structure of the Reynolds stress in a turbulent boundary layer, *J. Fluid Mech.*, 60, 481–511.
- Moin, P., and R. D. Moser (1989), Characteristic-eddy decomposition of turbulence in a channel, *J. Fluid Mech.*, 200, 471–509.
- Nakagawa, H., and I. Nezu (1981), Structure of space-time correlations of bursting phenomena in an open-channel flow, *J. Fluid Mech.*, 104, 1–43.
- Newson, M. D., and C. L. Newson (2000), Geomorphology, ecology and river channel habitat: Mesoscale approaches to basin-scale challenges, *Prog. Phys. Geogr.*, 24(2), 195–217.
- Nezu, I., and H. Nakagawa (1993), *Turbulence in Open-Channel Flows*, 281 pp., A. A. Balkema, Brookfield, Vt.
- Paola, C., G. Gust, and J. B. Southard (1986), Skin friction behind isolated hemispheres and the formation of obstacle marks, *Sedimentology*, 33, 279–293.
- Robert, A., A. G. Roy, and B. Deserres (1992), Changes in velocity profiles at roughness transitions in coarse-grained channels, *Sedimentology*, 39(5), 725–735.
- Roy, A. G., T. Buffin-Bélanger, H. Lamarre, and A. D. Kirkbride (2004), Size, shape and dynamics of large-scale turbulent flow structures in a gravel-bed river, *J. Fluid Mech.*, 500, 1–27.
- Shamloo, H., N. Rajaratnam, and C. Katopodis (2001), Hydraulics of simple habitat structures, *J. Hydraul. Res.*, 39(4), 351–366.
- Shvidchenko, A. B., and G. Pender (2001), Macroturbulent structure of open-channel flow over gravel beds, *Water Resour. Res.*, 37, 709–719.
- Smith, C. R. (1996), Coherent flow structures in smooth-wall turbulent boundary layers: Facts, mechanisms and speculation, in *Coherent Flow Structures in Open Channels*, edited by P. J. Ashworth et al., pp. 1–39, John Wiley, Hoboken, N. J.
- ter Braak, C. J. F., and P. Smilauer (2002), *Canoco Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination*, version 4.5, Microcomputer Power, Ithaca, N. Y.
- Thomson, J. R., M. P. Taylor, K. A. Fryirs, and G. J. Brierley (2001), A geomorphological framework for river characterization and habitat assessment, *Aquat. Conserv. Mar. Freshwater Ecosyst.*, 11(5), 373–389.
- Townsend, A. A. (1976), *The Structure of Turbulent Shear Flow*, Cambridge Univ. Press, New York.
- Tritico, H. M., and R. H. Hotchkiss (2005), Unobstructed and obstructed turbulent flow in gravel bed rivers, *J. Hydraul. Eng.*, 131(8), 635–645.
- Williams, J. J., P. D. Thorne, and A. D. Heathershaw (1989), Measurements of turbulence in the benthic boundary-layer over a gravel bed, *Sedimentology*, 36(6), 959–971.

R. W. J. Lacey and A. G. Roy, Département de Géographie, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7. (jay.lacey@umontreal.ca)

P. Legendre, Département de Sciences Biologiques, Université de Montréal, C.P. 6128, Succursale Centre-Ville, Montréal, Québec, Canada H3C 3J7.