entropy

MDPI

Article

Randomness Representation of Turbulence in Canopy Flows Using Kolmogorov Complexity Measures †

Dragutin Mihailović [1,*], Gordan Mimić [1], Paola Gualtieri [2], Ilija Arsenić [1] and Carlo Gualtieri [2]

[1] Faculty of Agriculture, University of Novi Sad, Dositej Obradovic Sq. 8, 21000 Novi Sad, Serbia; gordan.mimic@polj.uns.ac.rs (G.M.); ilija@polj.uns.ac.rs (I.A.)

[2] Department of Civil, Architectural and Environmental Engineering, University of Naples Federico II, Via Claudio 21, 80125 Naples, Italy; paola.gualtieri@unina.it (P.G.); carlo.gualtieri@unina.it (C.G.)

* Correspondence: guto@polj.uns.ac.rs; Tel.: +38-121-485-3203

† This paper is an extended version of our paper published in Proceedings of the 8th International Congress on Environmental Modelling and Software (iEMSs), 10–14 July 2016, Toulouse, France.

Received: 11 August 2017; Accepted: 25 September 2017; Published: 27 September 2017

Abstract: Turbulence is often expressed in terms of either irregular or random fluid flows, without quantification. In this paper, a methodology to evaluate the randomness of the turbulence using measures based on the Kolmogorov complexity (KC) is proposed. This methodology is applied to experimental data from a turbulent flow developing in a laboratory channel with canopy of three different densities. The methodology is even compared with the traditional approach based on classical turbulence statistics.

Keywords: turbulent flow; canopy flow; randomness; coherent structures; Shannon entropy; Kolmogorov complexity

1. Introduction

Randomness is narrowly related to the unpredictability of model or observation results and the inability to predict them with sufficient accuracy. Naturally, several questions arise: what in fact constitutes randomness; how can we quantify it; and why do we think about? In addition, randomness and volatility are oft-used agreeably in the same way that precision and accuracy receive the same colloquial treatment. Each means something on its own and merits consideration as such. Here, to the authors' knowledge, we count some papers, from the last two decades, discussing the complexity of the phenomenon from the perspectives of different scientific fields: cosmic radiation [1], statistical mechanics [2,3], chaotic behavior [4] and turbulence [5]. Kreinovich and Kunin [4] suggested the use of the ideas of the Kolmogorov complexity to distinguish singular and non-singular solutions. Lorenz [6] explored how the complexity of wind records (and more generally, of velocity descriptions of thermally-driven flow) relates to the available work and the frictional or viscous dissipation. Weck et al. [7] used the complexity to analyze fluctuating time series of different turbulent plasmas.

Some problems of turbulent flow remain unanswered because a more complete definition of turbulence has not been proposed yet [8]. In particular, randomness, one of the fundamental characteristics of a turbulent flow, is often verbally expressed in terms of the irregularity of the fluid flow, without its quantification. An overview of such definitions is reported by Ichimiya and Nakamura [8]. The only exception is Pope's definition of randomness related to turbulence [9]. Evaluation of the level of randomness in turbulence could be particularly interesting in canopy flows [10–14]. Vegetation in open channel flow considerably affects the flow and turbulence structure, so interactions between the lower layer with the vegetation (i.e., the vegetation layer) and the surface layer that is above the vegetation evolve [13]. In the case of submerged canopy flows, a roughness sublayer (RSL) is formed [15–17]. Within the RSL, three distinct zones exist (Figure 1a): (i) the deep

zone (RS1), primarily dominated by small eddies associated with the von Kármán streets; (ii) the RS2 zone, where the superposition of attached eddies and Kelvin–Helmholtz waves cause an inflection point in the mean velocity profile (Figure 1b), which develops when two co-flowing streams have different velocities [11]; in this turbulent mixing region, Kelvin–Helmholtz instability causes coherent turbulent structures that travel downstream in the fluid [15,16,18]; (iii) the RS3 zone, where the rough-wall boundary layer occurs. For a long time, there was an interest in the fluid mechanics for better understanding the physical processes involved in flow-canopy interaction, which includes interaction between the zones described above, e.g., the turbulent boundary layer and the outer laminar region, wall and free-shear turbulent flows exhibiting a coherent structure [8,19]. Based on the Kolmogorov complexity (KC) [20], Lempel and Ziv developed an algorithm for calculating the measure of randomness (LZA) [21].

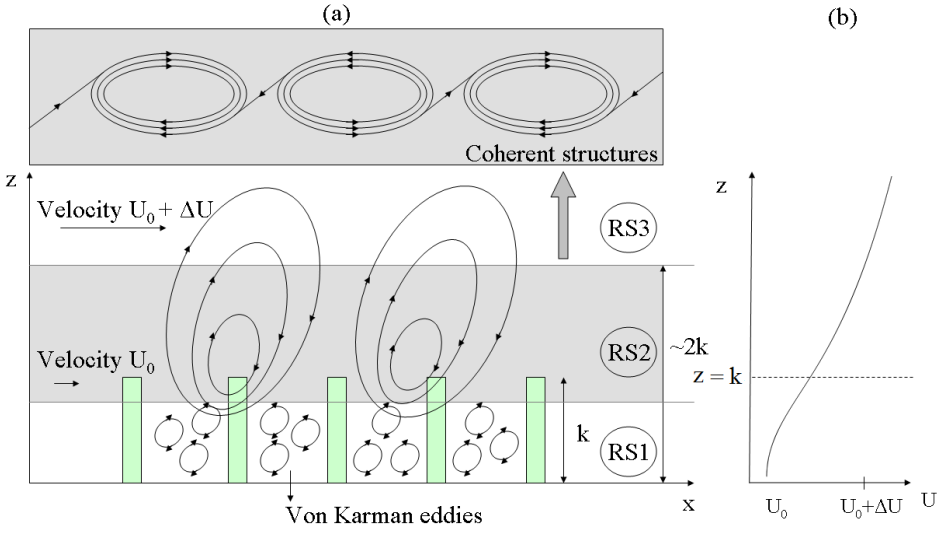

Figure 1. (**a**) A schematic diagram of the eddies structure over and within the canopy: (i) RS1 zone ($z/k < 1$; k is the height of the cylinders used for modeling canopies) where the flow field is primarily dominated by small eddies associated with the von Kármán streets; (ii) RS2 zone straddles the top portion of the canopy, and it is dominated by a mixing layer; and (iii) RS3 zone ($z/k > 2$) is the classical boundary-layer region dominated by eddies with length scales proportional to $(z - d)$, where d is the displacement height; (**b**) the mean velocity profile within the canopy $u(z) = u(k)\sqrt{-\beta_1 e^{-\beta_3 z} + \beta_2 e^{\beta_3 z}}$ is obtained from the solution of the partial differential equation $\partial u^2/\partial z = [(2C_d\lambda_d k^2)/(\sigma_s P_s)]u^2$ where: $u(k)$ is the velocity at the height k; β_1, β_2 are parameters depending on the morphological and aerodynamic characteristics of the canopy and $\beta_3 = [(2C_d\lambda_d k^2)/(\sigma_s P_s)]$; C_d is the drag coefficient; σ_s is the parameter of proportionality between the turbulent transport coefficient and velocity within the canopy; λ_d is the roughness density; and P_s is the shelter factor [22].

Before we clarify the intention of this paper, we concisely summarize some ideas by Ichimiya and Nakamura [8] in regards to randomness representation in turbulent flows with Kolmogorov complexity in mixing layer. They mentioned that theoretical books do not use phrases such as "the turbulence is the condition whose property is...", but they start from the distinction between mean and irregularly-fluctuating velocities [23,24], and also, discussing different definitions of turbulence, they state that the essential property of turbulence is verbally expressed in terms of irregular, random or chance. Thus, Pope [9] defines randomness in relation to turbulence [25]. This definition, which seems reasonable and practical, cannot be used to measure randomness, and it is difficult to relate it to Kolmogorov's probability axiom. Kolmogorov himself describes the application of probability theory, but does not mention random itself at that time [26].

In the treatment of turbulence based on values of turbulence intensity, it is broadly considered as a measure of randomness. However, this definition, which seems reasonable and practical, cannot be used as a complete measure of randomness [8] in the sense as it is defined in [9]. Let us note that Sato et al. [27] proposed a randomness factor defined as the ratio of the energy of the continuous spectrum to the total energy in the power spectrum distribution of the velocity fluctuations [9]. However, there is a problem that the randomness factor takes the same value in the two locations where the flow conditions are different. The goal of this paper is to examine the degree of turbulence in relation to the KC and measures based on that complexity. To explicate that goal, we use Figure 2. The left side of this figure depicts two time series of the velocity measured at two relative heights (see Figure 1) $z/k = 0.067$ (purple) and $z/k = 2.667$ (red), respectively, while its right side represents the corresponding distributions of frequencies of velocities that occur in certain intervals. Looking at the distributions of velocities (right panel), we can conclude at which relative height the turbulent motion is more random (red distribution) if we used the measure of randomness based on the values of turbulence intensity. In the further text, we will see that if randomness is evaluated using the KC measure, randomness is greater on the other relative height (purple), which could be in some sense confusing.

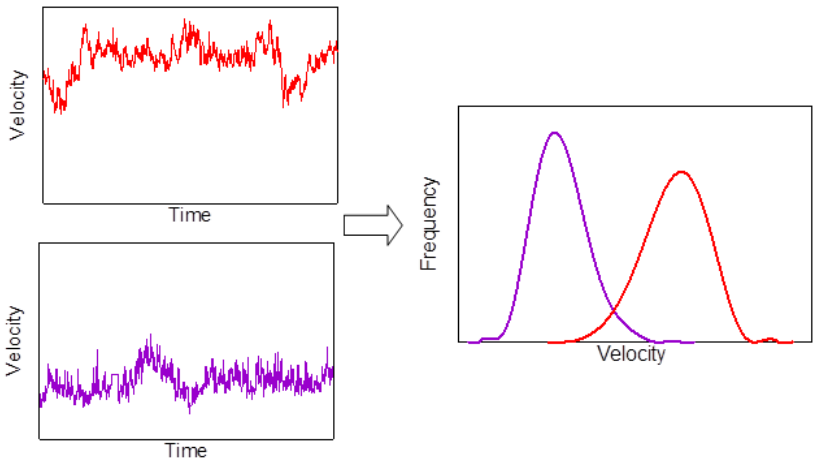

Figure 2. Towards quantification of the randomness of turbulence in canopy flows.

The proposed methodology has been tested with experimental measurements in a canopy flow with a simplified geometry, carried out in a laboratory channel of the Department of Civil, Architectural and Environmental Engineering of the University of Naples Federico II (Naples, Italy), since the focus of this paper is on the application of the KC complexity, not on the base flow. However, it should be pointed out that in the last decade, some authors performed measurements both on more realistic geometries and using more advanced techniques, at least 2C laser Doppler velocimetry (LDV) [28,29].

2. Kolmogorov Complexity and Information Measures Derived from It

2.1. Kolmogorov Complexity

The Kolmogorov complexity $K(x)$ of an object x is the length, in bits, of the smallest program that can be run on a universal Turing machine (U) and that prints object x. The complexity is maximized for random strings [30]. This information measure was developed by Kolmogorov [31], whose comprehensive elaboration can be found in [20]. The complexity $K(x)$ is not a computable function of x. However, a suitable approximation is considered as the size of the ultimate compressed version of x and a lower bound for which a real compressor can achieve [21,32]. The issue of non-computability and the LZA algorithm is comprehensively discussed by Kaspar and Schuster [33]. The calculation of the KC complexity of a time series $\{x_i\}, i = 1, 2, ..., N$ by the LZA algorithm includes

the following steps. (1) Encode the time series by constructing a sequence S of the characters 0 and 1 written as $s(i), i = 1, 2, ..., N$ according to the rule $s(i) = 0, x_i < x_*; 1, x_i \geq x_*$, where x_* is the mean value of the time series, selected as the threshold [34]. Other encoding schemes are also available [35,36]. (2) Calculate the complexity counter $c(N)$. The $c(N)$ is defined as the minimum number of distinct patterns contained in a given character sequence [37]. The complexity counter $c(N)$ is a function of the length of the sequence N. The value of $c(N)$ is approaching an ultimate value $b(N)$ as N is approaching infinity, i.e., $c(N) = O(b(N))$ and $b(N) = N/log_2 N$. (3) Calculate the normalized information measure $C_k(N)$, which is defined as $C_k(N) = c(N)/b(N) = c(N)log_2 N/N$. For a nonlinear time series, $C_k(N)$ varies between zero and one, although it can be larger than one [38]. The LZA algorithm is described in papers [21,33], but in a condensed form. However, here, we will explicate this algorithm through an example that is rather illustrative compared to the examples used in the above mentioned papers. Note that the pattern is a sequence in the coded time series, which is unique and non-repeatable. Recognition of the patterns with the Lempel–Ziv algorithm could be described in the following way: (1) the first digit, no matter if 0 or 1, is always first pattern; (2) define the sequence S, consisting of the digits contained in already recognized patterns; sequence S grows until the whole time series is analyzed; (3) define sequence Q, needed to examine the time series; it is formed by adding new digits until Q is recognized as the new pattern; (4) define the sequence SQ, by adding sequence Q to sequence S; (5) form the sequence $SQ\pi$ by removing the last digit of sequence SQ; (6) now examine if sequence $SQ\pi$ contains sequence Q; (7) if sequence Q is contained in sequence $SQ\pi$, add another digit to sequence Q and repeat the process until the mentioned condition is satisfied; (8) if sequence Q is not contained in sequence $SQ\pi$, then Q is a new pattern. Now, the new pattern is added to the list of known patterns called vocabulary R. Sequence SQ now becomes new sequence S, while Q is emptied and ready for further testing. Now, let us consider a short example with the binary sequence 1011010. The procedure obtained by the LZA can be simply described in this way:

- the first digit is always the first pattern, which implies $\rightarrow R = 1\cdot$
- $S = 1, Q = 0, SQ = 10, SQ\pi = 1, Q \notin v(SQ\pi) \rightarrow R = 1 \cdot 0\cdot$
- $S = 10, Q = 1, SQ = 101, SQ\pi = 10, Q \in v(SQ\pi) \rightarrow 1 \cdot 0 \cdot 1$
- $S = 10, Q = 11, SQ = 1011, SQ\pi = 101, Q \notin v(SQ\pi) \rightarrow R = 1 \cdot 0 \cdot 11\cdot$
- $S = 1011, Q = 0, SQ = 10110, SQ\pi = 1011, Q \in v(SQ\pi) \rightarrow 1 \cdot 0 \cdot 11 \cdot 0$
- $S = 1011, Q = 01, SQ = 101101, SQ\pi = 10110, Q \in v(SQ\pi) \rightarrow 1 \cdot 0 \cdot 11 \cdot 01$
- $S = 1011, Q = 010, SQ = 1011010, SQ\pi = 101101, Q \notin v(SQ\pi) \rightarrow R = 1 \cdot 0 \cdot 11 \cdot 010 \cdot .$

Finally, vocabulary R consists of the patterns 1, 0, 11 and 010, which means that in this particular case, the complexity of the counter is $c(N)$ = 4.

2.2. The Kolmogorov Complexity Spectrum and Its Highest Value

The quantification of the complexity of a system is one of the aims of nonlinear time series analysis. Due to artifacts in various forms (spurious experimental results, etc.), it is often not easy to get desirable and reliable information from a series of measurements. Note that the time series, obtained either by measurement or a modeling procedure, is the only source for evaluating the level of complexity of the environmental system (i.e., canopy flow in this paper, through its flow rate). However, the KC, as an information measure, has two drawbacks: (i) it cannot distinguish time series with different amplitude variations and similar random components; and (ii) in the conversion of a time series into a string, its complexity is hidden in the coding rules. Thus, in the procedure of establishing a threshold for a criterion for coding, some information about the structure of the time series can be lost. In time series analysis of the canopy flow, we use two information measures based on the KC complexity: (i) the Kolmogorov complexity spectrum and (ii) the Kolmogorov complexity spectrum's highest value, which are introduced in [39] and briefly described. According to Definition 1 in [39], the time series $\{x_i\}, i = 1, 2, ..., N$ consists of either measured or calculated values. The Kolmogorov complexity spectrum of time series $\{x_i\}$, i.e., the sequence $\{c_i\}, i = 1, 2, ..., N$ is obtained by the LZA

algorithm applied N times on time series, where thresholds $\{x_{t,i}\}$ are all elements in $\{x_i\}$. Namely, the original time series samples are converted into a set of 0-1 sequences $\{S_i^{(k)}\}, i = 1, 2, ..., N, k = 1, 2, ..., N$, defined by the comparison with a threshold $\{x_{t,k}\}$, i.e., $S_i^{(k)} = \{0, x_i < x_{t,k} or 1, x_i \geq x_{t,k}$. After the LZA algorithm is applied to each element of the series $\{S_i^{(k)}\}$, we get the KC complexity spectrum $\{c_i\}, i = 1, 2, ..., N$. This spectrum allows us to explore the range of amplitudes in a time series representing a physical system with highly enhanced stochastic components, i.e., with highest complexity. Note that for a large number of samples in the time series (as in this paper), the computation of the KC spectrum is computationally very consumable. Then, it is recommended to divide the range of the minimal and maximal values in the time series into subintervals, which are further used as thresholds. We used 100 thresholds to obtain each spectrum. The highest value K_{max}^C in this series, i.e., $K_{max}^C = max\{c_i\}$, is called the Kolmogorov complexity spectrum highest value (KCM). The KC complexity carries less information about the complexity of the time series than the KC spectrum, i.e., KCM, does. Namely, the KC gives average information about the complexity of the time series. In contrast, KCM carries the information about the highest complexity among all complexities in the spectrum. Therefore, this measure should be included to better understand the system's randomness and organization.

To illustrate the usefulness of the above measures, we consider two time series of the instantaneous velocity measured at $z/k = 0.067$ (purple) and $z/k = 2.667$ (red), respectively (see Figure 2). Figure 3 shows that the flow dynamics of turbulence varies with the relative height z/k and the Kolmogorov complexity spectrum's highest value (KCM). In fact, at $z/k = 0.067$ (purple curve), the KC is higher than at $z/k = 2.667$ (red curve). Therefore, we can conclude that at $z/k = 0.067$, the turbulence regime is more random, i.e., shows a more pronounced presence of stochastic components than at $z/k = 2.667$. In other words, these spectra can effectively point out the dissimilarities between the turbulent regimes for the two different relative heights.

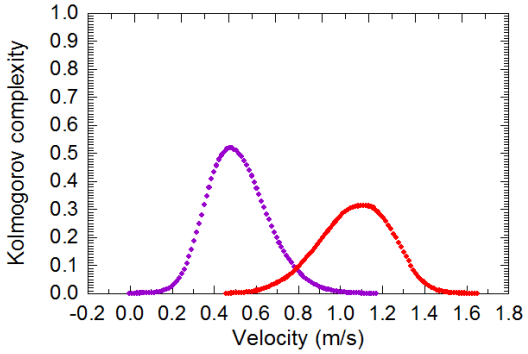

Figure 3. The Kolmogorov complexity spectra for two time series of the velocity measured at $z/k = 0.067$ (purple) and $z/k = 2.667$ (red), respectively. Both lines are fitting curves, which are obtained from the calculated discrete values of the Kolmogorov complexity (KC) spectra.

3. Experimental Details

The experiments were performed in a laboratory channel with a variable bed slope, which was 8 m long and 0.4 m wide. The channel slope is the difference in elevation between two points of the channel bed divided by the distance between them measured along the channel. Vegetation covered the bed of the channel and consisted of rigid cylinder rods of the same height and diameter (k = 0.015 m, d_c = 0.004 m), set in different aligned arrangements (rectangles or squares), with three different densities λ_d, and Q is the volumetric flow rate (Table 1).

Table 1. Experimental conditions.

Test	D1	D2	D3
λ_d ($m^2 \cdot m^{-2}$)	0.024	0.048	0.096
ϑ (°)	0.03	0.02	0.03
Q ($l\ s^{-1}$)	33	22	22
h_u (cm)	6.35	6.44	6.29
Cylinders per unit area	400	800	1600

Figure 4 depicts the experimental channel where the experiment measurements were carried out and the cylinders, modeling rigid aquatic plants in the literature on vegetated flows [11,13]. In Figure 5, sets of the cylinders for different densities are represented with the vertical cross section in the A-A direction (see figure).

Experimental measurements were carried out in a section set at 12 cm from the hydraulic left of the channel, at the center of a rectangle or a square of cylinders. Vegetation covered the entire length of the channel, and experimental measurements were carried out in a section where the flow attained the uniform condition, i.e., the slope of the free surface was equal to the channel slope ϑ. The friction velocity u_* was calculated as $\sqrt{g\vartheta(h_u - k)}$ where g is the gravitational acceleration, while h_u is the flow depth in uniform conditions. The validity of this method was confirmed comparing its results with the peak values of the Reynolds stress, in another condition of uniform flow on a vegetated bed. Vegetation density λ_d was evaluated as the total canopy frontal area per unit area. The vegetation was always fully submerged with submergence h_u/k of about four.

Figure 4. Experimental channel (**a**) and cylinders as the model of rigid vegetation (**b**) [40].

The measurements of the instantaneous velocity in the direction of the main stream were performed with a one-component Polytec LDV system used in back scatter mode. The system was accurately aligned [41]. Due to water characteristics, no artificial seeding of the channel was used. The light source was a laser diode with a maximum power of 65 mW at a wavelength between 810 nm and 840 nm. The sampling frequency for the LDV measurements was 2000 Hz, and in order to obtain a sufficient number of strong bursting events, the acquisition time in each measurement point was equal to 135 s, so the sampling data of N = 270,000 of instantaneous velocity were collected and analyzed for turbulence statistics. The sampling frequency and the sampling time were selected to accurately represent the investigated flow and to result in data files supported by the software package used for computation. For commercial LDV systems, processor information is converted to velocity through software using the equation $V_x = K\lambda f_D$, where: V_x is the component of the instantaneous velocity along the streamwise direction; $K\lambda$ is the calibration factor in nm; $K = 1/2\sin\theta$; θ is the half-angle of the beam intersection angle in the measurement point set in the focus of the lens; λ is the laser wavelength in nm; and f_D is the Doppler frequency in MHz. The uncertainty in the measurement of the Doppler frequency can be assumed negligible [42], and for modern signal processors, the uncertainty of the laser wavelength λ is negligible [43]. The uncertainty in the medium velocity has been evaluated

to be about 0.3%. For each measurement point, the analog signal from the processor, by means of an oscilloscope, was accurately checked, to verify the Doppler signal quality [11,44]. LDV provides flow velocity data having high quality, and therefore, it still remains the preferred measuring technique for complex turbulent flows' study [44], confirming the adequacy of fluid velocity measurements based on frequency acquisition and analysis [11,13,45,46]. In Nezu and Sanjou [13] and Nezu and Onitsuka [45], the most accurate measurement device was LDV, and the accuracy of PIV (particle image velocimetry) was evaluated by comparing PIV data with LDA data.

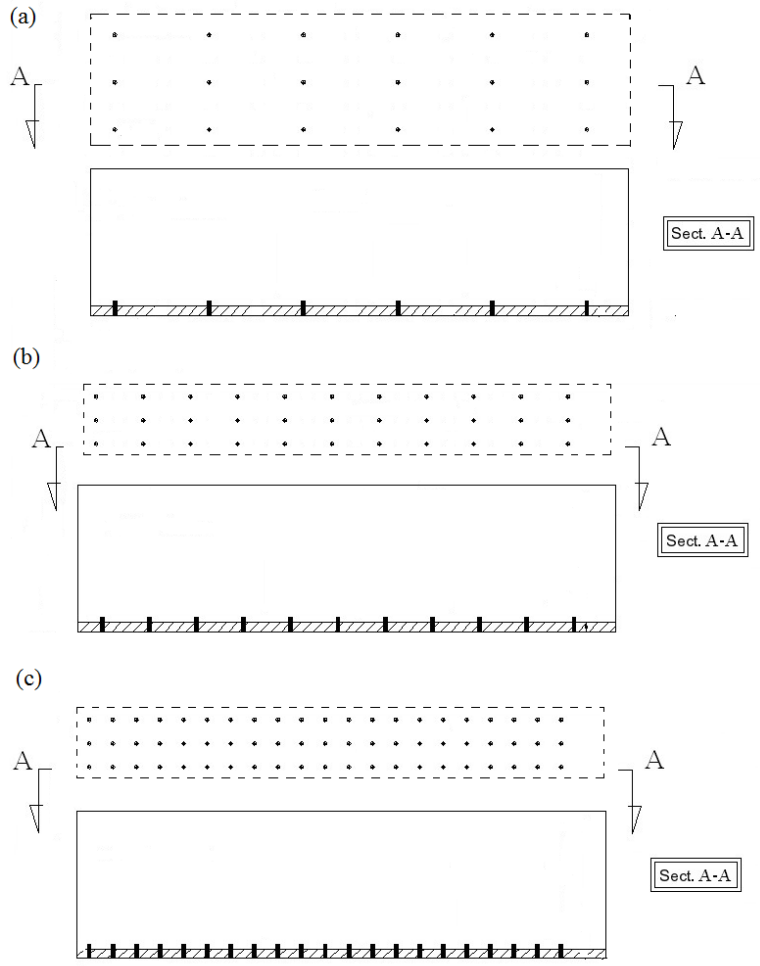

Figure 5. Sets of the cylinders for density equal to: (**a**) D1 = 0.024; (**b**) D2 = 0.048 and (**c**) D3 = 0.096.

4. Results

4.1. Basic Statistical Parameters

Furthermore, in order to deepen the canopy-flow interaction, it could be interesting to evaluate the distributions of skewness and kurtosis of the instantaneous streamwise velocities' distributions, for the different densities of the canopy. In fact, skewness quantifies how symmetrical the distribution of the instantaneous streamwise velocities is, and kurtosis quantifies whether the peak of the data distribution matches the normal distribution. In particular, if the distribution of the instantaneous streamwise velocities is symmetrical, the positive and negative values balance each other, and the skewness is close to zero. If the distribution is not symmetrical, the skewness is positive if the distribution is skewed to the right and negative if skewed to the left. In addition, the skewness values that are generally negative or positive denote the presence, respectively, of lower or higher velocity values from the mean

value. Positive kurtosis values indicate a relatively peaked distribution, while negative kurtosis values indicate a relatively flat distribution. Therefore, the distributions of skewness and kurtosis (Table 2) show the differences between the coherent eddy structure within and just above the canopy and the eddy structure in the surface layer above the canopy. In particular, skewness goes from positive inside the canopy to negative above the canopy, where it tends to a Gaussian value. Kurtosis, within the canopy, is positive or negative (for the lowest canopy density) and above the canopy is negative, tending to the Gaussian value. Within the canopy, denser canopies have larger positive skewness and kurtosis. The trend of skewness and kurtosis confirms the scheme of the RSL, proposed for mean velocity and turbulent intensities: in the RS3 zone, for all the densities, skewness and kurtosis tend to the Gaussian value, confirming that the turbulent structure is similar to the one in a boundary layer; in the RS1 zone, larger skewness and kurtosis for denser canopies show the role of the sweeps of high velocity fluid down into the dense canopy space; in the RS2 zone, the turbulent structure is analogous to a mixing layer, where coherent eddies are generated due to the inflection-point instability and govern the vertical transport of momentum [11,13].

Table 2. Basic statistical parameters of the instantaneous streamwise velocities at five relative heights.

D1				
z/k	**Mean**	**SD**	**Skew**	**Kurtosis**
2.667	1.49	0.1604	−0.3934	−0.1155
1.8	1.35	0.2035	−0.2483	−0.4652
1	1.19	0.2069	−0.0657	−0.3909
0.4	0.97	0.1878	0.1676	−0.2180
0.067	0.90	0.1715	0.1985	−0.1444
D2				
z/k	**Mean**	**SD**	**Skew**	**Kurtosis**
2.667	1.59	0.1412	−0.3644	−0.1181
1.8	0.93	0.1648	−0.1675	−0.4227
1	0.76	0.1677	0.0389	−0.3873
0.4	0.61	0.1490	0.2396	−0.0580
0.067	0.53	0.1313	0.3232	0.0796
D3				
z/k	**Mean**	**SD**	**Skew**	**Kurtosis**
2.667	1.65	0.1609	−0.2996	−0.1691
1.8	0.97	0.1871	−0.1501	−0.4304
1	0.73	0.1923	0.1616	−0.4566
0.4	0.58	0.1588	0.3942	0.0678
0.067	0.49	0.1369	0.4273	0.3019

The turbulence statistics applied on measured values of velocity quantify how the flow within and just above the canopy behaves as a perturbed mixing layer. Here, the turbulence intensity is expressed as $\bar{u} = \sqrt{\sum_{i=1}^{N} u_i^2 / N}$ and normalized by the friction velocity u_*, i.e., $\sigma_u = \bar{u}/u_*$. The number of the samples in experiments was N = 270,000. Figure 6 shows the measured profiles of σ_u for all three densities D1, D2 and D3. From this figure, it is seen that for $z/k < 1$, turbulence intensity σ_u is remarkably damped and is more pronounced for higher λ_d. In particular, σ_u is in the range from 1.69 (D1), for a sparse canopy, which is typical for rough-wall layers, to about 1.49 (D3), for a dense canopy, which is typical for mixing layers. Those values are close to the ones reported by Jimenez [10].

4.2. Shannon Entropy as a Measure of the Order or Disorder of the Flow

When stable laminar flows evolve toward the turbulence, they become high-order and complex, exhibiting irregular-like motions with organized dissipative arrangements. In order to precisely specify

their fields (velocity and displacement), more parameters are required than for the description of laminar flows, i.e., topological measures that quantify the order or disorder of the flow, as the Shannon entropy, which has been already used in the analysis of geophysical fluids [47,48]. The Shannon entropy SH is defined as $SH = -\sum p_i ln p_i$, where p_i is a discrete probability distribution satisfying the following conditions: $p_i \geq 0; \sum p_i = 1$ and $p_{i \cup j \cup ...} = p_i + p_j + ...$ [49]. In our calculations, p_i is defined as the probability that the velocity amplitude falls within interval $u_i + du$, where du was obtained by dividing the entire interval of velocity amplitudes into one million intervals, due to the fact that the precision of the measurements was up to the sixth decimal place. Figure 7 shows that SH is the highest in the mixing layer ($1 < z/k < 2$), where the turbulence intensity σ_u is the highest. However, it decreases towards the free surface. This trend of SH coincides with the conclusions by Wijesekera and Dillion [47]. The decrease of the SH, going to the rough bed, can be addressed to the occurrence of smaller eddies carrying a smaller amount of energy.

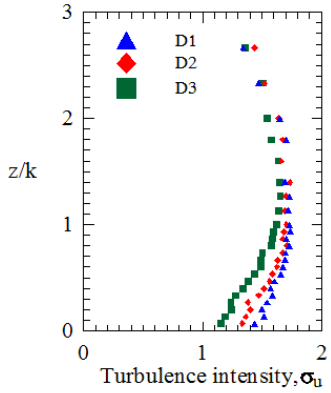

Figure 6. Turbulence intensity $\bar{u}$ normalized by u_* against the relative flow depth ratio of the vertical distance from bed z to the height k of the cylinders used for modeling canopies for all the densities [40].

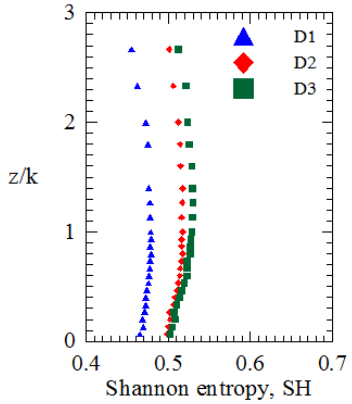

Figure 7. Shannon entropy (SH) versus the relative flow depth ratio of the vertical distance from bed z to the height of the cylinders used for modeling canopies k [40].

To avoid confusion in the following discussion, we make some comments. Namely, the term complexity in physical systems has the connotation of an explicit measure of the probability of the state of the system. It is a mathematical measure that should not be equalized with entropy in statistical mechanics [39]. Thus, Shannon's entropy refers to dissimilarities between amplitudes in a time series, while the Kolmogorov complexity refers to the apparent sequence disorder of amplitudes in a time series. This complexity that we intuitively understand as a measure range between uniformity and total randomness (Figure 3 in [39]) is of interest in this paper. Comparing Figures 7 and 8, they seem overall to have quite symmetrical trends. In Figure 7, the SH entropy weakly increases in the RS1 zone;

it has a constant value in the RS2 zone and then decreases in the RS3 zone. This trend is clearer for sparse bed roughness elements (D1), but anyway, the density of the bed roughness elements seems to affect SH entropy; in fact, lower SH entropy corresponds to sparse density (D1).

4.3. Kolmogorov Complexity

One of the main concerns in turbulence studies is the estimate of the temporal and, if possible, spatial characteristics of the existing eddy in the flow. Therefore, the integral length scale (L) is important in characterizing the structure of the turbulence. It is a measure of the longest correlation distance between the flow velocity at two points of the flow field. In Table 3 are given values of integral length scales (L) for different vegetation densities normalized by canopy height (k) at all relative heights.

Table 3. Integral length scales (L) for different vegetation densities normalized by canopy height (k) at all relative heights.

Integral Length Scale L/k			
z/k	**D1**	**D2**	**D3**
2.667	3.206	2.710	2.170
2.333	3.310	2.621	2.164
2.000	5.099	2.480	1.919
1.800	3.649	2.465	1.822
1.600	3.403	2.135	1.807
1.400	2.346	2.612	1.537
1.267	2.337	1.843	1.483
1.133	2.014	1.688	1.458
1.000	1.912	1.446	1.341
0.933	1.950	1.407	1.235
0.867	1.650	1.236	1.294
0.800	1.650	1.436	1.245
0.733	1.541	1.262	1.048
0.667	1.481	1.057	1.048
0.600	1.414	1.055	1.044
0.533	1.391	1.065	0.966
0.467	1.242	0.958	0.966
0.400	1.207	0.880	0.861
0.333	1.247	0.777	0.802
0.267	1.064	0.727	0.843
0.200	0.927	0.764	0.751
0.133	1.078	0.642	0.644
0.067	0.776	0.593	0.497

Some characteristics of the non-dimensional distributions of longitudinal integral length scales can be observed. For all the densities, the length scales are largest above the vegetation and are reduced progressively with depth into the vegetation. In the upper part of the vegetation, integral length scales are strongly influenced by the density, increasing as vegetation density decreases [50]. They are larger in sparser vegetation density because the confinement due to vegetation is looser. Near the canopy top, the turbulence length scales are of order of k. This is very similar to other reported measurements [51] and shows that the sweeps inferred from the skewness profiles are large turbulent structures, coherent in the streamwise direction over distances comparable with h. For all the densities, eddy scales are significantly compressed by the existence of the free surface [13].

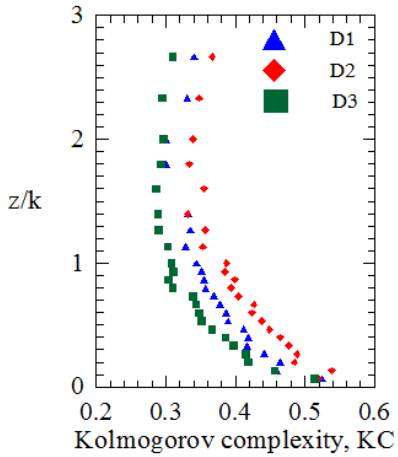

Figure 8. Kolmogorov complexity (KC) versus the relative flow depth ratio of the vertical distance from bed z to the height of the cylinders used for modeling canopies k [40].

We see that neither use of conventional measure for turbulence in canopy flow, nor the Shannon entropy give us quantitative information about the degree of the randomness. These two measures inform either about the intensity of the turbulence in canopy (σ_u) or about dissimilarities between amplitudes in a time series. Figure 9a shows the distributions of velocity frequencies for selected relative flow depths, for D3 canopy density (in further analysis, we use only this density for reasons of simplicity, but without losing the generality of the analysis).

Namely, going from $z/k = 0.067$ towards $z/k = 2.667$, the values of the peak velocity for all curves follow the profile of σ_u versus z/k in Figure 6. Looking at this figure, we can only say, in regard to randomness, that its highest value is expected for velocities around the peak of the distribution. There is no conclusion (neither quantitative, nor descriptive) that we can attain concerning which velocity distribution represents less or high randomness. However, the Kolmogorov complexity spectrum or more precisely its highest value (KCM) carries that information quantitatively through a value expressing the randomness of turbulence in the canopy. From Figure 9b, we see that randomness is higher for lower velocities, but decreases going towards $z/k = 1$. After this relative flow depth, it becomes higher.

Figure 8 shows that the value of KC decreases with height from the rough-wall to the mixing layer ($1 < z/k < 2$). This can be explained by the fact that in the RS1 zone, flow is dominated by smaller eddies (see Figure 1a) contributing to the higher randomness that becomes lower in the mixing layer having a constant value in this zone. This is because eddies in the RS2 zone are larger and coherently organized, without the possibility to introduce more randomness in the flow. Above the mixing layer (RS3 zone), KC slightly increases since eddies become smaller. Following this analysis, it seems that the KC could be related to eddy sizes.

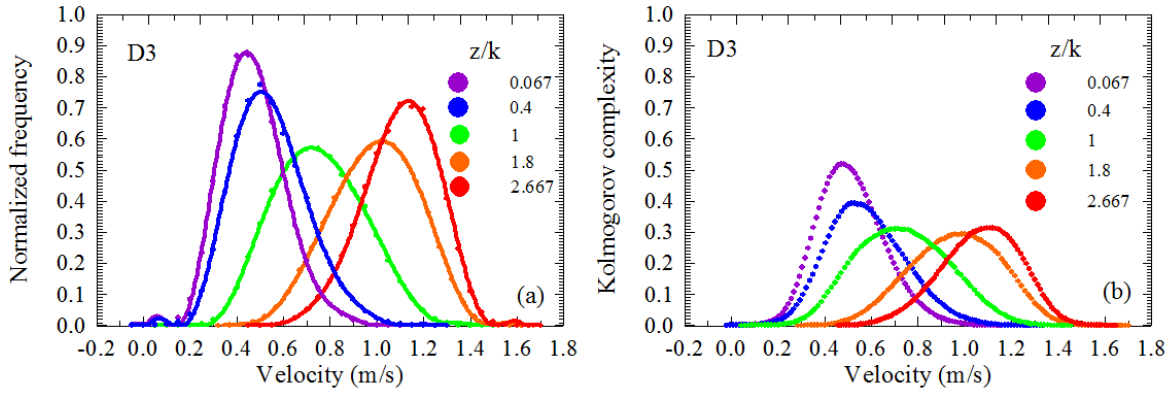

Figure 9. Distributions of velocity frequencies (**a**) and Kolmogorov complexity spectra (**b**) for selected relative flow depth, for D3 canopy density.

From Figure 9 is seen an interesting feature that the similarities between velocity spectra and Kolmogorov complexity spectra differ up to the constant. Whether this feature is a consequence either of physical sources, or the mathematical nature of the applied algorithm, or it is just the characteristics of these experiments remains to be deepened. Therefore, for a more reliable conclusion, there is a need for more experiments with different setups for the parameters.

5. Conclusions

In this paper, a methodology to quantify the randomness in turbulent flows with the canopy using experimental data collected in a laboratory channel with a variable bed slope and three different canopy densities is described. An analysis based on the Kolmogorov complexity and measures derived from it (the Kolmogorov complexity spectrum and its highest value) was proposed. First, for all the densities, the turbulence was analyzed with the classical turbulence statistics through the vertical profiles of turbulence intensity. In addition, the Shannon entropy was calculated for all densities, to describe the turbulence in terms of irregular or random flow. Finally, it was shown that the Kolmogorov complexity and measures based on it, for all densities, provides quantification of the degree of turbulence in relation to randomness rather than if it is descriptively expressed in terms of irregular or random. Finally, probably the most important result of this paper is that the value of KC is connected with the turbulent eddies' size, and it could be helpful in explaining the coherent structure in canopy turbulent flow.

Acknowledgments: This paper was realized as a part of Project No. 43007 financed by the Ministry of Education and Science of the Republic of Serbia for the period 2011–2017. Paola Gualtieri and Carlo Gualtieri acknowledge the support by the MIUR PRIN 2010–2011 research project "HYDROCAR".

Author Contributions: Paola Gualtieri and Carlo Gualtieri conceived and designed the experiments; Paola Gualtieri performed the experiments; Gordan Mimić and Ilija Arsenić analyzed the data; Dragutin Mihailović contributed analysis tools; Dragutin Mihailović wrote the paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gurzadyan, V.G. Kolmogorov Complexity as a Descriptor of Cosmic Microwave Background Maps. *EPL* **1999**, *46*, 114–117.
2. Gell-Mann, M.; Lloyd, S. Information Measures, Effective Complexity, and Total Information. *Complexity* **1996**, *2*, 44–52.
3. Gell-Mann, M.; Lloyd, S. Effective Complexity. In *Nonextensive Entropy: Interdisciplinary Applications*; Gell-Mann, M., Tsallis, C., Eds.; Oxford University Press: Oxford, UK, 2003.
4. Kreinovich, V.; Kunin, I.A. Kolmogorov Complexity and Chaotic Phenomena. *Int. J. Eng. Sci.* **2003**, *41*, 483–493.

5. Abarzhi, S. Review of theoretical modelling approaches of Rayleigh–Taylor instabilities and turbulent mixing. *Philos. Trans. R. Soc. A* **2010**, *368*, 1809–1828.
6. Lorenz, R.D. Maximum Frictional Dissipation and the Information Entropy of Windspeeds. *J. Non-Equilib. Thermodyn.* **2002**, *27*, 229–238.
7. Weck, P.J.; Schaffner, D.A.; Brown, M.R.; Wicks, R.T. Permutation entropy and statistical complexity analysis of turbulence in laboratory plasmas and the solar wind. *Phys. Rev. E* **2015**, *91*, 023101.
8. Ichimiya, M.; Nakamura, I. Randomness representation in turbulent flows with Kolmogorov complexity (in mixing layer). *J. Fluid Sci. Technol.* **2013**, *8*, 407–422.
9. Pope, S.B. *Turbulent Flows*; Cambridge University Press: Cambridge, UK, 2000.
10. Jimenez, J. Turbulent flows over rough walls. *Annu. Rev. Fluid Mech.* **2004**, *36*, 173–196.
11. Poggi, D.; Porporato, A.; Ridolfi, L.; Albertson, J.D.; Katul, G.G. The effect of vegetation density on canopy sub-layer turbulence. *Bound. Lay. Meteorol.* **2004**, *111*, 565–587.
12. Flack, K.A.; Schultz, M.P.; Shapiro, T. Experimental support for Townsend's Reynolds number similarity hypothesis on rough walls. *Phys. Fluids* **2005**, *17*, 035102.
13. Nezu, I.; Sanjou, M. Turbulence structure and coherent motion in vegetated canopy open-channel flows. *J. Hydro-Environ. Res.* **2008**, *2*, 62–90.
14. Cushman-Roisin, B.; Gualtieri, C.; Mihailović, D.T. Environmental Fluid Mechanics: Current issues and future outlook. In *Fluid Mechanics of Environmental Interfaces*, 2nd ed.; Gualtieri, C., Mihailović, D.T., Eds.; CRC Press/Balkema: Leiden, The Netherlands, 2012.
15. Raupach, M.R.; Finnigan, J.J.; Brunet, Y. Coherent eddies and turbulence in vegetation canopies: The mixing-layer analogy. *Bound. Lay. Meteorol.* **1996**, *78*, 351–382.
16. Hussain, A.K.M.F. Coherent Structures—Reality and Myth. *Phys. Fluids* **1983**, *26*, 2816–2850.
17. Nepf, H.M. Flow and transport in regions with aquatic vegetation. *Annu. Rev. Fluid Mech.* **2012**, *44*, 123–142.
18. Rogers, M.M.; Moser, R.D. Direct simulation of a self-similar turbulent mixing layer. *Phys. Fluids* **1994**, *6*, 903–923.
19. Tsuji, Y.; Nakamura, I. The fractal aspect of an isovelocity set and its relationship to bursting phenomena in the turbulent boundary layer. *Phys. Fluids* **1994**, *6*, 3429–3441.
20. Li, M.; Vitanyi, P. *An Introduction to Kolmogorov Complexity and Its Applications*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 1997.
21. Lempel, A.; Ziv, J. On the complexity of finit sequence. *IEEE Trans. Inf. Theory* **1976**, *22*, 75–81.
22. Sellers, P.; Mintz, Y.; Sud, Y.C.; Dachler, A. A Simple Biosphere Model (SIB) for Use within General Circulation Models. *J. Atmos. Sci.* **1986**, *43*, 505–531.
23. Batchelor, G.K. *The Theory of Homogeneous Turbulence*; Cambridge University Press: Cambridge, UK, 1956.
24. Fricsh, U. *Turbulence: The Legacy of A. N. Kolmogorov*; Cambridge University Press: Cambridge, UK, 1995.
25. Terwijn, S.A. The Mathematical Foundations of Randomness. In *The Challenge of Chance: A Multidisciplinary Approach from Science and the Humanities*; Landsman, K., van Wolde, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2016.
26. Kolmogorov, A.N. *Grundbegriffe der Wahrscheinlichkeitsrechnung*; Springer: Berlin/Heidelberg, Germany, 1933.
27. Sato, H. Laminar-Turbulent Transition in Free Shear Flow. In *Progress in Fluid Mechanics -Turbulent Flow*; Tani, I., Ed.; Maruzen: Tokyo, Japan, 1980. (In Japanese)
28. Pietri, L.; Petroff, A.; Amielh, M.; Anselmet, F. Turbulence characteristics within sparse and dense canopies. *Environ. Fluid Mech.* **2009**, *9*, 297–320.
29. Pietri, L.; Petroff, A.; Amielh, M.; Anselmet, F. Turbulent flows interacting with varying density canopies. *Mech. Ind.* **2009**, *10*, 181–185.
30. Feldman, D.P.; Crutchfeld, J.P. Measures of statistical complexity: Why? *Phys. Lett. A* **1998**, *238*, 244–252.
31. Kolmogorov, A. Logical basis for information theory and probability theory. *IEEE Trans. Inf. Theory* **1968**, *14*, 662–664.
32. Cerra, D.; Datcu, M. Algorithmic relative complexity. *Entropy* **2011**, *13*, 902–914.
33. Kaspar, F.; Schuster, H.G. Easily calculable measure for the complexity of spatiotemporal patterns. *Phys. Rev. A* **1987**, *36*, 842–848.
34. Zhang, X.S.; Roy, R.J.; Jensen, E.W. EEG complexity as a measure of depth of anesthesia for patients. *IEEE Trans. Biomed. Eng.* **2001**, *48*, 1424–1433.

35. Radhakrishnan, N.; Wilson, J.D. An alternate partitioning technique to quantify the regularity of complex time series. *Int. J. Bifurc. Chaos* **2000**, *10*, 1773–1779.
36. Small, M. *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance*; World Scientific: Singapore, 2005.
37. Ferenets, R.; Lipping, T.; Anier, A. Comparison of entropy and complexity measures for the assessment of depth of sedation. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 1067–1077.
38. Hu, J.; Gao, J.; Principe, J.C. Analysis of biomedical signals by the Lempel-Ziv complexity: The effect of finite data size. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 2606–2609.
39. Mihailović, D.T.; Mimić, G.; Nikolić-Djorić, E.; Arsenić, I. Novel measures based on the Kolmogorov complexity for use in complex system behavior studies and time series analysis. *Open Phys.* **2015**, *13*, 1–14, doi:10.1515/phys-2015-0001.
40. Mihailović, D.T.; Mimić, G.; Gualtieri, P.; Arsenić, I.; Gualtieri, C. Randomness representation in turbulent flows with bed roughness elements using the spectrum of the Kolmogorov complexity. In Proceedings of the 8th International Congress on Environmental Modelling and Software, Toulouse, France, 10–14 July 2016; pp. 4–11.
41. Annoni, M.; Monno, M.; Scarano, G. Uncertainty components interaction on the water jet velocity measurement by laser velocimetry. In Proceedings of the American WJTA Conference and Expo, Houston, TX, USA, 18–20 August 2009.
42. Annoni, M. Water jet velocity uncertainty in laser Doppler velocity measurements. *Measurement* **2012**, *45*, 1639–1650.
43. ITTC Recommended Procedure and Guidelines. *Uncertainty Analysis Laser Doppler Velocimetry Calibration*; ITTC: Kgs. Lyngby, Denmark, 2008.
44. Ristić, S.; Ilić, J.; Čantrak, D.; Ristić, O.; Janković, N. Estimation of laser Doppler anemometry measuring volume displacement in cylindrical pipe flow. *Therm. Sci.* **2012**, *16*, 1027–1042.
45. Nezu, I.; Onitsuka, K. Turbulent structure in partly vegetated open channel flows with LDA and PIV measurements. *J. Hydraul. Res.* **2011**, *39*, 629–642.
46. Ghannam, K.; Poggi, D.; Porporato, A.; Katul, G. The spatio-temporal statistical structure and ergodic behaviour of scalar turbulence within a rod canopy. *Bound. Lay. Meteorol.* **2015**, *157*, 447–460.
47. Wijesekera, H.W.; Dillion, T.M. Shannon entropy as an indicator of age for turbulent overturns in the ocean thermoclines. *J. Geophys. Res.* **1997**, *102*, 3279–3291.
48. Wesson, H.K.; Katul, G.G.; Siqueira, M. Quantifying organization of atmospheric turbulent eddy motion using nonlinear time series analysis. *Bound. Lay. Meteorol.* **2003**, *106*, 507–525.
49. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–623.
50. Green, S.R.; Grace, J.; Hutchings, N.J. Observations of turbulent air flow in three stands of widely spaced Sitka spruce. *Agric. For. Meteorol.* **1995**, *74*, 205–225.
51. Brunet, Y.; Finnigan, J.J.; Raupach, M.R. A wind tunnel study of air flow in waving wheat: Single-point velocity statistic. *Bound. Lay. Meteorol.* **1994**, *70*, 95–132.