www.czasopisma.pan.pl PAN POLSKA AKADEMIA NAUK www.journals.pan.pl

BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 67, No. 5, 2019
DOI: 10.24425/bpasts.2019.130876

The use of Lempel-Ziv complexity to analyze turbulence and flow randomness based on velocity fluctuations

S. BLONSKI, A. PREGOWSKA, T. MICHALEK, and J. SZCZEPANSKI*

Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawinskiego 5B, 02-106 Warsaw, Poland

Abstract. One of the mathematical tools to measure the generation rate of new patterns along a sequence of symbols is the Lempel-Ziv complexity (LZ). Under additional assumptions, LZ is an estimator of entropy in the Shannon sense. Since entropy is considered as a measure of randomness, this means that LZ can be treated also as a randomness indicator. In this paper, we used LZ concept to the analysis of different flow regimes in cold flow combustor models. Experimental data for two combustor's configurations motivated by efficient mixing need were considered. Extensive computer analysis was applied to develop a complexity approach to the analysis of velocity fluctuations recorded with hot-wire anemometry and PIV technique. A natural encoding method to address these velocity fluctuations was proposed. It turned out, that with this encoding the complexity values of the sequences are well correlated with the values obtained by means of RMS method (larger/smaller complexity larger/smaller RMS). However, our calculations pointed out the interesting result that most complex, this means most random, behavior does not overlap with the "most turbulent" point determined by the RMS method, but it is located in the point with maximal average velocity. It seems that complexity method can be particularly useful to analyze turbulent and unsteady flow regimes. Moreover, the complexity can also be used to establish other flow characteristics like its ergodicity or mixing.

Key words: turbulence, complexity, entropy, randomness.

1. Introduction

Last years extensive experimental, theoretical and numerical efforts have been done to understand important issues of fluid dynamics [1–3]. There is no commonly accepted theory developed for recognition of fundamental qualitative properties of the flow (turbulent – laminar, steady – unsteady, mixing – non-mixing), while in many practical applications there is a great need to find effective methods to assess these properties [4–6]. Computational fluid dynamics was also carried out to understand the combustion phenomena in the dual-fuel mode. In [7] emission of nitrous oxides and particle materials in a dual-fueled constant-speed engine were analyzed experimentally and numerically. Numerical simulation was performed by KIVA3V, and its results showed good agreements with the experimental results under cylinder pressure.

Turbulence is often expressed in terms of either irregular or random fluid flows [8]. Therefore, interesting problem is to study relations between turbulence and randomness of the flow [5, 10]. Nowadays, important issues regarding turbulence are connected with its identification, analysis, and classification of flow patterns [11]. Turbulent flow leads to fluctuations in the scalar field through turbulent convection and consequently affects the velocity field. This means that the type of fluctuations of velocity is one of the most important features to characterize the flow [12].

*e-mail: jszczepa@ippt.pan.pl

Manuscript submitted 2019-03-21, revised 2019-07-08, initially accepted for publication 2019-07-29, published in October 2019

The development of new methods being alternative to traditional approaches is of high importance. Recently, to describe turbulence authors try to apply concepts derived from Information Theory [13] like entropy, transfer entropy, permutation entropy and complexity [14]. The oil–gas–water three-phase flow in a vertical upward pipe with the use of complexity measures was studied in [15]. It turned out that the combination of Lempel–Ziv complexity and approximate entropy can serve as a unique classification criterion of three-phase flow patterns. In [16] the Lempel–Ziv algorithm and a multi-scaling approach were used to assess precipitation complexity. The methods allow characterizing precipitation complexity in the mountainous area and in the plain terrain.

To understand turbulence, production rate of entropy and complexity of fluid flows are extensively investigated numerically [17, 18, 20]. In [18] the directed co-flow effects on local entropy generation rate in turbulent and heated round jets were studied. It was shown that the directed co-flow with a positive angle enhances the mixing. In [19] transfer entropy was applied to study synthetic model of fluid turbulence, namely the Gledzer-Ohkitana-Yamada shell model. Using this tool the presence of a direct cascade along the scales in the shell model and the locality of the interactions in the space of wavenumbers were confirmed. Entropy generation with variable density in the turbulent plane jets was investigated in [20]. Computations were carried out with eddy viscosity model of turbulence. It turned out, that the high value of energy generation is correlated with higher inlet hot jet temperature. The results obtained indicated that the merit number increases progressively to reach an asymptotic value along the flow direction as the inlet jet temperature grows.

In this paper we apply Information Theory [13, 21–23] based concept, namely Lempel–Ziv complexity [24] to analyze turbulence in gas turbine combustor's configurations motivated by efficient mixing need. The results obtained were compared with the results obtained with the application of the standard indicator of turbulence, i.e. RMS indicator [25]. It turned out that the complexity values of the encoded sequences are well correlated with the values obtained by means of the RMS method (larger/smaller complexity larger/smaller RMS). However, our calculations pointed out also an interesting observation that most complex (that is most random) behavior does not overlap with the "most turbulent" point determined by RMS method but it locates in the point with maximal average velocity.

2. Materials and Methods

2.1. Description of the experimental data. The experiment was performed to simulate the flow in the gas turbine combustors [26] using transparent laboratory models. The first goal was to develop innovative combustor for small gas turbine, in which flameless oxidation of the fuel at high temperature ensures low emission of nitrogen oxides (NO_x). Nitrogen oxides are a group of harmful gases, responsible for air pollution (e.g. photochemical type of smog) and numerous respiratory diseases in humans. Tested combustors were designed to address this goal. The main target of the experimental investigations was to get physical insight into turbulent fluctuations of the main vortex, responsible for the efficient mixing of fuel and air [26–28]. Such type of measurements was performed for the laboratory models in order to deliver details about flow structures, important for optimization of combustor geometry and validation of numerical models (necessary information for predictions and optimization of real combustors). Two experimental configurations were tested (Figs 1a, 1b). For both configurations 26 uniformly distributed points x_k, $k = 1, 2, \ldots, 26$ with the distance 5 mm in between, along the central line were considered in the procedures described furthermore. First and last point were located 5 mm from chamber walls. Intuitively, mixing seems to be more intense in configuration B, due to its more complex shape of the channels.

In order to get complementary insight into the physical phenomena, we applied two methods of measurements. In the Particle Image Velocimetry method (PIV) average values of velocities in some small "neighborhood" of a point under consideration were measured, while in the second method ("hot wire") the values being precisely velocities at such points were recorded.

In PIV, which is an optical method, the investigated fluid was first seeded with the very small particles which we assume follow precisely the flow. Next, a light scattered by these particles is recorded by the digital camera. To acquire images high-speed video system (FASTCAM Ultima 40K, Photron) and high-resolution PIV system with 12-bit CCD camera (SensiCam, PCO Imaging) were used. Illumination was received with using CW 5 W argon laser (Argon-ion 120, ILA) and double-pulsed Nd:YAG 35 mJ laser (Solo PIV, NewWave Research Inc.) with minimum

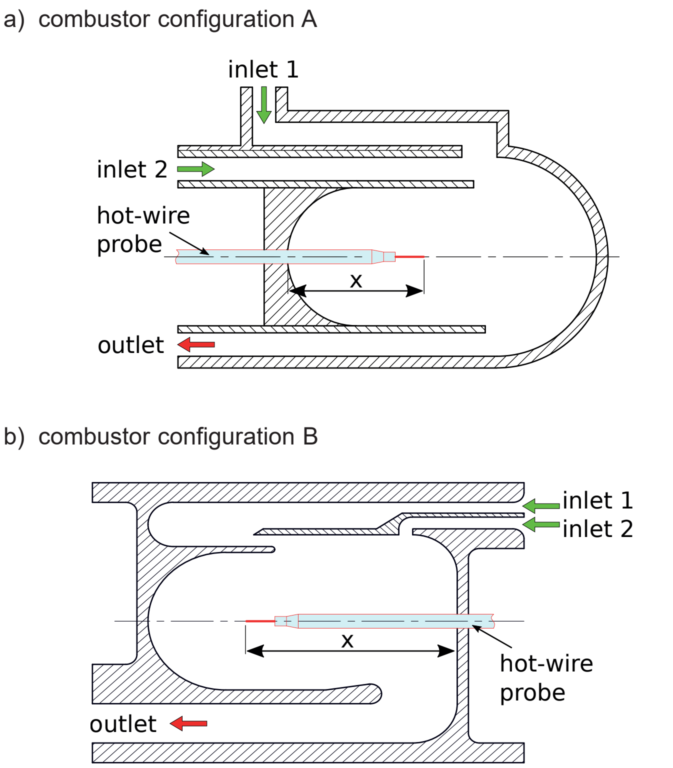

Fig. 1. Tested onfigurations of the combustors. They were designed to address efficient mixing of fluel and air, and simultaneously to ensure low emission of nitrogen oxides

time interval 200 ns between pulses. The typical investigated flow velocities were in the range 10–25 m/s. Observation of correlations of the successive images that include information about particles displacement allows for evaluation of the velocity field. The digital images are divided into small subareas called "interrogation areas". In our experiments, dimensions of the interrogation areas were 32×32 pixels what corresponds to 3.5×3.5 mm. The local displacement vectors for the images were determined for each interrogation area. For this reason, velocity field obtained from the PIV technique is the mean velocity, averaged over each interrogation area. The sampling rate of measurements in the PIV method was 4 Hz.

The second method, i.e. thermoanemometry (also called "hot-wire" method) is high-precision, indirect technique for point velocity measurement, especially useful to the highly unsteady flows. In the performed experiments we used hot-wire method with one-wire sensor. This method ensures precise measurement of velocity magnitude only, i.e. the length of the velocity vector in the plane perpendicular to the wire of the sensor, without information about vector direction. The velocity of the flow was evaluated by the measurement of the heat losses of the very thin metallic wire (in our experiment about 1 micrometer in diameter) that was heated up to about 200°C by the electric current and placed into the investigated flow. Thermal balance includes heat generation, heat conduction, and forced convection. Because the temperature is relatively low, heat radiation was neglected. To complete this description it is worth to stress that for extremely slow flows, natural convection has to be considered. After calibration, thermal balance allowed to calculate the instantaneous value of the fluid flow velocity at a point being considered. In this method, the sampling rate for each measurement was 10 kHz. From the conceptual point of view, the basic difference between these two methods is that the value of a measurement in a given point determined by PIV is, in fact, the average value of velocities in some small "neighborhood" of this point while hot-wire measurements correspond precisely to the velocity already at the given point. The velocities measurements obtained by these techniques were subsequently encoded using the procedure described in the next section and the complexity of these sequences was analyzed.

2.2. Lempel–Ziv complexity and encoding procedure. There are a lot of different kinds of complexity measures applied successfully in numerous fields [21, 23, 29, 30]. All of them share the common property of providing quantitative information coming from the structure of symbol sequence containing information about its source (in our case about the flow). In this paper, we propose the analysis of fluid flow by means of the Lempel–Ziv complexity as defined in [24]. This is a natural idea since normalized Lempel–Ziv complexity measures the generation rate of new patterns along symbol sequence. It is closely related to such important information-theoretic properties like entropy, randomness, compression ratio, and redundancy [24]. It was shown that for ergodic stochastic processes normalized complexity estimates entropy rate [24]. However, in contrast to the entropy concept, complexity is a property of individual sequences easy in implementation and calculation.

All versions of the Lempel–Ziv complexities follow the same basic idea which is to parse the sequence of symbols $b_1^n := b_1 b_2 \ldots b_n$ of length n into distinct phrases. In the case of the version being applied in [24] the parsing algorithm is as follows. We start with the first symbol in the sequence and it states the first phrase (i.e. pattern). To obtain the second phrase we consider the consecutive sequences' symbols up to the moment k when the phrase $b_2, \ldots, b_k$ obtained does not occur in the earlier sequence $b_1, \ldots, b_{(k-1)}$. To get the third phrase we repeat such procedure starting from the symbol $b_{(k+1)}$. To find next and next phrases we proceed similarly up to the moment we reach b_n. In this way, the sequence $b_1 b_2 \ldots b_n$ is decomposed into distinct phrases and Lempel–Ziv complexity $C_{LZ}(b_1^n)$ is the number of these phrases. The generation rate of new patterns along $b_1 b_2 \ldots b_n$ is measured by normalized complexity

$$c(b_1^n) = \frac{C_{LZ}(b_1^n)}{\frac{n}{\log_2 n}}. \tag{1}$$

It was proven [24] that for ergodic sources

$$\limsup_{n\to\infty} c(b_1^n) = h, \tag{2}$$

with probability 1, where h is entropy rate of the source [13]. Sequences with a repetitive or poor pattern structure (e.g. periodic, quasi-periodic or regular sequences) have a very small normalized complexity, close to 0. On the opposite end stand the random sequences, which unfold rich pattern diversity. For sequences coming from fully random sources normalized complexity is 1 with very high probability.

To apply the complexity approach we propose the encoding method that takes into account velocity fluctuation. To do this at each point x_k, $k = 1, 2, \ldots, 26$, the velocity average value $v_{avr}(x_k)$ was calculated and treated as the threshold in the encoding process. For a given sequence of measurements at a point x_k of velocity $(v_i(x_k))_{i=1}^n$ we define the sequence of bits

$$(b_i(x_k))_{i=1}^n = \begin{cases} b_i(x_k) = 1 & \text{if } v_i(x_k) \geq v_{avr}(x_k) \\ b_i(x_k) = 0 & \text{if } v_i(x_k) < v_{avr}(x_k) \end{cases} \tag{3}$$

where $v_{avr}(x_k) = \frac{1}{n}\sum_{i=1}^{n} v_i(x_k)$.

Such sequences of bits were next analyzed by means of the normalized complexity. The accuracy of the Lempel–Ziv estimator as a function of the length of time series was tested in our previous paper [29]. It was shown that for the sequences of 400 bits long the entropy estimation error is low, it is below 4 percent. The length of sequences which are considered in this paper satisfied this condition.

3. Results

Typical velocity profiles measured by Particle Image Velocimetry (PIV) for combustor A were presented in Fig. 2. It also illustrates how complex the velocity fluctuations were in centerline of the combustor.

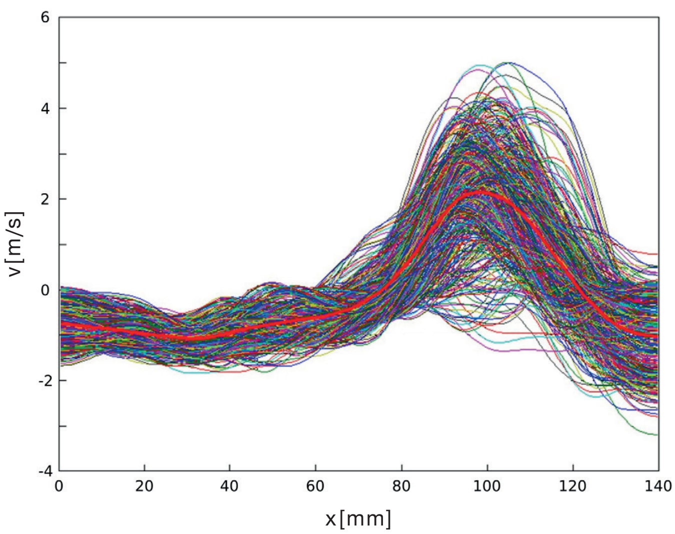

Fig. 2. Illustration of time variability of the flow in combustor A. 500 individual profiles of vertical component of the velocity, extracted along the centerline from two-dimensional instantaneous velocity fields measured by PIV with time interval 0.25 s are shown. Average velocity course is indicated by a red line. One can see that the flow, despite its stationarity, exhibits strong variation of velocity value

To present the results, first we briefly recall the concept of RMS value, which is a standard tool to measure turbulence [8, 25]. For a given sequence $\left(v_i(x_k)\right)_{i=1}^{n}$ of velocity measurements in some point x_k, $k = 1, 2, \ldots, 26$, RMS is defined as

$$RMS(x_k) := \left(\frac{1}{n}\sum_{i=1}^{n}\left(v_i(x_k) - v_{avr}(x_k)\right)^2\right)^{\frac{1}{2}}. \tag{4}$$

This indicator is being assumed a precursor of turbulence (it means that the increase of RMS indicates that the flow tends to a turbulent flow). To develop the complexity approach to characterize the type of flow, first we make natural assumption that the turbulent flow should be close to random. Therefore, since the complexity is a very good randomness indicator, we start our analysis by determining *a generic sampling frequency* of velocity for which the normalized complexity reaches the maximal value. Extensive numerical calculations have been done to find among all complexity curves (determined for various sampling frequency) the one with maximal complexity. To do this, for each combustor and for each point x_k, $k = 1, 2, \ldots, 26$, the range of sampling frequencies from the interval 0.1 Hz to 2 kHz has been considered with the step of 0.1 Hz.

The complexity analysis was applied to the velocity measurements obtained by hot wire method. After fixing generic sampling frequency the encoding procedure was used (see Section 2.2) to get sequences of bits and next to calculate the normalized complexities. Thus, in spite of the fact that sampling frequency in the experiments was 10 kHz, for the "complexity" purposes the generic frequency which we determined for the first experiment was 11.1 Hz, whereas in the second experiment it was 21.7 Hz.

The results of normalized complexity calculations, with the use of these generic frequencies are presented in Figs 3a and 3b. They are shown for 26 points placed uniformly along the central line in combustor for two configurations chambers (Figs 1a, 1b). For comparison the RMS analysis of velocity fluctuations both for data recorded with the use of thermoanemometry (Figs 3a, 3b) and with PIV method (Fig. 3a) are also given. One can observe qualitative similarity between the curves obtained with the use of complexity approach and with the use of RMS method (4).

The results presented in Fig. 3 also show that the complexity curves are well correlated with average velocity curves (at each point the average is taken over time). Moreover, one can also observe that for both combustor configurations the maximal normalized complexity is reached in the point which is located very close to the point with maximal average velocity. However, when in configuration B these points almost perfectly coincide, it turned out that for configuration A the most random point determined by complexity is at some distance from the maximal velocity point. Our hypothesis is that the points with maximal velocity and maximal normalized complexity better coincide for combustors with more mixing asymmetric property and this seems to be the B configuration.

On the other hand, another interesting observation is that for configuration A the maximal value of normalized complexity is close to 1 (Fig. 3a). This indicates the fully random character of the encoded sequence and consequently random character

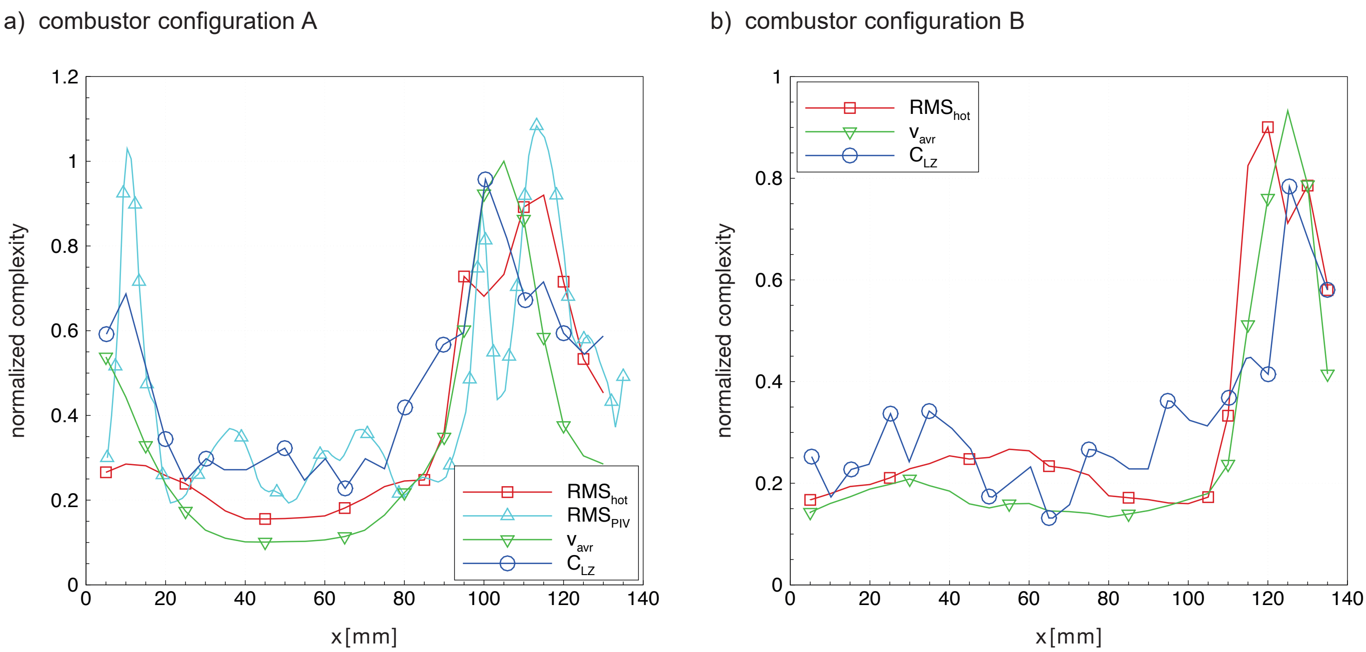

Fig. 3. Normalized complexity and RMS courses along selected line for the both configurations (a) and (b). Observe that the maximum of complexity (most random velocity fluctuations point) is very close to maximal average velocity while the most turbulent point determined with RMS is at some distance from this point

of the velocity fluctuations in this point. For configuration B the maximal normalized complexity value reached is about 0.8 what shows that for this configuration the flow is not able to get a random character, it is not fully developed.

4. Conclusions

A novel method based on the Information Theory was proposed and applied to the analysis of different flow regimes in gas turbine combustors. To do this we analyze velocity fluctuations by applying encoding procedure which addresses these fluctuations. An important feature of this approach is its strong mathematical background coming from the Information Theory [13, 24]. This way we can also study relations between turbulence regimes of the flow and entropy rates (estimated by the Lempel–Ziv complexity). It is known that entropy rate is a quantitative measure of randomness, thus, we can compare the turbulence levels with the levels of randomness.

For both combustor configurations our results show that, in general, entropy rates of velocity fluctuations are well correlated with the RMS values for $x \geq 20$ mm. On the other hand, in the case of configuration A we observed that close to the wall ($x < 20$ mm) this correlation weakens, which we conjecture is the result of some wall influence on the flow. Our results support the idea that the complexity methods can also be applied to determine and to classify flow zones.

On the other hand, we observed a subtle fact. It turned out that the most random velocity fluctuations are in the point with maximal velocity while the most "turbulent" point, as determined with the RMS method, is at a small distance from this point.

It seems that Information Theory method can be particularly useful to analyze turbulent and unsteady flow regimes and it can be treated as an alternative to other methods. Moreover, concepts like complexity can also be used to establish other flow characteristics like its non-regularity and mixing property. However, the important point that should be addressed in the future, in the context of application of Information Theory based methods, is to determine more effectively a generic sampling frequencies.

Acknowledgements. A part of the experimental work was performed within the EU 5th Framework Programme FLOXCOM. S.B. acknowledges financial support from the Polish National Science Centre under the grant 2014/14/E/ST8/00578.

References

[1] C. Argyropoulos and N. Markatos, "Recent advances on the numerical modelling of turbulent flows", *Applied Mathematical Modelling* 39, 693–732 (2015).

[2] M. Gad-El-Hak, "Coherent structures and flow control: genesis and prospect", *Bull. Pol. Ac.: Tech.* 67(3), 412–444 (2019).

[3] S. Gepner, J. Majewski, and J. Rokicki, "Parallel anisotropic mesh refinement with dynamic load balancing for transonic flow simulations", *Bull. Pol. Ac.: Tech.* 65(2), 195–207 (2017).

[4] X. Yu and J. Liu "The secondary instability in Goertler flow", *Physics of Fluids A: Fluid Dynamics* 3, 1845–1847 (1991).

[5] D. Mihailovic, G. Mimic, P. Gualtieri, I. Arsenic, and C. Gualtieri, "Randomness Representation of Turbulence in Canopy Flows Using Komogorov Complexity Measures", *Entropy* 18, e19100519 (2017).

[6] K. Gumowski, O. Olszewski, M. Pocwierz, and K. Zielonko-Jung, "Comparative analysis of numerical and experimental studies of the airflow around the sample of urban development", *Bull. Pol. Ac.: Tech.* 63(3), 729–737 (2015).

[7] E. Akbarian, B. Najafi, M. Jafari, S.F. Ardabili, S. Shamshirband, and K. Chau, "Experimental and computational fluid dynamics-based numerical simulation of using natural gas in a dual-fueled diesel engine", *Engineering Applications of Computational Fluid Mechanics* 12(1), 517–534 (2018).

[8] F. Nieuwstadt, B. Boersma, and J. Westerweel, Turbulence: Introduction to theory and applications of turbulent flows, Springer International Publishing: United States of America, 2016.

[9] M. Wojcik and M. Szukiewicz, "A simple method of determination of the degree of gas mixing by numerical Laplace inversion and Maple", *Bull. Pol. Ac.: Tech.* 67(2), 235–240 (2019).

[10] Q. Li and Z. Fu, "Permutation entropy and statistical complexity quantifier of nonstationarity effect in the vertical velocity records", *Physical Review E* 89, 012905 (2014).

[11] J. Zhan, B.C. Wang, L.H. Yu, Y.S. Li, and L. Tang, "Numerical investigation of flow patterns in different pump intake systems", *Journal of Hydrodynamics* 24, 873–882 (2012).

[12] O. Cadot, C. Touze, and A. Boudaoud, "Linear versus nonlinear response of a forced wave turbulence system", *Physical Review E* 82, 046211 (2010).

[13] T. Cover and J. Thomas, Elements of Information Theory; A Wiley-Interscience Publication, New York: United States of America, (1991).

[14] T. Bossomaier, L. Barnett, M. Harre, and J. Lizier, An introduction to transfer entropy; Springer-Verlag, New York: United States of America, (2016).

[15] Z. Wang, N. Jin, Z. Gao, Y. Zong, and T. Wang, "Nonlinear dynamical analysis of large diameter vertical upward oil-gas-water three-phase flow pattern characteristics", *Chemical Engineering Science* 65, 5226–5236 (2010).

[16] Q. Zhang, X. Liang, Z. Fang, and C. Xiao, "Complexity analysis of precipitation using the Lempel-Ziv algorithm and a multi-scaling approach: a case study in Jilin province, China", *Stochastic Environmental Research and Risk Assessment* 31, 1697–1707 (2017).

[17] M. Atlas, S. Hussain, and M. Sagheer, "Entropy generation and squeezing flow past a Riga plate with Cattaneo-Christov heat flux", *Bull. Pol. Ac.: Tech.* 66(3), 291–230 (2018).

[18] M. Gazzah and H. Belmabrouk, "Directed co-flow effects on local entropy generation in turbulent heated round jets", *Computers and Fluids* 105, 285–293 (2014).

[19] M. Materassi, G. Consolini, N. Smith, and R.D. Marco, "Information Theory Analysis of Cascading Process in a Synthetic Model of Fluid Turbulence", *Entropy* 16, 1272–1286 (2014).

[20] A. Elkaroui, M. Gazzah, N. Said, P. Bournot, and G.L. Palec "Entropy generation concept for a turbulent plane jet with variable density", *Computers and Fluids* 168, 328–341 (2018).

[21] A. Pregowska, J. Szczepanski, and E.Wajnryb, "Temporal code versus rate code for binary Information Sources", *Neurocomputing* 216, 756–762 (2016).

[22] A. Pregowska, A. Casti, E. Kaplan, E. Wajnryb, and J. Szczepanski, "Information processing in the LGN: a comparison of neural codes and cell types", *Biological Cybernetics* 113, 453–464, (2019).

[23] A. Pregowska, E. Kaplan, and J. Szczepanski, "How Far can Neural Correlations Reduce Uncertainty? Comparison of Information Transmission Rates for Markov and Bernoulli Processes", *International Journal of Neural Systems* 29(8) (2019). https://www.worldscientific.com/doi/10.1142/S01290657195000355.

[24] A. Lempel and J. Ziv, "On the complexity of an individual sequence", *IEEE Transactions on Information Theory* IT–22, 75–88 (1976).

[25] U. Frisch, Turbulence, Cambridge University Press: Cambridge, United Kingdom, (1995).

[26] T. Kowalewski, European Community Project of Fifth Framework Programme, FLOXCOM ENK5-CT-2000–00114.

[27] T. Kowalewski, F. Lusseyran, and S. Blonski, "Flow structure identification with PIV and high speed imaging in cold flow combustor model", *The 5th Euromech Fluid Mechanics Conference, Toulouse, France, August* 24–28, (2003).

[28] J. Szumbarski and S. Blonski, "Destabilization of laminar flow in a rectangular channel by transversely-oriented wall corrugation", *Archives of Mechanics* 4, 393–428 (2011).

[29] A. Pregowska, K. Proniewska, P. van Dam, and J. Szczepanski, "Using Lempel-Ziv complexity as effective classification tool of the sleep-related breathing disorders", *Computer Methods and Programs in Biomedicine* 182 (2019). https://doi.org/10.1016/j.cmpb.2019.105052.

[30] A. Pregowska, J. Szczepanski, and E.Wajnryb, "Mutual information against correlations in binary communication channels", *BMC Neuroscience* 16, 1–7 (2015).