TECHNICAL UNIVERSITY OF CLUJ-NAPOCA

ACTA TECHNICA NAPOCENSIS

Series: Applied Mathematics, Mechanics, and Engineering
Vol. 62, Issue IV, November, 2019

ENCODINGS FOR THE CALCULATION OF THE PERMUTATION HYPOENTROPY AND THEIR APPLICATIONS ON FULL-SCALE COMPARTMENT FIRE DATA

Flavia-Corina MITROI-SYMEONIDIS, Ion ANGHEL, Shigeru FURUICHI

Abstract: *Based on the data collected during a full-scale experiment, the order/disorder characteristics of a compartment fire are researched. We discuss methods, algorithms and the novelty of our entropic approach. From our analysis, we claim that the permutation type hypoentropies can be successfully used to detect unusual data and to perform relevant analysis of fire experiments.*
Key words: *full-scale fire experiment; compartment fire; permutation hypoentropy; time series analysis; statistical complexity; two length permutation hypoentropy*

1. INTRODUCTION

The purpose of this paper is to present an analysis of the evolution of the temperature during a full-scale fire experiment. We compare two known encoding algorithms and propose new statistical complexities, pointing out abnormal values and structure of the experimental time series. We apply our methods on a single experimental data set, as an illustration. For other recent studies using the entropic analysis of the fire phenomena see [16] and [20].

The experiment has been carried out using a container (single-room compartment) as shown in Fig.1. The container had the following dimensions: 12 m × 2.2 m × 2.6 m. A single ventilation opening was available, namely the front door of the container which remained open during the experiment. Parts of the walls and the ceiling of the container were furnished with oriented strand boards (OSB). The fire source has been a wooden crib, made of 36 pieces of wood strips 2.5 cm × 2.5 cm× 30 cm, on which has been poured 500 ml ethanol shortly before ignition. The fire bed was situated at 1.2 m below the ceiling. The measurement devices consisted in six built-in K-type thermocouples, fixed at fkey locations as shown in Fig.1, connected to a data acquisition logger. Flames were observed to impinge on the ceiling and exit through the front door opening, and we also noted the ignition of crumpled newspaper, stages of fire development which are known as indicators of flashover.

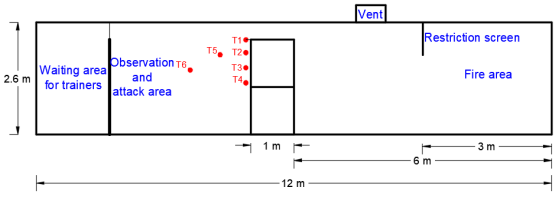

Fig. 1 The right-side view scheme of arrangement (instrumentation) of the flashover container

A thorough description of the experimental setup (materials and methods) and data analysis can be found in [15]. In Section 2 we present the theoretical background and algorithms used to analyze the experimental data collected during a full-scale fire experiment conducted at Fire Officers Faculty in Bucharest. Section 3 is dedicated to the analysis of the collected raw data.

2. THEORETICAL BACKGROUND AND REMARKS

2.1 Entropy and statistical complexity

The natural logarithm is used below, as elsewhere in this paper.

Shannon's entropy (Shannon, 1948) is defined as $H(P) = -\sum_{i=1}^{n} p_i \log p_i$, where $P = (p_1, \dots, p_n)$ is a finite probability distribution. It is nonnegative and its maximum value is $H(U) = \log n$, where

$$U = \left(\frac{1}{n}, \dots, \frac{1}{n}\right).$$

Throughout the paper we use the convention $0 \cdot \log 0 = 0$.

The Kullback-Leibler divergence [9] is defined by

$$D(P\|R) = \sum_{i=1}^{n} p_i(\log p_i - \log r_i),$$

where $P = (p_1, \dots, p_n)$ and $R = (r_1, \dots, r_n)$ are probability distributions. It is nonnegative and it vanishes for $P = R$.

If the value 0 appears in probability distributions $P = (p_1, \dots, p_n)$ and $R = (r_1, \dots, r_n)$, it must appear in the same positions for the sake of significance. Otherwise one usually considers the conventions $0 \log \frac{0}{b} = 0$ for $b \geq 0$ and $a \log \frac{a}{0} = \infty$ for $a > 0$. We remark that these are strong limitations and such conditions rarely occur in practice. In this paper we analyze some experimental fire data and discuss how to overcome this issue.

The Jensen-Shannon divergence (see [11] and [17] is

$$JS(P\|R) = \frac{1}{2} D\left(P \left\| \frac{P+R}{2}\right.\right) + \frac{1}{2} D\left(R \left\| \frac{P+R}{2}\right.\right) = H\left(\frac{P+R}{2}\right) - \frac{H(P)+H(R)}{2}.$$

The Jeffreys divergence [8] is defined by $D^J(P\|R) = D(P\|R) + D(R\|P)$. It holds $JS(P\|R) \leq \frac{1}{4} D^J(P\|R)$ (see (Lin, 1991) and (Crooks, 2008)).

The disequilibrium-based statistical complexity (LMC statistical complexity) introduced in [12] is defined as $C(P) = D(P)\frac{H(P)}{\log n}$, where $D(P)$, interpreted as disequilibrium, is the quadratic distance $D(P) = \sum_{i=1}^{n} \left(p_i - \frac{1}{n}\right)^2$.

The Jensen-Shannon statistical complexity [10], [22] is defined by $C^{(JS)}(P) = Q_{(JS)}(P)\frac{H(P)}{\log n}$, where the disequilibrium $Q_{(JS)}(P)$ is $Q_{(JS)}(P) = k \cdot JS(P\|U)$. Here $k = (\max_P JS(P\|U))^{-1}$ is the normalizing constant and $U = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)$. For the computation of the normalizing constant, the maximum is attained for P such that there exists $i, p_i = 1$.

By replacing the disequilibrium by $D^J(P\|U)$, called further *Jeffreys disequilibrium*, and define the *Jeffreys statistical complexity* formula as $C^J(P) \equiv D^J(P\|U)\frac{H(P)}{\log n}$, one encounters the following limitation: since U has no zero components, the probability distribution P must have only strictly positive components, a fact which depends on the algorithms used to determine the underlying probability distribution. This replacement was briefly mentioned as an alternative to the Jensen-Shannon complexity measure in [18]; here we follow this idea. For fire experiments, the collected data yields, by various algorithms, some underlying probability distributions which have zero components, hence there is a need for an alternative to the Kullback-Leibler divergence. Therefore, in an attempt to overcome the restrictions of the Kullback-Leibler divergence, our approach is based on the *Ferreri's hypodivergence at the level* λ [5], which solves these issues, defined as

$$K_\lambda(P\|R) = \frac{1}{\lambda}\sum_{i=1}^{n}(1+\lambda p_i)\log\frac{1+\lambda p_i}{1+\lambda r_i} \text{ with } \lambda > 0.$$

Remark 1 As mentioned by Ferreri, $K_\lambda(P\|R)$ is always non-negative. (It only vanishes for $P = R$.) The proof of the positivity can be done via the so-called *Log sum inequality,* from which it follows immediately. Another way is to use the same strategy as for the positivity of the Kullback-Leibler divergence: first we note that $\log a \leq a - 1$ for all $a > 0$, then we infer that

$$\lambda K_\lambda(P\|R) = -\sum_{i=1}^{n}\left(1+\lambda p_i\right)\log\frac{1+\lambda r_i}{1+\lambda p_i} \geq -\sum_{i=1}^{n}\left(1+\lambda p_i\right)\left(\frac{1+\lambda r_i}{1+\lambda p_i} - 1\right) = 0.$$

We use the *Jeffreys-Ferreri hypodivergence at the level* λ*, defined* by $D_\lambda^{JF}(P\|R) \equiv K_\lambda(P\|R) + K_\lambda(R\|P)$ (see formula (5.3) in [5]), to define the *Jeffreys-Ferreri disequilibrium at the level* λ by $D_\lambda^{JF}(P\|U)$ and the *Jeffreys-Ferreri statistical complexity at the level* λ *as* $C_\lambda^{JF}(P) \equiv D_\lambda^{JF}(P\|U)\frac{F_\lambda(P)}{F_\lambda(U)}$, where Ferreri's hypoentropy at the level λ is

$$F_\lambda(P) = \frac{1}{\lambda}(1+\lambda)\log(1+\lambda) - \frac{1}{\lambda}\sum_{i=1}^{n}(1+\lambda p_i)\log(1+\lambda p_i) \text{ with } \lambda > 0.$$

Ferreri's hypoentropy is positive and its maximum value is $F_\lambda(U)$, where $U = \left(\frac{1}{n}, \dots, \frac{1}{n}\right)$. See [6].

It is known that $\lim_{\lambda\to\infty} F_\lambda(P) = H(P)$, and $\lim_{\lambda\to\infty} K_\lambda(P\|R) = D(P\|R)$, and we emphasize that the second limit is finite when the value 0 does not appear in the probability distribution $R = (r_1, \dots, r_n)$ or it appears in the same positions as in $P = (p_1, \dots, p_n)$, otherwise $D(P\|R) = \infty$.

Moreover, it holds $\lim_{\lambda\to\infty} C_\lambda^{JF}(P) = C^J(P)$.

See in Fig. 2 a graphical comparison of $D^J(P\|1-P)$ and $D_\lambda^{JF}(P\|1-P)$, for $\lambda = 50{,}100, \dots, 250$ and $P = (t, 1-t)$, for $t = 0.01{,}0.03, \dots, 0.99$. It is important to mention that D_λ^{JF} approaches D^J when λ approaches infinity. We also mention that D_λ^{JF} is well defined in the entire range $0 \le t \le 1$, while D^J requires $0 < t < 1$.

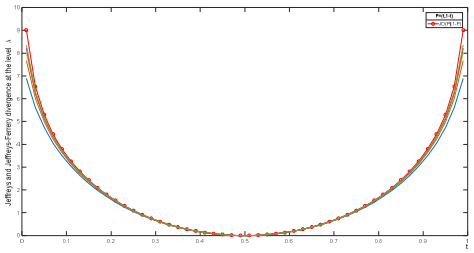

Fig. 2 Plots of the Jeffreys divergence (here denoted JD) and the Jeffreys-Ferreri divergences (with thinner line style) for $\lambda = 50{,}100, \dots, 250$

Seeing that $D_\lambda^{JF}(P\|R) < D^J(P\|R)$ holds for $R = 1-P, n = 2$ (Fig. 2), we check if this always holds. It is of mathematical interest and it clarifies our experimental results. The positive answer may be known, but since we have had no reference for it, we provide its proof as follows.

Proposition 1 The Ferreri divergence $K_\lambda(P\|R)$ is monotonically increasing as a function of λ.

Together with the known monotonicity of the Ferreri entropy, it yields $D_\lambda^{JF}(P\|R) < D^J(P\|R)$ for all $\lambda > 0$, positive integers n and probability distributions $P = (p_1, \dots, p_n)$ and $R = (r_1, \dots, r_n)$.

Proof. We establish that its derivative with respect to λ is non-negative. It suffices to see that

$$\frac{d}{d\lambda}K_\lambda(P\|R) = \frac{d}{d\lambda}\left[\sum_{i=1}^{n}\left(\frac{1}{\lambda}+p_i\right)\log\frac{1+\lambda p_i}{1+\lambda r_i}\right]$$

$$= -\sum_{i=1}^{n}\frac{1}{\lambda^2}\log\frac{1+\lambda p_i}{1+\lambda r_i} + \sum_{i=1}^{n}\left(\frac{1}{\lambda}+p_i\right)\frac{1+\lambda r_i}{1+\lambda p_i} \cdot \frac{p_i(1+\lambda r_i) - r_i(1+\lambda p_i)}{(1+\lambda r_i)^2}$$

$$= \sum_{i=1}^{n}\left[-\frac{1}{\lambda^2}\log\frac{1+\lambda p_i}{1+\lambda r_i} + \left(\frac{1}{\lambda}+p_i\right)\frac{1}{1+\lambda p_i}\cdot\frac{p_i-r_i}{1+\lambda r_i}\right]$$

$$= \sum_{i=1}^{n}\left(-\frac{1}{\lambda^2}\log\frac{1+\lambda p_i}{1+\lambda r_i} + \frac{1}{\lambda}\cdot\frac{p_i-r_i}{1+\lambda r_i}\right)$$

$$= \frac{1}{\lambda^2}\sum_{i=1}^{n}\left(\frac{\lambda(p_i-r_i)}{1+\lambda r_i} - \log\frac{1+\lambda p_i}{1+\lambda r_i}\right)$$

$$= \frac{1}{\lambda^2}\sum_{i=1}^{n}\left(\frac{1+\lambda p_i}{1+\lambda r_i} - 1 - \log\frac{1+\lambda p_i}{1+\lambda r_i}\right) \ge 0.$$

It vanishes only for the case $P = R$.

Remark 2 The Jeffreys-Ferreri statistical complexity is defined above without normalizing the disequilibrium, however the interested reader can also normalize it. The maximum of $D_\lambda^{JF}(P\|U)$ is attained for P such that there exists $i, p_i = 1$. The proof of this fact follows using the same recipe as in [14]. Straightforward computation yields

$$\max_P D_\lambda^{JF}(P\|U) = \left(1-\frac{1}{n}\right)\log(1+\lambda)$$

which shows that the disequilibrium $D_\lambda^{JF}(P\|U)$ is bounded from above by by $\log(1+\lambda)$. Its boundedness makes the normalization less interesting in our context. The disequilibrium of the Jeffreys statistical complexity cannot be normalized, and it is not defined for probability distributions with null components.

As in the case of the known complexities $C(P)$, $C^{(JS)}(P)$, the new one, $C_\lambda^{JF}(P)$, would be zero (minimum complexity) for $P = U$ or if there exists i such that $p_i = 1$. These two cases describe very different states of the system, considered simple, states with maximum, respectively minimum entropy.

The mutual information is defined by

$$I(X,Y) = D\big(P_{(X,Y)} \| P_X \otimes P_Y\big) = H(P_X) + H(P_Y) - H\big(P_{(X,Y)}\big).$$

Here X and Y are discrete random variables with probability distributions P_X and P_Y, the joint probability is $P_{(X,Y)}$ and the direct product is $P_X \otimes P_Y$. We introduce the *PYR statistical complexity* by

$$C^{(PYR)}(X,Y) \equiv I(X,Y)\frac{H(P_X)}{\log n}\cdot\frac{H(P_Y)}{\log m},$$

where n is the number of components of P_X, m is the number of components of P_Y. It is known that $0 \le I(X,Y) \le \min\{H(P_X), H(P_Y)\}$. To overcome the restrictions of the Shannon entropy and Kullback-Leibler divergence, our approach is based again on the hypoentropies and hypodivergences at the level λ. Therefore, we consider the PYR statistical complexity at the level λ as $C^{PYR}_{\lambda,i}(X,Y) \equiv I_{\lambda,i}(X,Y)\frac{F_\lambda(P_X)}{F_\lambda\left(\frac{1}{n},\dots,\frac{1}{n}\right)}\cdot \frac{F_\lambda(P_Y)}{F_\lambda\left(\frac{1}{m},\dots,\frac{1}{m}\right)}$. Here

$$I_{\lambda,1}(X,Y) \equiv K_\lambda\big(P_{(X,Y)} \| P_X \otimes P_Y\big) \text{ or}$$
$$I_{\lambda,2}(X,Y) \equiv F_\lambda(P_X) + F_\lambda(P_Y) - F_\lambda\big(P_{(X,Y)}\big).$$

The two definitions do not coincide, but it holds $\lim_{\lambda\to\infty} I_{\lambda,i}(X,Y) = I(X,Y)$, increasingly for for $i = 1$. Therefore $\lim_{\lambda\to\infty} C^{PYR}_{\lambda,i}(X,Y) = C^{(PYR)}(X,Y)$. The increasing monotonicity of the information $I_{\lambda,1}(X,Y)$ can be immediately inferred from Proposition 1 and we call $I_{\lambda,1}(X,Y)$ hypoinformation. It takes the value zero for independent random variables.

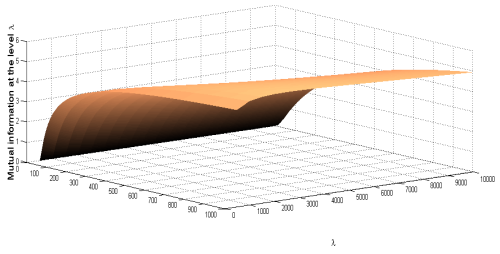

Fig. 3 $I_{\lambda,2}(X,Y)$, $n = 2,3,\dots,1000$

Remark 3 Applied on our experimental data, $I_{\lambda,2}(X,Y)$ decreases as a function of λ. See Fig and Fig. This fact does not hold true in general, that is $I_{\lambda,2}(X,Y)$ is not necessarily decreasing as a function of λ, for fixed random variables X, Y. A visual counterexample is the plot of $I_{\lambda,2}(X,Y)$ for the independent variables $X = Y = \left(\frac{1}{n},\dots,\frac{1}{n}\right)$ (Fig. 3).

2.2 Extraction of the underlying probability distribution

The permutation entropy PE [1] quantifies the randomness and the complexity of a time series based on the appearance of ordinal patterns, that is on comparisons of neighboring values of time series. Following the encoding steps in the PE-algorithm [1], which we include further for the sake of convenience, and using the hypoentropy instead of the Shannon entropy, we formulate the *PHE-algorithm* (*Permutation HypoEntropy algorithm*) as follows. For more details on the PE-algorithm see [15].

Let $T = (t_1,\dots,t_n)$ be a time series with distinct values.

Step 1. The increasing rearranging of the components of each j-tuple $(t_i,\dots,t_{i+j-1})$ as $\left(t_{i+r_1-1},\dots,t_{i+r_j-1}\right)$ yields a unique permutation of order j denoted by $\pi = (r_1,\dots,r_j)$, an encoding pattern that describes the up-and-downs in the considered j-tuple.

Example 1 For the 5-tuple (2.1, 1.3, 3.3, 1.1, 4.2) the corresponding permutation (encoding) is (4, 2, 1, 3, 5).

Step 2. The absolute frequency of this permutation (the number of j-tuples which are associated to this permutation) is

$$k_\pi \equiv \#\{i : i \le n-(j-1), (t_i,\dots,t_{i+j-1}) \text{ is of type } \pi\}.$$

These values have the sum equal to the number of all consecutive j-tuples, that is $n-(j-1)$.

Step 3. The *permutation hypoentropy of order j* is defined as

$$PHE(j) \equiv \frac{1}{\lambda}(1+\lambda)\log(1+\lambda) - \frac{1}{\lambda}\sum_\pi (1+\lambda p_\pi)\log(1+\lambda p_\pi),$$

where $p_\pi = \frac{k_\pi}{n-(j-1)}$ is the relative frequency of the permutation π, and $\lambda > 0$.

In [1] the measured values of the time series are considered distinct. The authors neglect equalities and propose to break them by adding small random perturbations (random noise) to the original series.

Another known approach is to rank the equalities according to their *order of emergence* (to rank the equalities with their sequential/chronological order). See for instance [2] and [4]. We use this method throughout the paper to compute $PHE(j)$ for j =3, 4, 5.

Applying the PHE-algorithm for experimental fire data, $C_{\lambda}^{JF}(P)$ cannot be zero. The number of the encoding patterns which occur is > 1 and these patterns are not equiprobable: some patterns may be rare or locally forbidden (that is, one encounters such patterns at some thermocouples, but not in all time series), as discussed in [15].

Following the encoding steps in the TLPE-algorithm (*Two-Length Permutation Entropy algorithm* [21] and using again the hypoentropy instead of the Shannon entropy, we formulate the TLPHE-algorithm (*Two Length Permutation HypoEntropy algorithm*) as follows. For other details on the TLPE-algorithm see [15].

Step 1. Given the j-tuple $T = (t_1, \dots, t_j)$, we start encoding the last $k \leq j$ elements $(t_{j-k+1}, \dots, t_j)$ according to the ordinal position of each element, that is, every t_s is replaced by a symbol which indicates the position occupied by t_s within the increasing rearranging of the considered k-tuple.

Next, we proceed by encoding each previous element t_m up to $m = 1$ according to the symbol provided by Step 1 applied to the k-tuple $(t_m, \dots, t_{m+k-1})$.

Example 2 Encoding obtained by the chronological ordering of equal values: $(3.1,\ 3.1,\ 3.1,\ 4,\ 3.1) \rightarrow (1,\ 1,\ 1,\ 3,\ 2)$ for k =3 and j=5.

Step 2 and **Step 3** coincide with Step 2 and Step 3 in the PHE-algorithm above.

This algorithm leads, after computing the relative frequencies of the encoding sequences, to *the two-length permutation hypoentropy* (TLPHE(k, j)).

Given the pair (k, j) of values, the number of symbolic (encoding) sequences of length j is $k!\, k^{j-k}$, a number which can be much smaller than $j!$, so this algorithm is faster, it involves a simplified computation and sometimes it makes the results more relevant for big values of j.

We deal with the equal values by using the same method as for PHE, that is we consider them ordered chronologically.

At our best knowledge, in the scientific literature there are no algorithms adjusted to determine probability distributions or statistical complexities conceived for fire data analysis. For the sake of completeness, we include here the description of the encoding step of the PYR-algorithm [13] that we recently developed for this purpose. By applying it repeatedly on our experimental data, we use its output to compute the PYR statistical complexity described in the previous section, a complexity that we claim that best fits the features of the fire data set environment.

Figure 4 shows the idealized fire curve that describes the evolution of the temperature values during a fire experiment in a compartment.

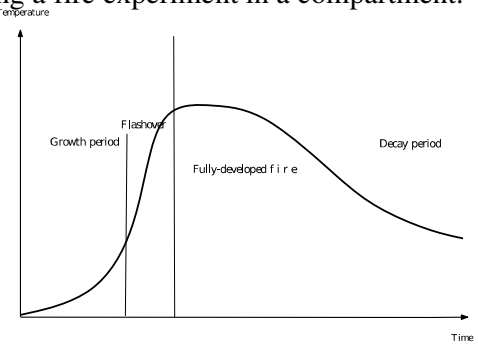

Fig. 4 Idealized time-temperature fire curve [7]

Encoding Step Let $mx(d) = \min\{s: t_s = \max(t_1, \dots, t_n)\}$ determined for the data collected at the thermocouple Td. The j-tuples with distinct elements are counted on behalf of a permutation as in the PHE-algorithm encoding step. The same for each j-tuple $(t_i, \dots, t_{i+j-1})$ that contains ties, after ordering the ties chronologically if $i \leq mx(d)$, or in reversed chronological order if $i > mx(d)$. This step is an adjustment of the counting procedure inspired by the evolution of the fire: a j-tuple is considered on the ascending trend before the maximum value of the temperature is reached (the growth period), respectively on the descending trend afterwards (the decay period).

Example 3 A 7-tuple which satisfies $t_{i+1} = t_{i+4} < t_{i+7} < t_{i+2} < t_{i+3} = t_{i+5} = t_{i+6}$, should be counted on behalf of the permutation $(1, 4, 7, 2, 3, 5, 6)$ if it occurs on a growth period,

or on behalf of the permutation $(4,1,7,2,6,5,3)$ if it appears during the decay period.

The resulting hypoentropy, computed following Step 2 and Step 3 in the PHE-algorithm above, is denoted by $PYRPHE(j)$.

In the next section we apply the above techniques and observe their capability to discern the changes of the Jeffreys-Ferreri statistical complexity of the experimental data.

3. RAW DATA ANALYSIS

The raw data set under consideration consists of measured temperatures during a compartment fire: six thermocouples T1, ..., T6 measure the temperatures every second during the experiment. Hence, we get six time-series consisting of 3046 entries (data points) and we aim to a better understanding of these results by modeling the time series using information theory, and to assess the performance of the discussed statistical complexities.

See in Fig.5 the plots of the discussed hypoentropies at the level $\lambda = 1000$, obtained using the experimental data.

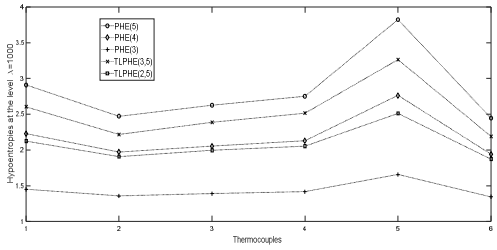

Fig. 5 Hypoentropies (PHE and TLPHE) at level λ=1000, for different embedding dimensions

In what follows we analyze the dynamical behavior of the temperature in the compartment fire from the viewpoint of the statistical complexity. We aim to observe the *features* captured by the statistical complexity. The novelty of our approach consists in investigating the *Jeffreys* and the *Jeffreys-Ferreri disequilibrium-based statistical complexities* ($C^J(P)$ and C_λ^{JF}) using the technique described above (that is plugging, whenever significant/finite, the entropies PE and TLPE in the $C^J(P)$ formula, respectively plugging the entropies PHE and TLPHE in the $C_\lambda^{JF}(P)$ formula). See the PE- algorithm [1] and the TLPE-algorithm [21].

In Fig.6 and Fig.7 one can see the plotted complexities C_λ^{JF} at $\lambda = 1000$ and $\lambda = 10^{12}$. By plotting the statistical complexity for each permutation type hypoentropy, we note that the convergence to C^J, as $\lambda \to \infty$, is faster for the algorithms with fewer encoding patterns (see Fig.9- Fig.13).

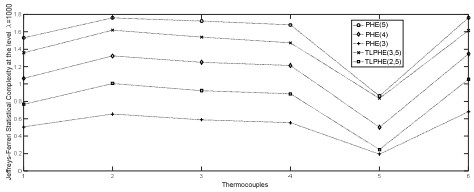

Fig. 6 The Jeffreys-Ferreri Statistical Complexity at level λ=1000

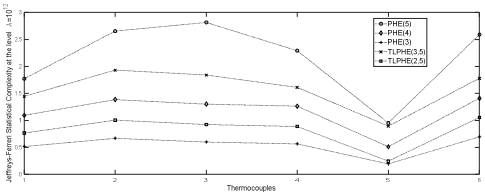

Fig. 7 The Jeffreys-Ferreri Statistical Complexity at level λ=10^12

A quick comparison of Fig.7 with the plot of the Jeffreys complexities in Fig.8 shows how the parametric (Jeffreys-Ferreri) complexity approach provides more useful insights for the fire data analysis, while some details are lost due to the infinite values of the Jeffreys complexity, for a bigger number of encodings (the plots corresponding to PE(5) and TLPE(3,5) in Fig.8; see also [15]).

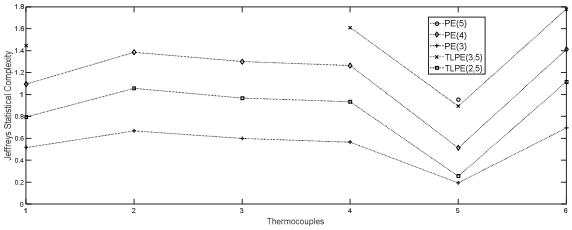

Fig. 8 The Jeffreys Statistical Complexity

It is known [5] that $F_\lambda(P)$ is a monotonically increasing function of λ. Using the experimental data, we conjecture that C_λ^{JF} is also monotonically increasing as a function of λ. See below (see Fig.9 - Fig.13) the plot of C_λ^{JF} at various levels λ.

When the Jeffreys complexity $C^J(P)$ is infinite (see Fig.8), plotting the Jeffreys-Ferreri complexities might be of interest, their plot (for instance for PHE(5) in Fig.12) being similar to the finite cases of $C^J(P)$, due to the relatively

slow convergence as $\lambda \to \infty$ of C_λ^{JF} to $C^J(P)$, for a bigger number of encodings.

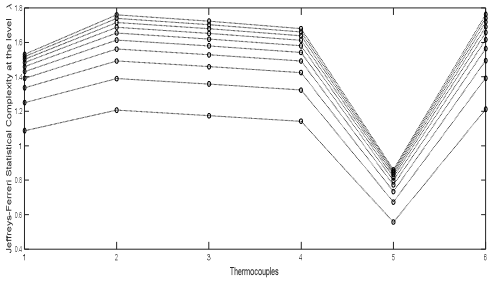

Fig. 9 Jeffreys-Ferreri Statistical Complexity for the PHE algorithm, $j = 5, \lambda = 100{,}200, \ldots, 1000$

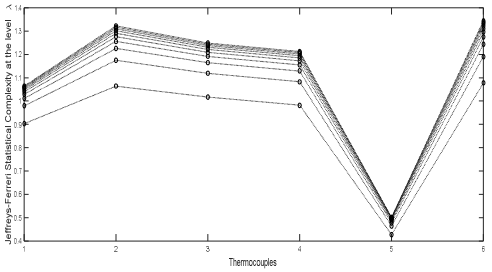

Fig. 10 Jeffreys-Ferreri Statistical Complexity for the PHE algorithm, $j = 4, \lambda = 100{,}200, \ldots, 1000$

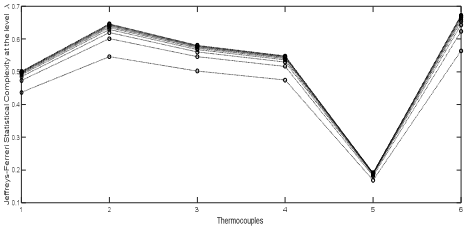

Fig. 11 Jeffreys-Ferreri Statistical Complexity for the PHE algorithm, $j = 4, \lambda = 40{,}90, \ldots, 490$

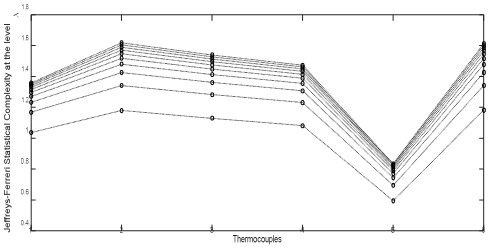

Fig. 12 Jeffreys-Ferreri Statistical Complexity for the TLPHE algorithm, $k = 3, j = 5$

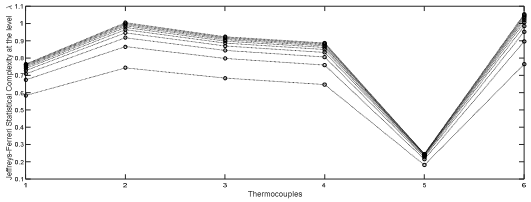

Fig. 13 Jeffreys-Ferreri Statistical Complexity for the TLPHE algorithm, $k = 2, j = 5$

In the hypoentropy - complexity plane we plot (see Fig. 14) the Jeffreys-Ferreri statistical complexity (for $\lambda \in \{100, 200, 300, \ldots, 2000\}$) against the hypoentropy. We use different markers to plot the results for each thermocouple (as indicated in the legend), for instance T1 is plotted with o, T2 with *… We use different colors to differentiate the plots corresponding to different embeding dimensions. The color code in Fig.14**Fig.** is: green for PHE(3), magenta TLPHE(2,5), blue PHE(4), black TLPHE(3,5) and red PHE(5). Once again we notice the unusual plotting of the time series at T5 (the squared plots situated below the rest of the "fireworks").

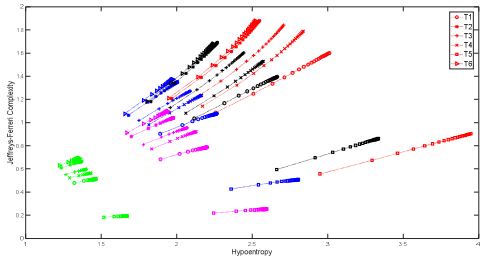

Fig. 14 - The hypoentropy - complexity plane

From the above analysis, we conclude that the PHE- and TLPHE-algorithms can be successfully used to detect unusual data collected in fire experiments. According to our findings, the Jeffreys-Ferreri statistical complexity is a valid tool for the analysis of the evolution of the temperature in a compartment fire data.

The new proposed Jeffreys and Jeffreys-Ferreri statistical complexities can complement and validate the information provided by the usual LMC and Jensen-Shannon statistical complexities. See [15]. See Fig.15**Fig** and Fig.16**Fig** for a quick comparison to the Jensen-Shannon and LMC statistical complexities [15]: we note that similar conclusions can be drawn, which confirm the Jeffreys-Ferreri statistical complexities as a valid tool to perform the analysis of experimental compartment fire data.

For the plots of the permutation hypoentropies $PYRPHE(j)$ and mutual informations $I_{\lambda,1}(X,Y)$, $I_{\lambda,2}(X,Y)$, see Fig. 17**Fig** – Fig.20 ($j = 3, 4,\ \lambda = 100, 200, \ldots, 1000$).

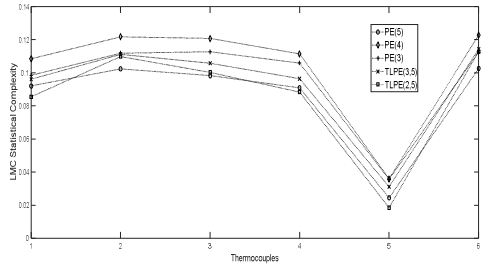

Fig. 15 LMC statistical complexity

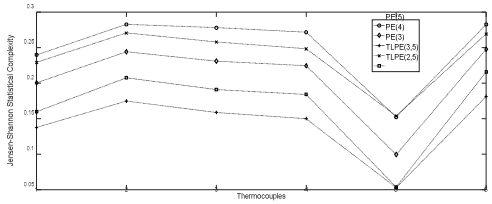

Fig. 16 Jensen-Shannon statistical complexity

We plotted the mutual information with squares, respectively circles). The plots of $I_{\lambda,1}(X,Y)$ and $I_{\lambda,2}(X,Y)$ increase (respectively decrease) with respect to λ (respectively decrease) to $I(X,Y)$.

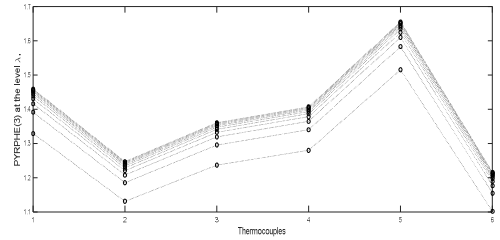

Fig. 17 $PYRPHE(3)$

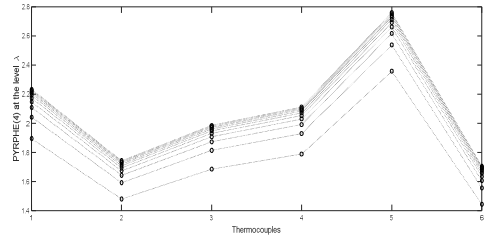

Fig. 18 $PYRPHE(4)$

The PYR algorithm has been used to compute the PYR statistical complexities (for the underlying marginal, joint and direct product probabilities). See Fig. 21 and Fig. 22. The plots of $C_{\lambda,1}^{PYR}(X,Y)$ and $C_{\lambda,2}^{PYR}(X,Y)$ increase (respectively decrease) with respect to λ, to $C^{(PYR)}(X,Y)$. We plotted the PYR complexities with squares, respectively circles.

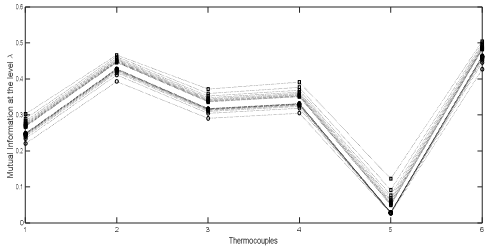

Fig. 19 Mutual information $I_{\lambda,1}(X,Y)$ and $I_{\lambda,2}(X,Y)$, $j=3$

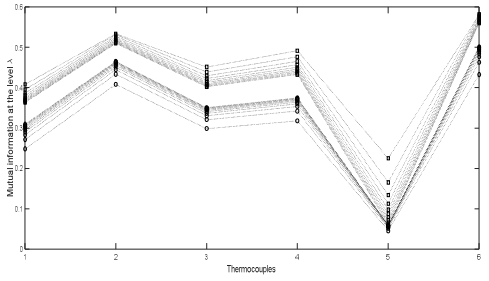

Fig. 20 Mutual information $I_{\lambda,1}(X,Y)$ and $I_{\lambda,2}(X,Y)$, $j=4$

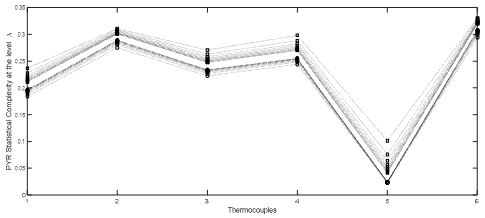

Fig. 21 PYR statistical complexity $C_{\lambda,1}^{PYR}(X,Y)$ and $C_{\lambda,2}^{PYR}(X,Y)$, j=3

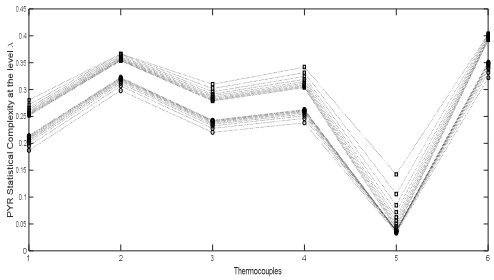

Fig. 22 PYR statistical complexity $C_{\lambda,1}^{PYR}(X,Y)$ and $C_{\lambda,2}^{PYR}(X,Y)$, j=4

The discrete random variables X and Y have been considered as follows: the range of X is the set of all permutations of order j, $\{\pi_1,\pi_2,\dots,\pi_{j!}\}$, the range of Y is $\{G,D\}$, where G (growth) and D (decay) are the only two possible positions of each j-tuple in each time series. A j-tuple $(t_i,\dots,t_{i+j-1})$ is on the growth period of the time series, at the thermocouple Td, if $i \le mx(d)$, otherwise it is on the decay

period. Therefore, the probabilities are (for $j = 3$ and 4)

$$P_X(\pi_i) = \frac{number\ of\ j-tuples\ of\ type\ \pi_i,\ counted\ with\ the\ PYR-algorithm}{3046-j+1},\ i = 1,2,\dots,j!,$$

$$P_Y(G) = \frac{number\ of\ j-tuples\ on\ the\ growth\ period}{3046-j+1} = \frac{mx(d)}{3046-j+1},\ P_Y(D) = 1 - P_X(G).$$

The joint probabilities are

$$P_{(X,Y)}(\pi_i, G) = \frac{number\ of\ j-tuples\ of\ type\ \pi_i,\ counted\ with\ the\ PYR-algorithm,\ on\ the\ growth\ period}{3046-j+1},$$

$$P_{(X,Y)}(\pi_i, D) = \frac{number\ of\ j-tuples\ of\ type\ \pi_i,\ counted\ with\ the\ PYR-algorithm,\ on\ the\ decay\ period}{3046-j+1},$$

for $i = 1,2,\dots,j!$.

4. CONCLUSIONS

We propose new formulas for the statistical complexity (based on the hypoentropy and hypodivergence, on Jeffreys divergence and mutual information). We apply these formulas using various algorithms to determine the probability distributions, including the PYR-algorithm, which we conclude that it is the most appropriate for the study of fire data. The newly proposed complexities are used to analyze a full-scale experimental data set collected from a compartment fire.

These results are improving and they agree with the current knowledge and understanding of the general patterns of the evolution of fire, contributing to the theoretical basis of fire research.

The level of the hypoentropy and hypodivergence not only increases the sensitivity of the statistical complexities, but it is an efficient way to avoid the constraints encountered for the Shannon entropy and Kullback-Leibler divergence.

Our results might also indicate a turbulence or a malfunction of the thermocouple T5, however, this is beyond the scope of the present paper to discuss it in detail.

5. ACKNOWLEDGEMENT

The authors (F.-C. M.-S.) and (I.A.) acknowledge the support of a grant of the Romanian Ministry of Research and Innovation, CCCDI - UEFISCDI, project number PN-III-P1-1.2-PCCDI-2017-0350 / 38PCCDI within PNCDI III. The author (S.F.) was partially supported by JSPS KAKENHI Grant Number 16K05257.

6. REFERENCES

[1] Bandt, C., Pompe, B., *Permutation entropy: a natural complexity measure for time series*. Physical review letters, 88(17), 174102 (2002).

[2] Cao, T., Tung, W. W., Gao, J. B., Protopopescu, V. A., Hively, L. M., *Detecting dynamical changes in time series using the permutation entropy*. Physical review E, 70(4), 046217 (2004).

[3] Crooks, G. E., *Inequalities between the Jenson-Shannon and Jeffreys divergences*. Tech. Note 004 (2008).

[4] Duan, S., Wang, F., Zhang, Y., *Research on the biophoton emission of wheat kernels based on permutation entropy*. Optik, 178, 723-730 (2019).

[5] Ferreri, C., *Hypoentropy and related heterogeneity, divergence and information measures.* Statistica, 2, 155-167(1980).

[6] Furuichi, S., Mitroi-Symeonidis, F. -C., Symeonidis, E., *On some properties of Tsallis hypoentropies and hypodivergences*. Entropy, 16(10), 5377-5399 (2014).

[7] Graham, T. L., Makhviladze, G. M., Roberts, J. P., *On the theory of flashover development.* Fire Safety Journal, 25(3), 229-259 (1995).

[8] Jeffreys, H., *An invariant form for the prior probability in estimation problems*. Proc. Roy. Soc. Lon. Ser. A 186, 453-461 (1946).

[9] Kullback, S., Leibler, L. A., *On information and sufficiency*. Ann. Math. Statistics, 22, 1, 79-86 (1951).

[10] Lamberti, P. W., Martin, M. T., Plastino, A., & Rosso, O. A., *Intensive entropic non-triviality*

measure. Physica A: Statistical Mechanics and its Applications, 334 (1-2), 119-131 (2004).

[11] Lin, J., *Divergence measures based on the Shannon entropy*. IEEE. IEEE - Transactions on Information Theory 37, 145-151 (1991).

[12] López-Ruiz, R., Mancini, H. L., Calbet, X. *A statistical measure of complexity*. Physics Letters A, 209 (5-6), 321-326 (1995).

[13] Mitroi-Symeonidis, F. -C., Anghel, I. *The PYR-algorithm for time series modeling of temperature values and its applications on full-scale compartment fire data.* Submitted 2019.

[14] Mitroi-Symeonidis, F. -C., Anghel, I., Minculete, N., *Parametric Jensen-Shannon statistical complexity and its applications on full-scale compartment fire data.* Submitted 2019.

[15] Mitroi-Symeonidis, F. -C., Anghel, I., Lalu, O., Popa, C., *The permutation entropy and its applications on fire tests data.* Submitted 2019. Available online at: https://arxiv.org/abs/1908.04274.

[16] Murayama, S., Kaku, K., Funatsu, M., Gotoda, H., *Characterization of dynamic behavior of combustion noise and detection of blowout in a laboratory-scale gas-turbine model combustor.* Proceedings of the Combustion Institute, 37(4), 5271-5278 (2019).

[17] Rao, C. R., Nayak, T. K. (1985). *Cross entropy, dissimilarity measures, and characterizations of quadratic entropy.* IEEE Trans. Inform. Theory, vol. IT-31. no. 5, 589-593.

[18] Ribeiro, H. V., Zunino, L., Lenzi, E. K., Santoro, P. A., Mendes, R. S., *Complexity-entropy causality plane as a complexity measure for two-dimensional patterns*. PloS one, 7(8), e40689 (2012).

[19] Shannon, C. E., *A mathematical theory of communication.* Bell System Technical Journal, 27(3), 379-423 (1948).

[20] Takagi, K., Gotoda, H., Tokuda, I. T., Miyano, T. *Dynamic behavior of temperature field in a buoyancy-driven turbulent fire.* Physics Letters A, 382(44), 3181-3186 (2018).

[21] Watt, S. J., Politi, A., *Permutation entropy revisited.* Chaos, Solitons & Fractals, 120, 95-99 (2019).

[22] Zunino, L., Soriano, M. C., Rosso, O. A., *Distinguishing chaotic and stochastic dynamics from time series by using a multiscale symbolic approach.* Physical Review E, 86(4), 046210 (2012).

Codificări pentru calculul hipoentropiei de permutare şi aplicaţii ale lor pe date obţinute din incendiu de compartiment la scară reală

Pe baza datelor colectate în timpul unui experiment la scară reală, sunt investigate caracteristicile de ordine/dezordine ale unui incendiu de compartiment. Discutăm metode, algoritmi şi noutatea acestei abordări bazată pe entropie. În urma analizei noastre, afirmăm ca hipoentropiile de tip permutare pot fi folosite cu success la detectarea unor date neobişnuite şi pentru a face investigaţii relevante ale experimentelor de foc.

Flavia-Corina MITROI-SYMEONIDIS, Research Assistant, Police Academy "Alexandru Ioan Cuza", Fire Officers Faculty, Str. Morarilor 3, Sector 2, Bucharest, Romania. E-mail address: fcmitroi@yahoo.com

Ion ANGHEL, Assistant Professor, Police Academy "Alexandru Ioan Cuza", Fire Officers Faculty, Str. Morarilor 3, Sector 2, Bucharest, Romania. E-mail address: ion.anghel@academiadepolitie.ro

Shigeru FURUICHI, Professor, Department of Information Science, College of Humanities and Sciences, Nihon University, 3-25-40, Sakurajyousui, Setagaya-ku, Tokyo, Japan. E-mail address: furuichi@chs.nihon-u.ac.jp