Cuesta Frau, D.; Miró Martínez, P.; Oltra Crespo, S.; Jordán Núñez, J.; Vargas-Rojo, B.; Vigil-Medina, L. (2018). Classification of glucose records from patients at diabetes risk using a combined permutation entropy algorithm. Computer Methods and Programs in Biomedicine. 165:197-204. https://doi.org/10.1016/j.cmpb.2018.08.018


# Classification of glucose records from patients at diabetes risk using a combined Permutation Entropy algorithm

D. Cuesta–Frau, P. Miró–Martínez, S. Oltra–Crespo,
J. Jordán–Núñez, B. Vargas, L. Vigil

PII: S0169-2607(18)30206-2
DOI: https://doi.org/10.1016/j.cmpb.2018.08.018
Reference: COMM 4779

To appear in: *Computer Methods and Programs in Biomedicine*

Received date: 12 February 2018
Revised date: 9 August 2018
Accepted date: 30 August 2018

Please cite this article as: D. Cuesta–Frau, P. Miró–Martínez, S. Oltra–Crespo, J. Jordán–Núñez, B. Vargas, L. Vigil, Classification of glucose records from patients at diabetes risk using a combined Permutation Entropy algorithm, *Computer Methods and Programs in Biomedicine* (2018), doi: https://doi.org/10.1016/j.cmpb.2018.08.018

**Highlights**

- First time Permutation Entropy is applied to glucose time series.
- Test of different customizations for Permutation Entropy in order to address equal values and amplitude variations.
- Prediction of evolution to diabetes based on a Permutation Entropy analysis of the glucose time series.

# Classification of glucose records from patients at diabetes risk using a combined Permutation Entropy algorithm

D. Cuesta–Frau, P. Miró–Martínez, S. Oltra–Crespo, J. Jordán–Núñez, B. Vargas, L. Vigil

*D. Cuesta–Frau, and S. Oltra–Crespo are with the Technological Institute of Informatics (ITI), Universitat Politècnica de València, Campus Alcoi (EPSA-UPV)*
*Plaza Ferrándiz y Carbonell, 2, 03801, Alcoi (Spain)*
*(dcuesta@disca.upv.es)*

*P. Miró–Martínez and Jordi Jordán–Núñez are with the Statistics Department at Universitat Politècnica de València,, Campus Alcoi*
*Plaza Ferrándiz y Carbonell, 2, 03801, Alcoi (Spain)*
*(pamimar@eio.upv.es)*

*B. Vargas and L. Vigil are with the Internal Medicine Service at the University Hospital of Móstoles*
*Río Júcar s/n, 28935 Móstoles, Madrid, Spain (Spain)*
*(mvarela.hmtl@salud.madrid.org)*

**Abstract**

*Background and objectives*: The adoption in clinical practice of electronic portable blood or interstitial glucose monitors has enabled the collection, storage, and sharing of massive amounts of glucose level readings. This availability of data opened the door to the application of a multitude of mathematical methods to extract clinical information not discernible with conventional visual inspection. The objective of this study is to assess the capability of Permutation Entropy(PE) to find differences between glucose records of healthy and potentially diabetic subjects.

*Methods*: PE is a mathematical method based on the relative frequency analysis of ordinal patterns in time series that has gained a lot of attention in the last years due to its simplicity, robustness, and performance. We study in this paper the applicability of this method to glucose records of subjects at risk of diabetes in order to assess the predictability value of this metric in this context.

*Results*: PE, along with some of its derivatives, was able to find significant differences between diabetic and non–diabetic patients from records acquired up to 3 years before the diagnosis. The quantitative results for PE were $3.5878 \pm 0.3916$ for the non–diabetic class, and $3.1564 \pm 0.4166$ for the diabetic class. With a classification accuracy higher than 70%, and by means of a Cox regression model, PE demonstrated that it is a very promising candidate as a risk stratification tool for continuous glucose monitoring.

*Conclusion*: PE can be considered as a prospective tool for the early diagnosis of the glucorregulatory system.

*Keywords:* Permutation Entropy, Continuous glucose monitoring, Signal classification, Diabetes

## 1. Introduction

Manual inspection of continuous blood glucose records [1] provides very useful information for patients' diagnosis or treatment decisions [2]. This manual assessment is usually carried out in terms of glucose level thresholding [3]. A more refined analysis of the blood glucose data involves the computation of a myriad of variability indices to better reflect the glycaemic control status [4].

More recently, other mathematical methods have been applied to these data [5, 6]. The general objective is also to characterise the glucose–insulin endocrine system for a more personalised and efficient treatment [7]. Among all these more advanced mathematical methods, those based on signal complexity, regularity, or predictability estimation are gaining momentum due to their ability of capturing the subtle differences among subjects. Approximate Entropy (ApEn) [8], Sample Entropy (SampEn) [9], Fuzzy Entropy (FuzzEn) [10], Dispersion Entropy [11], State–Space Correlation Entropy [12], Bubble Entropy [13], Lempel Ziv Complexity (LZC) [14], Detrended Fluctuation Analysis (DFA) [15], Distribution Entropy (DistEn) [16], and Permutation Entropy (PE) [17], are just a few of these methods that have been applied successfully in the context of biomedical records, including glucose time series in some cases [18, 19, 20, 21, 22, 23].

Specifically, Permutation Entropy (PE) [17] is a complexity measure that is receiving a lot of attention in the last years. It is conceptually simple, the algorithm is easy to implement and has a low computational cost, it is robust against observational and dynamical noise, does not require any model assumption, and window length and sampling frequency have very little influence on the results [17, 24]. It has been already used in a varied and diverse set of applications [25]. This is the measure chosen in this work.

Contrary to many other similar metrics, PE is based on temporal orders instead of amplitude differences. It requires the determination of three input

parameters [26]: length of the time series $N$, length of the subsequences under comparison $m$ (permutation order), and time delay $\tau$. For simplicity, we assume in this work $\tau = 1$, since other values are comparable to down–sampling [27]. $N$ was set at the acquisition stage, as described in Sec. 2.2, and $m$ was varied between 2 and 9, a little bit wider interval than that recommended in [17]. SampEn also needs three input parameters, $N$, $m$, and a threshold $r$, but it is much more sensitive to these input values than PE, since a suboptimal choice of these values can lead to incorrect results [28]. DFA is also quite unstable when the input parameters change [29].

This is a complete new approach to the analysis of blood glucose time series, where most of the methods have been based on DFA [30, 31] or Sample Entropy (Multiscale) [32, 33, 21]. Since these series often include consecutive equal values that may interfere with a correct PE computation, they were addressed as recommended in [17, 34]. The possible influence of subsequence amplitude differences was also studied and quantified [27]. As a result, a combined PE method was optimized to maximize the possible differences between diabetic and non–diabetic records. In addition, a comparative analysis of the impact of equal values and amplitude differences in PE segmentation performance was carried out.

## 2. Methods

### *2.1. Permutation Entropy*

The standard PE method was introduced in [17]. It is a simple complexity measure that can be applied to any time series, and it is also robust with respect to signal noise. This metric is based on sample order instead of sample amplitude.

PE can therefore be applied to glucose records. It assesses the temporal

structure of a sequence, inherits its causal information [35], and it is not affected by nonlinear monotonous transformations that could be introduced by the glucose monitors. It can be applied to deterministic or stochastic systems without any assumptions about the underlying process [36].

The mathematical definition is as follows. Given a discrete time series $y[j], j \in \mathbb{N}, y[j] \in \mathbb{R}$, of length $N$, $y[j] = \{y[0], y[1], \ldots, y[N-1]\}$, for each index $j$, a subsequence of length $m$ can be extracted from $y[j]$ as:

$$x_j[i] = \{y[j], y[j+1], \ldots, y[j+m-1]\} = \{x[0], x[1], \ldots, x[m-1]\}$$

The subsequence $x_j[i]$ can then be re–arranged in ascending order, resulting in $x_j^*[i] = \{x[(0)], x[(1)], \ldots, x[(m-1)]\}$, with $x[(0)] \leq x[(1)] \leq x[(2)], \ldots, x[(m-2)] \leq x[(m-1)]$. A list of the ordinal indices associated to the initial $x_j[i]$ is updated according to the changes performed in the subsequence during the sorting process. The resulting list of length $m$, $[(0), (1), \ldots, (m-1)]$, is then compared with all the $M = m!$ possible permutations of these values without repetition, $[\sigma(0), \sigma(1), \ldots, \sigma(m-1)]$ . When a coincidence is found at permutation $k$, a matches counter $c$ is increased, $c[k] = c[k] + 1, 0 \leq k < M$. Numerical examples of this process can be found in [37, 26, 25, 38]. Finally, a probability for each permutation is estimated as:

$$p[k] = \frac{c[k]}{N - m + 1} \tag{1}$$

PE can then be computed as:

$$\mathrm{PE}(y, m, N) = -\sum_{\forall k} p[k] \log p[k] \tag{2}$$

The computation of PE may involve an additional parameter, an embedding delay $\tau$. In such a case, the subsequences are extracted as:

$$x_j[i] = \{y[j], y[j+\tau], y[j+2\tau], \ldots, y[j+(m-1)\tau]\},$$

with $\tau \geq 1$. There are no general guidelines regarding how to select the input PE parameters $m$ and $\tau$ [39]. The higher is $m$, the more reliable is the value of PE [27, 26]. However, $m!$ should be smaller than $N$ to ensure a reasonable minimum for $c[k]$ in the computation of a stable $p[k]$ value [25]. A trade–off has to be found experimentally [27]. As a result, most PE studies, if not all, use an interval for $m$, such as the recommended $3, \ldots, 7$ interval [17, 40], or wider, between 2 and 15 for example, as in [41]. With regard to the embedding delay, many works recommend to assume $\tau = 1$ [17, 11, 27] (although additional information can be obtained with $\tau > 1$), since these values may lead to frequency aliasing. Other works propose to combine simultaneously different $\tau$ values [39].

The dissimilarity computation in PE does not take into account the amplitude differences between subsequences, only the order, as described above. Conceptually, this may lead to consider two subsequences equal despite having completely opposed amplitudes, and therefore impact negatively on the correct interpretation of the system dynamics under analysis. To address this problem, a number of approaches have been proposed in the scientific literature recently. For example, in [38], an additional parameter $q$ is introduced in the permutation type as an additional element, which quantifies the differences $d_j[i]$ between consecutive values in $x_j[i]$:

$$q = \left\lfloor \frac{\max(d_j[i])}{\mathrm{std}(d_j[i]) \times \alpha} \right\rfloor$$

where $d_j[i] = \{|y[j+1] - y[j]|, \ldots, |y[j+m-1] - y[j+m-2]|\}$. The closer the precision regulation factor $\alpha$ is to zero, the more permutation types can be generated. The authors use $\alpha = 1$ [38]. Other methods also employ parameters, such as in the Amplitude Aware Permutation Entropy (AAPE) method [27], but in this case only one parameter, $A \in [0, 1]$, is necessary to be defined. In AAPE, instead of increasing the corresponding histogram bin by 1, when an ordinal pattern match is found, a relative normalised probability is used. This

probability is based on the mean value:

$$\sum_{i=0}^{m-1}\left(\frac{1}{m}\left|x[i]\right|\right)$$

and differences between consecutive samples:

$$\sum_{i=1}^{m-1}\left(\frac{1}{m-1}\left|x[i]-x[i-1]\right|\right),$$

normalised by all the contributions. The $A$ parameter accounts for the relative weight of mean and differences. In signal classification applications, both terms are equally important, and therefore $A = 0.5$ or greater is recommended. To detect abrupt changes, $A << 0.5$ makes AAPE more sensitive [27]. Thus, $c[k]$ is updated with a new term:

$$c[k] = c[k] + \frac{A}{m}\left|x[0]\right| + \sum_{i=1}^{m-1}\left(\frac{A}{m}\left|x[i]\right| + \frac{1-A}{m-1}\left|x[i]-x[i-1]\right|\right) \tag{3}$$

with:

$$p[k] = \frac{c[k]}{\sum\limits_{\forall k} c[k]} \tag{4}$$

We term the addition of this factor to PE, Amplitude Included Permutation Entropy (AIPE), to avoid possible confusion with the complete AAPE method [27]. In the original paper, the authors coined the term AAPE, but they also introduced a modification in the definition of the permutation patterns that is not implemented here.

Another drawback of the standard PE algorithm is the ambiguity when there are equal values in the subsequence [26]. The standard PE method neglects equal values, and if present, proposes to add random perturbations to avoid them [17]. Nonetheless, glucose time series include many equal values due to the low resolution of the acquiring devices, and this is an issue that will have to be

properly addressed [35, 42, 43]. Specific methods to address this drawback have also been proposed. In [27], the number of all possible permutations of similar states are considered to be used as scaling factors of the contributions of motifs with equal states, increasing the algorithm complexity and computational cost. Another solution is described in [34]. In this case, equal values are mapped to the index of the first one. Therefore, the permutation pattern list must include both permutations without and with possible repetitions, increasing the memory requirements of the standard method.

The final method proposed in this paper to analyse the glucose records is based on the standard PE algorithm, including the amplitude correction of [27]. The ambiguity of equal values is not explicitly addressed in the algorithm in order to keep it simple and fast. As the results in Sec. 3 will confirm, amplitude seems to play a more major role than equal values in the classification performance of glucose time series. A detailed combined algorithm that implements this method is shown in Algorithm 1.

However, different configurations of the method proposed will be tested in order to characterise the possible influence of each drawback and the solutions adopted, including the equal values disambiguation of [34], and the addition of random perturbations [17].

Other metrics will also be tested. Specifically, SampEn [9] will be included for comparative purposes. SampEn is probably the most applied non–linear measure in the context of biomedical records, and it has also been used in glucose time series [44, 45]. In [30], the performance of other clinical metrics related to diabetes, for the same subjects, was assessed, and no one was found to achieve significant results.

**Algorithm 1** Calculation of AIPE$(y, m, N)$. Combined algorithm. All the necessary steps are included for completeness and to facilitate implementation in any high level programming language.

**Require:** Permutations list $h[0, \ldots, M-1][0, \ldots, m-1], N, m, M = m!, A$
  **for** $j = 0, \ldots, N-1$ **do**
    **for** $i = 0, \ldots, m-1$ **do**
      $x_j[i] \leftarrow y[j+i]$
      index$[i] \leftarrow i$
    **end for**
    bSorted$\leftarrow$ false
    **while** (bSorted=false) **do**
      bSorted$\leftarrow$ true
      **for** $i = 0, \ldots, m-2$ **do**
        **if** $(x_j[i] > x_j[i+1])$ **then**
          swap$(x_j[i], x_j[i+1])$
          swap(index$[i]$,index$[i+1]$)
          bSorted$\leftarrow$ false
        **end if**
      **end for**
    **end while**
    $k \leftarrow 0$
    **repeat**
      bEqual $\leftarrow$ true
      **for** $i = 0, \ldots, m-1$ **do**
        **if** index$[i] \neq h[k][i]$ **then**
          bEqual $\leftarrow$ false
          break
        **end if**
      **end for**
      **if** (bEqual=true) **then**
        $c[k] \leftarrow c[k] + \frac{A}{m}\,|x[0]| + \sum_{l=1}^{m-1} \left( \frac{A}{m}\,|x[l]| + \frac{1-A}{m-1}\,|x[l] - x[l-1]| \right)$
      **end if**
      $k \leftarrow k+1$
    **until** ($k \geq M$ OR bEqual=true)
  **end for**
  pe$\leftarrow$ 0
  $C \leftarrow \sum_{\forall k} c[k]$
  **for** $k = 0, \ldots, M-1$ **do**
    $p[k] \leftarrow c[k]/C$
    **if** $(p[k] > 0)$ **then**
      pe$\leftarrow$pe$-p[k] \ln p[k]$
    **end if**
  **end for**

*2.2. Experimental dataset*

The experimental dataset was composed of 206 blood glucose records sampled at 5 minutes during 24h (288 samples). The records were acquired at the Teaching Hospital of Móstoles, Madrid (Spain), from 262 subjects at risk of developing diabetes, according to any of the following criteria [30]:

- Essential hypertension.
- BMI$\geq 30 kg/m^2$.
- A first–degree relative diabetes diagnosis.

During the 3 year study, patients were followed up. At the end of this period, 18 out of 206 were considered to have become diabetic patients if at least two of the following criteria were met [30]:

- Fasting glucose$\geq 126 mg/dL$.
- $\text{HbA}_{1c} \geq 6.5\%$.
- Started on anti–diabetic drugs.

In case of contradicting or inconclusive results, tests were repeated. The remaining 56 subjects were excluded at some point of the study due to age, or interfering treatments. Further details of the experimental dataset can be found at [30].

Blood glucose was monitored for 3 days for each patient, but only a clean period (no artifacts, or less than 3 consecutive missing samples, which were interpolated) of 24h was considered for analysis. If possible, this period started at 8:00 on day 2. Although longer records would be desirable [1], subjects are reluctant to be monitored for more than a few days, and the longer the records, the more likely artifacts are.

Records of patients who finally were diagnosed of diabetes were termed $D$, whereas the remaining records were termed $ND$. An example of each class is shown in Fig. 1.

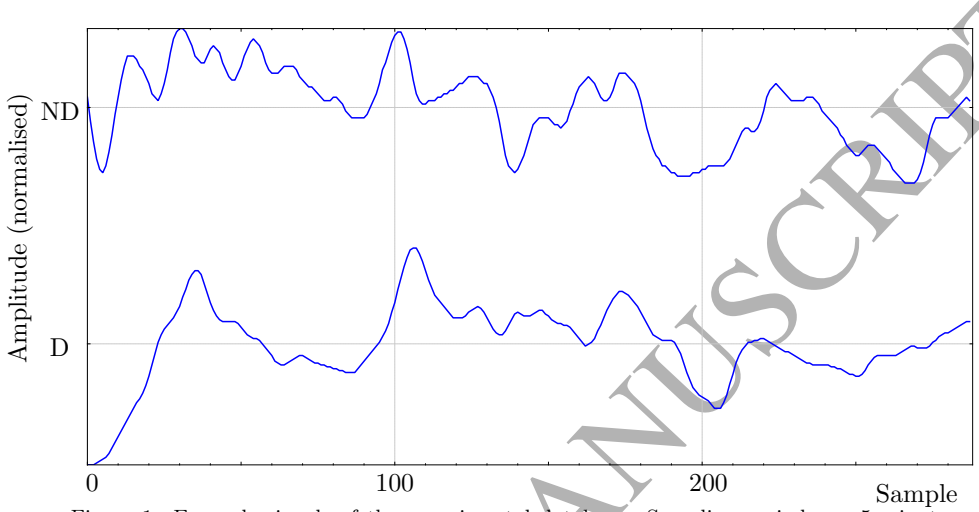


Figure 1: Example signals of the experimental database. Sampling period was 5 minutes. Duration 24h.

### *2.3. Statistical analysis*

The key items of the statistical analysis to assess the validity of the approach proposed are:

- *Statistical significance.* The results obtained with PE or any of its derivatives studied, were first analysed using a Shapiro–Wilk test to assess the distribution of the data. Since this procedure confirmed the normality of all the results, no further analysis was required in this regard. Then, a Student's t–test was applied to quantify the statistical significance of the possible differences between $D$ and $ND$ records. The threshold for significance was set at $p < 0.05$.

- *Classification performance.* The differences were studied using the Area Under Curve (AUC) of the Receiving Operating Characteristic (ROC)

as a generic performance measure [46]. The sensitivity was defined as the ratio of correctly classified $D$ records, and the specificity, the same for true negatives (the proportion of correctly classified non–diabetic records). The classification accuracy accounted for the correctly classified $D$ and $ND$ records.

- *Cross–validation.* A cross validation method was applied to assess the possible bias in the global classification results. The experimental dataset was randomly split into two sets of equal size, one used for training, and another one for testing. This validation was repeated 10 times, with replacement. The optimal PE threshold was obtained from the ROC analysis of the training set, and then it was applied to the test set. The results were quantified in terms of classification accuracy.

- *Survival analysis.* Since this classification also had an intrinsic time variability, namely, patients were diagnosed at quite different times during the follow–up period, we applied a Cox proportional hazard regression model for survival data analysis [47] to account for this variability. This model works with the hazard model formula. It finds a relationship for the hazard at time $t$, $h(t)$, for a patient according to a set of explanatory variables $z_i$:

$$h(t) = h_0(t)e^{\sum\limits_{\forall i}\beta_i z_i} \tag{5}$$

where $h_0(t)$ is the baseline hazard, when all $z_i = 0$, and it may vary with time, and $\exp(\beta_i)$ are the hazard ratios. If a ratio is greater than 1, the hazard associated to that variable increases and the survival or time of event becomes shorter.

## 3. Results

The capability of the basic PE method [17] was tested first. The AUC results for this experiment are shown in Table 1 for the entire dataset. Bandt and Pompe [17] recommended $m = 3, \ldots, 7$, and other works suggest $m! < N$ [27] to ensure a sufficient number of matches for a reliable estimation. However, there are studies that recommend to maximise $m$ to improve the resolution of differences in PE [26]. Therefore, we chose to explore a relatively wide range of $m$ values, from 2 up to 9. On the other hand, we tried several values for $\tau$, but the AUC dropped abruptly for $\tau > 1$. Consequently, only results with $\tau = 1$ are reported.

For $m = 2$, the results are not significant at all. For $m \geq 3$, there is a slightly growing trend with $m$ for AUC, confirmed by a decrease of $p$, well below the threshold for significance.

| | $m = 2$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| AUC | 0.531 | 0.712 | 0.729 | 0.725 | 0.720 | 0.728 | 0.745 | 0.753 |
| $p$ | 0.251 | 0.009 | 0.005 | 0.005 | 0.004 | 0.003 | 0.002 | 0.001 |

Table 1: AUC results for the standard PE method and $m$ ranging from 2 up to 9. Statistical significance was assessed using the Student's t–test.

A more detailed analysis of the ROC curves for PE, in terms of record classification performance, is summarised in Table 2. Sensitivity, specificity, and accuracy are quite stable for $m \geq 3$, similar to AUC. The results are the most homogeneous for $m \geq 7$, but with the highest sensitivity at $m = 3$, and the highest specificity and accuracy at $m = 5$. The interpretation of accuracy can be misleading in a few cases since it is very closely related to specificity due to the unbalanced classes (18 subjects for $D$, and 188 for $ND$).

For comparative purposes, the class separability analysis was repeated using the SampEn metric [9]. The highest AUC was obtained for $m = 1$, and $r = 0.26$, yielding AUC=0.667, $p = 0.0030$, Sensitivity=66.7%, Specificity=67.5%, and

| m | Sensitivity | Specificity | Accuracy |
|---|---|---|---|
| 9 | 72.2 | 72.3 | 72.3 |
| 8 | 72.2 | 72.3 | 72.3 |
| 7 | 72.2 | 73.4 | 73.3 |
| 6 | 77.8 | 63.8 | 65.0 |
| 5 | 61.1 | 80.9 | 79.2 |
| 4 | 72.2 | 67.0 | 67.5 |
| 3 | 77.8 | 56.9 | 58.7 |

Table 2: Classification results for PE with $m$ ranging from 3 up to 9.

Accuracy=67.4%. This is the maximum performance that could be achieved with SampEn. Other results with the usual recommended values of $m = 2$ and $r = 0.25$ were: AUC=0.597, Sensitivity=55%, Specificity=68%, not statistically significant, $p = 0.1745$. DFA was not used in the experiments because the standard DFA method did not find any significant differences in this dataset, a modified and customised version of the DFA algorithm is required, as described in [30], which is beyond the scope of this paper.

In order to try to improve the performance of PE when equal samples are contiguous, as is the case for glucose records, a small random noise level was added to the time series, as recommended in [17]. This level was 0.001% and 0.0001% of the peak–to–peak normalised amplitude of the input signal. Each test was repeated 100 times. The quantitative results are expressed as mean±SD (Standard Deviation) in Table 3, to account for the variability of the 100 random noise realisations.

| Noise | | $m = 3$ | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| 0.001% | AUC | $0.656 \pm 0.024$ | $0.674 \pm 0.019$ | $0.687 \pm 0.019$ | $0.704 \pm 0.015$ | $0.721 \pm 0.017$ | $0.732 \pm 0.015$ | $0.746 \pm 0.015$ |
| | $p$ | $0.025 \pm 0.031$ | $0.007 \pm 0.007$ | $0.004 \pm 0.006$ | $0.002 \pm 0.001$ | $0.002 \pm 0.001$ | $0.001 \pm 0.001$ | $0.001 \pm 0.001$ |
| 0.0001% | AUC | $0.652 \pm 0.026$ | $0.672 \pm 0.019$ | $0.690 \pm 0.018$ | $0.703 \pm 0.016$ | $0.717 \pm 0.016$ | $0.729 \pm 0.016$ | $0.744 \pm 0.016$ |
| | $p$ | $0.027 \pm 0.028$ | $0.008 \pm 0.007$ | $0.004 \pm 0.003$ | $0.002 \pm 0.002$ | $0.002 \pm 0.001$ | $0.002 \pm 0.001$ | $0.001 \pm 0.001$ |

Table 3: AUC results for the standard PE method with small random noise added to the experimental dataset to avoid consecutive equal values. Standard deviation for each parameter is included to provide an insight of the results' stability.

Noise amplitudes of 0.01% or higher, blurred the distinguishing features of the records, and no significant differences were found for any $m$ (AUC=$0.568 \pm$

0.041 for $m = 9$). Smaller amplitudes yielded similar performance as in Table 3, although slightly lower (AUC=0.699 $\pm$ 0.024 for $m = 9$ and 0.00001%). In any case, the addition of random perturbations does not seem to improve the performance of the standard PE method, despite consecutive equal values being a frequent anomaly in these records. Therefore, this approach is not used in the rest of the experiments.

Using the method proposed in [34] with $m = 4$, the result was AUC=0.699, with $p = 0.010$, statistically significant but worse than with the standard PE method, and at a higher memory and implementation cost. This modified PE algorithm uses additional symbol permutations to account for ties, assigning the same ordinal index to all the equal values, namely, repetitions are now possible in the $\sigma$ sequence. As a result, for $m = 4$, the $m! = 24$ possible patterns become 75, the corresponding Bell number [40], to include the patterns with ties, such as $[1, 2, 2, 2]$ and $[2, 2, 2, 1]$, among many others [34].

The amplitude differences between sequences were addressed implementing Eq. 3. The values tested for parameter $A$ were 0, 0.5, and 1. The results obtained in this case are shown in Table 4. As in the standard PE algorithm, AUC increases with $m$, with the maximum value at $m = 9$, but with higher values in AIPE.

| | | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $A = 0.0$ | AUC | 0.545 | 0.751 | 0.751 | 0.760 | 0.762 | 0.765 | 0.771 | 0.782 |
| | $p$ | 0.693 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| $A = 0.5$ | AUC | 0.500 | 0.733 | 0.751 | 0.748 | 0.750 | 0.754 | 0.765 | 0.775 |
| | $p$ | 0.673 | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |
| $A = 1.0$ | AUC | 0.522 | 0.721 | 0.753 | 0.749 | 0.745 | 0.751 | 0.760 | 0.769 |
| | $p$ | 0.411 | 0.003 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |

Table 4: Results obtained using the AIPE method for the same experiments.

The highest AUC was obtained for $A = 0.0$ and $m = 9$. The value of the $A$ parameter is not in agreement with the recommendations of the proposers of the

method [27]: medium $A$ values for classification purposes, and low $A$ values for spike detection. However, the differences are very small. Furthermore, $A = 0.0$ enables the simplification of Eq. 3.

The classification performance of the AIPE method in that case is shown in Table 5. The accuracy is also quite stable with $m$, with the highest value at $m = 6$ due to the maximum Specificity.

| m | Sensitivity | Specificity | Accuracy |
|---|---|---|---|
| 9 | 77.8 | 71.8 | 72.3 |
| 8 | 72.2 | 73.4 | 73.3 |
| 7 | 66.6 | 75.0 | 74.3 |
| 6 | 66.6 | 79.3 | 78.2 |
| 5 | 77.8 | 70.2 | 70.8 |
| 4 | 77.8 | 75.0 | 75.2 |
| 3 | 72.2 | 78.2 | 77.7 |

Table 5: Classification results for AIPE with $A = 0.0$, and $m = 3, ..., 9$.

According to the results in Table 5, the cross validation test was conducted for $m = 4$ and $A = 0.0$ using the AIPE method. The numerical average results of AIPE values were $3.5878 \pm 0.3916$ for the non–diabetic class, and $3.1564 \pm 0.4166$ for the diabetic class. The box plots in Figure 2 graphically show these results. They are in accordance with the hypothesis of decomplexification of pathological systems [48]. A healthy system has arguably a finer regulatory capability, and it does not allow big physiological excursions from normality, trying to mitigate them as soon as they are detected. Conversely, a pathological system has a more delayed response. As a result, records from healthy subjects are expected to exhibit frequent low amplitude oscillations (higher complexity), whereas records from a dysregulated system contain longer and larger oscillations (higher variability) [30].

The results of the cross validation test are shown in Table 6. Half of the records of each class were used for training, and the others for validation. The selection of the records took place randomly with replacement, 10 times. The

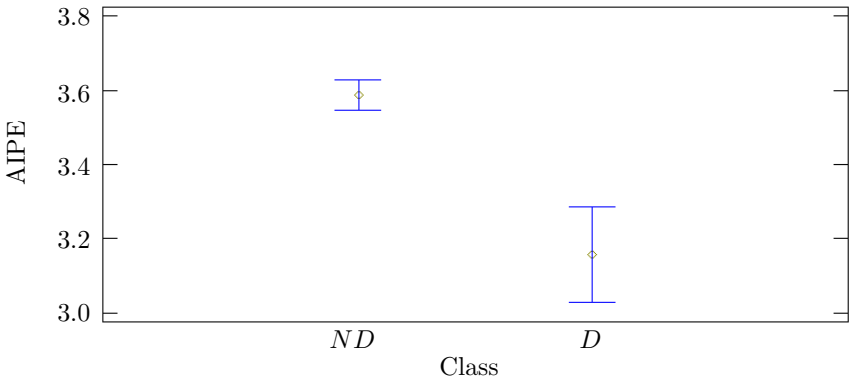


Figure 2: Box plots of the PE results for the classes $ND$ and $D$ (mean and 95%).

classification threshold was obtained from the ROC curves as the closest point to (0,1) [49]. Fig. 3 depicts the ROC curve obtained for the entire dataset.

Despite having half the number of subjects for analysis, the Shapiro–Wilk test confirmed the data normality, even for the $D$ training/validation class with only 9 samples. The differences were also still statistically significant, with $p = 0.022 \pm 0.014$.

| Training set | AIPE Threshold | Accuracy(%) |
|---|---|---|
| 1 | 3.247 | 74.8 |
| 2 | 3.405 | 72.8 |
| 3 | 3.234 | 76.7 |
| 4 | 3.338 | 76.7 |
| 5 | 3.305 | 74.8 |
| 6 | 3.405 | 67.0 |
| 7 | 3.364 | 71.8 |
| 8 | 3.247 | 79.6 |
| 9 | 3.405 | 68.9 |
| 10 | 3.303 | 73.8 |
| Mean±STD | 3.326±0.069 | 73.7±3.8 |

Table 6: Results of the 10 cross validation tests for $m = 4$ and $A = 0.0$.

The Cox proportional–hazards model was computed using the statistical

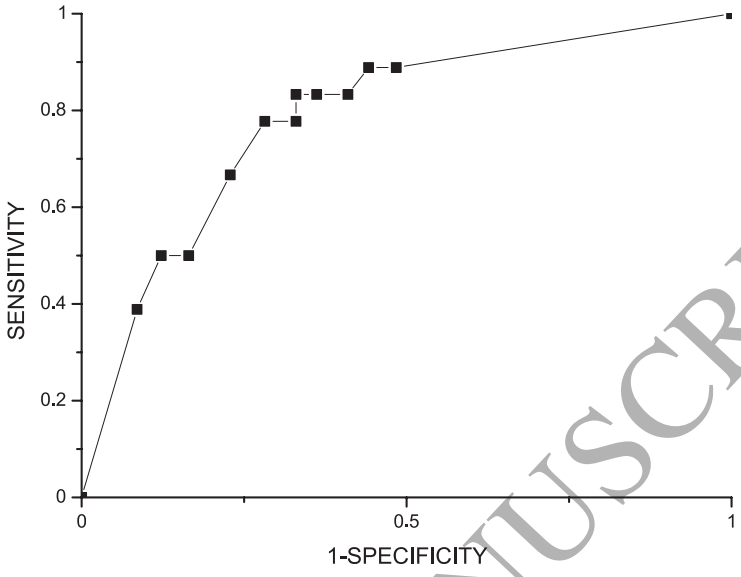


Figure 3: ROC curve for the entire dataset. AUC=0.782.

software R [50], package *survival*with 206 samples and 18 events. The results of this analysis are shown in Table 7 (using the same AIPE values as for Table 8).

| | coef. | exp(coef) | se(coef) | $z$ | Pr($>\|z\|$) |
|---|---|---|---|---|---|
| AIPE | -2.87342 | 0.05651 | 0.66261 | -4.337 | 0.0000145 |
| | exp(coef) | exp(-coef) | lower .95 | upper .95 | |
| AIPE | 0.05651 | 17.7 | 0.01542 | 0.2071 | |

Concordance=0.812 (se=0.073)
Rsquare=0.099 (max. possible 0.565)
Likelihood ratio test=21.45 on 1 df, $p = 0.000003641$
Wald test=18.81 on 1 df, $p = 0.00001448$
Score (logrank) test=21.5 on 1 df, $p = 0.000003535$

Table 7: Results obtained using the Cox regression analysis available in the R statistical software package.

These results confirm that AIPE variable seems to be a reliable tool to find differences between prospective $D$ and $ND$ records. Since the sign is negative, the hazard (becoming diabetic) is higher for subjects with lower values of that variable. The hazard ratio (exp(coef)) of developing diabetes is 0.05651 for

each AIPE increment of 1, or 17.7 for each AIPE decrement of 1, which again confirms the predictive value of AIPE. Moreover, the analysis of proportional hazards hypothesis yielded $p = 0.45$, in other words, we can safely conclude that the risks are not proportional. As an example, the time of event, in months, is shown for all records in class $D$ in Table 8, as well as the class assigned by the method.

| $D$ record | Month of event | AIPE | Class assigned |
|---|---|---|---|
| 1 | 12 | 4.016 | $ND$ |
| 2 | 24 | 3.303 | $D$ |
| 3 | 21 | 3.225 | $D$ |
| 4 | 10 | 3.516 | $ND$ |
| 5 | 10 | 2.532 | $D$ |
| 6 | 7 | 2.983 | $D$ |
| 7 | 31 | 4.030 | $ND$ |
| 8 | 12 | 3.405 | $ND$ |
| 9 | 2 | 2.761 | $D$ |
| 10 | 11 | 3.338 | $D$ |
| 11 | 24 | 3.243 | $D$ |
| 12 | 13 | 3.072 | $D$ |
| 13 | 17 | 2.800 | $D$ |
| 14 | 3 | 3.220 | $D$ |
| 15 | 16 | 2.631 | $D$ |
| 16 | 15 | 2.811 | $D$ |
| 17 | 9 | 2.915 | $D$ |
| 18 | 14 | 3.005 | $D$ |

Table 8: Time of event (diabetes diagnosis) for all the 18 records in class $D$.

## 4. Discussion

In this study, PE has been utilised to characterise the differences between $D$ and $ND$ blood glucose records before the disease was diagnosed. The original PE method provided significant differences for all the embedded dimensions tested except $m = 2$. Specifically, the best classification results were achieved for $m \geq 7$. In the original paper [17], authors recommended $m = 3, \ldots, 7$. We included two more cases as an attempt to find the maximum AUC, but

it seems to lie beyond $m = 9$. However, due to the computational cost and memory requirements, it was not possible for us to test the method for $m \geq 10$. Moreover, the classification accuracy is less dependent on $m$, overpassing the necessity of greater $m$ values. For illustrative purposes only, the running time of each test for $m = 3$ is a few seconds, whereas for $m = 9$ is 40 minutes, using a computer with an Intel©Core i7 processor at 2.6 GHz and 16 GB of RAM.

The results obtained with SampEn were relatively poor, with a maximum accuracy of 67.4%. Moreover, SampEn was very sensitive to the input parameters $m$, and $r$. In fact, the optimal values had to be found by a grid search. The records are probably too short for this metric, 288 samples, and that is why a more robust measure is necessary in this context.

The addition of noise to remove the ambiguities due to sample equalities did not seem to improve the performance of PE. Authors in [17] probably assumed equal values to be very rare in continuous distributions, but that is not the case in blood glucose records. For all the levels tested, the performance in terms of AUC decreased. Obviously, if the perturbations exceed a certain limit, signal differences are blurred by the noise and the method fails to find any significant segmentation between classes. This is the case for the noise amplitude of 0.01%. Once the equalities are broken without excessive signal distortion, smaller noise amplitudes does not seem to further improve performance. Other methods to address this drawback [34] did not improve the performance either.

Addressing amplitude differences with the AIPE method, more significant differences than with the PE method between $D$ and $ND$ records were found, in almost all embedding dimensions, from $m = 3$ up to $m = 9$. Only the case $m = 2$ seems to underperform again. The results are fairly similar in the range tested, with no significant changes in classification performance with $m$, with the highest AUC at $m = 9$, and the highest classification accuracy

at $m = 6$ (Table 5). This is a great advantage of PE and AIPE over other complexity or regularity estimators, very sensitive to the input parameter values. The additional parameter $A$ for AAPE also exerts a minimum influence on the results, being $A = 0.0$ slightly the best selection in this case (Table 4). The global performance of AIPE is better than that of PE, at the expense of a little bit more computational complexity.

The final cross validation test confirms the goodness of the approach based on AIPE. Although the performance is moderately lower in terms of classification accuracy, and the results are less statistically significant than for the entire groups, they are still valid. The deletion of the test instances causes a perturbation in the dataset that arguably decreases the accuracy of the classifier predictions, but except for two cases, the accuracy is well above 70%.

The Cox survival analysis also confirmed the applicability of AIPE as a metric to classify the individuals into $D$ and $ND$. The higher AIPE, the less risk of developing diabetes. This analysis employed the time–of–event data available during the three year study, but the accuracy of the method could arguably be improved if the follow-up had been longer, or more uniform (all the subjects followed–up till the end of the study). It can be reasonably hypothesized that some $ND$ patients became $D$ afterwards. Unfortunately, no specific data in this regard was available.

## 5. Conclusions

In this paper, we have described and compared several methods related to PE for glucose time series analysis. The results seem to confirm that series from patients that will eventually develop a diabetes may exhibit a lower complexity than healthier counterparts. The results also show that the performance of all the variations tested is quite similar, the influence of $m$ is almost negligible

for $m \geq 3$, and the amplitude differences are more representative than equal consecutive samples. The main limitations of the study are the relative small sample size, mainly for the $D$ class (18 subjects), and the short duration of the records, 24h.

The PE algorithm and some of its derivatives have been applied for the first time to glucose records. This metric was able to find statistically significant differences between records of future diabetic and non–diabetic patients, acquired up to three years before the diagnosis, with only 288 samples. The numerical results suggest that there is a correlation between the lower AIPE values, and the possibility of becoming diabetic. The method proposed, once confirmed in further studies, could be implemented on new preventive medical tools, which is of vital importance given the tremendous challenge that diabetes entails, and all the accompanying clinical complications. The core of these medical tools would be a continuous glucose monitoring and PE computation scheme, with a medium term analysis of the PE trend. In case this trend reflected a significant drop in PE values, countermeasures (medical, behavioural, dietary) should be applied, while suspected diabetes is still on the way.

The analysis of this kind of biomedical records using complexity or regularity measures is often a very difficult task. Not many databases are available, usually involving a few short records, and artifacts such as missing samples or saturated epochs are a recurrent phenomenon in these signals. Only works based on DFA [31], and to a less extent, on SampEn variations [33], have been successful so far in this context. However, in this case the results obtained with SampEn did not fulfil the expectations, probably due to the short length of the records. It is feasible that longer records would contribute to more consistent findings regardless of the metric employed. The recently approved International Consensus on Use of Continuous Glucose Monitoring [1] recommends to acquire at least two weeks

of data, and, if possible, reported in three time blocks (sleep, wake, and 24h). The availability of these data will surely foster new and improved methods of analysis.

The main advantages of the method proposed are its simplicity and stability, inherited from the underlying PE method. The algorithm can be easily implemented in any computer platform and programming language. If memory is a constraint, low values of $m$ also provide good classification results. In any case, different $m$ values yield a very similar performance. On the other hand, a disadvantage of the algorithm is the computation by default of all the $m!$ permutations, instead of creating them dynamically, as they are found in the input sequence. This is an open issue in all PE methods that should be addressed in future studies.

Although the method finally proposed achieved a promising performance on class recognition, the experimental dataset was drawn from a specific population, and the generalisation of the results requires further studies using other patient cohorts. Besides, the method is based on ordinal patterns, and seems to outperform amplitude based methods (SampEn), but a combination of both approaches could be even more sensitive, since each scheme alone does not probably provide the full picture of the glucose dynamics.

In addition to find better PE derivatives or preprocessing techniques (to avoid equalities in neighbouring values), it would be also very important to find more efficient PE algorithms in terms of memory requirements and permutation search cost. For relatively high values of $m$, it is very difficult to run the algorithm on state-of-the-art personal computers, and therefore this $m$ region remains unexplored in many cases.

## 6. Acknowledgements

The authors declare that there is no conflict of interest.

No funding was received to support this research work.

**References**

[1] T. Danne, R. Nimri, T. Battelino, R. M. Bergenstal, K. L. Close, J. H. DeVries, S. Garg, L. Heinemann, I. Hirsch, S. A. Amiel, R. Beck, E. Bosi, B. Buckingham, C. Cobelli, E. Dassau, F. J. Doyle, S. Heller, R. Hovorka, W. Jia, T. Jones, O. Kordonouri, B. Kovatchev, A. Kowalski, L. Laffel, D. Maahs, H. R. Murphy, K. Nørgaard, C. G. Parkin, E. Renard, B. Saboo, M. Scharf, W. V. Tamborlane, S. A. Weinzimer, M. Phillip, International consensus on use of continuous glucose monitoring, Diabetes Care 40 (12) (2017) 1631–1640. doi:10.2337/dc17-1600.

[2] J. Pettus, S. V. Edelman, Recommendations for using real-time continuous glucose monitoring (rtcgm) data for insulin adjustments in type 1 diabetes, Journal of Diabetes Science and Technology 11 (1) (2017) 138–147, pMID: 27530720. doi:10.1177/1932296816663747.

[3] G. P. Forlenza, D. Raghinaru, F. Cameron, B. Wayne Bequette, H. Peter Chase, R. Paul Wadwa, D. M. Maahs, E. Jost, T. T. Ly, D. M. Wilson, L. Norlander, L. Ekhlaspour, H. Min, P. Clinton, N. Njeru, J. W. Lum, C. Kollman, R. W. Beck, B. A. Buckingham, for the In-Home Closed-Loop (IHCL) Study Group, Predictive hyperglycemia and hypoglycemia minimization: In-home double-blind randomized controlled evaluation in children and young adolescents, Pediatric Diabetes (2017) n/a–n/adoi:10.1111/pedi.12603.

[4] K. Kohnert, S. E, Useful measures to assess glucose dynamics from continuous glucose monitoring data, Current research in diabetes and obesity journal 1 (3) (2017) 1–4.

[5] G. Fico, L. Hernández, J. Cancela, M. M. Isabel, A. Facchinetti, C. Fabris, R. Gabriel, C. Cobelli, M. T. A. Waldmeyer, Exploring the frequency

domain of continuous glucose monitoring signals to improve characterization of glucose variability and of diabetic profiles, Journal of Diabetes Science and Technology 11 (4) (2017) 773–779, pMID: 28627250. doi:10.1177/1932296816685717.

[6] M. D. Breton, D. P. Shields, B. P. Kovatchev, Optimum subcutaneous glucose sampling and fourier analysis of continuous glucose monitors, Journal of Diabetes Science and Technology 2 (3) (2008) 495–500, pMID: 19885217. doi:10.1177/193229680800200322.

[7] I. Iancu, E. Iancu, M. Mota, Spectral analysis of the blood glucose time series for automated diagnosis, in: 1st WSEAS International Conference on Sensors and Signals, 2008, pp. 88–93.

[8] S. Pincus, I. Gladstone, R. Ehrenkranz, A regularity statistic for medical data analysis, J. of Clin. Monit. and Comput. 7 (4) (1991) 335–345.

[9] J. Richman, J. R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am J Physiol Heart Circ Physiol 278 (6) (2000) H2039–2049.

[10] W. Chen, J. Zhuang, W. Yu, Z. Wang, Measuring complexity using FuzzyEn, ApEn, and SampEn, Medical Engineering & Physics 31 (1) (2009) 61–68.

[11] M. Rostaghi, H. Azami, Dispersion entropy: A measure for time-series analysis, IEEE Signal Processing Letters 23 (5) (2016) 610–614. doi:10.1109/LSP.2016.2542881.

[12] R. K. Tripathy, S. Deb, S. Dandapat, Analysis of physiological signals using state space correlation entropy, Healthcare Technology Letters 4 (1) (2017) 30–33. doi:10.1049/htl.2016.0065.

[13] G. Manis, M. Aktaruzzaman, R. Sassi, Bubble entropy: An entropy almost free of parameters, IEEE Transactions on Biomedical Engineering 64 (11) (2017) 2711–2718. doi:10.1109/TBME.2017.2664105.

[14] A. Lempel, J. Ziv, On the complexity of finite sequences, IEEE Transactions on Information Theory 22 (1) (1976) 75–81. doi:10.1109/TIT.1976.1055501.

[15] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, A. L. Goldberger, Mosaic organization of dna nucleotides, Phys. Rev. E 49 (2) (1994) 1685–1689. doi:10.1103/PhysRevE.49.1685.

[16] P. Li, C. Liu, K. Li, D. Zheng, C. Liu, Y. Hou, Assessing the complexity of short-term heartbeat interval series by distribution entropy, Medical & Biological Engineering & Computing 53 (1) (2015) 77–87. doi:10.1007/s11517-014-1216-0.

[17] C. Bandt, B. Pompe, Permutation entropy: A natural complexity measure for time series, Phys. Rev. Lett. 88 (2002) 174102. doi:10.1103/PhysRevLett.88.174102.

[18] J.-L. Chen, P.-F. Chen, H.-M. Wang, Decreased complexity of glucose dynamics in diabetes: evidence from multiscale entropy analysis of continuous glucose monitoring system data, American Journal of Physiology - Regulatory, Integrative and Comparative Physiology 307 (2) (2014) R179–R183. doi:10.1152/ajpregu.00108.2014.

[19] M. Engoren, T. A. Schwann, R. H. Habib, Hyperglycemia, hypoglycemia, and glycemic complexity are associated with worse outcomes after surgery, Journal of Critical Care 29 (4) (2014) 611 – 617.

[20] R. Brunner, G. Adelsmayr, H. Herkner, C. Madl, U. Holzinger, Glycemic

variability and glucose complexity in critically ill patients: a retrospective analysis of continuous glucose monitoring data, Critical Care 16 (5) (2012) R175. doi:10.1186/cc11657.

[21] L. Crenier, M. Lytrivi, A. Van Dalem, B. Keymeulen, B. Corvilain, Glucose complexity estimates insulin resistance in either nondiabetic individuals or in type 1 diabetes, The Journal of Clinical Endocrinology and Metabolism 101 (4) (2016) 1490–1497. doi:10.1210/jc.2015-4035.

[22] L. Vigil, E. Condés, M. Varela, C. Rodriguez, A. Colas, B. Vargas, M. Lopez, E. Cirugeda, Glucose series complexity in hypertensive patients, Journal of the American Society of Hypertension 8 (9) (2014) 630 – 636. doi:http://dx.doi.org/10.1016/j.jash.2014.05.008.

[23] M. Varela, C. Rodriguez, L. Vigil, E. Cirugeda, A. Colas, B. Vargas, Glucose series complexity at the threshold of diabetes, Journal of Diabetes 7 (2) (2015) 287–293. doi:10.1111/1753-0407.12182.

[24] L. Zunino, F. Olivares, O. A. Rosso, Permutation min-entropy: An improved quantifier for unveiling subtle temporal correlations, EPL (Europhysics Letters) 109 (1) (2015) 10005.

[25] M. Zanin, L. Zunino, O. A. Rosso, D. Papo, Permutation entropy and its main biomedical and econophysics applications: A review., Entropy 14 (8) (2012) 1553–1577.

[26] M. Riedl, A. Müller, N. Wessel, Practical considerations of permutation entropy, The European Physical Journal Special Topics 222 (2) (2013) 249–262. doi:10.1140/epjst/e2013-01862-7.

[27] H. Azami, J. Escudero, Amplitude-aware permutation entropy: Illustration in spike detection and signal segmentation, Computer methods and

programs in biomedicine 128 (2016) 40–51. doi:10.1016/j.cmpb.2016.02.008.

[28] J. M. Yentes, N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath, N. Stergiou, The appropriate use of approximate entropy and sample entropy with short data sets, Annals of Biomedical Engineering 41 (2) (2013) 349–365. doi:10.1007/s10439-012-0668-3.

[29] P. Carpena, M. Gómez-Extremera, C. Carretero-Campos, P. Bernaola-Galván, A. V. Coronado, Spurious results of fluctuation analysis techniques in magnitude and sign correlations, Entropy 19 (6) (2017),doi:10.3390/e19060261.

[30] C. Rodríguez de Castro, L. Vigil, B. Vargas, E. García Delgado, R. García Carretero, J. Ruiz-Galiana, M. Varela, Glucose time series complexity as a predictor of type 2 diabetes, Diabetes/Metabolism Research and Reviews 33 (2) (2017) e2831. doi:10.1002/dmrr.2831.

[31] H. Ogata, K. Tokuyama, S. Nagasaka, A. Ando, I. Kusaka, N. Sato, A. Goto, S. Ishibashi, K. Kiyono, R. Struzik, Y. Yamamoto, Long–range correlated glucose fluctuations in diabetes, Methods Inf Med 46 (2) (2007) 222–226.

[32] M. D. Costa, T. Henriques, M. N. Munshi, A. R. Segal, A. L. Goldberger, Dynamical glucometry: Use of multiscale entropy analysis in diabetes, Chaos: An Interdisciplinary Journal of Nonlinear Science 24 (3) (2014) 033139. doi:10.1063/1.4894537.

[33] J.-L. Chen, P.-F. Chen, H.-M. Wang, Decreased complexity of glucose dynamics in diabetes: evidence from multiscale entropy analysis of continuous glucose monitoring system data, American Journal of Physiology-

Regulatory, Integrative and Comparative Physiology 307 (2) (2014) R179–R183, pMID: 24808497. doi:10.1152/ajpregu.00108.2014.

[34] C. Bian, C. Qin, Q. D. Y. Ma, Q. Shen, Modified permutation-entropy analysis of heartbeat dynamics, Phys. Rev. E 85 (2012) 021906. doi:10.1103/PhysRevE.85.021906.

[35] L. Zunino, F. Olivares, F. Scholkmann, O. A. Rosso, Permutation entropy based time series analysis: Equalities in the input signal can lead to false conclusions, Physics Letters A 381 (22) (2017) 1883 – 1892. doi:https://doi.org/10.1016/j.physleta.2017.03.052.

[36] L. Zunino, C. W. Kulp, Detecting nonlinearity in short and noisy time series using the permutation entropy, Physics Letters A 381 (42) (2017) 3627 – 3635. doi:https://doi.org/10.1016/j.physleta.2017.09.032.

[37] C. Bandt, B. Pompe, Permutation entropy: A natural complexity measure for time series, Phys. Rev. Lett. 88 (2002) 174102. doi:10.1103/PhysRevLett.88.174102.

[38] L. Xiao-Feng, W. Yue, Fine-grained permutation entropy as a measure of natural complexity for time series, Chinese Physics B 18 (7) (2009) 2690.

[39] D. Li, Z. Liang, Y. Wang, S. Hagihira, J. W. Sleigh, X. Li, Parameter selection in permutation entropy for an electroencephalographic measure of isoflurane anesthetic drug effect, Journal of Clinical Monitoring and Computing 27 (2) (2013) 113–123. doi:10.1007/s10877-012-9419-0.

[40] E. Shumbayawonda, A. Fernández, M. P. Hughes, D. Abasolo, Permutation entropy for the characterisation of brain activity recorded with magnetoencephalograms in healthy ageing, Entropy 19 (6) (2017),doi:10.3390/e19040141.

[41] D. Jordan, M.Sc., G. Stockmanns, Ph.D., E. Kochs, M.D., S. Pilge, M.D., G. Schneider, M.D., Electroencephalographic order pattern analysis for the separation of consciousness and unconsciousness. an analysis of approximate entropy, permutation entropy, recurrence rate, and phase coupling of order recurrence plots, Anesthesiology 109 (6) (2008) 1014–1022. doi:10.1097/ALN.0b013e31818d6c55.

[42] Y. Zhuang, L. Yang, Y. Xia, C. Liu, Effect of equalities in rri time series on permutation entropy under different emotional states, in: 2017 Chinese Automation Congress (CAC), 2017, pp. 3380–3383. doi:10.1109/CAC.2017.8243363.

[43] D. Cuesta-Frau, M. Varela-Entrecanales, A. Molina-Picó, B. Vargas, Patterns with equal values in Permutation Entropy. Do they really matter for biosignal classification?, Complexity (2018) (In press).

[44] X. D. Zhang, D. Pechter, L. Yang, X. Ping, Z. Yao, R. Zhang, X. Shen, N. X. Li, J. Connick, A. R. Nawrocki, M. Chakravarthy, C. Li, Decreased complexity of glucose dynamics preceding the onset of diabetes in mice and rats, PLOS ONE 12 (9) (2017) 1–11. doi:10.1371/journal.pone.0182810.

[45] A. S. Leonardou, C. Karystianos, C. Argyropoulos, G. C. Nikiforidis, F. Kalfarentzos, T. K. Alexandrides, Restoration of high-frequency glucose-entrained insulin oscillations in obese subjects with type 2 diabetes after biliopancreatic diversion, Surgery for Obesity and Related Diseases 12 (8) (2016) 1539 – 1547. doi:https://doi.org/10.1016/j.soard.2016.04.009.

[46] A. P. Bradley, The use of the area under the roc curve in the evaluation

of machine learning algorithms, Pattern Recognition 30 (7) (1997) 1145 – 1159. doi:https://doi.org/10.1016/S0031-3203(96)00142-2.

[47] C. D. R., Regression models and life tables, Journal of the Royal Statistic Society B (34) (1972) 187–202.

[48] M. Varela, R. Ruiz-Esteban, M. J. Mestre De Juan, Chaos, fractals, and our concept of disease, Perspectives in Biology and Medicine 53 (2010) 584–595. doi:10.1353/pbm.2010.0003.

[49] D. Robinson, The statistical evaluation of medical tests for classification and prediction by m. sullivan pepe, Journal of the Royal Statistical Society: Series A (Statistics in Society) 169 (3) (2006) 656–656. doi:10.1111/j.1467-985X.2006.00430_9.x.

[50] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2013).
URL http://www.R-project.org/