



Improving Health Care Management Through Persistent Homology of Time-Varying Variability of Emergency Department Patient Flow

Maël Dugast, Guillaume Bouleux, Olivier Mory, Eric Marcon

► To cite this version:

Maël Dugast, Guillaume Bouleux, Olivier Mory, Eric Marcon. Improving Health Care Management Through Persistent Homology of Time-Varying Variability of Emergency Department Patient Flow. IEEE Journal of Biomedical and Health Informatics, 2018, 23 (5), pp.2174–2181. 10.1109/JBHI.2018.2882748 . hal-01927284

HAL Id: hal-01927284

<https://hal.science/hal-01927284>

Submitted on 19 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Improving Health Care Management Through Persistent Homology of Time-Varying Variability of Emergency Department Patient Flow

Maël Dugast*, Guillaume Bouleux*, Olivier Mory[†], and Eric Marcon*

Abstract

Excessive admissions at the Emergency Department (ED) is a phenomenon very closely linked to the propagation of viruses. It is a cause of overcrowding for EDs and a public health problem. The aim of this work is to give EDs' leaders more time for decision making during this period. Based on the admissions time series associated with specific clinical diagnoses, we will first perform a Detrended Fluctuation Analysis (DFA) to obtain the corresponding variability time series. Next, we will embed this time series on a manifold to obtain a point cloud representation and use Topological Data Analysis (TDA) through persistent homology technic to propose two early real-time indicators. One is the early indicator of abnormal arrivals at the ED whereas the second gives the information on the time index of the maximum number of arrivals. The performance of the detectors is parameter dependent and it can evolve each year. That is why we also propose to solve a bi-objective optimization problem to track the variations of this parameter.

Index Terms

Emergency Department, Epidemic, Persistent Homology, Topological Data Analysis, Variability, Non-stationary Processes, Detection, Change Point, Crowding.

I. INTRODUCTION

Respiratory diseases strongly impact Emergency Departments (EDs) and paediatric EDs with, in particular, excessive hospital admissions during the period of Influenza activity and Respiratory Syncytial Virus (RSV) circulation [28], [5], [21]. Even if in France, the ARS's (Regional Health Agency) and the French ministry have worked in the past to define a winter plan for EDs, with more beds, more staff, and subsequent grants for a maximum period of 3 months, this still inevitably causes overcrowding [18], [16], [29], and helping EDs to counter this effect is of major interest. To get EDs prepared for seasonal epidemics, queueing modeling [19], forecasting methods [14], change point detection [4] or Artificial Neural Network (ANN) methods [17] have been proposed with sometimes a very short time for decision making. Then, giving the ED leaders the most time for decision making is the challenge we aim at tackling in this work. In order to address this challenge, we will use the time series of the patients arriving at the ED and we will focus our analysis and actual testing on the paediatric ED of Saint Etienne, France. Similarly to change points detection methods [4], [15], we have based our work on the estimation of the first time index where variability of ED admissions (here for a specific pathology) substantially changes. Indeed, although the global evolution of the ED admissions could be roughly modeled, the strong and time-varying daily variability prevent this coarse modeling to be efficient for the estimation of early behavioral changes. This time-varying and daily variations are actually a picture of the virus circulation, which is by essence unpredictable. Our main concern will then be the representation and the analysis of daily admissions variability and not the admissions flow.

First, we will highlight the variability through a Detrended Fluctuation Analysis (DFA). We will next use Topological Data Analysis (TDA) to have a deep learning of the DFA time-varying variability. TDA has been largely employed [11], [12] through persistent homology techniques for it reveals intrinsic topologic properties of the process under study. It explains the global form of an manifold embedding of the process. According to our research, some studies have indeed been done on TDA and its variability process, but the coupling of this analysis with the detection of change points seems to be a totally new concept. This paper is organized as follows: in Section II, we will briefly introduce the DFA, and in more detail, the persistent homology. In Section III, we will present the results obtained by the persistent homology when applied to the patients diagnosed with bronchiolitis at the Saint Etienne paediatric ED. An early outbreak detector is proposed in Section IV and its automatic calibration for real-time implementation is then discussed. A conclusion is finally drawn in Section V.

* Univ Lyon, INSA Lyon, DISP, EA 4570, F-69621, Villeurbanne, France. Tel.: +33477448920, Fax: +33477448951, {mael.dugast,guillaume.bouleux,eric.marcon}@insa-lyon.fr

[†] Urgences Medico-Chirurgicales Pédiatriques (Pediatric ED), University Hospital (CHU) of Saint Etienne, 42055 Saint Etienne CEDEX 2, France. Tel: +33477828632, Fax: +33477828098, omory@chu-st-etienne.fr

II. PERSISTENT HOMOLOGY FOR DATA POINT CLOUD

A. Detrended fluctuation analysis as a measure of variability

We will now focus our study on the RSV epidemic which massively impacts paediatric EDs with excess hospital admissions of patients diagnosed with bronchiolitis. We propose to the reader to refer to Fig. 1 to have an idea of such a behavior. The (daily plot of Fig. 1 draw three consecutive years ?) at the paediatric ED of Saint Etienne, France. In addition, we note that the development presented thereafter is valid for any time series representative of epidemics.

In order to investigate the variability of the patients' admissions and consequently give information on the underlying epidemic, we proceed with a preliminary processing step, known as Detrended Fluctuation Analysis (DFA). We compute first the timely integrated time series $\{X_t\}_{t=1}^T$ of the admissions as

$$Z_t = \sum_{k=1}^t (X_k - \bar{X}_t) \quad (1)$$

where \bar{X}_t is the mean of $\{X_t\}_{t=1}^T$. Then, $\{Z_t\}_{t=1}^T$ is divided into $N_L = \lfloor T/L \rfloor$ (with $\lfloor \cdot \rfloor$ the integer part) non-overlapping segments of length L . We find next the best linear approximation $\{\tilde{Z}_k^L\}_{k=1}^{N_L}$ of each of the N_L segments by least-squares fit. The variability time series $\{Y_k\}_{k=1}^{N_L}$ is finally reached by concatenating the piecewise time series obtained by

$$Y_k = Z_k^L - \tilde{Z}_k^L \quad (2)$$

where Z_k^L is the k -th, $k = 1, \dots, N_L$, length L segment of $\{Z_t\}_{t=1}^T$.

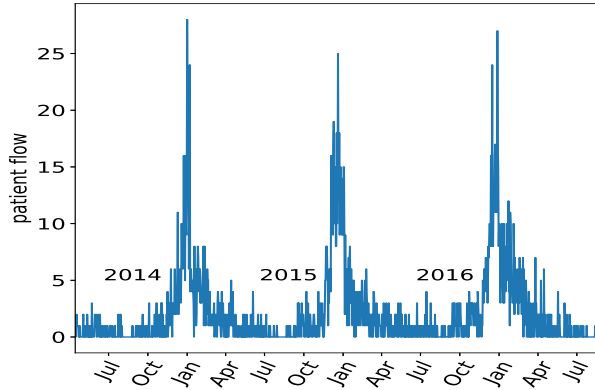


Fig. 1. Number of patients per day admitted at the paediatric ED of Saint Etienne, France for bronchiolitis symptoms during 2014-05-01 and 2017-08-31 period.

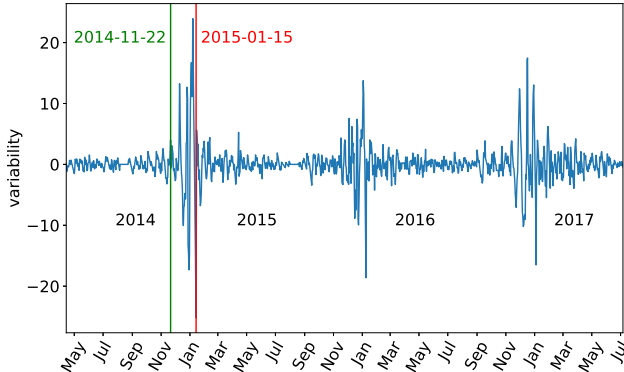


Fig. 2. Daily DFA representation of the admissions time series of Fig. 1, with $L = 20$.

Applying the DFA method on the admissions time series $\{X_t\}_{t=1}^{1096}$ represented by Fig. 1, gives the variability time series $\{Y_k\}_{k=1}^{36}$ plotted in Fig. 2. Some striking properties emerge from the analysis of this plot. We observe a clear periodicity in the variability process due to the seasonality of the RSV epidemic but still with unpredictable behavior. It reinforces the intuition

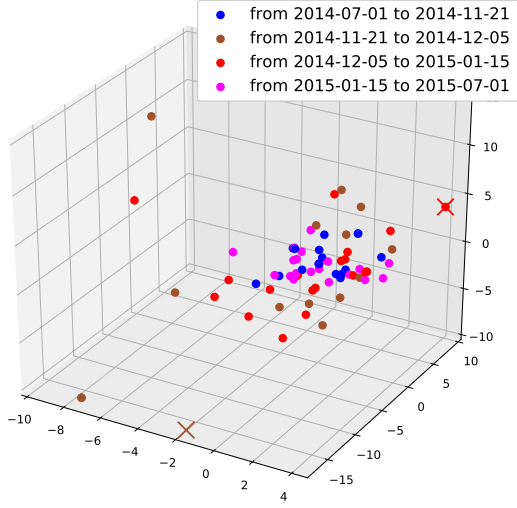


Fig. 3. The point cloud associated with the variability time series $\{Y_k\}_{k=1}^{36}$ plotted in Fig. 2. Different groups of points are differentiated by colors. We have particularly highlighted two points with a cross, they correspond respectively to day 2014-11-22 and 2014-12-05.

that early detection of an admissions can not be done by rough modeling of the admissions flow. Each epidemic period has its own behavior materialized by different time-evolving variability. As a consequence, it seems natural to assess the problem of epidemic detection by following the disruptions of the time-evolving variability. This corresponds to the use of second order statistical characteristics for change detection as many change detection methods such as CUSUM or the Shewhart chart implicitly do.

Time series analysis has long been used and hundreds of methods have been proposed. One way of dealing with this, is to embed the time series on a manifold [6] with for example Hankel-like embedding [3]. Inspired by [22], we use a sliding window of size N on the piecewise time series $\{Y_k\}_{k=1}^{N_L}$ determined by (2). Each of the $(N_L - N + 1)$ -sliding windows is then associated with a point in an N -euclidean space. The collection of the $(N_L - N + 1)$ points form the point cloud \mathcal{S} . For $N = 3$, we have plotted the embedding of $\{Y_k\}_{k=1}^{N_L}$ in Fig. 3. On this plot, we have colored four periods to highlight major facts. In particular, the periods from 2014-07-01 to 2014-11-21 and from 2015-01-15 to 2015-07-01 have points with the same behavior. They are concentrated inside a ball and all seem to be nearly equally spaced. This is really interesting because these two periods are associated with weak comings at the ED and there is no RSV epidemic at that time. On the contrary, for the period from 2014-11-21 to 2014-12-05, the points are no more nearly equally spaced but they spread off in the space. This period represents the beginning of the RSV epidemic and corresponds then to abnormally high arrivals at the ED. For the last period, which started at 2014-12-05 and ended at 2015-01-15, some of the points have high coordinates and some are concentrated inside a ball with weak coordinates. This is the period of the decrease of the epidemic and thus it indicates a decrease of the ED admissions. This point cloud representation finally tells everything about the time evolving of the ED admissions. This will be investigated in the rest of the paper by the use of Topological Data Analysis to (i) early detect abnormal arrivals associated with seasonal epidemics and (ii) provide an real-time indicator associated with the maximum number of arrivals; thus giving information about the decrease of the epidemic. The methodology is summarized in the following way:

- 1) We chose a time window of size L to perform the DFA on the time series of the admissions;
- 2) We chose next a window of size N to embed the obtained variability time series onto a system of N -coordinates;
- 3) The time varying point cloud thus obtained is finally analyzed thanks to TDA technique and we propose two indicators for decision making.

B. Homology and Persistent Homology

In this section, we introduce the Topological Data Analysis (TDA) of a point cloud and its application to time series. A point cloud is basically a collection of points in an n -dimensional space with a measure of the distance between them. The term "topological" refers to what remains of a structure when the geometry has changed. It is in a certain way the fundamental description of a geometrical object in terms of loops, holes, voids, etc. A useful example to illustrate the underlying concepts is to consider the case of an elastic circle: one can stretch, shrink, rotate, or deform this circle; it will still result in a circle,

assuming that the object is not torn apart, or broken. The fundamental description of a circle in this perspective consists of an object made up of one connected component forming a 1-dimensional hole. Although this description can seem to be straightforward, a rigorous mathematical framework is given in the literature to enable reliable numerical computations. This is the subject of the persistent homology theory. Persistent homology has been applied for time series classification, [27], image pattern recognition [8], biology [26], phylogenetic [7], science of language [23] and various other fields. To reveal topological features, the data point cloud under study is decomposed with the help of simplices whose forms are illustrated by Fig. 4. Next, a simplicial complex is obtained thanks to the collection of the simplices. The question of obtaining the simplices is answered in many ways but one of the easiest to compute is the Vietoris-Rips filtration technique. This method consists in fixing a number ρ , and simplices $K_\rho = (v_1, \dots, v_p)$ are formed as soon as $\|v_i - v_j\| < \rho$, $\forall i, j$. This operation is reproduced with an increasing ρ until a single simplex containing all the data is reached. This method is briefly illustrated in Fig. 5.

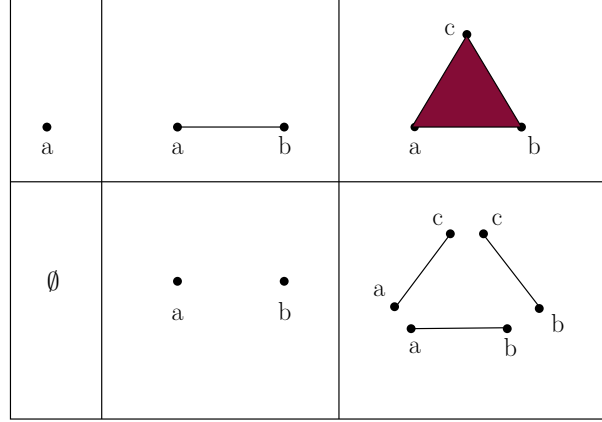


Fig. 4. A n -simplex is the convex hull of $(n+1)$ points in an n -dimensional space. It is the generalization of a triangle. Top (from left to right): A 0-simplex is a point, this is (a) into the plot. A 1-simplex is a segment, this is (a,b) (where a and b are the vertices that are at extremity of the segment), a 2-simplex is a triangle (a,b,c) in the plane. Bottom (from left to right): the corresponding faces.

Once the sequence of simplices is obtained, still remains the topological features to extract. We will not enter into technical considerations in this paper but details can be found in [11], [12] and [30]. The basic idea is to define a hole as a collection of simplices which forms a cycle, *i.e.* a set of simplices which starts and ends at the same point, but with the constraint that it is not the face of any other simplex. In Fig. 5, the big square-like figure is a hole whereas the triangles are not. During the construction of the Vietoris-Rips filtration, holes will appear and then later, disappear. This phenomenon is reported in a diagram called persistent diagram. An example of such a diagram based on the noisy circle of Fig. 6-(a) is provided in

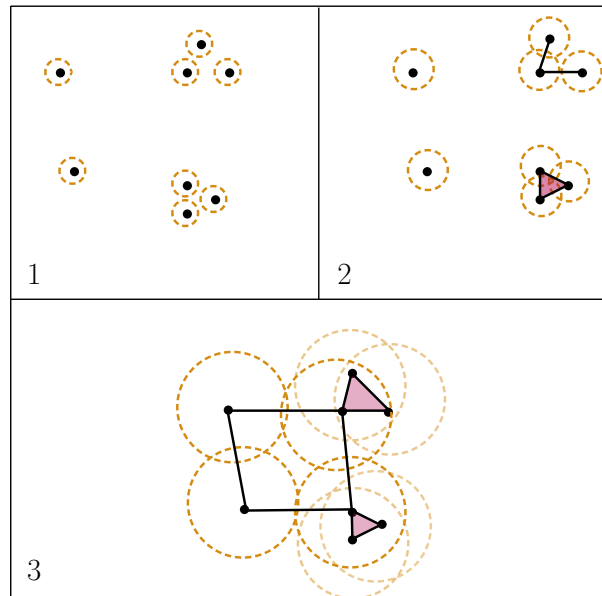


Fig. 5. A nested sequence of simplicial complexes is created by increasing the radius of balls represented in orange. As soon as two balls intersect, the points in their center are connected. When three balls intersect, we form the triangle, and so on. At stage 3 a hole appears. It is not yet the boundary of a simplex. This hole will disappear when its four vertices will be two by two connected.

Fig. 6-(b). A hole will naturally emerge at an early stage of the Vietoris-Rips filtration. It will survive longer than any other smaller holes that can appear. Regarding Fig. 6-(b), the lonely point (0.15, 1.3) indicates that this major hole was created for $\rho = 0.15$ and was destroyed for $\rho = 1.3$. It thus represents a main topological feature.

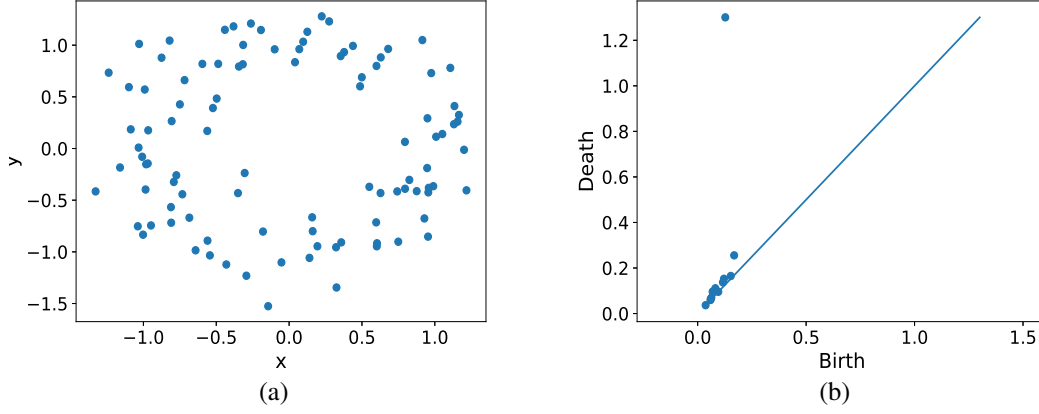


Fig. 6. (a) Noisy unit circle in \mathbb{R}^2 , (b) Persistent diagram of the noisy circle, the axes correspond to the steps ρ of the filtration.

III. PERSISTENT HOMOLOGY ON PATIENTS' ADMISSIONS

The following sections are now dedicated to the application of persistent homology on the time series of admissions for patients diagnosed with bronchiolitis.

A. Sequential Persistent Diagram

To track the evolution of an epidemic through the patients' admissions, we successively compute and store the point clouds $\{\mathcal{S}_k\}_{k=1}^{N_L-N+1}$. For each of the point cloud \mathcal{S}_k its associated persistent diagram $\mathcal{D}_k = \{(x_i, y_i)\}_{i=1}^{N_{p_k}}$, where (x_i, y_i) are the coordinates of the i -th point, is computed. The following two indicators

$$\text{Bar}_{\max}(k) = \max_{\mathcal{D}_k} \{y_i - x_i\}, i = 1, \dots, N_{p_k} \quad (3)$$

$$\text{PE}_{\text{norm}}(k) = -\frac{1}{\log(N_{p_k})} \sum_{j=1}^{N_{p_k}} p_j \log(p_j) \quad (4)$$

where $p_j = \frac{l_j}{L}$ with $l_j = y_j - x_j$, $L = \sum_{j=1}^{N_{p_k}} l_j$, are next used to summarize the information carried in the diagram. The indicator $\text{Bar}_{\max}(k)$ indicates the maximal survival time of a cycle in the diagram \mathcal{D}_k whereas the indicator $\text{PE}_{\text{norm}}(k)$ stands for the normalized entropy; some precursor definitions can be found in [10], [25]. It allows to understand the presence or not of particular structures in the persistent diagram or the point cloud itself. If all the points of one diagram \mathcal{D}_k represent quite the same survival time, meaning that all the cycles created and destroyed by the filtration have the same survival time, then the persistent entropy will reach its maximum value. This clearly means that the underlying point cloud is uniformly distributed. On the contrary, if the points of a diagram \mathcal{D}_k have many different survival times, the persistent entropy will reach a small value, meaning that the underlying point cloud has points separated with many different distances. Finally, since from one diagram to another, the number of points N_{p_k} can differ, the entropy is normalized by $\log(N_{p_k})$. In this way, the persistent entropy $\text{PE}_{\max}(k)$ takes its values between 0 and 1 for all persistent diagrams.

B. Results Analysis

Monitoring the evolution of $\text{Bar}_{\max}(k)$ and $\text{PE}_{\text{norm}}(k)$, with k understood as time, yields to a real-time analysis of the patients' admissions time series. This monitoring is for example highlighted thanks to the Fig. 7-(a) and Fig. 7-(b). These figures have been computed from the DFA of Fig. 2. To clearly point out important facts, we have plotted two lines for reference on Fig. 2, Fig. 7-(a), and Fig. 7-(b). The first one indicates the beginning of an abnormal episode of admissions for bronchiolitis at the ED and the second one materializes the time when the maximum number of admissions has been reached. The analysis of Fig. 7-(a) shows different behaviors for the variability of the patient's flow. This is captured by the evolution of the indicator $\text{Bar}_{\max}(k)$ which evolves in steps. Each step is associated with an increase in variability of the admissions, whereas the length of the steps is related to the time spent in this variability regime. When the maximum variability is reached, the $\text{Bar}_{\max}(k)$ indicator is also at its maximum. When the time series of the admissions grows in a brutal way, we see a

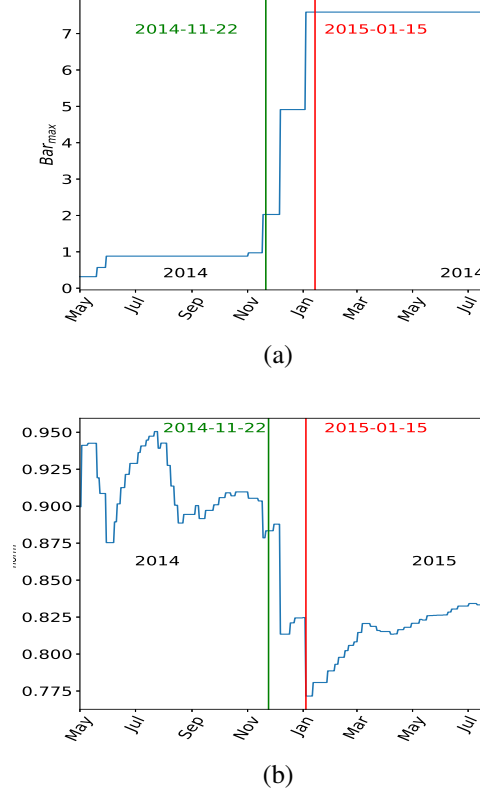


Fig. 7. Persistent homology results of the admissions for bronchiolitis at Saint Etienne ED between 2014-05-01 and 2015-08-31. (a) time evolution of $\text{Bar}_{\max}(k)$, (b) time evolution of $\text{PE}_{\text{norm}}(k)$. This computation has been carried out with $L = 30$ for the DFA in (2) and $N = 4$ for the persistent homology window size.

succession of close steps and the global shape of the $\text{Bar}_{\max}(k)$ indicator is very steep.

If we now turn to the $\text{PE}_{\text{norm}}(k)$ indicator, there is evidence that it reaches an extrema when a change of regime occurs. If there is a rise in variability followed by a mitigation, $\text{PE}_{\text{norm}}(k)$ reaches an extrema that corresponds to this change of level. We can observe from Fig. 7- (b) a minimum for the day 2015-01-15. After this day, $\text{PE}_{\text{norm}}(k)$ increases and we can say that the spread of the epidemic is decreasing. However, one must be careful with the interpretation of $\text{PE}_{\text{norm}}(k)$, indeed it may not correspond to the day with the greatest number of arrivals but to the day with the greatest variation in variability. Therefore, we assert that the peak of an epidemic is not characterized by a maximum number of cases, but by a maximum of variability. In other words, it is not when the maximum number of admissions is reached that the epidemic is really over, but when its variability begins to decrease. To stress the differences between the two indicators, observe that when the peak of the outbreak has been reached, $\text{Bar}_{\max}(k)$ remains at its maximum level, whereas $\text{PE}_{\text{norm}}(k)$ will be rising and thus enable the observer to track the end of the epidemic.

We observed the capacity of both the indicators to reveal useful information on the time series of the admissions with real-time considerations. They can be used obviously for early detect abnormal comings and give real-time information on the future decrease of the admissions. This is dealt with in the next section.

IV. DETECTORS

A. About the Detectors

As explained in the first section, the main parameter of the DFA method is the length L of the non-overlapping segments. In order to illustrate its influence on the $\text{Bar}_{\max}(k)$ and $\text{PE}_{\text{norm}}(k)$ indicators we have computed the persistent homology on the variability time series of Fig. 2 with $L = 5$ and $L = 50$. The results are given in Fig. 8-(a) and Fig. 8-(b) for respectively the $\text{Bar}_{\max}(k)$ and $\text{PE}_{\text{norm}}(k)$ indicators, with gold standard days plotted again. Obviously, for $L = 50$ the monitoring of the variability is inefficient. Only one big step is drawn in Fig. 8-(a), meaning that the method is unable to catch the variability in real-time. This inefficiency is less visible on Fig. 8-(b) but detection of the maximum of admissions is too optimistic and thus erroneous. Taking a small L value seems to be better. For $L = 5$ the $\text{PE}_{\text{norm}}(k)$ indicator is able to estimate the maximum of the admissions time-series with only few error days but the early detection of abnormal arrivals is still wrong due to over sensibility of the monitoring method. In the perspective of automatic detection for an implemented application, this parameter needs to be optimized and updated regularly so. This point is discussed next.

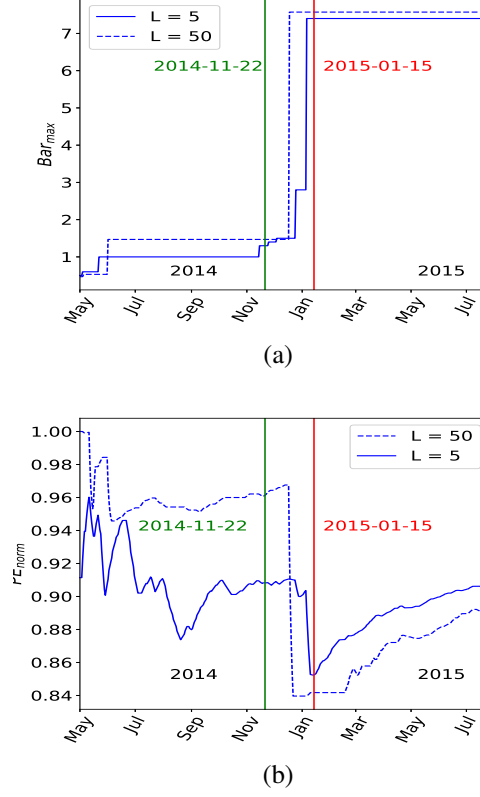


Fig. 8. Persistent homology results of the admissions for bronchiolitis at Saint Etienne ED between 2014-05-01 and 2015-08-31. (a) time evolution of $\text{Bar}_{\max}(k)$, (b) time evolution of $\text{PE}_{\text{norm}}(k)$. This computation has been carried out with $L = 50$ and $L = 5$ for the DFA in (2) and $N = 4$ for the persistent homology window size.

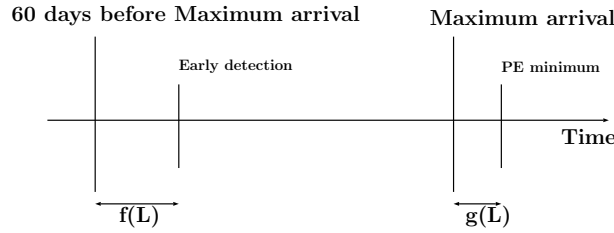


Fig. 9. Two criteria are minimized: $f(L)$, the early detection and $g(L)$ the estimation of maximum comings; L is the length of the DFA in (2).

B. Detector calibration

To pick the best L value, we propose to solve a bi-objective optimization problem, being that the parameter L influences both the early detection of the growing admissions and the precision of the persistent entropy. Two functions have to be minimized: the first one denoted $f(L)$ optimizes the early detection of increasing arrivals whereas the second one, named $g(L)$, optimizes the estimation of the maximum number of arrivals. With the aid of physicians and the analysis of past admissions data, it appeared that the early detection should be given nearly 60 days before the maximum arrivals. Fig. 9 illustrates this bi-objective problem. Several approaches have been proposed in the literature to solve it (see for example the discussion in [9]). Among those cited, we have retained the following method for its simplicity:

- For each year j used for calibration purposes, find the maximum number of arrivals at the ED and then compute $\alpha f_j(L) + \beta g_j(L)$ with $0 < \alpha, \beta < 1$ and $\alpha + \beta = 1$. The scalar values α and β allow to weigh each criteria for possibly giving more importance to one regarding the other.
- Sum up these values to obtain $\text{Obj}(L) = \sum_j (\alpha f_j(L) + \beta g_j(L))$.
- Repeat the operation for several L values, and pick the minimum value of $\text{Obj}(L)$.

In order now to find the best L value with this procedure, we used three consecutive years from 2013 to 2016. We have first estimated the three maximum of the admissions time series and then solved the bi-objective optimization with $\alpha = 0.6$ and $\beta = 0.4$, with the intention of giving slightly more importance to the detection of abnormal arrivals than to that of the maximum. As pointed out by Fig. 10-(a), the minimum value of the objective function $\text{Obj}(L)$ is reached for $L = 24$, which

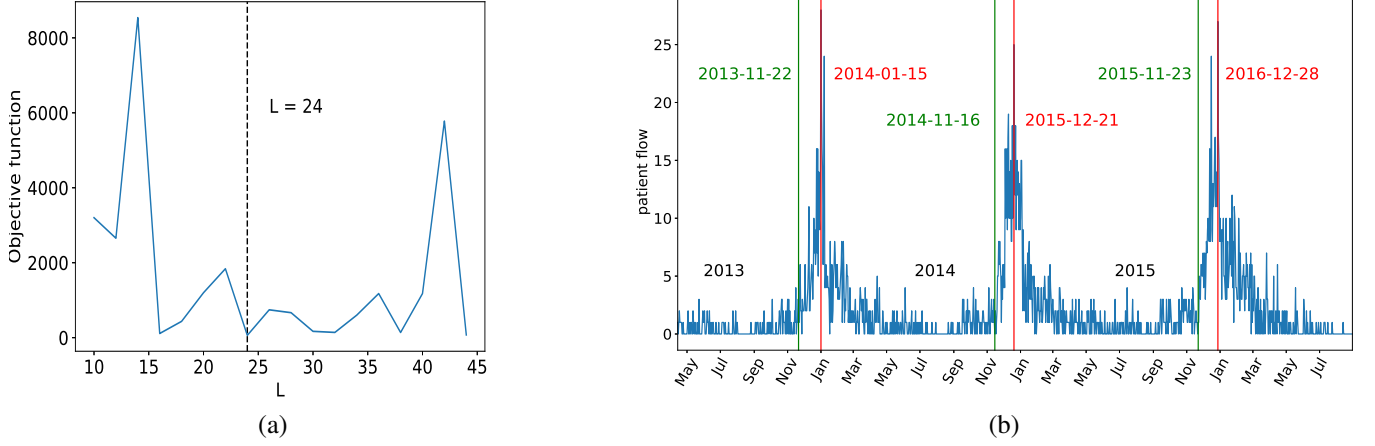


Fig. 10. (a) The objective function $Obj(L)$ obtained when calibrating the detector with three consecutive years from 2013-05-01 to 2016-08-31. (b) Dates of the early detection of abnormal comings and that of maximum comings with $L = 24$ for the three consecutive years.

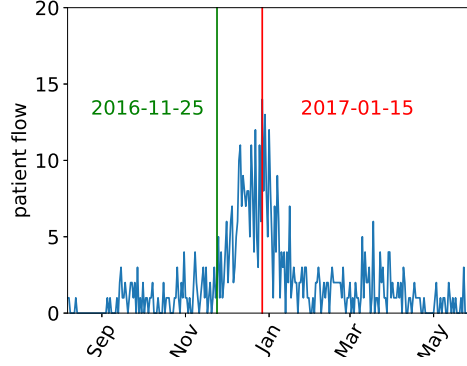


Fig. 11. Detection of abnormal comings (2016-11-25) and maximum comings estimation (2017-01-15) obtained with the calibrated system for year 2016-2017.

is consistent with the typical duration of the observed phenomenon. Owing to this parameter and still having a window size of $N = 5$, both detections have been added on the plot of Fig. 10-(b). We can clearly see that the detection of abnormal arrivals, the starting of the epidemic, is very early estimated. The detection of the maximum is rather well synchronized with the real value as well. The detector being now calibrated over the three consecutive and past years, we have tested its consistency on a new year. For the year 2016-2017, the results of the real-time detection have been reported on Fig. 11. At the date of 2016-11-25, abnormal arrivals at the paediatric ED were detected, and the maximum number of arrivals were estimated on the date of 2017-01-15, marking the decrease of the arrivals and consequently, the decrease of the associated epidemic.

C. Comparison with other methods

To assess the robustness, the accuracy and the early detection capacity of the $\text{Bar}_{\max}(k)$ indicator, we have benchmarked it with three other cutting-edge methods. The prominent methods commonly used for change detection are the CUSUM and the control (Shewhart) charts methods [20], [24], [2]. We have chosen to also compare the $\text{Bar}_{\max}(k)$ detector to the early detector given in [4] for which the detection is based on a local kurtosis measure. The difference between the indicator of [4] and the indicators proposed in this work is in the mathematical approach. In [4], a periodic model is first used to give an idea of the temporal evolving of the admissions. This rough model can serve as a prediction for the arrivals at the ED. Next, the difference between this model and the real arrivals is analyzed. The characterization of the induced distribution error allows to detect abnormal arrivals via the shape of the distribution. In [4], the arrivals are then directly used for the early detector whereas in the present work, the variability of the arrivals and the topological information given by its point cloud representation is used. The results are reported in Fig. 12 where the four methods were tested on the same epidemic event. In view of Fig. 12, the $\text{Bar}_{\max}(k)$ indicator performed as well as the CUSUM method for the detection of an early change in the ED admissions. However, many other detections are given by this control method. They are indicated by the red dots on the top plot of Fig. 12. It appears that the CUSUM method is quite simply too sensitive for the high variations of the admissions'

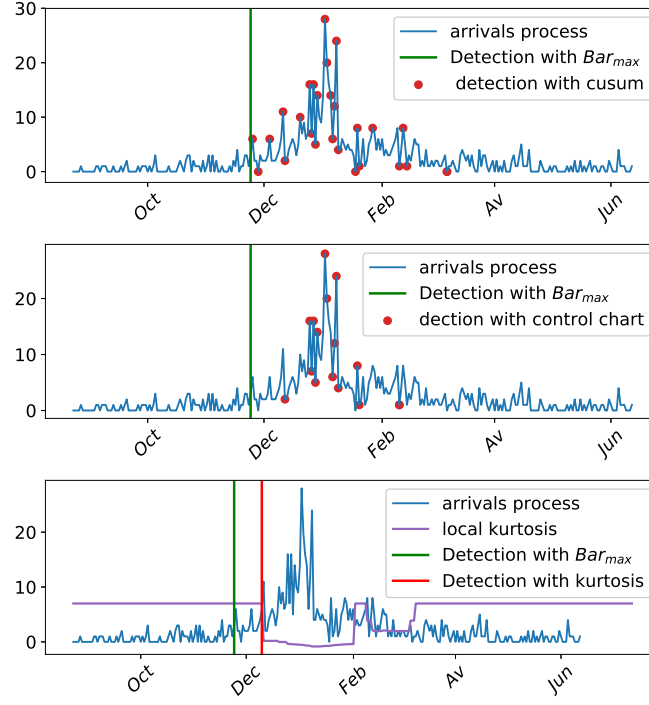


Fig. 12. Top: comparison between $\text{Bar}_{\max}(k)$ indicator and CUSUM change point detection. Middle: comparison between $\text{Bar}_{\max}(k)$ indicator and control (Shewhart) change point detection. Bottom: comparison between $\text{Bar}_{\max}(k)$ indicator and local kurtosis method.

numbers at the ED which renders it when all is said and done, useless for decisional aid. The control (or Shewhart) chart seems to be less sensitive to the high variations, but the first detection of abnormal arrivals is largely delayed and appears lately. In addition, there are still many detections obtained, yielding to the same inefficiency as that of the CUSUM method. The method which could be the most appropriate is the control chart-based method given by the local kurtosis analysis. Indeed, even if the detection comes later than that of the detection of the $\text{Bar}_{\max}(k)$ indicator, there is one and only one detection. We note furthermore that the local kurtosis method offers the possibility to dispose of an indicator for pointing out the end of an epidemic episode as well. It allows then to give some insights into the evolution of the admissions' outbreak as the couple $\text{Bar}_{\max}(k)$ and $\text{PE}_{\text{norm}}(k)$ indicators let do with the best performances.

V. CONCLUSION

To give to the Emergency Department leaders an early decisional aid to manage their departments is essential for financial gain and public health perspectives. We know that excessive admissions in the ED's occur when respiratory-like diseases spread off. The clinical symptoms of such diseases are sometimes well identified. This is particularly the case for the RSV which is associated with bronchiolitis symptoms by pediatricians. An understanding of the epidemic caused by the RSV can be then managed by focusing on the time series of the admissions for bronchiolitis in the paediatric ED's. In this work, we have proposed a new and original real-time method to automatically and (i) very early detect abnormal arrivals in the ED's, which gives the knowledge of the beginning of the associated epidemic and (ii) to precisely estimate the the maximum number of arrivals, which gives the knowledge of the decrease of the associated epidemic. We used for that the persistent homology technique on the variability time series obtained by Detrended Fluctuation Analysis (DFA) of the admissions time series for selected symptoms. The optimum choice for the parameter of the DFA is obtained by solving a bi-objective optimization problem. We finally specified that the method proposed in this work will be soon implemented in python language as a decisional aid for the paediatrics ED's of Grenoble and Saint Etienne, France, through the following web application: <http://predaflu.chu-st-etienne.fr>. We also point out that the method might have numerous other uses in clinics, especially when following quality of work, and even in analytics in certain laboratory tests such as those studying cardiac and nervous functions.

VI. ACKNOWLEDGMENT

We are sincerely grateful to the reviewers for their valuable comments which helped in improving the content of this paper.

REFERENCES

- [1] P.-A. Absil, R. Sepulchre, A. Bilge, and P. Gérard, “Nonlinear analysis of cardiac rhythm fluctuations using dfa method,” *Physica A: Statistical Mechanics and its Applications*, vol. 272, no. 1, pp. 235 – 244, 1999.
- [2] S. Aminikhanghahi and D. J. Cook, “A survey of methods for time series change point detection,” *Knowledge and Information Systems*, vol. 51, no. 2, pp. 339–367, May 2017.
- [3] G. Bouleux, “Oblique projection pre-processing and TLS application for diagnosing rotor bar defects by improving power spectrum estimation,” *Mech. Syst. Signal Process.*, vol. 41, no. 1, pp. 301–312, 2013.
- [4] G. Bouleux, E. Marcon, and O. Mory, “Early index for detection of pediatric emergency department crowding,” *IEEE journal of biomedical and health informatics*, vol. 19, no. 6, pp. 1929–1936, 2015.
- [5] F. T. Bourgeois, C. Valim, J. C. Wei, A. J. McAdam, and K. D. Mandl, “Influenza and Other Respiratory Virus-Related Emergency Department Visits Among Young Children,” *Pediatrics*, vol. 118, no. 1, pp. e1–e8, jul 2006.
- [6] D. S. Broomhead and G. P. King, “Extracting qualitative dynamics from experimental data,” *Phys. D Nonlinear Phenom.*, vol. 20, no. 2, pp. 217–236, 1986.
- [7] P. G. Cámara, A. J. Levine, and R. Rabadán, “Inference of Ancestral Recombination Graphs through Topological Data Analysis,” *PLOS Computational Biology*, vol. 12, no. 8, pp. e1005071, Aug. 2016.
- [8] G. Carlsson, T. Ishkhonov, V. de Silva, and A. Zomorodian, “On the Local Behavior of Spaces of Natural Images,” *International Journal of Computer Vision*, vol. 76, no. 1, pp. 1–12, Jan. 2008.
- [9] G. Chiandussi, M. Codegone, S. Ferrero, and F. E. Varesio, “Comparison of multi-objective optimization methodologies for engineering applications,” *Computers & Mathematics with Applications*, vol. 63, no. 5, pp. 912–942, Mar. 2012.
- [10] H. Chintakunta, T. Gentimis, R. Gonzalez-Diaz, M.-J. Jimenez, and H. Krim, “An entropy-based persistence barcode,” *Pattern Recognition*, vol. 48, no. 2, pp. 391–401, Feb. 2015.
- [11] Edelsbrunner, Letscher, and Zomorodian, “Topological Persistence and Simplification,” *Discrete & Computational Geometry*, vol. 28, no. 4, pp. 511–533, Nov. 2002.
- [12] H. Edelsbrunner and D. Morozov, *Persistent Homology: Theory and Practice*, 2013, Report Number: LBNL-6037E
- [13] A.K. Golińska, “Detrended fluctuation analysis (dfa) in biomedical signal processing: Selected examples”, *Studies in logic, grammar and rhetoric*, vol. 29, pp. 107–115, 01 2012.
- [14] N. R. Hoot and Epstein et al., “Forecasting emergency department crowding: an external, multicenter evaluation,” *Annals of Emergency Medicine*, vol. 54, no. 4, pp. 514–522.e19, Oct. 2009.
- [15] V. Isham, “Assessing the variability of stochastic epidemics,” *Mathematical Biosciences*, vol. 107, no. 2, pp. 209 – 224, 1991.
- [16] S. S. Jones, T. L. Allen, T. J. Flottesch, and S. J. Welch, “An independent evaluation of four quantitative emergency department crowding scales,” *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, vol. 13, no. 11, pp. 1204–1211, Nov. 2006.
- [17] K. L. Khatri and L. Tamil, “Early Detection of Peak Demand Days of Chronic Respiratory Diseases Emergency Department Visits Using Artificial Neural Networks,” *IEEE J. Biomed. Heal. Informatics*, pp. 1–1, apr 2017.
- [18] A. Kolker, “Process Modeling of Emergency Department Patient Flow: Effect of Patient Length of Stay on ED Diversion,” *Journal of Medical Systems*, vol. 32, no. 5, pp. 389–401, Oct. 2008.
- [19] M. Laskowski, R. D. McLeod, M. R. Friesen et al., “Models of Emergency Departments for Reducing Patient Waiting Times,” *PLOS ONE*, vol. 4, no. 7, pp. e6127, July 2009.
- [20] R.C. Lloyd, “Quality Health Care A Guide to Developing and Using Indicators”, *Jones and Bartlett publishers*, 2004.
- [21] L. J. Martin, C. Im, H. Dong, B. E. Lee, J. Talbot, D. P. Meurer, S. N. Mukhi, S. J. Drews, and Y. Yasui, “Influenza-like illness-related emergency department visits: Christmas and New Year holiday peaks and relationships with laboratory-confirmed respiratory virus detections, Edmonton, Alberta, 2004–2014,” *Influenza Other Respi. Viruses*, vol. 11, no. 1, pp. 33–40, jan 2017.
- [22] J. Perea and J. Harer, “Sliding Windows and Persistence: An Application of Topological Methods to Signal Analysis,” *arXiv:1307.6188 [math, stat]*, July 2013, arXiv: 1307.6188.
- [23] A. Port, I. Gheorghita, D. Guth, J. M. Clark, C. Liang, Shival Dasu, and Matilde Marcolli, “Persistent Topology of Syntax,” *Mathematics in Computer Science*, vol. 12, no. 1, pp. 33–50, Mar. 2018.
- [24] L. Provost, and S. Murray, “The Health Care Data Guide: Learning from Data for Improvement”, *Jossey-Bass publishers*, 2011.
- [25] M. Rucco, F. Castiglione, E. Merelli, and M. Pettini, “Characterisation of the Idiopathic Immune Network Through Persistent Entropy,” in *Proceedings of ECCS 2014*, Springer Proceedings in Complexity, pp. 117–128. Springer, Cham, 2016.
- [26] C. M. Topaz, L. Ziegelmeier, and T. Halverson, “Topological Data Analysis of Biological Aggregation Models,” *PLOS ONE*, vol. 10, no. 5, pp. e0126383, May 2015.
- [27] Y. Umeda, “Time Series Classification via Topological Data Analysis,” *Transactions of the Japanese Society for Artificial Intelligence*, vol. 32, no. 3, pp. D–G72_1–12, 2017.
- [28] F. H. Yap, P.-L. Ho, K.-F. Lam, P. K.S. Chan, Y.-H. Cheng, and J. S.M. Peiris, “Excess hospital admissions for pneumonia, chronic obstructive pulmonary disease, and heart failure during influenza seasons in Hong Kong,” *J. Med. Virol.*, vol. 73, no. 4, pp. 617–623, aug 2004.
- [29] F. H. Yap and P.-L. Ho, et al, “Excess hospital admissions for pneumonia, chronic obstructive pulmonary disease, and heart failure during influenza seasons in hong-kong,” *Journal of Medical Virology*, vol. 73, no. 4, pp. 617–623, Aug. 2004.
- [30] A. Zomorodian and G. Carlsson, “Computing Persistent Homology,” *Discrete Comput Geom*, vol. 33, no. 2, pp. 249–274, Feb. 2005.