METHODOLOGIES AND APPLICATION

Human-centric analysis and interpretation of time series: a perspective of granular computing

Witold Pedrycz · Wei Lu · Xiaodong Liu · Wei Wang · Lizhong Wang

Abstract In spite of the truly remarkable diversity of models of time series, there is still an evident need to develop constructs whose accuracy and interpretability are carefully identified and reconciled subsequently leading to highly interpretable (human-centric) constructs. While a great deal of research has been devoted to the design of nonlinear *numeric* models of time series (with an evident objective to achieve high accuracy of prediction), an issue of interpretability (transparency) of models of time series becomes an evident and ongoing challenge. The user-friendliness of models of time series comes with an ability of humans to perceive and process abstract constructs rather than dealing with plain numeric entities. In perception of time series, information granules (which are regarded as realizations of interpretable entities) play a pivotal role. This gives rise to a concept of granular models of time series or *granular* time series, in brief. This study revisits generic concepts of information granules and elaborates on a fundamental way of forming information granules (both sets—intervals as well as fuzzy sets) through applying a principle of justifiable granularity encountered in granular computing. Information granules are discussed with regard to the granulation of time series in a certain predefined representation space (viz. a feature space) and granulation carried out in time. The granular representation and description of time series is then presented. We elaborate on the fundamental hierarchically organized layers of processing supporting the development and interpretation of granular time series, namely (a) formation of granular descriptors used in their visualization, (b) construction of linguistic descriptors used afterwards in the generation of (c) linguistic description of time series. The layer of the linguistic prediction models of time series exploiting the linguistic descriptors is outlined as well. A number of examples are offered throughout the entire paper with intent to illustrate the main functionalities of the essential layers of the granular models of time series.

Keywords Information granularity · Granular time series · Principle of justifiable granularity · Linguistic description · Higher order granular models

W. Pedrycz (✉)
Department of Electrical and Computer Engineering,
University of Alberta, Edmonton, Canada
e-mail: wpedrycz@ualberta.ca

W. Pedrycz
Department of Electrical and Computer Engineering,
Faculty of Engineering, King Abdulaziz University,
Jeddah, Saudi Arabia

W. Pedrycz
Systems Research Institute, Polish Academy of Sciences,
Warsaw, Poland

W. Lu · X. Liu · W. Wang · L. Wang
School of Control Science and Engineering,
Dalian University of Technology, Dalian,
People's Republic of China

1 Introduction

Time series is a sequence of real-data, with each element in this sequence representing a value recorded at some time moment. With regard to processed data, time series is inherently associated with their large size, high dimensionality and a stream-like nature. Time series are ubiquitous in numerous domains. Recently, this data modality has attracted great attention in the data mining research community. In the plethora of currently available models of time series, the accuracy of the models has been a holy grail of the over-

all modeling. With the emergence of a more visible need for interpretable models that are easily comprehended by users and facilitate a more interactive style of modeling, arose an important need to develop models that are not only accurate but transparent (interpretable) Casillas et al. (2003) as well. Likewise, as a part of requirements, it is anticipated that the user should be at position to easily establish a sound trade-off between accuracy and interpretability. Similarly, effective mechanisms supporting a realization of relevance feedback become an important part of the overall research agenda.

Current research in time series mining, time series characterization, and prediction comes with a visible diversity and richness of the conceptual and algorithmic pursuits. In spite of their diversity, there is some evident and striking similarity; all of these approaches dwell upon the use and processing being realized at a numeric level.

The pursuits of perception, analysis and interpretation of time series, as realized by humans, are realized at a certain, usually problem-related, level of detail. Instead of single numeric entries—successive *numeric* readings of time series, developed are conceptual entities over time and feature space —information granules using which one discovers meaningful and interpretable relationships forming an overall description of time series. The granularity of information is an important facet being imperative to any offering of well-supported mechanisms of comprehension of the underlying temporal phenomenon. In all these pursuits, information granules manifest along the two main dimensions (as noted above). The first one is concerned with time granularity. Time series is split into a collection of time windows —temporal granules. One looks at time series in temporal windows of months, seasons, years. Time series is also perceived and quantified in terms of information granules being formed over the space of amplitude of the successive samples of the sequence; one arrives at sound and easily interpretable descriptors such as *low*, *medium*, *high* amplitude and alike. One can also form information granules over the space of changes of the time series. Combined with the temporal facets, the composite information granules arise as triples of entities, say *long* duration, positive *high* amplitude, *approximately* zero changes, etc. Once information granules are involved, long sequences of numbers forming time series are presented as far shorter, compact sequences of information granules-conceptual entities. As noted above, those are easier to understand and process. Let us note that while temporal granules are quite common (and they are of the same length), forming and processing composite information granules built over variable length time intervals call for detailed investigations.

Let us emphasize that information granules are sought as entities composed of elements being drawn together on a basis of similarity, functional closeness or spatial neighborhood. The quality of such granulation (abstraction) of data

is clearly related to the ability of this abstraction process to retain the essence of the original data (problem) while removing (hiding) all unnecessary details. With this regard, granularity of information (Bargiela and Pedrycz 2002, 2003, 2005, 2008, 2009; Apolloni et al. 2008; Srivastava et al. 1999; Ślęzak 2009; Pedrycz and Song 2011; Qian et al. 2011) plays a pivotal role and becomes of paramount importance, both from the conceptual as well as algorithmic perspective to the realization of granular models of time series. Subsequently, processing realized at the level of information granules gives rise to the discipline of granular computing (Bargiela and Pedrycz 2003). In granular computing, we encounter a broad spectrum of formal approaches realized in terms of fuzzy sets (Aznarte et al. 2010; Chen and Chen 2011; Kasabov and Song 2003; Lee et al. 2006; Pedrycz and Gomide 2007), rough sets (Pawlak 1991, 1985; Pawlak and Skowron 2007a,b), shadowed sets (Pedrycz 1998, 1999), probabilistic sets (Hirota and Pedrycz 1984), and others. Along with the conceptual setups, we also encounter a great deal of interesting and relevant ideas supporting processing of information granules. For instance, we can refer to the algorithmic perspective embracing fuzzy clustering (Bezdek 1981), rough clustering, and clustering being regarded as fundamental development frameworks in which information granules are constructed.

The key objective of this paper is to develop a general architecture of a system supporting human-centric analysis and interpretation of time series, both through their visualization as well as a linguistic description. In both cases, we carefully exploit concepts of information granules as being central to all faculties of description and interpretation of time series. In order to form an abstract, concise view at temporal data, the associated information granule is made justifiable, viz. strongly supported by experimental evidence, and semantically meaningful (of sufficient specificity, sound level of detail). We briefly discuss data sets used in Sect. 2. Section 3 recalls the underlying processing issues of time series. The granular framework of interpretation of time series is proposed in Sect. 4. A crux of a principle of justifiable granularity is outlined in Sect. 5. The principle gives rise to an information granulation algorithm, which takes a collection of temporal data and transforms them into a semantically sound information granule. It is shown how interval granules are built and how a collection of such nested granules are arranged into a fuzzy set. Section 6 is concerned with a formation of a collection of information granules. Composite information granules formed in the space of amplitude and changes of amplitude are studied. A suite of experimental studies comprising several real-world time series is discussed in Sect. 6. Conclusions are offered in Sect. 7.

A certain point has to be made with regard to the formalism of granular computing being used in this study. While most of the detailed investigations are carried out in the setting

of fuzzy sets (and in this way one could have contemplated the usage of the term *fuzzy* time series as a more suitable wording), which become instrumental in all the pursuits here, offering a conceptual framework and facilitating a realization of the sound algorithmic platform. It is emphasized that the introduced layered approach is equally suitable to deal with a variety of formalisms of information granules as well as the proposed constructs (such as e.g., the principle of justifiable granularity supporting a formation of a variety of information granules) and hence the notion of granular time series becomes well justified.

2 Experimental data sets—a summary

In this study, three publicly available time series coming from the finance domain are used to visualize a way how the granular constructs are designed and emphasize their essential aspects. More specifically, we consider here:

(1) The daily value of Taiwan Stock Exchange Weighted Stock Index (TAIEX) time series from 4th January 2000 to 29th December 2000, see Fig. 1a (sources: http://finance.

yahoo.com/q/hp?s=%5ETWII&a=00&b=1&c=2000&d=11&e=29&f=2000&g=d).

(2) The daily value of IBM common stock closing prices from 29th June 1959 to 30th June 1960, see Fig. 2a (sources: http://datamarket.com/data/set/2321/ibm-common-stocklosing-prices-daily-29th-jun%e-1959-to-30th-june-1960-n255#!display=table&ds=2321).

(3) The daily value of NASDAQ Composite Index (Nasdaq, for short) time series from 3th January 2000 to 31th December 2012, see Fig. 3a (sources: http://finance.yahoo.com/q/hp?s=%5EIXIC&a=00&b=1&c=2000&d=11&e=31&f=2012&g=d).

For each time series, the changes (first order differences) of the amplitude are smoothed with the use of the moving average filter (Brown 2004); the obtained results is presented in Figs. 1b, 2b and 3b using a dotted line. Besides, the temporal windows are also showed in Figs. 1, 2 and 3: the first two time series are split into 20 temporal windows of the same length and the third time series is split into 30 temporal windows of equal length. These values of the size of the windows are selected for illustrative purposes.
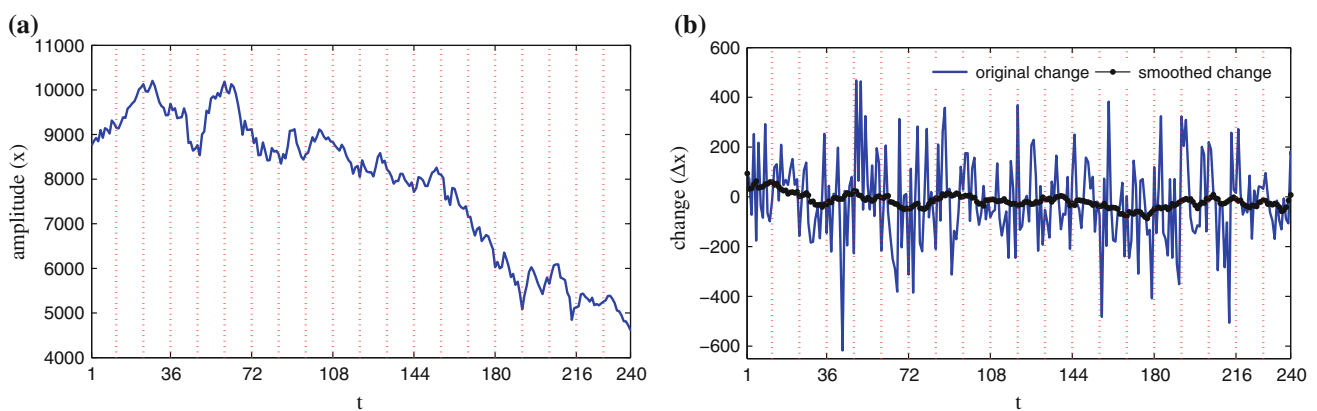

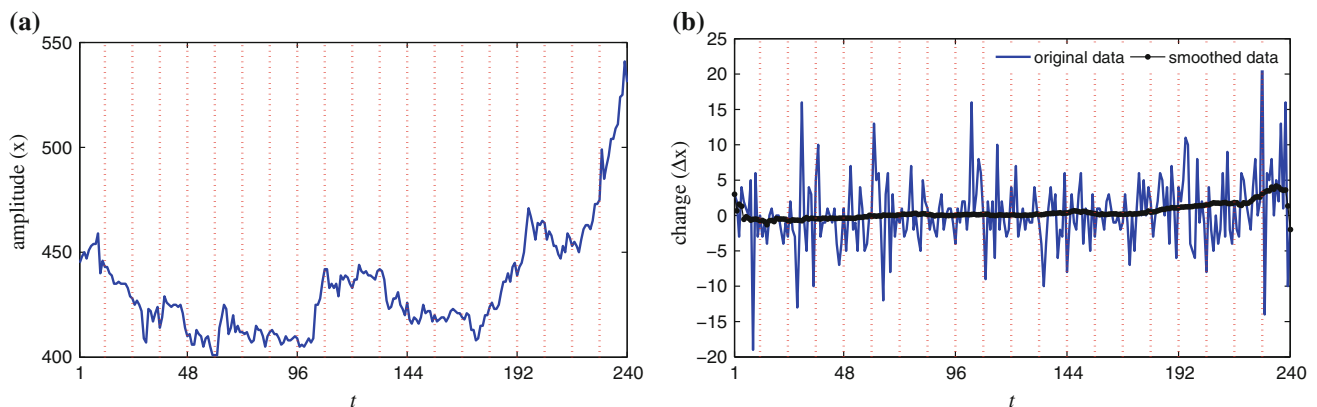**Fig. 1** TAIEX time series and the corresponding temporal windows

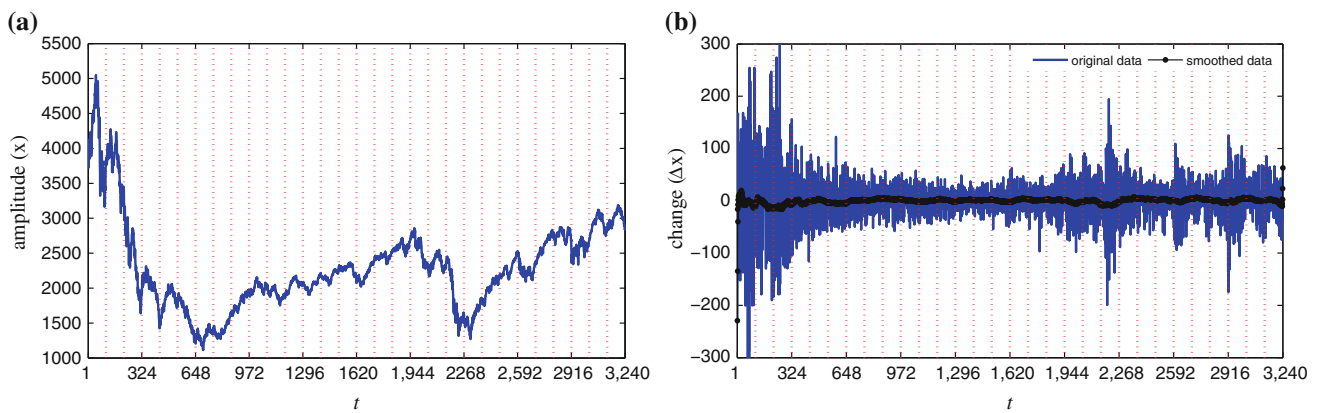**Fig. 2** IBM stock price time series and the corresponding temporal windows

Fig. 3 Nasdaq time series and the corresponding temporal windows

3 Time series-selected underlying processing issues

For the purpose of reduction of data and facilitating all mining algorithms, in most development schemes discussed is a segmentation step. Time series segmentation can be treated either as a preprocessing stage for a variety of ensuing data mining tasks or as an important stand-alone analysis process. An automatic partitioning of a time series into an optimal (or better to say, feasible) number of homogeneous temporal segments becomes an important problem. Quite often considered is a fixed-length segmentation process. Commonly encountered segmentation methods include the perceptually important points (PIP) (Fu et al. 2006; Jiang et al 2007), minimum message length (MML) (Oliver et al. (1998)) and minimum description length (MDL) segmentation (Fitzgibbon et al. 2002). A two-stage approach, in which one first uses a piecewise generalized likelihood ratio to carry out rough segmentation and then refines the results, has been proposed by Yager (2009). Keogh et al. (2001) adopted a piecewise linear representation method to segment time series. They focused on the problem of an-online segmentation of time series where a sliding window and bottom-up approach was proposed. Fuzzy clustering algorithms have showed a significant potential to address this category of problems. With this regard, Abonyi et al. (2005) developed a modified Gath-Geva algorithm to divide time-varying multivariate data into segments by using fuzzy sets to represent such temporal segments. Duncan and Bryant (1996) proposed dynamic programming to determine a total number of intervals within the data, the location of these intervals and the order of the model within each segment. Wang and Willett (2002) would be the segmentation problem as a tool for exploratory data analysis and data mining called the scale-sensitive gated experts (SSGE), which can partition a complex nonlinear regression surface into a set of simpler surfaces called "features". Fu (2011) presented a recent survey of mining time series.

4 A granular framework of interpretation of time series: a layered approach to the interpretation of time series

As noted so far, the notion of information granularity plays a pivotal role in all interpretation and analysis pursuits of time series. Our investigations of the description of time series being cast in the setting of granular computing are presented in a top-down fashion. We start with an overall view of the conceptual framework by stressing its key functionalities and a layered architecture and then move on with a detailed discussion by elaborating on the supported algorithmic aspects.

As commonly encountered in the investigations on time series, a starting point is a collection of numeric data (samples recorded in successive time moments) time series $\{x_1, x_2 \ldots x_N\}$. For the purpose of further analysis, we also consider the first-order dynamics of the time series by considering the sequences of differences (changes) observed there, namely $\{\Delta x_2, \Delta x_3 \ldots \Delta x_N\}$ where $\Delta x_i = x_i - x_{i-1}$. Quite often a smoothed version of these differences is sought. The space in which time series are discussed comes as a Cartesian product of the space of amplitude X and change of amplitude, namely $X \times \Delta X$.

The bird's-eye view of the overall architecture supporting processing realized at several layers which stresses the associated functionality is displayed in Fig. 4. Let us elaborate in detail on the successive layers at which consecutive phases of processing are positioned:

Formation of information granules The granules are formed over the Cartesian product of amplitude and change of amplitude, viz. the space $X \times \Delta X$. Here for each time slice T_i (note that the time variable is also subject to granulation as well; in addition to the same length time intervals, their size could be eventually optimized), we form an interval information granule (Cartesian product) $X_i \times \Delta X_i$. These information granules are constructed following the principle of justifiable granularity, as outlined in Sect. 5.1 and discussed further in Appendix A.
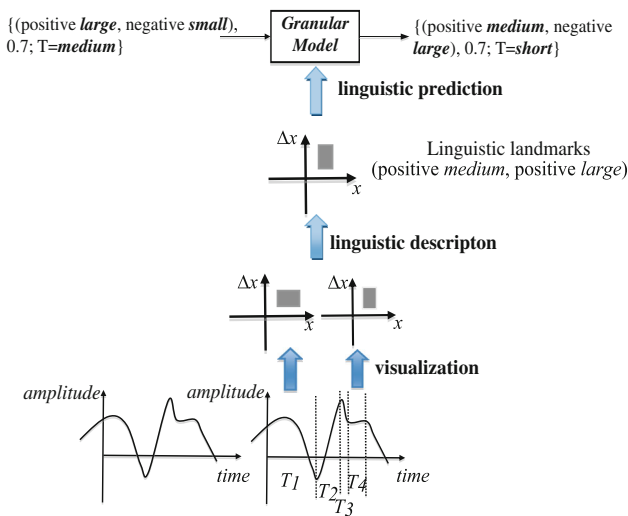
Fig. 4 Development of granular time series: visualization and linguistic interpretation. Emphasized is a layered hierarchical approach presented in this study

Visualization of information granules The results produced at the previous processing phase are visualized. In essence, for each time slice (segment) T_i, one can visualize a collection of Cartesian products of the information granules in $X \times \Delta X$ obtained in successive time slices, see Fig. 4.

Linguistic description of granular time series While the visualization of the granular time series could be quite appealing highlighting the main temporal tendencies observed in the time series, it is also worth arriving at the linguistic description of the time series, which is articulated in terms of some linguistic (granular) landmarks and ensuing levels of the best matching of these landmarks with the constructed information granules $X_i \times \Delta X_i$. Having a collection of the landmarks A_i, $i = 1, 2, \ldots, c$ where typically $c \ll p$ (where "p" stands for the number of time slices of the time series), the linguistic description comes as a string of the landmarks (for which the best matching has been accomplished) along with the corresponding matching levels $\mu_1, \mu_2, \ldots, \mu_p$. For instance, the linguistic description of the time series can read as the following sequence of descriptors:

{(*positive small*, *negative medium*})(0.7)}
{(*negative large*, *around zero*)(0.9)}...

In the above characterization, each granular landmark comes with its own semantics (e.g., *positive small*, *negative medium*, etc). The triples linguistically describe the amplitude and its change (expressed in terms of linguistic terms), and the associated matching level. To arrive at the linguistic description of this nature, two associated tasks are to be completed, namely (a) a construction of meaningful (semantically sound) granular landmarks, and (b) invocation of a matching mechanism, which returns a degree of matching

achieved. The first task calls for some mechanism of clustering of information granules while the second one is about utilizing one of the well-known matching measures encountered in fuzzy sets, say possibility or necessity measures.

Linguistic prediction models of time series The descriptions of the time series are useful vehicles to represent (describe) time series in a meaningful and easy to capture way. Per se, the descriptors are not models such as standard constructs reported in time series analysis. They, however, deliver all components, which could be put together to form granular predictive models. Denoting by $A_1, A_2, \ldots, A_c$ the linguistic landmarks developed at the previous phase of the overall scheme, a crux of the predictive model is to determine relationships present between the activation levels of the linguistic landmarks present for the current time granule T_k and those levels encountered in the next time granule T_{k+1}. The underlying form of the predictive mapping can be schematically expressed in the following way,

$$A_1(X_k), A_2(X_k), \ldots, A_c(X_k) \to A_1(X_{k+1}), \\ A_2(X_{k+1}), \ldots, A_c(X_{k+1}) \tag{1}$$

where $A_i(X_k)$ stands for a level of activation (matching) observed between A_i and the current information granule X_k. The operational form of the predictive model can be realized in the form of a fuzzy relational equation

$$A(X_{k+1}) = A(X_k) \circ R \tag{2}$$

with "∘" being a certain composition operator used in fuzzy sets (say, max-min or max-t composition where "t" stands for a certain t-norm) completed over information granules. $A_i(X_{k+1})$ is the activation levels of the linguistic landmarks A_i caused by the predicted information granule X_{k+1}. Overall, the vector $A(X_k)$ has "c" entries, namely $[A_1(X_k)A_2(X_k)\ldots A_c(X_k)]$. The granular relation R of dimensionality $c \times c$ captures the relationships between the corresponding entries of $A(X_k)$ and $A(X_{k+1})$. Note that as a result of using the model 2, X_{k+1} is not specified explicitly but rather through the degrees of activation of the linguistic landmarks. In other words, the predictive granular model returns a collection of quantified statements (the corresponding entries of $A(X_{k+1})$)

–predicted information granule is A_1 with degree of activation $\lambda_1 (= A_1(X_{k+1}))$
–predicted information granule is A_2 with degree of activation $\lambda_2 (= A_2(X_{k+1}))$ (3)
$\ldots$
–predicted information granule is A_c with degree of activation $\lambda_c (= A_c(X_{k+1}))$

which offer a certain intuitive view at the prediction outcome. Obviously, one can easily choose the dominant statement for which the highest level of matching has been reported, namely predicted information granule is A_i

with degree of activation $\lambda_i (= A_i(X_{k+1}))$ where $\lambda_i = arg\,max_{j=1,2,...,c} A_j(X_{k+1})$.

Note that there might be also another relational dependency as to the predicted length of the next time granule T_{k+1} (in case the sizes of temporal information granules vary across the description of the time series)

$$A_1(X_k), A_2(X_k), \ldots, A_c(X_k), T_k \rightarrow T_{k+1} \qquad (4)$$

which can be regarded as a certain granular relational equation describing dependencies among time granules T_k and T_{k+1}

$$T_{k+1} = A(X_k) \circ T_k \circ G \qquad (5)$$

In the above expression, G stands for a fuzzy relation, which links the input and output variables while "∘" denotes a sup-t composition operator. Alluding to expressions (1)–(5), we can view these predictive models as granular models of order-2 as they are described over the space of information granules rather than numeric entries.

5 Construction of information granules

The processing activities supporting the realization of information granules outlined in the previous section are now discussed in detail by presenting the successive phases of this process. We start with a brief elaboration on the principle of justifiable granularity, which forms a conceptual and algorithmic setup for the formation of information granules in the presence of some experimental evidence (numeric data).

5.1 The principle of justifiable granularity

We are concerned with a development of a single information granule (interval) Ω based on some numeric experimental evidence coming in a form of a collection of a one-dimensional (scalar) numeric data, $D = \{x_1, x_2 \ldots x_N\}$. The crux of the principle of justifiable granularity is to construct a meaningful information granule based on available experimental evidence (data) D. The require that such a construct has to adhere to the two intuitively compelling requirements:

(a) *Experimental evidence* The numeric evidence accumulated within the bounds of Ω has to be made as high as possible. By requesting this, we anticipate that the existence of the constructed information granule is well motivated (justified) as being reflective of the existing experimental data. The more data are included within the bounds of Ω, the better—in this way the set becomes more legitimate (justified).

(b) *Semantics of the construct* At the same time, the information granule should be made as specific as possible. This request implies that the resulting information granule comes with a well-defined semantics (meaning). In other words,

we would like to have Ω highly detailed, which makes the information granule semantically meaningful (sound). This implies that the smaller (more compact) the information granule (lower information granule) is, the better. This point of view is in agreement with our general perception of knowledge being articulated through constraints (information granules) specified in terms of statements such as "x is A", "y is B", etc. where A and B are constraints quantifying knowledge about the corresponding variables. Evidently, the piece of knowledge coming in the form "x is in $[1, 3]$" is more specific (semantically sound, more supportive of any further action, etc.) than another less detailed piece of knowledge where we know only that "x is in $[0, 12]$".

The further discusses for the principle of justifiable granularity is detailed in Appendix A.

5.2 Formation of fuzzy sets out of a family of interval information granules

For the data contained in the time interval T_i, we consider samples of time series $\{x_1, x_2 \ldots x_N\}$ and their differences $\{\Delta x_1, \Delta x_2 \ldots \Delta x_{N-1}\}$ out of which following the principle outlined in Sect. 5.1, we construct interval information granules X_i and ΔX_i and their Cartesian product. As these Cartesian products are indexed by the values of α, as a results of combining a family of the Cartesian products we form a fuzzy set-represented as a family of α-cuts $(X_i \times \Delta X_i)_\alpha$.

6 Clustering information granules—a formation of linguistic landmarks

Having a collection of information granules, they are clustered to form some meaningful representatives with well-expressed semantics. The fuzzy C-Means algorithm is one of viable alternatives to be used here. Schematically, we can illustrate the overall process as shown in Fig. 5.

The information granules, a collection of α-cuts formed within the time intervals are clustered to build meaningful semantically sound descriptors of the time series. There are two aspects of the formation of information, which deserve attention: (a) feature space in which clustering takes place. Given the nature of objects (information granules) to be clustered, which are undoubtedly more sophisticated than plain numeric entities, the granular data are expressed by a finite number of their α-cuts; for the pertinent details see Appendix B. The clustering is realized separately for each variable (namely amplitude and change of amplitude). This helps us order the obtained prototypes in a linear way and associate with them some semantics, say *low, medium, high*, etc. (b) The granular prototypes formed for the individual variables give rise to compound descriptors such as e.g., amplitude *high* and change of amplitude *close* to zero. This is accom-
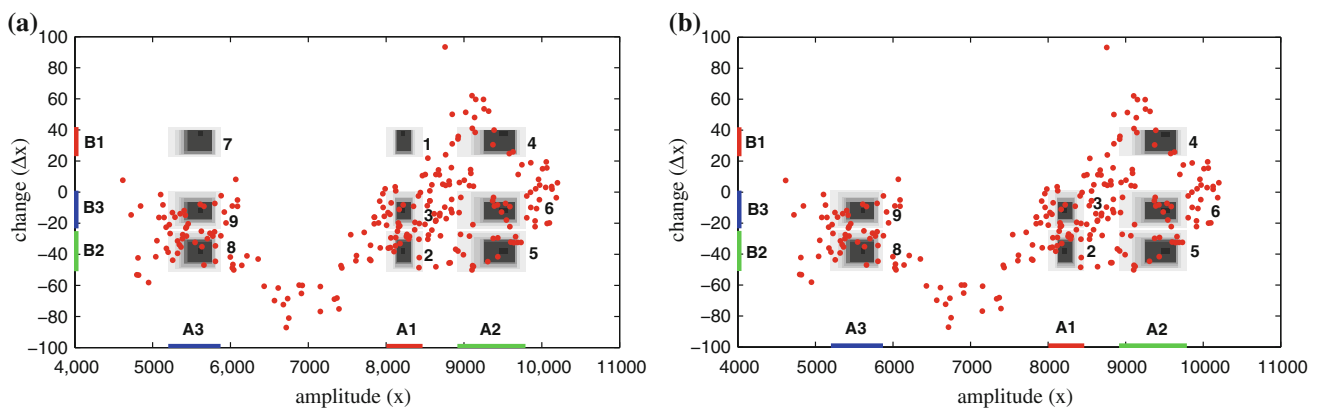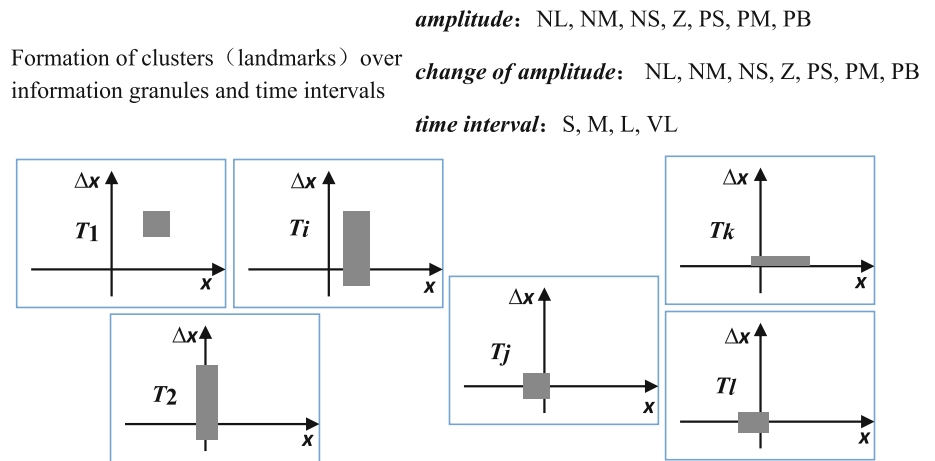
Fig. 5 From granular data to linguistic landmarks

Formation of clusters（landmarks）over information granules and time intervals

amplitude：NL, NM, NS, Z, PS, PM, PB

change of amplitude： NL, NM, NS, Z, PS, PM, PB

time interval： S, M, L, VL

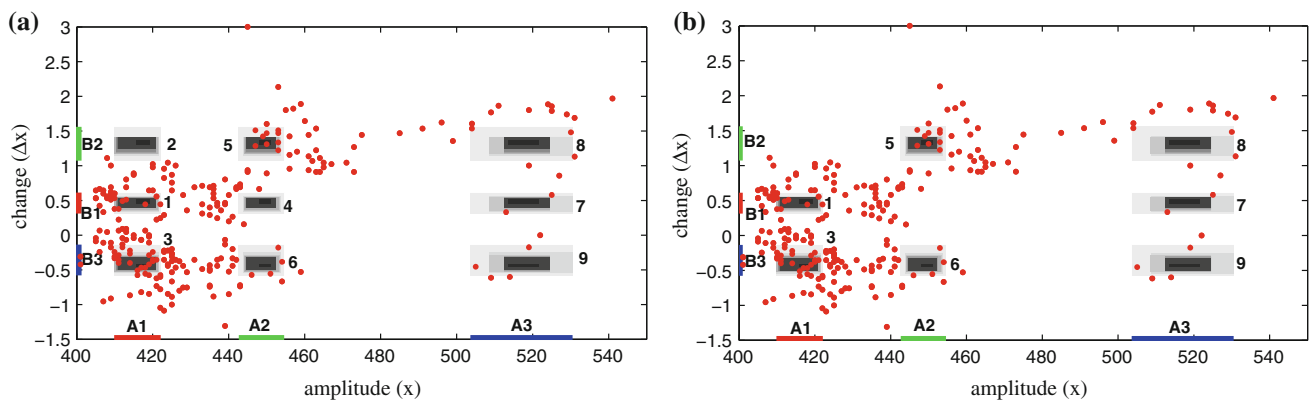**Fig. 6** Granular prototypes of TAIEX time series in the $x - \Delta x$ feature space

Fig. 7 Granular prototypes of IBM stock prices time series in the $x - \Delta x$ feature space

plished by forming Cartesian products of the granular prototypes, say $A_i \times B_j$ with A_i and B_j being the granular prototypes already constructed in the space of amplitude x and its change Δx. In general having r_1 and r_2 prototypes in the corresponding spaces, forming all prototypes constructed over the individual variables, we end up with $r_1 r_2$ Cartesian products, viz. compound information granules. Some of them could be *spurious* that is there is no experimental legitimacy behind them (a very few data associated with the cluster). Thus once the Cartesian product has been formed, its legitimacy vis-à-vis experimental data needs to be checked and quantified. The quantification is realized in the form of a weighted counting of the number of data "covered by granular prototype", for details to see Appendix B.
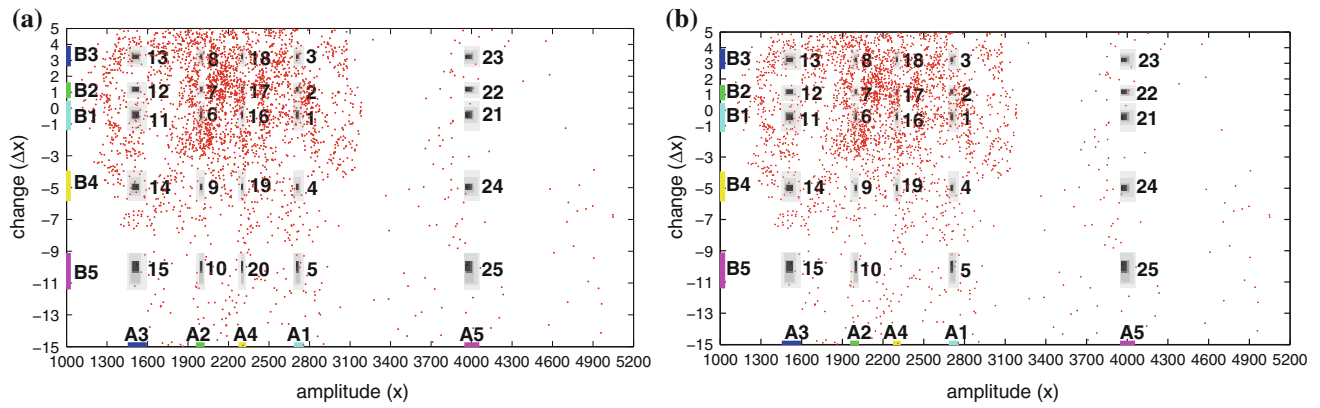
Fig. 8 Granular prototypes of Nasdaq time series in the $x - \Delta x$ feature space

Table 1 TAIEX time series: linguistic terms associated with granular prototypes

Granular label (amplitude)	Linguistic terms	Granular label (change of amplitude)	Linguistic terms
A1	Positive medium (PM)	B1	Positive medium (PM)
A2	Positive large (PL)	B2	Negative small (NS)
A3	Positive small (PS)	B3	Negative medium (NM)

Table 2 TAIEX time series: linguistic description of granular prototypes

Prototype no.	Prototype expressed in Cartesian product of x and Δx	Linguistic description of granular prototype	Legitimacy level μ
6	A2 × B3	PL × NM	27.7
5	A2 × B2	PL × NS	20.1
9	A3 × B3	PS × NM	17.0
4	A2 × B1	PL × PM	14.1
8	A3 × B2	PS × NS	14.0
3	A1 × B3	PM × NM	11.1
2	A1 × B2	PM × NS	10.4
1	A1 × B1	PM × PM	0.0
7	A3 × B1	PS × PM	0.0

Table 3 IBM stock prices time series: linguistic terms associated with granular prototypes

Granular label (amplitude)	Linguistic terms	Granular label (change of amplitude)	Linguistic terms
A1	Positive small (PS)	B1	Positive small (PS)
A2	Positive medium (PM)	B2	Positive medium (PM)
A3	Positive large (PL)	B3	Negative medium (NM)

Proceeding with the collection of the experimental data, the resulting granular prototypes by taking Cartesian product are displayed respectively in Figs. 6a, 7a and 8a for this three time series, where the number of temporal windows is 20 for the first two time series (see Figs. 1, 2) and 30 for the last time series (see Fig. 3). Besides, for the first two time series the number of clusters is set to 3 both for the amplitude and the change of amplitude while the last time series the number of clusters is set to 5 both for the amplitude and the change of amplitude. For each time series, linguistic terms

associated with granular prototypes (refer to Figs. 6, 7 and 8) are respectively reported in Tables 1, 3 and 5.

Along with the visualization of the granular prototypes, we provide with their characterization, viz. a linguistic description of each of them and experimental legitimacy level μ is reported in Tables 2, 4 and 6. Observe that there is a varying level of legitimacy of the prototypes and some of them could be completely eliminated as not being reflective of any data i.e. the spurious granular prototypes are removed when its value of legitimacy level is zero or close

Table 4 IBM stock prices time series: linguistic description of granular prototypes

Prototype no.	Prototype expressed in Cartesian product of x and Δx	Linguistic description of granular prototype	Legitimacy level μ
3	A1 × B3	PS × NS	42.7
1	A1 × B1	PS × PS	17.4
5	A2 × B2	PM × PM	12.5
6	A2 × B3	PM × NS	0.4
7	A3 × B1	PL × PS	0.2
8	A3 × B2	PL × PM	0.2
9	A3 × B3	PL × NS	0.2
2	A1 × B2	PS × PM	0
4	A2 × B1	PM × PS	0

Table 5 Nasdaq time series: linguistic terms associated with granular prototypes

Granular label (amplitude)	Linguistic terms	Granular label (change of amplitude)	Linguistic terms
A1	Positive large (PL)	B1	Zero
A2	Positive medium (PM)	B2	Positive small (PS)
A3	Positive small (PS)	B3	Positive medium (PM)
A4	Positive medium$^+$ (PM$^+$)	B4	Negative small (NS)
A5	Positive large$^+$ (PL$^+$)	B5	Negative medium (NM)

Table 6 Nasdaq time series: linguistic description of granular prototypes

Prototype no.	Prototype expressed in Cartesian product of x and Δx	Linguistic description of granular prototype	Legitimacy level μ
17	A4 × B2	PM$^+$ × PS	21.1
6	A2 × B1	PM × Z	16.2
13	A3 × B3	PS × PM	12.3
1	A1 × B1	PL × Z	10.1
7	A2 × B2	PM × PS	9.4
2	A1 × B2	PL × PS	8.4
11	A3 × B1	PS × Z	7.1
16	A4 × B1	PM × Z	6.9
25	A5 × B5	PL$^+$ × NM	4.5
4	A1 × B4	PL × NS	3.9
8	A2 × B3	PM × PM	3.9
3	A1 × B3	PL × PM	3.7
14	A3 × B4	PS × NS	3.3
18	A4 × B3	PM × PM	3.1
5	A1 × B5	PL × NM	1
19	A4 × B4	PM × NS	0.7
22	A5 × B2	PL$^+$ × PS	0.6
12	A3 × B2	PS × PS	0.5
15	A3 × B5	PS × NM	0.5
10	A2 × B5	PM × NM	0.4
24	A5 × B4	PL × NS	0.3
9	A2 × B4	PM × NS	0.2
21	A5 × B1	PL$^+$ × Z	0.1
23	A5 × B3	PL$^+$ × PM	0.1
20	A4 × B5	PM × NM	0

Table 7 TAIEX time series: description in terms of granular prototypes (in brackets shown are the prototypes with the second highest level of matching)

Temporal window ♯	Prototypes	Linguistic descriptor	Level of matching	Temporal window ♯	Prototypes	Linguistic descriptor	Level of matching
1	$A_2 \times B_1$	$PL \times PM$	0.37	11	$A_1 \times B_3$	$PM \times NM$	0.51
	$(A_2 \times B_3)$	$(PL \times NM)$	(0.25)		$(A_1 \times B_2)$	$(PM \times NS)$	(0.45)
2	$A_2 \times B_1$	$PL \times PM$	0.41	12	$A_1 \times B_3$	$PM \times NM$	0.51
	$(A_2 \times B_3)$	$(PL \times NM)$	(0.30)		$(A_1 \times B_2)$	$(PM \times NS)$	(0.42)
3	$A_2 \times B_3$	$PL \times NM$	0.36	13	$A_1 \times B_3$	$PM \times NM$	0.49
	$(A_2 \times B_2)$	$(PL \times NS)$	(0.36)		$(A_1 \times B_2)$	$(PM \times NS)$	(0.46)
4	$A_2 \times B_3$	$PL \times NM$	0.27	14	$A_1 \times B_2$	$PM \times NS$	0.39
	$(A_2 \times B_1)$	$(PL \times PM)$	(0.29)		$(A_1 \times B_3)$	$(PM \times NM)$	(0.36)
5	$A_2 \times B_3$	$PL \times NS$	0.37	15	$A_3 \times B_2$	$PS \times NS$	0.29
	$(A_2 \times B_2)$	$(PL \times NS)$	(0.29)		$(A_3 \times B_3)$	$(PS \times NM)$	(0.27)
6	$A_2 \times B_2$	$PL \times NS$	0.37	16	$A_3 \times B_2$	$PS \times NS$	0.51
	$(A_2 \times B_3)$	$(PL \times NM)$	(0.36)		$(A_3 \times B_3)$	$(PS \times NM)$	(0.44)
7	$A_1 \times B_2$	$PM \times NS$	0.42	17	$A_3 \times B_3$	$PS \times NM$	0.51
	$(A_1 \times B_3)$	$(PM \times NM)$	(0.38)		$(A_3 \times B_2)$	$(PS \times NS)$	(0.46)
8	$A_1 \times B_3$	$PM \times NM$	0.36	18	$A_3 \times B_3$	$PS \times NM$	0.51
	$(A_1 \times B_2)$	$(PM \times NS)$	(0.31)		$(A_3 \times B_2)$	$(PS \times NS)$	(0.46)
9	$A_2 \times B_3$	$PL \times NM$	0.25	19	$A_3 \times B_3$	$PS \times NM$	0.48
	$(A_2 \times B_2)$	$(PL \times NS)$	(0.23)		$(A_3 \times B_2)$	$(PS \times NS)$	(0.48)
10	$A_1 \times B_3$	$PM \times NM$	0.43	20	$A_3 \times B_2$	$PS \times NS$	0.50
	$(A_1 \times B_2)$	$(PM \times NS)$	(0.42)		$(A_3 \times B_3)$	$(PS \times NM)$	(0.43)

Table 8 IBM stock prices time series: description n terms of granular prototypes (in brackets shown are the prototypes with the second highest level of matching)

Temporal window ♯	Prototypes	Linguistic descriptor	Level of matching	Temporal window ♯	Prototypes	Linguistic descriptor	Level of matching
1	$A_2 \times B_3$	$PM \times NS$	0.84	11	$A_2 \times B_3$	$PM \times NS$	0.48
	$(A_2 \times B_2)$	$(PM \times PM)$	(0.15)		$(A_2 \times B_2)$	$(PM \times PM)$	(0.36)
2	$A_2 \times B_3$	$PM \times NS$	0.42	12	$A_1 \times B_3$	$PS \times NS$	0.49
	$(A_2 \times B_2)$	$(PM \times PM)$	(0.23)		$(A_1 \times B_1)$	$(PS \times PS)$	(0.35)
3	$A_1 \times B_3$	$PS \times NS$	0.53	13	$A_1 \times B_3$	$PS \times NS$	0.61
	$(A_1 \times B_1)$	$(PS \times PS)$	(0.39)		$(A_1 \times B_1)$	$(PS \times PS)$	(0.33)
4	$A_1 \times B_3$	$PS \times NS$	0.66	14	$A_1 \times B_1$	$PS \times PS$	0.59
	$(A_1 \times B_1)$	$(PS \times PS)$	(0.28)		$(A_1 \times B_3)$	$(PS \times NS)$	(0.39)
5	$A_1 \times B_3$	$PS \times NS$	0.54	15	$A_1 \times B_1$	$PS \times PS$	0.63
	$(A_1 \times B_1)$	$(PS \times PS)$	(0.38)		$(A_1 \times B_3)$	$(PS \times NS)$	(0.32)
6	$A_1 \times B_3$	$PS \times NS$	0.56	16	$A_1 \times B_1$	$PS \times PS$	0.38
	$(A_1 \times B_1)$	$(PS \times PS)$	(0.36)		$(A_1 \times B_3)$	$(PS \times NS)$	(0.24)
7	$A_1 \times B_3$	$PS \times NS$	0.47	17	$A_2 \times B_2$	$PM \times PM$	0.46
	$(A_1 \times B_1)$	$(PS \times PS)$	(0.46)		$(A_2 \times B_3)$	$(PM \times NS)$	(0.38)
8	$A_1 \times B_1$	$PS \times PS$	0.66	18	$A_2 \times B_2$	$PM \times PM$	0.58
	$(A_1 \times B_3)$	$(PS \times NS)$	(0.28)		$(A_2 \times B_3)$	$(PM \times NS)$	(0.35)
9	$A_1 \times B_1$	$PS \times PS$	0.51	19	$A_2 \times B_2$	$PM \times PM$	0.55
	$(A_1 \times B_3)$	$(PS \times NS)$	(0.40)		$(A_2 \times B_3)$	$(PM \times NS)$	(0.34)
10	$A_2 \times B_3$	$PM \times NS$	0.31	20	$A_3 \times B_2$	$PL \times PM$	0.61
	$(A_2 \times B_2)$	$(PM \times PM)$	(0.30)		$(A_3 \times B_1)$	$(PL \times PS)$	(0.21)

Table 9 Nasdaq time series: description in terms of granular prototypes (in brackets shown are the prototypes with the second highest level of matching)

Temporal window ♯	Prototypes	Linguistic descriptor	Level of matching	Temporal window ♯	Prototypes	Linguistic descriptor	Level of matching
1	$A_5 \times B_4$	$PL^+ \times NS$	0.23	16	$A_4 \times B_3$	$PM^+ \times PM$	0.13
	$(A_5 \times B_5)$	$(PL^+ \times NM)$	(0.21)		$(A_4 \times B_2)$	$(PM^+ \times PS)$	(0.13)
2	$A_5 \times B_4$	$PL^+ \times NS$	0.21	17	$A_4 \times B_2$	$PM^+ \times PS$	0.17
	$(A_5 \times B_5)$	$(PL^+ \times NM)$	(0.20)		$(A_4 \times B_1)$	$(PM^+ \times Zero)$	(0.16)
3	$A_1 \times B_5$	$PL \times NM$	0.15	18	$A_1 \times B_2$	$PL \times PS$	0.17
	$(A_1 \times B_4)$	$(PL \times NM)$	(0.14)		$(A_1 \times B_3)$	$(PL \times PM)$	(0.17)
4	$A_2 \times B_4$	$PM \times NS$	0.23	19	$A_1 \times B_4$	$PL \times NS$	0.22
	$(A_2 \times B_1)$	$(PM \times Zero)$	(0.20)		$(A_1 \times B_1)$	$(PL \times Zero)$	(0.19)
5	$A_2 \times B_1$	$PM \times Zero$	0.16	20	$A_4 \times B_1$	$PM \times Zero$	0.25
	$(A_2 \times B_2)$	$(PM \times PS)$	(0.16)		$(A_4 \times B_2)$	$(PM \times PS)$	(0.24)
6	$A_3 \times B_4$	$PS \times NS$	0.17	21	$A_3 \times B_5$	$PS \times NM$	0.16
	$(A_3 \times B_1)$	$(PS \times Zero)$	(0.16)		$(A_3 \times B_4)$	$(PS \times NS)$	(0.16)
7	$A_3 \times B_2$	$PS \times PS$	0.18	22	$A_3 \times B_3$	$PS \times PM$	0.25
	$(A_3 \times B_1)$	$(PS \times Zero)$	(0.18)		$(A_3 \times B_2)$	$(PS \times PS)$	(0.21)
8	$A_3 \times B_3$	$PS \times PM$	0.19	23	$A_2 \times B_3$	$PM \times PM$	0.32
	$(A_3 \times B_2)$	$(PS \times PS)$	(0.19)		$(A_2 \times B_2)$	$(PM \times PS)$	(0.23)
9	$A_2 \times B_3$	$PM \times PM$	0.13	24	$A_4 \times B_2$	$PM^+ \times PS$	0.28
	$(A_2 \times B_1)$	$(PM \times Zero)$	(0.13)		$(A_4 \times B_3)$	$(PM^+ \times PM)$	(0.28)
10	$A_2 \times B_1$	$PM \times Zero$	0.38	25	$A_4 \times B_1$	$PM^+ \times Zero$	0.25
	$(A_2 \times B_2)$	$(PM \times PS)$	(0.31)		$(A_4 \times B_2)$	$(PM^+ \times PS)$	(0.23)
11	$A_2 \times B_1$	$PM \times Zero$	0.20	26	$A_1 \times B_3$	$PL \times PM$	0.22
	$(A_2 \times B_1)$	$(PM \times Zero)$	(0.19)		$(A_1 \times B_2)$	$(PL \times PS)$	(0.20)
12	$A_2 \times B_2$	$PM \times PS$	0.20	27	$A_1 \times B_1$	$PL \times Zero$	0.23
	$(A_2 \times B_1)$	$(PM \times Zero)$	(0.19)		$(A_1 \times B_2)$	$(PL \times PS)$	(0.21)
13	$A_2 \times B_2$	$PM \times PS$	0.22	28	$A_1 \times B_2$	$PL \times PS$	0.17
	$(A_2 \times B_3)$	$(PM \times PM)$	(0.20)		$(A_1 \times B_3)$	$(PL \times PM)$	(0.17)
14	$A_4 \times B_2$	$PM^+ \times PS$	0.13	29	$A_1 \times B_1$	$PL \times Zero$	0.16
	$(A_4 \times B_1)$	$(PM^+ \times Zero)$	(0.12)		$(A_1 \times B_2)$	$(PL \times PS)$	(0.16)
15	$A_4 \times B_1$	$PM^+ \times Zero$	0.53	30	$A_1 \times B_2$	$PL \times PS$	0.15
	$(A_4 \times B_2)$	$(PM^+ \times PS)$	(0.27)		$(A_1 \times B_1)$	$(PL \times Zero)$	(0.21)

to zero, results of which are represented in Figs. 6b, 7b and 8b.

7 Matching information granules and a realization of linguistic description of time series

For the linguistic description of granular time series, it becomes essential to match the fuzzy relations $X_i \times \Delta X_i$ (their α-cuts) with the corresponding prototypes obtained during fuzzy clustering. The essence is to complete matching of $X_i \times \Delta X_i$ and the prototype $V_I \times \Delta V_I$ by computing the possibility measure $Poss(X_i \times \Delta X_i, V_I \times \Delta V_I)$. These calculations are straightforward as we find a maximal value of α for which the corresponding α-cuts of $X_i \times \Delta X_i$

and $V_I \times \Delta V_I$ overlap. We repeat the calculations for all $I = 1, 2 \ldots c$ and choose the index of the prototype λ_{I0} with the highest possibility value. Denote this value by λ_{I0}. The process is carried out for successive information granules $X_i \times \Delta X_i$ thus giving rise to a series of linguistic descriptors along with their levels of matching,

$$(V_{I1} \times \Delta V_{I1}, \lambda_1)(V_{I2} \times \Delta V_{I2}, \lambda_2) \ldots (V_{Ip} \times \Delta V_{Ip}, \lambda_p)$$

$$(6)$$

The details about the realization of the matching process are covered in Appendix C.

Proceeding with the time series discussed so far, in Tabels 7, 8 and 9, we report linguistic descriptions of the time series as a sequence of the linguistic terms, which match the segments of the data to the highest extent. Furthermore we report

those linguistic terms whose matching is the second the highest to gain a better sense as to the discriminatory capabilities of the linguistic terms used in the description of the time series.

8 Conclusions

In this study, we have proposed the concept of granular models of time series (granular time series, for short) and elaborated on the ensuing descriptions of the temporal data carried out at the level of information granules. Two conceptual levels of interpretation were clearly identified. It is shown how they endow the model with interpretation capabilities and facilitate interaction with the user. The first one, based on justifiable information granules is oriented towards supporting a graphic vehicle visualizing the nature of time series through a sequence of information granules reported in successive time windows.

In all endeavors, it has to be stressed that information granularity and information granules play a pivotal role by casting all temporal data in a setting, which is easily comprehended by users. The level of detail can be easily adjusted by specifying the number of information granules. The study is focused on the description of time series and the granular models aimed at prediction tasks are only highlighted as they occur at the highest level of the hierarchical structure in Fig. 4. The use of the temporal information granules of variable width (specificity) can bring a unique feature of predictive models when not only the results of prediction are generated but those are supplied with the time horizon over which they are valid.

While the study has offered some conceptual and methodological framework, it is apparent that in the overall discussion information granules are formalized as intervals and fuzzy sets. The principle of justifiable granularity comes with sufficient level of generality where one is able to engage other formalisms of information granularity thus making the overall framework sufficiently general. Furthermore the representation of time series, viz. the feature space could be very diversified. The study used on the simple temporal representation of the time series (information granules built for information granules in the space of amplitude and change of amplitude) however information granules can be equally well formed in some other feature spaces (such as e.g., those resulting from a spectral representation of time series).

Appendix A

In what follows, we briefly highlight the concept of justifiable information granularity concentrating on the computational aspects in case of interval and fuzzy set-based formalisms of information granules.

It is apparent that two requirements discussed in Sect. 5.1 are in conflict: the increase of the criterion of experimental evidence (justifiable) comes with a deterioration of the specificity of the information granule (specific). As usual, we are interested in forming a sound compromise. The requirement of experimental evidence is quantified by counting the number of data falling within the bounds of Ω. More generally, we may consider an increasing function of this cardinality, say $f_1(card\{x_k | x_k \in \Omega\})$, where f_1 is an increasing function of its argument. The simplest example is a function of the form $f_1(u) = u$. The specificity of the information granule Ω associated with its well-defined semantics (meaning) can be articulated in terms of the length of the interval. In case of $\Omega = [a, b]$, any continuous non-increasing function f_2 of the length of this interval, say $f_2(m(\Omega))$, where $m(\Omega) = |b - a|$ can serve as a sound indicator of the specificity of the information granule. The shorter the interval (the higher the value of $f_2(m(\Omega))$), the better the satisfaction of the specificity requirement. It is evident that two requirements identified above are in conflict: the increase in the values of the criterion of experimental evidence (justifiable) comes at an expense of a deterioration of the specificity of the information granule (specific). As usual, we are interested in forming a sound compromise between these requirements.

Having these two criteria in mind, let us proceed with the detailed formation of the interval information granule. We start with a numeric representative of the set of data D around which the information granule Ω is created. A sound numeric representative of the data is its median, $med(D)$. Recall that the median is a robust estimator of the sample and typically comes as one of the elements of D. Once the median has been determined, Ω (the interval $[a, b]$) is formed by specifying its lower and upper bounds, denoted here by "a" and "b", respectively. The determination of these bounds is realized independently. Let us concentrate on the optimization of the upper bound (b). The optimization of the lower bound (a) are carried out in an analogous fashion. For this part of the interval, the length of Ω or its non-increasing function, as noted above. In the calculations of the cardinality of the information granule, we take into consideration the elements of D positioned to the right from the median, that is $card\{x_k \in D | med(D) \leq x_k \leq b\}$. As the requirements of experimental evidence (justifiable granularity) and specificity (semantics) are in conflict, we resort ourselves to a maximization of the composite index in which a product of the two expressions governing the requirements. This is done

independently for the lower and upper bound of the interval, that is:

$$V(b) = card\{x_k \in D | med(D) \le x_k \le b\}$$
$$* f_2(|med(D) - b|) \tag{7}$$

$$V(a) = card\{x_k \in D | a \le_k \le med(D)\}$$
$$* f_2(|med(D) - a|) \tag{8}$$

We obtain the optimal upper bound b_{opt}, by maximizing the value of $V(b)$, namely $V(b_{opt}) = max_{b>med(D)} V(b)$. Among numerous possible design alternatives regarding functions f_1 and f_2, we consider the following alternatives

$$f_1(u) = u \tag{9}$$
$$f_2(u) = \exp(-\alpha u) \tag{10}$$

where α is a positive parameter delivering some flexibility when optimizing the information granule Ω. Under these assumptions, the optimization problem takes on the following form

$$V(b) = card\{x_k \in D | med(D) \le x_k \le b\}$$
$$* \exp(-\alpha|med(D) - b|) \tag{11}$$

Its essential role of the parameter α is to calibrate an impact of the specificity criterion on the constructed information granule. Note that if $\alpha = 0$ then the value of the exponential function becomes 1 hence the criterion of specificity of information granule is completely ruled out (ignored). In this case, $b = x_{max}$ with x_{max} being the largest element in D. Higher values of α stress the increasing importance of the specificity criterion.

The maximal value of α, say α_{max}, is determined by requesting that the optimal interval is the one for which $b_{opt} = x_1$, where x_1 is the data point closest to the median and larger than it. More specifically we determine α_{max} so that it is the smallest positive value of α for which the satisfaction of the following collection of inequalities holds,

$$1 * \exp(-\alpha|med(D) - x_1|) > 2 * \exp(-\alpha|med(D) - x_2|)$$
$$1 * \exp(-\alpha|med(D) - x_1|) > 3 * \exp(-\alpha|med(D) - x_3|)$$
$$\ldots$$
$$1 * \exp(-\alpha|med(D) - x_1|) > p * \exp(-\alpha|med(D) - x_p|)$$
$$\tag{12}$$

where the data $x_1, x_2 \ldots x_p$ form a subset of D and are arranged as follows $med(D) \le x_1 \le x_2 \le \cdots \le x_p$. Once the largest value of α_{max} has been determined, the range of these values $[0, \alpha_{max}]$ can be normalized to $[0, 1]$ and then the corresponding intervals $[a, b]$ indexed by α can be sought as a union of α-cuts of a certain fuzzy set of information granule A whose membership function reads as

$$A(x) = \bigcup_{\alpha \in [0,1]} \alpha \chi_{[a,b](\alpha)}(x) \tag{13}$$

where χ is the characteristic function of the interval $[a, b](\alpha)$. In this way, the principle of justifiable granularity gives rise to a fuzzy set.

The concept of justifiable information granule has been presented in its simplest, illustrative version. In case of multivariable data, each variable is treated separately giving rise to the corresponding information granules and afterwards a Cartesian product of them is formed. The principle of justifiable granularity can be applied to experimental data being themselves information granules rather than numeric data. In these situations, some modifications of the coverage criterion are required.

Appendix B

The FCM clustering algorithm has been commonly used to cluster numeric data. Here we are concerned with fuzzy sets represented as a family of α-cuts. To accommodate this format of objects to be clustered, it is essential to come up with their representation such that the geometry of the nested hyper boxes is captured to guide the clustering of the objects. The clustering is realized for a single variable, x and Δx. The organization of the feature space for the amplitude (x) involves a series of the lower and upper bounds of the intervals (α-cuts) in the form $[m^- \; m^+ \; x^-(\alpha_1) \; x^+(\alpha_1) \; x^-(\alpha_2) \; x^+(\alpha_2) \ldots x^-(\alpha_p) \; x^+(\alpha_p)]$. The representation of granular data in the x space as an example is illustrated in Fig. 9. The organization of the data in the Δx space is realized in the same way.

As the result of clustering done individually for the space of amplitude and change of amplitude, we arrive at a number of prototypes (as families of α-cuts) defined in the individual spaces. The prototypes can be associated with some semantics such as *low amplitude, medium amplitude, high amplitude, small positive change of amplitude*, etc. We form composite descriptors by taking Cartesian products of the prototypes. Those come with the semantics associated with the components of the products, say high amplitude and small negative change (for short, $High \times Small$). In the formation of these compound information granules the point has to be stressed, though. This concerns an emergence of so-called
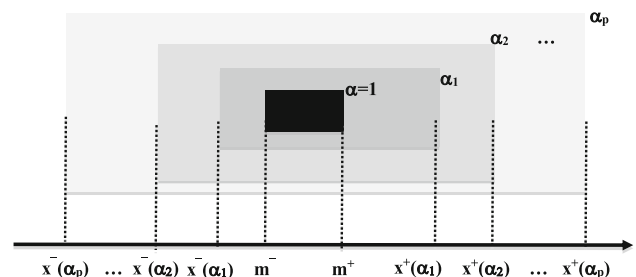

Fig. 9 Representation of granular data to be used in fuzzy clustering

spurious prototypes. While the granular prototypes formed in the individual spaces are well-supported (legitimized) by the granular data, the formation of the Cartesian products of some of them might not. This is caused by the fact that some Cartesian products arise in the region of the $x - \Delta x$ space where there are no granular data or a very limited number of them, see Fig. 10. Emergence of spurious granular prototypes in the regions sparsely populated by data. To account for this phenomenon and evaluate the legitimacy of the granular prototypes, we count the number of data contained in the α-cut of the prototype formed in the $x - \Delta x$ space and compute the expression (legitimacy level) μ:

$$\mu = \sum_{\alpha} (\alpha * \text{number of data covered by} A_i(\alpha) \times B_j(\alpha)) \tag{14}$$

where $A_i(\alpha)$ and $B_j(\alpha)$ are the α-cuts of the prototypes in the space of amplitude and change of amplitude and forming the Cartesian product $A_i \times B_j$ whose legitimacy is evaluated.

Appendix C

To quantify the closeness (matching) of two granular constructs, a certain prototype V_i and given information granule X, both described as a family of α-cuts, there are several alternatives that could be considered. Computing possibility measure, $Poss(X, V)$ is one among them. If 0-cuts of both X and V are disjoint, this measure returns zero and might be that in many situations we end up with a zero degree of matching. Another more sensible alternative is outlined below. The idea is outlined for a given α-cut of V_i and X. Let us assume that for the j-th variable the bounds of the granular prototype V_i form the interval $[v_{ij}^-, v_{ij}^+]$. For the j-th coordinate of X, $X_j = [x_j^-, x_j^+ |]$, we consider two situations: (i) $[x_i^-, x_i^+][v_{ij}^-, v_{ij}^+]$. In this case, it is intuitive to accept that the distance is equal to zero (as X_j is included in the interval of the V_{ij}) $dist_{ij} = 0$; (ii) $[x_i^-, x_i^+] \not\subset [v_{ij}^-, v_{ij}^+]$. Here we determine the minimal and the maximal distance between the bounds of the corresponding intervals, say

$$d_{ij}^- = min((x_j^- - v_{ij}^-)^2, (x_j^- - v_{ij}^+)^2, (x_j^+ - v_{ij}^-)^2, (x_j^+ - v_{ij}^+)^2)$$

$$d_{ij}^+ = max((x_j^- - v_{ij}^-)^2, (x_j^- - v_{ij}^+)^2, (x_j^+ - v_{ij}^-)^2, (x_j^+ - v_{ij}^+)^2)$$

The distance being computed on a basis of all variables $\|X - V_i\|^2$ is determined coordinatewise by involving the two situations outlined above. The minimal distance obtained in this way is denoted by $d_{min}(X, V_i)$ while the maximal one is
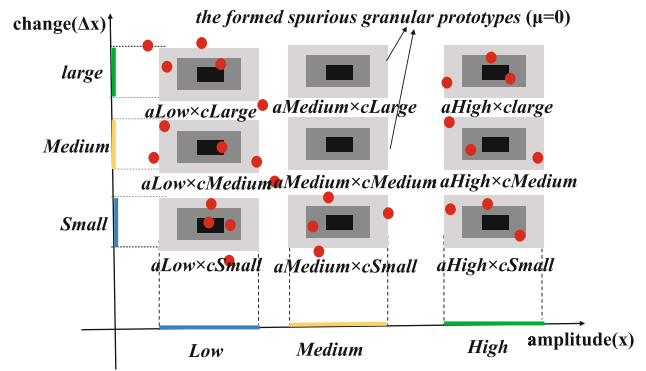

Fig. 10 Granular prototypes in $x - \Delta x$ space formed by taking Cartesian products of granular prototypes of x and Δx feature space

denoted by $d_{max}(X, V_i)$. More specifically we have,

$$d_{min}(X, V_i) = \sum_{j=1}^{n} d_{ij}^- \tag{15}$$

$$d_{max}(X, V_i) = \sum_{j=1}^{n} d_{ij}^+ \tag{16}$$

Having the distances determined, we compute the two expressions

$$w_{1(X)} = \frac{1}{\sum_{j=1}^{c} (\frac{d_{min}(X,V_i)}{d_{min}(X,V_j)})^{\frac{1}{(m-1)}}} \tag{17}$$

$$w_{2(X)} = \frac{1}{\sum_{j=1}^{c} (\frac{d_{max}(X,V_i)}{d_{max}(X,V_j)})^{\frac{1}{(m-1)}}} \tag{18}$$

Notice that these two formulas resemble the expression used to determine the membership grades in the FCM algorithm. In the sequel we compute the levels of matching as follows

$$\mu_i^-(X) = min(w_1(X), w - 2(X)) \tag{19}$$
$$\mu_i^+(X) = max(w_1(X), w - 2(X)) \tag{20}$$

In essence, we arrive at the interval-valued matching degree $U_i(X) = [u_i^-(X), u_i^+(X)]$.

The degree of matching λ is a function of α so in order to come up with a numeric representation, one determines a numeric representative of the interval, say its average $u_i^*(X) = [u_i^-(X) + u_i^+(X)]/2$ and aggregates these values over λ that is

$$\lambda = \int_0^1 u_i^*(X) d\alpha \tag{21}$$

References

Abonyi J, Fell B, Nemeth S, Arva P (2005) Modified gathgeva clustering for fuzzy segmentation of multivariate time-series. Fuzzy Sets Syst 149(1):39–56

Apolloni B, Bassisa S, Malchiodi D, Pedryczb W (2008) Interpolating support information granules. Neurocomputing 71(13–15):2433–2445

Aznarte JL, Manuel J, Benítez JM (2010) Equivalences between neural-autoregressive time series models and fuzzy systems. IEEE Trans Neural Netw 21(9):1434–1444

Bargiela A, Pedrycz W (2002) A model of granular data: a design problem with the tchebyschev-based clustering. In: Proceedings of the 2002 IEEE international conference on fuzzy systems, vol 1, pp 578–583

Bargiela A, Pedrycz W (2003) Granular computing: an introduction. Kluwer Academic Publishers, Dordrecht

Bargiela A, Pedrycz W (2005) Granular mappings. IEEE Trans Syst Man Cybern Part A 35(2):292–297

Bargiela A, Pedrycz W (2008) Toward a theory of granular computing for human-centered information processing. IEEE Trans Fuzzy Syst 16(2):320–330

Bargiela A, Pedrycz W (eds) (2009) Human-centric information processing through granular modelling. Springer, Berlin

Bezdek JC (1981) Pattern recognition with fuzzy objective function algorithms. Kluwer Academic Publishers, Dordrecht

Brown RG (2004) Smoothing, forecasting and prediction of discrete time series. Dover Publications, Dover Phoenix Ed edition, NY

Casillas J, Cordón O, Herrera F, Magdalena L (eds) (2003) Interpretability issues in fuzzy modeling. Springer, Berlin

Chen SM, Chen CD (2011) Taiex forecasting based on fuzzy time series and fuzzy variation groups. Exp Syst Appl 19(1):1–12

Duncan S, Bryant G (1996) A new algorithm for segmenting data from time series. In: Proceedings of the 35th IEEE conference on decision and control, vol 3, pp 3123–3128

Fitzgibbon LJ, Dowe DL, Allison L (2002) Change-point estimation using new minimum message length approximations. In: Proceedings of the seventh Pacific Rim international conference on artificial intelligence: trends in, artificial intelligence, pp 244–254

Fu T (2011) A review on time series data mining. Eng Appl Artif Intell 24(1):164–181

Fu T, lai Chung F, man Ng C (2006) Financial time series segmentation based on specialized binary tree representation. In: Proceedings of the 2006 international conference on data mining, pp 3–9

Hirota K, Pedrycz W (1984) Characterization of fuzzy clustering algorithms in terms of entropy of probabilistic sets. Pattern Recognit Lett 2(4):213–216

Jiang J, Zhang Z, Wang H (2007) A new segmentation algorithm to stock time series based on pip approach. In: Proceedings of the international conference on wireless communications, networking and mobile computing, pp 5609–5612

Kasabov NK, Song Q (2003) Denfis: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. IEEE Trans Fuzzy Syst 10(2):144–154

Keogh EJ, Chu S, Hart D, Pazzani MJ (2001) An online algorithm for segmenting time series. In: Proceedings of the 2001 IEEE international conference on data mining, pp 289–296

Lee LW, Wang LH, Chen SM, Leu YH (2006) Handling forecasting problems based on two-factor high-order fuzzy time series. IEEE Trans Fuzzy Syst 14(3):468–477

Oliver JJ, Baxter RA, Wallace CS (1998) Minimum message length segmentation. In: Proceedings of the second Pacific–Asia conference on knowledge discovery and data mining, pp 222–233

Pawlak Z (1985) Rough sets and fuzzy sets. Fuzzy Sets Syst 17:99–102

Pawlak Z (1991) Rough sets: theoretical aspects of reasoning about data. Springer, Berlin

Pawlak Z, Skowron A (2007a) Rough sets: some extensions. Artif Intell 177(1):28–40

Pawlak Z, Skowron A (2007b) Rudiments of rough sets. Inf Sci 177(1):3–27

Pedrycz W (1998) Shadowed sets: representing and processing fuzzy sets. IEEE Trans Syst Man Cybern Part b Cybern 28(1):103–109

Pedrycz W (1999) Shadowed sets: bridging fuzzy and rough sets, Springer, Berlin, pp 179–199

Pedrycz W, Gomide F (2007) Fuzzy systems engineering: toward human-centric computing. Wiley, NY

Pedrycz W, Song M (2011) Analytic hierarchy process (ahp) in group decision making and its optimization with an allocation of information granularity. IEEE Trans Fuzzy Syst 19(3):527–539

Qian Y, zhi Z Wu W, Dang C (2011) Information granularity in fuzzy binary grcmodel. IEEE Trans Fuzzy Syst 19(2):253–264

Ślęzak D (2009) Degrees of conditional (in) dependence: a framework for approximate bayesian networks and examples related to the rough set-based feature selection. Inf Sci 179(3):197–209

Srivastava A, Su R, Weigend A (1999) Data mining for features using scale-sensitive gated experts. IEEE Trans Pattern Anal Mach Intell 21(12):1268–1279

Wang ZJ, Willett PK (2002) Joint segmentation and classification of time series using class- specific features. IEEE Trans Syst Man Cybern Part B Cybern 34:1056–1067

Yager RR (2009) Participatory learning with granular observations. IEEE Trans Fuzzy Syst 17(1):1–13