Measuring the Temporal Stability of Fuzzy Linguistic Summaries about Time Series with Drifts

OSTROWSKI Marcin[1] and KACZMAREK-MAJER Katarzyna[1,2] and HRYNIEWICZ Olgierd[1]

[1] *Systems Research Institute, Polish Academy of Sciences, Newelska 6, 01-147 Warsaw, Poland*
E-mail: [marcin.ostrowski, k.kaczmarek]@ibspan.waw.pl
[2] *University of Ostrava, Institute for Research and Applications of Fuzzy Modeling, 30. Dubna 22, 701 03 Ostrava 1, Czech Republic*

1 Introduction

Fuzzy linguistic summaries provide insights about large numerical datasets in natural language. While their practical potential has been demonstrated with many applications across domains, effective monitoring of evolving sequences of such summaries remains a significant challenge. This limitation is especially evident in dynamic environments such as remote health monitoring, where new data are collected continuously and human-consistent monitoring approaches are needed.

The majority of the related works concentrate on assessing the quality of individual summary sentences, typically in terms of measures such as degree of truth, confidence, support, informativeness, or focus. The evaluation of sets or sequences of summaries is considerably more complex, particularly in real-world settings where data may arrive incrementally or remain partially incomplete, thereby introducing additional uncertainty into the assessment process. Interpretability of collections of linguistic summaries has been first studied in [1] and further extended in [2].

In this work, we aim to explore the properties of the sequences of fuzzy linguistic summaries, focusing on the construction of effective and easy-to-understand information granules that support communication about changes in the observed multivariate time series. For this purpose, we adapt the fuzzy linguistic summaries based on the concept of extended protoforms [3]. Let us now briefly recall the form of fuzzy linguistic summaries (FLS) [2]:

$$S : Q\ R_1 \star \ldots \star R_k\ x \text{ are } P_1 \diamond \ldots \diamond P_l, \tag{1}$$

where $\star, \diamond \in \{\text{AND, OR}\}$, Q is a linguistic quantifier, $R_1, \ldots, R_k$ are qualifiers and $P_1, \ldots, P_l$ are summarizes. This work also builds on previous research related to the theories of generalised and intermediate quantifiers [4] initiated by the work of Mostowski [5]. Formula [[1]] can be expressed also as the generalized quantifier of type $< 1, 1 >$ [6] being an operator Q binding

Acknowledgement Marcin Ostrowski and Katarzyna Kaczmarek-Majer acknowledge funding from the project "ExplainMe: Explainable Artificial Intelligence for Monitoring Acoustic Features extracted from Speech" (FENG.02.02-IP.05-0302/23) carried out within the First Team programme of the Foundation for Polish Science co-financed by the European Union under the European Funds for Smart Economy 2021-2027 (FENG). This work is supported from the project ,,Research of Excellence on Digital Technologies and Wellbeing CZ.02.01.01/00/22_008/0004583" which is co-financed by the European Union.

https://www.doi.org/10.15452/978-80-7599-515-5.2026.17

n-variables: $(Qx_, \ldots, x_n)(\varphi_1(x_, \ldots, x_n), \varphi_2(x_, \ldots, x_n))$. In this work, we consider selected examples of summaries with quantifiers *Majority* and *Around half*. The main focus of this work is on monitoring sequences of FLS. We propose a stability index and confront it with the selected well-established approaches of the statistical process control.

2 Stability assessment

Let us assume that we observe a stream of numerical data $X = x_1, \ldots, x_N$, N is large, e.g., voice features about speech collected sequentially. We assume that this stream X is divided into n segments (chunks). Fuzzy linguistic summaries (Eq.1) are constructed for consecutive segments. For further details and FLS definition, we refer to, e.g., [2].

Let $\mathcal{T}_2(S) = \{T_1^{(j)}(S), T_2^{(j)}(S)\}_{j=1}^n$ denote the sequence of degree of truth values and degree of support for a given linguistic summary S, where $T_1^{(j)}(S)$ represents degree of truth calculated for j-th segment of data. We consider the truth function (degree of truth) according to the following formula:

$$T_1(S) = \begin{cases} Q\left(\frac{\sum_{i=1}^n C(P(x_i), R(x_i))}{\sum_{i=1}^n R(x_i)}\right), & Q \text{ is relative,} \\ Q\left(n \cdot \frac{\sum_{i=1}^n C(P(x_i), R(x_i))}{\sum_{i=1}^n R(x_i)}\right), & Q \text{ is absolute,} \end{cases} \tag{2}$$

where S is a fuzzy linguistic summary, C is a conjunctive aggregation function. The degree of support is given by

$$T_2(S) = \frac{1}{n}|\{x_i : P(x_i) > 0 \wedge R(x_i) > 0\}|. \tag{3}$$

Let $\hat{T}_1^{(j)}(S)$ be its one-step predictor estimated via a modified exponentially weighted moving average:

$$\hat{T}_1^{(j)}(S) = (1-\lambda)\hat{T}_1^{(j-1)}(S) + \lambda T_1^{(j-1)}(S) + \beta d_{j-1}, \tag{4}$$

$$d_j = (1-\mu)d_{j-1} + \mu(T_1^{(t)}(S) - T_1^{(j-1)}(S)), \tag{5}$$

with initialization $\hat{T}_1^{(1)}(S) = T_1^{(1)}(S)$, $d_1 = 0$, and parameters $\lambda \in [0,1]$, $\mu \in [0,1]$, $\beta \in [0,1]$.

In previous work [7], a modified exponentially weighted moving average was considered for monitoring outcomes from the feature importance over time. In this research, we extend this idea and propose the summarization consistency index (SCI) for a fuzzy linguistic summary S to be defined as follows:

$$\mathrm{SCI}(\mathcal{T}_2(S)) = 1 - \sum_{j=2}^{n} T_2^{(j)}(S)\, \ell\left(r_S^{(j)}\right), \qquad r_S^{(j)} = \frac{T_1^{(j)}(S) - \hat{T}_1^{(j)}(S)}{s_j}, \tag{6}$$

where $T_2^{(j)}(S)$ is the degree of support calculated for fuzzy linguistic summary S on the j-th segment of data and

$$\ell(r) = \frac{r^2}{\tau^2 + r^2}, \qquad s_j = \mathrm{MAD}\left(\{T_1^{(j)}(S)\}_{j=1}^n\right) > 0, \qquad \tau > 0.$$

Intuitively, the approach captures how predictable the current degree of truth for a summary is given its historical trajectory and taking into account the corresponding degree of support. In Eq. 6, this intuition is formalized through a weighted aggregation of local deviations, where each term ℓ represents the loss function associated with the discrepancy between the observed and the expected value at time t.

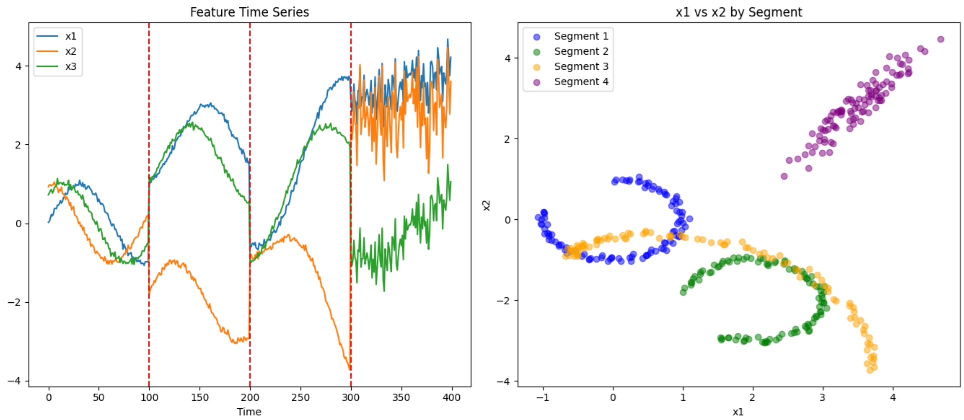

Figure 1: Simulated time series with various types of drift introduced at moment $t = 100$: sudden drift (abrupt change in mean); $t = 200$: incremental (gradual shift in mean); $t = 300$: variance change due to the increase in noise.

The SCI is bounded in $[0, 1]$, where higher values indicate smoother and more temporally consistent attributions. It is scale-invariant, achieved through robust normalization by the median absolute deviation (MAD), and adaptive to trends via exponential and momentum-based smoothing. By construction, SCI decreases monotonically with the accumulation of large, unexpected attribution shifts, making it sensitive to temporal irregularities while robust to small fluctuations.

For comparative purposes, we consider selected approaches from statistical process control. For further details, see e.g., the seminal work of Montgomery [8]. We concentrate on the methods for autocorrelated data. One approach could be that the original data are charted, but control limits are adjusted using the knowledge about the type of dependence. Alternatively, charts for monitoring residuals are constructed. To obtain the residuals, predictions are commonly calculated according to the stationary autoregressive model AR of the most appropriate order p: $X_t = \mu + a_1 X_{t-1} + \cdots + a_p X_{t-p} + Z_t$, where $X_i, i = t, t-1, \ldots$ are random observations of a process, $Z_i, i = t, t-1, \ldots$ is a series of independent random variables with constant variance and zero expectation, $a_1, \ldots, a_p$ are process parameters describing its correlation structure, and μ is a constant describing a process level.

3 Experiments

Experiments will be presented for the real-life and simulated time series with different types of introduced drift: (1) sudden (abrupt change in mean); (2) incremental (gradual shift in mean); (3) variance change (e.g., increase in noise level). Fig. 1 depicts exemplary simulated time series with these three types of drift; thus, we split the sequence into four segments of equal length. For the first segment, we consider the following three time series: $x_1(t) = \sin(0.05t) + \varepsilon_1(t)$; $x_2(t) = \cos(0.05t) + \varepsilon_2(t)$; $x_3(t) = \sin\big(0.05t + \frac{\pi}{4}\big) + \varepsilon_3(t)$ where $\varepsilon_i(t) \sim \mathcal{N}(0,\ 1)$. Then, at $t = 100$, we shift each variable by a constant (mean shift) of $c_1 = 2, c_2 = -2andc_3 = 1.5$. Then, at $t = 200$, incremental drift is introduced. At $t = 300$, the generation mechanism changes, and the noise is particularly bigger: $x_1(t) = 0.01\,t + \eta_1(t)$; $x_2(t) = 0.8\,x_1(t) + \eta_2(t)$; $x_3(t) = \cos(0.03\,t) + \eta_3(t)$, where $\eta_i(t) \sim \mathcal{N}(0,\ 0.4^2), i = 1, 2, 3$.

Next, FLS will be constructed for the consecutive segments in natural language, and the respective quality criteria and summarization stability index will be calculated. For reference, control charts for residuals will be considered. Such control chars about two exemplary sum-

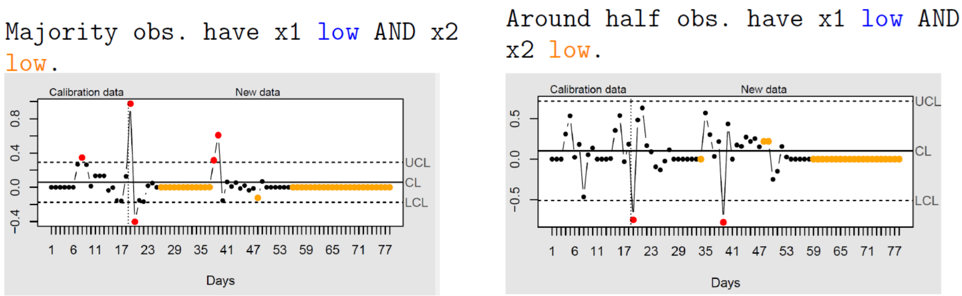

Figure 2: Two residual control charts for the linguistic summaries about simulated time series.

maries: *Majority observations have* x_1 *low and* x_2 *low.* and *Around half observations have* x_1 *low and* x_2 *low.* are illustrated in Fig. 2. As observed in this simple illustrative example, when monitoring the considered two summaries, alarms are generated in similar moments (around days 20 and 40), which correspond to the changes in the original time series generation process at $t = 100$ and $t = 200$.

In our contribution, we will thoroughly analyse the stability of sequences of FLSs with various quantifiers and various drifts, aiming to identify patterns between the statistical properties in observed multivariate time series and the calculated stability indexes. The secondary goal is to characterise the group of fuzzy linguistic summaries that may serve as promising explanations of changes detected in the original time series.

References

[1] M.-J. Lesot, G. Moyse, B. Bouchon-Meunier: *Interpretability of fuzzy linguistic summaries*, Fuzzy Sets and Systems **292** (2016) 307-317.

[2] K. Miś, K. Kaczmarek-Majer, M. Baczyński: *Fuzzy Linguistic Summaries and the Double Negation Property.* TechRxiv. July 24 (2025).

[3] L. A. Zadeh: *A computational approach to fuzzy quantifiers in natural languages.* Computers & Mathematics with Applications, **9(1)** (1983) 149–184.

[4] P. Murinova, V. Novak: *The theory of intermediate quantifiers in fuzzy natural logic revisited and the model of "Many".* Fuzzy Sets and Systems **388**, 56–89 (2020)

[5] A. Mostowski: *On a generalization of quantifiers.* Fund. Math. **44** (1957) 12–36.

[6] P. Lindsträm: *First-order predicate logic with generalized quantifiers* Theoria. **32** (1966) 186–195.

[7] M. Ostrowski, O. Hryniewicz: *Beyond static importance: Quantifying stability and distribution drift*, in: Proceedings of the Second Workshop on Explainable Artificial Intelligence for the medical domain- 25-30 October 2025, Bologna, Italy, co-located with 28th European Conference on Artificial Intelligence- ECAI 2025, volume — of CEUR Workshop Proceedings, (2025).

[8] D. C. Montgomery: *Introduction To Statistical Quality Control* **6th Ed.** (2011) Wiley.