#儘田　龍一[1), 藤本　晶子[1), 栗田　怜[2), 笠原　禎也[3), 松田　昇也[4), 松岡　彩子[5), 三好　由純[6), 中村　紗都子[7)
[1 九工大,[2 京都大学 生存研,[3 金沢大,[4 金沢大学,[5 京都大学,[6 名大 ISEE,[7IAR&ISEE

Electrostatic Solitary Waves detection model based on string compression and string-matching technology

#Ryuichi Mamada[1),Akiko Fujimoto[1),Satoshi Kurita[2),Yoshiya Kasahara[3),Shoya Matsuda[4),Ayako Matsuoka[5),Yoshizumi Miyoshi[6),Satoko Nakamura[7)
[1Kyushu Institute of Technology,[2Research Institute for Sustainable Humanosphere, Kyoto University,[3Emerging Media Initiative, Kanazawa University,[4Kanazawa University,[5Graduate School of Science, Kyoto University,[6Institute for Space-Earth Environement Research, Nagoya University,[7Nagoya University

Electrostatic Solitary Waves (ESWs), which are electric field solitary pulses, are frequently observed in the magnetosphere. ESWs play important roles in the wave-particle interaction as energy exchange between magnetospheric plasma particles. Methods on detection of ESW waveform have been proposed roughly in terms of image filtering (Kojima et al., 2000) and time-series filtering (Yagi, 2015). In Kojima et al. (2000), they used a bit-pattern comparison method to extract ESW waveforms. Computational cost of this method depends on a size of bit-pattern filter (a table of M x M matrices). In addition, detected ESW waveforms are limited due to pre-fixed filter. Method of Yagi (2015) can obtain ESW waveforms using median filters with two frequencies, in order to remove the high-frequency component superimposed on ESW and the low-frequency component of the background trend variation. This method is also limited to detect ESW waveforms because of frequencies of the median filters.

In this paper, we propose a new ESW detection model with low computational cost and few restrictions on detectable waveforms based on the string run-length compression and string-matching techniques. The proposed model consists of four components: a numeric value-run length compression string converter, an ESW waveform string searcher, a feature quantity calculator, and an ESW discriminator. A unique feature of this model is that it replaces the ESW waveform pattern recognition problem with a string-matching problem. The waveform length (=number of data points) of the typical ESW pattern is longer than the noise waveform length (=more data points). The values for ESW waveform length and waveform pattern are quantified as three different character string lengths and a character string length compression ratio. The ESW discriminator obtains discriminant conditions through classification tree learning by giving the above quantified variables as explanatory variables and ESW correct and incorrect labels as objective variables.

We use the electric field data (E_{jj}, parallel to the magnetic field) obtained by the Waveform Capture (WFC) of the Plasma Wave Experiment (PWE) onboard the Arase satellite, during 2017/8/414:28:04 to 2017/8/414:28:09 (5 seconds). The 5-second data are divided into 0.1-second intervals, training data are used randomly selected 10 dataset of 0.1-second intervals (1 second in total), and the remaining 4-second data is validation data. The recall (r), the precision (p), and the F-value (f) as model accuracies are calculated for the learning result of decision tree using the training data, and we select seven tree models with relatively high accuracy values as candidate ESW detection models. The performance experiment of these models applied to validation data and resulted in high accuracy with the recall (r=0.56), the precision (p=0.87), and the F-value (f=0.68) for the model of the character string length compression ratio 17 %. We will present more detailed specifications and improvements for the proposed model.