Recursive Least Square: RLS Method-Based Time Series Data Prediction for Many Missing Data

Kohei Arai[1]
Graduate School of Science and Engineering
Saga University, Saga City, Japan

Kaname Seto[2]
Former Student
Saga University, Saga City, Japan

Abstract—**Prediction methods for time series data with many missing data based on Recursive Least Square (RLS) method are proposed. There are two parameter tuning algorithms, time update and measurement update algorithms for parameter estimation of Kalman filter. Two learning methods for parameter estimation of Kalman filter are proposed based on RLS method. One is the method without measurement update algorithm (RLS-1). The other one is the method without both time and measurement update algorithms (RLS-2). The methods are applied to the time series data of Defense Meteorological Satellite Program (DMSP) / Special Sensor Microwave/Imager (SSM/I) data with a plenty of missing data. It is found that the proposed RLS-2 method shows smooth and fast convergence in learning process in comparison to the RLS-1.**

Keywords—*Special Sensor Microwave/Imager (SSM/I); Defense Meteorological Satellite Program (DMSP); Kalman filter; Recursive Least Square (RLS) method; missing data; parameter estimation*

I. INTRODUCTION

In general, earth observation satellites observe arbitrary points on the earth at unequal time intervals based on their orbital conditions. When this observation data is regarded as time-series data at equal time intervals, it can be regarded as time-series data including many unobserved and missing data.

One of the purposes of the time series analysis is to improve prediction accuracy of future data with the past data for the time series of data with a plenty of missing data. There is the famous method, so called, Kalman filter for future data prediction with the previously observed time series of data. There are the parameters for Kalman filter. It, however, is difficult to estimate the parameters.

Kalman filter is composed of an algorithm that updates the state with time (time update algorithm) and an algorithm that updates the observation process (observation update algorithm). Here, as a dynamic characteristic extraction method, we consider both the time series state and the observation process, and examine a method based on time series analysis using the Kalman filter [1], which is widely used because of its relatively high estimation accuracy.

As the parameter estimation method for adaptive filtering, the sequential least squares method (RLS method) [2], [3], which performs sequential learning on the assumption that the target time series is linearly stationary, is generally used.

Matsuoka and Tateishi reported on reflectance correction for remote sensing data including missing data actually observed using a time series model (BRDF model; Bidirectional Reflectance Distribution Function model) [a priori knowledge] [4]. However, when the a priori knowledge cannot be introduced when extracting the dynamic characteristics of the target time series, or when the a priori knowledge is used and the residual time series is used to improve the accuracy, the target is used. It is also conceivable to try to extract the dynamic characteristics. In such a case, a method for extracting the dynamic characteristics from only the time series data is required. In addition, there is no qualitative study on the method for extracting the dynamic characteristics from only the remote sensing data including the observed missing data.

The method proposed in this paper aims to estimate the data at an arbitrary time only from the time series data including such a large amount of missing data. In order to make this purpose possible, some method for extracting the dynamic characteristics of the target time series is required.

The following section describes research background followed by related research works. Then the proposed method is described followed by experiment. After that, conclusion is described together with some discussions.

II. RESEARCH BACKGROUND

As an example, Fig. 1 shows the SSM / 1 (Special Sensor of Microwave / Imager) microwave radiometer mounted on the DMSP (Defense Meteorological Satellite Program) satellite [quasi-return orbit] on June 2, 1998 (Japan time).
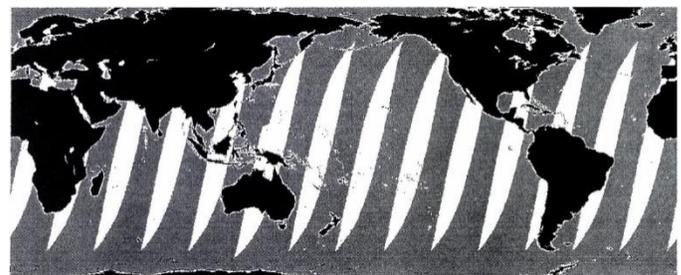
Fig. 1. An Example of the Observed Area with SSM/l Acquired on June 2 1998. The Land Areas are Colored in Black While the Ocean Areas are Colored in Gray and White. The Gray Colored Areas show the Observed Areas While the White Colored Areas show the Non-Observed Areas.

The obtained global observation area is shown. The black area is the land area, and the white / gray area is the sea area.

The gray area is the area observed in the sea area, and the white area is the area not observed in the sea area (missing area; missing data). Although there are missing areas in the land area, this time the land area was classified as the land area. From Fig. 1, it can be seen that there are many missing areas (missing data) in a wide area.

III. RELATED RESEARCH WORKS

As for the time series analysis, prediction method for time series of imagery data in eigen space is proposed [5]. Meanwhile, Geography Markup Language: GML based representation of time series of assimilation data and its application to animation content creation and representations is proposed [6]. On the other hand, recovering method of missing data based on the proposed modified Kalman filter when time series of mean data is known is proposed [7]. Time series analysis for shortened labor mean interval of dairy cattle with the data of BCS, RFS, Weight, Amount of Milk and Outlook is conducted [8].

Meanwhile, as for the Kalman filter related research, detecting algorithm for rainfall area movement based on Kalman filtering is proposed [9]. Rain flagging with SSM/I based on Kalman filtering with new parameter estimation is proposed [10] together with rain flagging for NSCAT with SSM/I through gap filling based on Kalman filter [11]. Also, rain flagging method with Kalman filtering for ADEOS/NSCAT is proposed [12]. On the other hand, comparative study on image prediction methods between the proposed morphing utilized method and Kalman filtering method is conducted [13].

On the other hand, time series analysis based on Kalman filter with parameter estimation considering non-linearity of system is conducted [14]. Then handling of missing data in parameter estimation of Kalman filter by RLS method is also conducted [15]. Furthermore, recovering method of missing data based on the proposed modified Kalman filter when time series of mean data is known is proposed [16].

Moreover, detecting algorithm for rainfall area movement based on Kalman filtering is proposed [17] together with rain flagging with SSM/I based on Kalman filtering with new parameter estimation [18]. Then rain flagging for NSCAT with SSM/I through gap filling based on Kalman filter is proposed [19] together with rain flagging method with Kalman filtering [20].

IV. PROPOSED METHOD

A. Theoretical Background: Kalman Filter

The Kalman filter was proposed by Kalman in 1960. This is an extension of the previous Wiener filter theory so that it can be applied even when time series data including signals and noise are described in a non-stationary process [14], [15]. Below, x represents the mean of x and x^T represents the transpose of x.

In the Kalman filter, the process of time series data is generally expressed by equations (1) and (2).

$$x_{t+1} = F_t x_t + u_t \qquad (1)$$

$$y_t = H_t x_t + w_t \qquad (2)$$

where, x_t indicates the state at time t, u_t is the system noise, y_t is the observation data at time t, w_t is the observation noise, F_t is the state transition matrix, and H_t is the observation matrix. Equations (1) and (2) are called state equations and observation equations, respectively, and the Kalman filter can be constructed by determining the state transitions of both equations and the elements of the observation matrix. It is possible to estimate missing data.

The learning algorithm for estimating the state x_t of the Kalman filter (Kalman filter learning algorithm) consists of the time update algorithm and the observation update algorithm of Eq. (3), (4).

$$\widehat{x_{t+1|t}} = F_t \widehat{x_{t|t}} \qquad (3)$$

$$P_{t+1|t} = F_t P_{t|t} F^T_t + Q_t \qquad (4)$$

$$K_t == P_{t|t-1} H^T_t (H_t P_{t|t-1} H^T_t + R_t)^+ \qquad (5)$$

$$\widehat{x_{t|t}} = \widehat{x_{t|t-1}} + K_t (y_t - H_t \widehat{x_{t|t-1}}) \qquad (6)$$

$$P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1} \qquad (7)$$

where, x represents the estimated value of x, and $(H_t P_{t|t-1} H^T_t + R_t)^+$ represents the Moore-Penrose generalized inverse matrix of $(H_t P_{t|t-1} H^T_t + R_t)$. t | t-1 means a variable that transitions from t-1 to t. However, the following equations (8) to (14) are assumed.

$$E[u_k] = \overline{u_k} = 0 \qquad (8)$$

$$E[(u_k - \overline{u_k})(u_l - \overline{u_l})^T] = \delta_{k,l} Q_k \qquad (9)$$

$$E[w_k] = \overline{w_k} = 0 \qquad (10)$$

$$E[(w_k - \overline{w_k})(w_l - \overline{w_l})^T] = \delta_{k,l} R_k \qquad (11)$$

$$E[x_0] = \overline{x_0} \qquad (12)$$

$$E[(x_0 - \overline{x_0})(x_0 - \overline{x_0})^T] = P_0 \qquad (13)$$

$$\delta_{k,l} = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \qquad (14)$$

B. Autoregressive Model and Sequential Least Squares Method (RLS Method)

Consider the n-dimensional autoregressive model of Eq. (15).

$$y(t) = \sum_{j=1}^{d} A_j y(t-j) + v(t) \quad (t = 0,1,\dots) \qquad (15)$$

When estimating the coefficient matrix A_j ($j = 1,2,\dots, d$) in Eq. (15), the Kalman filter can be constructed as follows.

$$z(t+1) = z(t) \qquad (16)$$

$$y(t)^T = H_t z(t) + v(t)^T \qquad (17)$$

However, the following equations (18) and (19) are used.

$$z(t) = \begin{bmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_{d-1}^T \\ A_d^T \end{bmatrix} \qquad (18)$$

$$H_t = [y \, (t\text{-}1)^{\,T}, \, y \, (t\text{-}2)^{\,T}, \ldots, y \, (t\text{-}d)^{\,T}] \tag{19}$$

Also, the fact that the equations (20), (21), and (22) are used is used.

$$y \, (t)^{\,T} = (y1 \, (t), y2 \, (t), \ldots, yn \, (t)) \tag{20}$$

$$z \, (t) == [x \, (t, 1), x \, (t, 2), \ldots, x + (t, n)] \tag{21}$$

$$v \, (t)^{\,T} = (v1 \, (t), v2 \, (t), \ldots, vn \, (t)) \tag{22}$$

Therefore, it can be expressed as Eqs. (23) and (24).

$$x \, (t+1, i) = x \, (t, i) \tag{23}$$

$$y_i \, (t) = H_t x \, (t, i) + v_i \, (t) \tag{24}$$

where, $i = 1, 2, \ldots, n$. The RLS method applies the Kalman filter learning algorithm to Eqs. (16) and (17).

C. The Proposed Method

Fig. 2 shows the configuration of the Kalman filter. Adaptive filtering is a method of estimating the parameters used in these algorithms (time update algorithm, observation update algorithm) to realize the Kalman filter.

There is. However, when the RLS method is applied as a dynamic characteristic extraction method for time-series data containing a large amount of missing data, the RLS method performs sequential learning for the target time series, so the missing data in the sequential learning process, the coping method becomes a problem. Here, as a countermeasure for missing data in the learning process of the RLS method, we propose a method that uses only the time update algorithm without using the observation update algorithm and a method that does not use both the observation update algorithm and the time update algorithm. The former will be called RLS method 1 and the latter will be called RLS method 2. Fig. 3 and 4 show the difference between RLS method 1 and RLS method 2.
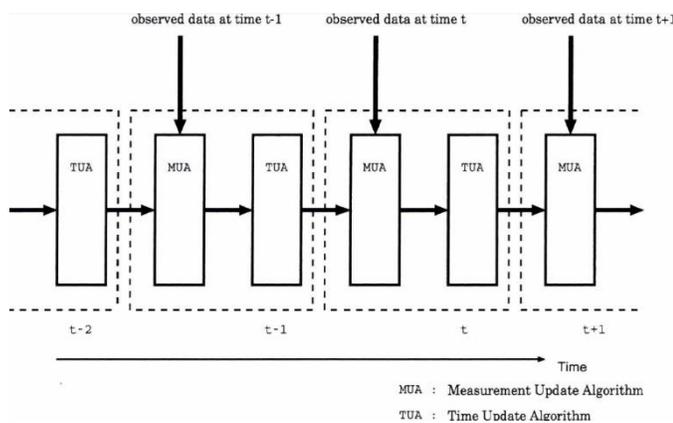
Fig. 2. Kalman Filter Consists of the Measurement Update Algorithm [Equation (5), (6), and (7)] and the Time Update Algorithm [Equation (3), and (4)].
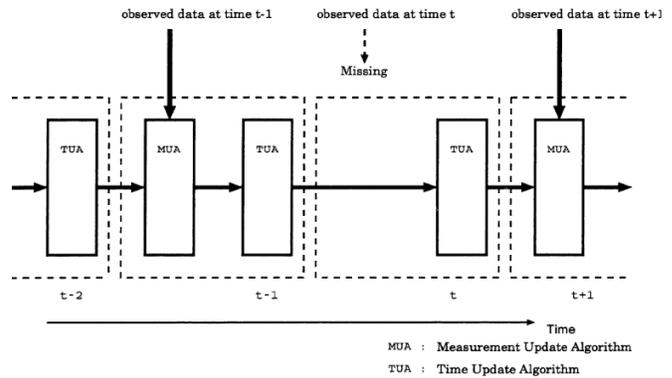
Fig. 3. The RLS Method #1 (MUA is not Applied to the Missing Data).
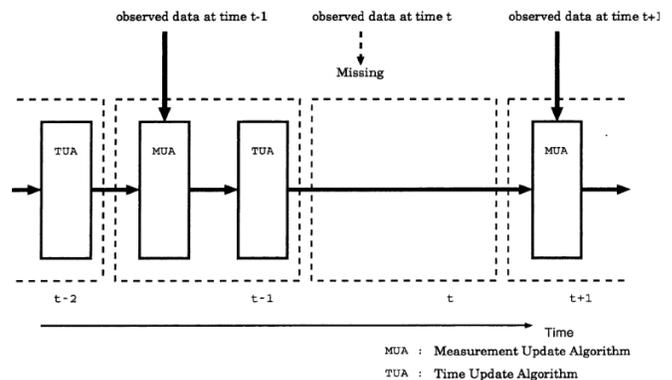
Fig. 4. The RLS Method #2 (Both MUA and TUA are not Applied to the Missing Data).

In this paper, we compare the learning process for past observation data and the prediction behavior for future data in the Kalman filter with parameter estimation based on those proposed methods (RLS method 1, RLS method 2), SSM / 1 observation brightness temperature.

The data (actual observation data) was used. By making these qualitative comparisons, it is possible to investigate the measures and application limits when the RLS method is used as a dynamic characteristic extraction method for time series containing a large amount of missing data. For time series containing many missing data, RLS method 2. The superiority of the proposed method is confirmed.

V. EXPERIMENT

A. The Data Used

Here, a simulation experiment is performed by imitating the SSM / I 19.3GHz vertical polarization observation luminance data as an n-dimensional time series. Experiments will be conducted with the time axis in the traveling direction of the DMSP satellite equipped with SSM / 1 and the dimensional axis in the scanning direction. In other words, the experiment is performed using the path-shaped observation luminance data.

SSM / I is a passive microwave radiometer launched by the United States for earth observation, and has 4 frequencies and 7 channels. The breakdown is 19.3GHz vertical polarization, 19.3GHz horizontal polarization, 22.235GHz vertical polarization, 37.0GHz vertical polarization, 37.0GHz

horizontal polarization, 85.5GHz vertical polarization, 85.5GHz horizontal polarization.

B. Creation of Missing Data

For the above n-dimensional time series, generate a uniform random number $l(k)$ with an integer from 0 to q, and use it as missing data for the unit time of the number $l(k)$ that appears (k-1, 2). , 3, ...). In addition, we will introduce an operation to move (shift) in the positive direction in the time axis direction for a unit time by ρ for an n-dimensional time series for time series data generation. In other words, p is the period of continuous observation [the length of discrete time, and q is the period of non-continuous observation (missing)).

Represents the maximum length of [discrete time]. It is considered that it is possible to generate time-series data including arbitrary missing data by introducing these (p,). Examples of time-series data including the missing data generated in this way are shown in Fig. 5 and 6. As shown in Fig. 5, multidimensional time series data including missing data was displayed using the time series of images. Fig. 6 helps to grasp the time axis and dimension axis in this experiment. When p = ∞ and q = 0, the time series does not include missing data (complete observation time series).

C. Evaluation Function

The evaluation function for past data (t <t0) and future data (t> t.) with the current time as t0 is as follows. As for the past data, Eq. (25) and (26) are used as evaluations to examine the learning rate for past observation data.

$$J1(t) = \sqrt{\frac{1}{N_t}\sum_{i=1}^{t} I(y(i)) \frac{1}{n}\|y^*(i) - y(i)\|^2} \qquad (25)$$

$$I(y(i)) \equiv \begin{cases} 1 \ y(i) \ is \ observed \ data \\ 0 \ y(i) \ is \ missing \ data \end{cases} \qquad (26)$$

However, t is an arbitrary time, and N_t is the number of observed data up to time t. Further, $y^*(i)$ is the estimated data (n-dimensional) at time i, and y (i) is the correct answer data (n-dimensional) at time i. Since missing data cannot be evaluated, the value of J1 (t) is negative.

In general, the value of J1 (t) is expected to decrease as the time t increases.

On the other hand, as for the future data, in order to verify the result of learning using the past data, the future data is predicted using the estimated parameters, and the behavior of the prediction accuracy is investigated. Equation (27) is used for evaluation.

$$J2(m) = \sqrt{\frac{1}{n}\|y^*(t_0 + m) - y(t_0 + m)\|^2} \quad (m>0) \qquad (27)$$

Further, $y^*(i)$ is the estimated data (n-dimensional) at time i, and y (i) is the correct answer data (n-dimensional) at time i.

D. Experimental Results

In this experiment, we assume an autoregressive model with a degree of 1, and change p and q with respect to the SSM / I19.3GHz vertical polarization observation brightness data with dimension n = 5 and current time t0-100. We compared the learning process for past observation data and

the prediction behavior for future data between RLS method 1 and RLS method 2.

By changing p and q from one multidimensional time series, it is possible to generate multiple multidimensional time series including arbitrary missing data. This time, we generated four types of five-dimensional time series: (p, q) = (5,5), (5,10), (10,5), (10,10). The following similar experiments were performed on these four 5-dimensional time series. Based on these experimental results, it is possible to make a qualitative comparison between RLS method 1 and RLS method 2. RLS method 1 and RLS method 2 were used to train until the current time t0, respectively. J1 (t) was used as the evaluation function. Then, in order to verify the learning results of RLS method 1 and RLS method 2, future data was predicted using each parameter estimated at the current time t0, and the behavior of prediction accuracy was investigated. J2 (m) was used as the evaluation function. The prediction was made up to 50 years ahead (m = 50).

This experiment is for qualitative comparative verification of prediction methods (RLS method 1, RLS method 2) based on the RLS method for time series data including missing data. Therefore, it is considered that only 19.3 GHz vertical polarization observation luminance data (1 channel) is enough, and it is sufficient to consider an autoregressive model having an order of 1.
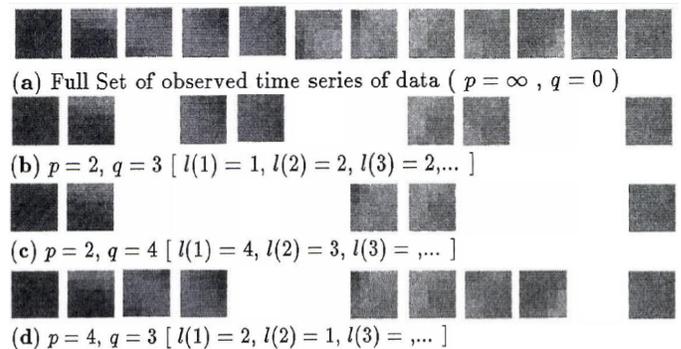
Fig. 5. Examples of the Time Series Including the Missing Data. The Parameter p is the Length of the Period when it is Observed. The Value *l(k)(k=1, 2, 3, …)* is the Length of the Non-Observed Period. (0<*l(k)*<q).
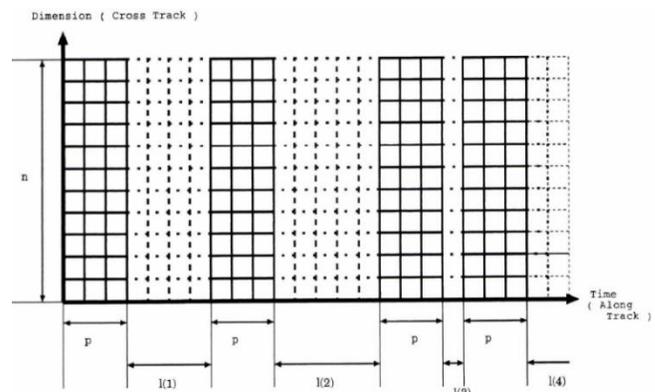
Fig. 6. The Representation of the Observed and the Non-Observed Data. (The Horizontal Axis shows Time Slots While the Vertical Axis shows the Number of Dimensionality which are Corresponding to the Along Track and the Cross Track Directions.) In this Example, Data are Observed for three Time Slots (p=3).

Fig. 7 shows the experimental results when p and q are changed, respectively. The figure on the left shows the transition of J1 (t), and the figure on the right shows the transition of J2 (m). In the figure on the left, the horizontal axis is displayed in the range of [0,100] and the vertical axis is displayed in the range of [-20,120]. In the figure on the right, the horizontal axis is displayed in the range of [1,50] and the vertical axis is displayed in the range of [0,160].
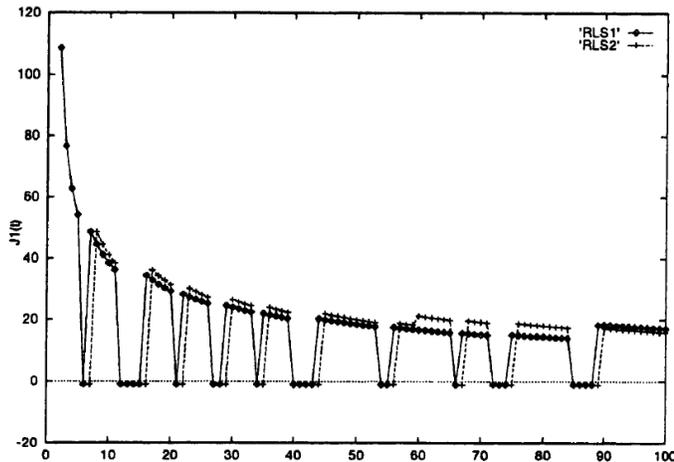
(a) J1: p=5, q=5.
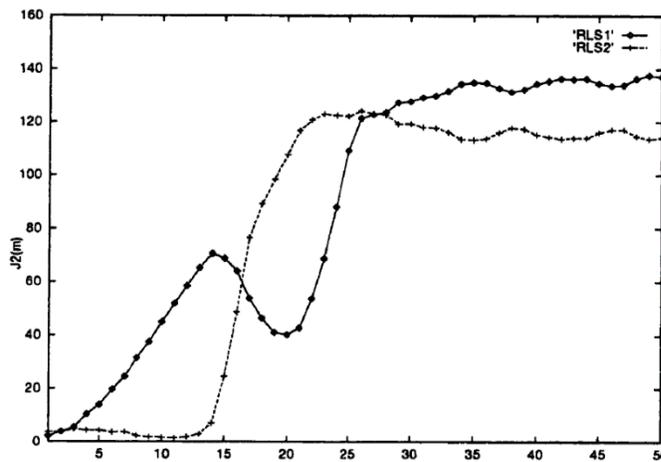
(b) J2: p=5, q=5.
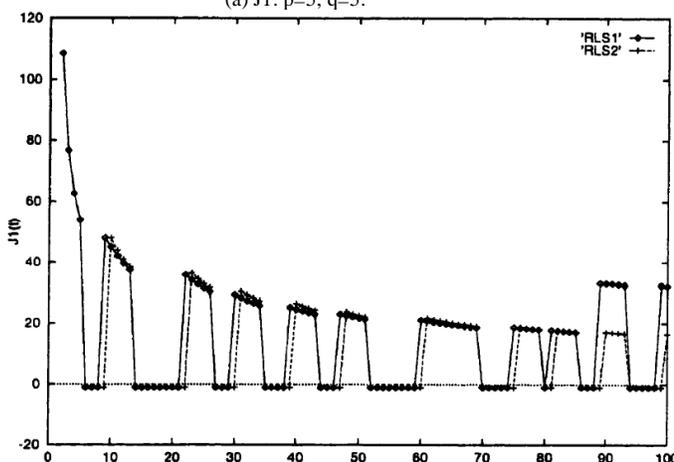
(c) J1: p=5, q=10.
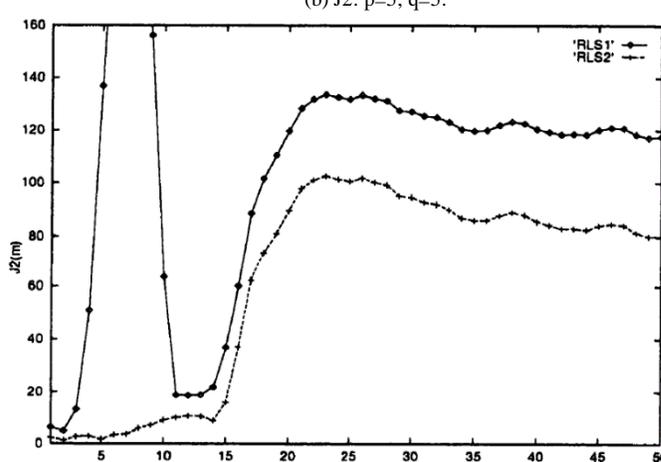
(d) J2: p=5, q=10.
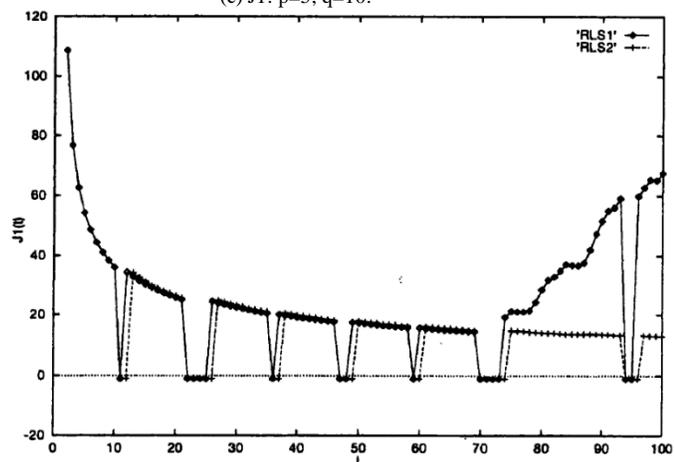
(e) J1: p=10, q=5.

(f) J2: p=10, q=5.

(g) J1: p=10, q=10.

(h) J2: p=10, q=10.
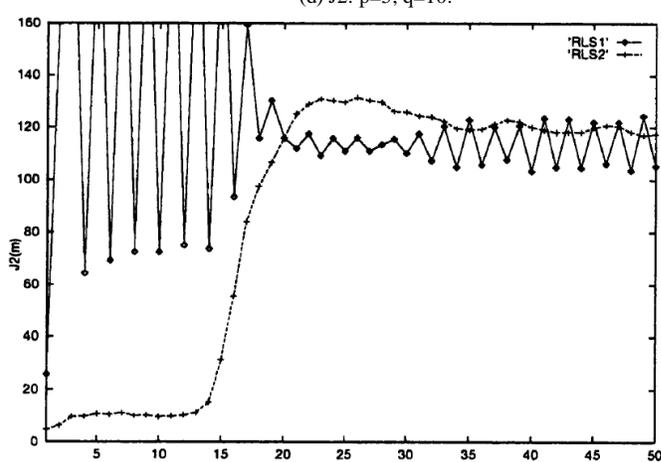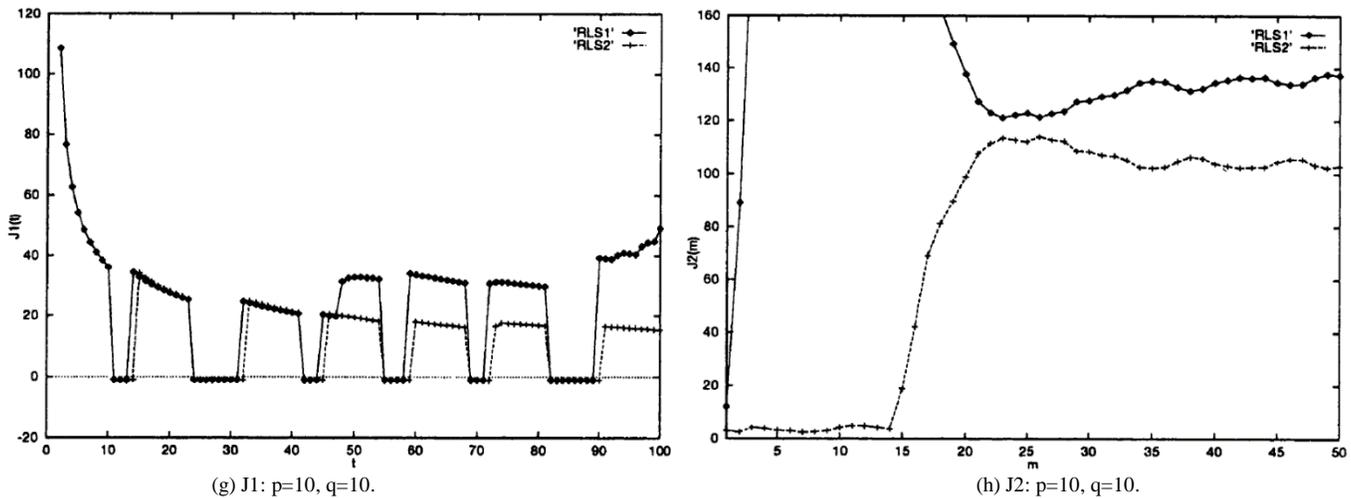
Fig. 7. The Transient Responses of J1(t) and J2(m) with the RLS Method #1 and the RLS Method #2 for (p, q)=(5,5), (5, 10), (10, 5), (10,10).

E. Remarks

In the comparison of the evaluation function J1 (t) between RLS method 1 and RLS method 2, the value of J1 (t) in RLS method 2 decreased in all four cases as t increased, whereas RLS decreased. Since the value of J1 (t) in method 1 may increase in some cases, the superiority of RLS method 2 can be confirmed.

Evaluation between RLS method 1 and RLS method 2 In the comparison of J2 (m), RLS method 2 shows almost the same characteristics of J2(m) with respect to rn in four cases, whereas the RLS method In some cases, 1 has drawbacks such as the behavior of J2 (m) with respect to rn causing vibration.

The reason why the behavior of J2 (m) with respect to m in RLS method 1 causes vibration is that the value of the evaluation relation J1 (t) of the learning process for the past observation data increased. For the above reasons, it can be said that RLS method 2 is superior to RLS method 1.

VI. Conclusion

In this paper, we proposed a method based on the Kalman filter as a method for estimating missing data from a multidimensional time series including missing data, or for estimating data at any time. At that time, we proposed two transformation methods of the sequential least squares method (RLS method), which has been widely used in the past as a parameter estimation learning method for the Kalman filter, and compared the two methods (RLS method 1, RLS method 2). The learning process for the past observation data and the prediction behavior for the future data were compared using the observation brightness temperature data by SSM / 1. In order to generate a multidimensional time series including arbitrary missing data and perform qualitative comparison verification between RLS method 1 and RLS method 2, the time axis is set in the direction of travel of the satellite with respect to the path-shaped observation data. The experiment was conducted with the dimension axis in the scanning direction.

In RLS method 1, learning is performed while performing a time update algorithm on past missing data, so it is difficult to perform effective learning when there are many missing data in the past. It is considered that there are cases.

In RLS method 2, learning is performed on past missing data without performing a time update algorithm, so it is possible to prevent a rapid increase in error due to estimation of many past missing data. Conceivable.

It was confirmed that RLS method 2 is suitable as a learning method by RLS method for a time series in which many missing data are continuously present.

VII. Future Research Works

Further research works are required for the other missing data consideration. There are some interpolation and extrapolation methods of the alternative methods for the proposed Kalman filter based method with RLS method.

References

[1] R.E. Kalman: A New Approach to Linear Filtering and Prediction Problems, Trans. ASME, J. Basic Eng., 82D-1, 35/45, (1960).

[2] Toru Katayama: "Applied Kalman Builta", Asakura Shoten, (1983).

[3] Eweda Eweda: Comparison of RLS, LMS, and Sign Algorithms for Tracking Randomly Time-Varying Channels, IEEE Trans. Signal Processing, 42-11, 2937/2944, (1994).

[4] Shinyo Matsuoka, Ryutaro Tateishi: "Study on Correction of Reflectance in AVHRR Data Using Time Series Model", Photogrammetry and Remote Sensing, 37-2, 4/14, (1998).

[5] Kohei Arai Prediction method for time series of imagery data in eigen space, International Journal of Advanced Research in Artificial Intelligence, 2, 1, 12-19, (2013).

[6] Kohei Arai, Geography Markup Language: GML based representation of time series of assimilation data and its application to animation content creation and representations, International Journal of Advanced Research in Artificial Intelligence, 2, 4, 18-22, 2013.

[7] Kohei Arai, Recovering method of missing data based on the proposed modified Kalman filter when time series of mean data is known, International Journal of Advanced Research in Artificial Intelligence, 2, 7, 18-23, 2013.

[8] Kohei Arai, Osamu Fukuda, Hiroshi Okumura, Kenji Endo, Kenichi Yamashita, Time Series Analysis for Shortened Labor Mean Interval od Dairy Cattle with the Data of BCS, RFS, Weight, Amount of Milk and Outlook, IJACSA, 9, 7, 108-115, 2018.

[9] Kohei Arai, Detecting Algorithm for Rainfall Area Movement based on Kalman Filtering, Proceedings of the NSAT/SWT Symposium, Kyoto, Nov. 1995.

[10] Kohei Arai, Rain Flagging with SSM/I Based on Kalman Filtering with New Parameter Estimation, Proceedings of the NSCAT Science Workshop, (1997).

[11] Kohei Arai, Rain Flagging for NSCAT with SSM/I Through Gap Filling Based on Kalman Filter, Proc. of the NASA Scatterometer Science Symposium, pp. 161-163, 1998.

[12] Kohei Arai and Kaname Seto Rain Flagging Method with Kalman Filtering for ADEOS/NSCAT, Proc. of the COSPAR Congress, A0.1-0009, 1998.

[13] Kohei Arai, Comparative Study on Image Prediction Methods between the Proposed Morphing Utilized Method and Kalman Filtering Method, International Journal of Research and Reviews in Computer Science (IJRRCS) Vol. 3, No. 6, 1875-1880, December 2012, ISSN: 2079-2557, 2012.

[14] Kohei Arai, Kaname Seto: "Time series analysis based on Kalman filter with parameter estimation considering non-linearity of system", Proceedings of the 24th Annual Meeting of the Remote Sensing Society of Japan, (1998).

[15] Kohei Arai, Kaname Seto: "Handling of missing data in parameter estimation of Kalman filter by RLS method", Proceedings of the 26th Annual Meeting of the Remote Sensing Society of Japan, (1999).

[16] Kohei Arai, Recovering method of missing data based on the proposed modified Kalman filter when time series of mean data is known, International Journal of Advanced Research in Artificial Intelligence, 2, 7, 18-23, 2013.

[17] Arai,K., Detecting Algorithm for Rainfall Area Movement based on Kalman Filtering, Proceedings of the NSAT/SWT Symposium, Kyoto, Nov. 1995.

[18] Kohei Arai, Rain Flagging with SSM/I Based on Kalman Filtering with New Parameter Estimation, Proceedings of the NSCAT Science Workshop, (1997).

[19] Kohei Arai, Rain Flagging for NSCAT with SSM/I Through Gap Filling Based on Kalman Filter, Proc. of the NASA Scatterometer Science Symposium, pp. 161-163, 1998.

[20] Kohei Arai and Kaname Seto Rain Flagging Method with Kalman Filterinmg for ADEOS/NSCAT, Proc. of the COSPAR Congress, A0.1-0009, 1998.

AUTHOR'S PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is a Science Council of Japan Special Member since 2012. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Science Commission "A" of ICSU/COSPAR since 2008 then he is now award committee member of ICSU/COSPAR. He wrote 55 books and published 620 journal papers as well as 450 conference papers. He received 66 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. http://teagis.ip.is.saga-u.ac.jp/index.html