Journal of Arid Environments 73 (2009) 557–562

Contents lists available at ScienceDirect

Journal of Arid Environments

journal homepage: www.elsevier.com/locate/jaridenv

NDVI prediction over Mongolian grassland using GSMaP precipitation data and JRA-25/JCDAS temperature data

H. Iwasaki*

Faculty of Education, Gunma University, Maebashi, 4-2 Aramaki, Maebashi, Gunma, 371-8510, Japan

ARTICLE INFO

Article history:
Received 31 March 2008
Received in revised form
11 December 2008
Accepted 12 December 2008
Available online 11 January 2009

Keywords:
Multiple linear regression equation
Plant litter
Reanalysis data
Semi-arid region
Steppe

ABSTRACT

An algorithm to predict the NDVI (Normalized Difference Vegetation Index) distribution over Mongolia, which is based on a stepwise multiple linear regression analysis, has been developed using global precipitation data obtained from satellites and global surface air temperature data obtained from the reanalysis data during the period 1998–2005. This algorithm can predict the NDVI value up to 1–3 months in advance for a grid with a spatial resolution of $0.25^\circ \times 0.25^\circ$.

In order to validate the algorithm, the NDVI distribution was predicted for the period from May to November 2006 using 1 to 3-month prediction algorithms. The distributions of the predicted normalized anomalies agreed well with those of the observed normalized anomalies. It was found that these algorithms were effective for arid and semi-arid regions, despite its low accuracy for August and regions with high vegetation activity.

1. Introduction

Mongolia is located in transition zone for vegetation, where it ranges from taiga forest in the north to desert in the south, corresponding to a decrease of annual precipitation amount. As shown in Fig. 1, the rangelands in Mongolia occupying 97% of the country are classified into five subtypes: a high mountainous region (4.5%), a forest steppe (23%), a steppe (28%), a desert steppe (28%), and a desert (16%) (Bolortsetseg et al., 2000). These rangelands are the main source of forage for nomadic livestock. Poor biomass production due to the drought in summer is the main cause of the large-scale loss of livestock in winter (the so-called dzud) (Morinaga et al., 2004). The seasonal prediction of biomass production would be of a great benefit to the Mongolian society since the productivity of grass in the rangeland intensively influences the lives of its people.

The relationship between meteorological parameters and NDVI (Normalized Difference Vegetation Index), which is a proxy for the biomass production, in arid region has been investigated by several researchers. Some studies have elucidated that a positive lag correlation exists between the precipitation anomaly in the antecedent month and the NDVI anomaly for the Mongolian grassland (Schultz and Halpert, 1995; Miyazaki et al., 2004; Iwasaki, 2006a,b; Shinoda et al., 2007). Iwasaki (2006a) demonstrated that the distribution of rainfall amount observed by weather radar corresponded well to the distribution of NDVI 1–2 months later over the arid region of Mongolia. Iwasaki (2006b) has shown that winter temperature is also correlated with the NDVI value in summer at several meteorological stations. His study indicated that the NDVI value in the developing stage and matured stage can be expressed as a linear function of the precipitation amount and air temperature at several meteorological stations using multiple linear regression equations. However, it will be useful if the predicted NDVI distribution over different types of regions, even the regions far away from the meteorological stations, is provided. In order to predict the NDVI distribution, instead of surface observation data, the use of global precipitation data obtained from satellites and the global surface air temperature obtained from the reanalysis data might be effective. The purpose of this study is to examine the possibility of NDVI prediction for the entire Mongolian grassland using the precipitation data from the satellites and the temperature from the reanalysis data.

2. Data

The algorithm to predict the NDVI distribution is based on a multiple linear regression analysis using the data from SPOT NDVI and GSMaP (Global Satellite Mapping of Precipitation) and the surface air temperature obtained from the JAR-25 (Japanese

* Tel./fax: +81 27 220 7297.
E-mail address: iwasaki@edu.gunma-u.ac.jp

0140-1963/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jaridenv.2008.12.007

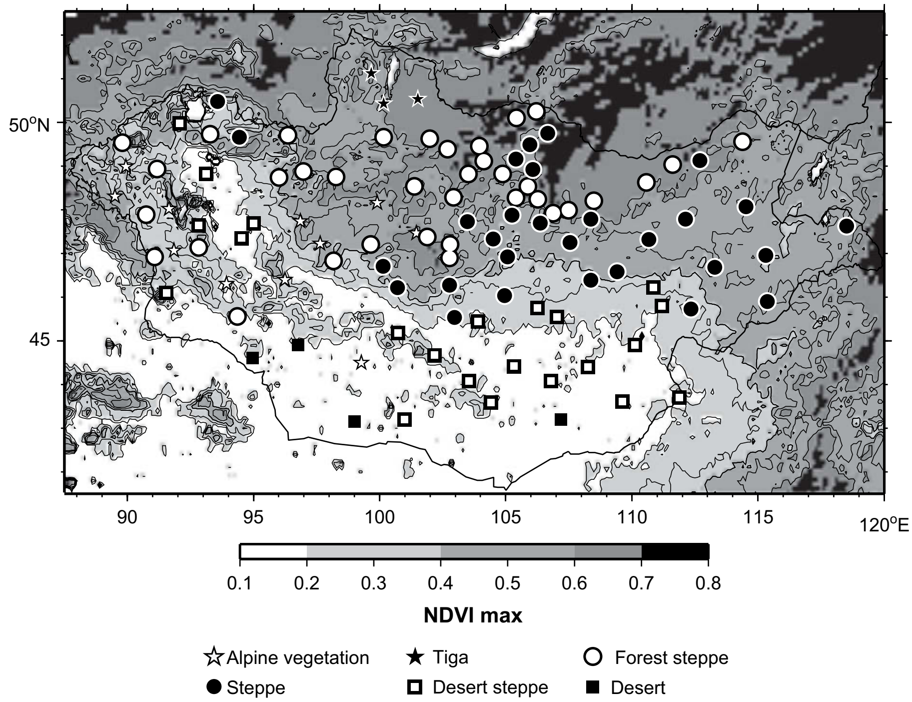

Fig. 1. Distribution of the maximum NDVI (NDVI_max) for the period 1998–2005 and the location of meteorological stations and vegetation type.

25-year reanalysis)/JCDAS (JMA Climate Data Assimilation System) reanalysis data for the period 1998–2005. The validation of the algorithm for NDVI prediction employs NDVI, GSMaP precipitation, and JCDAS temperature in 2006.

2.1. SPOT NDVI

The SPOT/VEGETATION data compiled every 10 days with a spatial resolution of 1 km are used as dependent variables for the multiple linear regression equations and validation data of the algorithm. The NDVI data are generally found to be well correlated to a fraction of photosynthetically active radiation absorbed by green vegetation and is a good indicator of the biomass production (e.g., Tucker, 1979). The NDVI value is used as a proxy for the biomass production. High-frequency noises due to clouds and signal noises in the 10-day composite NDVI data are removed using the BISE (Best Index Slope Extraction) method (Viovy et al., 1992). Since the small NDVI values can produce large errors in regions where the mean NDVI values are close to zero, only the regions with NDVI values of more than 0.1 are used in the calculation.

The SPOT NDVI data are compiled every month with a spatial resolution of $0.25^\circ \times 0.25^\circ$ to match the spatial resolution of the GSMaP precipitation data.

2.2. GSMaP precipitation data

The GSMaP_MWR product (Ver. 4.8.4) is used to estimate the global precipitation amount with a spatial resolution of $0.25^\circ \times 0.25^\circ$ by the application of the retrieval algorithm for brightness temperatures (Tbs) from satellite-borne microwave radiometers. The basic aim of the algorithm is to find the optimal precipitation amount for which the Tbs calculated by the radiative transfer model fits perfectly with the observed Tbs (Kubota et al., 2007). The algorithm has been improved based on the classification in terms of rain or no rain over the land (Seto et al., 2005), and the scattering algorithm has been improved using the 85- and 37-GHz polarization-corrected temperatures. The estimation of precipitation amount over the land is based on the scattering algorithm that employs the emission from liquid water and the scattering due to ice at 85 GHz.

The monthly precipitation data are used as independent variables for the multiple linear regression equations and validation data of the algorithm. Since the precipitation amount is one of limiting factors for NDVI value over the Mongolian grassland, the accuracy of NDVI prediction depends on the accuracy of GSMaP precipitation data. Fig. 2 shows the scatter diagram between the monthly GSMaP precipitation data and the monthly raingauge precipitation data obtained at the 97 meteorological stations shown in Fig. 1. The correlation coefficient and bias are 0.61 and +3.2 mm, respectively. The accuracy of the GSMaP precipitation data in an arid region is not good since the GSMaP algorithm has been developed for precipitation over the tropics and subtropics. It may cause a decrease in the accuracy of NDVI prediction.

2.3. Surface air temperature obtained from JRA-25 and JCDAS

The JRA-25 data are reanalysis data provided by JMA (Japan Meteorological Agency) and Central Research Institute of Electric Power Industry; the data was obtained during the period 1979–2004. JCDAS takes over the same system as JRA-25 and the data assimilation cycle, and are available from 2005 up to the present (Onogi et al., 2007). The surface air temperature at 2 m AGL obtained from the two types of reanalysis data is compiled every month with a spatial resolution of $0.25^\circ \times 0.25^\circ$, and its correlation coefficient with the data from the observation exceeds 0.98. The monthly data are used as the independent variables for the multiple linear regression equations and validation data of the algorithm.

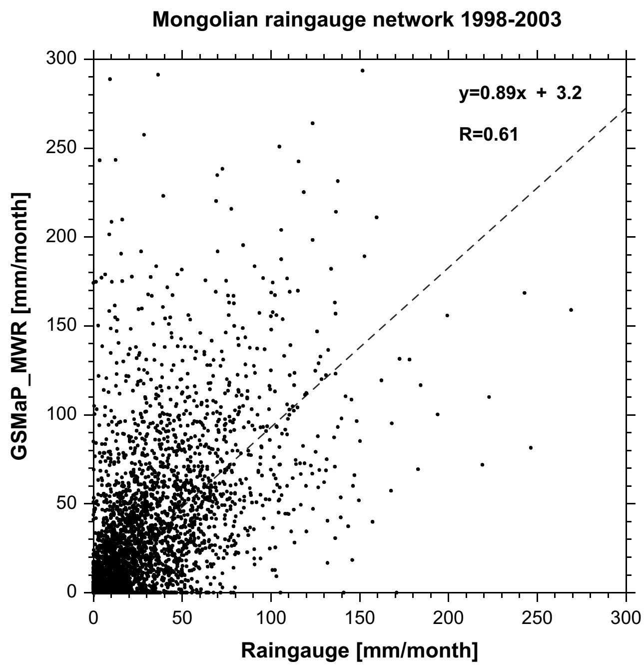

Fig. 2. Scatter diagram between monthly GSMaP precipitation data and raingauge precipitation data at the meteorological stations shown in Fig. 1 for the period 1998–2003.

3. Algorithm to predict NDVI distribution

3.1. Determination of multiple linear regression equations

The algorithm to predict the NDVI distribution is based on a multiple linear regression analysis using a stepwise method. The analysis region including the Mongolian grassland is divided into 6912 ($=144 \times 48$) grids with a spatial resolution of $0.25° \times 0.25°$. The NDVI value at each grid is expressed as a function of the precipitation amount and surface air temperature. For example, the multiple linear regression equations for 1-month prediction of August 2006 are determined by the stepwise method, using the NDVI value for August during the period 1998–2005 as dependent variables, and the GSMaP precipitation data and surface air temperature from November to July during the period 1998–2005 as independent variables (Fig. 3a).

In order to predict the NDVI value for one month ahead, the multiple linear regression equations are calculated using the independent variables obtained one month before the required period. The regression equations at the 95% significant level were adopted for this algorithm (the adjusted multiple correlation > about 0.7). Equation (1) is an example of the regression equation for a certain grid ($i = 21$, $j = 7$), and NDVI for August can be predicted using precipitation in July and temperature in May.

$$\mathrm{NDVI\_Aug} = 0.04 \times \mathrm{Prec\_Jun} + 0.002 \times \mathrm{Temp\_Mar} + 1.01 \quad (1)$$

Since the regression equations were determined by the stepwise method, precipitation terms and temperature terms with lower contribution or with high multiple collinearity were removed.

The multiple linear regression equations for 2 and 3-month predictions are also calculated by the same method.

3.2. Prediction of NDVI distribution

The NDVI distribution is predicted using the multiple linear regression equations and the data of precipitation and surface air temperature (Fig. 3b). For example, the NDVI values for August 2006 for a month ahead are calculated by substituting the GSMaP precipitation data and surface air temperature from November 2005 to July 2006 for the multiple linear regression equations for 1-month prediction of August. The NDVI distribution is obtained by calculating the NDVI values for all grids. The NDVI values for August 2006 for 2 and 3 months ahead are also calculated using multiple linear regression equations for 2 and 3-month predictions of August by the same method as 1-month prediction.

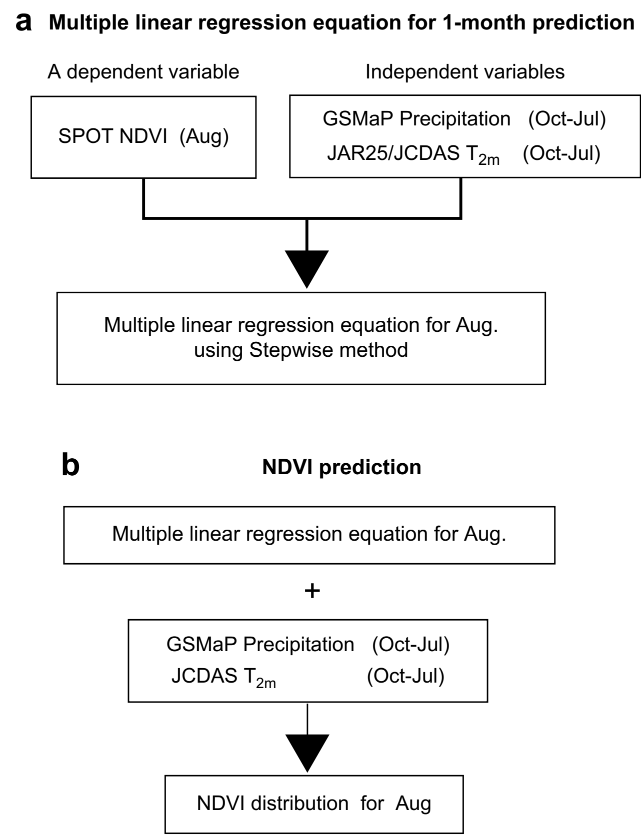

Fig. 3. Flowchart for NDVI prediction algorithm.

Occasionally, NDVI values that are significantly large or small are also predicted. If the predicted NDVI value is more (less) than the maximum (minimum) NDVI value for the period 1998–2005, the maximum (minimum) NDVI value is adopted as the predicted value. Except for this empirical constraint, all the predicted values are simply calculated by the process shown in Fig. 3b.

4. Results of NDVI prediction

The NDVI prediction experiment from May to November in 2006 over Mongolia was carried out using the multiple linear regression equations determined from the data obtained during the period 1998–2005 so as to validate the 1 to 3-month prediction algorithms. The distribution of observed and predicted NDVI for 1-month prediction is shown in Fig. 4a. The gray region in Fig. 4 indicates the grid in which a significant regression equation is not determined, and Fig. 5 shows the time series of the number of the valid grids for three cases. The number of the valid grids increases with an increase in time. As for 1-month prediction the significant regression equation is determined for 6032 grids (about 87%), however, the valid grids for 2 and 3-month predictions are lower in comparison with 1-month prediction, especially in May and June.

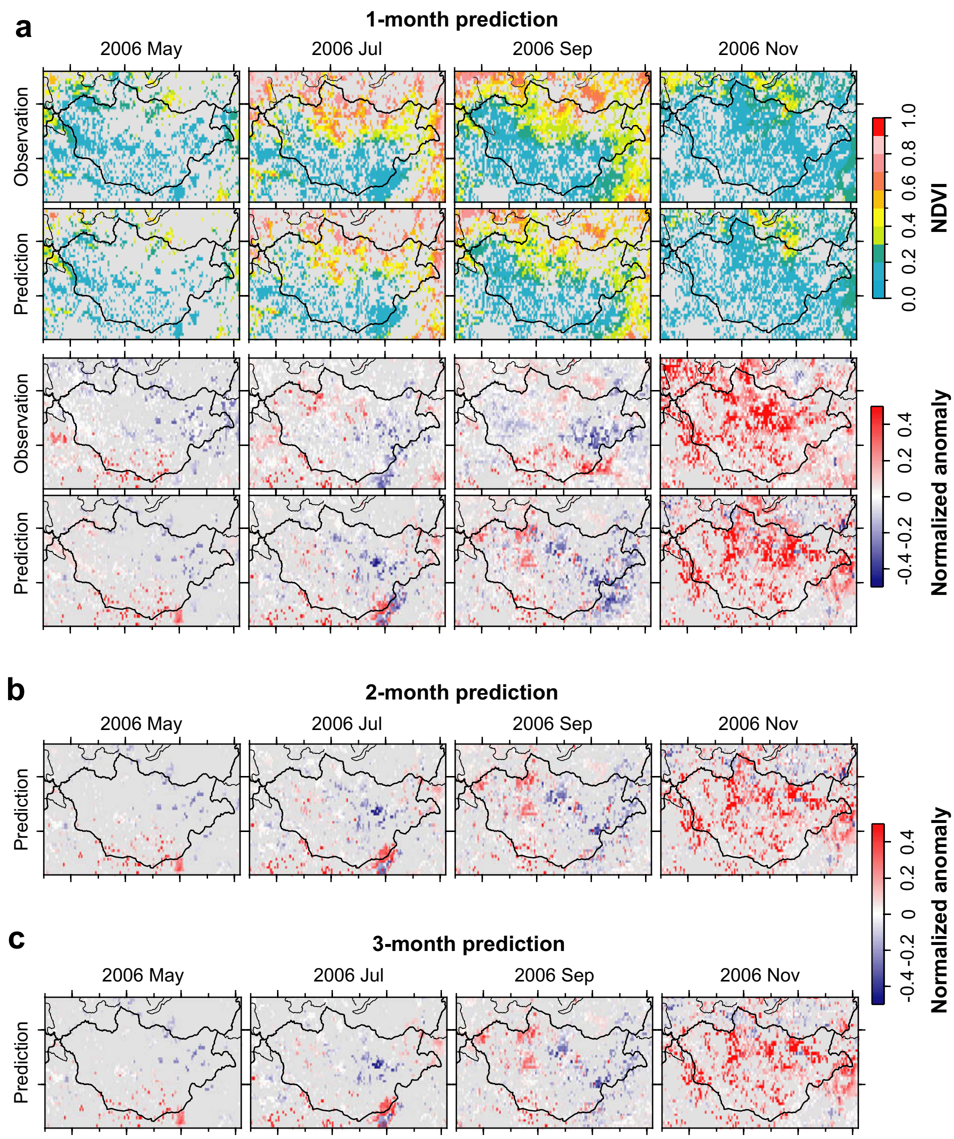

Fig. 4. Results of NDVI prediction from May to November 2006. (a) The observed and predicted NDVI distributions (top panels) and the observed and predicted normalized anomalies (bottom panels) for 1-month prediction. (b) The observed and predicted normalized anomalies for 2-month prediction. (c) The observed and predicted normalized anomalies for 3-month prediction. The gray region indicates the grid where a significant regression equation is not determined.

The difference between the NDVI values of the south and north Mongolia associated with the transition zone of vegetation and the seasonal development of the NDVI distribution are well predicted in Fig. 4a. Fig. 6a shows the relationship between observed NDVI and predicted NDVI for all grids from May to November, 2006. The correlation coefficient between the observed and predicted NDVI values for 1-month prediction is very high, that is, 0.97. Since the correlation coefficients for 2 and 3-month predictions are also 0.97 (Table 1), there is no significant difference in the accuracy of three prediction periods.

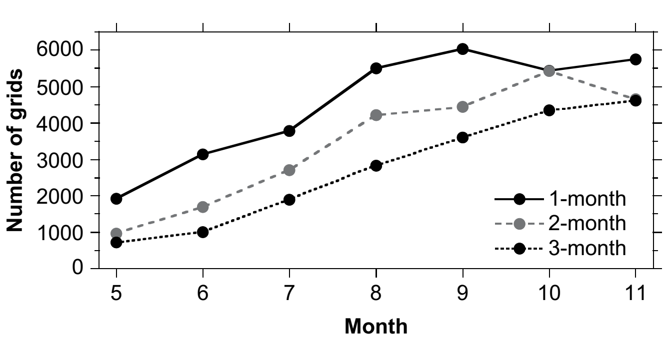

Fig. 5. Time series of the number of grids where significant multiple linear regression equations were obtained for 1 to 3-month predictions. Total number of grids for analysis region is 6912.

The accuracy of NDVI prediction should be evaluated using the NDVI anomaly. The bottom panels of Fig. 4a–c show the distribution of the observed and predicted anomalies, which are normalized by the monthly NDVI value averaged over 1998–2005. In spite of considerable difference in the number of valid grids (Fig. 5), qualitative features are common to the results of three prediction periods. That is, the positive anomalies around the southern parts of the analysis region from May to November and the predominant positive anomaly over entire Mongolia in November are well predicted. The negative anomalies in the eastern part of Mongolia from May to September are also well reproduced. On the other hand, the predicted anomaly is not consistent with the observed anomaly in

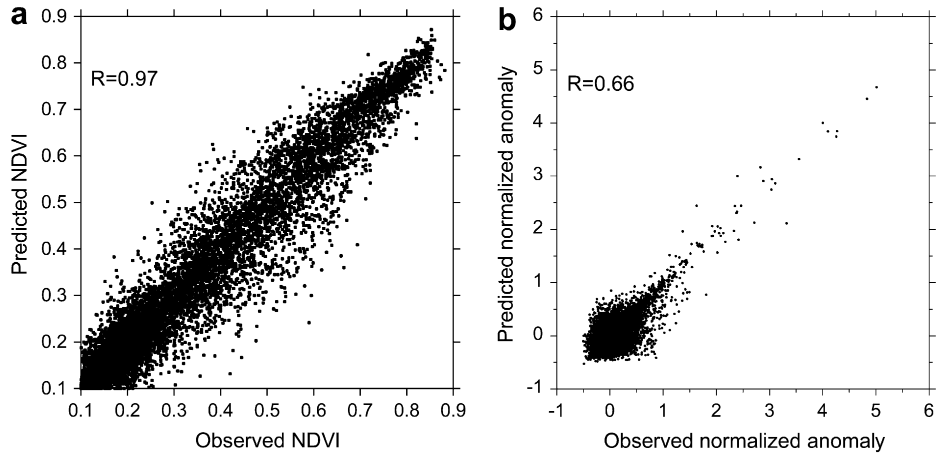

Fig. 6. Scatter diagram between the (a) observed and predicted NDVI and (b) observed and predicted normalized anomalies of NDVI.

the northern part of Mongolia from July to September, where the NDVI value is relatively high.

Fig. 6b shows the relationship between observed NDVI anomaly and predicted NDVI anomaly for all grids from May to November, 2006. For all the prediction region and months, the correlation coefficient between the observed and predicted normalized anomalies is found to be 0.66 for 1-month prediction, which is a significantly high value. The correlation coefficients for 2 and 3-month predictions are also high (Table 1), which means that the NDVI values can be predicted up to a few months in advance.

The accuracy depends on the climatological NDVI value. For simplicity, the analysis regions are classified into two regions using the maximum NDVI value for the period 1998–2005 (NDVI_max); the low-NDVI region where NDVI_max is less than 0.4 and the high-NDVI region where NDVI_max exceeds 0.4. As shown in Fig. 1, the low-NDVI region almost corresponds to the vegetation zone of a desert and a desert steppe, and the high-NDVI region almost corresponds to the vegetation zone of a forest steppe and the taiga forest. The vegetation zone of a steppe belongs to both these regions. Fig. 7 shows time series of the correlation coefficients between observed NDVI and the predicted NDVI using 1-month prediction for the high- and low-NDVI region. It is found that this NDVI prediction algorithm is very effective for the low-NDVI regions of the arid and semi-arid regions. This feature is also recognized in the results obtained by 2 and 3-month predictions (not shown), and it is consistent with results obtained by Iwasaki (2006b).

The accuracy also depends on the season. The correlation coefficients in spring and autumn are high; however, it is extremely low in August (Fig. 7). A possible cause for the low correlation in August is that the NDVI value in the matured stage (almost August) has a negative correlation with the air temperature in the matured stage (Iwasaki, 2006b). That is, the influence of summer temperature on the NDVI value does not have a large time lag. Since this NDVI prediction is based on the lag correlation, the NDVI value in August cannot be predicted with good accuracy. The poor biomass production due to summer drought is one of main causes for the large-scale loss of livestock in winter, so that, the low accuracy in August is a serious drawback and should be improved.

Table 1
Correlation coefficients between the observed value and the predicted value for 1 to 3-month predictions.

	1-month	2-month	3-month
NDVI value	0.97	0.97	0.97
Normalized NDVI anomaly	0.66	0.71	0.73

5. Discussion

5.1. Anomaly of NDVI in November

As shown in Fig. 4, the predicted normalized anomalies agree well with those of the observed normalized anomalies in November. It is hard to regard the NDVI values as the biomass production of green vegetation because monthly mean air temperature in November is less than −10 °C even in the southern part of Mongolia. What the anomaly of NDVI in November means will be discussed.

Since the difference of the reflectance in red and near infrared, which is used for calculating NDVI, is generally greater for plant litter than soils, typical NDVI values are also larger for litter (0.14–0.45) than soils (0.08–0.16) (McMurtrey et al., 1993). However, it is difficult to reliably distinguish plant litter from soils by using NDVI values, because litter may be brighter or darker than a particular soil and the reflectance spectra of soils depend on moisture conditions (Nagler et al., 2000).

Nagler et al. (2003) showed that as the coverage of plant litter increased on a specific soil with a constant reflectance, the difference of the reflectance in red and near infrared also increased; that is, it leads to the increase in the NDVI value even for the dead biomass. Since the interannual variation of the soil reflectance

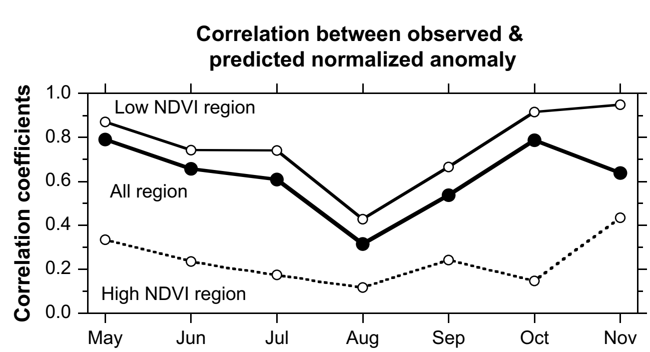

Fig. 7. Time series of correlation coefficient between the predicted and observed normalized anomalies for the entire analysis region (solid line), low-NDVI region (thin solid line), and high-NDVI region (dotted line).

spectra for each grid in November would be negligible under the low temperature condition, the amount/coverage of grass litter can contribute to the variation of NDVI anomaly in winter. The increase in the amount/coverage of grass litter over Mongolian grassland in November, 2006 is thought to be observed as the positive anomaly of NDVI, and this feature could be reproduced using the NDVI prediction algorithm.

5.2. Accuracy of GSMaP precipitation

According to the results of many studies, precipitation is a limiting factor for the vegetation activity over the Mongolian grassland. Iwasaki (2006b) has studied the effect of precipitation amount and air temperature on NDVI for the period 1998–2003, and shows that number of the meteorological stations which have significant lag correlations between precipitation and NDVI value is a few times larger than number of the meteorological stations which have significant lag correlations between air temperature and NDVI value.

A similar analysis has been carried out using the GSMaP precipitation data for the period 1998–2005. Fig. 8 shows the time series of the number of the grids with a high lag correlation with the precipitation amount and temperature. It is found that temperature is a limiting factor for NDVI in the analysis using the GSMaP precipitation data. This feature is contradictory to the results obtained by Iwasaki (2006b); this suggests strongly that the GSMaP precipitation data is not sufficiently accurate.

The low accuracy of the GSMaP precipitation data, as shown in Fig. 2, will decrease the accuracy of the NDVI prediction. However, the GSMaP project was started in November 2004, and the GSMaP algorithm has been developed for precipitation over the tropics and subtropics. It is expected that GSMaP precipitation algorithm for arid regions will be improved in the near future through the GPM (Global Precipitation Measurement) project, and it would lead to a high accuracy in the NDVI prediction.

6. Summary

An algorithm to predict the NDVI distribution up to 1–3 months in advance over Mongolia, which is based on multiple linear regression analysis using the stepwise method, has been developed. The analysis region is divided to 6912 grids with a spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$. A multiple linear regression equation for each and every grid was determined using the SPOT NDVI data for the period 1998–2005, which were used as dependent variables, and the GSMaP precipitation data and the surface air temperature obtained from the reanalysis data for the period 1998–2005, which were used as independent variables.

The distribution of the monthly mean NDVI from May to November 2006 was predicted up to 1–3 months in advance in order to validate the prediction algorithm. The distribution of the predicted normalized anomalies agreed well with that of the observed normalized anomalies. A comparison between the observed and predicted normalized anomalies for each and every grid showed a correlation coefficient of 0.66, 0.71 and 0.73 for 1, 2 and 3-month predictions, respectively. The accuracy depended on the season and the climatological NDVI value; the correlation coefficients were low in August and for areas with a high-NDVI value. It was found that the algorithm was effective for arid and semi-arid regions since the correlation coefficients for low-NDVI regions were significantly high.

Because this algorithm uses the global data set of NDVI, precipitation data, and air temperature, the technique can be easily applied for any arid and semi-arid regions.

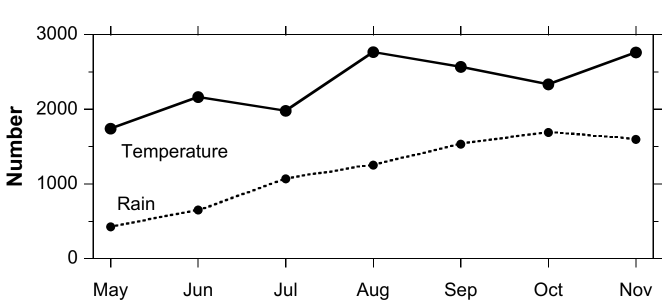

Fig. 8. Time series of the number of grids with a high lag correlation with precipitation data and temperature.

Acknowledgments

This research has been supported by the Sumitomo Foundation. The author would like to thank all members of the JRA-25 and JCDAS execution teams for providing the reanalysis data, all members of the GSMaP project for providing the GSMaP precipitation data, and Dr. O. Kouzan of Kyoto University for providing the BISE code.

References

Bolortsetseg, B., Bayasgalan, S.H., Dorj, B., Natsagdorj, L., Tuvaansuren, G., 2000. Impact on agriculture. In: Batima, P., Dagvadorj, D. (Eds.), Climate Change and its Impacts in Mongolia. JEMR Publishing, Ulaanbaatar, pp. 96–198.

Iwasaki, H., 2006a. Study on influence of rainfall distribution on NDVI anomaly over the arid regions in Mongolia using an operational weather radar. SOLA 2, 168–171.

Iwasaki, H., 2006b. Impact of interannual variability of meteorological parameters on vegetation activity over Mongolia. Journal of the Meteorological Society of Japan 84, 745–762.

Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y.N., Nakagawa, K., Iwanami, K., Ushio, T., Kachi, M., Okamoto, K., 2007. Global precipitation map using satelliteborne microwave radiometers by the GSMaP Project: production and validation. IEEE Transactions on Geoscience and Remote Sensing 45, 2259–2275.

McMurtrey, J.E., Chappelle, E.W., Daughtry, C.S.T., Kim, M.S., 1993. Fluorescence and reflectance of crop residue and soil. Journal of Soil and Water Conservation 48, 207–213.

Miyazaki, S., Yasunari, T., Miyamoto, T., Kaihotsu, I., Davaa, G., Oyunbaatar, D., Natsagdorj, L., Oki, T., 2004. Agrometeorological conditions of grassland vegetation in central Mongolia and their impact for leaf area growth. Journal of Geophysical Research 109, D22106.

Morinaga, Y., Bayarbaator, L., Erdenetsetseg, D., Shinoda, M., 2004. Zoo-meteorological study of cow weight change in a forest steppe region of Mongolia. Proceedings of the Sixth International Workshop on Climate Change in Arid and Semi-Arid Region of Asia, Ulaanbaatar, Mongolia, pp. 100–108.

Nagler, P.L., Daughtry, C.S.T., Goward, S.N., 2000. Plant litter and soil reflectance. Remote Sensing of Environment 71, 207–215.

Nagler, P.L., Inoue, Y., Glenn, E.P., Russ, A.L., Daughtry, C.S.T., 2003. Cellulose absorption index (CAI) to quantify mixed soil-plant litter scenes. Remote Sensing of Environment 87, 310–325.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., Taira, R., 2007. The JRA-25 Reanalysis. Journal of the Meteorological Society of Japan 85, 369–432.

Schultz, P., Halpert, M.S., 1995. Global analysis of the relationship among a vegetation index, precipitation, and land surface temperature. International Journal of Remote Sensing 16, 2755–2777.

Seto, S., Takahashi, N., Iguchi, T., 2005. Rain/no-rain classification methods for microwave radiometer observations over land using statistical information for brightness temperatures under no-rain conditions. Journal of Applied Meteorology 44, 1243–1259.

Shinoda, M., Ito, S., Nachinshonhor, G.U., Erdenetsetseg, D., 2007. Phenology of Mongolian grasslands and moisture conditions. Journal of the Meteorological Society of Japan 85, 359–367.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment 8, 127–150.

Viovy, N., Arino, O., Belward, A.S., 1992. The Best Index Slope Extraction (BISE): a method for reducing noise in NDVI time-series. International Journal of Remote Sensing 13, 1585–1590.