Integrated model for predicting rice yield with climate change**

Jin-Ki Park[1], Amrita Das[2], and Jong-Hwa Park[2] *

[1]National Institute of Crop Science, Rural Development Administration, Mirayang-si, Gyeongnam, South Korea, 50424
[2]Department of Rural and Agricultural Engineering, Chungbuk National University, Seowon-gu, Cheongju,
Chungbuk, South Korea, 28644

A b s t r a c t. Rice is the chief agricultural product and one of the primary food source. For this reason, it is of pivotal importance for worldwide economy and development. Therefore, in a decision-support-system both for the farmers and in the planning and management of the country's economy, forecasting yield is vital. However, crop yield, which is a dependent of the soil-bio-atmospheric system, is difficult to represent in statistical language. This paper describes a novel approach for predicting rice yield using artificial neural network, spatial interpolation, remote sensing and GIS methods. Herein, the variation in the yield is attributed to climatic parameters and crop health, and the normalized difference vegetation index from MODIS is used as an indicator of plant health and growth. Due importance was given to scaling up the input parameters using spatial interpolation and GIS and minimising the sources of error in every step of the modelling. The low percentage error (2.91) and high correlation (0.76) signifies the robust performance of the proposed model. This simple but effective approach is then used to estimate the influence of climate change on South Korean rice production. As proposed in the RCP8.5 scenario, an upswing in temperature may increase the rice yield throughout South Korea.

K e y w o r d s: rice yield, artificial neural network, normalized difference vegetation index, spatial interpolation, crop yield models

INTRODUCTION

There is an increasing demand for a crop yield model (CYM) among the policy makers and farmers in South Korea so as to forecast production before the actual harvest in the current year, as well as in the distant future under the cloud of climate change. Rice, which is the chief crop in South Korea, and also the staple food of all Koreans, regard-less of income level, is crucial for the country's economic health and perspective. Therefore, South Korean management specialists need a simple but accurate estimation procedure for rice yield. Statistical CYMs are traditional and widely accepted techniques for forecasting the yield of numerous crops throughout the world (Gornott and Wechsung, 2016 Houseman, 1942; Rose, 1936; Thompson, 1962). These provide a quantitative estimation of expected yield, in advance, based on historical statistical information on climate and crop yield. The two most frequently used datasets for CYMs are climate data and certain vegetation related information (Shi *et al.*, 2013). The NDVI or normalized difference vegetation index has been taken as an input variable in many agricultural and ecological studies to reflect the healthy growth of plants, as well as the soil conditions (Halder and Patnaik, 2010; Prasad, 2007; Sarma *et al.*, 2008; Savin and Isuev, 2010). The NDVI is also very popular in large area crop management and yield studies (Bouman, 1995; Ko *et al.*, 2015). The degree of correlation between NDVI and crop yield mainly depends on crop types and time of acquisition of NDVI data (Mosleh *et al.*, 2015). Major climatic variables, used as model inputs for rice, are air temperature, humidity, precipitation and available solar radiation in various combinations (Crasta and Cox, 1996; Wang *et al.*, 2014). However, the use of meteorological data has several implications that include the uneven spatial distribution of ground stations and differences in units of time, *i.e.* some are available on an hourly basis, some daily, others as a monthly average. Unfortunately, all the data necessary for modelling crop yield are seldom available at the same spatial and temporal scale. Moreover, diverse topography creates a strong micro-climatic variation in the

Korean peninsula that increases the requirement for proper data pre-processing and scaling-up techniques. Past studies have shown that a systematic error of ± 1°C day^{-1} during the course of a single crop cycle could lead to errors of 0.5 to 1 t ha^{-1} in yield prediction by crop models (Monestiez *et al.*, 2001). GIS and spatial interpolation techniques are two basic methods that can be used to address the issue (Priya and Shibasaki, 2001; Strand, 1981). The primary role of GIS lies in the bridging of the remote sensing derived vegetation indices, with ground-based information and statistical data (Nuarsa *et al.*, 2012; Patel *et al.*, 1991). It acts as an aid to extract pixel based information from remote sensing imagery and to homogenise the input parameters in terms of spatial resolution before modelling (Prasad *et al.*, 2007; Shen *et al.*, 2009). Linear regression has been used in several past studies for CYMs. However, the main disadvantage of regression-based studies is their inefficiency to address the non-linearity in the system which is apparent in crop response to agro-climatological conditions. The artificial neural network (ANN) is thus considered to be one of the best computational methods for extracting information from imprecise and non-linear conditions. ANN models are a more versatile modelling approach and allow designing much more complex and non-linear relationships without the rigorous computations and assumptions regarding the distribution and interconnections of input variables. ANN models can also identify, as well as simultaneously learn, correlation patterns between input parameters and corresponding target values through self-adaptive training (Bishop, 1995; Zhang *et al.*, 1998).

Therefore, the primary objective of the present study is to build an accurate rice forecasting model using ANN for South Korea with special emphasis on scaling up input variables in the light of spatial interpolation and GIS techniques. An attempt has also been made to evaluate the effect of climate change in rice yield over South Korea.

DATA AND METHODOLOGY

Historical records of rice production and climatological variables over South Korea and MODIS NDVI data were used to develop the yield model. Land cover data published by the Ministry of Environment were used to extract paddy field areas. Topographical parameters were required for data pre-processing in the cases of temperature and solar radiation. The details of the data are given in Table 1. A new set of algorithms developed by Takeuchi and Yasuoka (2004), has been successfully used to reduce the effect noise in MODIS-NDVI image. In creating these, they had analysed MODIS spectrum data for channel 1 to 7 using the radiative transfer simulation model and found that the blue wavelength is mostly subjected to the atmospheric effect, moreover, pixels with cloud shadow in them are cooler than normal pixels. These two criteria, *i.e.* minimum blue and maximum thermal criteria (TminB) were used to identify and modify the reflectance value of other channels in clouded pixels. Filtered MODIS-NDVI data and image (10 days composite) covering almost the whole of East Asia are available on the WebMODIS website (http://web-modis.iis.u-tokyo.ac.jp).

Four meteorological factors (temperature, rainfall, solar radiation and sunshine duration) were used for the present study. There are 456 automatic weather stations and 81 meteorological stations continuously measuring the temperature and rainfall over South Korea. Sunshine duration and solar radiation data are, however, only available in 81 and 22 meteorological stations, respectively. All the historical

Table 1. Details of the Variable used in present study

Parameters		Unit	Source	Time period	
				Model development	Validation
	NDVI	NA	WebMODIS	2001-2011	2012
Meteorological data	Temperature	°C	Korea Meteorological Administration	2001-2011	2012
	Rainfall	mm			
	Sunshine hours	NA			
	Solar radiation	MJ m^{-2}			
Land cover		NA	Ministry of Environment	NA	NA
Topography		NA	National Geographic Information Institute	NA	NA
Rice yield		kg 10^{-1}a	Korean Statistical Information Service	2001-2011	2012

meteorological data were made available for the purposes of the study by the Korea Meteorological Administration (http://www.kma.go.kr).

The land cover map of South Korea was provided by the Ministry of Environment (https://egis.me.go.kr) based on satellite imagery of Landsat, IRS and Kompsat 2 and 3. The data was found to be more than 95% accurate in both cases during on-site verification of randomly chosen sample pixels and when compared to reference data. DEM data for the whole country in a 30 m grid is available on the website of NGII (http://www.ngii.go.kr). Three topographic data parameters (slope, aspect and elevation) are extracted and used in the present study.

Annual rice production data have been collected from each province after harvesting and published by the National Statistics Office on every 5th November since 1965. The rice production data are collected over randomly distributed 931 thousands of sample area, each consisting 2 ha of rice field, from all over the country. Yield data is provided as provincial average and presented on the website of the National Statistical Office (http://kosis.kr/).

The artificial neural network (ANN) is a method which mimics the capacity of learning through a biological neural system and performs better than a regression system in identifying the non-linearity between variables. It finds relationships by observing a large number of input and output examples and develops a formula that can be used for prediction. In our study, we used a 3-layer ANN model with one hidden layer. The input and output layer contains nodes that correspond to input and output variables, respectively. Data moves between layers across weighted connections. The transfer function used here is a sigmoidal function for the hidden and output layers and a linear transfer function for the input layers.

For comparison with the ANN model, we also used the same input parameters for a linear regression analysis through employing the following equation:

$$Y_t = \beta_0 + \sum_{i=1}^{l} \beta_i Z_{i,t} + \varepsilon_t, \tag{1}$$

here: Y_t is the rice yield for year t, β_0 is the intercept, β_i is the regression coefficient for the Z_i variable and ε is residual. The estimated yield and error parameters are compared with the results of the ANN model.

The Republic of Korea occupies the southern portion of the Korean Peninsula. It lies between latitudes 33°N and 39°N, and longitudes 124°E and 130°E, covering a total area of 100,032 km². South Korea's terrain is mostly mountainous and most of it is not arable. Slope average in South Korea is 14.3°, and lowland area with less than 5° of slope is only 28.2%. The country has a wet temperate climate with a distinctive seasonal trend. Cool and dry wind from north-west decreases the ambient temperature throughout the winter season, while south-east wind brings ample amount of moisture causing high rainfall during summer. Although precipitation and cloudy skies can be observed throughout the year, rainfall reaches the highest values during the monsoon months from June to September. Temperature reaches the lowest values during January and highest during July. In Korea, crop yield data are available at the provincial level, whereas weather data only exist for weather stations distributed unevenly throughout the country. Moreover, whereas weather data are available on a daily basis, NDVI data that reflect crop condition are available as a 10 day composite. During the last few decades, GIS and remote sensing based methods were developed for forecasting rice production, particularly, in Asiatic countries (Chen *et al.*, 2011; Jing-feng *et al.*, 2002; Rahman *et al.*, 2012; Wang *et al.*, 2010). Most of these methods were developed on the basis of exploiting the remote sensing based spectral indices like NDVI. However, many are lacking in pre-processing the climatic data. Scaling up all the input parameters on the spatial and temporal levels is very important for successful modelling. Depending on agricultural and climatic conditions, Choi and Yun (1989), divided South Korea into 19 agro-meteorological zones, which Na *et al.* (2012) modified to obtain 9 major zones as shown in Fig. 1. Each of these zones. consisting of several smaller provinces, has its distinctive weather and soil characteristics. As historical rice yield data had been integrated on the provincial level, all the input parameters were scaled up
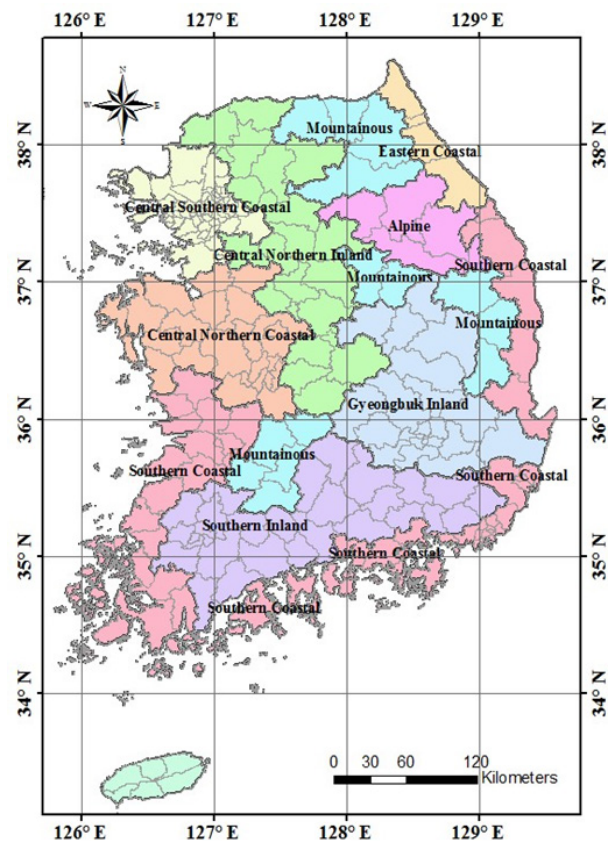

Fig. 1. Agro-climatic zones of South Korea.

for each provincial zone and then an individual ANN based model was generated for each of the climatic zones using spatial interpolation. To homogenise on a temporal level, average and/or summed up data for August 1 to October 10, were used as model primary input. The time period was chosen based on the NDVI study described later. The schematic diagram of the model is given in Fig. 2. A regression based model has also been studied with the same input parameters to compare the performance of the ANN-based model. The details of the methods applied to the input variables are given below.

Due to a huge difference in the resolutions of MODIS-NDVI (250 m) and the land cover image (30 m), in case of mixed pixel covering other features along with paddy fields, NDVI value was found to be different from the pure pixel with only the paddy field in it. For example, as shown in Fig. 3, the NDVI value of grid 1 with rice field alone is 0.68. This is much lower than grids 2, 3 and 4 with impurities in them. Such effect can surely introduce error in the measurement of average NDVI. To resolve this issue, the pixels, with only paddy fields in them, in MODIS-NDVI images, were extracted using the land cover map as reference and the NDVI values of only those pixels were used for model development. For example, in Fig. 3, from pixels 1 to 5, only pixel 1 will be chosen. Next, the extracted pixels are analysed for noise and error. Although Web-MODIS data is already filtered for inherent noise, there are fluctuations with noteworthy high and low values due to the presence of cloud and atmospheric turbidity. Scientists used a number of methods *viz.* Fourier-based method,

Asymmetrical Gaussian function-fitting, Savitzky-Golay filter, and Mean Value Iteration (MVI) to reduce noise in NDVI imagery. Each of the above-mentioned approaches holds its own advantages and disadvantages (Hird and McDermid, 2009). Most have complex and time consuming algorithms. Among them, MVI is preferred due to its simplicity and accuracy (Ma and Veroustraete, 2006). In the present study, NDVI data of ith date was replaced with the average of $(i+1)^{th}$ and $(i-1)^{th}$ data if it was lower than the average as given in the following equation, as well as depicted in Fig. 4:

$$NDVI_i = \begin{cases} \frac{NDVI_{i-1} + NDVI_{i+1}}{2}, & NDVI_i < \frac{NDVI_{i-1} + NDVI_{i+1}}{2} \\ NDVI_i & NDVI_i \geq \frac{NDVI_{i-1} + NDVI_{i+1}}{2} \end{cases} \quad (2)$$

The MVI method was used for each of the 26,198 pixels in each 10-day composite NDVI image from 2001 to 2012. The mean NDVI for South Korea increased from 0.26 ± 0.17 before correction, to 0.61 ± 0.09 after applying MVI filtration. The variation of average corrected and non-corrected NDVI from 2001 to 2012 is depicted in Fig. 5. It is clear from the figure that the sudden low in NDVI is due to the presence of clouds. When, other noises are removed, the corrected NDVI becomes a better reflection of the rice cultivation cycle over Korea.

In a temperate climate, the NDVI value has been found to vary with the maturity of plant leaves in progressive seasons. Therefore, in many crop phonological studies, a cyclic change in NDVI value has been observed. Likewise,
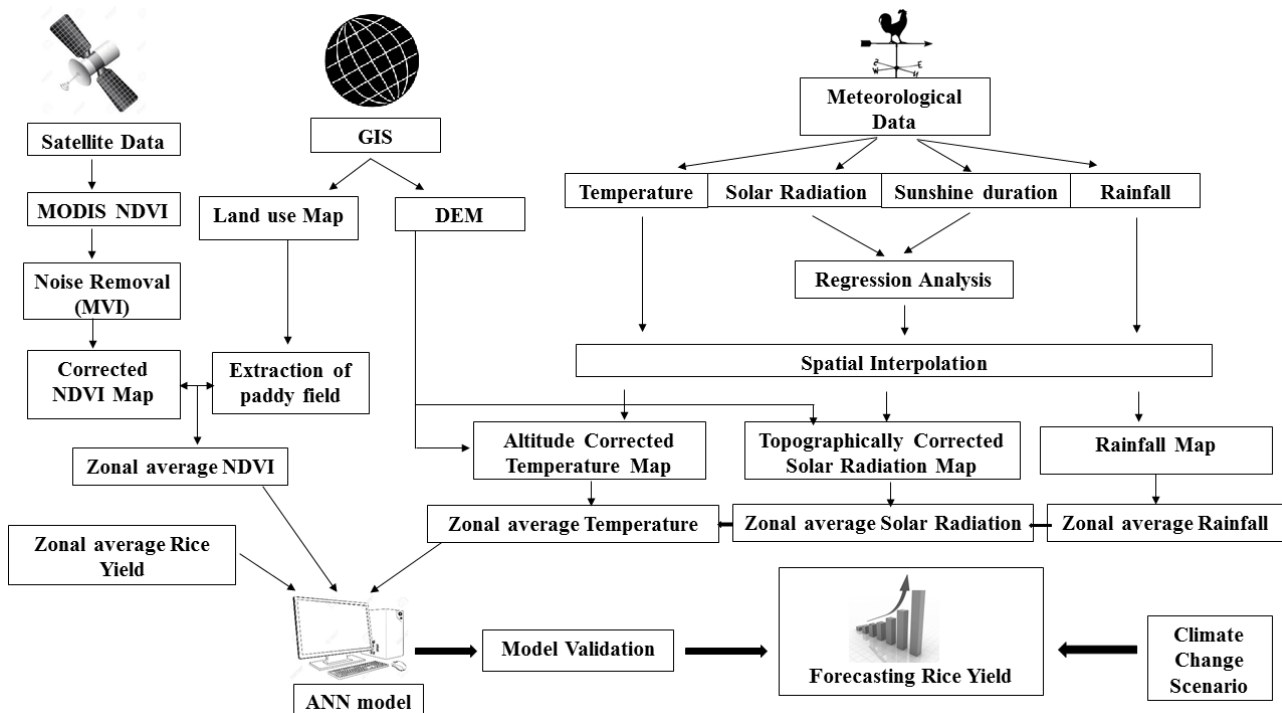

Fig. 2. Schematic diagram of model development.

the relationship between NDVI and rice yield strongly depends on the acquisition time of the NDVI image (Chen *et al.*, 2012). Thus, before being used as a model input, the NDVI over South Korea at different time periods was correlated with the rice yield. In South Korea, rice cultivation starts in spring *i.e.* in the middle of March, and seedlings are transplanted to freshly flooded fields during the rainy season, usually in May. Generally, the harvesting season starts with the beginning of October. The highest correlation, 0.33 with a p value less than 0.01, has been found with average NDVI from August 1 to October 10, and so it was used as one of the input variables in the ANN model. The NDVI integral over the growing season has also been successfully used as an indicator by many researchers in the past (Rasmussen, 1992; Tucker *et al.*, 1980).
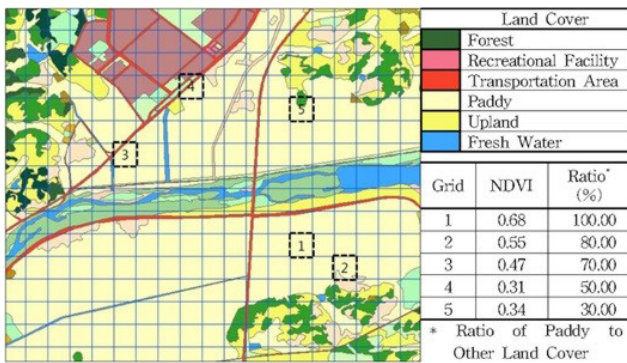

Fig. 3. Land-use map and different value of NDVI by ratio of paddy to other land cover.
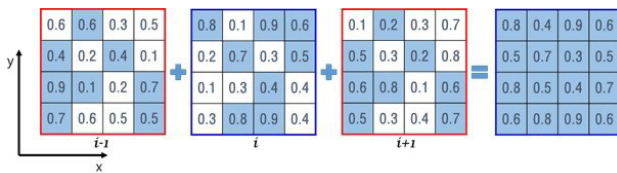

Fig. 4. Corrected NDVI using the MVI method.

Temperature is one of the meteorological factors that is extremely influenced by the elevation. It also acts as a limiting and deterministic factor in agricultural yield. Lower average temperatures indicate less moisture drawn out of the soil through evaporation, but also more extreme cold during winter. Lookingbill and Urban (2003), suggested a local model for estimating temperature changes in a complex topography. Ishida and Kawashima (1993), used the cokriging method to predict the changes in surface air temperature with elevation. In general, studies on thermal and radiation environment across a divergent topography used the DEM data (McCune, 2007). Maximum, minimum and accumulated temperatures from 536 stations across the country have been taken for our study and an altitude corrected map of maximum, minimum and accumulated temperatures were made using the spatial interpolation technique (Fig. 5). Initially, the measured temperature (T_m) was converted to sea level temperature (T_{sl}) using the following equations (Yun *et al.*, 2001):

$$T_{sl} = T_m + (Station\ altitude)\ \text{km} \times 8.1. \qquad (3)$$

These data were used in turn to create a sea level temperature map of South Korea by applying the Ordinary Kriging (OK) method. Then with DEM (altitude) data, the sea level temperature was again converted to corrected temperature or estimated temperature (T_{est}) (Fig. 6). The process was applied to all three temperature variables *i.e.* accumulated, average minimum and average maximum temperatures:

$$T_{est} = T_{sl} - (DEM.altitude)\ \text{km} \times 8.1. \qquad (4)$$

There are only 22 solar radiation measuring stations in South Korea. Such small number can increase the error when used for spatial interpolation. Therefore, to increase the accuracy of the available solar radiation map, an empirical model to estimate global solar radiation (GSR) from
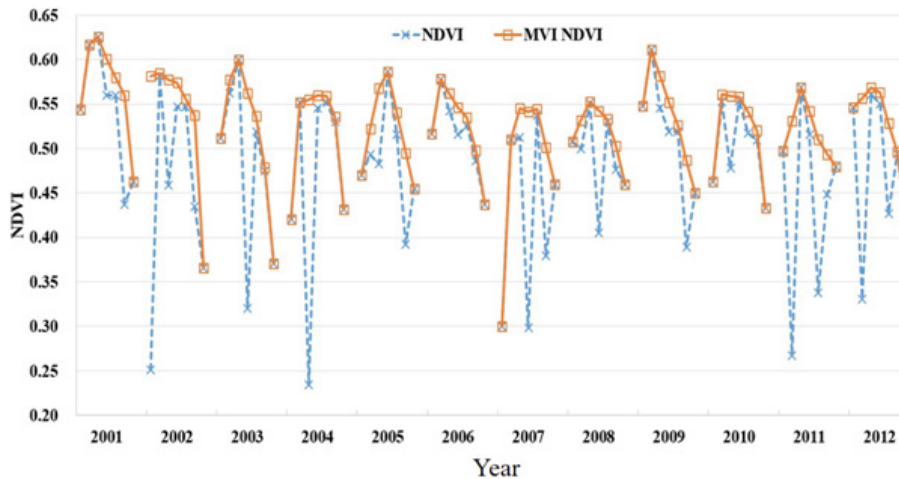

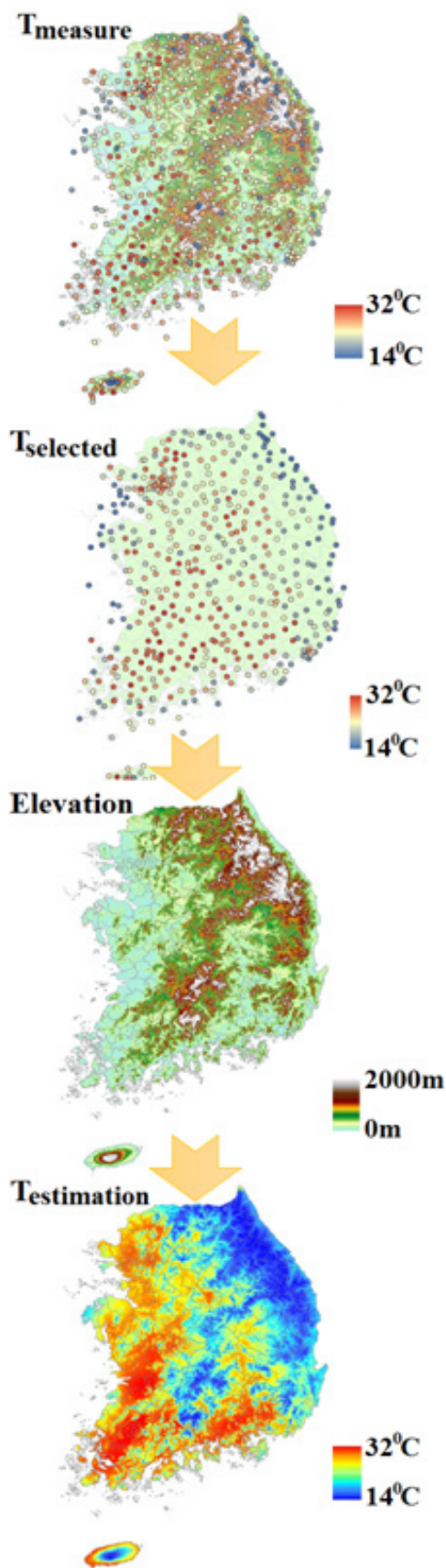**Fig. 5.** Time series average NDVI before and after applying MVI.

Fig. 6. Procedure of spatial distribution of temperature.

sunshine hours and topographic factors was introduced. The details of the model are discussed by Park *et al.* (2015). The average solar radiation of each province extracted from the map was used for model input.

Inaccuracy in the spatial distribution of rainfall data can also significantly increase the errors in the end results of a CYM. In a study, Pizza (2011) compares the different techniques that can be used for the spatial interpolation of rainfall data, and concluded univariate methods like Ordinary Kriging were the best for prediction. Ly *et al.*, 2013 reviewed different studies on spatial interpolation of rainfall and showed the importance of choice of interpolation method. For monthly and annual rainfall, the kriging method was found to be superior to others (Tabis and Salas, 1985; Abtew *et al.*, 1993) and hence it was used in the present study.

In order to meet the worldwide population expansion, the yields of food crops such as rice need to increase according to the rising demand. However, the changes in the global climate may make this goal more challenging to meet. Different approaches have been employed to evaluate the impacts of climate change on crop yields. In most cases, they have examined the effect of the future climate change on crop yields using a crop model and climate change scenarios. This approach has also been adopted in this study. Herein, the climatic variables of temperature, rainfall and solar radiation of RCP 4.5 and 8.5 are taken for forecasting rice yield in South Korea. In our work, a representative concentration pathway (RCP) of 4.5 is used as a stabilisation scenario and assumes that climate policies, like the introduction of a pricing method for greenhouse gas emissions, will achieve the goal of limiting emissions and concentrations, and, hence, radiative forcing (Clarke, 2007). In contrast, an RCP of 8.5 corresponds to a high greenhouse gas emissions pathway compared to literature (Riahi, 2007), and, therefore, also represents the upper bound of the RCPs. The greenhouse gas emissions and concentrations in the RCP 8.5 scenario increase considerably over time, leading to a radiative forcing of 8.5 W m^{-2} at the end of the century.

RESULTS AND DISCUSSION

Due to non-homogeneity in available meteorological stations, the climatic variables were first subjected to spatial interpolation. Mean values of each of the 137 unit areas have been taken as input for the ANN model. Instead of using one variogram model, different types of variograms have been tested for each variable in each year. Among these, the best performing model was selected based on the least RMSE as shown in Table 2. The highest percentage error has been found in the case of accumulated rainfall, but it is less than 25% and thus is within acceptable range. The larger error can be attributed to the patchy distributional pattern of rainfall over South Korea. On the other

Table 2. Error parameters after Kriging interpolation

Year	Variogram	MAPE (%)	RMSE	Variogram	MAPE (%)	RMSE
	Accumulated temperature (°C)			Mean maximum temperature (°C)		
2012		2.15	44.96	Gaussian	2.30	0.86
2011		2.16	44.21	Gaussian	2.45	0.92
2010		2.07	74.17	Circular	2.63	1.94
2009		2.00	43.44	Gaussian	1.93	0.69
2008		1.97	41.36	Circular	1.97	071
2007		2.13	38.24	Spherical	1.76	0.66
2006	Exponential	1.75	44.15	Spherical	2.14	0.77
2005		1.96	43.32	Gaussian	1.97	0.72
2004		1.92	43.92	Spherical	2.01	0.72
2003		2.50	47.06	Circular	2.08	0.74
2002		2.35	47.93	Spherical	2.16	0.73
2001		2.20	47.15	Gaussian	2.15	0.77
	Mean minimum temperature (°C)			Accumulated rainfall (mm)		
2012	Exponential	3.45	0.90	Gaussian	12.00	123.19
2011	Exponential	6.86	1.84	Exponential	17.92	75.89
2010	Exponential	3.39	1.07	Exponential	15.97	151.87
2009	Exponential	3.99	0.95	Exponential	13.31	47.78
2008	Spherical	3.70	0.90	Gaussian	18.90	66.54
2007	Exponential	2.60	0.71	Gaussian	13.38	127.29
2006	Exponential	3.82	0.98	Gaussian	25.38	68.19
2005	Spherical	2.85	0.76	Gaussian	15.78	97.55
2004	Spherical	3.58	0.90	Exponential	14.60	124.94
2003	Spherical	3.41	0.86	Exponential	13.83	120.77
2002	Exponential	3.72	0.89	Exponential	15.01	135.43
2001	Exponential	3.80	0.89	Spherical	15.98	57.31

hand, percentage error is relatively low in the case of all temperature related parameters. It is also noteworthy that the standard deviation of accumulated rainfall is more than 50% from its national average, whereas it is only 5% in case of accumulated temperature. Accumulated solar radiation was estimated using the sunshine based model of Park *et al.* (2015).

The ANN model 'build up' steps include the creation of a training and testing dataset, training multiple networks with varied model parameters and analysing the results. The model was fed with data from 2001 to 2011, 70% of which were used for training and the remaining 30% for testing purposes. The training and testing data were chosen automatically and randomly by machine. Data of

Table 3. Importance of input data according to weighting coefficients in different agro-climatic zones

Zones	NDVI	Acc_Rad	Max_Temp	Min_Temp	Acc_Temp	Range_Temp	Acc_Rain
GI	7	3	4	2	1	5	6
SI	2	6	1	5	4	3	7
SC	6	7	3	2	5	1	4
EC	5	3	7	2	6	1	4
AP	4	7	3	2	6	1	5
MT	6	5	3	1	4	2	7
CNC	2	6	1	4	5	3	7
CNI	6	7	5	2	3	1	4
CSC	7	1	6	3	2	4	5

Weighting coefficient = high to low: 1 to 7, GI – Gyeongbuk Inland, SI – Southern Inland, SC – Southern Coastal, EC – Eastern Coastal, AP – Alpine, MT – Mountainous, CNC – Central Northern Coastal, CNI – Central Northern Inland, CSC – Central Southern Coastal.

2012 were used as holdout set for model validation. The IBM SPSS ANN multilayer perceptron (MLP) model uses a supervised learning technique and has a feedforward architecture. MLP procedure, although it takes a longer time, is better equipped to find complex relationships. It predicts the error for each input parameters and then propagates the output error backward in the network and repeats the process until it reaches a model of minimum error (Jain *et al.*, 1996).

In network architecture, the hyperbolic tangent method was used to calculate the weighting coefficient for both input and output layers. No lower boundary of learning rate was set, but the threshold of error was kept to 0.0001, with an error ratio 0.001. MLP with different hidden layers were tested, but ANN with one hidden layer showed as being the best procedure for each climatic zone when analysed in terms of required time and prediction error. All climatic variables have their own influence in grain yield and are rigorously discussed in literature (Chen *et al.*, 2015; Deng *et al.*, 2015; Krishnan *et al.*, 2011; Yoshida, 1981), thus, the weighting coefficients of ANN models are different in case of different zones (Table 3).

It is clear from the table that different variables are taking the main role in influencing the ultimate production of rice. For example, for the mountainous region (MT), the rice yield is mostly influenced by minimum temperature and least affected by rainfall. In fact, in most of the zones, temperature range and minimum temperature play the main role in predicting yield. With advanced technology, the availability of fertilisers and better irrigation practices, the inter-annual variation of rice yield along with spatial gradient are less attributable to plant growth (NDVI) and rainfall. This is reflected in smaller weighting coefficients

in Table 3. The results also support our approach of exploring zonal models instead of one general model, otherwise the information about the variability in climatic influence would lose its importance.

The best ANN model (Zone: EC) structure is given in Fig. 7. The performance results of all ANN based and linear regression based models are given in Table 4. The performance of ANN-based models are far better than linear regression-based models, although there were a number of studies on predicting yield using regression analysis in the past that had shown good results. Still, it is also a fact that the relationship between grain yield and climatological conditions is hardly linear. Drummond *et al.* (1995) compares several statistical models for predicting yield and concludes that less complex models such as multiple linear regression (MLR) are not enough to explain the yield variability. Moreover, in recent studies comparing these two methods, *i.e.* ANN and MLR, ANN has shown the better performance in most of the cases – as in our case (Ji *et al.*, 2007; Paswan *et al.*, 2013; Yam, 2015).

It is clear from Table 4 that not only did the ANN model predict rice yield more precisely than regression analysis, but a low error and high R^2 also validate the novel approach of combining ANN with spatial interpolation that is presented in this paper. Spatial interpolation using DEM to generate meteorological surface has already been tested for rice yield in South Korea and has shown significant improvement in yield prediction (Yun, 1999, 2003). However, this study was lacking in two ways, it did not include any indicator for plant growth or health, and the used CYM was more complex, time consuming and required special experienced handlers. ANN models are more simple and easier to reproduce. When integrated with
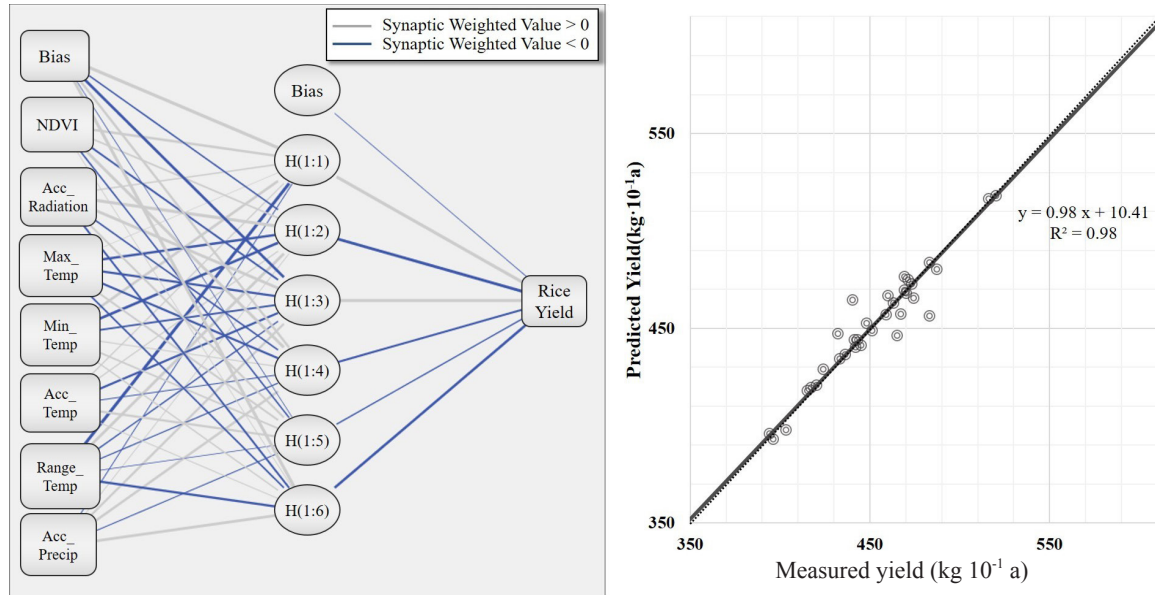
Fig. 7. Structure and yield estimation of best ANN model over EC.

Ta b l e 4. Results of ANN models and regression analysis in estimating rice yield within South Korea

Agro-climatic zone	Cross-validation			Independent variable (2012)		
	RMSE	MAPE	R^2	RMSE	MAPE	R^2
Error parameters for ANN model						
GI	20.52	3.20	0.76	20.49	3.23	0.74
SI	21.07	3.53	0.68	21.53	3.60	0.66
SC	28.40	4.64	0.58	29.57	4.91	0.60
EC	8.44	1.14	0.97	8.14	1.11	0.98
AP	6.32	0.96	0.94	6.14	0.93	0.94
MT	20.76	3.49	0.78	20.47	3.43	0.78
CNC	17.74	2.91	0.73	17.57	2.88	0.72
CNI	20.40	3.48	0.70	20.34	3.46	0.69
CSC	17.32	2.68	0.73	17.21	2.67	0.73
Error parameters for regression analysis						
GI	32.75	5.29	0.39	47.15	8.88	0.29
SI	29.42	4.80	0.37	32.25	5.44	0.18
SC	34.95	5.83	0.37	59.50	10.80	0.19
EC	35.92	7.27	0.54	48.17	11.59	0.96
AP	24.71	4.29	0.19	27.17	5.40	0.002
MT	36.66	6.23	0.33	30.90	5.34	0.30
CNC	30.73	5.21	0.19	13.01	2.41	0.03
CNI	30.21	5.04	0.34	24.83	4.29	0.08
CSC	27.63	4.30	0.32	26.88	4.11	0.01

MAPE in % and RMSE in kg 10^{-1} a.

remote sensing-GIS based methods, such statistical methods are seen to be more accurate (Prasad, 2007; Sarma, 2008; Savin and Isaev, 2010). The percentage error is lowest in the Alpine zone and highest in the Southern Coastal zone. Moreover, the regression coefficient is highest in the case of the Eastern Coastal zone.

The impact of global climate change on agriculture has been studied extensively for various crops at many different scales (Dhungana *et al.*, 2006; Fu *et al.*, 2016; Kurukulasuriya and Shane, 2013). However, relatively few studies have focused on rice, the foremost food crop in Asia. Tow field-based crop models ORIZA and SIMRIW predict a decrease in rice yield with an increase in temperature and an increase in yield when only the carbon-dioxide level becomes raised (Kim *et al.*, 2013; Vaghefi *et al.*, 2013). The present model is thus run with two standard climate change scenarios (RCP4.5 and RCP8.5) to evaluate the effect on rice production. Changes in different climatic parameters in South Korea, according to the two scenarios, are depicted in Fig. 8.

There is actually no significant trend that is observable in the case of solar radiation as well as precipitation. Obviously, temperature shows a remarkable rise with time especially in case of scenario 8.5. However, increasing temperature acted differently in different agricultural zones

in influencing rice yield as seen in Fig. 9a. Still, the trend is not statistically significant in many cases. The influence of precipitation and solar radiation nullifies the effect of increased temperature, and, therefore, the forecasted yield as shown in Fig 9b is almost random rather than showing an increasing or decreasing pattern. Yet, in the long term, a decreasing trend in rice production can be observed in the case of scenario 4.5. Furthermore, a higher increase in minimum temperature and accumulated temperature can increase the rice production of the country – as observed in the case of scenario 8.5. Similar to the past studies (Lee *et al.*, 2012), we saw that changes in yield are more effective in the eastern part of the country than in other regions.

It should be underlined that the effect of enhanced CO_2 is not considered here, which can complicate the forecasting system. It is very difficult to anticipate the effect of climate change in any biological system. This is because its interaction with the physio-chemical environment is extensively complex and hard to reproduce (Kurukulasuriya and Shane, 2013; Perkins, 2010). Still an attempt is made here to do so via ascertaining, albeit in an oversimplifying manner, if there is a statistical relationship between climate and crop production - if only to view a glimpse of the future and take necessary steps to adopt to the changes. So far,
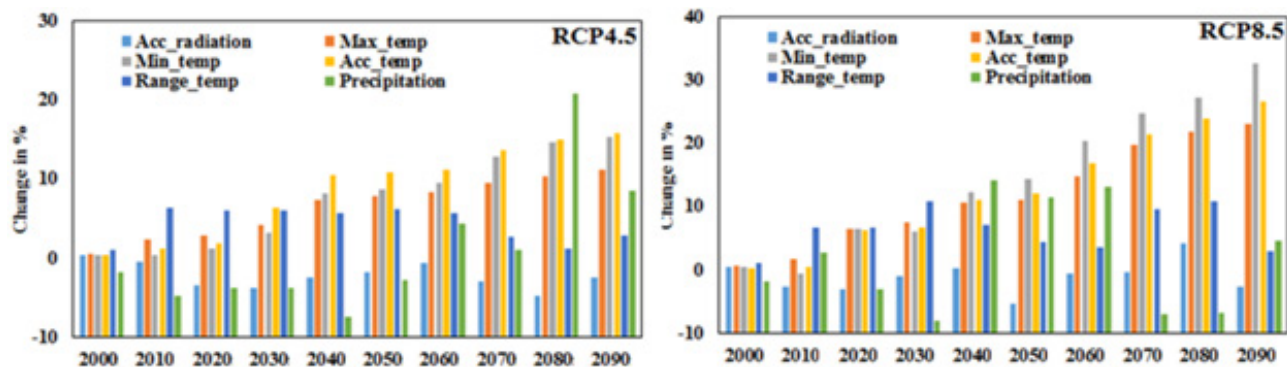

Fig. 8. Changes in input parameters from measured average according to RCP4.5 and RCP8.5 scenarios.
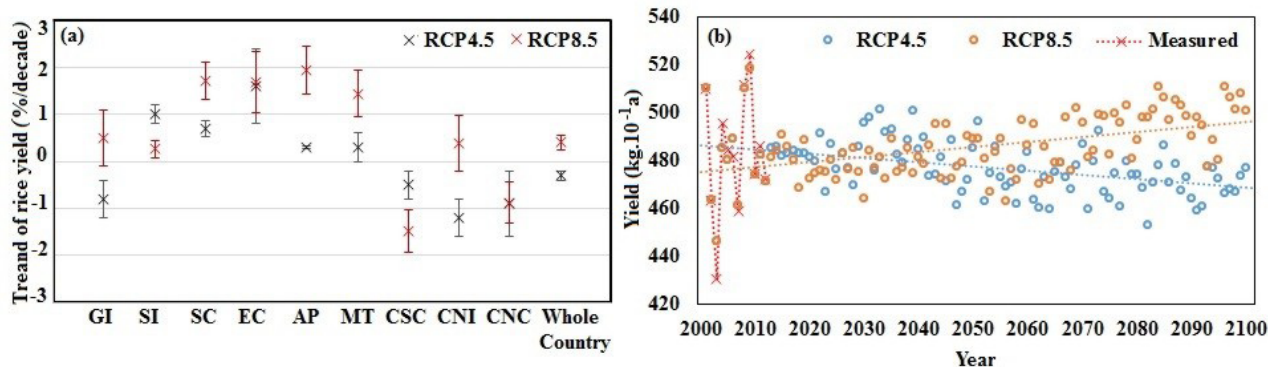

Fig. 9. Trend of rice yield in different agricultural zones (a) and forecasted average yield for the whole country in the case of climate change scenarios 4.5 and 8.5 (b).

although random, it can be said that increasing temperature at a higher rate may have a positive effect on future rice production in South Korea.

CONCLUSIONS

1. In the present study, ANN-based localised models have been developed to estimate rice production within South Korea. The main feature of our study is the combination of the spatial interpolation technique with the statistical crop model.

2. Scaling-up all climatic input is very important for the precision of model output specifically for countries like South Korea that feature high topographical variability. In addition to the complex terrain, spatial variation in soils, climate and management practice is far more complex in South Korea than in other major rice producing areas of the world. Hence, instead of one general model, the whole country is divided into 9 agro-climatic zones, and specific regional models were developed for each zone. Two major issues, inhomogeneity in climate and uneven distribution of the weather stations are addressed using GIS techniques.

3. Again to eliminate noise, the NDVI value of pixels with only paddy fields in them were chosen for study, and then again, the NDVI data of each of these pixels for each 10 days composite from 2001 to 2012 are analysed and treated using the MVI method.

4. Thus, when designing the model, every precautions has been taken to reduce the sources of noise and error in each input variable and in the full model development procedure. The efforts were reflected in the low percentage error and the high correlation between observed and modelled rice yield. It also justified the validity of the novel approach presented in this paper in forecasting the yield.

5. There is a growing concern among the scientific community, as well as in in the farming community and the general population about the influence of changing climate on growth and production of major world crops such as rice. Scientists argue, however, that with adapting mechanisms like alternative management practice and developing or shifting to new crop genotypes can reduce the effect of climate change on agricultural production. Still, it is a prerequisite to have an idea of the pathway of changes in the future before incorporating any adaptive practice. Therefore, the present model is used against the future climate change scenarios 4.5 and 8.5 to predict the changes in rice yield over South Korea. According to our model, higher temperature as postulated by scenario 8.5 may have a positive effect on rice yield within South Korea.

6. While the model was seen to be successful for the present study, for prediction of yield with changing climate, more improvement in methodologies is required. The model should include more sophisticated issues such as soil characteristics and changes in CO_2 level when it comes to forecasting yield for the far future.

Conflict of interest: The Authors do not declare conflict of interest.

REFERENCES

Abtew W., Obeysekera J., and Shih G., 1993. Spatial analysis for monthly rainfall in South Florida. J. Am. Water Resour. Assoc., 29(2), 179-188.

Basso B., Ritchie J.T., Pierce F.J., Braga R.P., and Jones J.W., 2001. Spatial validation of crop model for precision agriculture. Agric. Sys., 68, 97-112.

Bishop C.M., 1995. Neural networks for patterns recognition, Oxford University Press, London, 1995.

Bouman B.A.M., 1995. Crop modelling and remote sensing for yield prediction. Netherlands J. Agr. Sci., 43(2), 143-161.

Chen C., Quilang E.J.P., Alosnos E.D., and Finnigan J., 2011. Rice area mapping, yield, and production forecasting for the Province of Nueva Ecija using RDARSAT imagery. Canadian J. Remote Sensing, 37, 1-16.

Chen C.R., Chen C.F., and Son N.T., 2012. Detecting rice crop phenology from time series MODIS data. 33rd Conf. Remote Sensing, November 26-30, Pattiya, Thailand.

Chen H., Wu W., and Liu H.B., 2015. Assessing the relative importance of climate variables to rice yield variation using support vector machines. Theor. App. Climatol., 1-7.

Choi D.H. and Yun S.H., 1989. Agricultural climate zone classification and meteorological disaster characteristics (in Korean). Crop Science, 34, 13-33.

Clarke L., Edmonds J., Jacoby H., Pitcher H., Reilly J., and Richels R., 2007. Scenarios of Greenhouse Gas Emissions and Atmospheric Concentrations. Sub-report 2.1A of Synthesis and Assessment Product 2.1 by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research. Department of Energy, Office of Biological and Environmental Research, Washington, DC, USA.

Crasta O.R. and Cox W.J., 1996. Temperature and soil water effects on maize growth, development yield and forage quality. Crop Sci., 36(2), 341-348.

Deng N., Ling X., Sun Y., Zhang C., Fahad S., Peng S., Cui K., Nie L., and Huang J., 2015. Influence of temperature and solar radiation on grain yield and quality in irrigated rice system. Eur. J. Agron., 64, 37-46.

Dhungana P., Eskridge K.M., Weiss A., and Baenziger P.S., 2006. Designing crop technology for a future climate: An example using response surface methodology and the CERES-Wheat model. Agric. Syst., 87, 63-79.

Drummond S.T., Sudduth K.A., and Birrell S.J., 1995. Analysis and correlation methods for spatial data. ASAE Paper No. 951335, St. Joseph, MI, USA.

Fu T., Ha1 B., and Ko J., 2016. Simulation of CO_2 enrichment and climate change impacts on soybean production. Int. Agrophys., 30, 25-37.

Gornott C. and Wechsung F., 2016. Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany, Agri. Forest Meteorol., 217, 89-100.

Halder D. and Patnaik C., 2010. Synergistic use of multi-temporal Radarsat SAR and AWiFS data for rabi rice identification. J. Indian Society of Remote Sensing, 38, 153-160.

Hird J.N. and McDermid G.J., 2009. Noise reduction of NDVI time series: An empirical comparison of selected techniques. Remote Sensing of Environment, 113, 248-258.

Houseman E.E., 1942. Methods of Computing a Regression of Yield on Weather, Agricultural Marketing Service. United States Department of Agriculture Bulletin, 302, Iwoa, USA.

Ishida T. and Kawashima S., 1993. Use of cokriging to estimate surface air temperature from elevation. Theor. App. Climat., 47, 147-157.

Jain A.K., Mao J., and Mohiuddin K.M., 1996. Artificial Neural Networks: A Tutorial", Michigan State University, USA.

Jing-Feng H., Shu-chuan T., Abou-Ismail O., and Ren-chao W., 2002. Rice yield estimation using remote sensing and simulation model. J. Zhejiang University Science, 3(4), 461-466.

Kim H.Y., Ko J., Kang S., and Tenhunen J., 2013. Impacts of climate change on paddy rice yield in a temperate climate, Global Change Biology, 19, 548-562.

Ko J., Jeong S., Yeom J., Kim H., Ban J., and Kim H., 2015. Simulation and mapping of rice growth and yield based on remote sensing. J. Appl. Remote Sens., 9(1), 960-967.

Krishnan P., Ramakrishnan B., Raja Reddy K., and Reddy V.R., 2011. High-Temperature effects on rice growth, yield, and grain quality" in Ch. 3 in Advances in Agronomy (Eds D.L. Sparks) Academic Press, Burlington.

Kumar R. and Silva L., 1973. Light ray tracing through a leaf cross- section. Appl. Optics., 12(12), 2950-2954.

Kurukulasuriya P. and Shane R., 2013. Climate change and agriculture: A review of impacts and adaptations. environment department papers, No. 91. Climate Change Series. World Bank, Washington, DC, USA.

Lee T.S., Choi J.Y., Yoo S.H., Lee S.H., and Oh Y.G., 2012. Analyzing consumptive use of water and yields of paddy rice by climate change. J. Korean Soc. Agric. Eng., 54, 47-54.

Lookingbill T.R. and Urban D.L., 2003. Spatial estimation of air temperature differences for landscape-scale studies in Montane environments. Agric. Forest Meteor., 114, 141-151.

Ly S., Charles C., and Degré A., 2013. Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review. Biotechnol. Agron. Soc. Environ., 17(2), 392-406.

Ma M. and Veroustraete F., 2006. Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China, Advances in Space Research, 37, 835-840.

McCune B., 2007. Improved Estimates of Incident Radiation and Heat Load Using Non-Parametric Regression against Topographic Variables. J. Vegetation Sc., 18, 751-754.

Monestiez P., Courault D. Allard D., and Ruget F., 2001. Spatial interpolation of air temperature using environmental context: Application to a crop model, Environ. Eco. Stat., 8, 297-309.

Mosleh M.K., Hassan Q.K., and Chowdhury E.H., 2015. Application of remote sensors in mapping rice area and forecasting its production: a review. Sensing, 15, 769-791.

Na S.I., Park J.H., and Park J.K., 2012. Development of Korean paddy rice yield prediction model (KRPM) using meteorological element and MODIS NDVI. J. Korean Soc. Agric. Eng., 54(3), 141-148.

Nuarsa I.W., Nishio F., and Hongo C., 2012. Rice yield estimation using Landsat ETM+ data and field observation. J. Agric. Sc., 4 (3), 45-56.

Patel N.K., Ravi N., and Navalagund R.R., 1991. Estimation of rice yield using IRS-1A digital data in costal tract of Orissa. Int. J. Remote Sens., 12, 2259-2266.

Park J.K., Das A., and Park J.H., 2015. A new approach to estimate the spatial distribution of solar radiation using topographic factor and sunshine duration in South Korea. Energy Conversion Manag., 101, 30-39.

Perkins S., 2010. Agriculture: Complex effects. Nature Climate Change, Editor, Press, city, country.

Piazza D., Conti F.L., Noto L.V., Viola F., and La Loggia G., 2011. Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. Int. J. App. Earth Obs. Geoinformation, 13(3), 396-408.

Prasad A.K., Singh R.P., Tare V., and Kafatos, M., 2007. Use of vegetation index and meteorological parameters for the prediction of crop yield in India. Int. J. Remote Sens., 28, 5207-5235.

Priya S. and Shibasaki R., 2001. National spatial crop yield simulation using GIS-based crop production model. Ecological Modelling, 135, 113-119.

Rahman A., Khan K., Krakauer N.Y., Roytman L., and Kogan F., 2012. Use of remote sensing data for estimation of Aman rice yield. Int. J. Agriculture Forestry, 2(1), 101-107.

Rasmussen M.S., 1992. Assessment of millet yields and production in northern Burkina Faso using integrated NDVI from the AVHRR. Int. J. Remote Sens., 13(18), 3431-3442.

Riahi K., Gruebler A., and Nakicenovic N., 2007. Scenarios of long-term socio-economic and environmental development under climate stabilization. Technological Forecasting Social Change, 74(7), 887-935.

Rose J.K., 1936. Corn Yield and Climate in the Corn Belt. Geog. Rev., 26, 88-102.

Sarma A.A.L.N., Kumar T.L., and Koteswararao K., 2008. Development of an agroclimatic model for the estimation of rice yield. J. Ind. Geophys. Union, 12(2), 89-96.

Savin I.Y. and Isaev V.A., 2010. Rice yield forecast based on satellite and meteorological data. Russ. Agric. Sci., 36 (6), 424-427.

Shen S.H., Yang S.B., Li B.B., Tan B.X., Li Z.Y., and Le T.T., 2009. A scheme for regional rice yield estimation using ENVISAT ASAR data. Sc. China Ser. D: Earth Sc., 52 (8), 1183-1194.

Shi W., Tao F., and Zhang Z., 2013. A review on statistical models for identifying climate contributions to crop yields, Geographical Sci., 23(3), 567-576.

Strand B.W., 1981. Spatial scale of crop-yield models: a review of the relationship between scale of models and accuracy. ESS Stuff Report No. AGESS8l0320. United States Department of Agriculture, Economics and Statistics Service, Statistical Research Division.

Tabios G.Q. and Salas J.D., 1985. A comparative analysis of techniques for spatial interpolation of precipitation. Water Resour. Bull., 21, 265-380.

Takeuchi W. and Yasuoka Y., 2004. Development of compositing algorithms for MODIS data. J. Japan Soc. Photogrammetry, 43, 36-48.

Thompson L.M., 1962. Evaluation of Weather Factors in the Production of Wheat, Jour. of Soil and Water Conservation, 17, 149-156.

Tucker C.J., Holben B.N., Elgin J.H., and McMurtrey J.E., 1980. Relationship of spectral data to grain yield variation. Photogrammetric Eng. Remote Sens., 46(5), 657-666.

Vaghefi N., Shamsudin M. N., Radam A., and Rahim K.A., 2013. Modelling the impact of climate change on rice production: An overview. J. App. Sc., 13, 5649-5660.

Wang X., Peng L., Zhang X., Yin G., Zhao C., and Piao S., 2014. Divergence of climate impacts on maize yield in Northeast China. Agric. Ecosys. Environ., 196, 51-58.

Wang Y.P., Chang K.W., Chen R.K., Lo J.C., and Shen Y., 2010. Large-area rice yield forecasting using satellite imageries. Int. J. Applied Erath Observation Geoinformation, 12(1), 27-35.

Yoshida S., 1981. Climatic environment and its influence. Ch. 2 in Fundamentals of rice crop science. Int. Rice Res. Institute, Laguna, Philippines.

Yun J.I., 2003. Predicting regional rice production in South korea using spatial data and crop-growth modeling. Agric. Sys., 77, 23-38.

Yun J.I., Choi J.Y., and Ahn J.H., 2001. Seasonal trend of elevation effect on daily air temperature in Korea (in Korean). Korean J. Agric. Forest Meteor., 3, 96-104.

Yun J.I., Lee D.S., Choi J.Y., Cho S.I., Park E., and Hwang H., 1999. Elevation-corrected spatial interpolation for near-real time generation of meteorological surfaces from point observations. AgroInformatics J., 1, 28-33.

Zhang G.P., Patuwo E.P., and Hu M.Y., 1998. Forecasting with artificial neural networks: The state of the art. Inter. J. Forecasting, 14, 35-62.