“Statistical and Neural Method for Site - Specific Yield Prediction”

THESIS

Submitted to the

Jawaharlal Nehru KrishiVishwaVidyalaya, Jabalpur

In partial fulfillment of the requirements for

the Degree of

MASTER OF TECHNOLOGY

In

AGRICULTURAL ENGINEERING
(SOIL AND WATER ENGINEERING)

By

PRAMOD KUMAR MEENA

Department of Soil and Water Engineering

College of Agricultural Engineering

Jawaharlal Nehru KrishiVishwaVidyalaya

Jabalpur (M.P.)

2012

CERTIFICATE - I

*This is to certify that the thesis entitled **"Statistical and Neural Method for Site –Specific Yield Prediction"** submitted in partial fulfilment of the requirement for the degree of **Master of Technology** in Agricultural Engineering in **SOIL AND WATER ENGINEERING** of Jawaharlal Nehru KrishiVishwaVidyalaya, Jabalpur is a record of the bonafide research work carried out by **Mr. PRAMOD KUMAR MEENA**under my guidance and supervision. The subject of the thesis has been approved by the Student's Advisory Committee and the Director of Instruction.*

No part of the thesis has been submitted for any other degree or diploma (Certificate awarded etc.) or has been published. All the assistance and help received during the course of the investigation has been acknowledged by him.

Place : Jabalpur (**Dr.DEVA KANT)**

Date : Chairman, Advisory Committee

THESIS APPROVED BY THE STUDENT'S ADVISORY COMMITTEE

Chairman : Dr. Deva Kant ……………………………………

Member :Dr. M. K. Hardaha ……………………………………

Member : Dr. Manish Bhan ……………………………………

CERTIFICATE – II

This is to certify that the thesis entitled ***“Statistical and Neural Method for Site –Specific Yield Prediction”*** submitted by ***Mr.PRAMOD KUMAR MEENA*** to the Jawaharlal Nehru KrishiVishwaVidyalaya, Jabalpur in partial fulfillment of the requirements for the degree of ***Master of Technology***in Agricultural Engineering in the Department of ***SOIL AND WATER ENGINEERING***has been, after evaluation, approved by the External Examiner and by the Student’s Advisory Committee after an oral examination on the same.

Place : Jabalpur (**Dr. DEVA KANT)**

Date : Chairman Advisory Committee

MEMBERS OF THE ADVISORY COMMITTEE

Member	Dr. Deva Kant	………………..........
Member	Dr. M. K. Hardaha	………………..........
Member	Dr. Manish Bhan	………………..........
Head of Deptt.	Dr. Dev Kant	……………….........
Director of Instruction	Dr. P.K. Mishra	……………….........

DEPARTMENT OF SOIL AND WATER ENGINEERING
COLLEGE OF AGRICULTURAL ENGINEERING
JNKVV, JABALPUR

No./CAE/SWE/M.Tech./2012/ Date: 29/06/2012

To,

The Director of Instruction,
JNKVV, Jabalpur (M.P.)

Through: - Dean Agril. Engineering JNKVV, Jabalpur.

Subject: - M. Tech. (SWE) thesis submission of Mr. Pramod Kumar Meena.

Please find enclosed two copies of the thesis alongwith proforma-I, Abstract submitted by Mr. Pramod Kumar Meena, M.Tech, Soil and Water Engineering Department. M.Tech. Thesis entitle as **"Statistical and Neural Method for Site –Specific Yield Prediction."** The student delivered result seminar satisfactory on 14.06.2012 and suggestion have been incorporated in the thesis.

Forwarded for necessary action please.

HOD SWE

DEAN
COLLEGE OF AGRICULTURAL
ENGINEERING

Encl:
1. Thesis (2 Copies)
2. Proforma (2 Copies)

ACKNOWLEDGEMENT

The word at my command are really inadequate to express my deep sense of gratitude to Major Advisor and Chairman of the Advisory Committee, Dr. Deva Kant, Professor & Head, Department of Soil and Water Engineering, College of Agricultural Engineering, JNKVV, Jabalpur for his able guidance, constructive criticism and valuable suggestions and his painstaking efforts during all phases of this M. Tech research work. His everlasting and praiseworthy cooperation and encouragement during entire research work.

I express my deepest sense of reverence and indebtedness to the esteemed members of my Advisory Committee, Dr. M.K. Hardaha, Professor, Soil and Water Engineering Department and Dr. Manish Bhan, Agro meteorologist, Department of Physics & Agro-meteorology, for their valuable suggestions and encouragement at various stages of investigation and research work.

I express sincere gratitude to Dr. Deva Kant, Head Department of Soil and Water Engineering, CAE, JNKVV, Jabalpur for extending the necessary facilities to carry out the research work.

I acknowledge my sincere thanks to Dr. T.K. Bhattacharya, Dean, Faculty of Agricultural Engineering, JNKVV, Jabalpur for providing all the necessary facilities for the present study.

I am also very grateful to Prof. R.K. Patel, Prof. S.K. Gupta, Prof. K.L. Mishra, Dr. M.K. Awasthi, Dr. R.K. Nema, Dr. S.K. Pyasi, Prof. Y K. Tiwari, Dr. S.K. Sharma and Prof. R.N. Shrivastava, Department of Soil and Water Engineering, CAE, JNKVV, Jabalpur for their involvement, guidance and help during the course of investigation.

Last but not the least, I am at loss to find words to express my heartfelt gratitude to my respected parents, my father Shri K.L. Meena and my mother Smt. M. Meena for their support and encouragement which have been a beacon of light in all the undertakings.

Place: Jabalpur

Date: (PRAMOD KU. MEENA)

CONTENTS

LIST OF TABLE

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Abbreviation/Symbol	Stand for
%	Percent
Ag.	Agriculture
Agril.	Agricultural
ANN	Artificial Neural Network
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
ASCE	American Society of Civil Engineering
Avg.	Average
C.A.E.	College of Agricultural Engineering
Engg.	Engineering
Eq.	Equation
et al.	And others
Etc	Et cetra
FFBP	Feed Forward Back Propagation
FFNN	Feed Forward Neural Network
Fig.	Figure
GRNN	Generalized Regression Neural Network
HDT	High Day Temperature
HRU	Hydrological Response Unit
HYMAS	Hydrological Model Application System
I.A.R.I.	Indian Agricultural Research Institute
IASRI	Indian Agricultural Statistical Research Institute
i.e.	That is
IPPC	Intergovernmental penal on climate change
I.S.A.E.	Indian Society of Agricultural Engineers
Inst.	Institution
J.	Journal

J.N.K.V.V.	Jawaharlal Nehru KrishiVishwaVidyalaya
Kg/ha	Kilogram per hectors
L	Length of data
LM	Levenberg-Marquardt
LMBP	Levenberg-Marquardt back propogation
Lph	Litre Per Hour
mm	Milimeter
MT	Million tone
M7	Minimum temperature of seventh week
M24	Minimum temperature of twenty four week
M25	Minimum temperature of twenty five week
M27	Minimum temperature of twenty seven week
M28	Minimum temperature of twenty eight week
M29	Minimum temperature of twenty nine week
M31	Minimum temperature of thirty one week
M34	Minimum temperature of thirty four week
M35	Minimum temperature of thirty five week
M36	Minimum temperature of thirty six week
M38	Minimum temperature of thirty eight week
M44	Minimum temperature of forty four week
M45	Minimum temperature of forty five week
M46	Minimum temperature of forty six week
M47	Minimum temperature of forty seven week
M49	Minimum temperature of forty nine week
Mn	Average minimum temperature at n^{th} week, ^{o}c
M.P.	Madhya Pradesh
MAE	Mean Absolute Error
Max.	Maximum
mg/L	Milli-gram per litre
Min.	Minimum
MISO	Multiple-input single-output
MLR	Multiple linear Regression
mm.	Millimeter
MSE	Mean Square Error
MSL	Mean Sea level

MUSLE	Modified Universal Soil Loss Equation
N	Number of neurons
Pc_y	Paddy crop yield
PEs	Processing elements
PME	Paddy model from Enter method of regression
PMS	Paddy stepwise regression model
pp	page
Proc.	Proceeding
Pub.	Publication
R	Correlation Coefficient
R3	Rainfall of third week
R5	Rainfall of fifth week
R6	Rainfall of six week
R7	Rainfall of seven week
R8	Rainfall of eight week
R10	Rainfall of tenth week
R11	Rainfall of eleventh week
R12	Rainfall of twelfth week
R13	Rainfall of thirteen week
R24	Rainfall of twenty four week
R25	Rainfall of twenty five week
R30	Rainfall of thirty week
R31	Rainfall of thirty one week
R32	Rainfall of thirty two week
R37	Rainfall of thirty seven week
R38	Rainfall of thirty eight week
R39	Rainfall of thirty nine week
R44	Rainfall of forty four week
R45	Rainfall of fourty five week
R47	Rainfall of fourty seven week
R48	Rainfall of fourty eight week
R50	Rainfall of fifty week
Rn	Rainfall during n^{th} week, mm
R^2	Coefficient Of Determination
RBF	Radial Basis Function
Res.	Research
Rev.	Review

RMSE	Root mean square error
S.W.E.	Soil water engineering
SCS	Soil conservation services
SE	Standard Error
SEC	Standard error of calibration
SEP	Standard error of prediction
SNN	Single Neural Network
SMW	Standard metrological week
SPSS	Statistical package the Social Science
SSM	Site specific management
SWMLR	Stepwise multiple linear regression
T_c	with respect to
T1	Maximum temperature of first week
T2	Maximum temperature of second week
T3	Maximum temperature of third week
T24	Maximum temperature of twenty four week
T26	Maximum temperature of twenty six week
T28	Maximum temperature of twenty eight week
T30	Maximum temperature of thirty week
T33	Maximum temperature of thirty three week
T34	Maximum temperature of thirty four week
T37	Maximum temperature of thirty seven week
T44	Maximum temperature of forty four week
T46	Maximum temperature of forty six week
T48	Maximum temperature of forty eight week
T52	Maximum temperature of fifty two week
Tn	Average maximum temperature at n^{th} week, ^oc
Trg	Training
trainbfg	Quasi- newton algorithm
traincgf	Fletcher-reeves Update
traincgp	Polak- Ribiere Update
traingd	Batch gradient decent
traingda	Variable learning rate
traingdm	Batch gradient decent with momentum
trainlm	Levenvergmarquardt algorithm
trainoss	Quasi- newton- one step secant
trainrp	Resilient back propagation
trainscg	Scaled conjugated gradient

Univ.	University
Val	Validation
Viz	Like
Vs	Versus
Wc_y	Wheat crop yield
WME	Wheat model from Enter method of regression
WMS	Wheat stepwise regression model
x	Observed yield
y	Estimated yield

INTRODUCTION

Crop yield is a key element for rural development and an indicator for national food security. Accurate, objective, reliable, and timely predictions of crop yield over large areas are critical for national food security through policy making on import export plans and prices. In the recent years, a variety of mathematical models relating to crop yield have been proposed.

Rice (*Oryza sativa*) is one of a major food crop of India covering an area of 43.8 million hectares (Mha) with a production of about 93.36 Million Tons (MT).In Madhya Pradesh (M.P.), rice is being constituted by around 3.79% of national hectares with a production of 1.37 MT, or contributes 1.47% of national production, and productivity of 8.24 quintal/hectare. The population growth rate per capita rice consumption is increasing and required additional 10 MT by the end of 2015(Directorate of Economics and Statistics, 2010). Rice yield is directly correlated with the amount of rainfall occurred in that area. If there is less rainfall, rice production will be less and vice-versa.

Wheat (*Triticum spp.)* is a major crop of rabi season in India. It is cultivated in 28.52 Mha with the total production of about 80.71 million tons. The area, production and productivity of wheat in the states of Madhya Pradesh are 1.815 (Mha), 7.2796 MT, and 18.35 q/ha respectively. The Yield of this crop is very sensitive to temperature variation when soil moisture is less during the critical stages, then it will affect the productivity of wheat.

Yield patterns in fields may change annually, due to spatial variations in soil properties and weather. Climatic factors like temperature, solar radiation and rainfall affect crop yield. Changes in climatic variables like rise in temperature and decline in rainfall have been reported by Intergovernmental Panel on Climate Change (IPCC, 2007). Pre- and post – anthesis high temperature and heat had massive impacts on wheat growth, whereas stress reduced its photosynthetic efficiency (Wang *et al.,*2011).You et al. (2009) observed a significant reduction in yield caused by a rise in temperature byAn increase is temperature1.8°C caused 3 to 10% reduction in wheat yield. A few days of temperature above a threshold value, if coincident with anthesis, can significantly reduce yield by affecting subsequent reproductive processes (Wheeler *et al*., 1996).

For the past few years, different approaches have been used in site-specific management studies to estimate the magnitude of various yield-limiting factors on yield. Early attempts used statistical regression (Khakural*et al*., 1996) approaches but not successful in predicting yield responses to stress in independent years, and they did not describe much of the variations in yield.

A variety of nonlinear techniques for investigating yield response have also been investigated, including boundary line analysis (Kitchen et al., 1999), state-space analysis Bayesian networks, and regression trees (Adams et al., 1999). However, many nonlinear methods can be difficult to implement, and comparison of the results from these vastly different methods is problematic. Clearly, nonlinear methods that are relatively easy to implement and can be readily compared to one another would be highly desirable. A relatively new branch of nonlinear techniques, artificial neural networks (ANN or NN), has been applied not only to artificial intelligence (Rumelhart and McClelland, 1986) and classification applications (Burks *et al*., 2000) but also as general, non–parametric "regression" tools. A neural network consists of layers of highly interconnected processing units, each containing a small amount of local "memory." The network is trained using an iterative method to adjust the weights of connections between these units. Network types, topologies, and training techniques vary considerably, but a rudimentary explanation of the critical aspects of back propagation neural networks is contained in Burks *et al*. (2000).

Models are classified based on their comprehensiveness in representing the physical processes involved. With increasing comprehensiveness; models are classified as black-box models, conceptual models, and physically based distributed models. The last of the three can be considered the better choice in a theoretical sense.

Present study area has been carried out for Jabalpur district. It's located at the Agro-Climatic zone Kymore Plateau and Satpura hills. The major crops planted in this area are ricein kharif season while wheat inrabi season. In view of this, astudy has been carried out to Estimate the yield of paddy and wheat crop in Jabalpur district based on climatic factor i.e. temperature and rainfall.

To establish the true merits of ANNs relative to simpler statistical techniques, comparisons are also made between the forecasting skill of the two ANN and a stepwise multiple linear regression models with the following specific objectives:

1. To develop multiple linear regression model forpaddy andwheatproductivity for Jabalpur district.
2. To develop an Artificial Neural Network model for paddy and wheat productivity for Jabalpur district.

REVIEW OF LITRATURE

This chapter deals with the significant contributions made by the various researchers in the field of Multiple Regression and Artificial Neural Network (ANN) in general and crop yield estimation in particular.

2.1 Regression analysis and crop yield response

Regression analysis is the relation between dependent variable Y and one or more independent variables Xi. Use of regression model in general for making forecasts/predictions/estimates for Y, investigation of functional relationship between Y and Xi, filling in missing data in Y-series, validation of Y-series.

Asana (1966)made physiological analysis of yield in wheat in relation to water-stress and temperature and found that it is often difficult to distinguish the main cause of yield reduction when both stresses overlap, as there are many apparent similarities in the response of kernel filing to drought and heat.

Wardlaw (1970) studied on the early stages of grain development in wheat response to light and temperature in a single variety and found; high temperatures increase the grain growth rate in the early ripening period, but reduce the duration of grain growth and ultimately result in a decrease in final grain weight.

Katz (1977) assessed the impact of climate change on food production. Climate Change Regression models have been criticised, since underlying mechanisms which transform climatic input into yield are not explicitly described and the hierarchical structure of the underlying physiological processes is not taken into account.

Nicolas *et al.,*(1984)studied the effects of drought and high temperature on grain growth in wheat and reported that rate of grain growth and the final grain weight have been reported to be closely associated with water deficit, shading and high temperature. Wheat grain dry weight loss was highly correlated with the number of endosperm cells being reduced by high temperature and/or drought.

Tashiro and Wardlaw (1989).compared the effect of hightemperature on graindevelopment in wheat and ricechronic high temperatures, up to a mean of approximation. 27 °C (day time maximum 30 °C), during kernel filling appear to have a similar effect to drought. There is often only a marginal increase in the rate of kernel filling as mean temperatures increase from 18 to 27 °C and a significant reduction in the duration of kernel filling, resulting in smaller kernels at maturity.

Cerrato and Blackmer(1990) reported a critical consideration in the evaluation of regression techniques is the selection of a fair means of comparing the methods chosen for analysis. The most commonly used criterion of model performance has been the coefficient of determination (R^2), which is the ratio of the variance explained by the model to the total variance in the data. However, a number of authors have concluded that R^2 is not a good means of comparison between models representing yield response. Numerous reasons have been cited as to why R^2 is not a suitable measure; however, a primary consideration is the fact that it gives no indication of how well a model performs when applied to data that were not used to create the model. Over fitting, best described as “good performance on calibration data, but poor performance on test data,” is often the result.Commonly used measure of accuracy is the root mean square error (RMSE), more commonly referred to as the standard error (SE). A major advantage of using SE over R^2 for model evaluation is that it can be calculated not only on the calibration data (termed the standard error of calibration, or SEC) but also on any “new” data set not used in developing the model (termed the standard error of prediction, or SEP) to estimate true predictive ability. While training accuracies in terms of R^2 and/or SEC have almost always been reported, predictive statistics (such as SEP) are often not reported nor even considered. When they are considered, various rules–of–thumb are applied in an attempt to avoid over fitting. A common means of measuring prediction accuracy is to use a “split–sample” approach, in which a subset of the data is withheld from training. A measure of the accuracy of prediction on this validation set is then reported.

Kitchen *et al.,* (1999) developed model with multiple linear regression techniques using polynomial and interaction terms have also been considered with some improvement over strictly linear models.

Landau *et al.,* (2000)developeda parsimonious, multiple-regression model of wheat yield response to environment. A database of nearly 2000 yield observations from winter wheat crops grown in UK trials between 1976 and 1993 was used to develop a new model of effects of weather on wheat yield. The intention was to build a model which was parsimonious (i.e., has the minimum number of parameters and maximum predictive power), but in which every parameter reflected a known climate effect on the UK crop-environment system to allow mechanistic interpretation. At this end, the model divided the effects of weather into phases which were predicted by a phenology model. A maximum set of possible weather effects in different phonological phases on yield was defined from previous knowledge. Two-thirds of the database was used to select which effects were necessary to include in the model and to estimate parameter values. The final model was tested against the independent data in the remaining third of the data set (246 aggregated yield observations) and showed predictive power (r=0.41), which was improved when comparing against mean annual yields (r=0.77). The final model allowed the relative importance of the 17 explanatory variables, and the weather effects they represent (defined before fitting), to be assessed. The most important weather effects were found to be: (1) negative effects of rainfall on agronomy before and during anthesis, during grain-filling and in the spring (2) winter frost damage (3) a positive effect of the temperature-driven duration of grain-filling and (4) a positive effect of radiation around anthesis, probably due to increased photosynthesis.

Kayam*et al*., (2000) analysed the impact of climatic change on wheat production of Aegean region in Turkey not been considered. The preliminary study would indicate that the reduction of 5-10% in the rainfall will have a small effect on the wheat yield in a region with 500-600mm of rainfall. The increase in temperature on the other hand by 1-2^{o}C will reduce yields by 7.4% per 1^{o}C.

Raiand Chandrahas (2000) used time series data on weather parameters and also developed technique of discriminant function analysis for crop yield prediction, using weather data of these groups; linear/quadratic discriminant functions were fitted.These functions were used to find weather scores for each year at different phases of crop growth and were used as repressors in forecast model.

Wardlaw(2000) reported on interaction between drought and chronic high temperature during kernel filling in wheat in a controlled environment and found that wheat plants (*Triticumaestivum L. `Lyallpur'*), limited to a single culm, were grown at day/night temperatures of either 18/13°C (moderate temperature), or 27/22°C (chronic high temperature) from the time of anthesis, effect of water stress on kernel dry weight at maturity of plants with either the full complement or reduced numbers of kernels, and subjected to low and high temperatures following anthesis, indicate that the effect of drought on kernel dry weight may be reduced, in both absolute and relative terms, rather than enhanced, at high temperature. It is suggested that where high temperature and drought occur concurrently after anthesis there may be a degree of drought escape associated with chronic high temperature due to the reduction in the duration of kernel filling, even though the rate of water use may be enhanced by high temperature.

Agrawal*et al.,* (2001) developed model on agro-climatic zone basis by combining the data of various districts within the zone so that a long series could be obtained in a relatively shorter period. Previous years yield, moving averages of yield and agricultural inputs were taken as the variables taking care of variation between districts within the zone. Year variable was included to take care of technological changes. Different strategies for pooling district level data for the zone were adopted. Results revealed that reliable forecasts can be obtained using this methodology at 12 weeks after sowing i.e. about 2 months before harvest. The data requirement reduced to 10-15 years as against 25-30 years for district level models. The approach has been successfullyused for forecasting yields of rice, wheat and sugarcane for Uttar Pradesh.

Saksena*et al.,* (2001) developed models for rainfed crop using weighted stress indices. In this approach, water deficit/surplus has been worked out at different phases of crop growth and using suitable weights, accumulated weighted stress index has been developed for each year which was used as regressor in the forecast model.

Drummond *et al*., (2003) Stepwise Multiple Linear Regression (SMLR) may help improved our understanding of the complex relationship between soil properties topography and crop yields. However leaves one site-year out (LOO) experiments showed clear sing of over lifting, even when climatological conditions were reduced to a single variable a large number of additional site years of data representing a range of climatological.

Morita *et al*., (2005) studied on grain growth and endosperm cell size under high night temperatures in rice (*Oryzasativa L*.). The results show that HNT(high night temperature) compared with HDT(high day temperature) reduced the final grain weight by a reduction in grain growth rate in the early or middle stages of grain filling, and also reduced cell size midway between the central point and the surface of endosperm.

Uno *et al.*,(2005)studied on artificial neural networks to predict corn yield from compact airborne andestimated various soil and crop parameters for precision farming from airborne hyper spectralimagery. The potential of artificial neural networks (ANNs) for the development of in-season yield mapping and forecasting systems was examined. Hyper spectral images of corn (Zea mays L.) plots in eastern Canada, subjected to different fertilization rates and various weed management protocols, were acquired by a compact airborne spectral imager. Statistical and ANN approaches along with various vegetation indices were used to develop yield prediction models. Principal component analysis was used to reduce the number of input variables. Greater prediction accuracy (about 20% validation RMSE) was obtained with an ANN model than with either of the three conventional empirical models based on normalized difference vegetation index, simple ratio, or photochemical reflectance index. No clear difference was observed between ANNs and stepwise multiple linear regression models. Although the high potential usefulness of ANNs was confirmed, particularly in the creation of yield maps, further investigations are needed before their application at the field scale can be generalized.

Lal*et al*., (2008) reported that agricultural sector is one of the sensitive areas which would be influenced by the projected global warming and associated climate change. In spite of the uncertainties about the precise magnitude of climate change on regional scales, an assessment of the possible impacts of changes in key climatic elements on our agricultural resources is important for formulating response strategies. In this study, vulnerability of wheat and rice crops in northwest India to the projected climate change is examined. CERES wheat and rice models adopted for the study were validated for their ability to reproduce yields at the selected NW Indian stations. The sensitivity experiments with these models showed higher yields for both wheat and rice (28% and 15% respectively for a doubling of CO_2) under elevated CO_2 levels. A 3°C (2°C) rise in air temperature nearly cancels out the positive effect of elevated CO_2 on the wheat (rice) yields. While the wheat crops are found to be sensitive to increase in maximum temperature, the rice crops are vulnerable to increase in minimum temperature. The combined effect of enhanced CO_2 and imposed thermal stress on the wheat (rice) crop is 21% (4%) increase in yield for the irrigation schedule presently practiced in the region. While the adverse impacts of likely water shortage on wheat crops would be minimized to a certain extent under elevated CO_2 levels, they would largely be maintained for the rice crops resulting in about 20% net decline in rice yields. In general, acute water shortage conditions combined with the thermal stress should adversely affect both the wheat and more severely the rice productivity in NW India even under the positive effects of elevated CO_2 in the future.

Shin *et al.*, (2009) made a comprehensive evaluation of crop yield simulations with various seasonal climate data is performed to improve the current practice of crop yield projections. The El Nino Southern Oscillation (ENSO)-based historical data are commonly used to predict the upcoming season crop yields over the southeast United States. In this study, eight different seasonal climate data are generated using the combinations of two global models, a regional model, a statistical downscaling technique, and two convective schemes. These data are linked to maize and peanut dynamic models to assess their impacts on crop yield simulations compared to the ENSO-based approach. Improvement of crop yield simulations with the climate model data is varying, depending on the model configuration and the crop type. While the global climate model data provide no improvement, the dynamically and the statistically downscaled data show increased skill in the crop yield simulations. A statistically downscaled operational seasonal climate model shows a statistically significant (5%) inter annual predictability in the peanut yield simulation. Since the yield amount simulated by the dynamical crop model is highly sensitive to wet/dry spell sequences (water stresses) during the growing season, a proper parameterization of precipitation physics is essential in climate models to improve the crop yield projection.

Alva *et al.*, (2010) developed a crop simulation model for predicting yield and rate of nitrogen in Irrigated potato rotation cropping system and concludedthat development of regression models provides a valuable tool to analyse the behaviour of wheat crop under a wide range of production conditions.

Qayyum and Khalid (2010) concluded on Impact of weighted rainfall on the yield of wheat in the Punjab, observed that as the rainfall in the important months of Rabi season is increased; the average yield of wheat also increases. But at a certain level, excess rainfalls cause a decline in the yield. The suitable quadratic models were fitted independently between yield and amount of rainfalls for each month of the season. In each case explained variation R^2 was used as a proposed weight to the rainfall of the relevant Rabi month. The empirical study supported the use of weighted rainfalls instead of total rainfalls of Rabi season. The concept of weighted rainfall was also proved effective in estimating the crop production through a statistical model.

Ahmed *et al.*, (2011) concludedregression models are regarded as valuable tools for the evaluation of temperature, solar radiation and photo thermal quotient effects on wheat yield to bring its resilience to climatic vulnerability. The objective of this study was to evaluate sole and cumulative impact of temperature and solar radiation on spring wheat (*Triticumaestivum* L.) yield using regression modelling approach. Scatter-plot regression model was developed at 95% confidence interval with crop data and climate variables. Results indicate direct relationship of yield with solar radiation, cumulative effect of temperature and solar radiation, whereas yield had an inverse relationship with temperature alone. Crop-yield under increased climatic vulnerability.

2.2 ANN model

The development of Artificial Neural Network (ANN) began approximately 60 years ago, inspired by a desired to understand the human brain and emulate its functioning. It has experienced a huge resurgence due to the development of more sophisticated algorithms and emergence of more powerful computers. Extensive research has been devoted to investigating the potential of ANNs as computational tool that acquire, represent and compute a mapping from one multivariate input space to another. Mathematically, an ANN is often viewed as a universal approximate. The ability to identify a relationship from given pattern makes it possible for ANNs to solve large-scale complex problems like pattern recognition, nonlinear modelling, classification, association and control.

An ANN is described as an information processing system that is composed of many non-linear and interconnected processing element or neurons. ANN acquires the knowledge through a learning process that involves finding an optimal set of weights for the connection and threshold value for the nodes. ANN has been proven to understand problem that deal with noise or involve pattern recognition and where input is incomplete or ambiguous by nature. These are calibrated using automatic calibration techniques, thus eliminating the lengthy cycle. These properties suggest the ANNs may be well suited to the problem of estimation and prediction in atmospheric and hydrology, which is complicated because of non-linearity of physical processes, and uncertainty in parameter estimation. ANN is simply a sophisticated regression, which has a network of many simple elements or processers or ‘neurons’. The elements are connected by communication channels or ‘connecters’, which usually carry numeric data, encoded by a variety of means and often organized into subgroups or layer. A neural network can perform a particular function when certain values are assigned to the connection or ‘weights’ between elements. To describe a system, there is no assume structure of the model, instead the networks are adjusted or ‘trained’, so that a particular input leads to a specific target output, which is called supervised learning. Objective of the training is to minimize the residual sum of squares between measured and predicted output. After appropriate training, Artificial Neural Network are found to generate satisfactory results for many prediction problems.

May *et al.*, (1993) indicated thatthe choice of input variables is a fundamental, and yet crucial consideration in identifying the optimal functional form of statistical models. The task of selecting input variables is common to the development of all statistical models, and is largely dependent on the discovery of relationships within the available data to identify suitable predictors of the model output. In the case of parametric, or semi-parametric empirical models, the difficulty of the input variable selection task is somewhat alleviated by the a priori assumption of the functional form of the model, which is based on some physical interpretation of the underlying system or process being modelled. However, in the case of artificial neural networks (ANNs), and other similarly data-driven statistical modelling approaches, there is no such assumption made regarding the structure of the model. Instead, the input variables are selected from the available data, and the model is developed subsequently. The difficulty of selecting input variables arises due to (i) the number of available variables, which may be very large; (ii) correlations between potential input variables, which create redundancy; and (iii) variables that have little or no predictive power. Variable subset selection has been a longstanding issue in fields of applied statistics dealing with inference and linear regression and the advent of ANN models has only served to create new challenges in this field. The non-linearity, inherent complexity and non-parametric nature of ANN regression make it difficult to apply many existing analytical variable selection methods. The difficulty of selecting input variables is further exacerbated during ANN development, since the task of selecting inputs is often delegated to the ANN during the learning phase of development. A popular notion is that an ANN is adequately capable of identifying redundant and noise variables during training, and that the trained network will use only the salient input variables. ANN architectures can be built with arbitrary flexibility and can be successfully trained using any combination of input variables (assuming they are good predictors). Consequently, allowances are often made for a large number of input variables, with the belief that the ability to incorporate such flexibility and redundancy creates a more robust model. Such pragmatism is perhaps symptomatic of the popularization of ANN models through machine learning, rather than statistical learning theory. ANN models are too often developed without due consideration given to the effect that the choice of input variables has on model complexity, learning difficulty, and performance of the subsequently trained ANN.

Haykin (1994) reported thatANNs can solve highly non-linear problems without the need to define the explicit relationship existing between inputs and outputs. ANNs simulate the relationship by mapping the varying values of inputs and outputs. Any parameter which remains constant is not considered by ANNs, even though these constant parameters are important to conceptual models. For instance, when ANNs simulate the hydrologic conditions at only one location, the hydrology conductivity would not vary and can be discarded by ANNs. ANNs thus require fewer input variables than do conceptual models. Also, ANNs can run in a short time once they are trained.

Mcclendon*et al.*, (1996) applied ANNs to Control and Optimize the Irrigation Management of Peanuts. ANNs have the capability for the simulation and real-time control of water table management systems under subsurface drainage and sub irrigation.

Yang *et al.*, (1997) analysed on an artificial neural network to estimate soil temperature his study was undertaken to develop an artificial neural network (ANN) model for transient simulation of soil temperature at different depths in the profile. The capability of ANN models to simulate the variation of temperature in soils was investigated by considering readily available meteorological parameters. The ANN model was constructed by using five years of meteorologic data, measured at a weather station at the Central Experimental Farm in Ottawa, Ontario, Canada. The model inputs consisted of daily rainfall, potential evapotranspiration, and the day of the year. The model outputs were daily soil temperature sat the depths of 100, 500 and 1500 mm. The estimated values were found to be close to the measured values, as shown by a root-mean-square error ranging from 0.59 to 1.82°C, a standard deviation of errors from 0.61 to 1.81°C, and a coefficient of determination from 0.937 to 0.987. Therefore, it is concluded that ANN models can be used to estimate soil temperature by considering routinely measured meteorological parameters.

Tokar and Johnson (1999) employed a FFNN methodology to forecast daily runoff as a function of daily precipitation, temperature, and snowmelt for the little Patuxent river watershed in Maryland. They compared FFNN rainfall-runoff model with results obtained using the statistical regression and a simple conceptual model and found that the FFNN performs better than the others.

Mather (2000) developed a new artificial neural networks (ANN) model simulating yield using NDVI should represent non-linear relationships between input variables and desired variables. The ability of ANN models to associate complicated spectral information with target attributes without any constraints of sample distribution makes them ideal for describing the complex non-linear relationships between canopy-level spectral signatures and various crop conditions

Burks *et al.*, (2000) applied artificial neural networks (ANN or NN) to artificial intelligence and classification applications and also as general, non–parametric “regression” tools.

Liu *et al.*, (2001) found that standard back propagation neural network can be used to estimate corn yield over a number of year of small plot data, which included soil, weather, and management factor. Their result were promising, with predictive error reported to be approximately 20 % of actual yield although only a single validation set was used.

Dawson and Wilby (2001) suggested that research might focus on the extraction of hydrological rules from ANN weights, and on the development of standard performance measures that penalize unnecessary model complexity.

Chang and Chen (2001) used a fuzzy-neural network model for real time stream flow prediction. They compared the model results with those of the autoregressive moving average exogenous variables model (ARMAX) and found that the fuzzy-neural network model could offer a higher degree of reliability and accuracy than the ARMAX in stream flow forecast.

Fisher *et al.*, (2002) analysed the long-term hydrologic response of a small (7.8) zero-order southern piedmont watershed form 1940-1984. Land use and rainfall variability influenced runoff characteristics. Row cropping produced the greatest runoff, percentage runoff, and peak flow. Bermuda grass reduced runoff more than row cropping but not as much as kudzu or kudzu mixed with rescue grass. Peak flows increased during the cropping phase. Monthly rainfall-runoff regression relationships were developed and R^2 greater than 0.60 were found for row cropping.

Agarwal*et al.*, (2004) developed the daily ANN models worked well in runoff forecasting, whereas other conventional model failed in calibration as well as in both the verification periods. The daily ANN model showed an improvement over the former in calibration, but was slightly poor in cross-validation as well as in verification.

Jeong and Kim (2005) calibrated ANN models were compared with each other first. The results show that the ENN (Ensemble Neural Network) is less sensitive to the input variable selection and the number of hidden nodes than the SNN (Single Neural Network) is, and the ENN, in general, produced smaller RMSEs than the corresponding SNN, which implies that the ENN can reduce the generalization error more efficiently than the SNN.

.Sarangi*et al.*,(2005) developed artificial neural network and regression models using watershed-scale geomorphologic parameters to predict surface runoff. The coefficient ofdetermination (R^2) and model efficiency factor (E) wereestimated to ascertain the model performance. ANN model validation statistics resulted in R^2 values ranging from 0.85 to 0.95 and E values from 0.74 to 0.82 for peak runoff rate.

Li *et al.*, (2007) estimating crop yield from multi-temporal satellite data using multivariate regression and neural network techniques, developed a new methodology using an artificial neural network (ANN) to estimate and predict corn and soybean yields on a county-by-county basis, in the “corn belt” area in the Midwestern and Great Plains regions of the United States. The historical yield data and long time-series NDVI derived from AVHRR and MODIS are used to develop the models. A new procedure is developed to train the ANN model using the SCE-UA optimization algorithm. The performance of ANN models is compared with multivariate linear regression (MLR) models and validation is made on the model’s stability and forecasting ability. The new algorithms can effectively train ANN models, and the prediction accuracy can be as high as 85 per cent.

Carter and Page (2009)studied on The human brain book, Barnes & Noble and found that Humans have an innate predisposition for ambulation (walking). The motor neuron stimulation involved in ambulation is generated by a natural neural network located.

Kaczmarczyk*et al.*, (2009) found ANN analysis to be superior to qualitative variable analysis for classifying post stroke patients’ gait patterns into three gait types, as well as superior to the analysis of min/max joint angle values.

Chen *et al.*,(2011) analysed onpredictionmethod of crude oil output based on artificial neural network and found the prediction of oil field output is a complex and nonlinear problem. According to shortcomings of some current prediction methods, the paper applied the improved artificial neural network method to the prediction of oil field output and got good results, also improved prediction precision.

Othman and Mahdi (2011) presented an approach for forecasting of long term reservoir inflow using monthly inflow available data. A Levenberg-Marquardt Back Propagation (LMBP) algorithm has been used to develop the ANN models. In developing the ANN models, different networks with different numbers of neuron hidden layers were evaluated. A total of 21 years of historical data were used to train and test the networks. The optimum ANN network with 4 inputs, 5 neurons in hidden layer and one output was selected. To evaluate the accuracy of the proposed model, the Mean Squared Error

(MSE) and the Correlation Coefficient (CC) were employed. The network was trained and converged at MSE = 0.0188 by using training data subjected to early stopping approach. The network could forecast the testing data set with the accuracy of MSE = 0.0283. Training and testing process showed the correlation coefficient of 0.7282 and 0.7228 respectively.

From the forgoing discussion. It can be stated that yield estimation by linear regression as well as ANN base models are in use. The model with limited predictor variables are preferred over those use number of variables especially those difficult to measure accurately. Models based on MLR and ANN has been developed for specific location i.e., Jabalpur district in this study.

MATERIAL AND METHODS

This chapter deals with the detailed description of methodology adopted for estimation of Crop yield and development of neural networks models strategy applied at Jabalpur District.

3.1 Study area

Present study was carried out at Jabalpur district. It's located at latitude $23^0$09'36''N to 23^0.37'N, longitude $79^0$57'E to 79^0.95'E, and average MSL of 408 m of the Agro-Climatic zone of Kymore Plateau and Satpura hills. Location of the district is depicted in Fig. 3.1.Jabalpur district has a humid subtropical climate, typical of North-Central India. Summer starts in late March and last up to June. May is the hottest month with average temperature reaching up to and beyond 45^0C. They are followed by monsoon season, which lasts until early October, with a normal precipitation of nearly 1386 mm. Winter starts in late November and last until early March. They peak in January with average temperature near 7^0C. Jabalpur gets moderate rainfall of 890 to 970 mm during July-September due to the southwest monsoon.

3.2 Collection of data

3.2.1Crop yield

Yield data of wheat and paddy crop of Jabalpur district was collected from Department of Agricultural Economics & Farm Management, JNKVV, Jabalpur for the year 1969-2010. Yield data form 1969 to 1998 includes the yield of Katni district at present the productivity of these crops has be computed from yield and area under specific crop on the district.

3.2.2. Rainfall and temperature

The weekly data of rainfall, maximumand minimum temperature of Jabalpur werecollected from The Department of Physics & Agro-meteorology, College of Agricultural Engineering,JNKVVJabalpur for the years 1969-2010.

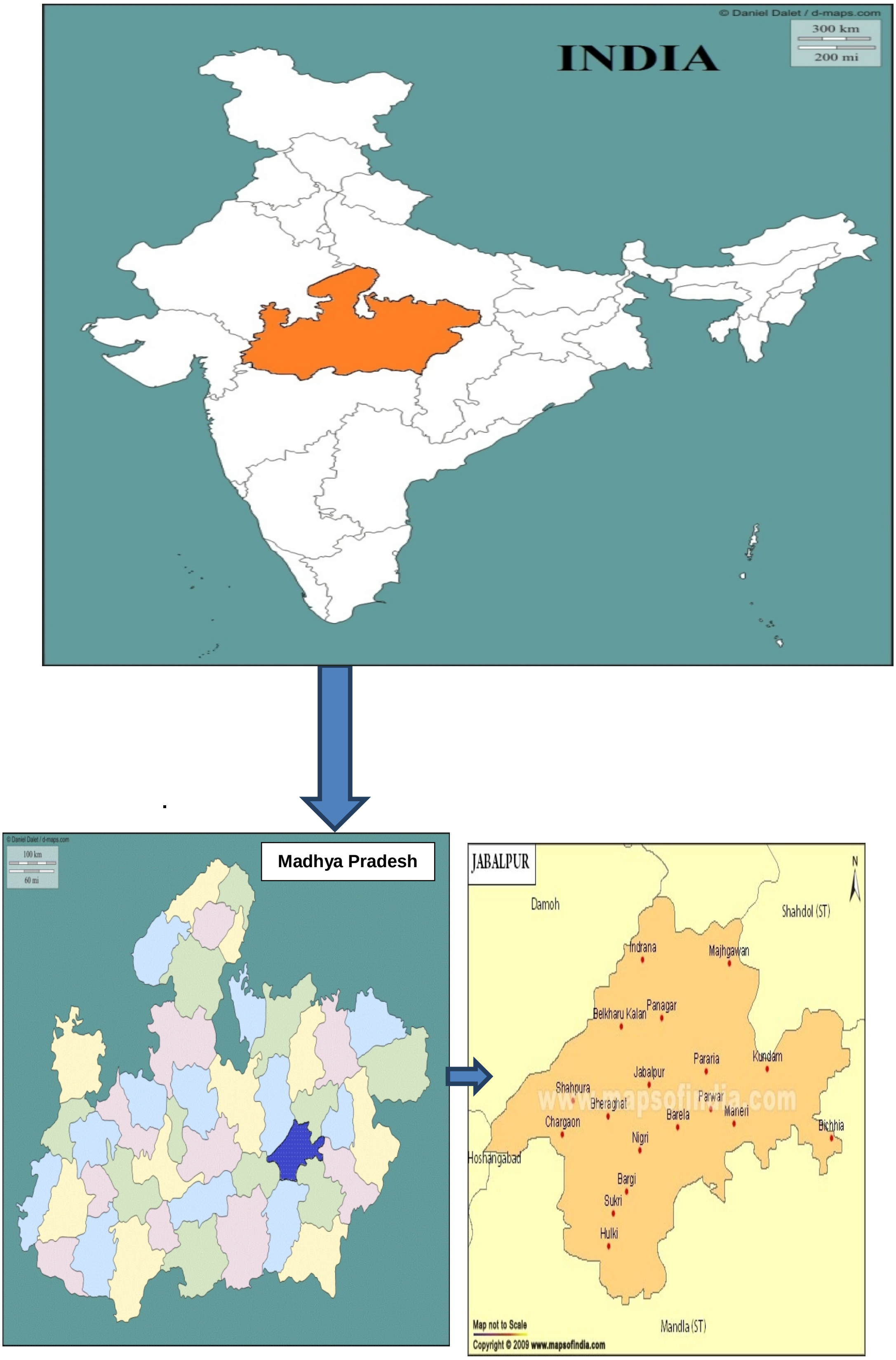


Fig. 3.1: Location of Jabalpur district

3.2.3 Randomization of data

Out of 41 years of data considered for the analysis, initially 25 years of data have been used for model development and rest for validation of the model.Due to separation of Katnidistrict, from the Jabalpur in the year 1999,large variations in the productivity have been observed. In order to reduce the temporal effect on productivity, the total 41 years data have been randomized for the purpose of development and validation of the model.

3.3 Design of regression model

Regression analysis is used for explaining or modelling the relationship between a single variable Y, called the response, output or dependent variable, and one or more predictor, input, independent or explanatory variables, X_1……....X_n. When n=1, it is called simple regression but when n > 1 it is called multiple regression or sometimes multivariate regression.

3.3.1 Linear regression model

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 \ldots\ldots\ldots\ldots\ldots\ldots + \beta_n X_n + \varepsilon \qquad \ldots\ldots (3.1)$$

$$Y = \alpha + \sum_{i=1}^{n} \beta n X n + \varepsilon \qquad \ldots\ldots (3.2)$$

Where,

Y = Dependable variable.

β= Partial slope coefficient (also called partial regression coefficient, metric coefficient). It represents the change in E(Y) associated with a one-unit change in X i when all other input variables are held constant.

α = The intercept. Geometrically, it represents the value of E(Y) where the regression surface (or plane) crosses the Y axis. Substantively, it is theevent of Y when all the IVs equal 0.

ε = The deviation of the value Y from the mean value of the distribution given X. This error term may be conceived as representing (1) the effect on Y of variables explicitly included in the equation.

In most situations, we are not in a position to determine the population parameters directly. Instead, we must estimate their values from a finite sample from the population. The sample regression model is written as

$$Y = a + b_1X_1 + b_2X_2 + \ldots\ldots\ldots.. + b_nX_n + \varepsilon \qquad \ldots\ldots\ldots\ldots\ldots\ldots(3.3)$$

Where, a is sample estimate of α and b_n is the sample estimate of β_n.

3.3.2 Multiple variable models

While choosing a predictor variable you should select one that might be correlated with the criterion variable, but not strongly correlated with the other predictor variables. However, correlations among the predictor variables are not unusual. The term multicollinearity (or collinearity) is used to describe the situation when a high correlation is detected between two or more predictor variables. Such high correlations cause problems when trying to draw inferences about the relative contribution of each predictor variable to the success of the model. SPSS provides you with a means of checking for this and we describe this below.

SPSS (originally, Statistical Package for the Social Sciences) was released in its first version in 1968 after being developed by Norman H. Nie and C.Hadlai Hull. SPSS is among the most widely used programs for statistical analysis.It is used by market researchers, health researchers,surveycompanies, government, education, researchers, marketing, organizations and others. In addition to statistic, data management (case selection, creating derived data) and data documentation (a metadata dictionary is stored in the datafile) are features of the base software.Wewere using SPSS 16.0.2 for regression model.window of linear regression statistics of SPSS in show in Fig. 3.2.

Enter method

There are different ways that the relative contribution of each predictor variable can be assessed. In the “simultaneous” method (which SPSS calls the Enter method), the researcher specifies the set of predictor variables that make up the model. The success of this model in predicting the criterion variable is then assessed.

In contrast, “hierarchical” methods enter the variables into the model in a specified order. As each variable is entered into the model its contribution is assessed. If adding the variable does not significantly increase the predictive power of the model then the variable is dropped. In “statistical” methods, the order in which the predictor variables are entered into (or taken out of) the model is determined according to the strength of their correlation withthe criterion variable.

If we have no theoretical model in mind, then it is probably safest to use Enter, the simultaneous method. Statistical procedures used with caution and only when you have a large number of cases. This is because minor variations in the data due to sampling errors have been a large effect on the order in which variables are entered.

All variables must pass the tolerance criterion to be entered in the equation, regardless of the entry method specified. The default tolerance level is 0.0001. Also, a variable is not entered if it would cause the tolerance of another variable already in the model to drop below the tolerance criterion.

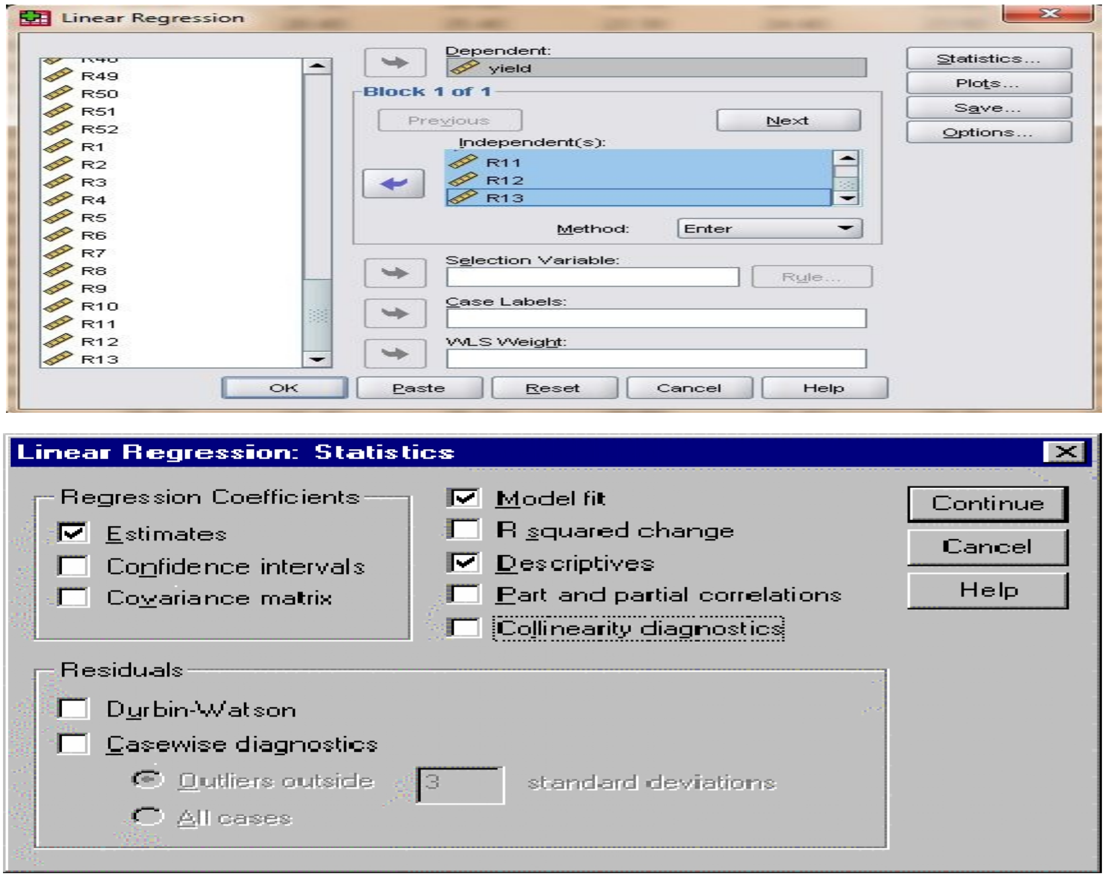


Fig. 3.2:SPSS Window for multiple regression model

3.3.3 Stepwise regression analysis

A good theory is the end result of a winnowing process. A comprehensive model that includes all conceivable, testable influences on the phenomena under investigation, then a test has been carried out to find the components of the initial comprehensive model, to identify the less comprehensive sub models that adequately account for the phenomena under investigation.Finally, from these submodels, a single out the simplest sub model, which by the principle of parsimony we take to be the “best” explanation for the phenomena under investigation.

A simple model is preferred not just for philosophical but also for practical reasons. Simple models are easier to put to test again in replication and cross–validation studies.Simple models are less costly to put into practice in predicting and controlling the outcome in the future.The philosophical reason for preferring simple models should not be downplayed. However simpler models are easier to understand and appreciate, and therefore have a “beauty” that their more complicated counterparts often lack.The entire winnowing process described above is encapsulated in the model-building techniques of stepwise and best–subset regression.The use of these model-building techniques begins with specification of the design for a comprehensive “whole model”.Less comprehensive sub-models are then tested to determine if they adequately account for the outcome under investigation. Finally, the simplest of the adequate is adapted as the “best”.

The initial model in stepwise regression

The initial model is designated the model at step 0. The initial model always includes the regression intercept (unless the No intercept option has been specified). For the backward stepwise and backward removal methods, the initial model also includes all effects specified to be included in the design for analysis. The initial model for these methods is therefore the whole model.

The forward stepwise method

The forward stepwise method employs a combination of the procedure used in the forward entry and backward removal methods. At Step1 the procedures for forward entry are performed. At any subsequent step where 2 or more effects have been selected for entry into the model forward entry is performed if possible and backward removal is performed if possible, until no procedure can be performed and stepping is terminated. Stepping is also terminated if the maximum number of steps is reached.

The backward stepwise method

The backward stepwise method employs a combination of the procedures used in the forward entry and backward removal method. At Step 1 the procedures for backward removal are performed. At any subsequent step where 2 or more effects have been selected for entry into the model, forward entry is performed is possible, and backward removal is performed if possible, until neither procedure can be performed nor stepping is terminated. Stepping is also terminated if the maximum number of step is reached.

Entry and removal criteria

Either critical f values critical p values or (0.05) can be specified to be used to control entry and removal of effects from the model. If p values are specified, the actual values used to control entry and removal of effects from the model is1 minus the specified p values. The critical value for model entry must exceed the critical value for removal from the model. A maximum numbers of steps can also be specified, if not previously terminated, stepping stops when the specified maximum number of Steps is reached.

Method selection allows you to specify how independent variables are entered into the analysis. Using different methods, you can construct a variety of regression models from the same set of variables.All variables must pass the tolerance criterion to be entered in the equation, regardless of the entry method specified. The default tolerance level is 0.0001. Also, a variable is not entered if it would cause the tolerance of another variable already in the model to drop below the tolerance criterion.

3.4Predictor variables for MLR

During modelling process emphases given that the model should include those predictor variables which are essay and less costly to observe.Hencerainfall, minimum temperature,andmaximum temperature are considered as predictor variables since large variation in temperature and rainfall observe during the growing period of crop, daily oratleast weekly rainfall and temperature should be considered as predictor variables. Daily rainfall, daily maximum and daily minimum temperature will account large number of predictor variables therefore weekly rainfall, average weekly minimum, and average weekly maximum temperature has been considered as predictor variable in this study.

3.4.1 Predictor variables for paddy

There are large variations in the duration of paddy crop in the district. To generalise the model it is assumed that the paddy crop has average crop period of 16^{th} week that is from 24^{th} standard metrological week (SMW) to 39^{th} SMW. The total number of predictor variable for this period becomes 48, hence to reduce the number of predictor variables model was developed with stepwise regression (PMS) as well as enter method (PME). The predictor variables selected by enter method and stepwise regressionmethod for paddy crop shown in Table 3.1.

3.4.2 Predictor variables for wheat

There are large variations in the duration of wheat crop in the district. To generalise the model it is assumed that the wheat crop has average crop period of 22^{th} week that is from 44^{th} standard metrological week (SMW) to 13^{th} SMW. The total number of predictor variable for this period becomes 66, hence to reduce the number of predictor variables model was developed with stepwise regression (WMS) as well as enter method (WME). The predictor variables selected by enter method and stepwise regression method for wheat crop shown in Table 3.2.

Table 3.1 Multipleregression models with varying input variables for paddy

Model	Predictor	Output(Kg/ha)
Enter Method (PME)	T26 T28 T30 T32 T33 T34 T37 M24 M25 M27 M28 M29 M31 M34 M35 M36 M38 R25 R30 R31 R32 R37 R38 R39	Paddy crop yield (PCy)
Stepwise regression (PMS)	T24 M25 R24 R31	paddy crop yield (PCy)

PCy=Paddy crop yield (kg/ha)

Tn= Average maximum temperature of n^{th} week (^oC)

Mn=Average minimum temperature of n^{th} week (^oC)

Rn= Rainfall during nth week (mm)

Table 3.2 Multipleregression models with varying input variables for wheat

Model	Predictor	Output (Kg/ha)
Enter Method (WME)	T44 T46 T52 T1 T2 T3 M46 M49 M7 R44 R45 R46 R47 R48 R50 R3 R5 R6 R7 R8 R10 R11 R12 R13	Wheat crop yield (WCy)
Stepwise regression (WMS)	T11 T48 M44 M45 M47 M49 R8 R50	Wheat crop yield (WCy)

WCy= Wheat crop yield (kg/ha)

Tn= Average maximum temperature of n^{th} week (^oC)

Mn =Average minimum temperature of n^{th} week (^oC)

Rn=Rainfall during nth week (mm)

3.5 Artificial Neural Network (ANN)

An ANN is a massively parallel-distributed information processing system that has certain performance characteristics resembling biological neural networks of the human brain (Haykin, 1994). ANN is a network of large number of processing elements (PEs - also called as nodes or artificial neurons) connected by unidirectional communication channels (connections), each with its own small sphere of knowledge and access to data in its local memory. Typically, a neural network is characterised by its architecture that represent s the pattern of connection between neurons, its method of determining the connection weights and the activation function (Fausett, 1994). ANNs can be categorised by the number of layer (single, bilayer or multilayer).

These can also be categorised based on the direction of information flow and processing (feed forward or recurrent). In most networks, the input layer receives the input variables for the defined problem and the last layer consist of values predicted by the network and thus represents model output. The number of hidden layers and the number of neurons on each layer are usually determined by a trial and error procedure.

3.5.1 The analogy to the brain

The basic processing elements of artificial neural network are the artificial neuron, which is analogous to the biological neurons. The biological neuron is a specific type of cell of the human brain. Each of these neurons can connect with up to 200000 other neurons. The power of brain comes from the number of these basic components and the multiple connections between them. A typical structure of the biological neurons is depicted in Fig. 3.3.

All biological neuron have five basic components, which are the dendrites, the soma,thehillock, the axon and the synapses.The soma is the body of neurons, receives the synaptic information from the dendrites, which are the extensions of soma and from as input channels, and perform further processing of the information. The soma provides aggregation of input signals and decides when and how to respond to inputs. If the weighted aggregation of signals exceeds certain threshold, soma will produce an output. These output potentials are transmitted along the axon, the connection between the soma and the synapses to other neurons for further processing. The strength of the synapses is a representation of the stored knowledge. The synaptic operation assigns a relative weight to each incoming signal according to the past experience stored in the synapses.

The artificial neurons are analogous to the biological neuron in its structure as well as function. Artificial neurons are much simpler than the biological neuron. The functional diagram of an artificial neuron is shown in Fig. 3.4. There are weighted input connections to the artificial neurons, analogous to the dendrites. These input signals get added up, and are fed into the activation function,which determines if the neurons should at all react like soma in biological neuron. The reaction signals of the neuron would then pass through a transfer function, which decided the strength of the out signal similar to hillock.

Finally, the output signal is send through all the output connections to other neurons as through synapses I case of biological neuron.

$$y_j = \int\{W_j \times X_i\} - \theta_j \qquad \text{…………. (3.4)}$$

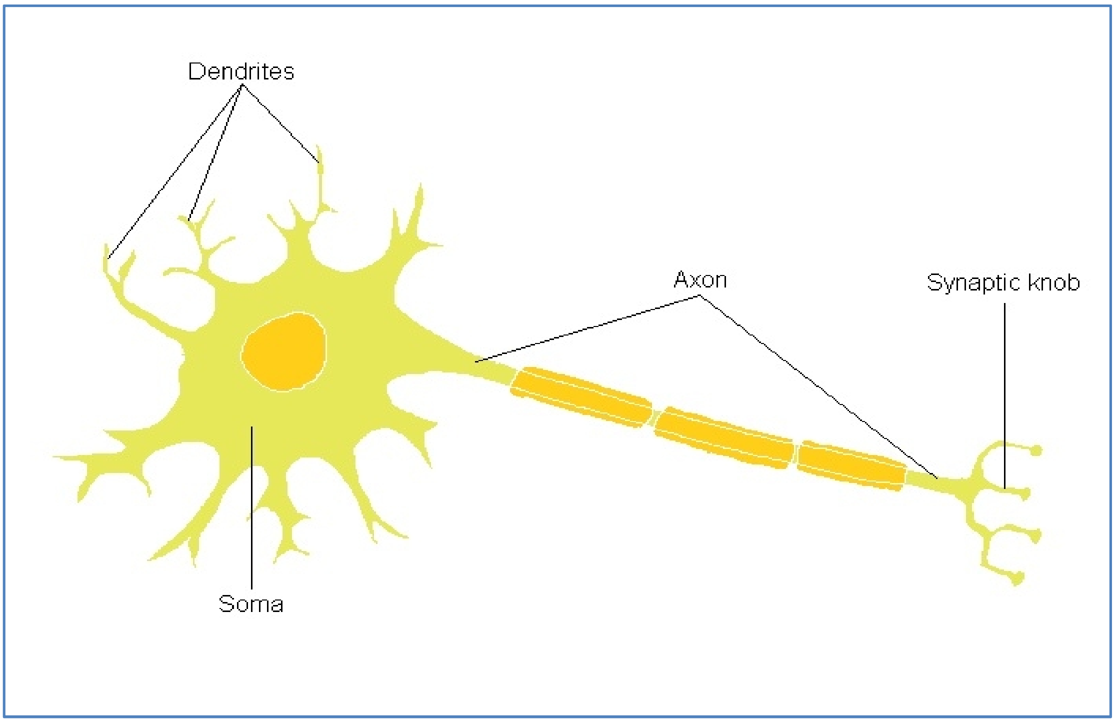


Fig.3.3: A biological neuron and its components

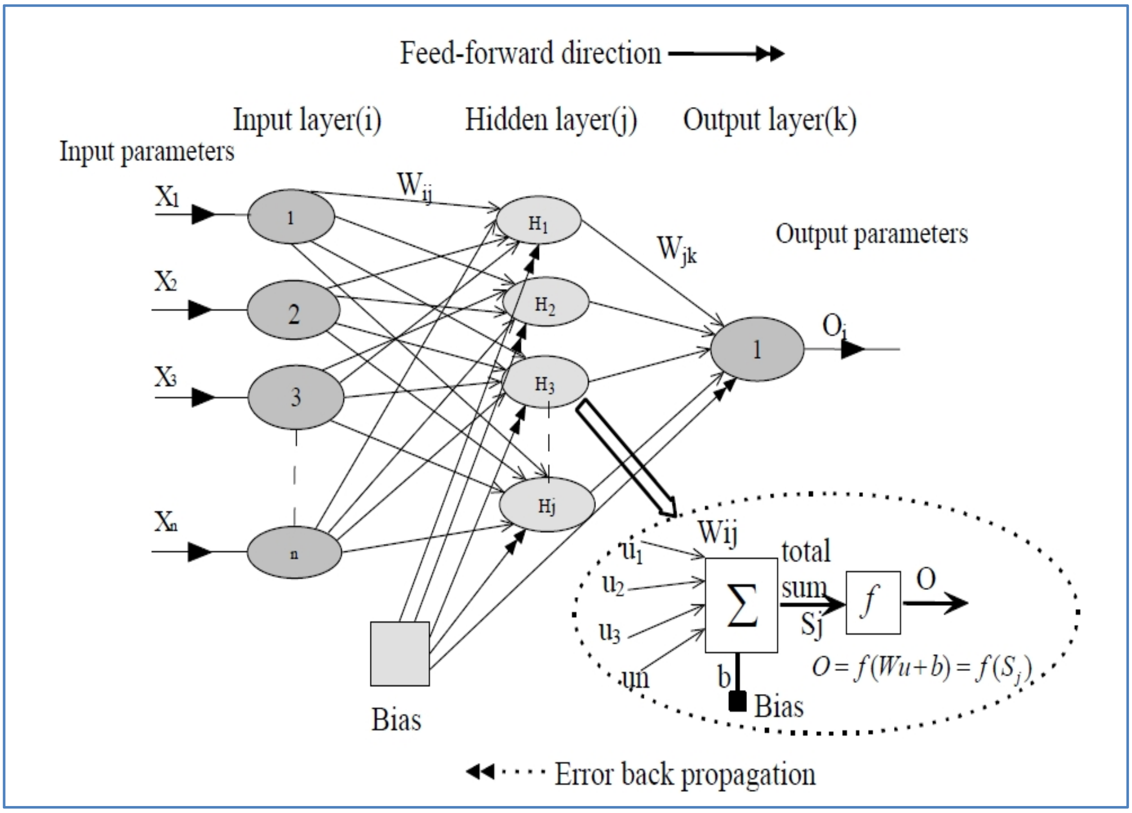


Fig.3.4:An artificial neuron showing its functions

The function *f*(x)is called as an activation function, the activation function enable a network to map any non-linear process. The most commonly used function is the sigmoidal function expressed as:

$$f(x) = \frac{1}{1+ e^{(-x)}} \qquad \text{…………..} \ (3.5)$$

3.5.2 Design of ANN model

ANNs are mathematical models that consist of simple, densely interconnected elements known as neuron, which are typically arranged in layer. An ANN received signals through the input units and these signals are propagated and transformed through the network towards the output neuron(s).

An ANN is characterized by its Architecture that represents the pattern of connection between nodes, its method of determining the connection weights and the activation function. The steps were used for sediment yield modelling by using the concept of ANN. Neural Network Architecture considered for crop yield is shown in Fig. 3.4.

3.5.2.1 Feed-forward neural network (FFNN)

Feed-forward neural network, where the data flow input to output units is strictly feed forward. The data processing can extend over multiple (layer of) units, but no feedback connections are present, that is, connections extending from outputs of units in the same layer or previous layers. In this study, so-called Feed Forward ANN (FANN) is used, in which information always travels in the direction of the ANN output without delay. A conceptual diagram of FANN has shown in Fig. 3.5.

3.5.2.2 Architecture of artificial neural network (ANN)

This important step, involves the determination of the ANN architecture and selection of a training algorithm. An optimal architecture may be considered the one yielding the best performance in term of error minimization while retaining a simple and compact structure. In absence of a unified theory for the design of an optimal ANN architecture, a trial and error procedure is generally applied to decide the number of hidden layer and number of nodes in each hidden layer. Since most of the past research reported the sufficiency of one hidden layer architecture,it has been adopted for the design of network.

In this study, network-growing technique (Gallant, 1986; kwok and Yeung, 1995) is used to arrive at the final architecture of network. This typically starts with a small network and adds node when a suitable chosen measure stops decreasing. The initial number of neuron is taken equal to number of input vectors and the same number have been added for growing the network keeping maximum neurons equal ti ten times of the input vector. A maximum of 3000epochs/cycle is allowed to judge the convergence of the network.

In order to generate an output vector that is as close possible to target vector from an ANN, a training process (also called as learning) is employed to find optimal weight matrices and bias vectors that minimize a predetermined error function. About 90% of the work on ANN based modelling in hydrology reported the use of multilayer feed forward neural network trained by the standard back propagation (Coulibaly et al., 1999). In this study, supervised training procedure is employed. The three training algorithms i.e., gradient-descent back propagation (BP), Levenberg-Marquardt (LM) and radial basis functions (RBF) are employed to evaluate their suitability. A program is written in MATLAB 6.5; Neural Network Toolbox to design the ANN architecture and for defining training algorithm, error goal etc. In the present study, the number of hidden layers is taken as one, and the number of neurons in the hidden layer is varied. The ANN algorithms, architecture and other parameters used for modelling the runoff.

3.5.3 Normalization of data

Due to nature of the algorithm, large values slowdown training because the gradient of the sigmoid function at extreme values approximate zero. Also, because of the nature of the logistic activation function used in the output layer, outputs from the network are constrained to the range [0, 1]. In addition, because each predictor can cover a different range of values, it is prudent to rescale each input to common range so that one predictor does not dominate all other.

To avoid this problem input and output data in both the training and testing sets were rescaled using an appropriate transformation. The transformation may be done either by standardization or normalization. This is done to restrict their range within the interval of 0-1 because the processing elements, PE’s (number of neurons) of the middle layer are assigned a Log-Sigmoid activation function. The shape of Log-Sigmoid function plays an important role in ANN learning.

Mean and Stand. Dev. (mapstd), an approach for scaling network inputs and targets is to normalize the mean and standard deviation of the training set. The function mapstd normalizes the inputs and targets so that they will have zero mean and unity standard deviation. The normalized inputs and targets (rn and sn) that are returned will have zero means and unity standard deviation. They effectively become a part of the network, just like the network weights and biases. After that these outputs converted back into the same units.

3.5.4 Selection of input vector

The selection of an appropriate input vector that may allow an ANN to map the desired output is an important prerequisite for successful application of artificial neural networks. Selection of key input parameters help in avoiding loss of information, whereas inclusion of spurious variable tends to confuse the training process. A sensitivity analysis can be used to determine the relative importance of a variable when sufficient data is available. The input variables that do not have significant effect on the performance of an artificial neural network can be trimmed, resulting in a more compact network.

3.5.5 Selection of optimum number of neuron

ANN models have been developed and evaluated for their performance had 20 numbers of neurons. Model with more number of neurons takes more time for training. Also very high number of neurons does not help in generalization of the model. Hence, to determine the optimum number of neurons, at which network should have to perform is best; trial and error method is applied. Selection of optimum number of neurons is essential part ANN model development. Neurons in the hidden layer have been varied from 1 to 20 number of neurons that performed better was considered for further improvement of ANN model.

3.5.6 Training algorithms (Back-propagation network)

During the training of the network, the weights and biases of the network are iteratively adjusted to minimize the network performance function. These entire algorithms use the gradient of the performance function to determine how to adjust to weight to minimize performance. The gradient is determined using a technique called backward through the network. The basic back propagation algorithms used here are, in which weight are moved in the direction of the negative gradient. Following (Table 3.3) back propagation training algorithm have been used for evaluate the performance of ANN. The algorithm that performed better during training and validation was selected for further improvement.

Table 3.3 ANN Models with varying training algorithms

S.No.	**Acronyms**	**Training Algorithms**
1	LM-trainlm	Levenverg Marquardt Algorithm
2	GD-traingd	Batch Gradient Decent
3	GDM-traingdm	Batch Gradient Decent With Momentum
4	GDA-traingda	Variable Learning Rate
5	GDX-traingdx	Variable Learning Rate
6	RP-trainrp	Resilient Back Propagation
7	CGF-traincgf	Fletcher-reeves Update (Conjugated Gradient)
8	CGP-traincgo	Polak- Ribiere Update (Conjugated Gradient)
9	CGB-traincgb	Powell/Beale Restarts (Conjugated Gradient)
10	SCG-trainscg	Scaled Conjugated Gradient
11	BFGS-trainbfg	Quasi- Newton Algorithm
12	OSS-trainoss	Quasi- Newton- One Step Secant

There are different types of ANNs. Some of the most popular types include the multi-layer perceptron, which is generally trained with the back propagation of error algorithm, radial basis function. ANNs can also be classified as feed forward or recurrent (feedback) depending on how data is processed.

Networks can also be classified as supervised or unsupervised. The networks, which are used in this study, are described below. All ANNs were trained using supervised training algorithms that tried to minimize a performance measure (often termed objective function from the point of view of validation), namely the Mean Squared Error (MSE). The merits of using a good algorithm are threefold:

- Better accuracy leads to better ANN performance
- Faster convergence leads to smaller calculation times, and
- Lower spread in the performance makes it easier and more honest to evaluate and compare ANNs.

The simplest implementation of back-propagation learning updates the network weights and biases in the direction in which the performance function decreases most rapidly, the negative of the gradient. One iteration of this algorithm can be written:

$$X_{k+1} = X_k - \alpha_k g_k \quad \text{………….. (3.6)}$$

Where,

X_k = Vector of current weights,

g_k = the current gradient,

α_k = Learning rate.

3.5.6.1 Classification of networks (training algorithms)

There are different types of ANN. Some of the most popular types include the multi-layer perceptron, which is generally trained with the back propagation of error algorithm, learning vector quantization, radial basis function, and Hopfield and Kohonen networks. ANNs can also be classified as feed forward or recurrent (feedback) depending on how data is processed. Networks can also be classified as supervised or unsupervised. The networks, which are used in this study, are described below.

3.5.6.2 Back-propagation network

Back-propagation (BP) networks are the most widely used ANNs. It is a gradient descent technique that minimizes the network error function. In fact, the name back-propagation comes from the error term, which is propagated through the network during learning and used to change the weights on equation:

$$\Delta w_{ij} = -\eta . \frac{\delta E}{\delta w_{ij}} + m . \Delta w_{ij}(n-1) \quad \text{…………….. (3.7)}$$

A similar equation can be written for correction of bias values. In Eq^n. (3.7), η and m are called as learning rate and momentum coefficient, respectively. The momentum factor speeds up training in very flat regions of the error surface and helps prevent oscillating in the weights. A learning rate is used to increase the chance of avoiding the training process being trapped in local minima instead of global minima. Since BP is a first order method based on gradient descent with the direction vector being set equal to the negative of the gradient vector, the solution often follows a zigzag path while trying to reach a minimum error position, which may slow down the training process. It is also possible for the training process to be trapped in the local minimum despite the use of a learning rate.

3.5.6.3 Levenverg-Marquardt algorithm (trainim)

Like the quasi-Newton methods, the Levenverg-Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares (as is typical in training feed forward networks), then the Hessian matrix can be approximated as

$$H=J^TJ \qquad \text{…………..} \ (3.8)$$

Andthe gradient can be computed as

$$g= J^Te \qquad \text{…………..} \ (3.9)$$

Where,

J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and e is a vector of network errors. The Jacobian matrix can be computed through a standard backpropagation technique that is much less complex than computing the Hessian matrix. The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$Xk+1=xk-[J^TJ+uI]-1J^Te \qquad \text{…………..} \ (3.10)$$

When the scalar μ is zero, this is just Newton’s method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. Newton’s method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible. Thus, μ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm.

Epochs, show, goal, time, mm_grad, max_fail, mu, mu_dec, mu_inc, mu_max, and mem_reduc. The first six parameters were discussed earlier the parameter mu is the initial value for μ. This value is multiplied by mu_dec .,whenever the performance function is reduced by a step. It is multiplied by i_inc whenever a step would increase the performance function. If mu becomes larger than mu_max, the algorithm is stopped The parameter mem_reduc is used to control the amount of memory used by the algorithm parameters show and epochs are set to 5 and 300, respectively In practice LM is much faster and better for a variety of the problems (Coulibaly et al., 1999).

3.5.6.4 Batch gradient descent (traingd)

In batch mode the weights and biases of the network are updated only after the entire training set has been applied to the network. The gradients calculated at each training example are added together to determine the change in the weights and biases.

The batch steepest descent training function is traingd. The weights and biases are updated in the direction of the negative gradient of the performance function. There are seven training parameters associated with traingd: Epochs, show, goal, time, mm_grad, max_fail, Ir.

The learning rate Ir is multiplied times the negative of the gradient to determine the changes to the weights and biases. The larger the learning rate, the bigger the step. If the learning rate is made too large, the algorithm becomes unstable. If the learning rate is set too small, the algorithm takes a long time to converge.The training status is displayed for every showiterations of the algorithm. (If show is set to NaN, then the training status is never displayed). The other parameters determine when the training stops. The training stops if the number of iterations exceeds epochs, if the performance function drops below goal, if the magnitude of the gradient is less than mingrad, or if the training time is longer than time seconds, max_fail, which is associated with the early stopping technique.

3.5.6.5 Batch gradient descent with momentum (traingdm)

In addition to traingd, there is another batch algorithm for feed forward networks that often provides faster convergence: traingdm, steepest descent with momentum. Momentum allows a network to respond not only to the local gradient, but also to recent trends in the error surface. Acting like a lowpass filter, momentum allows the network to ignore small features in the error surface. Without momentum a network can get stuck in a shallow local minimum. With momentum a network can slide through such a minimum.

Momentum can be added to backpropagation learning by making weight changes equal to the sum of a fraction of the last weight change and the new change suggested by the backpropagation rule. The magnitude of the effect that the last weight change is allowed to have is mediated by a momentum constant, mc, which can be any number between 0 and 1. When the momentum constant is 0, a weight change is based solely on the gradient. When the momentum constant is 1, the new weight change is set to equal the last weight change and the gradient is simply ignored. The gradient is computed by summing the gradients calculated at each training example, and the weights and biases are only updated after all training examples have been presented. If the new performance function on a given iteration exceeds the performance function on a previous iteration by more than a predefined ratio, max_perl_inc, (typically 1.04), the new weights and biases are discarded, and the momentum coefficient mc is set to zero. The batch form of gradient descent with momentum is invoked using the training function traingdm. The traingdm function is invoked using the same steps shown above for the traingd function, except that you can set the mc, Ir, and max_perf_inc learning parameters.

3.5.6.6 Variable learning rate (traingda, traingdx)

With standard steepest descent, the learning rate is held constant throughout training. The performance of the algorithm is very sensitive to the proper setting of the learning rate. If the learning rate is set too high, the algorithm can oscillate and become unstable. If the learning rate is too small, the algorithm takes too long to converge. It is not practical to determine the optimal setting for the learning rate before training, and, in fact, the optimal learning rate changes during the training process, as the algorithm moves across the performance surface.

The performance of the steepest descent algorithm can be improved if the learning rate is allowed to change during the training process. An adaptive learning rate attempts to keep the learning step size as large as possible while keeping learning stable. The learning rate is made responsive to the complexity of the local error surface. An adaptive learning rate requires some changes in the training procedure used by traingd. First, the initial network output and error are calculated. At each epoch new weights and biases are calculated using the current learning rate. New outputs and errors are then calculated.

As with momentum, if the new error exceeds the old error by more than a predefined ratio, max_pen_inc (typically 1.04), the new weights and biases are discarded. In addition, the learning rate is decreased (typically by multiplying by Ir_dec = 0.7). Otherwise, the new weights, etc., are kept. If the new error is less than the old error, the learning rate is increased (typically by multiplying by Ir_inc = 1.05).

This procedure increases the learning rate, but only to the extent that the network can learn without large error increases. Thus, a near-optimal learning rate is obtained for the local terrain. When a larger learning rate could result in stable learning, the learning rate is increased. When the learning rate is too high to guarantee a decrease in error, it is decreased until stable learning resumes.

3.5.6.7 Resilient backpropagation (trainrp)

Multilayer networks typically use sigmoid transfer functions in the hidden layers. These functions are often called “squashing” functions, because they compress an infinite input range into a finite output range. Sigmoid functions are characterized by the fact that their slopes must approach zero as the input gets large. This causes a problem when you use steepest descent to train a multilayer network with sigmoid functions, because the gradient can have a very small magnitude and, therefore, cause small changes in the weights and biases, even though the weights and biases are far from their optimal values.

The purpose of the resilient back propagation (Rprop) training algorithm to eliminate these harmful effects of the magnitudes of the partial derivatives. Only the sign of the derivative is used to determine the direction of the weight update; the magnitude of the derivative has no effect on the weight update. The size of the weight change is determined by a separate update value. The update value for each weight and bias is increased by a factor delt_inc whenever the derivative of the performance function with respect to that weight has the same sign for two successive iterations. The update value is decreased by a factor delt_dec whenever the derivative with respect to that weight changes sign from the previous iteration. If the derivative is zero, then the update value remains the same. Whenever the weights are oscillating, the weight change is reduced. If the weight continues to change in the same direction for several iterations, then the magnitude of the weight change increases

The training parameters for trainrp are epochs, show, goal, time, mm grad, max_fail, delt_inc, deltdec, deltao, and deltamax. The first eight parameters have been previously discussed. The last two are the initial step size and the maximum step size, respectively. The performance of Rprop is not very sensitive to the settings of the training parameters. For the example below, most of the training parameters are left at the default values; show is reduced below its previous value, because Rprop generally converges much faster than the previous algorithms.

3.5.6.8 Conjugate gradient algorithms

The basic backpropagation algorithm adjusts the weights in the steepest descent direction (negative of the gradient), the direction in which the performance function is decreasing most rapidly. It turns out that, although the function decreases most rapidly along the negative of the gradient, this does not necessarily produce the fastest convergence. In the conjugate gradient algorithms a search is performed along conjugate directions, which produces generally faster convergence than steepest descent directions. This section presents four variations of conjugate gradient algorithms.

Fletcher-Reeves update (traincgf)

All the conjugate gradient algorithms start out by searching in the steepest descent direction (negative of the gradient) on the first iteration. A line search is then performed to determine the optimal distance to move along the current search direction. Then, the next search direction is determined so that it is conjugate to previous search directions. The general procedure for determining the new search direction is to combine the new steepest descent direction with the previous search direction. The various versions of the conjugate gradient algorithm are distinguished by the manner in which the constant k is computed.

The training parameters for traincgf are epochs, show, goal, time, mm_grad, max_fail, srchFcn, scal_tol alpha, beta, delta, gamalow_lim, up_urn, maxstep, minstep, and bmax.).

Polak-Ribiéreupdate (traincgp)

Another version of the conjugate gradient algorithm was proposed by Polak and Ribiére. For the Polak-Ribiére update, the constant k is computed which is the inner product of the previous change in the gradient with the current gradient divided by the norm squared of the previous gradient. The training parameters for traincgp are the same as those for traincgf the parameters show and epochs are set to the same values as they were for traincgf.

Powell-Beale restarts (traincgb)

For all conjugate gradient algorithms, the search direction is periodically reset to the negative of the gradient. The standard reset point occurs when the number of iterations is equal to the number of network parameters (weights and biases), but there are other reset methods that can improve the efficiency of training One such reset method was proposed by Powell, based on an earlier version proposed by Beale This technique restarts if there is very little orthogonally left between the current gradient and the previous gradient.

Scaled conjugate gradient (trainscg)

Each of the conjugate gradient algorithms discussed so far requires a line search at each iteration. This line search is computationally expensive, because it requires that the network response to all training inputs be computed several times for each search The scaled conjugate gradient algorithm (SCG), developed by Moller, and was designed to avoid the timeconsuming line search This algorithm combines the model-trust region approach (used in the Levenberg-Marquardt algorithm, described in Levenberg-Marquardt), with the conjugate gradient approach.

3.5.6.9 Quasi-newton algorithms bfgs algorithm (trainbfg)

Newton's method is an alternative to the conjugate gradient methods for fast optimization. Newton's method often converges faster than conjugate gradient methods. Unfortunately, it is complex and expensive to compute the Hessian matrix for feed forward neural networks. There is a class of algorithms that is based on Newton's method, but which doesn't require calculation of second derivatives. These are called quasi-Newton (or secant) methods. They update an approximate Hessian matrix at each iteration of the algorithm. The update is computed as a function of the gradient. The quasi-Newton method that has been most successful in published studies is the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) update. This algorithm is implemented in the trainbfg routine. The training parameters for trainbfg are the same as those for traincgf. This algorithm requires more computation in each iteration and more storage than the conjugate gradient methods, although it generally converges in feweriterations. The approximate Hessian must be stored, and its dimension is n x n, where n is equal to the number of weights and biases in the network. For smaller networks, however, trainbfg can be an efficient training function.

One step secant algorithm (trainoss)

Because the BFGS algorithm requires more storage and computation in each iteration than the conjugate gradient algorithms, there is need for a secant approximation with smaller storage and computation requirements. The one step secant (OSS) method is an attempt to bridge the gap between the conjugate gradient algorithms and the quasi-Newton (secant) algorithms. This algorithm does not store the complete Hessian matrix; it assumes that at each iteration, the previous Hessian was the identity matrix. This has the additional advantage that the new search direction can be calculated without computing a matrix inverse. The training parameters for trainoss are the same as those for traincgf. This algorithm requires less storage and computation per epoch than the BFGS algorithm. It requires slightly more storage and computation per epoch than the conjugated gradient algorithms. It can be considered a compromise between full quasi-Newton algorithms and conjugate gradient algorithms.

3.5.7 Distribution and format of training data

Yield, maximum and minimum temperature and rainfall data which are collected are used to train the network and to calibrate the network. This is normally considered as a satisfactory method selection of the training data. To assess the effect of training data on the performance of network, training is performed with different number of experimental dataset. Firstly a relatively large set of data including all ranges of input variables is used to train the network. The performance of the network during validation is monitored and observation made on the level of accuracy obtained during validation. Subsequently the same above mentioned procedure is repeated using lesser number of data during training and the performance of the network during validation is monitored. In this way the sensitivity of the performance of the network to the size of training data is evaluated. Different combinations are used to train and to calibrate the data, according to total length of data shown in Table 3.4.

Table 3.4 ANN Models calibrated and validated with varying length of records.

S.N.	Model	Calibrated with % of total length of	Validated with % of total length of data
1.	M50	50	50
2.	M60	60	40
3.	M70	70	30
4.	M80	80	20

3.6 Design of ANN model for crop yield

The neural network utility file is edited in MATLAB (6.5 Version). The input variables selection, input data source file, network option, training function, setting for the data for training, validation, plotting the predicting values and saving the network is created and run in MATLAB software.

The model was first implemented into an Excel spread sheet to serve as the guideline to the MATLAB program defining the model. The spread sheet contained an elaborate scheme of links that would allow a direct real-time visual manipulation of the governing parameters to yield the best graphical fit.

When the input variables were selected, then the source data input file is imported for training and validation. In this study, models with various input variables were developed and their performance was evaluated. Various models developed with varying inputs are given in Table 3.5. The variables were selected according to the model P-24, P-4, W-24, W-8 for developing and evaluating the ANN models. The ANN model architecture is single layer feed forward network, which is the most commonly used neural network for the prediction of the nonlinear process. The number of the hidden layer is one. The transfer function from input to hidden layer is Tan-Sigmoid Transfer Function (Tansig) and from hidden layer to output layer is Linear Transfer function (Purelin). Figure of transfer function for input and output layer shown in Fig 3.5. The Back propagation training function has been selected, which is the most common and accurate as reported by many workers. The performance function for training and testing the networks used are MSE (Mean Squared Error). The variouscombinations of hidden nodes and training function is done to arrive at optimum combinations to give less error. The network iterations (Epochs) were kept at 500.

Table 3.5 ANN models with varying input variable

Model	Predictor	Output(Kg/ha)
P-24	T26 T28 T30 T32 T33 T34 T37 M24 M25 M27 M28 M29 M31 M34 M35 M36 M38 R25 R30 R31 R32 R37 R38 R39	PCy
P-4	T24 M25 R24 R31	PCy
W-24	T44 T46 T52 T1 T2 T3 M46 M49 M7 R44 R45 R46 R47 R48 R50 R3 R5 R6 R7 R8 R10 R11 R12 R13	WC_Y
W-8	T11 T48 M44 M45 M47 M49 R8 R50	WC_Y

PCy=paddy crop yield (kg/ha)

WCy= Wheat crop yield (kg/ha)

Tn= Average maximum temperature of n^{th} week (^oC)

Mn = Average minimum temperature of n^{th} week (^oC)

Rn =Rainfall during nth week (mm)

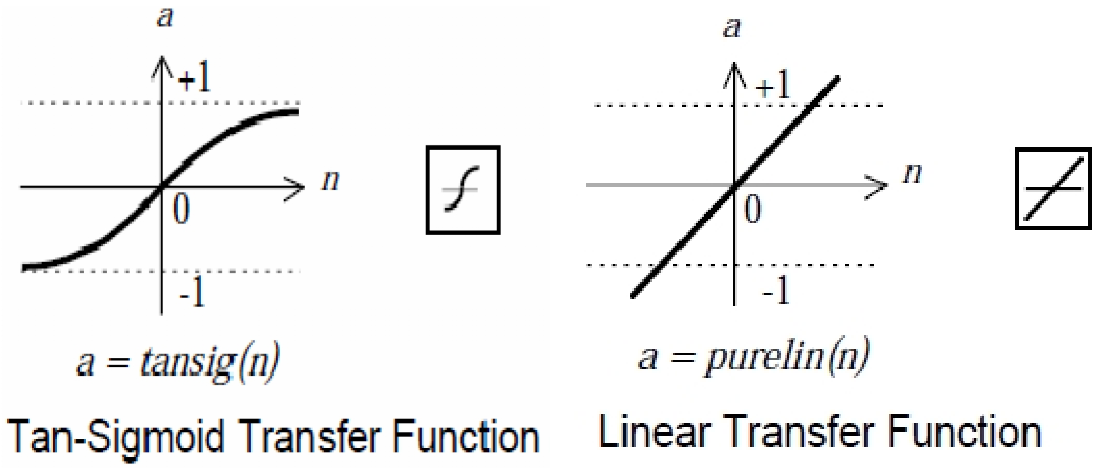


Fig. 3.5:Transfer function used in ANN Model

Table 3.6 ANN algorithms, architecture and parameters used in ANN model

Algorithm	**Back- propagation- L.M.**
Training Functions	Trainlm
Number of Neurons	7
Number of hidden layers	1
Scaling method	Normalization
Activation function of hidden layer	Tan-sigmoid
Activation function of output layer	Purelin
Number of Epochs	500

The network training process is displayed on the screen with the minimum error in comparison with the set goal (0). The program gives the minimum value of MSE is saved as a model with particular identification name. Architecture for ANN models shown in Table 3.6. The model is saved with name P-24, P-4, W-24, W-8Matlab window of ANN model is shown in Fig.3.6.

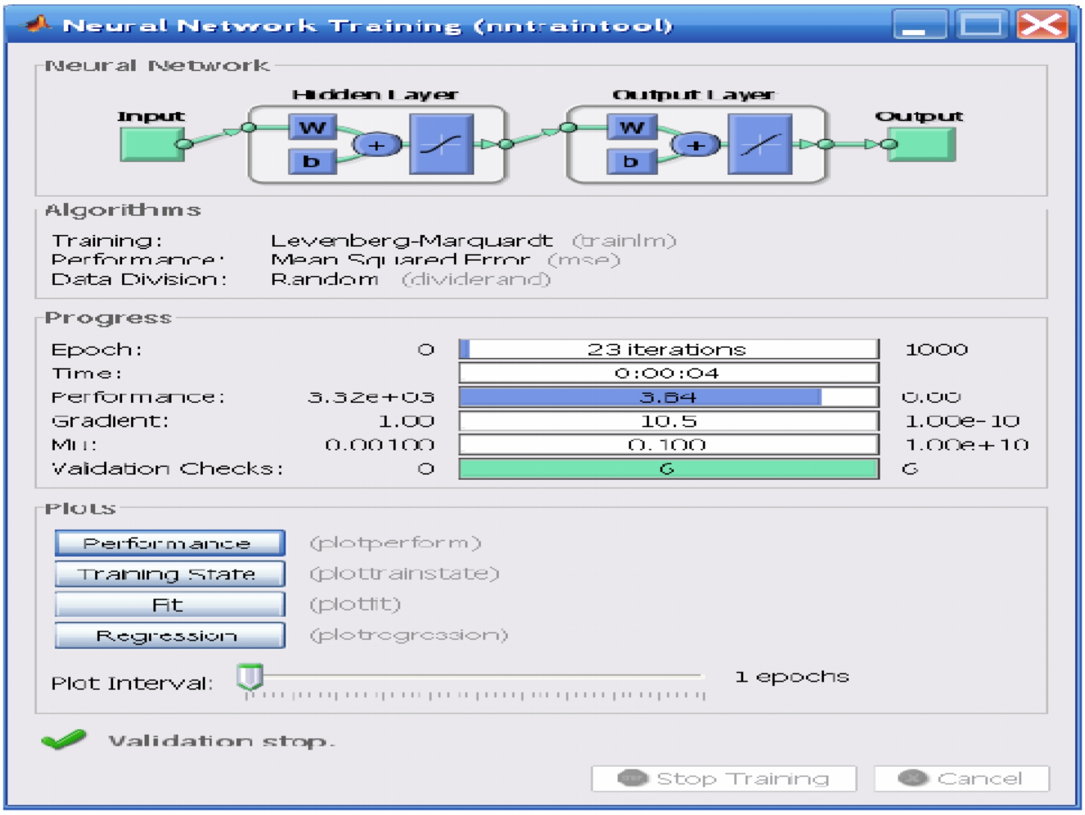


Fig.3.6: MATLAB window of ANN model

3.7 Performance Indicators

Coefficient of determination (R^2), Correlation coefficient (R), Mean squared error (MSE), Root mean squared error (RMSE) and Mean Absolute error (MAE), were used to as the model development parameters as well as the evaluation criteria's. The Correlation coefficient(R) measures the degree to which two variables are linearly related. MSE and MAE provided different types of the information about the predictive capabilities of the model. The Coefficient of determination (R^2) measures the goodness of fit relevant to sediment yield value in the watershed, whereas the MAE, yields a more balanced perspective of the goodness of fit as moderate output values (Karunanithi*et al*., 1994).

Correlation Coefficient (R): The Correlation Coefficient is given by the following equation.

$$R = \frac{\sqrt{\sum_{i=1}^{n}\{O_i - O_{avg}\}\{S_i - S_{avg}\}}}{\left[\sum_{i=1}^{n}(O_i - O_{avg})^2\right]\left[\sum_{i=1}^{n}(S_i - S_{avg})^2\right]} \quad \text{..................(3.11)}$$

Coefficient of Determination (R^2): The value of R^2 is given by:

$$R^2 = \frac{\left[\sum_{i=1}^{n}\{O_i - O_{avg}\}\{S_i - S_{avg}\}\right]}{\sqrt{\left[\sum_{i=1}^{n}(O_i - O_{avg})^2\right]\left[\sum_{i=1}^{n}(S_i - S_{avg})^2\right]}} \quad \text{................ (3.12)}$$

Root Mean Square Error (RMSE): is given by

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(O_i - S_i)^2} \quad \text{................. (3.13)}$$

Mean Absolute Error (MAE): The mean absolute error (MAE) is defined as

$$MAE = \frac{1}{n}\sum_{i=n}^{n}(|Q_{obs} - Q_{est}|) * 100 \quad \text{.................. (3.14)}$$

Mean Square Error (MSE): The mean Square error (MSE) is defined as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(O_i - S_i)^2 \text{.................. (3.15)}$$

Where,

O_i = Observed data,

O_{avg} = Mean of the observed data,

S_i = Simulated value,

S_{avg} = Mean of the simulated value,

n = Number of observations

RESULT AND DISCUSSION

The chapter deal with the result obtained for model development and for estimation of paddy and wheat yield by adopting the procedure as discussed in chapter III and comparison of developed models has been discussed in this chapter.

4.1 Development of multiple regression model

Crop yield models have been developed for estimation of paddy and wheat crop using multiple linear regression(MLR).As discussed in chapter III, Three variable maximum temperature, minimum temperature and rainfall weekly values at sixteen week during growing season have been used as predictor variables. In order to reduce the predictor variables, enter method and stepwise method of regression analysis has been used.

4.1.1 Model for paddy yield

When enter method of MLR used, the result obtained with 24 predictor variables during training of the Model with 60 length of dataset is depicted in Fig. 4.1. The correlation coefficient (R) and coefficient of determination (R^2) obtained for the Model is 0.8776 and 0.7702 respectively. The model can be expressed as:

PC_y=3009.38+39.891T26-72.747T28+41.924T30-22.633T32+46.861T33-4.765T34+50.372T37-185.277M24+127.872M25-113.283M27+222.558M28-10.582M29+80.608M31-70.747M34+70.364M35+17.159M36350.778M38+2.963R25-0.47R30+1.096R31-1.33R32+3.472R37-3.176R38+1.357R39. ……(4.1)

The above model when validation with remains 40% of data, the coefficient of determination (R^2) is comparatively quite low i.e.0.0316. The relationship between observed paddy yield predictor paddy yield is depicted in Fig.4.2.

The relationship between predicted yield of paddy (y) and observed yield (x) during model training and validation is given in equation 4.2 and 4.3 respectively.

$y = 1.0206x - 65.274$ …………………………………… (4.2)

$y = -0.3919x + 962.3$……………………………………….(4.3)

When stepwise regression analysis has made, the model contains four predictor variable i.e. rainfall during 24th SMW, 31th SMW, minimum temperature of 25th SMW and maximum temperature of 24th SMW the model is expressed in equation 4.4. Relationship between observed and predicted yield is depicted in fig. 4.3 and 4.4 during training and validation of the model respectively. The mathematical relation between observed and predicted yield of paddy by the model is given in equation 4.5 and 4.6 during training and validation respectively. Performance of the two models for paddy with enter (PME) and stepwise regression method (PMS) presented in Table 4.1.

PC_y=-1809.575+5.093R24+132.51M25-30.014T24+0.675R31 … ..(4.4)

$y = 0.6948x + 202.8$ ……… (4.5)

$y = -0.2477x + 760.32$ …….. (4.6)

Table 4.1 Performance of paddy regressionmodels (training with 60% of data).

S.No.	Model	R	R^2	RMSE	MAE	R^2	
						Trg.	Val.
1	PME	0.8776	0.7702	450.01	291.71	0.7702	0.0316
2	PMS	0.8277	0.6851	252.6	189.05	0.6851	0.0666

It can observed from the Table 4.1 that during training of both the model coefficient of determination(R^2)is satisfactory. However, during validation (R^2) is 0.0316 for enter method model and 0.0666 for stepwise method. The performance of models during validation is not satisfactory.

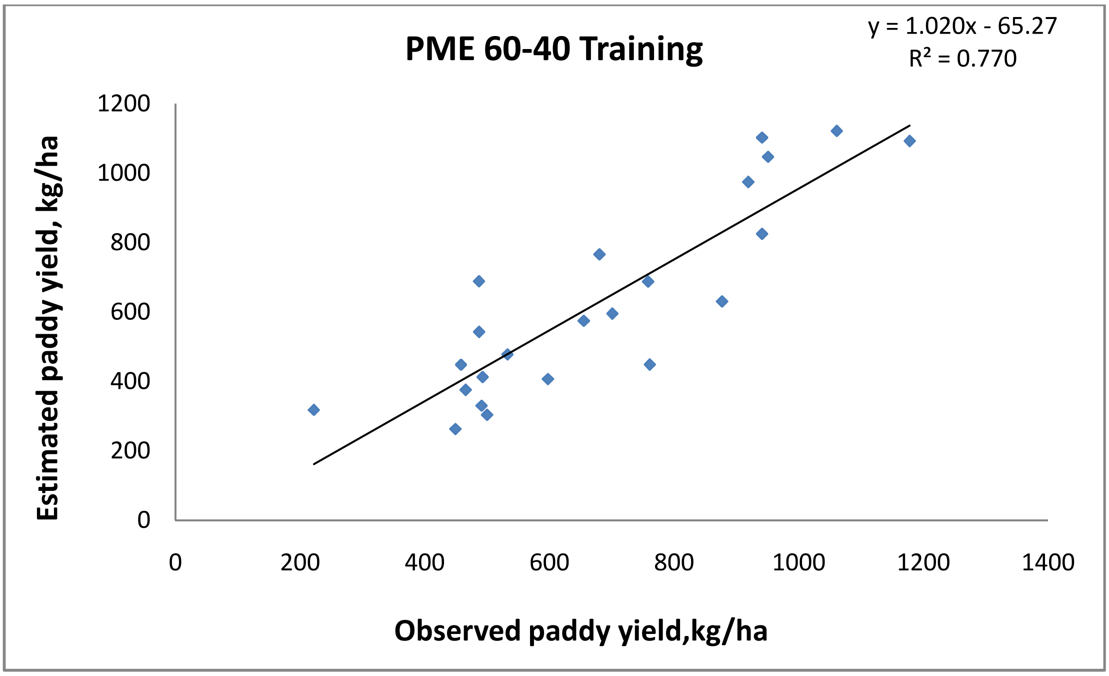


Fig. 4.1Relationship between observed and estimated paddy yield during training of the model MLR-Enter

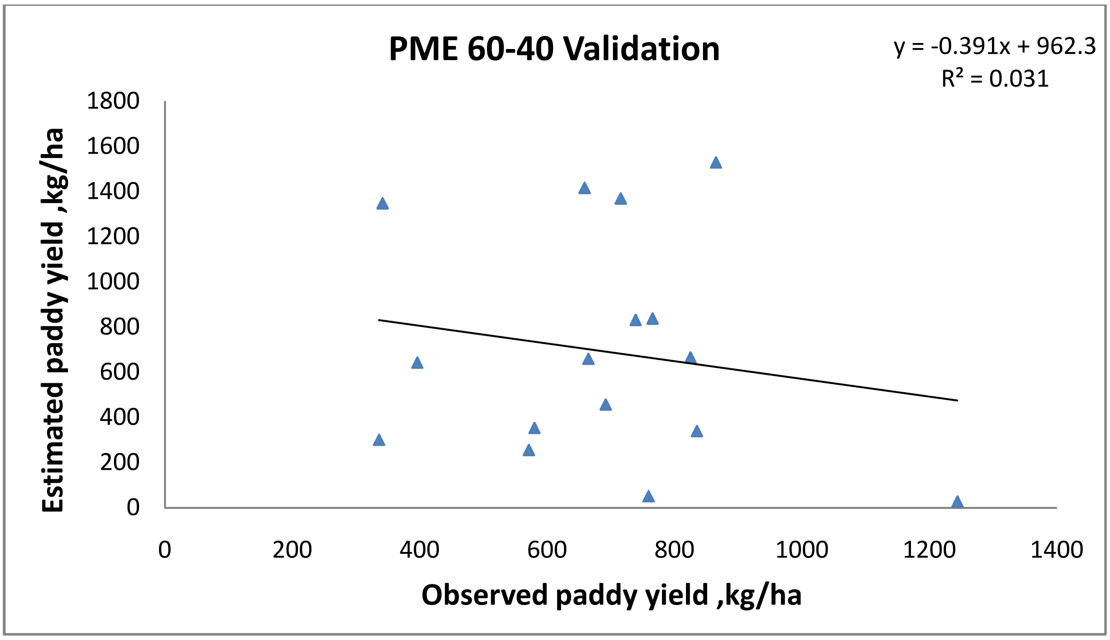


Fig. 4.2 Relationship between observed and estimated paddy yield during validation of the model MLR-Enter

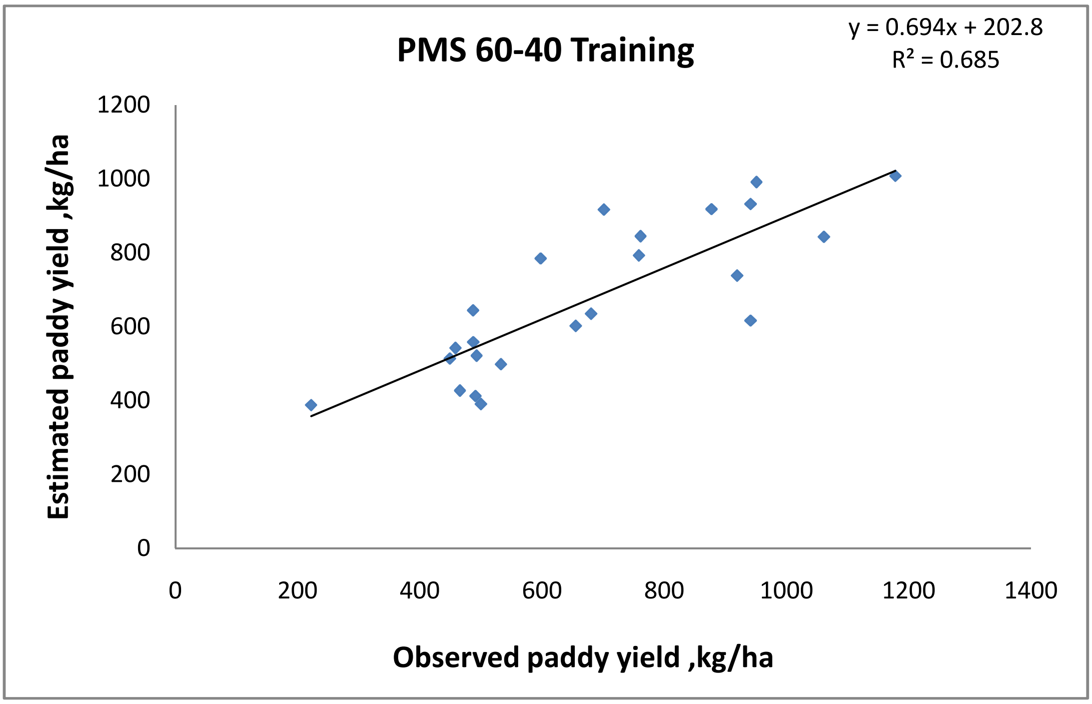


Fig. 4.3 Relationship between observed and estimated paddy yield during training of the model MLR-stepwise

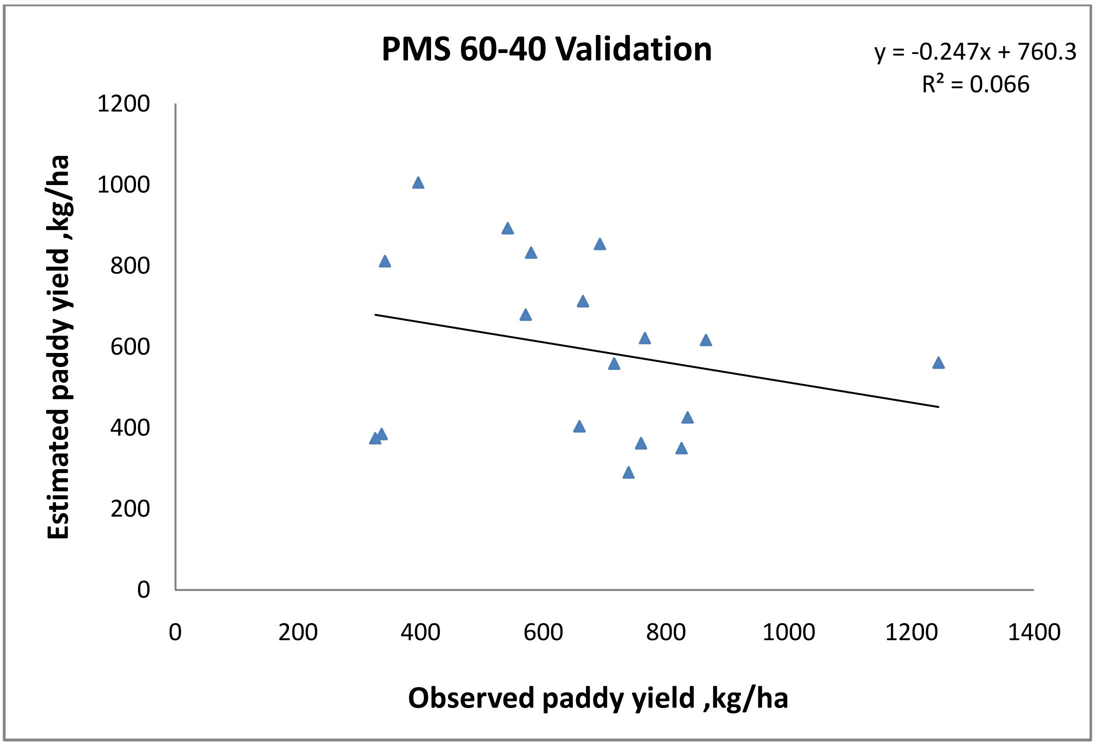


Fig. 4.4 Relationship between observed and estimated paddy yield during validation of the model PMLR-stepwise

4.1.2 MLR model for wheat yield

Temperature and rainfall are two important climatic parameter effecting the yield of wheat crop. Duration of wheat crop is considered as twenty two week starting with 44^{th} SMW to 13^{th} SMW. Initially all 66 predictor variables with weekly magnitude have been taken into consideration. Number of predictor variables have been reduced by adopting enter method and stepwise method of multiple liner regression.

Enter method reduced into selection of twenty four predictor variables. The multiple liner regression model is expressed in equation 4.7. The relationship between observed wheat yield and predicted wheat yield have been shown in fig. 4.5 during model training with 60% data and in Fig 4.6 during model validation with remaining data . It can be observed that R^2 during model training (0.7998) is satisfactory. The R^2 (0.1206) during model validation is quite low and show the negative relation between observed and estimated yield of wheat. The regression equation between observed and estimated wheat yield are expressed in equation 4.8 and 4.9 respectively.

WC_y=2396.715+46.592T44+80.755T46-72.02T52+68.754T1-133.186T2+32.948T3-129.923M46+57.87M49-64.132M7-5.606R44-2.384R45+7.372R46-4.827R47-8.324R48-37.573R50+2.931R3-0.469R5+3.522R6+5.149R7+7.183R8+9.17R10-8.289R11-40.191R12 …….(4.7)

$$y = 0.867x + 294.34 \quad (4.8)$$

$$y = -0.3248x + 2202.4 \quad (4.9)$$

Where,

y=predicted wheat yield and x=observed wheat yield.

When stepwise regression technique was used, predictor variables reduced to 8 in the model as expressed in equation 4.10. The relationship between observed yield and estimated by the model during model development and its validation are depicted in Fig4.5 and 4.6 respectively .The regression equation between observed and e.stimated yield during model training is expressed an equation 4.11 and during validation as equation 4.12.

WC_y=-208.444-203.579M47-59.201R50+189.175M49+85.735M45+15.493R8+33.136T11+69.893M44+58.93T48……………………………..….(4.10)

$$y = 0.9859x - 517.21 \quad \text{….. (4.11)}$$

$$y = -0.5232x + 2272.4 \quad \text{…...(4.12)}$$

Where,

y=predicted wheat yield and x=observed wheat yield.

Performance of the two models developed for wheat yield is summarized in Table 4.2. The MLR model with enter method (WME) resulted RMSE of 571.84 and MAE of 245.98, where as model MLR with stepwise method (WMS) resulted RMSE of 794.91 and MAE of 619.80. It shows that stepwise MLR model is superior than enter model.

Table 4.2 Performance of wheat regression models (training with 60% of data)

S.N.	Model	R	R^2	RMSE	MAE	R^2	
						Trg.	Val.
1	WME	0.8943	0.7998	571.84	245.98	0.7998	0.1206
2	WMS	0.9404	0.8844	794.91	619.80	0.8844	0.1909

From the above result, it can be inferred that MLR model train then well but their performance with independent data set is not satisfactory.

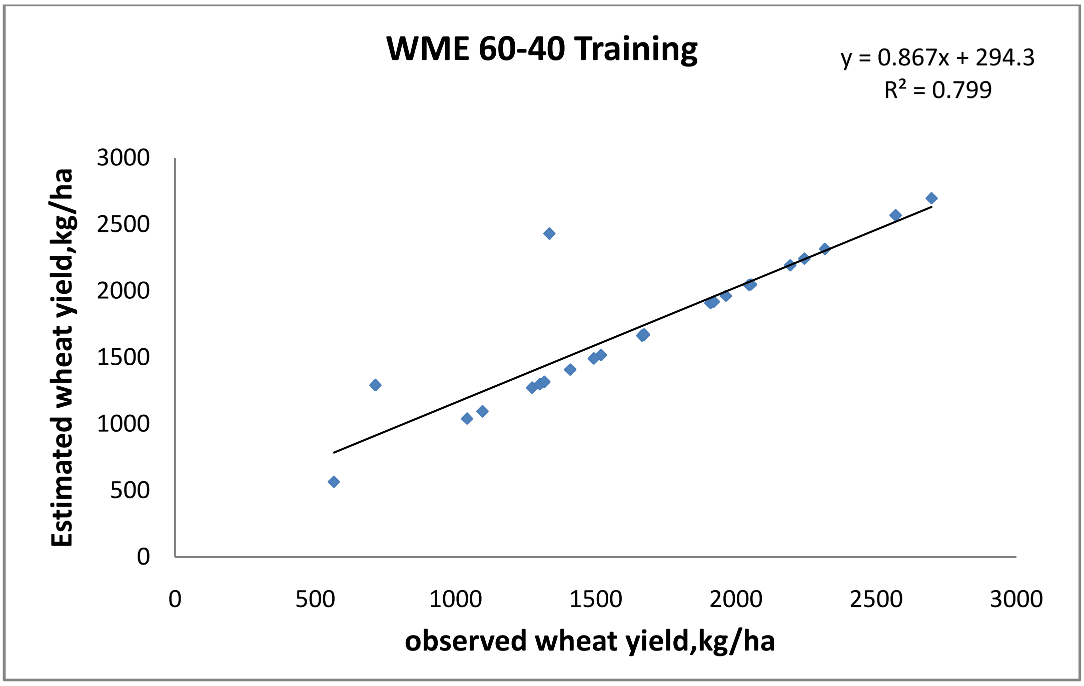


Fig. 4.5 Relationship between observed and estimated wheat yield during training of the model MLR-Enter method

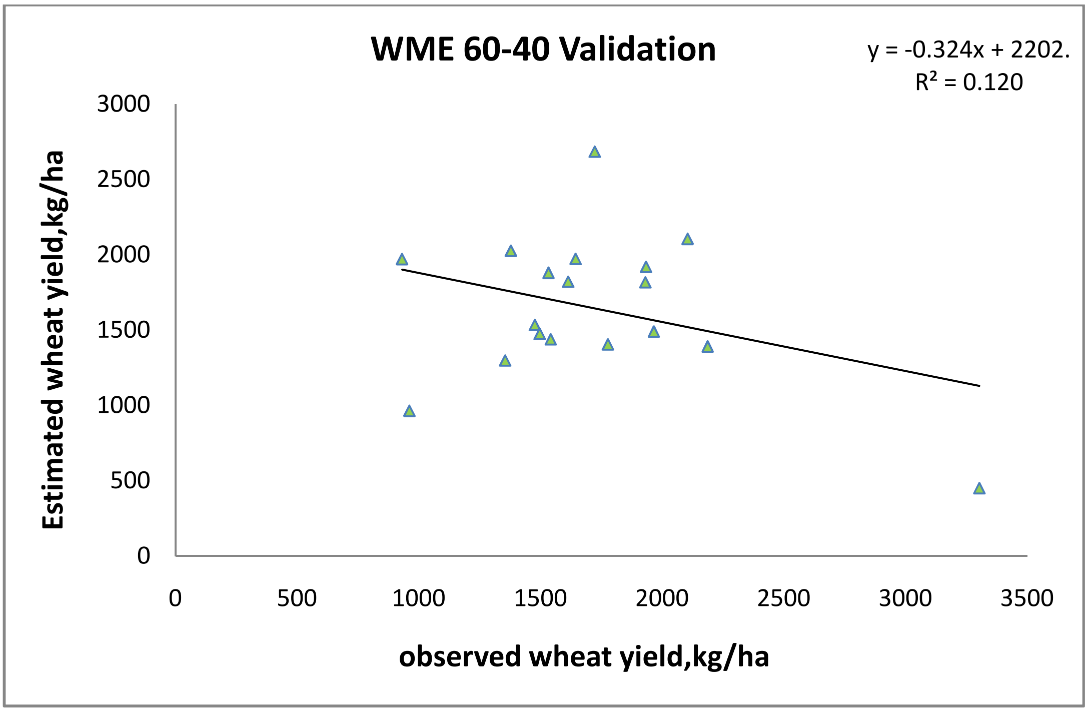


Fig. 4.6 Relationship between observed and estimated wheat yield during validation of the model MLR-Enter method

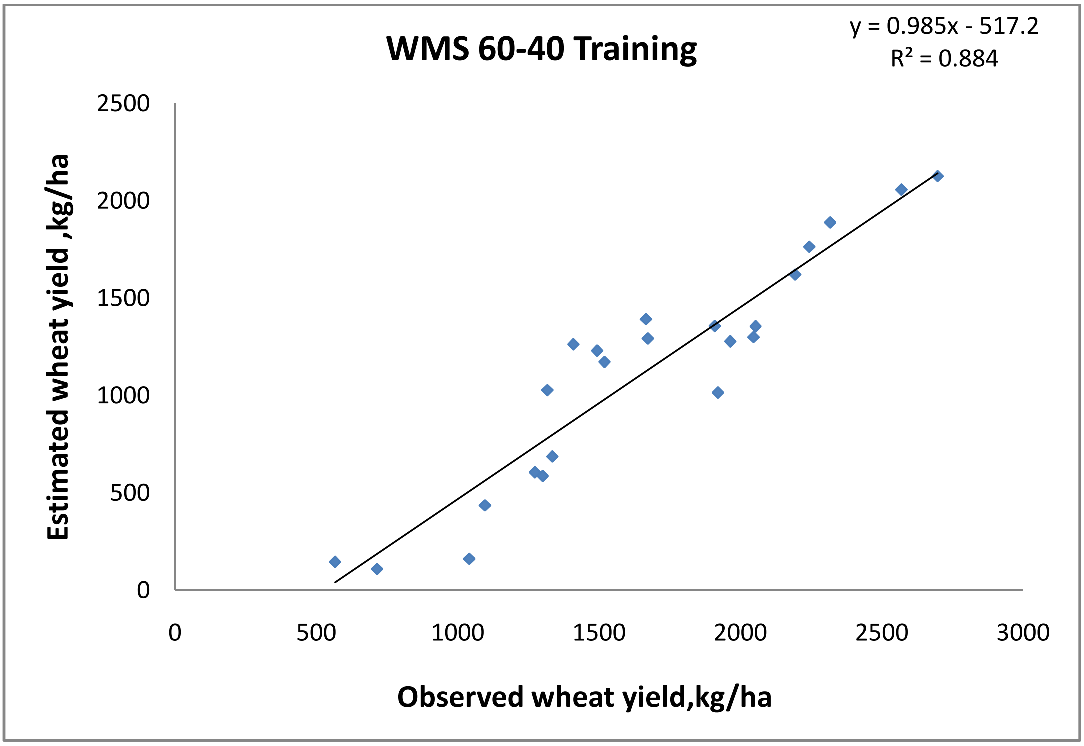


Fig. 4.7Relationship between observed and estimated wheat yield during training of the model MLR-stepwise

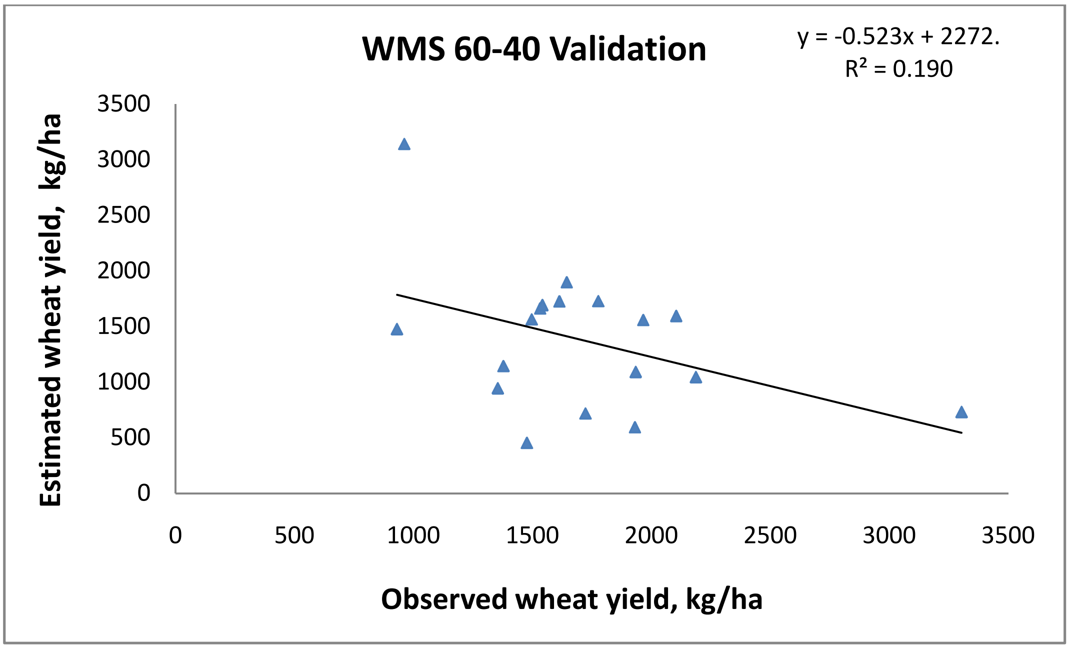


Fig. 4.8 Relationship between observed and estimatedwheat yield during validation of the model MLR-stepwise

4.2 Artificial Neural network based model

As seen in section 4.1, MLR models performance is poor during validation, hence attempt has been made to develop Artificial Neural Network (ANN) based model for estimation for yield of paddy and wheat.

4.2.1 ANN model for paddy

Initial an ANN base model has been developed with LevenbergMarquardt training algorithm. 7 Number of neurons and single hidden layer. Two models with 4 and 24 predictor variables have been developed.

4.2.1.1- Models with different predictor variable

An ANN based model (P-4) with 4 input variable has been developed. The input variables are those selected by stepwise MLR method that is T24,M25,R24,and R31. Fig 4.9 show the performance it can be observed that correlation coefficient between network output and target during training and validation are found to be 0.42618and 0.27351. The RMSE and MAE during training are estimated as 326.8153 and 251.8655respectively and during validation are found to be 432.9648 and 236.4837respectively.

Another model (model P-24) with 24 predictor variables as selected by enter method of MLR has been developed. These variables are T26,T28,T30,T32,T33,T34,T37,M24,M25,M27,M28,M29,M31,M34,M35,M36, M38,R25,R30,R31,R32,R37,R38,R39.Performance of the model P-24 indicated in Fig 4.10. The MSE during training and validation is $7.39\times10^{-32}$and 0.2367 respectively. Correlation coefficient (R) between network output and target during training and validation are found to be 0.9867 and 0.2205. The RMSE and MAE during training are estimated as 47.558 and 37.807 respectively and during validation are found to be 253.99 and 212.55 respectively.

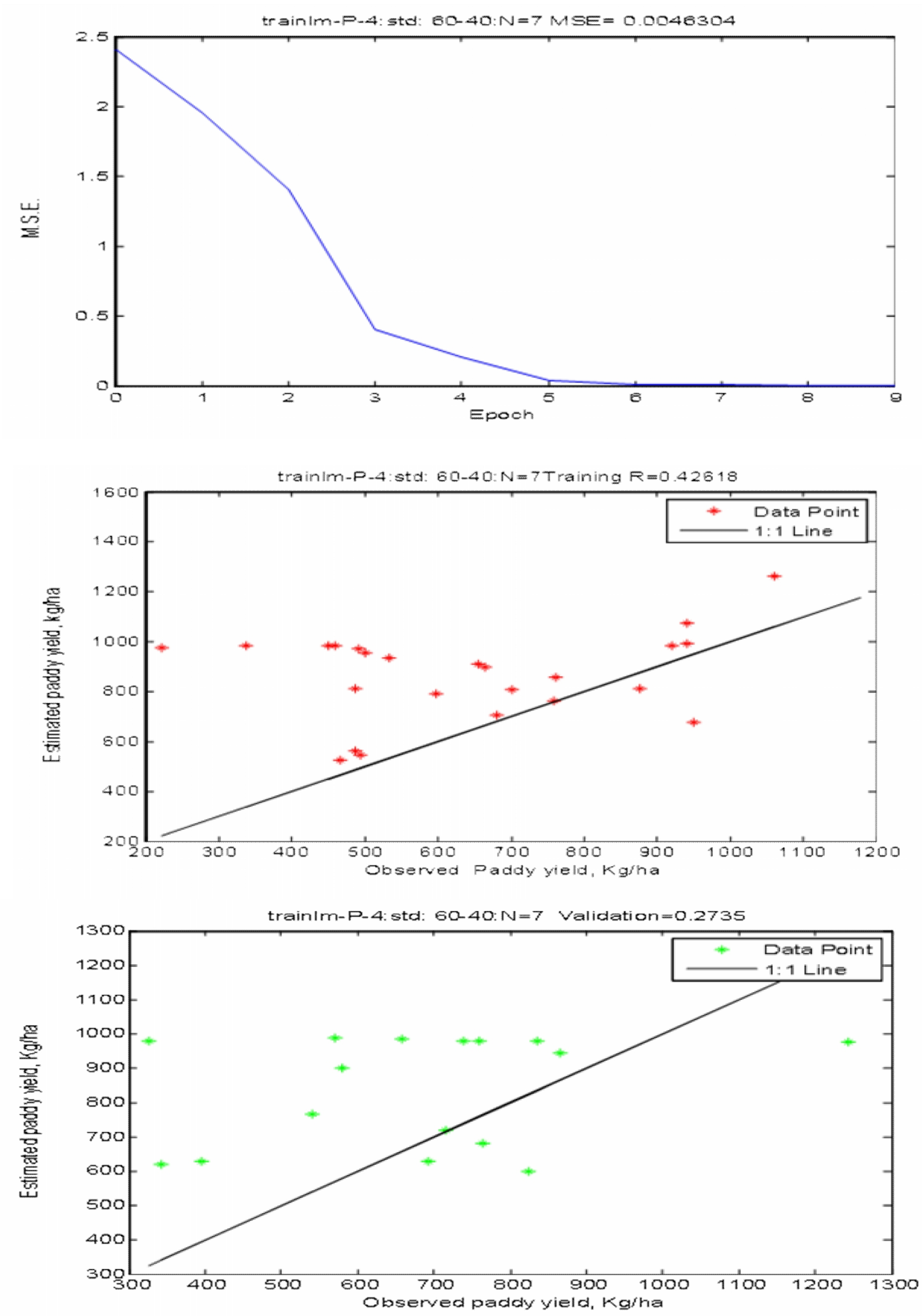


Fig. 4.9 Performance of ANN Model P-4 for paddy estimation.

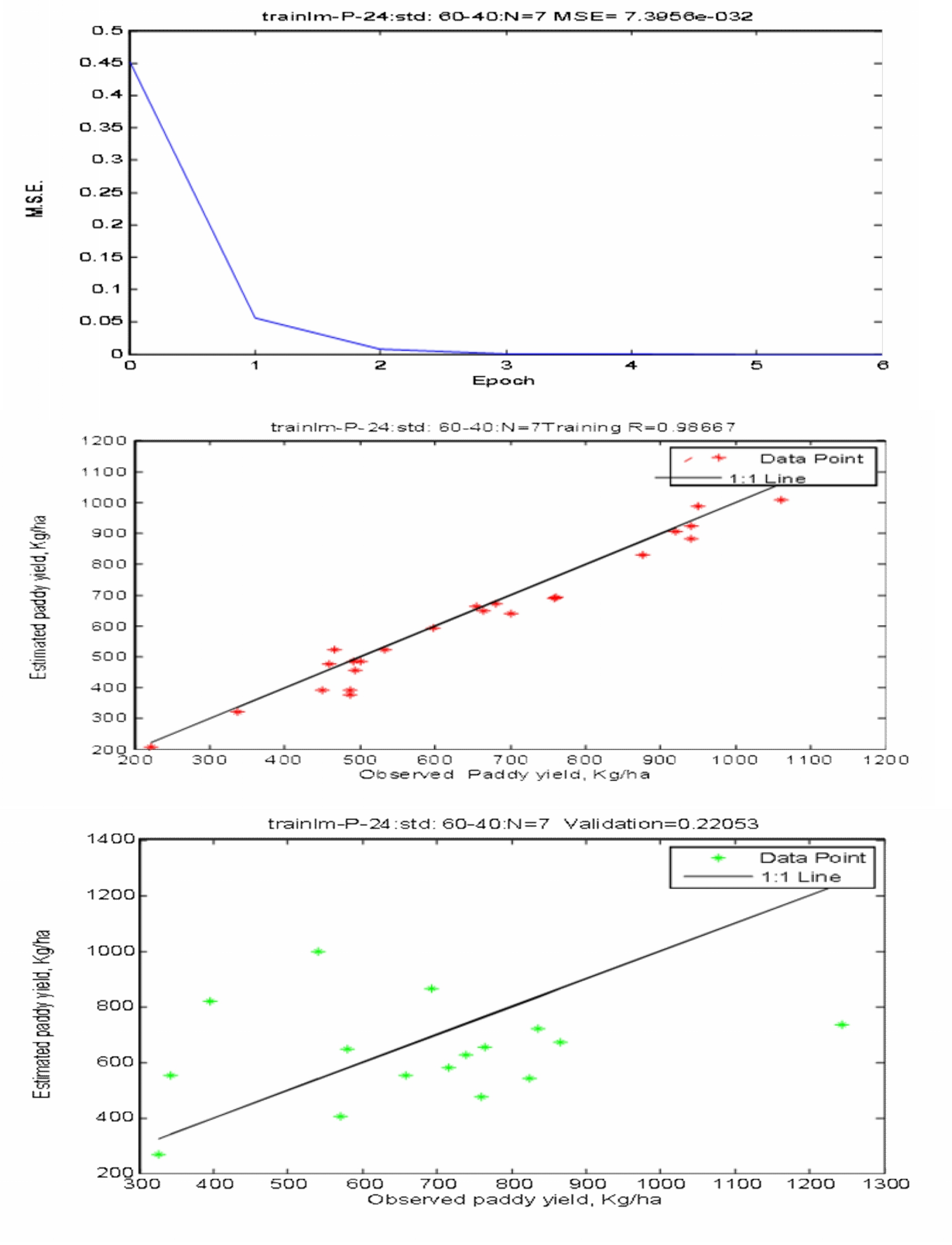


Fig. 4.10Performance of ANN Model P-24 for paddy estimation.

Table 4.3 Performance of various paddy ANN models (training with 60% of data).

S.No.	Model	R		MSE		RMSE		MAE	
		Trg	Val	Trg	Val	Trg	Val	Trg	Val
1	P-4	0.42618	0.2735	0.0046	0.6008	326.8153	432.9648	251.8655	236.4837
2	P-24	0.98667	0.2205	7.391×10^{-32}	0.2367	47.5582	253.9873	37.8073	212.5458

(trg = training, val= validation)

Performance of these two model has been summarized in Table 4.3. It can be observed that value of R during training of the model P-24 is much higher than model P-4 (0.426). However value R during validation of the model P-24 (0.2205)is slightly lower than model P-4 (0.2773) .MSE, RMSE AND MAE for the model P-24 are lower than model P-4, hence model P-24 has been selected for farther refinement.

4.2.1.2 Training algorithms for paddy yield model

The P-24 model as discussed in section 4.2.1.1has been Testedwith the different training algorithm. For developing the ANN based paddy yield model, performance of 12 training algorithms were evaluated. The model P-24 was developed using Levenverg Marquardt Algorithm (trainlm).The best training algorithm in the hidden layer of ANN model can be determined by trial and error, at which a model perform better.

4.2.1.2.1trainlm

"trainlm" is the syntax used for Levenberrg training algorithm. The value of the MSE obtained during the training of model is found as $7.39\times10^{-32}$the performance of model is depicted in Fig. 4.14.Parameters of the model performance for trainlm has been discussed in section 4.2.1.1.

4.2.1.2.2 traingd

Batch gradient decent (traingd) is also used for crop yield modelling.The training algorithm traingd does not perform well, as MSE during training is found to be 0.0536 (Table 4.4). Correlation coefficient between estimated paddy yield and observed paddy yield is 0.8681 during training.It has very poor during validation as indicated by R= -0.1466, RMSE= 377.71 and MAE=269.23.

4.2.1.2.3traingdm

The traingd is syntax used for training algorithm Batch Gradient Descent with Momentum (traingdm). Performance of this training algorithm is presented inFig 4.12. The algorithm completed training 9 epochs. That algorithm resulted well, performance as indicates Coefficient of correlation between estimated paddy yield and observed paddy yield is 0.52588and 0.4314during training and validation, respectively. RMSE has been worked out as 264.1855 and 301.9908; and MAE as 213.4292and 167.2342during training and validation respectively.this algorithm performed better than traingd; but poorer than trainlm.

4.2.1.2.4traingda

The method of ANN back propagation training with adaptive learning rate is implemented as traingd.It can be observed that MSE = 0.2358 during the training of model correlation coefficient between estimated and targeted paddy yield -0.14215; and during validation it is -0.1265.RMSE and MAE during training are found to be 289.38 and 217.35 respectively where during validation of the model these are 362.41 and 240.53 respectively (Table 4.4). This algorithm also does not show any improvement over algorithm trainlm.

4.2.1.2.5traincgf

The “trincgf” syntax is used for Conjugate gradient-fletcher-reeves update algorithm. The model gats trainedtraincgf algorithm with value of MSE as 0.0041 and 0.2769 during validation. The model resulted into correlation coefficient between estimated paddy yield and observed paddy yield as 0.95581 and 0.07365, for training and validation, respectively (Table 4.4). The model performance is very poor during validation and not suitable for paddy yield modelling.

4.2.1.2.6traincgb

Another version of the conjugate gradient algorithm was proposed by Polak and Ribierrend its syntax is “traincgp”. For the polke –ribiere update, the constant k is computed which is the inner product of the previous change in the gradient with the current gradient divided by the norm squared of the previous gradient .The correlation coefficient between estimated paddy yield and observed paddy yield as .46037and -0.2045, for traingn and validation, respectively (Table 4.4).This algorithm also does not show any improvement over algorithm trainlm.

4.2.1.2.7traincgp

In Powell-Beale Restarts(traincgb) training algorithms Fig 4.11. This algorithm resulted moderate performance as indicated MSE of scaled output and target is 0.0312and 0.1675during training and validation, respectively. Coefficient of correlation between estimated paddy yield and observed paddy yield is 0.42561 and 0.39841during training and validation, respectively. RMSE has been worked out as 215.5091 and 246.7089; and MAE as 181.8986and 153.2342during training and validation respectively.

4.2.1.2.8 tringdx

Ann training algorithm –back propagation variable learning rate is know as “traigdx”. It can be observed that MSE 0.1262 during training and 0.3144 during validation. The model results into coefficient of correlation between estimated and observed paddy yield during training as 0.59211 and during validation it is -0.2045.RMSE of estimated and observed paddy yield is 190.74 and 331.028 during training and validation, respectively (Table4.4).

4.2.1.2.9trainrp

The “trainrp” is an algorithm used for Resilient Back Propagation network. Fig.4.13 presents the performance of trainrp algorithm for training neural network model for paddy yield estimation. MSE during training of the model is found to be 0.0597 at 17 epochs. Coefficient of correlation between estimated and observed paddy yield during training as 0.7447 and during validation it is -0.2689.RMSE of estimated and observed paddy yield is 165.58 and 134.98 during training and validation respectively;where was MAE is 134.98 and 274.26 respectively. The algorithm performed well during training as well as validation of the model.

4.2.1.2.10 trainbfg

Newton's method is an alternative to the conjugate gradient method for fast optimization. Newton's method often converges faster than conjugate gradient method. Unfortunately, it is complex and expensive to compute the hessian matrix for feed forward neural networks. The alternate is quasi newton (or secant) method and algorithm is given by syntax "trainbfg". The resulted into MSE=0.316 and performed well as R during training and validation is 0.5476 and 0.2586, respectively. Value of RMSE and MAE for model is 201.13and 166.80during training and 295.71and 203.35 during validation, respectively (Table 4.4). The training algorithm performed well during training.

4.2.1.2.11 trainoss

The method of traininf of ANN paddy model is one secant and algorithm as known as "trainoss". MSE during training of model is found as 0.195 during training and validation of the model R is 0.22437 and -0.1860, respectively (Table 4.4). The algorithm trainoss does performed for training ANN model for paddy yield estimation.

4.2.1.2.12 trainscg

The scaled conjugate gradient algorithm (SCG), has syntax "trainscg". It require less time for training the model. The model performed well during the training (R=0.93511) but its performance is poor during validation of the model (R=-0.0417) as listed in Table 4.4.

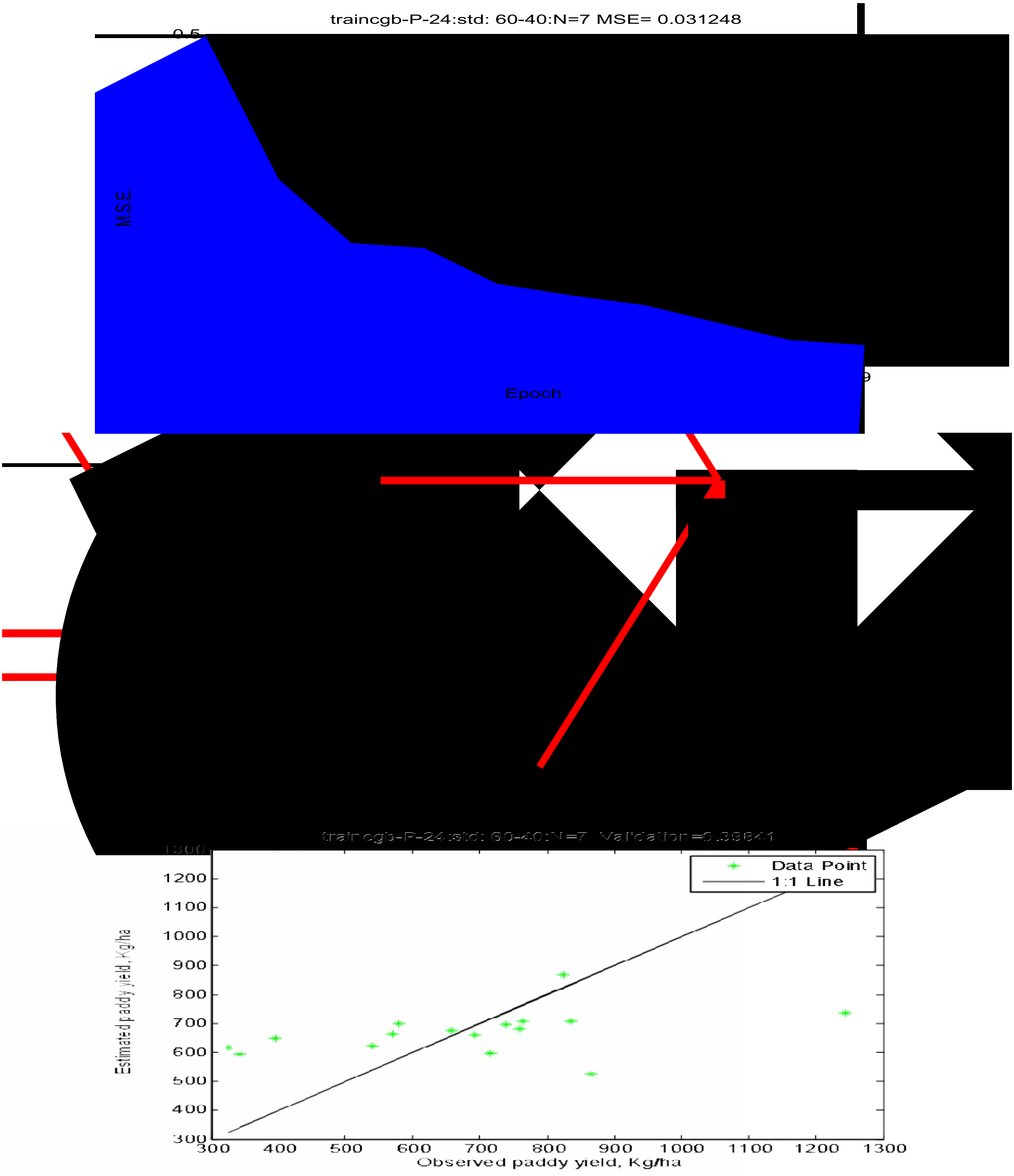


Fig. 4.11 Performance of traincgbtraining algorithm for paddy yield modelling.

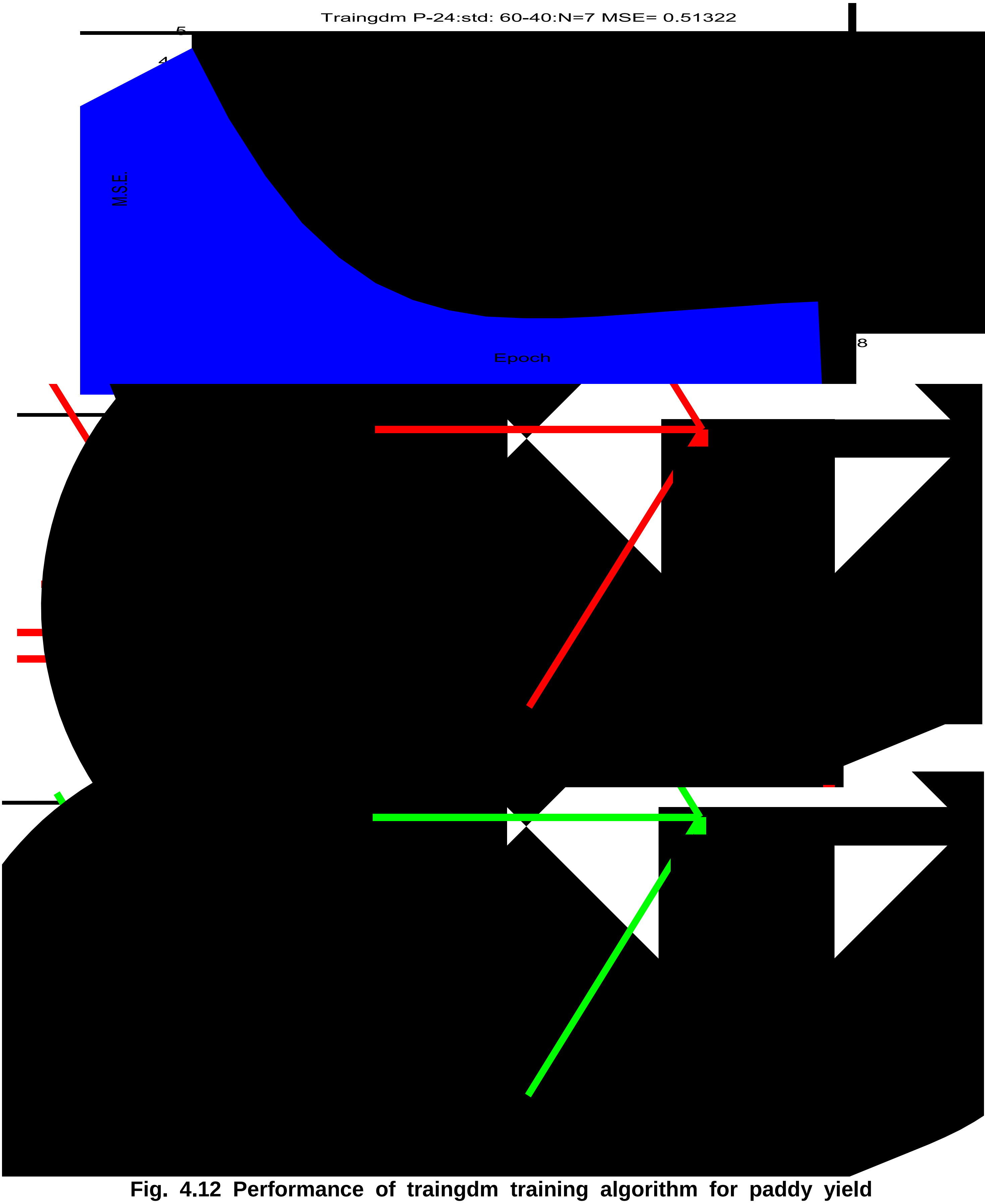
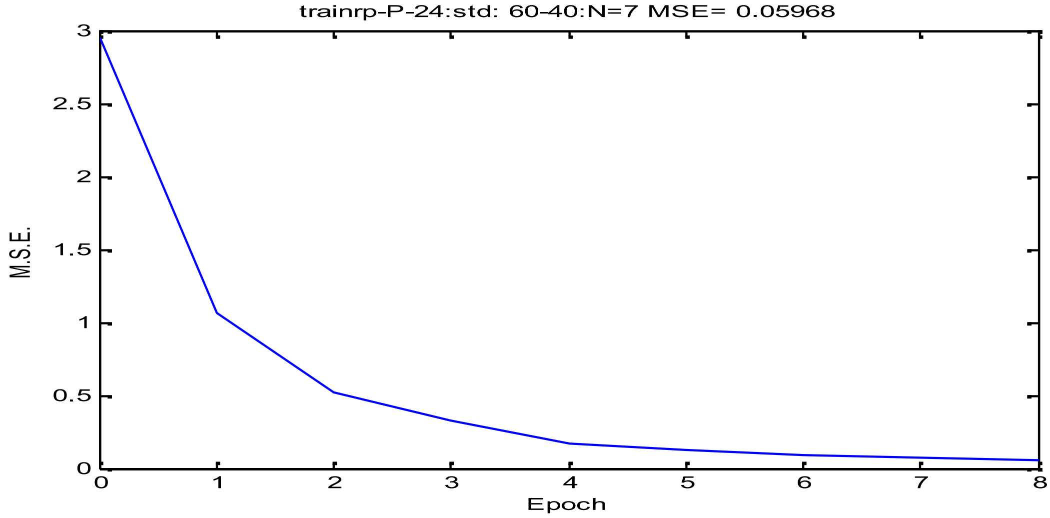
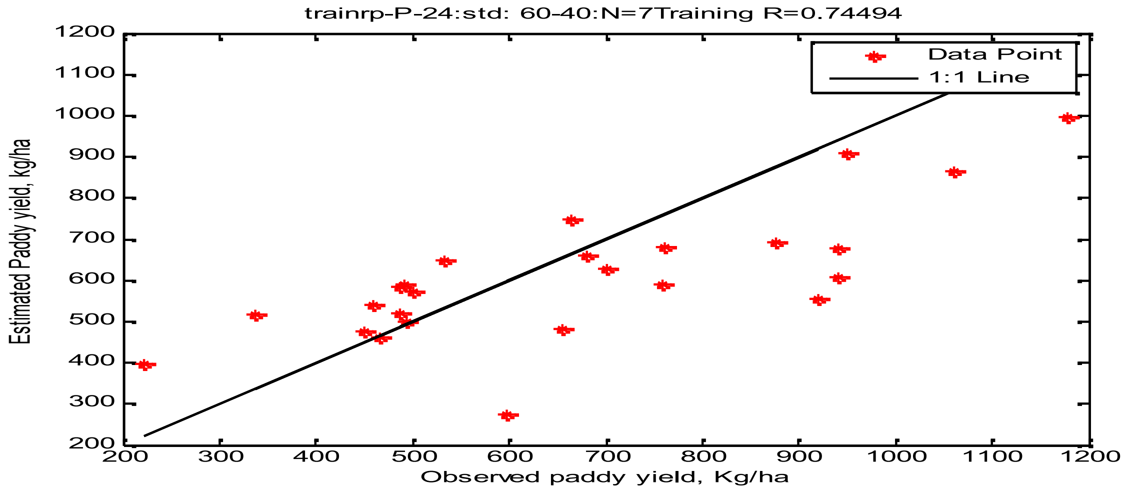


Fig. 4.12 Performance of traingdm training algorithm for paddy yield modelling.





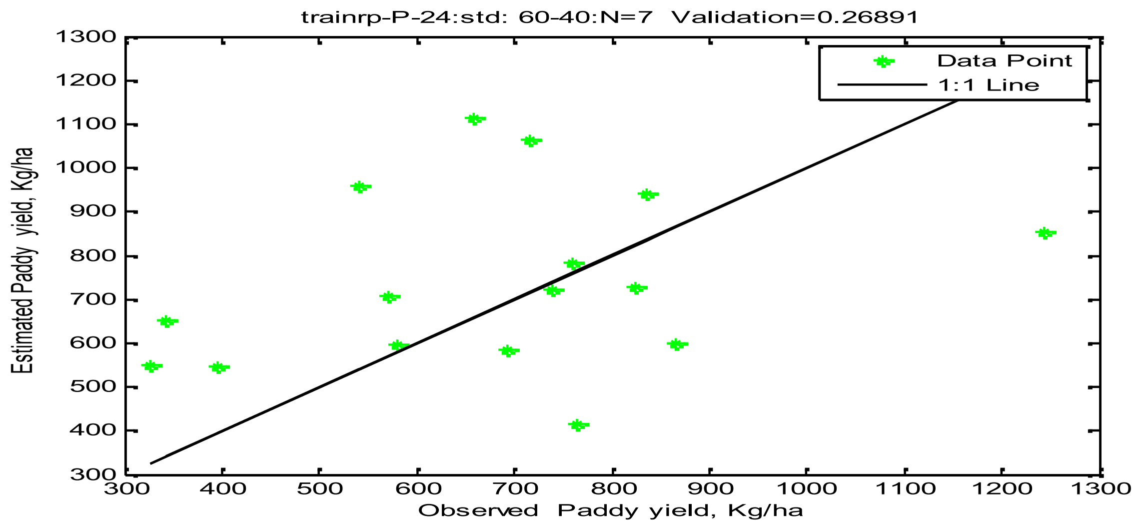


Fig. 4.13Performance of trainrp training algorithm for paddy yield modelling.

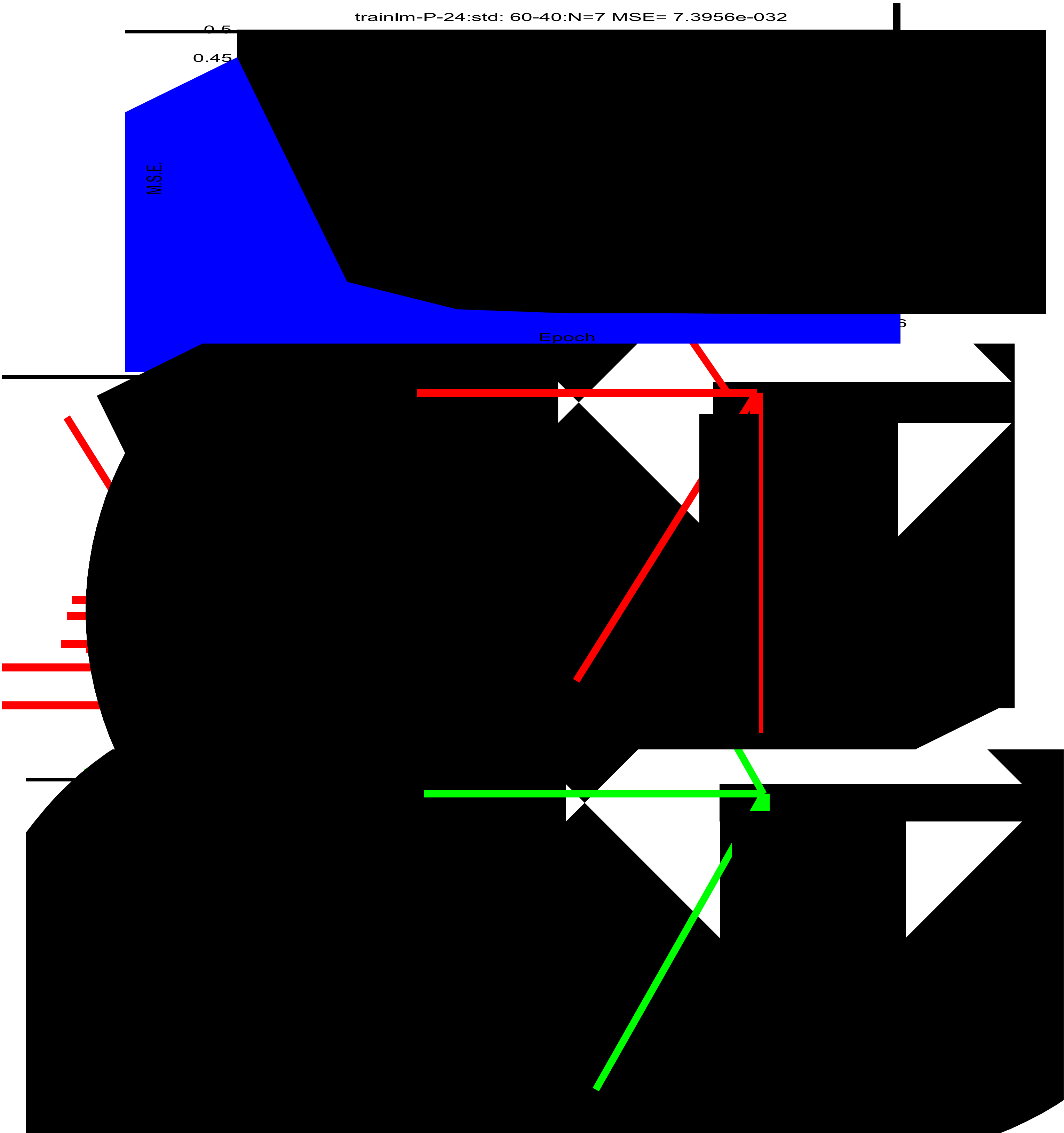


Fig. 4.14Performance of trainlm training algorithm for paddy yield modelling.

Table:-4.4 Performance of different training algorithm methods for ANN based paddy yield modelling.

Algorithm	R		MSE		RMSE		MAE	
	trg	val	trg	val	trg	val	trg	val
Trainlm	0.98667	0.2205	7.39×10^{-32}	0.2367	47.55	253.98	37.80	212.54
Traingd	0.86817	-0.1466	0.0536	0.4927	119.64	377.71	91.08	269.23
Traingdm	0.52588	0.4314	0.5132	0.3098	264.18	301.99	213.42	167.23
Traingda	-0.14215	-0.1265	0.2358	0.3107	289.38	362.41	217.35	240.53
Traingdx	0.5921	-0.1528	0.1262	0.3144	190.74	331.02	159.13	239.78
Traincgf	0.9558	0.0736	0.0041	0.2769	69.51	272.10	48.08	223.23
Traincgp	0.4603	-0.2045	0.0288	0.3077	210.08	328.92	167.80	233.14
Traincgb	0.4256	0.3984	0.0312	0.1675	215.50	246.7089	181.8986	153.2342
Trainscg	0.9351	-0.0407	0.0024	0.4096	83.60	308.3401	60.987	247.1489
Trainbfg	0.5476	0.2586	0.0316	0.2826	201.13	295.7151	166.8056	203.3561
Trainoss	0.2243	-0.1860	0.195	0.2582	229.72	285.24	187.4298	179.887
Trainrp	0.7449	0.2689	0.0597	1.8602	168.58	272.10	134.98	274.26

(trg = training, val= validation)

Table 4.4 indicates that training algorithm "trainlm" resulted model withhighest value of correlation coefficient as 0.98667 and 0.2205 during training and validation, respectively. Model performance indicators; MSE with scaled estimate and target is lowest as 7.39×10^{-32} and 0.2367 during training and validation, respectively. RMSE has been worked out as 47.5582 and

253.9873; and MAE as 37.8073and 4212.5458 during training and validation, respectively. Hence ANN wheat model trained with trainlm performed best for paddy yield estimation.

Hence the ANN paddy model with 24 input variables “mapstd” method of normalization, “trainlm” training algorithm at 7 neurons performed best amongst all twelve algorithms used for training.

4.2.1.3 Selection of optimum number of neurons in the hidden layer for the paddy yield

Increasing the number of neurons in the hidden layer, the network get an over fit, that is the net have problem to generalize. To determine the optimum number of neuron, at which network should have to perform its best, trial and error method is applied. Selection of optimum number of neurons is essential part wheat ANN model development. The model P-24 with learning function “trainlm” and normalization function “mapstd” trained with 60 percent of data has been evaluated for optimum number of neurons. Neurons in the hidden layer have been varied from 1 to 20.

Performance of ANN model with N=6 is shown in Fig.4.15 MSE of scaled output and target is 1.158×10^{-27} and 1.0645 during training validation, respectively.Coefficient of correlation between estimated paddy yield and observed paddy yield is 0.9400 and 0.3778 during training and validation, respectively.RMSE has been worked out as 100.588 and 228.947; and MAE has been worked out as 82.074and 2180.266 during training and validation respectively.

Performance of ANN model with N=8 is shown in Table 4.5. MSE of scaled output and target is $4.30\times10^{-32}$and 0.6365during training validation, respectively.Coefficient of correlation between estimated paddy yield and observed paddy yield is 0.8052and 0.0758during training and validation, respectively.RMSE has been worked out as 207.7666and 348.4246; and MAE has been worked out as 163.4328and 243.5206during training and validation respectively.

In model P-24 with“trainlm” algorithm, the model with number of neurons taken as N=9 is shown in Fig. 4.16. MSE of scaled output and target is $3.28\times10^{-24}$and 0.4848during training validation, respectively. Coefficient of correlation between estimated paddy yield and observed paddy yield is 0.9986and 0.0699during training and validation, respectively. RMSE has been worked out as 13.1703and 366.5343; and MAE as 10.3292 and 297.7527during training and validation respectively.

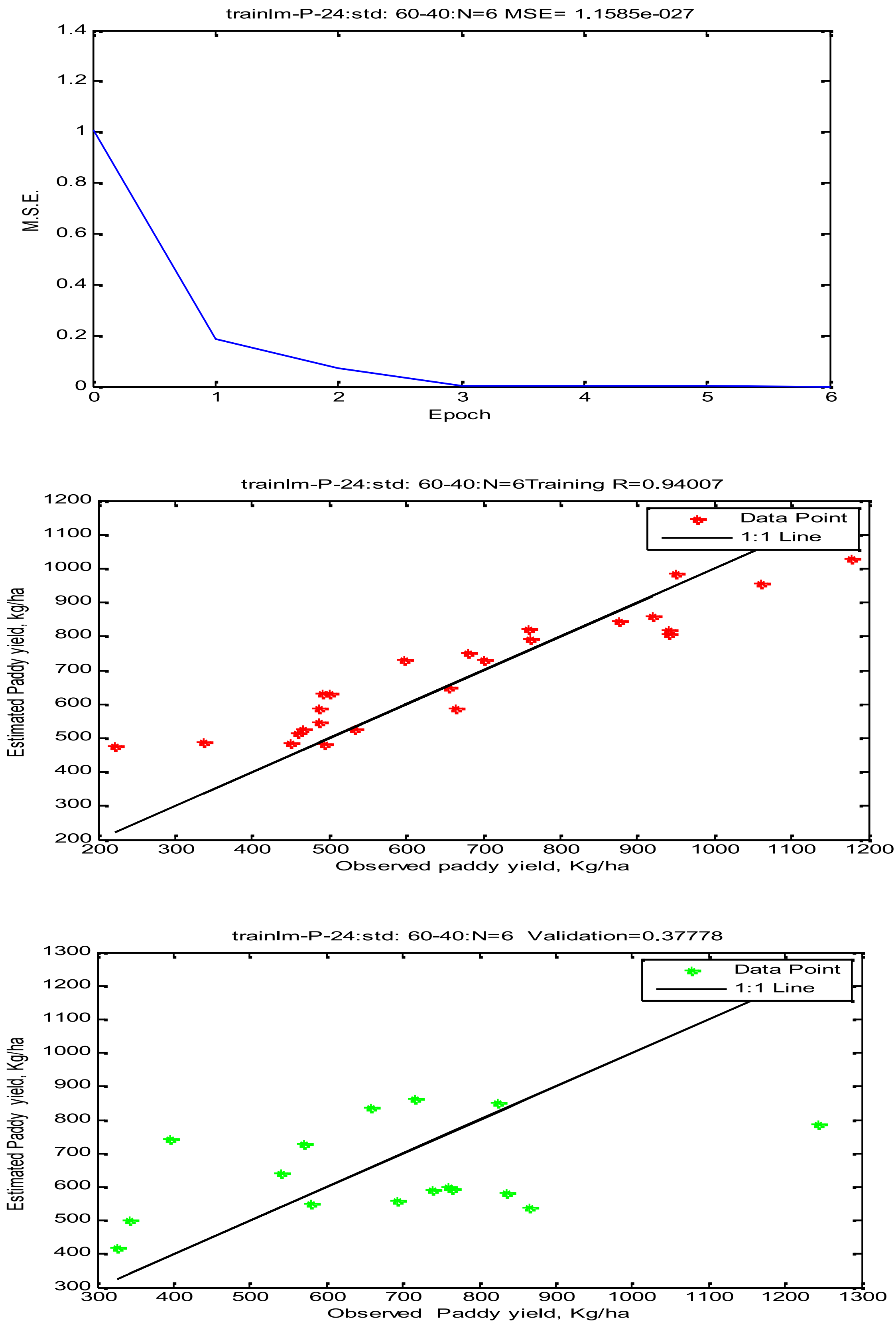


Fig. 4.15Performance of ANN Model for paddy yield estimation at 6 neurons.

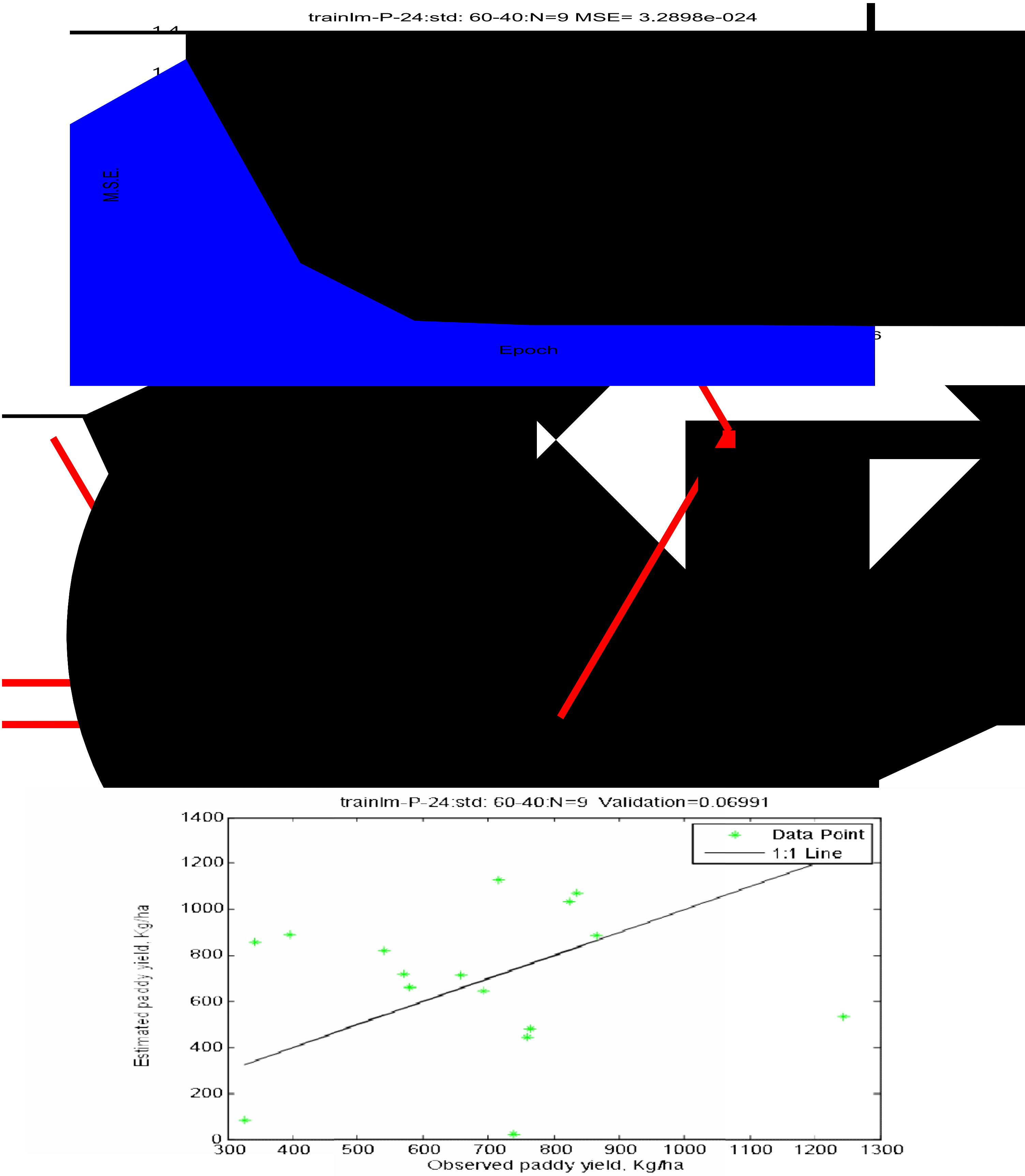


Fig. 4.16 Performance of ANN Model for paddy yield estimation at 9 neurons.

Table 4.5 Performance of neural network with different number of neurons for paddy yield estimation

S.N.	Model	R		MSE		RMSE		MAE	
		trg	val	trg	val	trg	val	trg	val
1	N5	0.7687	-0.1733	5.33×10^{-32}	0.5490	190.3380	310.1735	164.6667	196.6063
2	N6	0.9400	0.3778	1.15×10^{-27}	1.0645	100.5888	228.9471	82.0749	180.2660
3	N7	0.9866	0.2205	7.39×10^{-32}	0.2367	47.5582	253.9873	37.8073	212.5458
4	N8	0.8052	0.0758	4.30×10^{-32}	0.6365	207.7666	348.4246	163.4328	243.5206
5	N9	0.9986	0.0699	3.28×10^{-24}	0.4848	13.1703	366.5343	10.3292	297.7527
6	N10	0.9266	-0.1129	5.89×10^{-32}	0.4129	126.6834	308.9712	99.9917	244.5761
7	N13	0.7557	-0.3468	1.34×10^{-23}	0.4628	158.4285	341.6473	117.7207	263.6176
8	N15	0.9999	-0.2121	4.08×10^{-24}	0.5478	0.8393	389.7250	0.6443	344.3954
9	N20	0.5062	-0.1851	2.44×10^{-31}	0.9078	330.3143	539.6797	272.0260	381.1725

(trg = training, val= validation)

On comparison of performance parameters presented in Table 4.5, it can be stated that model P-24 trained with “trainlm” algorithm, “mapstd” normalization function and 6 neurons performed best. This network architecture has been further improved varying length of data for training and validation.

4.2.1.4 Effect of length of record on performance of ANN model for paddy yield estimation

ANN models developed and discussed in section 4.2.1.2 and 4.2.1.3, 60 percent of total length record has been used for model training and validation of the model. An attempt has made to evaluate the performance of ANN model with varying length of data set for model training and validation.

ANN model training with 50 data and tested with 50 percent data resultedcorrelation coefficient is 0.8426 and 0.0850 during training and validation respectively.It is found that MSE of estimated and target paddy (scaled) is $1.49998\times10^{-27}$and 1.8989 for training and validation, respectively (Fig.4.17). RMSE and MAE were calculated 157.138 and 108.604, respectively during training of model and, 267.945 and 204.678 during validation. The performance of model training with 60 percent has discussed in section 4.2.1.3 and presented in Fig. 4.15.

When 70 of the total data have been used for training of the model, it is found that MSE of estimated and target paddy yield (scaled) is $4.2190\times10^{-27}$and 1.7510for training and validation, respectively (Fig.4.18). RMSE and MAE were calculated 288.531 and 239.840, respectively during training of model and 374.817 and 160.254during validation. It can also be observed that correlation coefficient is -0.02244 and 0.4516 during training and validation respectively.

When 80 of the total data have been used for training of the model, it is found that MSE of estimated and target wheat (scaled) is$8.2041\times10^{-14}$and 2.1824for training and validation respectively (Fig.4.19). RMSE and MAE were calculated 138.21 and 112.27, respectively during training of model and, 394.52 and 258.41 during validation. It can also be observed from Table 4.6 that correlation coefficient is 0.7916 and -0.1508during training and validation respectively. Model trained with 80 percent of data does not perform well during validation.

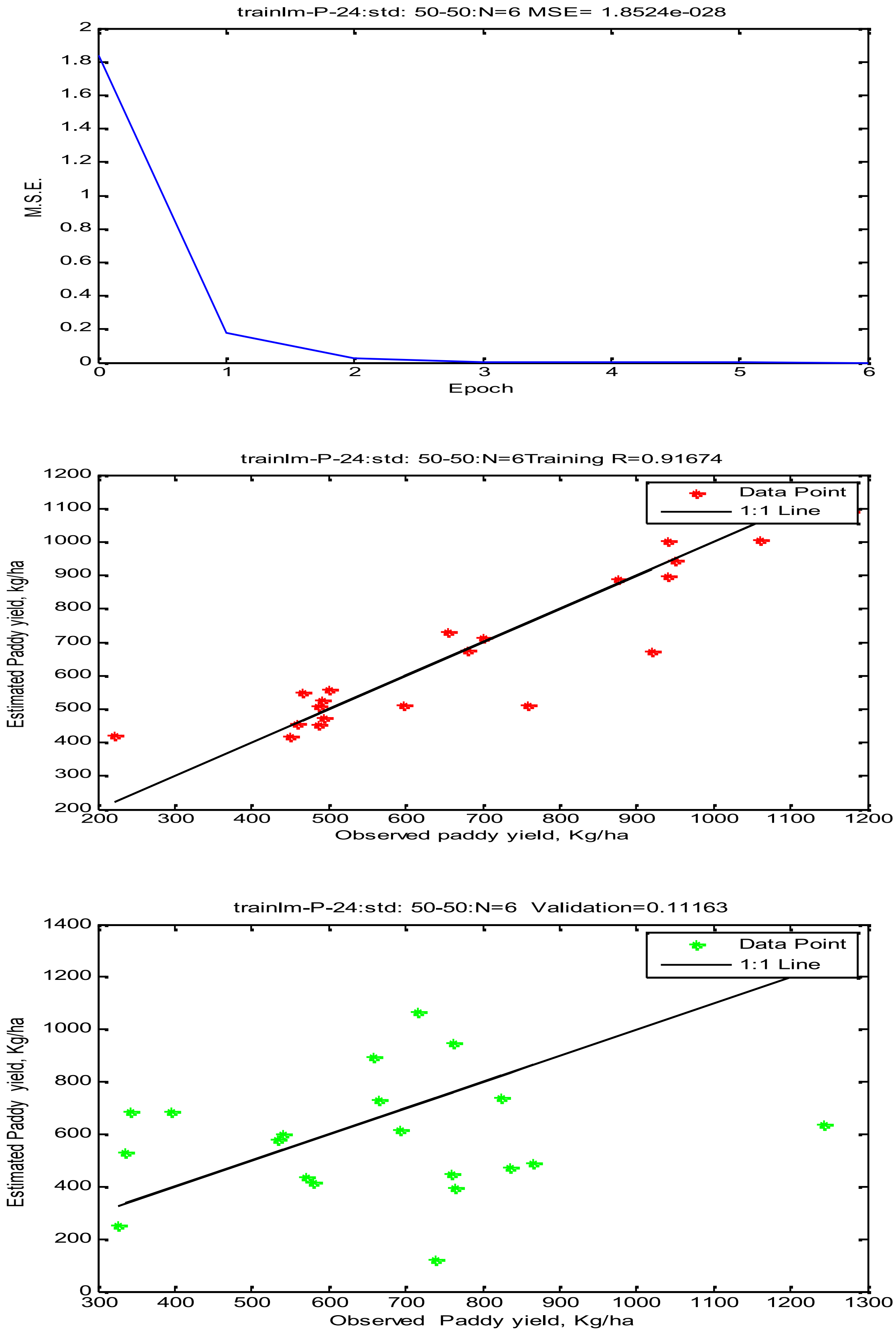


Fig. 4.17 Performance of ANN model for paddy yield estimation trained with 50percent data set.

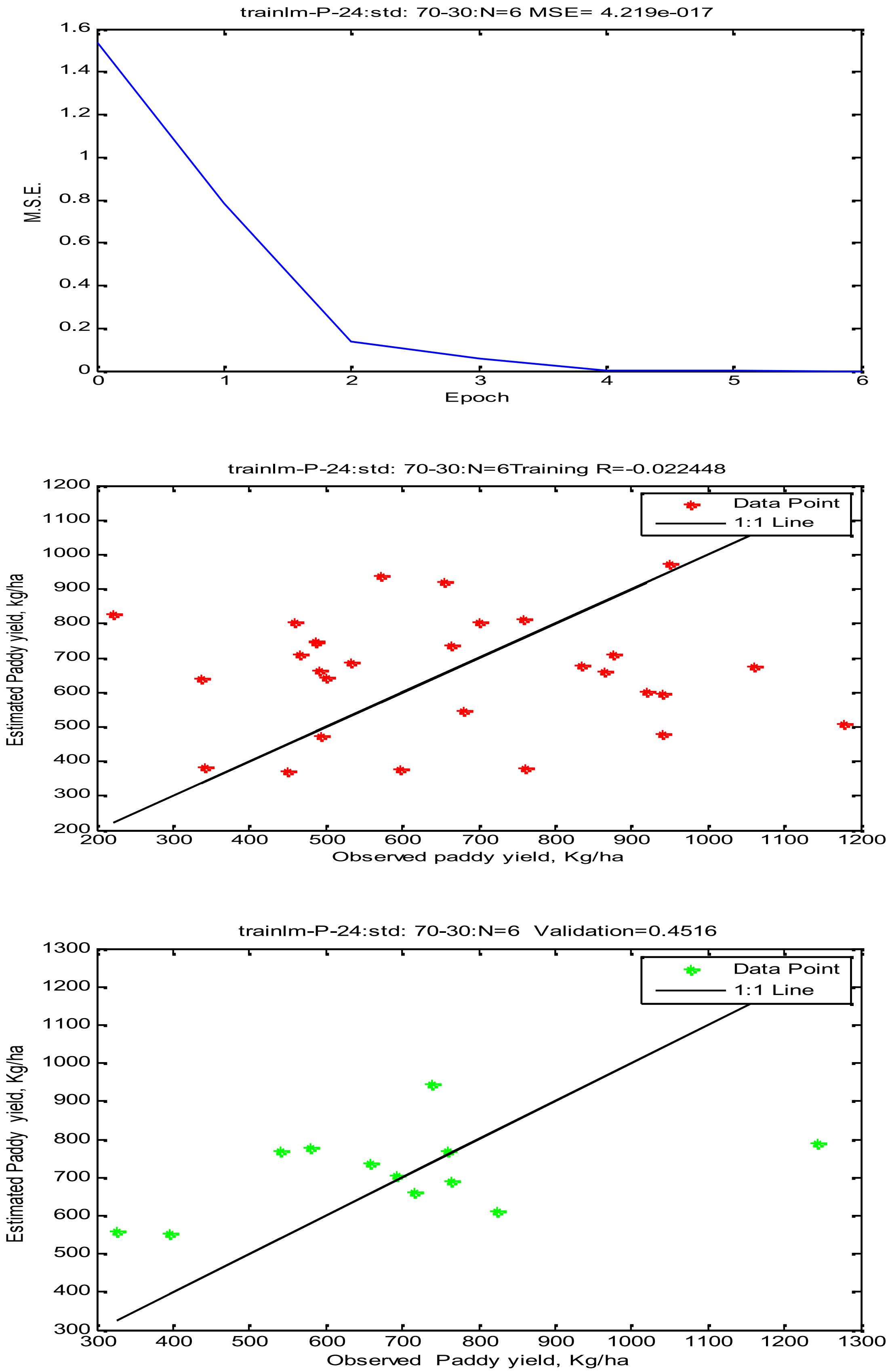


Fig. 4.18Performance of ANN model for paddy yield estimation trained with 70 percent data set.

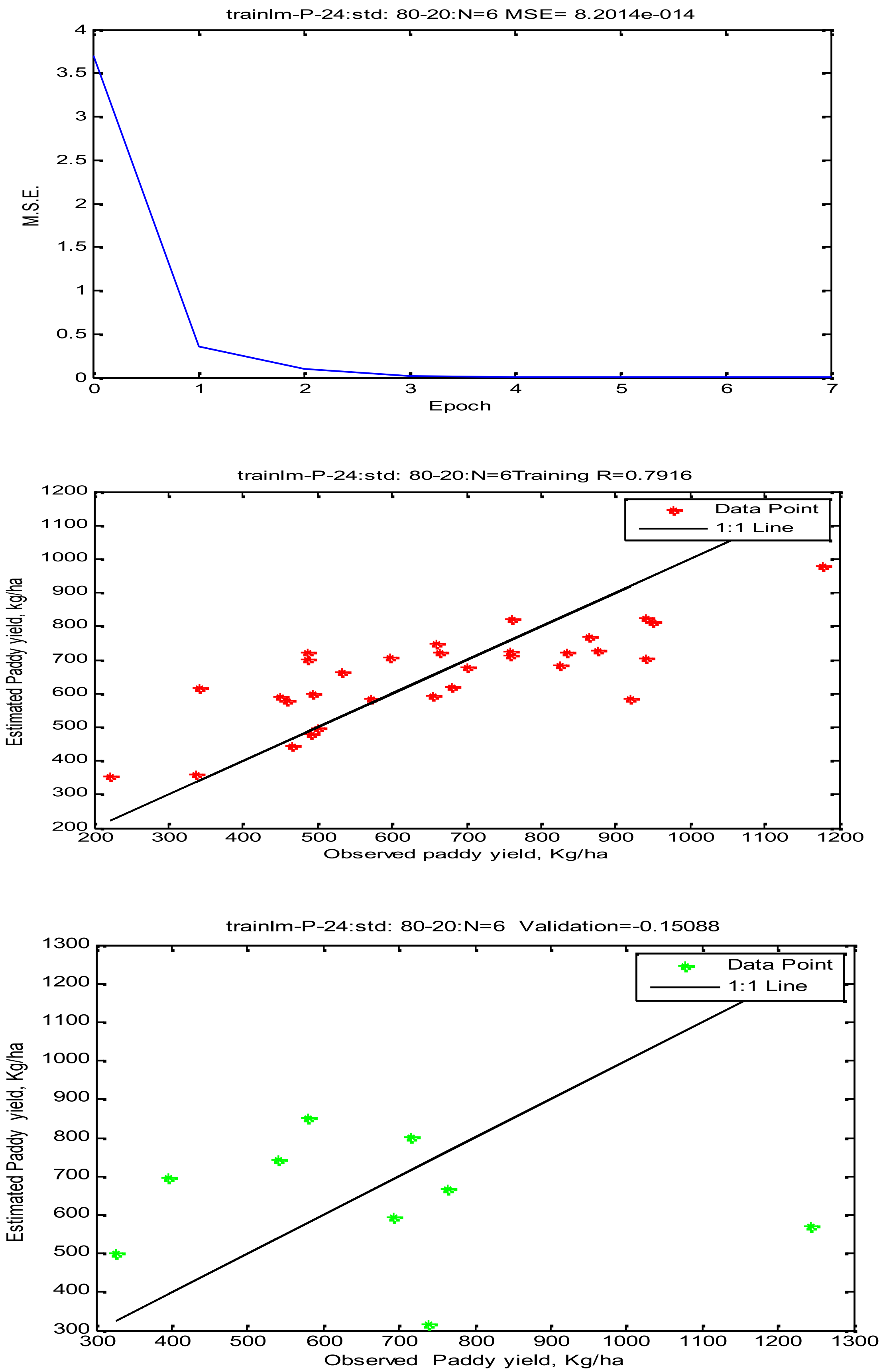


Fig. 4.19Performance of ANN model for paddy yield estimation trained with 80 percent data set.

Table 4.6 Performance of Neural Network with varying length of data set for paddy yield ANN modelling

S. No.	Model	R		MSE		RMSE		MAE	
		Trg	Val	Trg	Val	Trg	Val	Trg	Val
1	L-50	0.8426	0.0850	1.49×10-27	1.898	157.13	267.94	108.60	204.67
2	L-60	0.9400	0.3778	1.15×10-29	1.064	100.58	228.94	82.07	180.26
3	L-70	-0.0224	0.4551	4.21×10-27	1.751	288.53	374.81	239.84	160.25
4	L-80	0.7916	-0.1508	8.20×10-14	2.182	138.21	394.52	112.27	258.41

(trg = training, val= validation)

On the basis of model performance parameters evaluated for model trained with different length of data set (Table 4.6), it can stated that when model trained with 60 percent of data set, model efficiency is highest amongst the varies length of record tested.

4.2.2 ANN model for wheat

Initial an ANN base model has been developed with Levenverg-Marquardt training algorithm. 7 Number of neurons and single hidden layer. Two models with 24 and 8 predictor variables have been developed.

4.2.2.1 Models with different predictor variable

An ANN based model (W-8) with 8 input variable has been developed. The input variables are those selected by stepwise MLR method that is T11,T48,M44, M45,M47,M49,R8,R50Fig 4.20 show the MSE during training 3.60×10^{-5}at 6 epoch. Correlation coefficient (R) between network output and target during training and validation are found to be 0.4279and -0.1064. The RMSE and MAE during training are estimated as 540.9318and 438.4340respectively and during validation are found 676.5467and 459.7540 respectively.

Another model (model W-24) with 24 predictor variables as selected by enter method of MLR has been developed. These variables are T44,T46,T52,T1,T2,T3,M46,M49,M7,R44,R45,R46,R47,R48,R50,R3,R5,R6,R7,R8,R10,R11,R12,R13.Performance of the model W-24 indicated in Fig. 4.21.The MSE during training and validation is 440.2543and 415.4397respectively. Correlation coefficient (R) between network output and target during training and validation are found to be 0.2001and 0.3508. The RMSE and MAE during training are estimated as 578.7521and 440.2543respectively and during validation are found to be 651.0224and 415.4397respectively.

Table 4.7 Performance of wheat ANN models (training with 60% of data).

S.N	Model	R		MSE		RMSE		MAE	
		Trg	Val	Trg	Val	Trg	Val	Trg	Val
1	W-8	0.4279	-0.1064	3.60×10^{-5}	2.3066	540.932	676.547	438.434	459.754
2	W-24	0.2001	0.3508	3.53×10^{-15}	2.2306	578.752	651.022	440.254	415.44

(trg = training, val= validation)

Performance of these two models has been summarized in Table 4.3. It can be observed that value of R during validation of the model W-24 is much higher than model W-4 (-0.1064). However value R during training of the model P-24 (0.2001) is slightly lower than model P-4 (0.4279). MSE, RMSE AND MAE for the model P-24 are lower than model P-4, hence model P-24 has been selected for farther refinement.

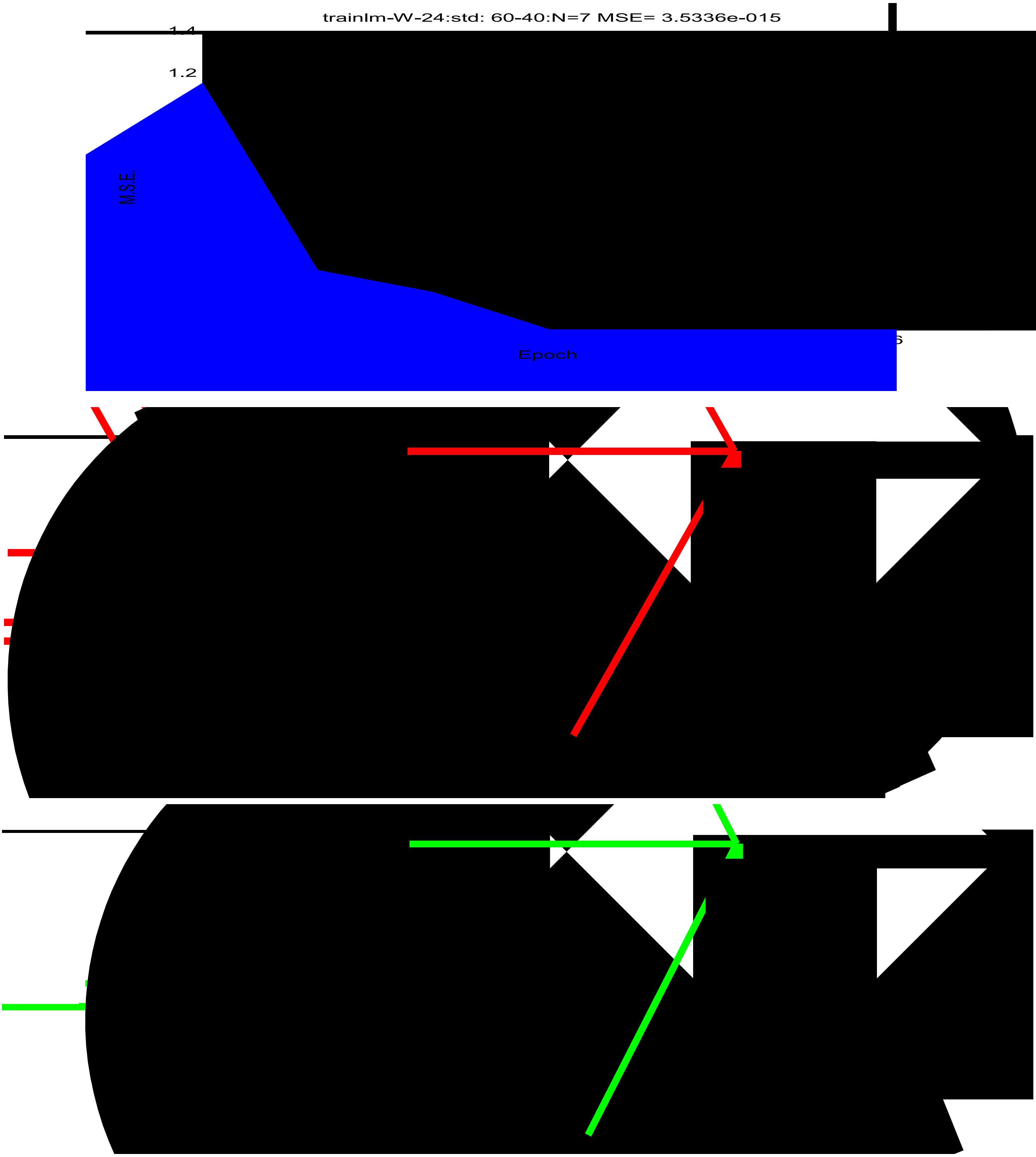


Fig. 4.20Performance of ANN Model W- 24 for wheat estimation.

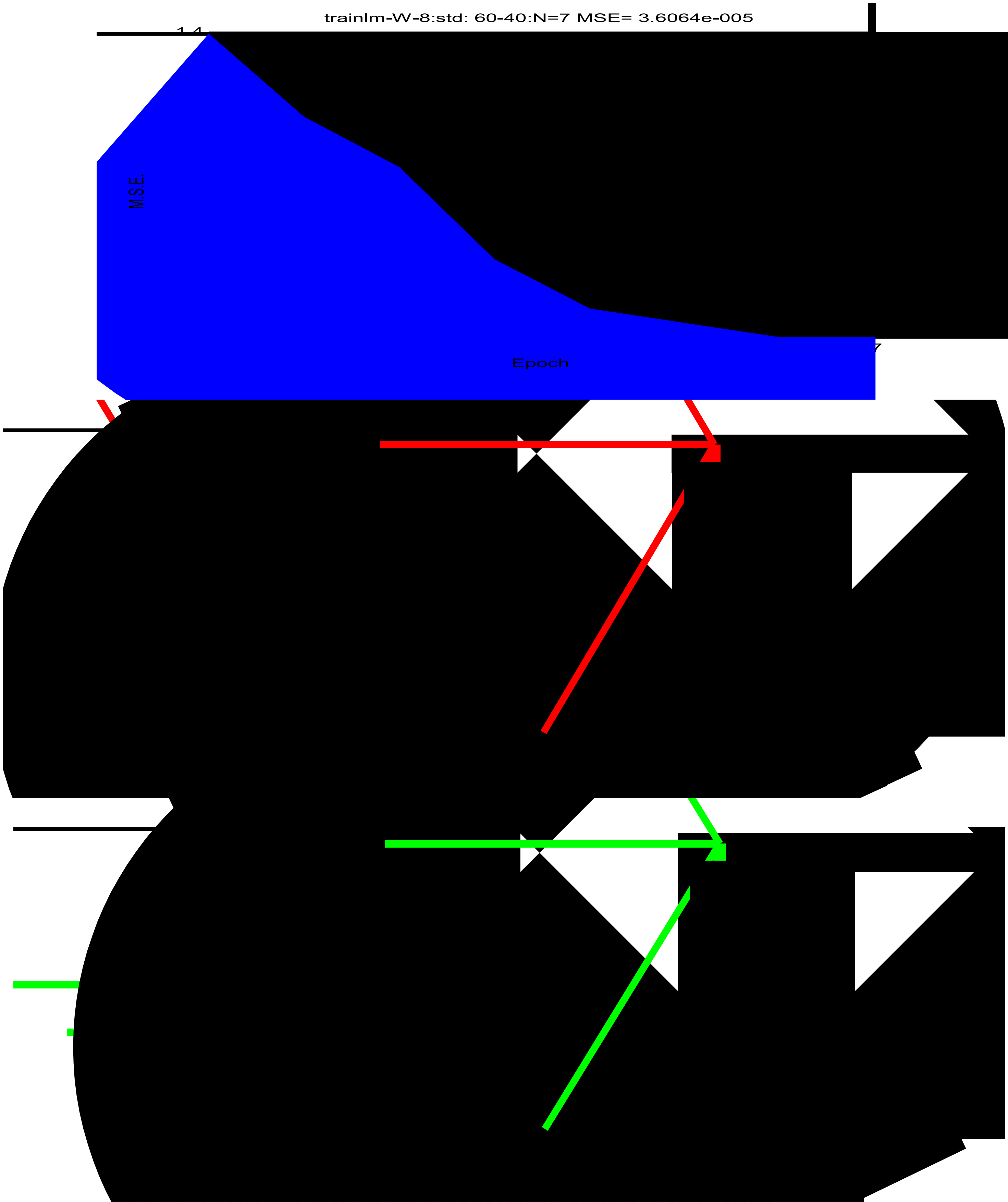


Fig. 4.21Performance of ANN Model W- 8 for wheat estimation.

4.2.2.2 Training Algorithms for wheat yield model

The W-24 model as discussed in section 4.2.2.1has been tested with the different training algorithm. For developing the ANN based wheat yield model, performance of 12 training algorithms were evaluated. The model W-24 was developed using Levenverg Marquardt Algorithm (trainlm).The best training algorithm in the hidden layer of ANN model can be determined by trial and error, at which a model perform better. All algorithm have been discussed in case of paddy, hence 3 algorithm which are well performing in case of wheat are been discussed here and rest are being tabulated (Table 4.8).

In “traingd” training algorithms (Fig. 4.22indicates) the MSE of scaled output and target is 0.6689and 0.8195during training and validation, respectively. Coefficient of correlation between estimated wheat yield and observed wheat yield is 0.4718 and 0.2719during training and validation, respectively. RMSE has been worked out as 481.5343 and 616.8720; and MAE as 399.4440and 380.7634during training and validation respectively.

In “trainrp” training algorithms (Fig. 4.23) the MSE of scaled output and target is 0.0168and 1.6542during training and validation, respectively. Coefficient of correlation between estimated wheat yield and observed wheat yield is 0.8599and 0.1193during training and validation, respectively. RMSE has been worked out as 316.1764 and 744.1776; and MAE as 269.3000and 528.7487during training and validation respectively.

In “trainscg” training algorithms (Fig. 4.24) the MSE of scaled output and target is 0.0066 and 1.8321during training and validation, respectively. Coefficient of correlation between estimated wheat yield and observed wheat yield is 0.7484and 0.3877during training and validation, respectively. RMSE has been worked out as 365.2947 and 551.2655; and MAE as 282.9732and 402.2662during training and validation respectively.

.

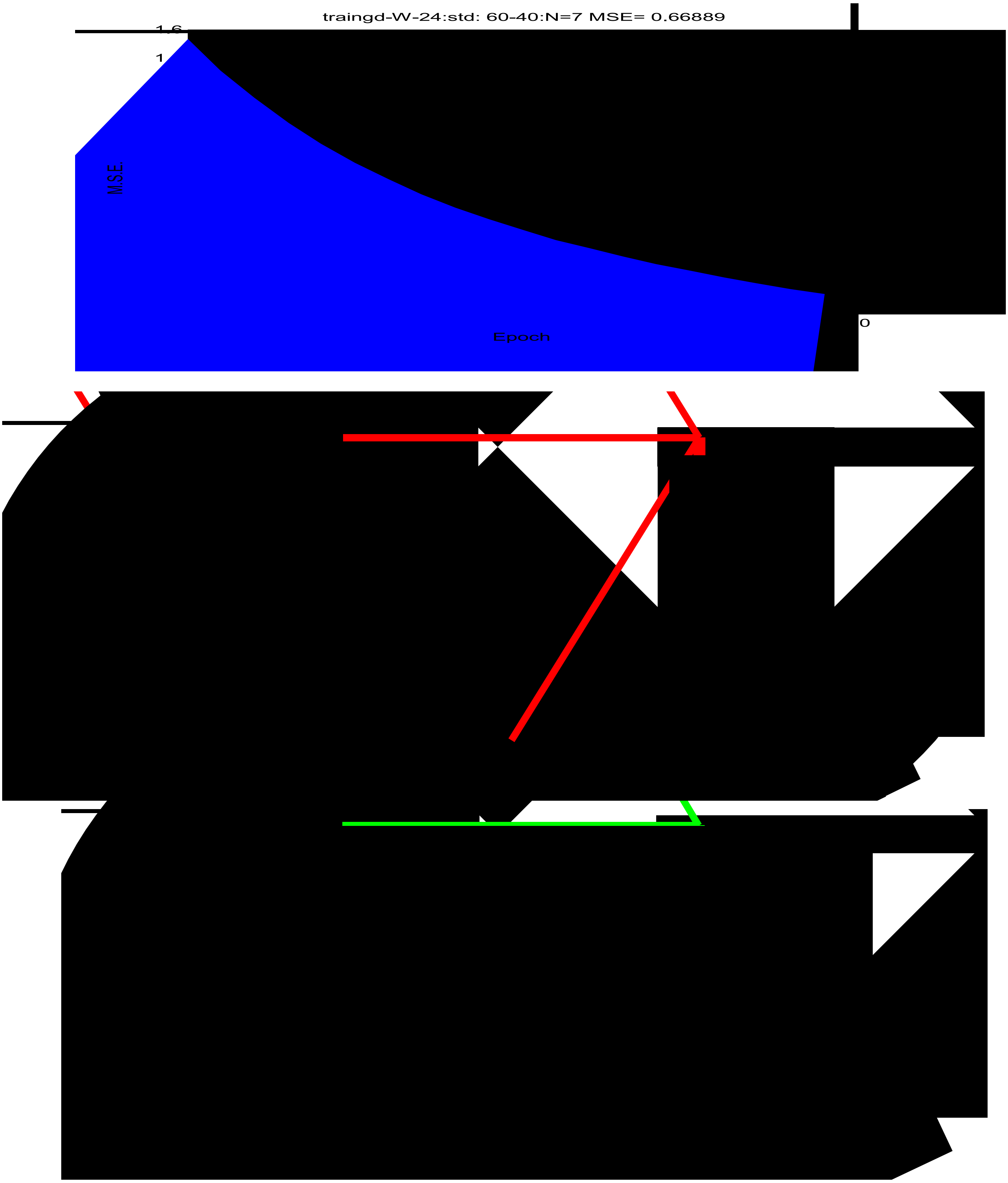


Fig.4.22 Performance of traingd training algorithm for wheat yield modelling.

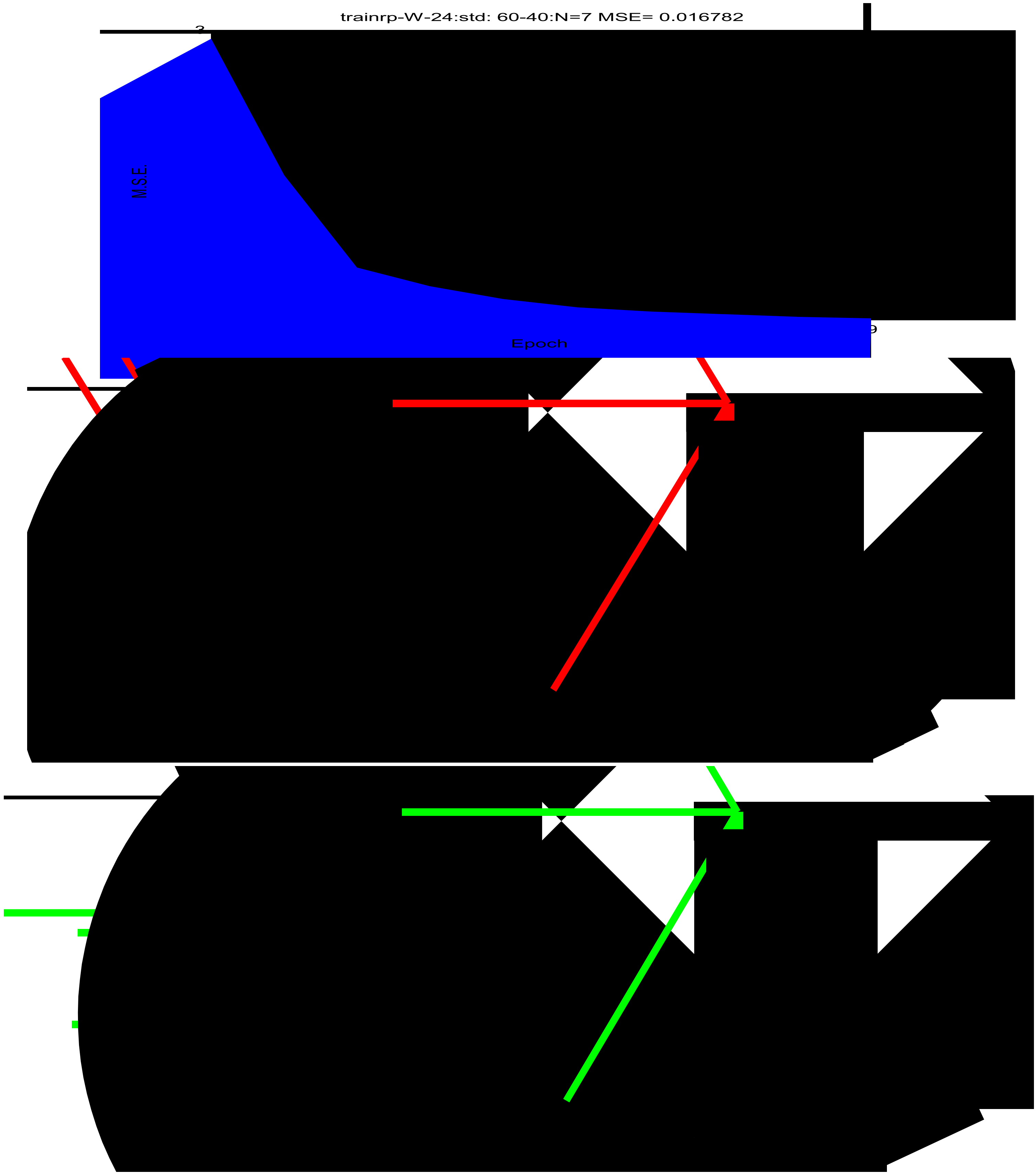


Fig. 4.23Performance of trainrp training algorithm for wheat yield modelling.

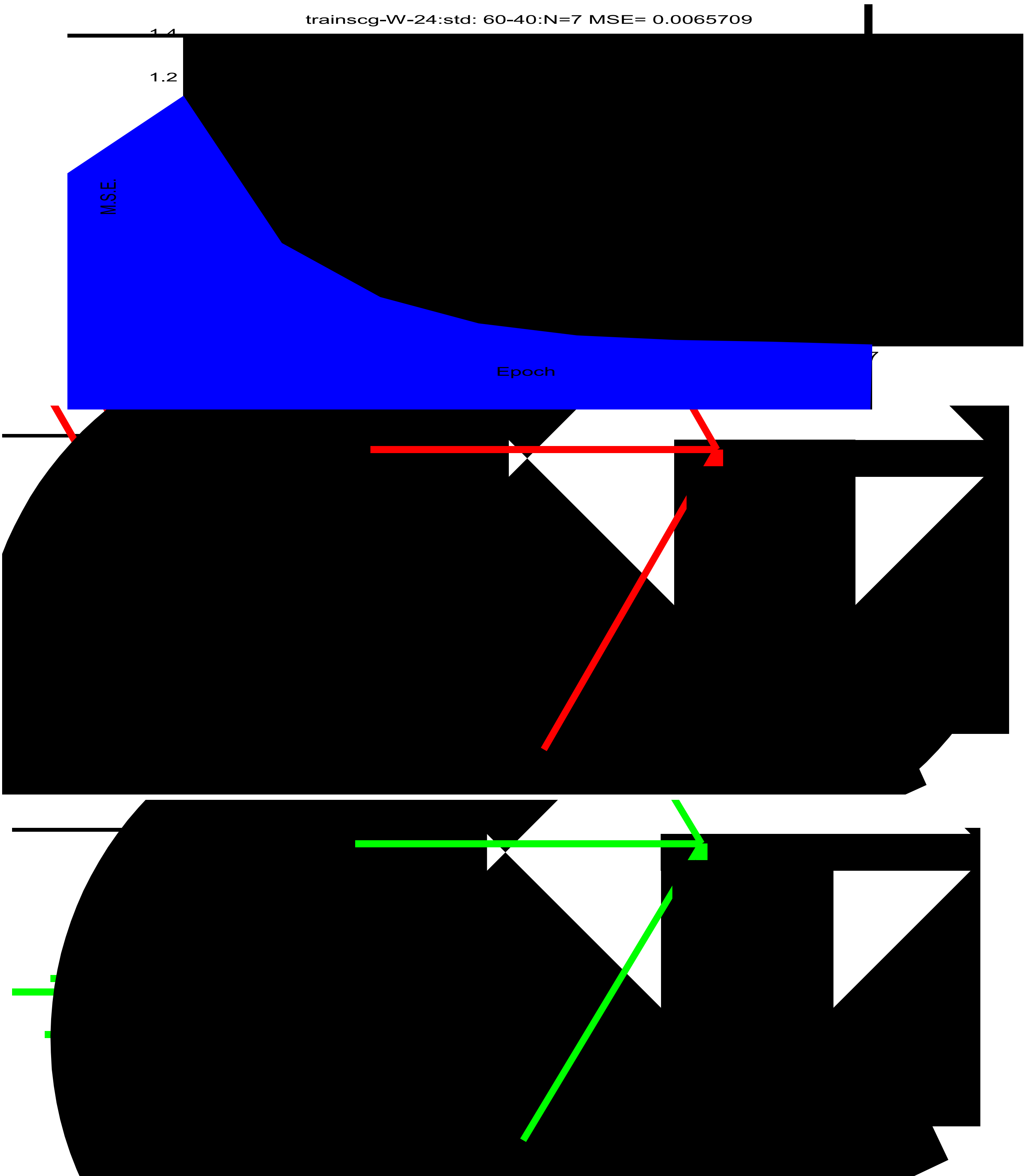


Fig. 4.24Performance of trainscg training algorithm for wheat yield modelling.

Table: 4.8Performance of different training algorithm methods for ANN based Wheat yield modelling

Algorithm	R		MSE		RMSE		MAE	
	Trg	Val	trg	val	trg	val	trg	val
Trainlm	0.2001	0.3508	3.5336×10^{-15}	2.2306	578.7521	651.0224	440.2543	415.4397
Traingd	0.4718	0.2719	0.6689	0.8195	481.5343	616.8720	399.4440	380.7634
traingdm	0.3291	0.1077	1.1286	0.9963	536.5373	679.9928	427.2391	420.0453
traingda	0.8487	-0.3090	0.2652	1.3848	333.6259	716.2079	280.5983	423.7550
traingdx	-0.0769	-0.06025	1.0294	0.9823	579.7674	674.4028	474.9673	395.0
traincgf	0.3428	-0.5223	0.0411	1.2554	511.3547	664.9457	428.3093	390.2071
traincgp	0.8270	-0.2042	0.0404	1.7493	308.4897	663.1753	240.5436	470.6702
traincgb	0.1963	0.0272	0.0135	1.0015	532.2125	654.8752	441.8909	419.5207
trainscg	0.7484	0.3877	0.0066	1.8321	365.2947	551.2655	282.9732	402.2662
trainbfg	0.3867	0.3350	0.0216	1.3740	598.9163	704.0446	478.5202	412.3882
trainoss	0.9096	-0.0936	0.0115	1.3154	325.3623	661.9669	276.3559	465.7902
Trainrp	0.8599	0.1193	0.0168	1.6542	316.1764	744.1776	269.3000	528.7487

(trg = training, val= validation)

Table 4.8 indicates that training algorithm “trainscg” resulted model withhighest value of correlation coefficient as 0.7484 and 0.3877 during training and validation,respectively. Model performance indicator; MSE with scaled estimate and target is lowest as 0.0066 and 1.8321 during training and validation, respectively. RMSE has been worked out as 365.2947 and 551.2655; and MAE as 282.9732and 402.2662during training and validation, respectively. Hence ANN wheat model trained with trainscg performed best for wheat yield estimation .the next model which performs comparatively better is traingdwith value R as 0.4718 and 0.2719 during training and, validation, respectively. The training algorithm traingd resulted intoRMSE has been worked out as 481.5343 and 616.8720; and MAE as 399.4440and 380.7634during training and validation respectively. The training algorithm “trainrp” also perfomed well R as 0.8599 and 0.1193 during training and validation.

Hence the ANN wheat model with 24 input variables “mapstd” method of normalization, “trainscg” training algorithm at 7 neurons performed best amongst all twelve algorithms used for training.

4.2.2.3 Selection of optimum number of neurons in the hidden layer for the wheat yield

Increasing the number of neurons in the hidden layer, the network get an over fit, that is the net have problem to generalize.To determine the optimum number of neuron, at which network should have to perform its best, trial and error method is applied. Selection of optimum number of neurons is essential part wheat ANN model development. The model W-24 with learning function “trainscg” and normalization function “mapstd” trained with 60 percent of data has been evaluated for optimum number of neurons. Neurons in the hidden layer have been varied from 1 to 20.

Performance of ANN model with N=6 is shown in Fig 4.25 indicates that MSE of scaled output and target is 0.0252and 1.7291during training validation, respectively.Coefficient of correlation between estimated wheat yield and observed wheat yield is 0.4326and 0.6096during training and validation, respectively.RMSE has been worked out as 502.3941 and 530.5772; and MAE has been worked out as 423.9089 and 347.6618during training and validation respectively.

Performance of model with 9 neurons is shown in Fig. 4.26 indicates MSE of scaled output and target is 0.0184and 1.1184during training validation, respectively. Coefficient of correlation between estimated wheat yield and observed wheat yield is 0.6167and 0.5151during training and validation, respectively. RMSE has been worked out as 436.0321 and 496.1517; and MAE has been worked out as 347.0187 and 371.6917during training and validation respectively.Thus, the model performed better than model with 6 neurons.

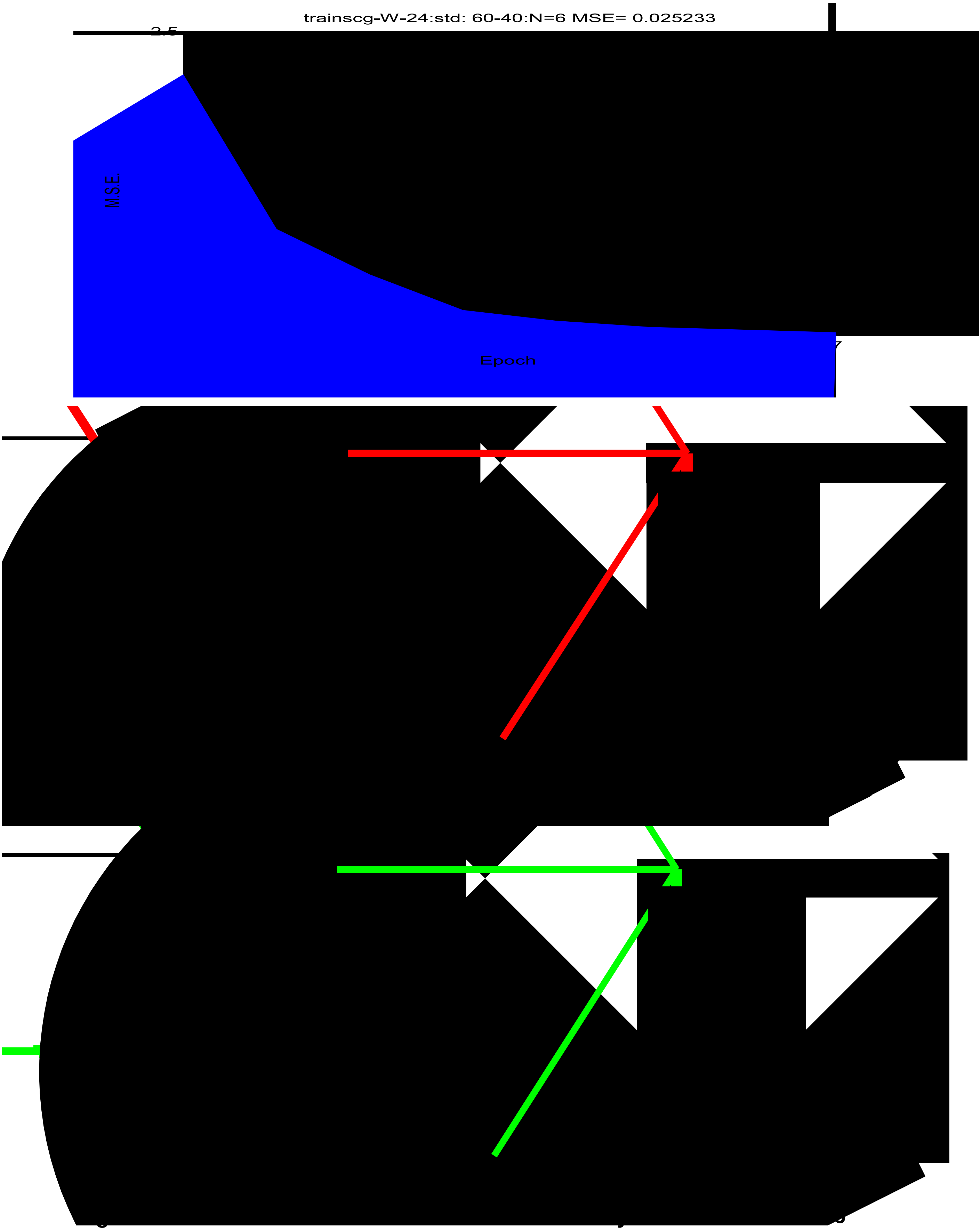
Performance of model with 10 neurons has been depicted in Fig. 4.27.the model gets trained at 8 epochs with MSE= 0.0266. The correlation coefficient between estimated wheat yield and observed wheat yield is 0.6635 and 0.6231 during training and validation, respectively; RMSE and MAE during training is 420.73 and 335.33, respectively; whereasduring validation, RMSE and MAE is 501.86 and 352.85 respectively. The model performed better than model with 9 neurons.

Table 4.9 Performance of neural network with different number of neurons with wheat yield ANN modelling

S.N.	Model	R		MSE		RMSE		MAE	
		Trg	val	trg	val	trg	val	Trg	val
1	N5	0.3016	0.0431	0.2063	1.5613	520.2894	677.0386	411.6595	396.9013
2	N6	0.4326	0.6096	0.0252	1.7291	502.3941	530.5772	423.9089	347.6618
3	N7	0.7484	0.3877	0.0066	1.8321	365.2947	551.2655	282.9732	402.2662
4	N8	0.8335	-0.4225	0.0170	1.8484	337.3418	681.1088	267.8740	426.4330
5	N9	0.6167	0.5151	0.0184	1.1184	436.0321	496.1517	347.0187	371.6917
6	N10	0.6635	0.6231	0.0266	1.1837	420.7317	501.8679	335.3337	352.8562
7	N13	0.5785	-0.2718	.0088	2.0710	452.7904	766.0056	356.8693	532.4971
8	N15	0.8760	-0.3573	0.0021	2.9875	275.5138	910.9472	215.7080	642.5361
9	N20	0.9563	-0.2533	0.0024	2.0662	161.6833	782.9499	120.0632	595.9301

(trg = training, val= validation)

On comparison of performance parameters presented in Table 4.9, it can be stated that model W-24 trained with “trainscg” algorithm, “mapstd” normalization function and 10 neurons performed best. This network architecture has been further improved varying length of data for training and validation.



neurons.

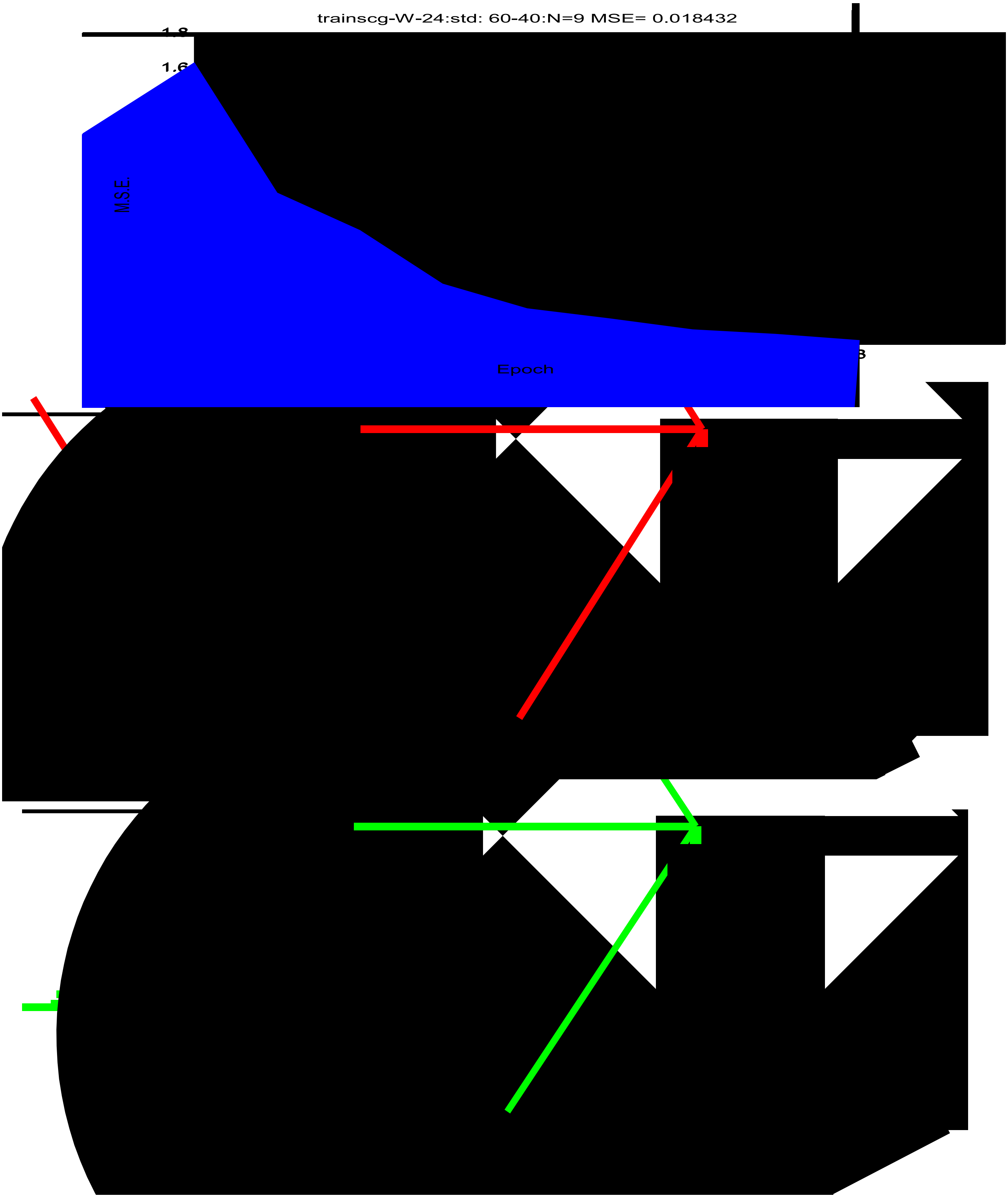
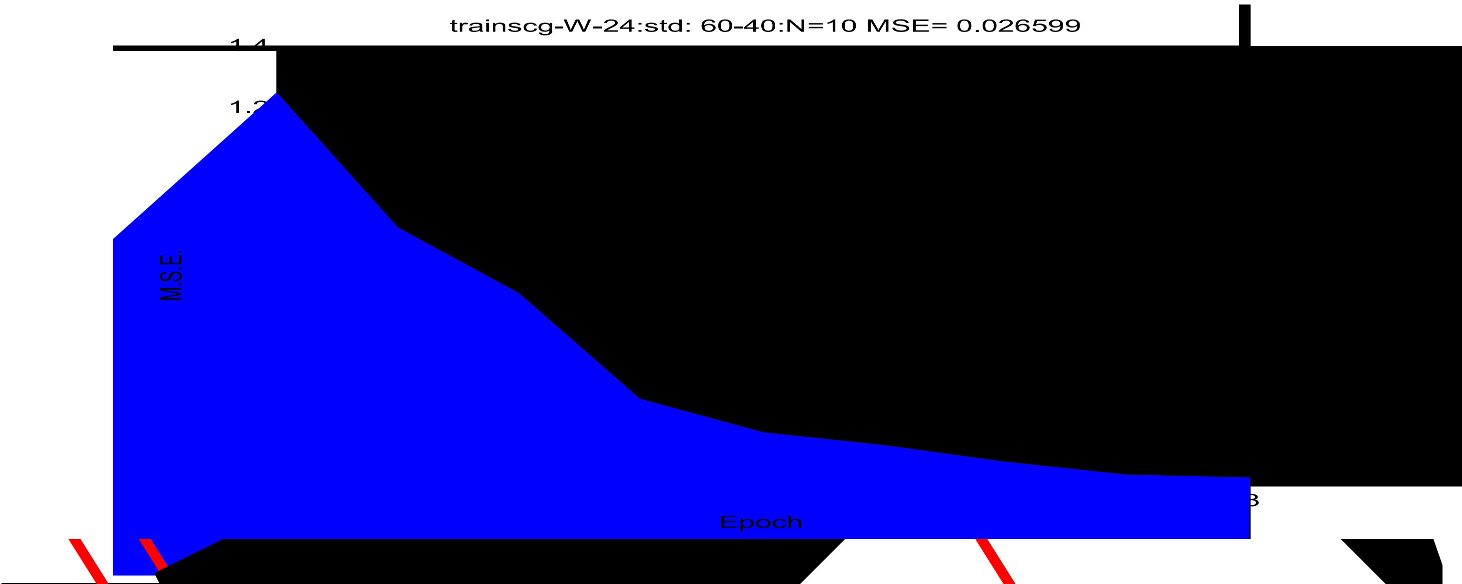
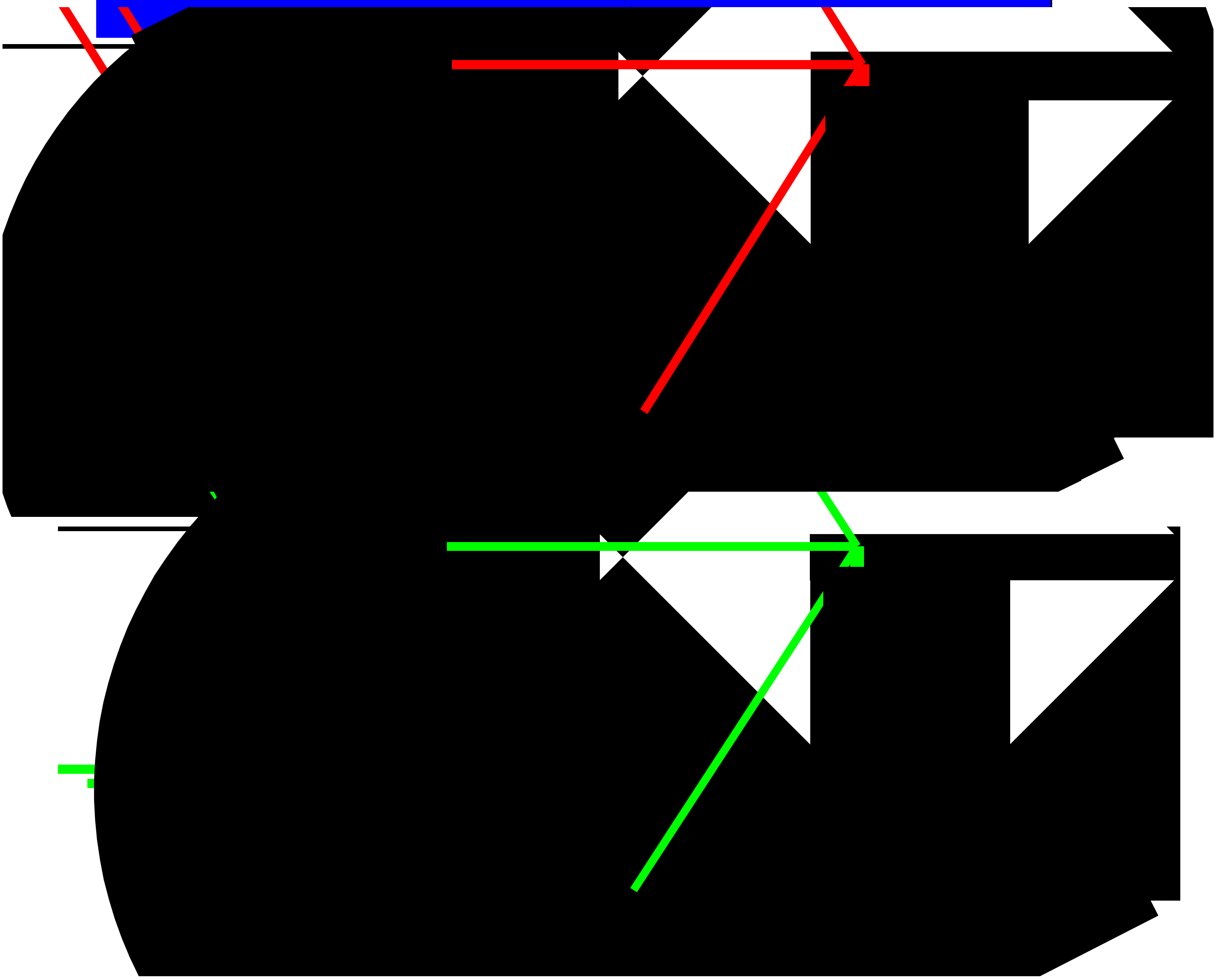


Fig. 4.26Performance of ANN model for wheat yield estimation at 9 neurons.





neurons.

4.2.2.4 Effect of length of record on performance of ANN model for wheat yield estimation

ANN models developed and discussed in section 4.2.2.2 and 4.2.2.3, 60 percent of total length record has been used for model training and validation of the model. An attempt has made to evaluate the performance of Ann model with varying length of data set model training and validation.

ANN model traing with 50 data and tested with 50 percent data resulted correlation coefficient is 0.7985and 0.2174during training and validation respectively. It is found that MSE of estimated and target wheat (scaled) is 0.0145and 1.2037for training and validation, respectively (Fig. 4.28).RMSE and MAE were calculated 370.6431 and 291.8614, respectively during training of model and, 520.9975 and 420.3118during validation. The performance of model training with 60 percent has discussed in section 4.2.2.3 and presented in Fig. 4.27.

When 70 of the total data have been used for training of the model, it is found that MSE of estimated and target wheat yield (scaled) is 0.0051and 1.3496for training and validation, respectively(Fig.4.29). RMSE and MAE were calculated291.6169 and 230.7964, respectively during training of model and 623.0150 and 403.7124during validation. It can also be observed that correlation coefficient is 0.8749and 0.1455during training and validation respectively.

When 80 of the total data have been used for training of the model, it is found that MSE of estimated and target wheat (scaled) is0.0709and 1.6361for training and validation respectively (Fig.4.30). RMSE and MAE were calculated 319.4343 and 248.4652, respectively during training of model and, 874.5481 and 581.4148during validation. It can also be observed from Table 4.10 that correlation coefficient is 0.7742and 0.2195 during training and validation respectively.Model trained with 80 percent of data does not perform well during validation.

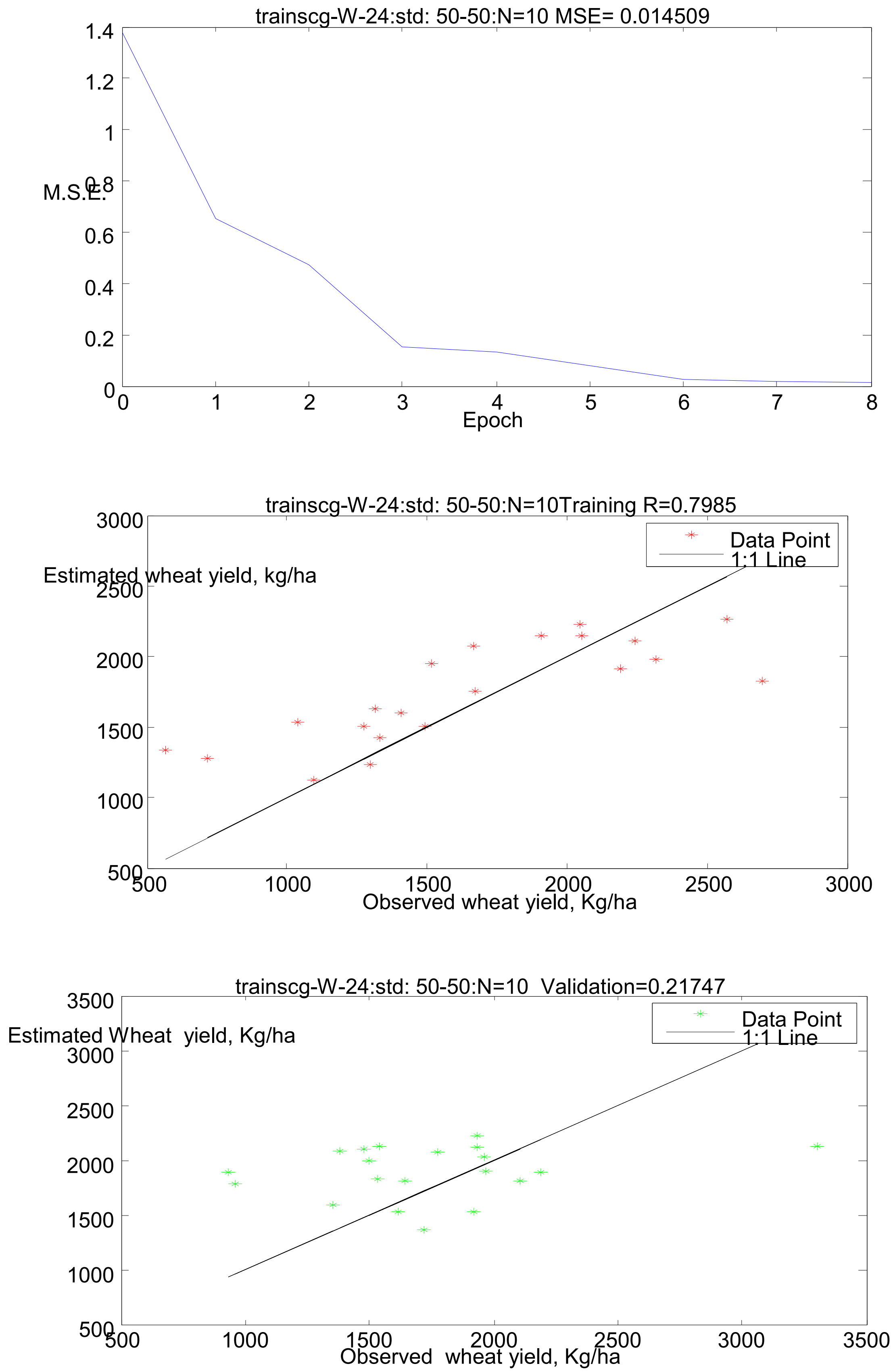
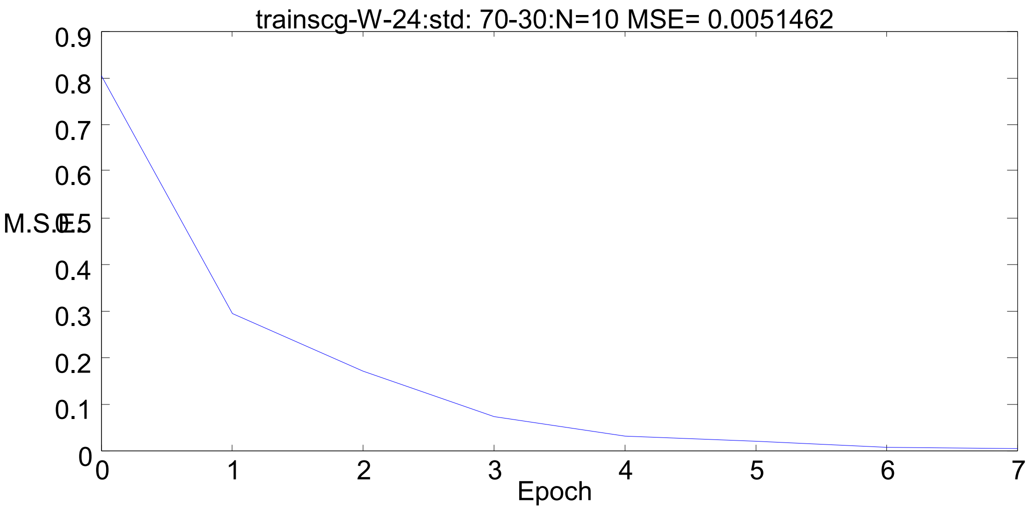
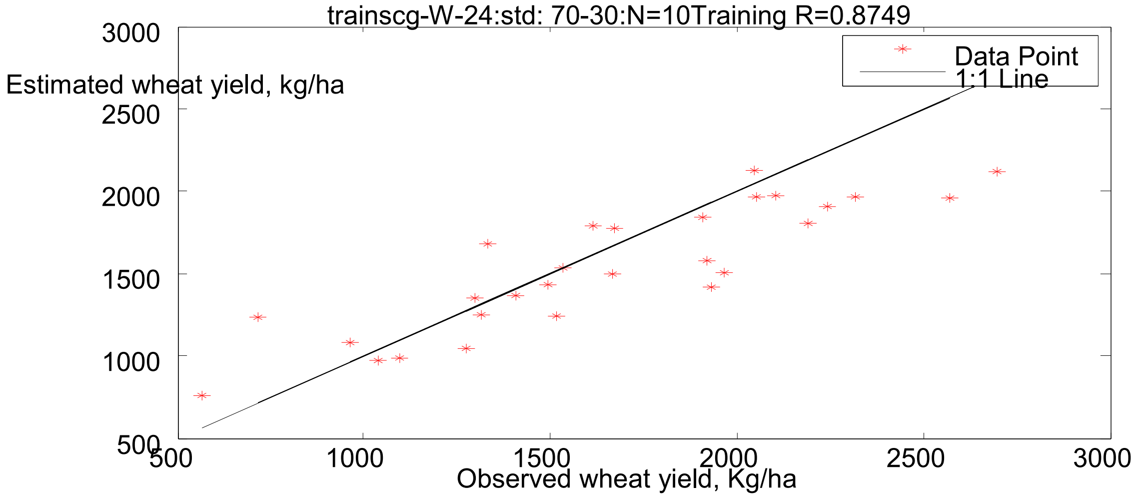


Fig. 4.28 Performance of ANN model for wheat yield estimation trained with 50 percent of dataset





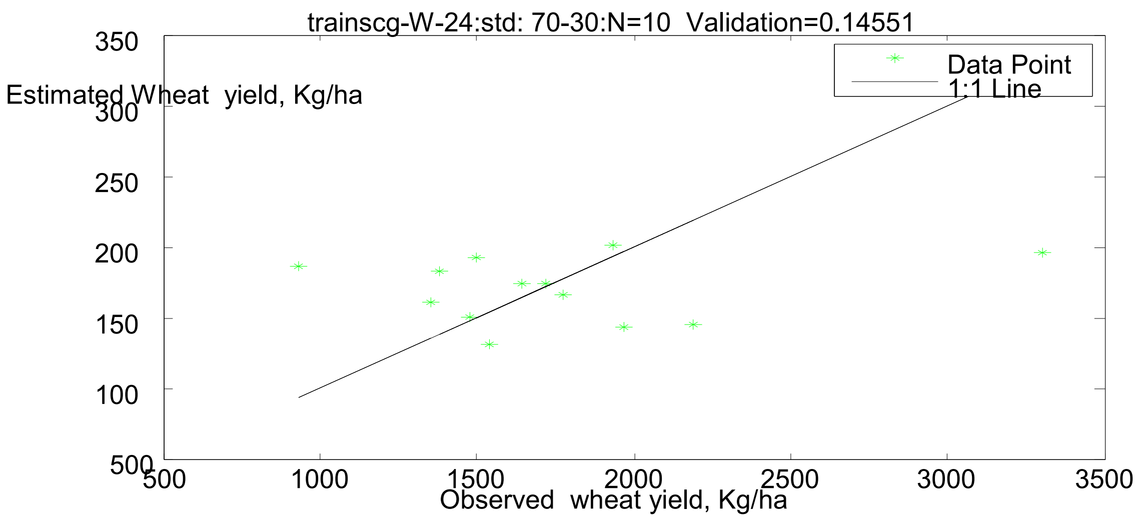
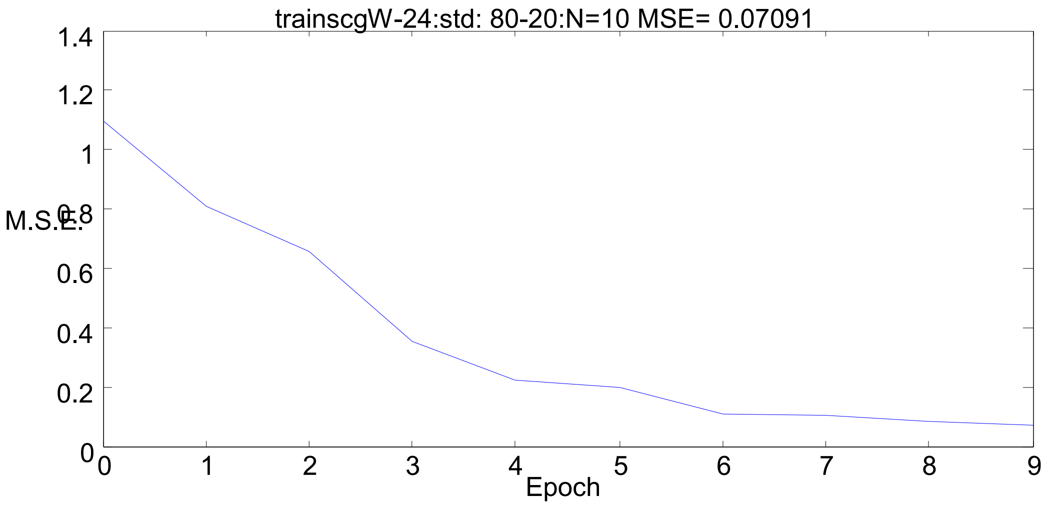
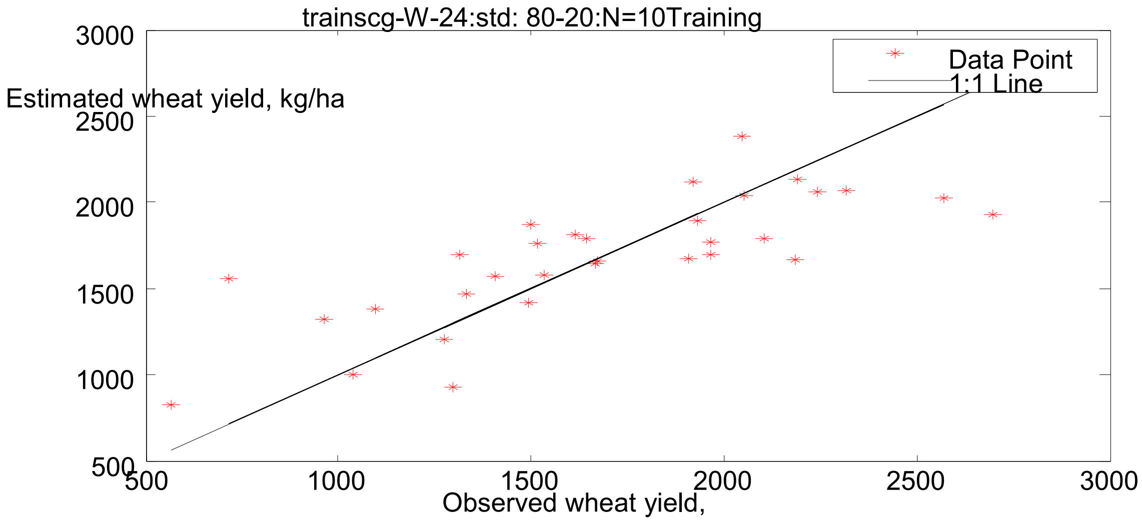


Fig. 4.29 Performance of ANN model for wheat yield estimation trained with 70 percent of dataset





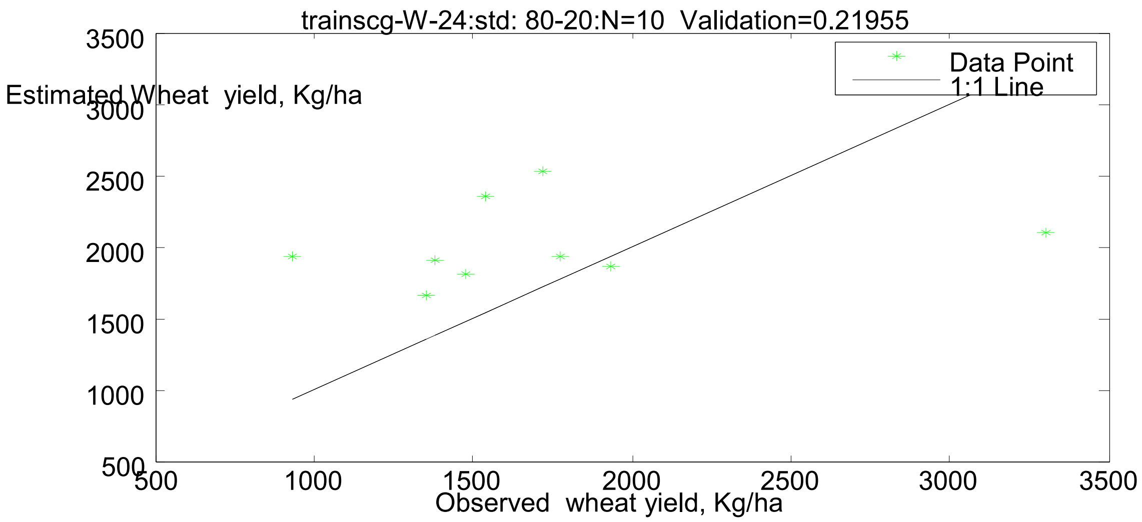


Fig. 4.30 Performance of ANN model for wheat yield estimation trained with 50 percent of dataset

Table 4.10 Performance of Neural Network with varying length of data set for wheat modelling

S.N.	Model	R		MSE		RMSE		MAE	
		Trg	Val	Trg	Val	Trg	Val	Trg	Val
1	L-50	0.7985	0.2174	0.0145	1.2037	370.64	521	291.86	420.31
2	L-60	0.6635	0.6231	0.0266	1.1837	420.73	501.87	335.33	352.86
3	L-70	0.8749	0.1455	0.0051	1.3496	291.62	623.02	230.8	403.71
4	L-80	0.7742	0.2195	0.0709	1.6361	319.43	874.55	248.47	581.41

(trg = training, val= validation)

On the basis of model performance parameters evaluated for model trained with different length of data set (Table 4.10), it can stated that when model trained with 60 per cent of data set, model efficiency is highest amongst the various length of record tested.

.

SUMMARY AND CONCLUTIONS

A study on“Statistical and Neural Method for Site –Specific Yield Prediction” was undertaken for Jabalpur district. Predictor variables were weekly maximum temperature, minimum temperature and weeklyrainfall, whereas response variables were yield of paddy and wheat crops of the district. Attempt was made to build parsimonious model (i.e., has the minimum number of predictor parameters and maximum predictive power). In this study multiple regression model for estimation of crop yield and ANN based model have been developed with following specific objectives.

1. To develop multiple linear regression model for paddy and wheatproductivity for Jabalpur district.
2. To develop an Artificial Neural Network model for paddy andwheat productivity for Jabalpur district.

5.1 Multipleregressionmodels for crop yield estimation

The maximum and minimum temperature, rainfall and yield data were collected from the study area anddeveloped multiple regression modal,was evaluated by comparing estimated values of response parameter and observed parameter, in terms of correlation coefficient (R), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Regression Module of SPSS (16.0.2 version, 2008), was used for developing multiple regression model.

5.2 ANN Architecture

Artificial Neural Network (ANN) model for estimation of crop yield was developed. The model hasone input layer, one hidden layer and one output layer. Method normalizes the data was “std” which transform data such that mean is zero and standard deviation is unity.

The input dataset for crop yield modelling includes weekly rainfall and maximum and minimum temperature since 1969 to 2009.ANN models weredeveloped in Neural Network Module of MATLAB (6.5 version, 2006). Model performance has been evaluated in terms of R, MSE, RMSE and MAE.

5.3 Optimization of the ANN based crop yield model

The basic ANN architecture was optimized in term of training algorithm, number of neurons in the hidden layer, input variables and length of data set for training of the model. Twelve algorithms for training the neural network have been evaluated. Various combinations of input variables have been studied to improve the model efficiency. Performance of the model was evaluated with number of neurons varied from 1 to 20 in the hidden layer.

5.4 Comparison of regression model and ANN model

There is very poor agreement between observed crop yield and estimated crop yield by regression model. The ANN based crop yield developed models based on recorded crop yield performed well during training as well as Validation with independent data set. It was found that ANN model for crop yield estimation performed much better than regression model. The performance of regression model and ANN model are summarized in Table 5.1 and Table 5.2.

Table 5.1 Performance parameters for paddy yield models developed for Jabalpur district

Performance parameter	Regression model (24 predictor variables)	ANN model (24 predictor variables)
RMSE	450.01	100.58
MAE	291.71	82.07
R (Trg)	0.8776	0.9400
R (Val)	0.1778	0.3778

Table 5.2 Performance parameters for wheat yield models developed for Jabalpur district

Performance parameter	Regression model (8 predictor variables)	ANN model (24 predictor variables)
RMSE	794.91	420.73
MAE	619.80	335.33
R (Trg)	0.9404	0.6635
R (Val)	0.4369	0.6231

5.5 Conclusions

From the study carried out following salient points emerged.

1. Estimation of crop yield is possible through the use of regression analysis.
2. Stepwise regression model for paddy includes four input variables i.e.T24, M25, R24, R31.
3. Stepwise regression model for Wheat includes eight input variables i.e.T11, T48, M44, M45, M47, M49, R8, and R50.
4. Enter regression model for paddy includes twenty-four input variables i.e.,T26,T28,T30,T32,T33,T34,T37,M24,M25,M27,M28,M29,M31,M34,M35,M36,M38,R25,R30,R31,R32,R37,R38,R39.
5. Enter regression model for wheat includes twenty-four input variablesi.e.,T44,T46,T52,T1,T2,T3,M46,M49,M7,R44,R45,R46,R47,R48,R50,R3,R5,R6,R7,R8,R10,R11,R12,R13.

6. Highest value of correlation coefficient between estimated and observed paddy yield by regression model 0.8776 and 0.1778 during training and validation.
7. Highest value of correlation coefficient between estimated and observed wheat yield by regression model 0.9404 and 0.4369 during training and validation.
8. Highest value of correlation coefficient between estimated and observed paddy yield0.9400 and 0.3778 during training and validation by ANN.
9. Highest value of correlation coefficient between estimated and observed wheat yield 0.6635 and 0.6231 during training and validation by ANN.
10. “Trainlm” training algorithm of back propagation performed better than other twelve algorithms evaluated in ANN paddy model.
11. Sixneurons in the hidden layer of ANN paddy model performed better.
12. Training with 60percent length of record and validation with rest of data is performed best in ANN paddy model.
13. “Trainscg” training algorithm of back propagation performed better than other twelve algorithms evaluated in ANN wheat model.
14. Ten neurons in the hidden layer of ANN wheat model performed better.
15. Training with 60 percent length of record and validation with rest of data is performed best in ANN wheat model.

Based on the above findings following conclusions can be drawn.

1. Regression model with twenty four input variables is found most suitable for paddy yield estimation and with eight input variables is found most suitable for wheat yield estimation.
2. The ANNmodel with 24 input variables,“trainlm” algorithm, 6 number of neurons, 60 percent and 40percent length of record for training and validation is found best model for paddy yield estimation, and for wheat yield estimation ANN model with 24 input.

with “trainscg” algorithm, 10 number of neurons, 60 percent and 40 percent length of record for training and validation is found best.

3. ANN based models for estimation of crop yield is superior to regression models.

SUGGETIONS FOR FURTHER WORKS

Following suggestions have been made to highlight the future scope of the problem:

1. Artificial Neural Network model can be developed using different weather parameter (humidity, sunshine hours) for yield modelling.
2. Other neural networks like radial basis function may be tried for the estimation of yield.
3. ANN models may be developed to estimate yield form "mapminmax" normalization method.

.

REFERENCES

Adams M.L.,S. E.,Cook, P. A. Caccetta., andM. J. Pringle (1999).Machine learning methods in site–specific management research: An Australian case study. In Proc. 4th Int. Conf. on Precision Agriculture, 1321–1333

AgarwalA. and R. D. Singh, (2004).Runoff modeling through back propagation artificial neural network with variable rainfall–runoff data; Water Resources Management 18285–300.

Agrawal R.,R.C. Jain, and S.E. Mehta, (2001). Yield forecast based on weather variables and agriculturalinput on agroclimatic zone basis. Ind. J. Agri. Sci., 71(7),487-490.

Agricultural statistics at a glance (2010) Directorate of Economics and statistics, GOI.

AhmedM.A., M. H. Fayyaz-ul, andM. A. Aslam (2011).Regression model for the study of sole and cumulative effect of temperature and solar radiation on wheat yield. African Journal of Biotechnology Vol. 10(45), pp. 9114-9121.

Alva A.K.,J. Marcos,,Stockle, C.,V.R. Reddy,andD. Timlin, (2010).A crop simulation model for predicting yield and fate of nitrogen in Irrigated potato rotation cropping system. J. Crop Improv., 24(2), 142-152.

Asana R.D. (1966). Physiological analysis of yield in wheat in relation to water-stress and temperature. Journal of the Post Graduate School(Delhi) 4: 17±31.

Burks,T.F., S.A. Shearer, R.S. Gates, and K.D. Donohue, (2000)Backpropagation neural network design and evaluation for classifying weed species using color image texture. Trans. ASAE 43(4): 1029–1037.

Cerrato, M. E., and A.M.Blackmer, (1990).Comparison of models for describing corn yield response to nitrogen fertilizer.Agron. J. 82(1): 138–143.

Chen, X. and Q. Yan, (2011). prediction Method of Crude Oil Output Based on Artificial Neural Network Computational and Information Sciences (ICCIS), 2011 International Conference on Page(s): 702 – 704.

Coulibaly, P.,F. Anctil, andB. Bobee, (1999).Daily reservoir inflow forecasting using artificial neural network with stopped training approach, Journal of Hydrology, Vol. 230, 244-257.

Drummond S.T.,K.A. Sudduth,A. Joshi,S.J.Birrell, andN.R.Kitchen, (2003).Statistical and Neural Method For Site Specific yield prediction. Vol.46(1) XXX-XXX

Fausett L. (1994). Fundamentals of neural networks.Prentice Hall. Englewood Cliffs, New Jersey.

Fisher, D.,J. Steiner, andD.Endale, (2002). Long-Term Rainfall and Runoff characteristics of a small southern piedmont watershed. Annual research report.USDAARS, Southern piedmont conservation Research Center, Watkinsville, GA 2002, 497-502.

Gallant, S.l. (1986). Three constructive algorithms for network learning, Proceedings of 8th Annual Conference of Cognition Science Society, Michigan.652-660.

Haykin, S. (1994). Neural networks: A comprehensive foundation.Macmillan College Publishing Company, New York, NY.696 pp.

IPCC (2007).Summary for policymakers.In Parry M., Canziani, O. Palutik, J. vander linden,Handson C. Climate change 2007.Impacts adaptation & vulnerability.Contribution of working group to fourth Assessment Report of the Intergovernmental panel on Climate Change II. Cambridge: CambridgeUniversity Press,7-22.

Jeong, D. and Y. Kim, (2005). Rainfall-runoff models using artificial neural networks for ensemble stream flow prediction. Hydrol.Process. 19, 3819-3835.

Kaczmarczyk K.,A Wit,M. Krawczyk& J. Zaborski,(2009).Gait classification in poststroke patients using artificial neural networks.Gait & Posture, 30, 207-210, ISSN0966-6362.

Karunanithi, N.G., D. Whitley, and K. Bovee, (1994). Neural networks for river flow prediction. ASCE J. Comp. Civil Eng. 8(2), 201-220.

Katz R.W. (1977). Assessing the impact of climate change on food production. Climate Change 1, 85–96.

KayamY.,U. Ozsoy,J.Lomas,O.Oden, , M.Mandel, M. Gurbuz, (2000).The impact of climatic change on wheat production of aegeanregion in turkey. The effect of a reduction of rainfall and the increase of temperature on wheat yields (www.fao.Org/sd/climagrimed/pdf/ws01_26.pdf).

Khakural B.R., P.C. Robert, and D.J.Mulla,(1996). Relating corn/soybean yield to variability in soil and landscape. In Proc.3rd International Conference for Precision Agriculture,117-128. Robert, P.C. et al., eds. Madison, Wisc. ASA-CSSA-SSSA.

KitchenN. R., K. A.Sudduth, and S.T. Drummond, (1999).Electrical conductivity as a crop productivity measure for claypan soils. J. Prod. Agric. 12(4): 607–617.

LalM., K.K.Singh, L.S.Rathore, B.G.Srinivasanc, S.A. Saseendran, (2008).Vulnerability of rice and wheat yields in NW India to future changes in climate Agricultural and Forest MeteorologyVolume 89, Issue 2, February 1998, Pages 101–114.

Landau,S., R.A.C. Mitchell, V. Barnett, J.J. Colls, J. Craigon , R.W. Payne (2000). A parsimonious, multiple-regression model of wheat yield response to environment Agricultural and Forest Meteorology 101 (2000) 151–166.

LiA., S. Liang, A. Wang, and J. Qin(2007). Estimating Crop Yield from Multi-temporal Satellite Data Using Multivariate Regression and Neural Network Techniques Photogrammetric Engineering & Remote Sensing Vol. 73, No. 10, pp. 1149–1157.

Liu J.,C. E. Goering, andL. Tian,(2001). A neural network for setting target yields. Trans. ASAE 44(3): 705–713.

Mather P.M. (2000). Computer Processing of Remotely-sensed Images: An Introduction, John Wiley & Sons, Chichester, U.K.

May Robert,D. Graeme and V. H. Maier, (1993) Review of Input Variable Selection Methods for Artificial Neural Networks University of Adelaide Australia.

McClendon R. W., G. Hoogenboon andI. Seginer, (1996).Optimal control and neural networks applied to peanut irrigation management. Trans. Am. Soc. Agric. Eng. 39: 275–279.

Morita, S.,H. Shiratsuchi, J. Takanashi,and K. Fujita, (2002). Effect of high temperature on ripening in rice plants-Comparison of the effects of high night temperatures and high day temperatures.Japanese Journal of Crop Science 71:102–109.

Nicolas, M.E., R.M.Gleadow, and M.J. Dalling(1984).Effects of drought and high temperature on grain growth in wheat. Australian Journal of Plant Physiology 11:553–566.

Othman F. andN. Mahdi, (2011).Reservoir inflow forecasting using artificial neural networkInternational,Journal of the Physical Sciences Vol. 6(3),pp. 434-440, 4 Available online at http://www.academicjournals.org/IJPS ISSN 1992 - 1950 ©2011 Academic Journals.

Qayyum A. and K.P. Muhammad (2010).Impact of weighted rainfalls on the yield of wheat in the Punjab, Pakistan. African Journal of Agricultural Research Vol. 5(22), pp. 3017-3025.

RaiT. and Chandrahas, (2000). Use of discriminant function of weather parameters for developing forecast model for rice crop. (IASRI Publication).

Rumelhart D. E., and J. L. McClelland,(1986) Parallel Distributed Processing. Vol. 1. Boston, Mass.: MIT Press.

Saksena, A., R.C. Jain, and R.L.Yadav,(2001).Development of early warning and yield assessment model for rainfed crops based on agrometeorological indices (IASRI publication).

Sarangi, A., Madramootoo C. A.,P. Enright, S. O.Prasher, andR. M Patel, (2005). Performance evaluation of ANN and geomorphology-based models for runoff and sediment yield prediction for a Canadian watershed. Current Science, vol. 89, no. 12, 25 December (2005).

ShinD.W., G.A.Baigorria, Y.K. Lim,S. Cocke, R. La., T.E. James, J. O'Brien and W. James, (2009).Assessing Crop Yield Simulations with Various Seasonal Climate DataScience and Technology Infusion Climate Bulletin NOAA's National Weather Service 7th NOAA Annual Climate Prediction Application Science Workshop Norman.

Tashiro T. and I.F. Wardlaw, (1989).A comparison of the effect of high temperature on grain development in wheat and rice. Annals of Botany 64: 59±65.

Tokar A. S. &P. A. Johnson, (1999). Rainfall-runoff modeling using artificial neural networks. ASCE, J. Hydrol. Eng. 4(3), 232–239.

Uno Y., Prasher, S.O.,Lacroix, R.,Goel, P.K.,Karimi, Y. A.,Viau, Patel, R.M. (2005). Artificial neural networks to predict corn yield from Compact Airbornea Computers and Electronics in Agriculture Volume 47, Issue 2, May 2005, Pages 149–161.

WangX.,J. Cai, D. Jiang, F. Liu, and T. Dai(2011).Pre-anthesis high-temperature acclimation alleviates damage to theflag leaf caused by post-anthesis heat stress in wheat. J. Plant Physiol., 168(6): 585-593.

WardlawI.F. (1970). The early stages of grain development in wheat:response to light and temperature in a single variety. Australian Journal of Biological Science 23: 765–774.

WardlawLan. (2000) Interaction Between Drought and Chronic High Temperature During Kernel Filling in Wheat in a Controlled Environment Annals of Botany (90:469±476,2002doi:10.093/aob/mcf219,available www.aob.oupjournals.org).

Wheeler T.R.,T.D. Hong, R.H. Ellis,G.R. Batts,J. Morison and P.Hadley (1996).The duration and rate of grain growth, and harvest index ofwheat (*Triticumaestivum*L.) in response to temperature and CO_2. J. Exp. Bot., 47: 623–630.

Yang C.C., O.Prasher, and G. R. Mehuys (1997).An artificial neural network to estimate soil temperature. Can. J. Soil Sci. 77: 421–429.

You M.W.,S. Wood, and D. Sun, (2009).Impact of growing season temperature on wheat productivity in China. Agric. Forest *Mrteor*., 149:1009-1014.

REFERENCES

Adams M.L.,S. E.,Cook, P. A. Caccetta., andM. J. Pringle (1999).Machine learning methods in site–specific management research: An Australian case study. In Proc. 4th Int. Conf. on Precision Agriculture, 1321–1333

AgarwalA. and R. D. Singh, (2004).Runoff modeling through back propagation artificial neural network with variable rainfall–runoff data; Water Resources Management 18285–300.

Agrawal R.,R.C. Jain, and S.E. Mehta, (2001). Yield forecast based on weather variables and agriculturalinput on agroclimatic zone basis. Ind. J. Agri. Sci., 71(7),487-490.

Agricultural statistics at a glance (2010) Directorate of Economics and statistics, GOI.

AhmedM.A., M. H. Fayyaz-ul, andM. A. Aslam (2011).Regression model for the study of sole and cumulative effect of temperature and solar radiation on wheat yield. African Journal of Biotechnology Vol. 10(45), pp. 9114-9121.

Alva A.K.,J. Marcos,,Stockle, C.,V.R. Reddy,andD. Timlin, (2010).A crop simulation model for predicting yield and fate of nitrogen in Irrigated potato rotation cropping system. J. Crop Improv., 24(2), 142-152.

Asana R.D. (1966). Physiological analysis of yield in wheat in relation to water-stress and temperature. Journal of the Post Graduate School(Delhi) 4: 17±31.

Burks,T.F., S.A. Shearer, R.S. Gates, and K.D. Donohue, (2000)Backpropagation neural network design and evaluation for classifying weed species using color image texture. Trans. ASAE 43(4): 1029–1037.

Cerrato, M. E., and A.M.Blackmer, (1990).Comparison of models for describing corn yield response to nitrogen fertilizer.Agron. J. 82(1): 138–143.

Chen, X. and Q. Yan, (2011). prediction Method of Crude Oil Output Based on Artificial Neural Network Computational and Information Sciences (ICCIS), 2011 International Conference on Page(s): 702 – 704.

Coulibaly, P.,F. Anctil, andB. Bobee, (1999).Daily reservoir inflow forecasting using artificial neural network with stopped training approach, Journal of Hydrology, Vol. 230, 244-257.

Drummond S.T.,K.A. Sudduth,A. Joshi,S.J.Birrell, andN.R.Kitchen, (2003).Statistical and Neural Method For Site Specific yield prediction. Vol.46(1) XXX-XXX

Fausett L. (1994). Fundamentals of neural networks.Prentice Hall. Englewood Cliffs, New Jersey.

Fisher, D.,J. Steiner, andD.Endale, (2002). Long-Term Rainfall and Runoff characteristics of a small southern piedmont watershed. Annual research report.USDAARS, Southern piedmont conservation Research Center, Watkinsville, GA 2002, 497-502.

Gallant, S.l. (1986). Three constructive algorithms for network learning, Proceedings of 8th Annual Conference of Cognition Science Society, Michigan.652-660.

Haykin, S. (1994). Neural networks: A comprehensive foundation.Macmillan College Publishing Company, New York, NY.696 pp.

IPCC (2007).Summary for policymakers.In Parry M., Canziani, O. Palutik, J. vander linden,Handson C. Climate change 2007.Impacts adaptation & vulnerability.Contribution of working group to fourth Assessment Report of the Intergovernmental panel on Climate Change II. Cambridge: CambridgeUniversity Press,7-22.

Jeong, D. and Y. Kim, (2005). Rainfall-runoff models using artificial neural networks for ensemble stream flow prediction. Hydrol.Process. 19, 3819-3835.

Kaczmarczyk K.,A Wit,M. Krawczyk& J. Zaborski,(2009).Gait classification in poststroke patients using artificial neural networks.Gait & Posture, 30, 207-210, ISSN0966-6362.

Karunanithi, N.G., D. Whitley, and K. Bovee, (1994). Neural networks for river flow prediction. ASCE J. Comp. Civil Eng. 8(2), 201-220.

Katz R.W. (1977). Assessing the impact of climate change on food production. Climate Change 1, 85–96.

KayamY.,U. Ozsoy,J.Lomas,O.Oden, , M.Mandel, M. Gurbuz, (2000).The impact of climatic change on wheat production of aegeanregion in turkey. The effect of a reduction of rainfall and the increase of temperature on wheat yields (www.fao.Org/sd/climagrimed/pdf/ws01_26.pdf).

Khakural B.R., P.C. Robert, and D.J.Mulla,(1996). Relating corn/soybean yield to variability in soil and landscape. In Proc.3rd International Conference for Precision Agriculture,117-128. Robert, P.C. et al., eds. Madison, Wisc. ASA-CSSA-SSSA.

KitchenN. R., K. A.Sudduth, and S.T. Drummond, (1999).Electrical conductivity as a crop productivity measure for claypan soils. J. Prod. Agric. 12(4): 607–617.

LalM., K.K.Singh, L.S.Rathore, B.G.Srinivasanc, S.A. Saseendran, (2008).Vulnerability of rice and wheat yields in NW India to future changes in climate Agricultural and Forest MeteorologyVolume 89, Issue 2, February 1998, Pages 101–114.

Landau,S., R.A.C. Mitchell, V. Barnett, J.J. Colls, J. Craigon , R.W. Payne (2000). A parsimonious, multiple-regression model of wheat yield response to environment Agricultural and Forest Meteorology 101 (2000) 151–166.

LiA., S. Liang, A. Wang, and J. Qin(2007). Estimating Crop Yield from Multi-temporal Satellite Data Using Multivariate Regression and Neural Network Techniques Photogrammetric Engineering & Remote Sensing Vol. 73, No. 10, pp. 1149–1157.

Liu J.,C. E. Goering, andL. Tian,(2001). A neural network for setting target yields. Trans. ASAE 44(3): 705–713.

Mather P.M. (2000). Computer Processing of Remotely-sensed Images: An Introduction, John Wiley & Sons, Chichester, U.K.

May Robert,D. Graeme and V. H. Maier, (1993) Review of Input Variable Selection Methods for Artificial Neural Networks University of Adelaide Australia.

McClendon R. W., G. Hoogenboon andI. Seginer, (1996).Optimal control and neural networks applied to peanut irrigation management. Trans. Am. Soc. Agric. Eng. 39: 275–279.

Morita, S.,H. Shiratsuchi, J. Takanashi,and K. Fujita, (2002). Effect of high temperature on ripening in rice plants-Comparison of the effects of high night temperatures and high day temperatures.Japanese Journal of Crop Science 71:102–109.

Nicolas, M.E., R.M.Gleadow, and M.J. Dalling(1984).Effects of drought and high temperature on grain growth in wheat. Australian Journal of Plant Physiology 11:553–566.

Othman F. andN. Mahdi, (2011).Reservoir inflow forecasting using artificial neural networkInternational,Journal of the Physical Sciences Vol. 6(3),pp. 434-440, 4 Available online at http://www.academicjournals.org/IJPS ISSN 1992 - 1950 ©2011 Academic Journals.

Qayyum A. and K.P. Muhammad (2010).Impact of weighted rainfalls on the yield of wheat in the Punjab, Pakistan. African Journal of Agricultural Research Vol. 5(22), pp. 3017-3025.

RaiT. and Chandrahas, (2000). Use of discriminant function of weather parameters for developing forecast model for rice crop. (IASRI Publication).

Rumelhart D. E., and J. L. McClelland,(1986) Parallel Distributed Processing. Vol. 1. Boston, Mass.: MIT Press.

Saksena, A., R.C. Jain, and R.L.Yadav,(2001).Development of early warning and yield assessment model for rainfed crops based on agrometeorological indices (IASRI publication).

Sarangi, A., Madramootoo C. A.,P. Enright, S. O.Prasher, andR. M Patel, (2005). Performance evaluation of ANN and geomorphology-based models for runoff and sediment yield prediction for a Canadian watershed. Current Science, vol. 89, no. 12, 25 December (2005).

ShinD.W., G.A.Baigorria, Y.K. Lim,S. Cocke, R. La., T.E. James, J. O'Brien and W. James, (2009).Assessing Crop Yield Simulations with Various Seasonal Climate DataScience and Technology Infusion Climate Bulletin NOAA's National Weather Service 7th NOAA Annual Climate Prediction Application Science Workshop Norman.

Tashiro T. and I.F. Wardlaw, (1989).A comparison of the effect of high temperature on grain development in wheat and rice. Annals of Botany 64: 59±65.

Tokar A. S. &P. A. Johnson, (1999). Rainfall-runoff modeling using artificial neural networks. ASCE, J. Hydrol. Eng. 4(3), 232–239.

Uno Y., Prasher, S.O.,Lacroix, R.,Goel, P.K.,Karimi, Y. A.,Viau, Patel, R.M. (2005). Artificial neural networks to predict corn yield from Compact Airbornea Computers and Electronics in Agriculture Volume 47, Issue 2, May 2005, Pages 149–161.

WangX.,J. Cai, D. Jiang, F. Liu, and T. Dai(2011).Pre-anthesis high-temperature acclimation alleviates damage to theflag leaf caused by post-anthesis heat stress in wheat. J. Plant Physiol., 168(6): 585-593.

WardlawI.F. (1970). The early stages of grain development in wheat:response to light and temperature in a single variety. Australian Journal of Biological Science 23: 765–774.

WardlawLan. (2000) Interaction Between Drought and Chronic High Temperature During Kernel Filling in Wheat in a Controlled Environment Annals of Botany (90:469±476,2002doi:10.093/aob/mcf219,available www.aob.oupjournals.org).

Wheeler T.R.,T.D. Hong, R.H. Ellis,G.R. Batts,J. Morison and P.Hadley (1996).The duration and rate of grain growth, and harvest index ofwheat (*Triticumaestivum*L.) in response to temperature and CO_2. J. Exp. Bot., 47: 623–630.

Yang C.C., O.Prasher, and G. R. Mehuys (1997).An artificial neural network to estimate soil temperature. Can. J. Soil Sci. 77: 421–429.

You M.W.,S. Wood, and D. Sun, (2009).Impact of growing season temperature on wheat productivity in China. Agric. Forest *Mrteor*., 149:1009-1014.

VITA

The author of the thesis is is **Mr. Pramod Kumar Meena** s/o Mr. K.L. Meena was born on 8th sep 1988 at Gadarwara (M.P.) He has completed his Higher Secondary School Examination from Govt. Higher Secondary School, Narsinghpur (M.P.) in the year of 2006 with 82% under the M.P.Board, Bhopal.

He was admitted to B.Tech. Degree course in Agricultural Engineering at College of Agricultural Engineering, J.N.K.V.V. Jabalpur (M.P.) in the year of 2006 and successfully completed the degree in the year of 2010. With a CGPA of 7.3 out of 10.00 point scale.

In the year 2010 he joined **"Master of Technology"** two year Post Graduation degree programme in the Department of Soil and Water Engineering at College of Agricultural Engineering, J.N.K.V.V. Jabalpur. (M.P.). After completing the entire prescribed course work successfully, he has submitted the thesis entitled "*Statistical and Neural Method for Site –Specific Yield Prediction* " in partial fulfillment of the requirements for the award of degree of "Master of Technology".