International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

A SMALL SURVEY ON EFFECT OF WEATHER CONDITIONS ON AGRICULTURE USING VARIOUS TECHNIQUES

A.MANASA
Assistant Professor, Department of Computer Science and Engineering,
Siddhartha Institute of Technology and Sciences, Narapally, Hyderabad, Telangana, India

ABSTRACT: Agriculture is highly dependent on the spatial and temporal distribution of monsoon rainfall. This paper presents an analysis of crop–weather condition relationships for Andhra Pradesh, using important production statistics for major crops like rice, cotton, tobacco, etc.,. It is observed that the crop prediction is affected due to weather changes. Andhra Pradesh is a rained farming area where rice, drought-tolerant crops and several types of vegetables and many more are cultivated. The effects of weather change were estimated by inputting meteorological data that reflects the influence of weather change. The crop yield prediction is summarized and displayed as output.

Keywords— Precipitation, Vapour Pressure, Temperature, Crop Yield, Regression, Prediction

I. INTRODUCTIONTHE EFFECTS OF CLIMATE CHANGE (ON TEMPERATURE, PRECIPITATION, AND VAPOUR PRESSURE) CAN SIGNIFICANTLY AFFECT AGRICULTURE. MOST COUNTRIES IN THE ASIAN MONSOON REGION ARE AGRICULTURAL AREAS, AND THE IMPACT OF CLIMATE CHANGE ON THIS REGION IS VERY SIGNIFICANT. RAIN-FED AREAS, IN PARTICULAR, ARE VERY VULNERABLE TO EXTREME WEATHER EVENTS, SUCH AS FLOODS AND DROUGHTS. ON THE OTHER HAND, IN MOUNTAIN REGIONS, IT IS EXPECTED THAT AIR TEMPERATURE CHANGES WILL BE AMPLIFIED. MOREOVER, AN AMENDMEN T IN THE ONSET AND PERIOD OF THE RAINY SEASON WILL HAVE SERIOUS EFFECTS ON YIELDS AND FARMING PLAN. TO PREDICT THE CROP YIELD BASING ON THE CLIMATIC CHANGES SUCH AS TEMPERATURE, PRECIPITATION ,VAPOUR PRESSURE WE USE LINEAR REGRESSION ANALYSIS IN R PROGRAMMING. IT IS OBSERVED THAT THE ACCURATE AND TIMELY WEATHER FORECASTING IS A MAJOR CHALLENGE FOR THE SCIENTIFIC COMMUNITY. CROP YIELD PREDICTION MODELLING COMPRISES A COMBINATION OF COMPLETE MODELS. OBSERVATION AND KNOWLEDGE OF TRENDS AND PATTERNS [1]. WHILE APPLYING THESE METHODS THE PREDICTION OF CROP YIELD STUDYING DIFFERENT CLIMATIC CHANGES. REGRESSION IS STATISTICAL EMPIRICAL TECHNIQUE AND IS WIDELY USED IN BUSINESS, THE SOCIAL AND BEHAVIOURAL SCIENCES, THE BIOLOGICAL SCIENCES, CLIMATE PREDICTIONS, AND IN MANY OTHER AREAS. IN INDIA CROP YIELD VARIES SIGNIFICANTLY FROM REGION TO REGION. THE PROPOSED STUDY ANALYSES THE IMPACT OF VARIOUS GLOBAL CLIMATIC PARAMETERS SUCH A S TEMPERATURE, PRECIPITATION AND VAPOUR PRESSURE OF KRISHNA DISTRICT OF ANDHRA PRADESH. IN THIS ANALYSIS, CORRELATION AND LINEAR REGRESSION APPROACH ARE USED FOR PREDICTION OF CROP YIELD OVER KRISHNA DISTRICT.

Ii. Data analysisthe crop yield prediction is analyzed using the methods of networks, and data mining in some of the journals. Artificial intelligence and neural networks are more complicated when compared to data mining because artificial intelligence involves some artificial neural networks are computational models inspired by animals' central nervous systems (in particular the brain) that are capable of machine learning and pattern recognition. In data mining, some of the functionalities are used i.e. Classification, clustering, regression or prediction, association, etc. Finally, the prediction methods are taken into consideration and are done by applying the regression approach. In the process of regression, kerl pearson correlation coefficient is used in finding the measure of crop yield in kg's/hectare in the particular region. It is an obligation to predict the crop yield in the coming years, using the linear regression approach. While using the data, the crop yield is computed with the help of correlations coefficient. The correlation coefficient is compared with the predicted data with the help of regression approach. Crop yield is measured regarding kilogram/hectare(kg/hec) years, and crop yield is given in the graph with x and y-axis respectively. To use the input data, the output is given as soon as possible. Therefore, the output is predicted for the future year's crop yield[3]. It is observed that the output data is approximate because the prediction of the harvest yield for the coming years calculated using regression approach. In this method, some of the predictor variables are used which is useful to predict the crop yield during 2000 to 2013 years+. The data utilized in the present study is collected from the chief planning officer of krishna district, andhra pradesh.

Iii. Regression

Regression analysis is a statistical process for estimating the relationships among variables. It includes many techniques for modeling and analyzing several variables, when the focus is on the link between a dependent and one or more independent variables. Regression is of two methods namely simple linear regression and multiple regression models.

Standard linear regression. Regression is a statistical measure that attempts to determine the strength of the relationship between one dependent variable (i.e. The label attribute) and a series of other changing variables known as independent variables (regular attributes). Predicting specific labels. In the standard linear regression model, we consider linear regression models of the form[6].

$$y = f(x_1, x_2, \dots, x_k) + \varepsilon = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon,$$

Least squares usually achieve the estimation of the parameters. The least squares estimate (α, β1,…. , βk) minimize the sum of the squared differences between the observations and the values that are implied by the model,

$$D(\alpha, \beta_1, \dots, \beta_k) = \sum_{i=1}^{n} \left[y_i - (\alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik})\right]^2.$$

The expression below is called fitted value

$$\widehat{y}_i = \widehat{\alpha} + \widehat{\beta}_1 x_{i1} + \widehat{\beta}_2 x_{i2} + \cdots + \widehat{\beta}_k x_{ik}$$

And the difference $y_i - \widehat{y}_i$ is known as the residual. The minimizing value $\widehat{D} = D(\widehat{\alpha}, \widehat{\beta}_1, \dots, \widehat{\beta}_k) = \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2$ determines the estimate of $\mathrm{Var}(\varepsilon_i) = \sigma^2$, , the r-square, and the f-statistic for testing the overall significance of the regression. The unbiased estimate of σ2 is given by

$$\widehat{\sigma}^2 = \frac{\widehat{D}}{n - k - 1}.$$

the r-square,[14].

$$R^2 = 1 - \left[\frac{\hat{D}}{\sum (yi - \overline{y})^2}\right]$$

Expresses the proportion of variation that is explained by the regression model.

The *f*-statistic,[14].

$$F = \frac{\left[\sum_{i=1}^{n} (yi - \overline{y})^2 - \hat{D}\right] / k}{\hat{D} / (n-k-1)}$$

The adjusted r-square for a model with k regressors and k + 1 estimated coefficients,

$$R^2_{\text{adj}} = 1 - \frac{\widehat{D}/(n-k-1)}{\sum (y_i - \overline{y})^2/(n-1)},$$

given a set of predictions for m new cases, we can evaluate the predictions according to their (me and rmse[10],mape[11])

Mean error: $$\text{ME} = \left(\frac{1}{m}\right)\sum_{i=1}^{m}(y_i - \hat{y}_i),$$

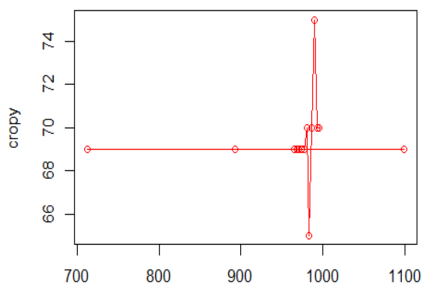
Root mean square error: $$\text{RMSE} = \sqrt{\left(\frac{1}{m}\right)\sum_{i=1}^{m}(y_i - \hat{y}_i)^2},$$ and

Mean absolute percent error: $$\text{MAPE} = \frac{100}{m}\sum_{i=1}^{m}\frac{|y_i - \hat{y}_i|}{y_i}.$$

Description:

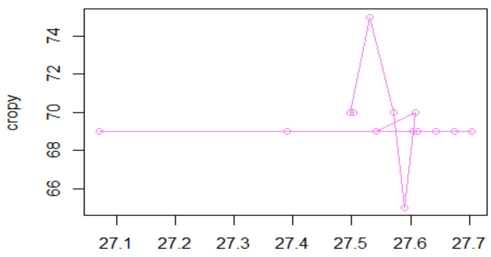
first, create some data sets like a year, precipitation vapour pressure, temperature and crop yield by assigning some values to them. Next, we plotted the graphs for them based on the given values till the year 2013 by the help of the following code.

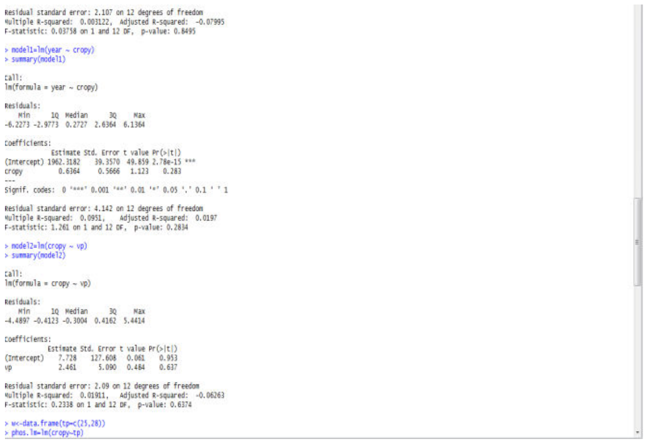
- Plot(precipitation,cropy,type= "o",col="red")



precipitation

- Plot(tp,cropy,type= "o",col="violet")

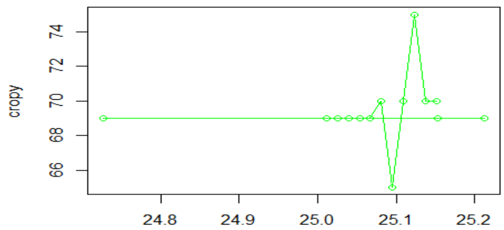




```
Residual standard error: 2.107 on 12 degrees of freedom
Multiple R-squared:  0.003122,  Adjusted R-squared:  -0.07995
F-statistic: 0.03738 on 1 and 12 DF,  p-value: 0.8495

> model1=lm(year ~ cropy)
> summary(model1)

Call:
lm(formula = year ~ cropy)

Residuals:
    Min      1Q  Median      3Q     Max
-6.2273 -2.9773  0.2727  2.6364  6.1364

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1962.3182    39.3570  49.859 2.78e-15 ***
cropy          0.6364     0.5666   1.123    0.283
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.142 on 12 degrees of freedom
Multiple R-squared:  0.0951,   Adjusted R-squared:  0.0197
F-statistic: 1.261 on 1 and 12 DF,  p-value: 0.2834

> model2=lm(cropy ~ vp)
> summary(model2)

Call:
lm(formula = cropy ~ vp)

Residuals:
    Min      1Q  Median      3Q     Max
-4.4897 -0.4123 -0.3004  0.4162  5.4414

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    7.728    127.608   0.061    0.953
vp             2.461      5.090   0.484    0.637

Residual standard error: 2.09 on 12 degrees of freedom
Multiple R-squared:  0.01911,   Adjusted R-squared:  -0.06263
F-statistic: 0.2338 on 1 and 12 DF,  p-value: 0.6374

> w<-data.frame(tp=c(25,28))
> phos.lm=lm(cropy~tp)
```

tp

- Plot(vp,cropy,type= "o",col="green")



Vp

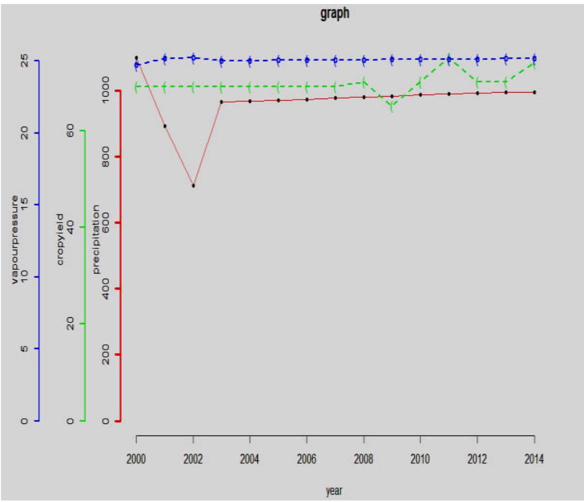
- Plot(year,cropy,type= "o",col="brown")

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Prediction graph:

we often have to plot multiple time-series with different scale of values for comparative purposes, and although placing them in different rows are useful, put on the same graph is still useful sometimes

Red=precipitation
Green=crop yield
Blue=vapour pressure



Iv .conclusion

In this paper, we applied multiple linear regression approach to extract knowledge from krishna district climate and crop yield. The dataset includes thirteen years period from 2000 to 2013 climate and crop yield observation. We have gone through all prediction process and applied many prediction algorithms like linear regression. Multiple linear regressions provide a very useful and accurate knowledge in a form of rules, models, and visual graphs as shown in the figures. This knowledge can be used to obtain useful prediction and support the decision making for different sectors.

v. Acknowledgment

The authors thank the meteorological department, machilipatnam, and krishna district for providing the opportunity for sharing information and receiving useful comments on climatic data and crop yield.

Vi. References

[1] m.kannan, rainfall forecasting using data mining technique

[2] empirical method technique to make short term forecast of rainfall for a specific region

[3] a programming environment for data analysis and graphics version 3.2.3 page no.-59

[4] http://tryr.codeschool.com/levels/1/challenges/

[5] abraham, b. And ledolter, j.: statistical methods for forecasting. New york: john wiley & sons, inc., 1983

[6] further testing issues testing multiple restrictions and mis-specification da v i d c. Br o a d s t o c k research institute of economics & management, china.

[7] data mining and business analytics with r, first edition. Johannes ledolter.2013 john wiley & sons, inc. Published 2013 by john wiley & sons, inc.

[8] http://www.r-tutor.com/elementary-statistics/simple-linear-regression/significance-test-linear-regression lm=..

[9] http://www.statmethods.net/stats/correlations.html

[10] root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature t. Chai and r. R. Draxler

[11] using the mean absolute percentage error for regression models arnaud de myttenaere , boris golden , b´en´edicte le grand & fabrice rossi

[12] abraham, b. And ledolter, j.: introduction to regression modeling. Belmont, ca: duxbury press, 2006.

[13] hawkins, d.m.: the problem of overfitting. Journal of chemical information and computer science, vol. 44 (2004), 1–12.

[14]data_mining_and_business_analytics_with_r__johannes_ledolter,johannes ledolter

[15] ledolter, j. And abraham, b.: parsimony and its importance in time series forecasting. Technometrics,vol. 23 (1981), 411–414.

[16] shmueli, g., patel, n.r., and bruce, p.c.: data mining for business intelligence. Second edition.hoboken, nj: john wiley & sons, inc., 2010.

[17] brinkman, n.d.: ethanol fuel—a single-cylinder engine study of efficiency and exhaust emissions.sae transactions, vol. 90 (1981).

[18] cleveland, w.s. And devlin, s.j.: locally weighted regression: an approach to regression analysis by local fitting. Journal of the american statistical association, vol. 83 (1988),596–610.

[19] craven, p. And wahba, g.: smoothing noisy data with spline functions. Numerische mathematic,vol. 31 (1979), 377–403.

[20] epanechnikov, v.a.: nonparametric estimates of a multivariate probability density. Theory of probability and its applications, vol. 14 (1969), 153–158.

[21] fan, j. And gijbels, i.: local polynomial modelling and its applications. London: chapman and hall, 1996.