

# Assessing the Role of Precipitation in Wheat and Barley Germination in Mosul, Iraq

<sup>1</sup>Lamia Khalil Ibrahim, <sup>2</sup>Bashar A. Al-Talib

<sup>1,2</sup>Department of Statistics and Informatics, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq  
E-mails: [Lamia.23csp131@student.uomosul.edu.iq](mailto:Lamia.23csp131@student.uomosul.edu.iq), [bashar.altalib@uomosul.edu.iq](mailto:bashar.altalib@uomosul.edu.iq)

**Abstract** - The present research evaluates the influence of precipitation and other associated climatic factors within wheat and barley germination in Mosul, Iraq, between the 2012 and 2024. The analysis encounters such issues as incomplete data, small sample size ( $n=13$ ), and some outliers. Multiple Imputation by Chained Equations (MICE) was used to impute the missing values to ensure the variability was realistic, as well as to maintain the inter-variable relationships. Ordinary least squares (OLS) and robust estimators (Huber, Tukey, Hampel, S, and MM) regression tests were performed to measure the importance of rainfall and other climatic distinguishing variables including temperature, humidity, wind, and sun radiant. In wheat, OLS model had the  $R^2$  value of 0.479, as compared to adjusted  $R^2$  value of -0.25, the level of error in the residual was about 515,500 tons as compared to Tukey M- estimator whose error level in the residual was about 104,000 tons which showed the significance of robust methods when dealing with outliers. In barley, OLS gave  $R^2 = 0.442$ , the standard error of the residual = cannot be smaller than 991, 300 tons, whereas robust estimators gave a higher degree of stability. Five observations of wheat were detected as strong outliers and a few observations of barley were strong outliers, and the apparent significant association between climatic variables (e.g., wind direction vs. wheat: -0.792) is some suggestion of the presence of multicollinearity. Monte Carlo simulations showed that methods that are robust are more effective than OLS in the presence of Y contamination and LTS do not perform the same in the presence of X contamination. These results demonstrate the importance of precipitation in crop germination and the need to apply efficient statistical methods, outlier test, and prudent management of this multicollinear weather variable to come up with effective predictions of Mosul wheat and barley products.

**Keywords:** Wheat germination, Barley germination, Precipitation impact, Robust regression, Climatic factors.

## I. INTRODUCTION

The Iraqi economy is still largely based on agricultural food as agriculture is the mainstay of food security with cereal foods, barley and wheat being the primary food staple. Climatic conditions especially the amount of rainfall are crucial to the successful growth of these crops [1]. The most determinant factor in germination of plant seeds has been identified to be rainfall. Indicatively, in semi-arid regions of Mosul in the north of Iraq, there is a high variation in the level of rainfall. Such circumstances become a challenge to the agricultural sector regarding the successful germination of the seeds to grow into the desired volumes of produce [2].

There is a complicated connection between the cereal crops and rainfall and their productivity. On the one hand, the heavy rainfall guarantees the homogenous germination and the maximum growth. Conversely, excessive rainfall can cause low concentration of aeration of soil through waterlogging. Moreover, other complications, such as fungal diseases, also check in with a higher probability of occurrence [3]. Quite the contrary, low quantities of water fall can lead to low levels of germination. Meanwhile, it is also possible that growth will be influenced. Moreover, other factors of the climate also could influence the development of the plants. These influences are considered to be temperature, humidity, solar radiation, wind speed and direction [4].

Regression analysis methods have been widely utilized in the agricultural research in order to determine the influence of climatic index on crop yield. Linear regression methods, which are standard methods used in analysing the regression analysis, such as ordinary least squares regression analysis are used to analyse the effect of several predictor climatic indices on crop production. In real-life examples of these regression methods, the values used in the climatic indices under measure might be missing [5]. The datasets can have missing data in a farm survey or weather observation. The absence of a weather observation record or the outcome of a climate observation error are some of the causes of missing values. Also, the points of the dataset could be beyond the range of the predicted values. These arguments can be presented in the form of severe climatic conditions [6].

Since it is impossible to effectively work with the missing values in the dataset, the research outlined to the application of the Multiple Imputation by Chained Equations (MICE) approach coupled with the missForest methodology. The imputation is performed using the missForest methodology that involves the utilization of a random forest algorithm [7]. The principal strength of the MICE method of handling the issue of missing values in a set of data is that multiple complete data sets could be generated through a series of reiterated prediction of the missing values utilizing the ones that were present in the dataset. It is due to this that the uncertainty of the process of imputation is now able to be accurately measured. On the other hand, the missForest algorithm considers the effective utilization of numerous decision trees to forecast the missing values correctly [8].

Once the imputation process has been done to fill in the missing values, the research then utilizes the robust regression analysis to establish the role of climatic variables in crop production. M-estimators (Huber, Hampel, Tukey), S-estimators and MM-estimators are some of the powerful regression methods. These methods are also less affected by the outliers than the normal least-squares regression analysis [9]. The use of robust regression analysis in the modeling of the effect of the climatic factors on the crop production in the agricultural research is especially important because the values of the outliers in the agricultural research may indicate the abnormal climatic occurrences. The other model that was used in the study to identify the multivariate outliers in the data is the Robust Mahalanobis Distance. In order to test the performance of the models further, the study employs Monte Carlo simulation to model the dataset with the controlled amount of contamination [10].

This simulation allows for a number of comparisons to be made between ordinary least squares regression analysis and robust regression. Additionally, the simulation helps in the evaluation of the efficiency of the different models in predicting the production of wheat and barley. There are four main aims of the proposed research. These aims are to: (1) analyze the effect of rainfall and climatic factors on germination and output for wheat and barley in Mosul; (2) impute missing values in the dataset utilizing the MICE approach in combination with missForest in order to guarantee the robustness of statistical analysis; (3) implement robust regression analysis side by side with traditional ordinary least squares analysis to analyze the robustness of results in the context of the presence of outliers; and (4) offer evidence-based recommendations to farmers and policymakers for agricultural development in regions of semi-arid climatic conditions. This research aims to supply a complete framework to analyze the complex relationships between climate factors and the output of cereal crops in semi-arid

regions by utilizing innovative tools for handling missing values in the dataset.

## II. RELATED WORKS

According to Salami et al. [11], agro-meteorology deals with how the various conditions of weather determine farming. Factors that are normally thought to run parallel with crops, like temperature and rainfall, become the actual driving factors in crop production and are utilized in yield forecasting. Climate change is manifested by changes in these factors of air warmth and rainfall and has lately attracted public interest because of its socio-economic consequences, especially due to erratic rainfall over the years. Droughts and floods disrupt farming systems and water supplies, influencing surface water and groundwater recharge over the areas concerned.

Weather fluctuations ripple via many sectors: heating and cooling, irrigation, and lighting for energy; livestock raising; the spread of disease vectors; flood risk; and even tourism-among others. In the Kurdistan Region of Iraq, studies have indicated that rainfall significantly varies by place and season. The northeast gets the most rain every year, with a decline as one goes south. When looking at seasons, more rainfall is usually associated with higher agricultural output, thus underlining how important rain-fed farming is there [12].

For Central Europe, the number of humid years increased during the last decades of the 20th century, but it was not uniformly distributed. There was great variability in precipitation from year to year with respect to the annual mean. In some weather stations an increased number of rainy days per year was associated with a large number of very wet days, resulting in large variability in the extremes of rainfall. On the contrary, the regions characterized by less rainfall were associated with unsteadiness in the extreme rainfall trends. Mottet et al. [13] mapped the maximum rainfall in Europe over 1901-2000 and identified regional differences in seasonality and intensity.

In Iraq, droughts have been particularly disastrous in cereal production. Ahmad et al. [14], studying wheat at the Kurdistan Region, observed a fall in wheat yield with sharp declines in rainfall, especially during growth stages. Better seed germination occurs under normal water availability and with an increase in water stress. Similarly,

Meanwhile, winter rainfall has increased across Europe in terms of average strength and in the incidence of heavy events, with summer rainfall largely unchanged. Long-term studies in the UK revealed wetter winters and drier summers during the later 20th century. Similarly, western Germany has been experiencing winters and transitional seasons with increased intensity and frequency of heavy rainfall, with relatively

similar summer rainfall extremes. Such patterns are indicative of how precipitation variability directly impacts water resources and farming.

Rainfall is one of the key elements of the hydrological cycle but it possesses immense variability in space and time, both year to year and century to century. European studies have trailed the high rainfall occurrence of various countries and also the various stations within Europe, and the entire continent, in order to get to know the regional reactions of climatic conditions. Alterations of the precipitation show the external forcing and internal climate feedbacks. The rainfall is essential to agriculture in that it provides the water to the soil and plants directly. Despite its variation, there is a tendency of rainfall patterns to take asymmetric cycles that reflect the earlier established results concerning irregular extremes in rain distribution. The exploration of the variability of rainfalls plays an important role in the management of water, agriculture and environment. An example is in Iraq where the annual rainfall has reduced drastically since 1969 with harsh droughts between the year 2000 and 2008 reducing agricultural land and vegetation cover mostly in Kurdistan. Such prolonged droughts, coupled with low precipitation, have had a severe impact on the yields of crops and food security on a regional scale.

### III. METHODOLOGY

This paper uses a systematic methodological procedure to test the relationship between precipitation and other climatic variables and wheat and barley germination in Mosul in Iraq between 2012 and 2024. Since the dataset has a limited number of annual observations (13) and many missing data, especially in the initial observations of the series, it was necessary to ensure data preparation was done carefully to prevent biased findings. To estimate the missing values, Multiple Imputation by Chained Equations (MICE) was adopted since it maintains the relationships between the climatic variables as well as offering realistic estimates that could be used to supply the time-varying agricultural data.

Once the dataset was complete, analysis was done using a mixture of ordinary least squares (OLS) and robust regression techniques, such as Huber, Tukey bisquare, Hampel, S-estimation and MM-estimation. The choice of these models was due to the presence of influential points, high variability, and multicollinearity that are easy to find in climate crop data. Model assumptions and outliers or leverage points were assessed with diagnostic plots including; residuals versus fitted values, Q-Q plots, scale-location plots, leverage analysis and others. Then Monte Carlo simulations were carried out in order to compare stability of OLS and robust estimators in scenarios of controlled contamination. This combined

methodological design will make sure that estimated impacts of precipitation on germination are real agricultural trends and not an illusion created by noise, empty values, or extreme values.

### 3.1 Data Collection & Study Variables

The research explores the effect of climatic indicators on the production of wheat and barley in the Mosul Region of Iraq from the year 2012 to 2024. The research adopted a methodology of gathering information from agricultural statistics in the country and climatic change indicators. These indicators include the total amount of rainfall in a year, average temperature, humidity, total solar radiation, wind speed and direction, and the total amount of evaporation in a year as shown in Table (1). Agricultural production was measured in metric tons for both wheat and barley.

Table 1: Dataset variables

Variable	Description	Units
Year	Calendar year of observation	Year
Wheat.production.tons	Annual wheat production	Metric tons
Barley.production.tons	Annual barley production	Metric tons
Rainfall	Total annual rainfall	mm
Average.temperature	Mean annual temperature	°C
Average.relative.humidity	Mean relative humidity	%
Total.solar.radiation	Total solar radiation	MJ/m <sup>2</sup>
Average.wind.speed	Mean wind speed	m/s
Wind.direction	Average wind direction	Degrees
Evaporation	Annual evaporation	Mm

With the overall dataset at hand, a complete analysis of the interactions between climate conditions and agriculture can be made. In general, monthly instead of annually derived values would have been propagated in order to account for long-period variation, while the effect of shorter-period variation would have had no significant influence on the overall yield.

### 3.2 Missing Data Imputation

Missing values in the dataset occur in agricultural and environmental research when the information was either not

recorded or not accurately measured. In the research, the missing values problem was solved by incorporating two techniques.

Multiple Imputation by Chained Equations (MICE) was used to carry out the prediction of values in a sequence of equations based on the conditional distribution of the set of variables. MICE help create multiple complete data sets to account for the uncertainty of the missing information. Predictive Mean Matching (PMM) was applied to numerical variables to match the distribution of the original values. PMM also avoids the prediction of implausible values. Some of the variables whose values had high missing rates or seemed to be irrelevant in the analysis would not be considered in the prediction.

To serve as a robustness test, MissForest, a non-parametric random forest approach to missing data imputation, was also used. MissForest performs particularly well when there is multicollinearity or a nonlinear association between the predictor variables. Unlike traditional regression techniques, MissForest relies on a combination of ensemble decision trees to make predictions about the missing information without assuming anything about the independence of the predictor variables. All the techniques mentioned above have the effect of creating a complete and correct dataset for further analysis.

### 3.3 Data Diagnostics

Prior to the modeling process, the dataset was reviewed for possible problems that may have resulted in biased analysis.

1. Detection of Multicollinearity: Correlation matrices for all numeric variables in the dataset have been computed to look for the presence of highly correlated predictors. Highly correlated predictors may result in higher standard errors for regression coefficients. Such highly correlated predictors may need to be removed or consolidated.
2. Detection of Outliers: Multivariate outliers in the dataset have been identified by robust Mahala Nobis distance. These points take into consideration the multivariate distribution of all the predictors. Data points that stray a long way from the multivariate center have been identified for further analysis. The need to explore these points cannot be underestimated since they may be contributors to distorted estimations in regression analysis.

Table 2: Diagnostics parameters

Diagnostic	Input Data	Method	Purpose
Multicollinearity	Numeric variables	Pearson correlation matrix	Identify highly correlated predictors
Outlier detection	Predictor matrix X	Robust Mahala Nobis distance	Identify extreme observations in multidimensional space
Outlier threshold	$RD > \text{mean} + 2 \times SD$	Empirical	Flag potential influential points

### 3.4 Regression Analysis

To analyze the connection between climatic variables and the production of crops, robust regression analysis was used. The criterion variables were the annual production of wheat and barley. The predictor variables were rainfall, temperature, humidity, solar radiation, wind speed, wind direction, and evaporations. Ordinary Least Squares (OLS) regression was taken as a benchmark model. But the problem with OLS regression is that it is highly affected by the presence of outliers in the dataset. Agricultural datasets are more prone to have outliers in them. To overcome such problems, the following robust regression techniques:

1. M-estimators (Huber, Hampel, Tukey/Bisquare): These techniques down weight the influence of unusual residuals by using weighting functions. They permit the modeling of the central trend in a dataset.
2. S-estimators: These offer highly robust estimates for parameters that are very resistant to outliers.
3. MM-estimators: These have a high breakdown point, similar to S-estimators, but higher efficiency. They offer a combination of robustness and statistical efficiency.

These techniques together allow for the overall assessment of the effect of climatic factors on crop output while incorporating the consideration of the influence of the presence of outliers in the dataset. Table (3) shows Regression models and parameters.

Table 3: Regression models and parameters

Model	Method	Robust Function
OLS	lm()	None
M-estimator	rlm()	Huber psi
M-estimator	rlm()	Hampel psi
M-estimator	rlm()	Tukey/Bisquare psi
S-estimator	lmrob()	Method="S"
MM-estimator	lmrob()	Method="MM"

### 3.5 Monte Carlo simulation for model validation

To assess the quality of regression models in contaminated environments, a simulation experiment was conducted. A series of test sets was artificially created to resemble real-world variation. Randomly distributed points representing contaminated values in the target variable were also taken into consideration. On these test sets, regression models were applied in order to calculate their quality in terms of the Root Mean Square Error. The simulation enabled a comparison between the results of both the Ordinary Least Squares Regression technique and the robust regression techniques in a controlled contamination experiment. The outcome of the simulation gave a clear insight into the robustness of the modeling techniques in their abilities to deal with anomalous values in the dataset. Table (4) indicates Monte Carlo simulation parameters.

Table 4: Monte Carlo simulation parameters

Parameter	Value	Description
Number of simulations (num_sim)	100	Repetitions for averaging RMSE
Sample size (n)	50	Number of rows per simulation
Number of predictors (p)	7	Climate variables used as independent variables
Outlier contamination rate	0.1 (10%)	Fraction of response values artificially perturbed
Response generation	Linear combination of X + noise	Beta coefficients randomly sampled 1-5
Performance metric	RMSE	Root mean squared error

### 3.6 Practical Considerations

There are a number of practical considerations in terms of methodology. Firstly, the approach of multiple imputation together with robust regression makes sure that the analysis performed under the code is robust to missing values and also to the effect of points of high leverage in the dataset. There may be some advantage in the case of missing values in a

sequence in a time series. Additionally, there may be problems of multicollinearity. Finally, the methodology contains a reproducible framework for the analysis of the impacts of climate on crop productivity. The combination of the processes of imputation, robust estimation, detection of outliers, and simulation provides a solid framework for drawing inferences that can give results in terms of the main climatic factors affecting crop productivity in the region.

## IV. RESULTS AND DISCUSSIONS

The dataset used in the analysis was the annual crop production and corresponding climatic factors in the Kurdistan Region of Iraq for the years 2012-2024. They're existed values that were missing in the dataset, specifically more values at the start of the dataset. Methods of replacing the values in the dataset without the need for analysis would either ignore the trend or underestimate the variance between the climatic factors affecting the agriculture in the region. To account for all the factors in accurately modeling the influence of climatic factors on agricultural productivity in the region without compromising the values of the dataset or changing the nature of the original dataset, the approach of \*Multiple Imputation by Chained Equations\* was adopted. This was in combination with the \*Predictive Mean Matching\* approach for numeric variables to produce values close to reality while retaining the variance in the relationships of the crop production factors. The complete dataset was comprised of the production of wheat, barley crops, the amount of rainfall, temperatures, humidity levels, solar radiation values, wind speed direction, and rates of evaporation.

### 4.1 Wheat Production Least Squares Method (OLS)

The diagnostic plots in Figure (1) of the OLS model show many deviations from the basic assumptions. The Residuals vs. Fitted plot presents the case of non-constant variability with strong nonlinear relationships. Also, the points of high influence, including points 3-11, have high values of residuals. The Q-Q plot presents a significant deviation in the residuals from the normal distribution, including the tails. Their deviations also establish the existence of strong points of high influence or strong outliers. Additionally, the Scale-Location plot presents a visual representation of the instability in the variance of the residuals for different values. Finally, the Residuals vs. Leverage points have many points close to the boundaries of the CDeven indices of high influence. The OLS technique fails to perform efficiently in the given dataset.

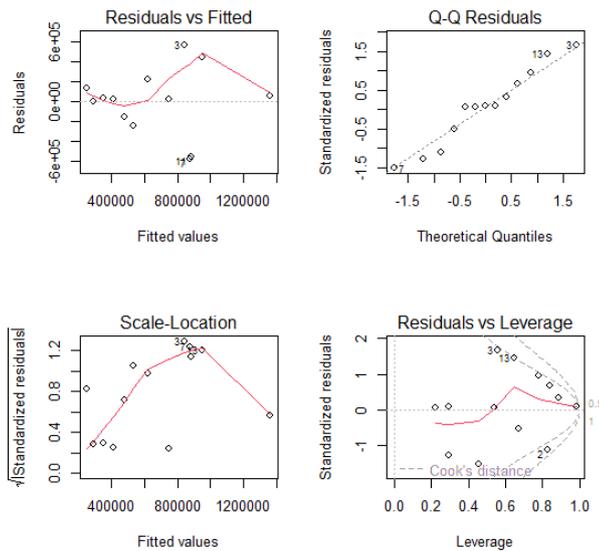


Figure 1: Wheat Production Least Squares Method (OLS)

#### 4.2 Wheat Production Huber M-Estimator Diagnostic

The Huber M-estimator makes the model more robust compared to the ordinary least squares approach. This has been shown in the diagnostic plots in Figure (2). Looking at the Residuals vs. Fitted plot, the values look more robust compared to the previous model. There seems to be less effect of extreme values. In the Q-Q plot also, the deviation seems to have reduced compared to the previous figure. But there seems to be some effect of extreme values. In the Scale-Locations plot, there seems to be more uniformity in the values of the standardized residuals. There seems to be less effect of heteroscedasticity compared to the ordinary least squares method. Additionally, in the Residuals vs. Leverage plot, the effect of leverage seems to have been reduced compared to the previous model.

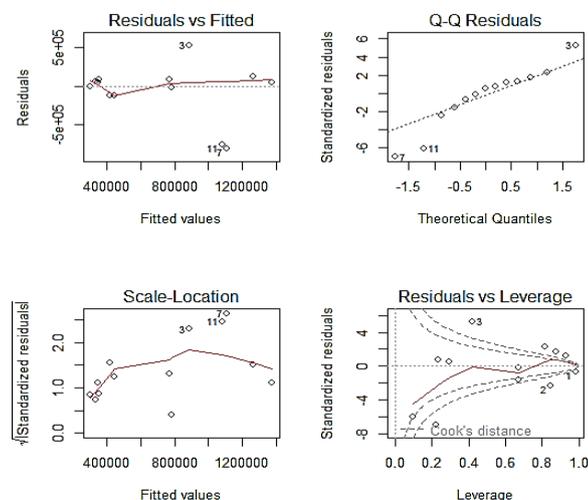


Figure 2: Wheat Production Huber M-Estimator Diagnostic

#### 4.3 Wheat Production Hampel M-Estimator

The Hampel M-estimator performs the best in terms of robustness in the Hampel family of estimators. There is a significant lack of the influence of points of high leverage in the Residual vs. Leverage plot. There also seems to be a lack of points in the Residual vs. Leverage plot around critical values of the Cook's distances. There seems to be a significant improvement in the Residual vs. Fitted plot. The points in the Q-Q plot have improved significantly. There seems to be homogeneity in the Scale-Location plot. There seems to be a lack of influence of points of high leverage in the Residual vs. Leverage plot. The Hampel M-estimator seems to offer a robust fit as indicated in Figure (3).

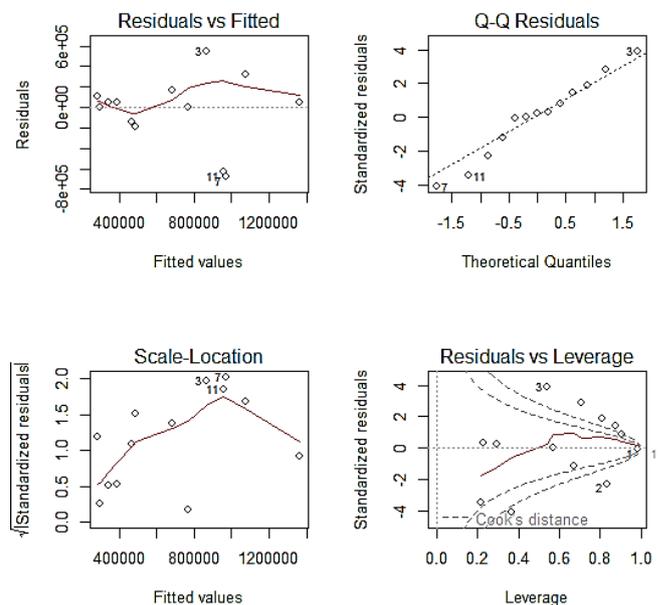


Figure 3: Wheat Production Hampel M-Estimator

#### 4.4 Barley Production Ordinary least squares (OLS)

The diagnostic plots of the Least Squares regression in Figure (4) show many irregularities in the classic assumptions of the Linear Model. In the Residual vs Fitted plot, the strong curvature of the residuals together with the uneven spread of residuals express the lack of homoscedasticity in the residuals. In the Residual vs Fitted plot, the points representing residuals 17 & 30 show strong deviation from the rest of the points.

Additionally, the graphical illustration in the Q-Q plot clearly demonstrates the lack of normality in the residuals of the model. Finally, the Scale-Location plot further develops the illustration of the homoscedasticity displayed in the residuals of the model. According to the graphical analysis in the Residual vs Leverage plot of the model, points 13 & 30 show strong deviations in their Leverage values when approaching the points of the graphical illustration for the

corresponding Cook's distances. In total, the four graphical illustrations reveal many characteristics of non-normality, lack of homoscedasticity in the residuals of the model. The graphical illustrations also indicate the points of high influence in the OLS analysis.

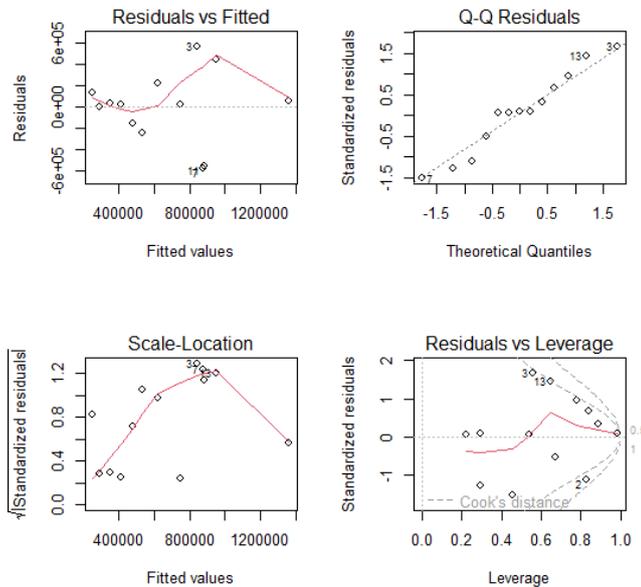


Figure 4: Barley Production Ordinary least squares (OLS)

#### 4.5 Barley Production Huber M-Estimates

The diagnostic plots for the Huber M-estimator in Figure (5) indicate a significant enhancement in the modeling pattern compared to the OLS approach but with some remaining problems. In the Residuals vs. Fitted plot, the values of the residuals become more constant, thereby reducing their curviness in the plot. Also, the Residual vs. Fitted plot indicates a significant effect in reducing the existing non-linearity in the residual values. The influence of points 3 and 11 has also reduced.

The Residual vs. Fitted plot of the Huber M-estimator also suggests a significant smoothing effect in the distribution of the residuals. The Quantile-Quantile Probability plot for the Huber M-estimator suggests a significant enhancement in the compliance of the residuals to the normal distribution compared to the OLS approach. Though there lies some deviation in the tail of the theoretical values of the Quantile-Quantile Probability plot for the Huber M-estimator. The Scale-Location plot of the Huber M-estimator helps to determine the effect of the robust approach in reducing the heteroscedasticity.

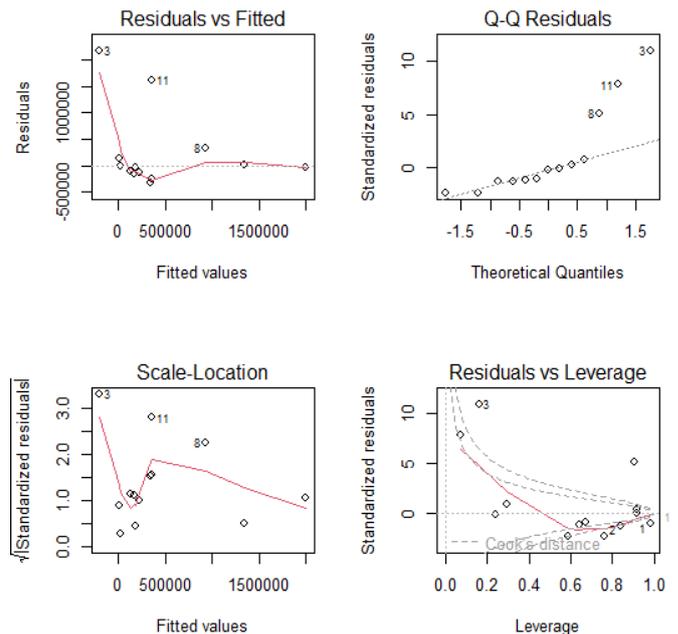


Figure 5: Barley Production Huber M-Estimates

#### 4.6 Discussion of Results

This study compared annual output of wheat and barley during the years 2012 to 2024 and found that there are significant methodological and data challenges. The first is the unusually small sample size,  $n = 13$ , which will compare to the seven climate variables that will render making sound statistical inferences quite challenging and expose the overfitting risk. Multicollinearity is evident between climate variables such as temperature, solar radiation, relative humidity, evaporation, and rainfall and hides the effect of each of the variables individually.

In the case of wheat, the ordinary least squares gave an  $R^2$  of approximately 0.479, but the adjusted  $R^2$  was negative at approximately -0.25 and this actually indicates that we actually have more predictors than we have data to argue with and some of the coefficients might be accidental. The standard error of the residual was very high with the value of approximately 515,500 that revealed that there was a significant unexplained variation with the production and poor predictability of this OLS model. Stronger regression methods that utilize M-estimators such as Huber, Tukey, Hampel have more stable coefficients and reduced residuals. Indicatively, Tukey bisquare minimized the residual standard error to approximately 104,000, thereby indicating that such methods have the ability to tame outliers. Nonetheless, S and MM estimators presented messages of exact fit with zero standard errors indicating poor ability of their estimates due to the extremely small sample size and high collinearity. Five wheat data- specifically, data numbers 1, 3, 6, 11, and 12- seem to be a very strong outlier, and should be checked against entry

errors or odd circumstances; such as an unusual weather pattern or agriculture.

The same trends in barley indicated that the OLS provided a large  $R^2$  of nearly 0.442 and a huge standard error of the residue of nearly 991,300. Whereas some of the variables appeared to be significantly different on the basis of some M-estimators, the results were found to be extremely dependent on the weighting function selected, which highlights the precariousness of the results. There were also exact-fit issues and outliers reported by MM and S estimators, suggesting that either the sample would have to increase, or the number of predictors would have to be decreased to make reliable inference. Some of the variables have a high correlation as indicated by the correlation matrix like in the example of wind direction and wheat yield (-0.792) which may cause multi-collinearity. Mighty Mahalanobis distance calculations led to negative values because of the peculiarities of computation or transformation, and standardization is required. The behavior of robust methods in simulation settings was expected based on Monte Carlo simulations to be most consistent under Y-contamination but not under X-contamination, suggesting that misconfiguration of LTS might occur. Generally, the findings are leaning towards the sensibility of the model conclusions with respect to the nature of data and value used in the application of sound techniques in agricultural climatic studies.

## V. CONCLUSIONS

The take-home is that the small sample size, when compared to the number of predictors, seriously reduces the reliability of statistical inferences, especially for standard regression. Robust regression methods like M-estimators reduce outlier influence and result in more stable residuals, but their results depend strongly on the choice of the weighting function. Other methods, such as S and MM, may be misleading for a small sample or when multicollinearity is high. Due to the strong correlations among climate variables, reduction of dimensionality is required, for instance, by PCA, in order to form composite climate indices. Verification of the outliers flagged by these methods is important because they could be data entry errors, or show unusual environmental events or other exogenous factors. Simulations illustrated that robust methods generally outperform OLS when facing Y-contamination; however, with respect to LTS, its performance was not always reliable, so setting parameters might be relevant.

## REFERENCES

[1] W. Husain and A. Al-Heali, "CHALLENGES FACING FOOD SECURITY IN IRAQ, CEREAL CROPS AS MODEL," *Plant Arch.*, pp. 5130–5134,

Jan. 2020.

- [2] A.T. Muhammad and A. S. Shukr, "The economic impact of climate and environmental changes on wheat and rice crops in Iraq for the period 2000-2023 using the Ricardo approach .," *Euphrates J. Agric. Sci.*, vol. 17, no. 2, pp. 161–170, 2025.
- [3] S. M. N. Manik, G. Pengilly, G. Dean, B. Field, S. Shabala, and M. Zhou, "Soil and Crop Management Practices to Minimize the Impact of Waterlogging on Crop Productivity," *Front. Plant Sci.*, vol. 10, p. 140, Feb. 2019.
- [4] K. Muthiah, K. G. Arunya, V. Sridhar, and S. K. Patakamuri, "Heavy Rainfall Impact on Agriculture: Crop Risk Assessment with Farmer Participation in the Paravanar Coastal River Basin," *Water*, vol. 17, no. 5, p. 658, 2025.
- [5] Y.-U. Kim *et al.*, "Mechanisms and modelling approaches for excessive rainfall stress on cereals: Waterlogging, submergence, lodging, pests and diseases," *Agric. For. Meteorol.*, vol. 344, p. 109819, Jan. 2024.
- [6] Z. A. Ismaeel and T. T. Yousif, "Development of AI Models for Predicting Dispersion and Confinement Loss in PCFs," *Int. J. Microw. Opt. Technol.*, vol. 20, no. 6, pp. 648–658, 2025.
- [7] O. R. Olaniran and A. R. R. Alzahrani, "Bayesian Random Forest with Multiple Imputation by Chain Equations for High-Dimensional Missing Data: A Simulation Study," *Mathematics*, vol. 13, no. 6, p. 956, 2025.
- [8] L. O. Joel, "On the Performance of Imputation Techniques for Missing Values on Healthcare Datasets arXiv : 2403 . 14687v1 [ cs . LG ] 13 Mar 2024," *arXiv:2403.14687v1*, pp. 1–20, 2024.
- [9] U. Shahzad, N. Al-Noor, M. Hanif, I. Sajjad, and M. M. Anas, "Imputation based mean estimators in case of missing data utilizing robust regression and variance-covariance matrices," *Commun. Stat. Simul. Comput.*, Mar. 2020.
- [10] M. R and R. Myilsamy, "M-Estimators in Regression Models," *J. Math. Res.*, vol. 2, pp. 23–27, Oct. 2010.
- [11] Z. Brixiova Schwidrowski, "Smallholder Agriculture in East Africa: Trends, Constraints and Opportunities," Jan. 2010.
- [12] M. Aljaff, "Trend and variability of rainfall in recent years in Iraq: A physical perspective," vol. 2, pp. 1118–1128, Feb. 2025.
- [13] A.Mottet, S. Ladet, N. Coqué, and A. Gibon, "Agricultural land-use change and its drivers in mountain landscapes: A case study in the Pyrenees," *Agric. Ecosyst. Environ.*, vol. 114, no. 2, pp. 296–310, 2006.

- [14] N. Ahmad, K. Mustafa, and D. Ahmad, "Early Screening of Some Kurdistan Wheat (*Triticum aestivum* L.) Cultivars under Drought Stress," *J. Agric. Sci.*, vol. 9, p. 88, Jan. 2017.

**Citation of this Article:**

Lamia Khalil Ibrahim, & Bashar A. Al-Talib. (2025). Assessing the Role of Precipitation in Wheat and Barley Germination in Mosul, Iraq. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 9(12), 131-139. Article DOI <https://doi.org/10.47001/IRJIET/2025.912021>

\*\*\*\*\*