Hindawi
Advances in Materials Science and Engineering
Volume 2023, Article ID 9759731, 1 page
https://doi.org/10.1155/2023/9759731

Retraction

Retracted: Application of Regression Analysis to Identify the Soil and Other Factors Affecting the Wheat Yield

Advances in Materials Science and Engineering

Received 26 December 2023; Accepted 26 December 2023; Published 29 December 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] A. Hayat, M. Amin, S. Afzal, A. H. Muse, O. M. Egeh, and H. S. Hayat, "Application of Regression Analysis to Identify the Soil and Other Factors Affecting the Wheat Yield," *Advances in Materials Science and Engineering*, vol. 2022, Article ID 7793187, 10 pages, 2022.

Hindawi
Advances in Materials Science and Engineering
Volume 2022, Article ID 7793187, 10 pages
https://doi.org/10.1155/2022/7793187

Research Article

Application of Regression Analysis to Identify the Soil and Other Factors Affecting the Wheat Yield

Azhar Hayat[1] **Muhammad Amin**[2] **Saima Afzal**[1] **Abdisalam Hassan Muse**[3,4] **Omer Mohamed Egeh**[4] **and Hafiz Saqib Hayat**[5]

[1]Department of Statistics, Bahauddin Zakariya University, Multan, Pakistan
[2]Department of Statistics, University of Sargodha, Sargodha, Pakistan
[3]Department of Statistics, Pan African University, Institute for Basic Sciences,Technology and Innovation, Nairobi, Kenya
[4]Department of Mathematics and Statistics, Amoud University, Borama, Somalia
[5]Department of Agronomy, Bahauddin Zakariya University, Multan, Pakistan

Correspondence should be addressed to Omer Mohamed Egeh; 1933@amoud.edu.so

Received 21 February 2022; Revised 12 April 2022; Accepted 19 April 2022; Published 10 May 2022

Academic Editor: Palanivel Velmurugan

In farming and related fields, numerous connections exist that should be distinguished quantitatively. Several factors affect the various crop yields in different dimensions. These factors may have relation with farmer's practices or with quality of soil. In this study, our main focus is to explore the effect of soil and other factors on the wheat yield. Regression modeling plays an important role in the identification of such factors that greatly affect the crops yield. For reliable and valid results, one has to check the data for outliers and other critical results. In this study, we have fitted the regression models with and without satisfying some regression assumptions to determine the factors affecting yield of wheat. For analysis purposes, the required data were collected from the district Multan. It was observed that when the regression assumptions were satisfied, then coefficient of determination (R^2) was improved from 45% to 48%, R^2 (adjusted) was improved from 40% to 46%, and the standard error of the estimates was reduced from 2.772 to 2.649. These results indicate that the soil characteristics, such as saturation, electrical conductivity, organic matter, phosphorus, potassium, calcium carbonate, and micronutrients (zinc, copper, iron, manganese, and boron), are the significant factors for wheat yield. While among all other factors, urea, chemical coating of seed, use of compost, and previously sown crops are the significant factors for wheat yield.

1. Introduction

The agriculture sector plays a key role in the economy of Pakistan as it contributes 19.2% to gross domestic product (GDP) and 38.5% of the labor force is engaged in this sector [1]. It also has a major contribution to the foreign exchange earnings of the country and the development of other sectors. However, nowadays, Pakistan is facing the issues of water stress, pollution, and environmental degradation along with institutional and socio-economic problems; these issues have severely affected the agricultural productivity of the country. Consequently, there is a need for strong strategies and their implementation to improve productivity in the agriculture sector [2].

Wheat is the main food crop of Pakistan, and it dominates all crops in production. According to world population review (2022), Pakistan is ranked at seventh place among the highest wheat-producing countries in the world in terms of yield per acre, area, and production. The mean per capita consumption of wheat in Pakistan is about 125 kg per year and is about 60% of daily food for an average person [3].

The earnings of almost 80% of the farmers in Pakistan are dependent on wheat. It is grown at an area of 9 million hectares that is roughly 40% of the total cultivated land in the country. The wheat crop added about 8.7% to the total value of agriculture and contributes about 2.1% to the GDP [4]. Wheat is considered a staple crop for the countries in temperate zones and is increasing in demand in the countries that are undergoing industrial and urban development. It is the main source of starch and energy, further it provides essential nutrients like vitamins, dietary fiber, and proteins, which are necessary for human growth [5]. That is why most researchers are strenuous to know factors that increase and decrease wheat production due to its high importance.

Worldwide, several analysts examined the impact of various factors on the yield of wheat. Barzegar et al. [6] investigated the impact of wheat straw, composted sugarcane residue on properties, and yield of wheat in Iran. They applied the split-plot design and found that organic materials increased wheat yield, aggregate stability, water retention, and decrease soil bulk density. Martre et al. [7] proposed mechanistic hypotheses for grain nitrogen accumulation in cereal. The four cultivars were tested in France. The proposed hypotheses were tested with the help of linear regression. It was found that under normal conditions, Arche and Rectal cultivars had the highest grain per year while Tamaro cultivar had the lowest grain per year. Plaut et al. [8] conducted experiments on Suneca and Batavia wheat varieties in Australia. A mechanistic model was developed and showed that kernel number was not affected by water deficit, but high temperature caused a significant reduction in both varieties. Water deficit and the high temperature were also found to increase the relative contribution of transported dry matter to kernels. Leilah and Al-Khateeb [9] explored the association between wheat yield and its components under drought conditions in Saudi Arabia. Multiple statistical techniques including correlation, linear regression, path analysis, factor analysis, cluster analysis, and principal components analysis were used and it was revealed that weight of grain, number of spikes per square meter, and biological yield had the most influence on wheat grain yield indicating that breeding material with these qualities can produce a high yield in drought conditions. Moriondo et al. [10] adapted the use of Normalized Difference Vegetation Index (NDVI) data obtained from the simulation model and satellite platforms for the estimation of wheat yield. This methodology was applied in two high wheat-yielding provinces of Italy. They found that the proposed methodology produced a better accuracy in the estimation of wheat yield. Khan et al. [11] studied the effect of zinc fertilizer at different levels on the yield of wheat in Multan, Pakistan. The difference among these treatments was compared using the least significant differences (LSD), polynomial curve fitting, and coefficient of determination, and it was found that application of 5 kg ha^{-2} of zinc sulphate gave the highest marginal rate of return as compared to other amounts of application of zinc sulphate. Whalley et al. [12] discussed the relation between soil strength and yield of wheat for two different soil types in England. They applied the factorial design and found that soil strength is a good predictor of crop yield irrespective of soil type, water status, and concluded that soil strength seemed to limit crop productivity. Abbas et al. [13] evaluated the influence of trace elements in nitrogen, phosphorus, and potassium absorption on wheat yield. With the help of randomized complete block design (RCBD) and Duncan's multiple range test, they showed that limited use of iron increased wheat yield. If the application of iron was increased, the effect on the yield was insignificant or undesirable. Hassan et al. [14] conducted a study to find significant factors that were influencing the wheat yield in mixed cropping zones of Punjab. The data were collected from four different districts of Punjab. They applied the linear regression and found that seed rate, cost weedicide that was used, education of farmer, use of nitrogenous fertilizer, rotavator use, and sowing time are the significant factors for wheat yield. Gul et al. [15] explored the effect of foliar application of micronutrients on Ghazanive-98 variety of wheat in Peshawar, Pakistan. They applied the RCBD and found that foliar treatment of micronutrients affected the growth of the wheat variety without causing any effect on the time of growth.

Nadim et al. [16] evaluated the yield characteristics and physiology of the Gomal-8 variety of wheat for different levels of micronutrients in Dera Ismail Khan. With the help of RCBD, they found that the application of Boron produced a higher leaf area index while the application of copper produced a maximum number of tillers causing an increase in wheat yield. Rezaei and Hemati [17] studied the effect of soil properties on wheat yield in Iran. They applied the RCBD, and the results showed that a balanced percentage of sand, clay, and silt provide favorable conditions for the improvement of wheat yield. El-Lethy et al. [18] discussed the potassium impacts on wheat plants under saline conditions in Giza, Egypt. With the help of analysis of variance (ANOVA) and LSD, they concluded that the yield of wheat is decreased significantly in saline conditions, and potassium fertilizer reduces the undesirable effects of salinity. Alam and Salahin [19] conducted a series of experiments to study the effect of soil density and moisture retention on wheat yield in Bangladesh. They applied the multiple regression models and found that wheat yield almost doubled from lowest soil depth to highest soil depth. Muarya et al. [20] conducted a field experiment during the Rabi season in Kanpur to study the influence of potassium levels on wheat yield and growth. Using factorial RCBD, it was observed that the application of 80 kg ha^{-1} of potash (K_2O) produced the highest grain yield, straw yield, and biological yield compared to 0, 40, and 60 kg ha^{-1} application of K_2O. Limon-Ortega and Martinez-Cruz [21] studied the impact of nitrogen on wheat yield in Mexico. The data were analyzed through ANOVA. It was found that nitrogen sources impacted wheat yield and the number of spikes based on soil reaction while the fungicide spray had a positive influence on the wheat yield.

Ghadikolayi et al. [22] studied the influence of crop residue and nitrogen on wheat yield in Iran. They applied the RCBD and found that 135 kg/ha of nitrogen gave the highest soil organic matter (OM). It was also found that all of the residues used in the experiment reduced the yield of wheat but the reduction with the use of sunflower residue was lowest compared to other crop residue tested. Sarto et al. [23] investigated the effects of silicate application on soil chemical properties in Parana, Brazil. With the help of regression models and it was concluded that the calcium/magnesium silicate in acid clayey soil improves the yield of wheat. However, the soil with pH higher than 5.3 and high silicon does not impact the grain yield of wheat.

Arshad et al. [24] conducted an experiment in Peshawar, Pakistan, to study the interactive impact of zinc and phosphorus on wheat. With the help of LSD test, it was found that 5 kg ha^{-1} of zinc gave maximum straw yield, but 10 kg ha^{-1} zinc significantly increased all other indicators of yield. As for phosphorus, 90 kg ha^{-1} was observed to produce the best results. Mehmood et al. [25] studied the input factors that positively influence wheat production in Bahawalnagar, Punjab. The researchers used the linear regression and found that the sowing method, use of fertilizer, that is, nitrogen and phosphorus, variety of wheat, use of weedicides spray and irrigation mode have a significant effect on the wheat yield. Chairi et al. [26] conducted a study on the genetic gain in yield for durum wheat in three experiment stations in Spain. They considered the linear regression to determine variability in locations for absolute and relative genetic gain (AGG and RGG). The result of the analysis showed that the rate of genetic progress in durum wheat yield has been low in the past decade. Rajicic et al. [27] carried out tests on wheat plants in soil with low pH by applying nitrogen along with phosphorus and potassium fertilizers in Serbia. They applied correlation analysis and found that nitrogen had a significant impact on wheat yield and the treatments where the highest amount of nitrogen was applied with other combinations of phosphorus, and potassium had a high yield compared to the lower application of nitrogen. Polisetty and Paidipati [28] estimated the trends in the production of wheat using data from four states of India. The trend analysis was conducted using nonparametric methods including Pettitt's Standard Normal Homogeneity, Buishand's range test, and Mann Kendall test. The outcomes of the trend analysis showed that all states under consideration had an upwards trend and indicated an improvement in wheat production. Sial et al. [29] investigated the effect of waste-derived-biochars of milk tea and fruit peels on growth, yield, root traits, soil enzyme activity, and nutrient status of the wheat crop in Shaanxi, China. Eight treatments were analyzed with one-way ANOVA and it was found that plant height, dry weight of root and shoot, chlorophyll amount, grain yield significantly increased using the treatment of milk tea biochar and chemical fertilizer. Ashraf et al. [30] analyzed the input-output flow for wheat production to identify energy-efficient ways through data from Mailsi, Pakistan. It was found through multiple regression models that higher inputs, large fields, high fertilizer application, and tillage operation provided the highest energy outputs with high productivity and efficiency in energy. Zhou et al. [31] studied the relationship between the depth distribution of wheat roots and soil macroporosity in United Kingdom with six varieties of wheat. Two-way ANOVA was used, and it was determined that there was no significant difference in wheat genotypes, and the wheat root system was more affected by the soil macropore system. Recently, Hayat et al. [32] explored the effect of soil properties and other factors on the cotton yield of Pakistan. They observed that EC, pH, saturation, OM, P, Zn, Cu, Fe, and B are the significant factors for cotton yield. They also found that fertilizer (Nitrophos, nitrogen, and urea), previously sown crops (wheat and corn), type of seed, chemical coating of seed, type of water, way of cultivation, and use of compost are also a significant factors for cotton yield.

As we have seen in the literature that worldwide, mostly researchers studied the effect of individual soil characteristic or two or other factors on the wheat yield separately. No one still studied the joint influence of soil characteristics and other factors on the wheat yield. So, in this study, our main focus is to explore the joint influence of these factors on the wheat yield. To explore the influence of these factors on the wheat yield, we will consider the linear regression model and will identify which factors contribute a significant role in the wheat yield. Moreover, we will also evaluate some of the regression assumptions to obtain the reliable results. These assumptions include no multicollinearity, constant error variance, no autocorrelation, no outlier, and influential observation [33, 34]. In this article, we have paid special attention to the regression model diagnostics and its impact on the wheat yield model for the identification of factors. Regression diagnostics include outliers and influential observations analysis which can affect the model estimates and predicted values. With the presence of these values, the fitted model results may indicate the significance of the factors which are playing no role in the response variables and vice versa.

2. Materials and Methods

The major food producing province of Pakistan is Punjab and is titled as "bread basket" to feed more than 220 million Pakistanis. Wheat is the main cereal crop produced in Punjab.

The investigation zone is in the bond that is made by five rivers of Punjab situated at 30.157 degrees North and 71.524 degrees South with an average elevation of 122 m above ocean level. The region as a zone of about 132.1 km square is partitioned into various tehsils. The normal temperature in the area fluctuates from 20°C to 45°C with normal precipitation of 175 mm a year. This demonstrates that a particular crop cannot be planned to cultivate in the soil for the entire year in the area because climate conditions change drastically. The conditions of this zone in winter are ideal for wheat production.

The population size is 2620 acres of this area, where one acre is taken as a single sampling unit. The sample was taken as 655 and computed using appropriate formula. Systematic random sampling technique was used to collect the soil samples. The sample was collected from every fourth acre using 4 as sampling interval.

It is necessary to maintain the specific concentrations of organic and inorganic matter in the soil for a good yield of crops. Five hundred grams soil was taken from each selected acre of land to measure the variables specified in Table 1. Then these samples were sent to laboratory of the Agriculture Department of Punjab under Agri-Smart project for the computations of the soil characteristics, where they used the several apparatus. The details of apparatus to measure variables related to soil characteristics are presented in

Table 1: Apparatus for measuring the soil characteristics

Sr. No	Variable	Apparatus
1	Saturation percentage	1. Polythene sheets 2. Spade 3. Soil auger 4. Moisture boxes/cans 5. Balance 6. Oven 7. Ring stand 8. Funnel (glass or plastic) 9. Tubing (to attach to bottom of funnel) 10. Clamp (to secure tubing) 11. Filter paper (to line funnel) 12. Beakers (250 mL) 13. Graduated cylinder 14. Stirring rod (long)
2	Electrical conductivity	1. Conductivity bridge 2. Vacuum filtration system
3	Power of hydrogen	1. pH meter with combined electrode 2. Beakers: Preferably use polyethylene or TFE beakers 3. Mechanical stirrer, with inert plastic coating 4. Wash bottle, plastic
4	Organic matter	1. Magnetic stirrer and teflon-coated magnetic stirring bar 2. Glassware and pipettes for dispensing and preparing reagents 3. Titration apparatus (burette)
5	Phosphorus	1. Spectrophotometer or colorimeter 2. Standard laboratory glassware: beakers, volumetric flasks, pipettes, and funnels
6	Potassium	1. Flame photometer with accessories 2. Beakers 3. Pipettes and volumetric flasks, as required for dilution and tests of interference effects
7	Calcium carbonate	1. Hot plate 2. Burette 3. Erlenmeyer flask 4. Volumetric pipette
8	Zinc	1. Atomic absorption spectrophotometer 2. Mechanical shaker, reciprocal
9	Copper	1. Atomic absorption spectrophotometer 2. Mechanical shaker, reciprocal
10	Iron	1. Atomic absorption spectrophotometer 2. Mechanical shaker, reciprocal
11	Manganese	1. Atomic absorption spectrophotometer 2. Mechanical shaker, reciprocal
12	Boron	1. Porcelain crucibles 2. Spectrophotometer 3. Polypropylene test tubes

Table 1. The factors estimated for each sample of soil attributes incorporate OM, phosphorous (P), potassium (K), and calcium carbonate (CC). The significant micronutrients in the soil that were estimated from each sample include zinc (Zn), copper (Cu), iron (Fe), manganese (Mn), and boron (B).

The information related to the utilization of fertilizer, such as DAP, potash, urea, Nitrophos, and nitrogen manures in the land per section of land is also collected from the farmers. The data about the recurrence of pesticides and water systems and strategy for the water system, that is, tubewell water or trench water were gathered. The information about the crop sown before wheat, seed type, that is, coated or uncoated, the technique for cultivating, that is, penetrating or drilling sewing strategy, and utilization of fertilizer in the land were also collected. Further details and descriptions about these factors are given in Table 2.

3. Results

Descriptive analysis of all considered variables is given in Table 3. The descriptive statistics include average, standard deviation, coefficient of variation, maximum, minimum, range skewness, and Kurtosis. The average wheat yield in this area is found to be 41.87 mund/acre with a standard deviation of 3.66. The average saturation is 37.75% which

Table 2: Description of wheat yield and its associated factors.

Name of variable	Notation	Nature of variable	Unit of variable
Saturation (%)	—	Continuous	—
Electrical conductivity (dsm-1)	EC	Continuous	—
pH	—	Continuous	—
Organic matter (%)	OM	Continuous	—
Phosphorus	P	Continuous	Ppm
Potassium	K	Continuous	Ppm
Calcium carbonate (%)	CC	Continuous	—
Zinc	Zn	Continuous	Ppm
Copper	Cu	Continuous	Ppm
Iron	Fe	Continuous	Ppm
Manganese	Mn	Continuous	Ppm
Boron	B	Continuous	Ppm
Nitrophos	—	Discrete (frequency)	50 kg per acer
Nitrogen	—	Discrete (frequency)	50 kg per acer
DAP	—	Discrete (frequency)	50 kg per acer
Potash	—	Discrete (frequency)	50 kg per acer
Urea	—	Discrete (frequency)	50 kg per acer
Last crop	—	Discrete (frequency)	—
Water frequency	—	Discrete (frequency)	—
Pesticide frequency	—	Discrete (frequency)	—
Chemical coating of seed	—	Categorical	—
Type of water	—	Categorical	—
Way of cultivation	—	Categorical	—
Usage of compost	—	Categorical	—
Wheat yield per acre	—	Continuous	Munds

Note: Ppm = "parts per million".

Table 3: Descriptive statistics of the consider characteristics.

Variables	Average	SD	CV	Minimum	Maximum	Range	Skewness	Kurtosis
Wheat yield	41.87	3.66	0.09	29.00	50.00	21.00	−1.41	−2.00
Saturation (%)	37.75	3.82	0.10	21.00	45.00	24.00	−4.81	−1.98
EC (dsm-1)	4.24	3.97	0.94	0.20	22.00	21.80	22.18	25.53
pH	8.26	0.29	0.03	7.20	9.00	1.80	−1.13	−2.66
Org. M. (%)	0.55	0.15	0.27	0.12	1.09	0.97	3.73	1.51
Phosphorus (ppm)	8.20	3.26	0.40	1.40	23.50	22.10	13.90	14.43
Potassium (ppm)	176.10	74.81	0.42	78	380.00	302	1.081	−0.058
CC (%)	6.36	1.43	0.23	2.80	13.20	10.40	5.37	6.84
Zinc (ppm)	0.78	0.22	0.28	0.10	1.00	0.90	−12.34	0.86
Copper (ppm)	0.18	0.02	0.11	0.05	0.32	0.27	−10.10	59.30
Iron (ppm)	4.08	0.68	0.17	0.42	4.93	4.51	−25.64	36.10
Manganese(ppm)	0.89	0.15	0.17	0.07	1.87	1.80	−20.26	51.24
Boron (ppm)	0.46	0.01	0.03	0.39	0.49	0.10	−6.93	2.83
Nitrophos	0.41	0.58	1.41	0.00	2.00	2.00	11.16	0.78
Nitrogen	0.51	0.71	1.40	0.00	3.00	3.00	11.97	1.68
DAP	1.01	0.50	0.50	0.00	3.00	3.00	1.65	8.23
Potash	0.31	0.49	1.60	0.00	2.00	2.00	12.42	1.08
Urea	1.92	0.68	0.35	0.00	4.00	4.00	0.14	3.37
Water frequency	3.40	0.70	0.21	2.00	5.00	3.00	1.61	−0.81
Pesticide frequency	1.93	0.82	0.43	1.00	6.00	5.00	7.32	2.54

means that the soil of this area is loam which indicated that this soil is suitable for wheat. The average value of EC is 4.24 dS/m depicting that there is slight saline soil which is a favorite for wheat. The average pH of the soil is 8.26 indicating that it is difficult for the plant to obtain phosphorus from the soil. The average amount of OM in the soil is 0.55%, which shows that there is a lack of organic components in the soil. The average amount of phosphorus in the soil is 8.20 ppm which is normal. The average amount of potassium in the soil is 176.10 ppm which is good for the soil to be called fertile. The average zinc is about 0.78 ppm in the soil indicating an intermediate amount of the nutrient. The average copper in the soil was found to be 0.18 ppm, which is a satisfactory amount for the fertility of the soil. The average iron in the soil is 4.08. This indicates the soil's high fertility. The manganese in the soil on average is 0.89 ppm which is a

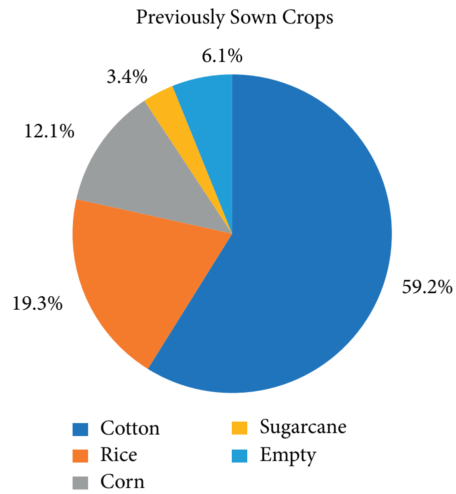

Figure 1: Pie chart for previously sown crops.

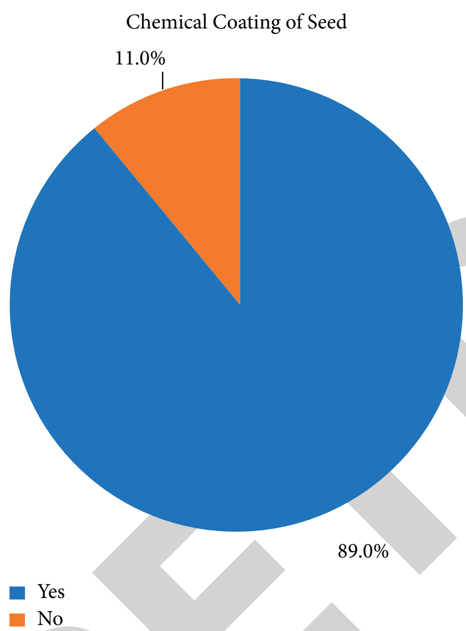

Figure 2: Pie chart for chemical coating of a seed.

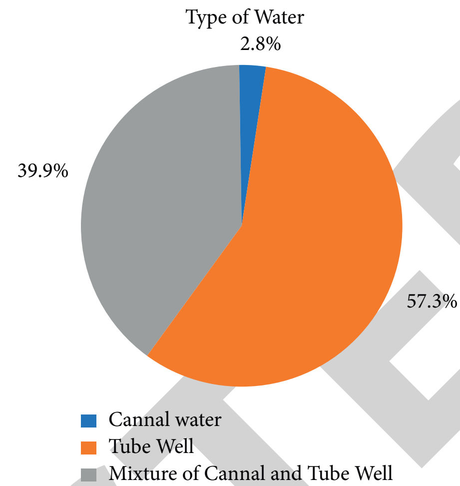

Figure 3: Pie chart for the type of water.

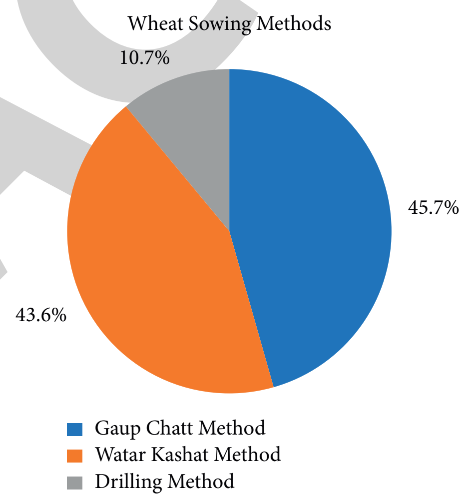

Figure 4: Pie chart for wheat sowing methods.

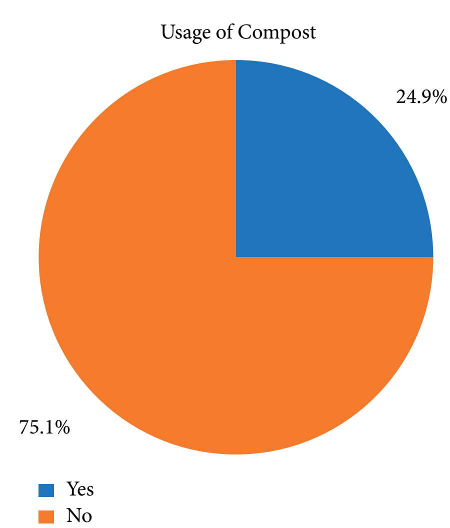

Figure 5: Pie chart for using compost.

good amount for the fertility of the soil. Finally, the average amount of boron in the soil is 0.46 ppm indicating enough nutrients in the soil. The skewness of most data distribution was positive, and the highest values of the skewness coefficients were EC and phosphorus. Visconti et al. [33] studying soil saturation extracts in Spain, also found higher positive skewness for the soil potassium and attributed this to the fact that fertilizers may have been applied at a higher concentration at some locations. Negative skewness was observed in some factors such as iron (ppm), manganese (ppm), and zinc (ppm).

The summary of some qualitative factors is shown in Figures 1 to 5. Figure 1 shows that cotton was the most common previously sown crop as compared to other crops. From Figure 2, we have found that most of the seeds were chemically coated. Figure 3 displays that most farmers use

Table 4: Regression analysis with and without outlier.

Term	Full data				After deleting outliers			
	Beta	SE	*T*	*P* value	Beta	SE	*T*	*P* value
Constant	69.350	6.2800	11.04	0.0000	68.420	6.0800	11.26	0.0000
Saturation (%)	−0.154	0.0335	−4.59	0.0000	−0.145	0.0324	−4.46	0.0000
EC (dsm-1)	0.072	0.0297	2.44	0.0150	0.080	0.0288	2.76	0.0060
pH	2.345	0.4440	5.28	0.0000	2.340	0.4330	5.41	0.0000
Org. M. (%)	1.454	0.7920	1.83	0.0670	1.862	0.7730	2.41	0.0160
Phosphorus (ppm)	−0.263	0.0388	−6.78	0.0000	−0.285	0.0375	−7.6	0.0000
Potassium (ppm)	−0.010	0.0019	−5.18	0.0000	−0.008	0.0019	−4.33	0.0000
CC (%)	−0.494	0.0829	−5.96	0.0000	−0.558	0.0804	−6.93	0.0000
Zinc (ppm)	−4.740	0.5950	−7.97	0.0000	−4.892	0.5810	−8.43	0.0000
Copper (ppm)	−10.740	5.7700	−1.86	0.0630	−11.500	5.6600	−2.03	0.0430
Iron (ppm)	−0.337	0.1670	−2.02	0.0440	−0.357	0.1610	−2.22	0.0270
Manganese (ppm)	−2.137	0.7890	−2.71	0.0070	−1.936	0.7580	−2.55	0.0110
Boron (ppm)	−55.210	9.2800	−5.95	0.0000	−52.210	8.9500	−5.83	0.0000
Nitrophos	0.092	0.2110	0.44	0.6610	0.001	0.2020	0.01	0.9950
Nitrogen	0.140	0.1690	0.83	0.4070	0.069	0.1640	0.42	0.6730
DAP	−0.117	0.2430	−0.48	0.6310	−0.188	0.2360	−0.8	0.4250
Potash	−0.052	0.2360	−0.22	0.8270	0.036	0.2300	0.16	0.8750
Urea	0.298	0.1750	1.70	0.0900	0.525	0.1970	2.66	0.0080
Water frequency	−0.171	0.1680	−1.01	0.3110	−0.277	0.1630	−1.7	0.0900
Pesticide frequency	0.139	0.1360	1.02	0.3080	0.279	0.1430	1.96	0.0510
Last crop								
Rice	0.785	0.2940	2.67	0.0080	0.713	0.2840	2.51	0.0120
Corn	0.299	0.3680	0.81	0.4170	0.493	0.3590	1.37	0.1700
Sugarcane	0.840	0.6280	1.34	0.1820	0.791	0.6010	1.32	0.1890
Empty	0.591	0.4760	1.24	0.2150	0.542	0.4560	1.19	0.2350
Seed chemical coating								
No	−1.031	0.3700	−2.79	0.0050	−0.959	0.3580	−2.68	0.0080
Water type								
Tube well	−0.172	0.6960	−0.25	0.8040	−0.054	0.6660	−0.08	0.9360
Both (tube well and canal water)	−0.350	0.7000	−0.50	0.6170	−0.229	0.6690	−0.34	0.7320
Way of cultivation								
Water kashat method	0.009	0.2410	0.04	0.9700	0.003	0.2340	0.01	0.9900
Drilling method	0.369	0.3940	0.94	0.3490	0.308	0.3830	0.8	0.4230
Compost use								
No	−1.218	0.2570	−4.74	0.0000	−1.167	0.2470	−4.72	0.0000
F (P value) = 17.75 (0.0000), S = 2.772, R^2 = 45.24%, R^2(adj) = 40%					F (P value) = 19.59 (0.0000), S = 2.649, R^2 = 48.38%, R^2 (adj) = 45.91%			

tube-well water in their wheat fields. Figure 4 indicates that most of the farmers use Gaup Chatt sowing method to grow wheat in this area. From Figure 5, we have found that most farmers did not use compost in their fields before sowing wheat.

After the analysis of descriptive statistics, the next step was to identify the contribution of significant variables by fitting a linear regression model to the data of the soil chemical and other characteristics which influence the wheat yield. The results obtained from the linear regression model on full data showed that R^2 of the fitted model explained 45.24% of the variability in wheat yield per acre due to the considered soil and other characteristics. The results of the F-test given in Table 4 show that considered variables contribute a 45.24% role in the wheat yield per acre. The adjusted R-squared statistic, which is a more suitable safeguard, used for not overfitting of the regression model with multiple factors, was 40%. The role of the individual factors in the wheat yield was also explored. To determine the contribution of variables and identification of most significant variables, we use the t-test, and the results are reported in Table 4. It can be seen that saturation (%), EC, pH, phosphorus, potassium, CC, zinc, iron, manganese, and boron were found to be statistically significant factors at a 5% level of significance. Of these factors, zinc was found to be the most significant factor for wheat yield (see t values and its associated p values). The t value on absolute for this factor is larger than the t values of the other considered factors, and it is already indicated in the literature that zinc contributes a significant role in various crop yields [35–38]. The second-most important factor is soil phosphorous, which contributes a significant role to increase or decrease the wheat yield per acre. The factor previously sown crops (rice), chemical coating of seed (no), and use of compost (no) are statistically significant factors that are increasing the wheat yield. The nutrients organic matter and copper respectively are not play

Table 5: Residual analysis for the regression based on full data.

Observation number	Wheat yield	Fit	Resid	Std resid
27	38	43.32	−5.32	−2.01
82	36	41.308	−5.308	−2.02
135	38	43.36	−5.36	−2.05
153	35	40.494	−5.494	−2.06
213	47	41.73	5.27	2.01
302	47	41.454	5.546	2.03
315	37	42.057	−5.057	−2.02
319	36	41.725	−5.725	−2.11
366	35	40.427	−5.427	−2.01
368	38	43.278	−5.278	−2.01
370	38	43.521	−5.521	−2.05
388	36	41.558	−5.558	−2.09
477	50	44.39	5.61	2.07
506	37	42.438	−5.438	−2.0
517	35	40.519	−5.519	−2.03
533	38	43.473	−5.473	−2.08
614	38	43.517	−5.517	−2.05
617	36	41.291	−5.291	−2.03

a significant role in the growth of wheat yield. As well as the impact of some fertilizers namely Nitrophos, nitrogen, DAP, potash, and urea are not beneficial to some extent as compared to the factors discussed earlier.

The regression results are reliable if there, is no outlier in the residual. Checking the presence and identification of outliers is the next step. For this purpose, we used standardized residuals. The evaluation of these residuals is given in Table 5. Table 5 indicates that the observation number 27, 82, 135, 153, 213, 302, 315, 319, 366, 368, 370, 388, 477, 506, 517, 533, 614, 617 are identified as outlier.

After identification of these outliers, the output showed the results of a linear regression model to describe the relationship between wheat yield per acre and 13 independents (qualitative and quantitative) variables. Since the p value of the F-test in Table 4 is less than 0.05, there was a statistically significant relationship between the wheat yield and associated factors with the exclusion of substantial values observations. The R^2 indicated that the model as fitted explained 48.38% of the variability in wheat yield per acre. The adjusted R-squared statistic, which was more suitable for comparing models with different numbers of independent variables, was 45.91%. The standard error of the estimate showed the standard deviation of the residuals to be 2.649 which was smaller as compared to model fitted variability with full data. Now we explore the role of the individual factors in the wheat yield. It can be seen from Table 4 that saturation (%), EC, pH, organic matter, phosphorus, potassium, CC, zinc, copper, iron, manganese, boron, and urea were found to be statistically significant factors at a 5% level of significance. One more thing is noted, the role of organic matter, copper, and urea was hidden in the full data set due to the outliers, but after removing the outliers, they are statistically contributing a significant role in the wheat yield.

Now we discuss the effect of farmer-related factors on wheat yield. One factor that is previously sown crop where several crops were indicated. Of these previously sown crops, rice was the most significant as compared to other previously sown crops which increases the wheat yield. Other previously sown crops also increases wheat yield but not significantly. The second-most important farmer-related factor is chemical coating of the wheat seed which also contributes a significant role in the wheat yield. From Table 4, we observed that the wheat yield of the chemical coating seed was maximum as compared to the wheat yield where the seed is not chemical coated. The use of compost also increases wheat yield. Results indicate that the wheat yield of the land was higher where compost was used than where it was not used. In the similar way, we also study the impact of some fertilizers namely urea, Nitrophos, nitrogen, DAP, and potash on the wheat yield. Table 4 results show that urea fertilizer has a direct influence on the wheat yield while other types of fertilizers are not beneficial to some extent of this land area. Furthermore, water frequency has negative impact on wheat yield, which means that this crop need small water frequency. If we increase the water frequency, then wheat yield may be decreased. Pesticide frequency also has a direct impact on wheat yield. This indicated that several times use of pesticides increases the wheat yield.

After removing all outliers, the results of the fitted model as shown in Table 4 were entirely different as compared to the results of the model fitted to the full data model. On comparing the fitted models as given in Table 4, significant results satisfying the regression assumption outliers were found. It was found that the R^2 and adj R^2 (Adjusted) are increased from (45.24% to 48.38%) and (40% to 45.91%), respectively (Table 4). The standard error of the estimates decreased from (2.772 to 2.649) and indicated better results with satisfying regression assumptions.

4. Discussion

All previous studies for the wheat yield model did not give any attention to regression assumptions and all these factors simultaneously. This study focused on the effect of outliers on the regression analysis to identify the factors which affects the wheat yield. Here, we also studied the effect of outliers on the regression assumptions. We have seen that outliers have strongly affected the model estimates. Also, it can be observed that insignificant factors have turned out to be significant after deleting all these outlying values. After considering the regression assumptions, the significant role of organic matter, copper, and urea from fertilizers can be observed. These variables were insignificant when regression assumptions were not satisfied. These results coincide with previous studies [35–42].

The outliers may also be the cause of multicollinearity, heteroscedasticity, and autocorrelation. It is also observed that after deleting outlying observations, the multicollinearity, heteroscedasticity, autocorrelation, and error variance were reduced substantially. Based on our results, it is found that saturation (%), EC, pH, phosphorus, potassium, CC, zinc, iron, manganese, boron, chemical coating of seed, previously sown crops are the significant factors to change the wheat yield. Based on these results, it may be suggested that in all agriculture-related fields, the identification and exclusion of outliers is a crucial step for obtaining better results of regression analysis.

5. Conclusion, Limitations, and Future Research

There are various factors which affect the crops yields. These factors include soil characteristics and farmer-related characteristics. As our interest is to identify the soil and other factors which influence the wheat yield. Also identify which one factor(s) are most critical to increase wheat yield which can be helpful for farmer. For the identification of such factors, we considered the regression analysis with the evaluation of outliers. When we ignore the presence of outlier, then the role of some important factors is found to be statistically insignificant. When we consider the outlier and exclude from the data, then these insignificant factors now indicate the significant role in the wheat yield. From the results, we found that saturation (%), EC, pH, phosphorus, potassium, CC, zinc, iron, manganese, boron, chemical coating of seed, and previously sown crops are the significant factors to change the wheat yields. Of these significant factors, more important six factors with order of importance are zinc, phosphorous, CC, pH, boron, and saturation.

This study considered some factors related to wheat yield, there may be other factors which may be related to wheat yield. These factors may be seed variety, land type, bushels, and so on. As our interest was mainly on soil characteristics and some other additional factors, we are unable to collect the information about these factors. This study can also be extended for other crops, such as maize, rice, and others to identify soil related factors and influence on yield. Moreover, such analysis can also be consider by considering some other regression models such as generalized linear models, nonlinear models, and nonparametric regression model.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Government of Pakistan, "Pakistan Economic Survey 2020-21. Finance Division," 2021, https://www.finance.gov.pk/survey/chapters_21/PES_2020_21.pdf.

[2] M. Aslam, "Agricultural productivity current scenario, constraints and future prospects in Pakistan," *Sarhad Journal of Agriculture*, vol. 32, no. 4, pp. 289–303, 2016.

[3] A. Azam and M. Shafique, "Agriculture in Pakistan and its impact on economy. A review," *International Journal of Advanced Science and Technology*, vol. 103, pp. 47–60, 2017.

[4] Government of Pakistan, "Pakistan Economic Survey 2019-20. Finance Division," 2020, http://www.finance.gov.pk/survey/chapter_20/PES_2019_20.pdf.

[5] P. R. Shewry and S. J. Hey, "The contribution of wheat to human diet and health," *Food and Energy Security*, vol. 4, no. 3, pp. 178–202, 2015.

[6] A. R. Barzegar, A. Yousefi, and A. Daryashenas, "The effect of addition of different amounts and types of organic materials on soil physical properties and yield of wheat," *Plant and Soil*, vol. 247, no. 2, pp. 295–301, 2002.

[7] P. Martre, J. R. Porter, P. D. Jamieson, and E. Triboï, "Modeling grain nitrogen accumulation and protein composition to understand the sink/source regulations of nitrogen remobilization for wheat," *Plant Physiology*, vol. 133, no. 4, pp. 1959–1967, 2003.

[8] Z. Plaut, B. J. Butow, C. S. Blumenthal, and C. W. Wrigley, "Transport of dry matter into developing wheat kernels and its contribution to grain yield under post-anthesis water deficit and elevated temperature," *Field Crops Research*, vol. 86, no. 2-3, pp. 185–198, 2004.

[9] A. A. Leilah and S. A. Al-Khateeb, "Statistical analysis of wheat yield under drought conditions," *Journal of Arid Environments*, vol. 61, no. 3, pp. 483–496, 2005.

[10] M. Moriondo, F. Maselli, and M. Bindi, "A simple model of regional wheat yield based on NDVI data," *European Journal of Agronomy*, vol. 26, no. 3, pp. 266–274, 2007.

[11] M. A. Khan, M. P. Fuller, and F. S. Baloch, "Effect of soil applied zinc sulphate on wheat (Triticum aestivum L.) grown on a calcareous soil in Pakistan," *Cereal Research Communications*, vol. 36, no. 4, pp. 571–582, 2008.

[12] W. R. Whalley, C. W. Watts, A. S. Gregory, S. J. Mooney, L. J. Clark, and A. P. Whitmore, "The effect of soil strength on the yield of wheat," *Plant and Soil*, vol. 306, no. 1-2, pp. 237–247, 2008.

[13] G. Abbas, M. Q. Khan, M. J. Khan, F. Hussain, and I. Hussain, "Effect of iron on the growth and yield contributing parameters of wheat (Triticum aestivum L.)," *The Journal of Animal and Plant Sciences*, vol. 19, no. 3, pp. 135–139, 2009.

[14] I. Hassan, M. B. Chattha, T. H. Chattha, and M. A. Ali, "Factors affecting wheat yield: a case study of mixed cropping zone of Punjab," *Journal of Agricultural Research*, vol. 48, no. 3, pp. 403–408, 2010.

[15] H. Gul, A. Said, B. Saeed, F. Mohammad, and I. Ahmad, "Effect of foliar application of nitrogen, potassium and zinc on wheat growth," *ARPN Journal of Agricultural and Biological Science*, vol. 6, no. 4, pp. 56–59, 2011.

[16] M. A. Nadim, I. U. Awan, M. S. Baloch et al., "Effect of micronutrients on growth and yield of wheat," *Pakistan Journal of Agricultural Sciences*, vol. 48, no. 3, pp. 191–196, 2011.

[17] M. Rezaei and Z. Hemati, "Reaction of wheat yield to soil physical and chemical Characteristics in Arak Fields," *World Applied Sciences Journal*, vol. 20, no. 8, pp. 1183–1187, 2012.

[18] S. R. El-Lethy, M. T. Abdelhamid, and F. Reda, "Effect of potassium application on wheat (Triticum aestivum L.) cultivars grown under salinity stress," *World Applied Sciences Journal*, vol. 26, no. 7, pp. 840–850, 2013.

[19] M. K. Alam and N. Salahin, "Changes in soil physical properties and crop productivity as influenced by different tillage depths and cropping patterns," *Bangladesh Journal of Agricultural Research*, vol. 38, no. 2, pp. 289–299, 2013.

[20] P. Maurya, V. Kumar, K. K. Maurya, N. Kumawat, R. Kumar, and M. P. Yadav, "Effect of potassium application on growth

and yield of wheat varieties," *The Bioscan*, vol. 9, no. 4, pp. 1371–1373, 2014.

[21] A. Limon-Ortega and E. Martinez-Cruz, "Effects of soil pH on wheat grain yield and quality," *Communications in Soil Science and Plant Analysis*, vol. 45, no. 5, pp. 581–591, 2014.

[22] A. K. Ghadikolayi, S. A. Kazemeini, and M. J. Bahrani, "Wheat yield and soil properties as influenced by crops residues and nitrogen rates," *Australian Journal of Crop Science*, vol. 9, no. 9, pp. 853–858, 2015.

[23] M. V. M. Sarto, M. d. C. Lana, L. Rampim, J. S. Rosset, and J. R. Wobeto, "Effects of silicate application on soil fertility and wheat yield," *Semina: Ciências Agrárias*, vol. 36, no. 6, pp. 4071–4082, 2015.

[24] M. Arshad, M. Adnan, S. Ahmed et al., "Integrated effect of phosphorus and zinc on wheat crop," *American-Eurasian Journal of Agricultural & Environmental Sciences*, vol. 16, no. 3, pp. 455–459, 2016.

[25] Q. Mehmood, M. Riaz, M. H. Sail, and M. Moeen, "Identifying key factors for maximizing wheat yield: a case study from Punjab (Pakistan)," *Pakistan Journal of Agricultural Research*, vol. 31, no. 4, pp. 361–367, 2018.

[26] F. Chairi, O. Vergara-Diaz, T. Vatter et al., "Post-green revolution genetic advance in durum wheat: the case of Spain," *Field Crops Research*, vol. 228, pp. 158–169, 2018.

[27] V. Rajičić, J. Milivojević, V. Popović et al., "Winter wheat yield and quality depending on the level of nitrogen, phosphorus and potassium fertilization," *Journal of Agriculture and Forestry*, vol. 65, no. 2, pp. 79–88, 2019.

[28] K. Polisetty and K. K. Paidipati, "Statistical assessment of trend analysis on production of wheat crop over India," *Sarhad Journal of Agriculture*, vol. 36, no. 1, pp. 178–184, 2019.

[29] T. A. Sial, Z. Lan, L. Wang et al., "Effects of different biochars on wheat growth parameters, yield and soil fertility status in a silty clay loam soil," *Molecules*, vol. 24, no. 9, pp. 1798–1817, 2019.

[30] M. N. Ashraf, M. H. Mahmood, M. Sultan et al., "Investigation of input and output energy for wheat production: a comprehensive study for tehsil Mailsi (Pakistan)," *Sustainability*, vol. 12, no. 17, pp. 6884–6905, 2020.

[31] H. Zhou, W. R. Whalley, M. J. Hawkesford et al., "The interaction between wheat roots and soil pores in structured field soil," *Journal of Experimental Botany*, vol. 72, no. 2, pp. 747–756, 2021.

[32] A. Hayat, M. Amin, and S. Afzal, "Statistical investigation to explore the impact of soil and other characteristics on cotton yield," *Communications in Soil Science and Plant Analysis*, vol. 51, no. 11, pp. 1434–1442, 2020.

[33] I. Pardoe, *Applied Regression Modeling*, John Wiley and Sons, Inc, Hoboken, New Jersey, 2nd edition, 2012.

[34] M. Amin, A. Akbar, and M. A. Manzoor, "Fitting regression model with some assumptions for the identification of cotton yield factors," *Pakistan Journal of Life and Social Sciences*, vol. 13, no. 2, pp. 86–90, 2015.

[35] F. Visconti, J. M. de Paz, and J. L. Rubio, "Principal component analysis of chemical properties of soil saturation extracts from an irrigated Mediterranean area: implications for calcite equilibrium in soil solutions," *Geoderma*, vol. 151, no. 3-4, pp. 407–416, 2009.

[36] B. Y. Li, D. Zhou, L. Cang, H. L. Zhang, X. H. Fan, and S. W. Qin, "Soil micronutrient availability to crops as affected by long-term inorganic and organic fertilizer applications," *Soil and Tillage Research*, vol. 96, no. 1-2, pp. 166–173, 2007.

[37] D. Sarkar, B. Mandal, and M. C. Kundu, "Increasing use efficiency of boron fertilisers by rescheduling the time and methods of application for crops in India," *Plant and Soil*, vol. 301, no. 1-2, pp. 77–85, 2007.

[38] N. Grotz and M. L. Guerinot, "Molecular aspects of Cu, Fe and Zn homeostasis in plants," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1763, no. 7, pp. 595–608, 2006.

[39] S. Hussain, M. A. Maqsood, Z. Rengel, and T. Aziz, "Biofortification and estimated human bioavailability of zinc in wheat grains as influenced by methods of zinc application," *Plant and Soil*, vol. 361, no. 1-2, pp. 279–290, 2012.

[40] B. M. Khan, B. R. Khan, A. Razzaq et al., "Effect of different tillage implements on the yield of wheat," *Pakistan Journal of Agricultural Research*, vol. 7, pp. 141–147, 1986.

[41] M. Nabi, "Relationship between crop productivity and input usage," *Journal of International Development*, vol. 8, pp. 68–88, 1991.

[42] K. Bakhsh, I. Hassan, and A. Maqbool, "Factors affecting cotton yield: a case study of Sargodha (Pakistan)," *Journal of Agriculture and Social Sciences*, vol. 1, no. 4, pp. 332–334, 2005.