STATISTICS, VALIDATION & MODELING of TEMPERATURE, RELATIVE HUMIDITY AND WET BULB IN HYGROSCOPIC CHEMICALS' STORAGE ENVIRONMENT- with A Case Study

Reyhan ATA*

Department/Research Institute, University, Country Tekirdağ Namık Kemal University, Çerkezkoy Vocational School, Tekirdag 59860, Turkey

*(rata@nku.edu.tr) *Email of the corresponding author*

Abstract – Validation is an obligatory for human health products such as food, medicine, cosmetics and auxiliary products due to international and national requirements. Hygroscopic chemicals, are the products require appropriate air humidity levels to be stored and validation should be made in accordance to relevant legislation. In this study, temperature (T), Relative Humidity (%RH) and Wet Bulb Temperature (WB) were measured in summer, autumn and winter in a company storage area of glycerol, glycose, fructose, galactose, riboflavin, niacin, folic acid, panthenoic acid, sulfuric acid, many fertilizer chemicals. Results of statistical analysis, histogram, detrended plot, skewnsess and kurtosis revealed that measurement data-set did not exhibited normal distribution pattern. Therefore, Mann-Whitney-U test was used to define significant difference in data-set. A corelation analysis was applied to determine the effect of temperature, %RH and WB on each other. A regression analysis and modeling was carried out to explain the relationship. Finaly, Fstatistic, tstatistic , %Sr values were calculated for the validation after removing outliers. In order to protect the shelf life of hygroscopic chemicals, the relative humidity increase rate has been provided without increasing the temperature by the developed modeling.

Key words; Validation, Hygroscopic Chemicals, Relative Humidity, Wet-Bulb, Correlation- Regression-Modeling

I. INTRODUCTION

Hygroscopic chemicals react to changes in air humidity in the environment due to their characteristic features. If the weather conditions of the storage ambient is quite dry, such stored materials may lose their moisture. If appropriate weather humidity levels are kept, these kinds of chemicals can be stored and retain their quality longer, losing less weight and maintaining shelf life. The common storage conditions of the hygroscopic substances, such as ellulose fibers (such as cotton and paper), glycerol, glycose, fructose, galactose, riboflavin, niasin, folik asit, pantenoik asit, sulfuric acid, many fertilizer chemicals and many salts (like calcium chloride, bases like sodium hydroxide etc.), and a wide variety of other substances even tea, baker's yeast, are 18-22oC temperature, 60-80% %RH and 14-20oC WB temperatures suitable for these conditions [1]. Additionaly to all these, in accordance with the requirements of the USA Food and Drug Adminitration, European Drug Agency, ISO/IEC 17025 General requirements for the competence of testing and calibration laboratories standards, ISO 22001:2018 Food Management Systems, Turkish Good Manufacturing Practices (GMP) for Medicinal Products in Human Use, validation, verification and qualification are obligatory for these kind of companies manufacturing products for human health such as food, medicine, cosmetics and auxiliary products and services (raw materials, packaging materials, logistics support). Validation is the determination of compliance with performance criteria and the provision of objective evidence that certain requirements are met for a particular intended use according to the ISO/IEC 17025 standard [2-4]. Validation is a ind of guarantee that a product or service meets the needs of the customer and other stakeholders. Validation is assurance that a product, service or system meets the needs of the customer and other identified stakeholders. It usually includes acceptance and compliance with customers or pre-defined criteria specified in MSDS's (Material Safety Data Sheet), Technical Data Sheets, relevant standards or in the reports of stability tests. With verification, it is also aimed to check whether a product, service or system meets the specifications determined during the research and development stage [3-11]. If it were to be expressed with the questions if the right product being produced for manufacturing companies, and if the right service being presented for service companies and it corresponds to that questions reveal the customers' needs are explained by validation process. On the other hand, the questions of that, if the product is produces with the right methods for the manufacturing companies, if the service is offered correctly for the companies providing services, were responded by verification process to check whether the specifications of the product or service were applied appropriately in that company. The validation process should be carried out by consultants, validation team (Quality Assurance, R&D, Production, Quality Control or Engineering Departments), expert, experienced, knowledgeable people in GMP and validation. A separate department or team can be established for validation works in enterprises. A separate specified department or team can be established for validation works in the companies. Statistical data analysis is one of the important parts of the validation process. In this paper, validation study starts with calculating the central tendency measurements of temperature, %RH and WB values of the company such as arithmetic mean, mode, median of the measurement values collected with the selected measurement methods. Afterwards, the relative positions of the results, regarding these kind of hygroscopic chemicals stored, such as closeness to each other, significance and difference, standard deviation, variance distribution (spread, scattering) have been calculated.

-Arithmetic mean ($\bar{X}$); It is the value obtained by summing the results from a group of measurement and dividing by the number of analyzes and is calculated with Eq. (1) [12].

-Arithmetic mean ($\bar{X}$); It is the value obtained by summing the results from a group replicate analysis (measurement) and dividing by the number of analyzes and is calculated with Eq. (1) [12].

$$\bar{x} = \frac{\sum x_i}{n}$$

Eq. (1) where, X_i = results of the measurements, n = number of measurements

-Median (μ) ; As the results in a data set are arranged in ascending order, the value in the middle of the row is called the median. If the median value is that one in odd-numbered analysis results, it is two in even-numbered analysis results. In such cases, the average of the two middle values is taken as the median value. Median value is calculated by the formula given in Eq. (2) [12].

$\mu = \{(n + 1) \div 2\}^{th}$ value. Eq. (2). where, n = number of items (measurements)

-Mod (Mo) ; The most repeated value in a data set is called the mode. Mod value is calculated by the formula given in Eq. (3) [12].

$Mo = L_0 + [(f_1-f_0) / (2f_1-f_0-f_2)]\times h$. Eq.(3). where, L_0= lower limit of modal class,

h=class interval size, f1=frequency of modal class, f_0=frequency of proceeding to modal class, f_2=frequency of succeeding to modal class.

-Standard Deviation (σX); It is the measurement of the distribution of the results obtained from a group analysis around the mean value. It is calculated by the formula given in Eq.(4).

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

Eq.(4). where, Xi=Results of the measurements, n=number of items, $\bar{X}$= arithmetic mean

-Variance (σ2); It is the square of the standard deviation calculated by Eq.(4) [12].

-Relative Standard Deviation (%Sr); It is a comparable value which it is the standard deviation divided by the mean value and calculated with the formula given in Eq.(5) [12].

$\%Sr = (\sigma X / \bar{X}).100$ Eq. (5). where,

$\bar{X}$=arithmetic mean , σX = standart deviation

-Error (Bias, Deviation) (ϵ); It is the difference between the measured quantity value and the reference quantity value.

-Degree of freedom (df); It is usually called 1 minus the number of measurements calculated with the formula given in Eq.(6) [12].

df=n-1 Eq.(6) where, n= number of items

-Skewness; It is the third standardized moment, defined as Eq. (7).

$$\mathrm{skew}(X) = \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$ Eq. (7). where, μ3 = third central moment

σ = standard deviation.

-Kurtosis; is the fourth standardized moment, defined as Eq. (8). [12].

$$\mathrm{kurt}(X) = \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$$ Eq. (8) [12]. Where, μ3 = fourth central moment

σ = standard deviation.

The validation study continued with statistical significance tests for the evaluation of the analytical data (measurement results). In validation study accuracy is used to show the closeness of the measurement results to the true value and to each other. The accuracy parameter consists of two main components: trueness and precision which is the closeness of the results found to each other, is called precision indicates consistency. In order to define the trueness, it is necessary to know the correct value or a reference value that is accepted as correct (in the stability test results or the value specified in the MSDS or the relevant standard), i.e. the real value. With significance tests;

- The value considered correct can be compared with the experimentally determined mean value. With this comparison, it is checked whether a method has a systematic error.
- The mean values calculated from two separate measurement series can be compared and it is determined whether there is a difference in terms of their validity.
- Standard deviations, calculated from two separate measurement series, can be compared, so that their precision can be compared.
- If one or more of a group of measurement results are significantly different from the others, it can be decided whether to exclude it from the group (regarding the number of available data).
- The range, in which the correct value will be found, can be calculated.

Commonly used significance tests are F-Test, T-Test, ANOVA Test (Test of Variance) in order to evaluate the measurement results in terms of accuracy (precision and trueness).

-F-test; It is also known as "significance tests of precision", which is used to compare the precisions (standard deviations) of two data series and is calculated with the formula given in Eq. (9) [12]. For the F-test, the squares of the standard deviations ($(\sigma_X)^2$), of the two data-sets to be compared, are divided into large value/small value and the results are compared with the $F_{critical}$ value obtained from the tables at the desired confidence level.

$$F_{statistical} = \frac{(\sigma_{X1})^2}{(\sigma_{X2})^2}$$ Eq. (9) where $(\sigma X)^2$ = the squares of the standard deviations

If $F_{statistical} < F_{critical}$, it is decided that the difference between the standard deviations is not statistically significant.

-t-Test; One of the frequently used hypothesis tests for two data series is the t-test. It is used to compare means. It is used to compare the mean value of a data series with a predetermined value (obtained by reference method, certificate value or specified in the Material Safety Data Sheet (MSDS), standard or Technical Data Sheet etc.) and to test whether the difference between the means is significant at a certain confidence level. There are three types of t-tests, they are;

a. One-Sample t-test,
b. Independent Samples t-test,
c. Paired-samples t-test.

These three types of t-tests can be used for different purposes in different fields. A t-value is calculated for all t-tests and this ($t_{statistical}$) is compared with $t_{critical}$ obtained from the t-table to the chosen confidence interval and degrees of freedom. If $t_{statistical} < t_{critical}$, it is judged that the difference between the compared mean values, are not significant.

-ANOVA Test (Analysis of Variance); The ANOVA test is used to compare three or more means. ANOVA analysis calculates the variance between groups and the variance within the groups, and makes a decision based on the size of these variances. In ANOVA analysis, different data sets are considered collectively, only whether the difference is significant or not. ANOVA analysis does not give the source of the difference, if any. Conditions that must be met for the ANOVA test to be performed correctly;

- data have a normal distribution
- equal group variances
- data are independent of each other.

ANOVA analysis, depending on the number of influencing independent variable (factor), It is classified as;

- One Way ANOVA (One Independent Variable, One Dependent Variable)

-Two-Way ANOVA (Two Independent Variables, One Dependent Variable).

No matter how many groups are compared in the ANOVA, there is a single F ($F_{statistical}$) value and a corresponding single p value. These $F_{statistical}$, $F_{critical}$, and p values in the results table were considered in order to decide whether there is a statistically significant difference between the multiple groups compared as a result of the analysis.

-p value: The p value in the ANOVA table is called the significance level, and a threshold value for the p significance level (this value is *0.05* for the *95%* confidence level and *0.01* for the *99%* confidence level) is selected while performing the analysis. If the p value is less than the significance level that is determined, it indicates that there is a significant difference between the group means, and if the p value is greater than the significance level that is determined, it indicates that there is no significant difference between the group means. Accordingly, as a result of the ANOVA analysis (for *95%* confidence level), statistically;

- There is a significant difference between the $F_{statistical} > F_{critical}$ and $p < 0.05$ groups.
- There is not a significant difference between the $F_{statistical} < F_{critical}$ and $p > 0.05$ groups [13].

Correlation analysis; It is a statistical method that provides information about the relationship between variables, the direction and severity of this relationship. The correlation coefficient (r) is a ratio, means a value between -1 and +1. If the coefficient is positive, it means one of the variables increases as the other increases; If it is negative, it means that as one variable increases, the other decreases. Hypotheses for the significance level of correlation coefficient; For H_0: $p=0$ and H_1: $p>0$

$t_H=r/\sqrt{(1-r^2/n-2)}$ calculated with Eq.(10) [12]. where

r=Correlation Coefficient,

n-2= degrees of freedom

For the given significance level (α); When,

$t_H>t\alpha{:}n\text{-}2$, H_0 is rejected and H_1 is accepted.

It is seen whether the correlation coefficient r is statistically significant and there is a relationship between the variables [14].

Regression analysis is an analysis method used to measure the relationship between two or more quantitative variables. Regression analysis consists of several types including linear, nonlinear and multilinear. But the most useful types of regression are simple linear and multiple linear regression. Simple linear regression allows to test whether there is a linear relationship between two variables. The general equation for a linear regression model is Eq. (11) [15].

$F(x)=Y=\beta_0+\beta_1X_1+\epsilon$ Eq.(11)[12]. Where,

Y=dependent varaible,

X= independent variable,

β0=intersection point,

β= the slope of the line,

ϵ=error

A multiple linear regression gives the opportunity to test whether there is a linear relationship between more than two variables. The general equation for a multiple linear regression model is Eq (12).

$F(x)=Y=\beta_0+\beta_1X_1+\beta_2X_1+\ldots B_nX_n+\epsilon$ Eq. (12) [12].

The mathematical relationship between two or more variables is analyzed by "Regression Analysis" and the direction and degree of the relationship is examined by "Correlation Analysis". In this study, the storage ambient conditions were validated by using the Temperature, %RH and WB temperature measurement values of the hygroscopic chemicals storage areas having a heating, cooling and ventilation system in a company. The results of this study will also be used in a sort of prospective validation to decide on the necessity of steam humidification of storage areas. The relative humidity (%RH) mentioned in this study is the ratio of the partial pressure of the water vapor in the air to the equilibrium vapor pressure of the water at the same temperature in the storage area of the company. In other words, relative humidity indicates what percentage of the amount of moisture the air can carry at a given temperature. Relative humidity can change when the water vapor content of the air or the air temperature changes. Wet-Bulb, basically, temperature is the temperature of adiabatic saturation. This is the temperature indicated by a moistened thermometer bulb exposed to the air flow. Wet Bulb temperature can be measured by using a thermometer with the bulb wrapped in wet muslin. It is identical with 100% relative humidity. In this study, temperature, %RH and WB values were measured using a device and recorded with a data logger. The chemicals, in the storage house of the company, where the study is carried out, are called hygroscopic chemicals group such as glycerol, glycose, fructose, galactose, riboflavin, niacin, folic acid, panthenoic acid, sulfuric acid, many fertilizer chemicals and similar substances. The formula structure of the hygroscopic chemicals available in Figure 1. One of the chemicals in the storage area, glycerol (glycerine) is a polar organic trihydroxy hygroscopic chemical in liquid form. It is a slightly sweet, non-toxic liquid and miscible with water and ethanol [16]. Glucose, that of others, is overall the most abundant monosaccharide, a subcategory of carbohydrates having a 6 carbon [17] and fructose has a 6-carbon polyhydroxy ketone bonded to glucose to form the disaccharide sucrose [18]. Also, galactose is a 6 carbon carbohydrates monosaccharide like glycose, fructose [19]. Riboflavin composed from the pentose which are ribitol and lumichrome, has $C_{17}H_{20}N_4O_6$ chemical formula, also known as vitamin B2 [20].

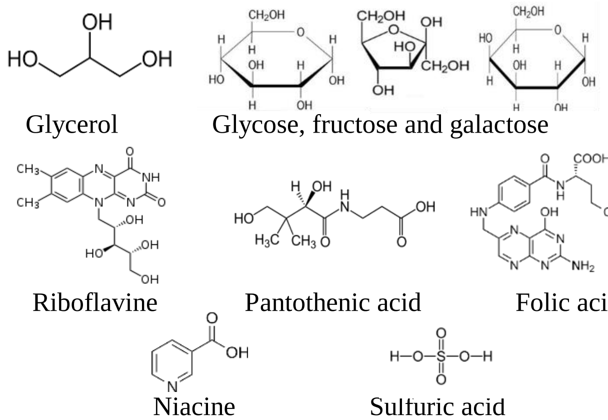


Figure 1. Formula structure of chemicals available in the storage of the company [16-25].

Additionally niacin, known as nicotinic acid, is an organic compound in a form of vitamin B3, that has the formula $C_6H_5NO_2$ and belongs to the group of the pyridinecarboxylic acids [21]. Other chemicals, folic acid, known as Folate, also known as vitamin B9 and folacin [6], is one of the B vitamins with the formula of $C_{19}H_{19}N_7O_6$ [22,23]. Moreover Pantothenic acid, also called vitamin B5 is a water-soluble B vitamin and therefore an essential nutrient has a formula of $C_9H_{17}NO_5$ [24]. And at last sulfuric acid, is a mineral acid composed of the elements sulfur, oxygen and hydrogen, with the molecular formula H_2SO_4 [25].

2. Case Study- Material and Method– Obtaining validation data

In this study, Temperature, %RH and WB temperature were measured in the hygroscopic chemical storage area of the company (Fig. 2.), where the study was carried out, at 08:00 in the morning and 20:00 in the evening during July, August, September, October, November and December. For this purpose, Extech RH490 Precision Hygro–Thermometer–Multifuctional High Precision Temperature and Humidity Meter was used. The measurements, performed from the critical points on the storage area according to the settlements of the chemicals, were recorded via calibrated datalogger. A 368 temperature, 368 %RH and 368 WB temperature measurement data-set (totally1104), exhibiting the seasonal changes in the summer (July, August), autumn (September, October), winter (November, December), daily at 08:00 in the morning (AM) and 20:00 in the evening (PM) climatic conditions of a 120 square meter storage area operated with a heating, cooling and ventilation system, were used in this study. Regarding the MSDSs and stability test results of these chemicals were examined, the reference was taken as 7.90569415 for the relative standart sapma ($\%Sr_{crt}$) at storage conditions not exceeding 24 to 25^{o}C temperature, at interval of %60 to 75 relative humidity and 14-20^{o}C WB temperatures. In this context, The Lower Limit (LL) and Upper Limit (UL) is accepted as 18^{o}C to 22^{o}C respectively for the temperature values, as %60 to %75 for the %RH values and as 14 to 20^{o}C for the WB temperature values (LL_T =18^{o}C, UL_T =22^{o}C, $LL_{\%RH}$ =%60, $UL_{\%RH}$ =%75, LL_{WB}=14^{o}C, UL_{WB} =20^{o}C).

SPSS program was used for the statistical analysis and the validation study. The storage area plan is as seen in Figure 2. There is a mechanical ventilation system that has been used for many years, including industrial type ventilation ducts, air filters and vents, where heating and cooling processes were carried out in that storage plant.

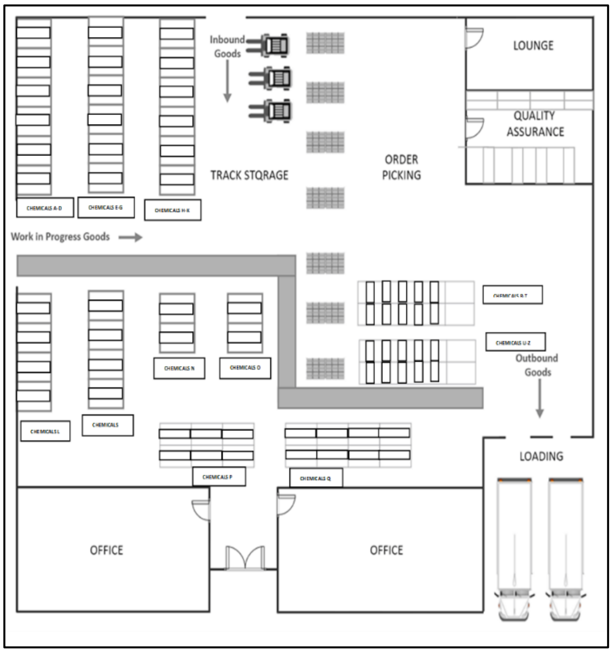


Figure 2. Storage area plan

3. Results and Discussion

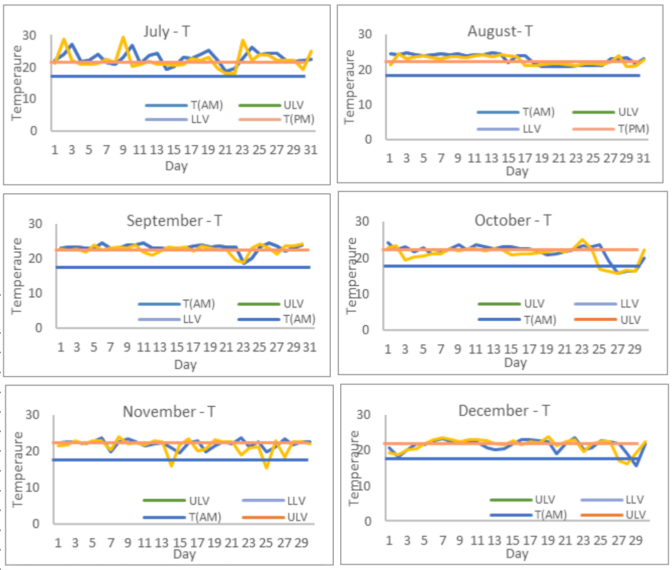
The study started with the statistical descriptive analysis of the measured temperature values presented in Table 1, for the months of July, August, September, October, November and December.

Table 1. Descriptive statistical analysis of temperature data-set

Descriptive Statistics

	N	Min.	Max.	Diff.	$\bar{X}$	Std. Dev.	σ^2	Skew.	Std. Err.	Kurt.	Std. Error
JAMT	31	18,5	27,3	8,8	22,8	2,06	4,3	0,1	0,42	0,1	0,82
JPMT	31	18	29,4	11,4	22,2	2,67	7,1	1,3	0,42	2,1	0,82
AAMT	31	20,8	24,9	4,1	23,0	1,49	2,2	-0,4	0,42	-1,5	0,82
APMT	31	20,8	24,6	3,8	22,6	1,23	1,5	-0,1	0,42	-1,6	0,82
SAMT	30	18,6	24,6	6	23,2	1,18	1,4	-2,5	0,43	8,4	0,83
SPMT	30	19	24,3	5,3	22,7	1,23	1,5	-1,5	0,43	2,7	0,83
OAMT	31	15,7	24,2	8,5	21,6	2,15	4,6	-1,7	0,42	2,6	0,82
OPMT	31	15,8	25,0	9,2	20,9	2,29	5,2	-1,1	0,42	0,6	0,82
NAMT	30	19,4	23,8	4,4	22,0	1,17	1,3	-0,8	0,43	0,1	0,83
NPMT	30	15,3	23,9	8,6	21,4	2,01	4,0	-1,8	0,43	3,5	0,83
DAMT	31	15,7	23,4	7,7	21,1	1,95	3,8	-1,5	0,42	2,2	0,82
DPMT	31	16,2	23,7	7,5	21,3	2,02	4,0	-1,2	0,42	0,7	0,82
N	30										

As seen in the Table 1., mean values for the AM and PM temperatures were measured as 22,82°C, 22,16°C and 23,04°C ve 22,62°C in summer months in July and August respectively. In the autumn months, in September, mean values were measured as 23,20°C and 22,71°C, in October as 21,68°C and 20,96°C respectively. Moreover, in winter months, in November they were measured as 22,01°C and 21,47°C, and in December 21,16°C and 21,34°C as close to each other, respectively. Maximum and minimum temperature values determined, for the AM to PM hours, respectively as 18,5°C, 18°C to 27,3°C, 29,4°C in July, as both 20,8°C to 24,9°C, 24,6°C, in August, as 18,6°C, 19°C to 24,6°C, 24,3°C in September, as 15,7°C, 15,8°C to 24,2°C, 25°C in October, as 19,4°C, 15,3°C to 23,8°C, 23,9°C in November and finally 15,7°C, 16,2°C to 23,4°C, 23,7°C in December. These measurements reveal that the months, with the highest maximum and minimum temperature differences, are in the evening of July, then in the morning of July and in the evening of October. The highest mean temperature value was 23,04°C meaured in August AM and the lowest mean temperature value was 20,96°C in October PM time. Standard deviation values were generally close to each other and the highest standard deviation was 2,67 measured in July PM time, when the temperature distributions were relatively higher compared to other months. Since the skewness and kurtosis values were greater than 1.5, it means that temperature measurements data-set did no exhibit a normal distribution patern [26]. Daily temperature changes by months are illustrated in Graphic 1. in detailed.



Graph 1a. July Temperature Changes
Graph 1b. August Temperature Changes
Graph1c. September Temperature Changes
Graph 1d. October Temperature Changes
Graph 1e. November Temperature Changes
Graph 1f. December Temperature Changes

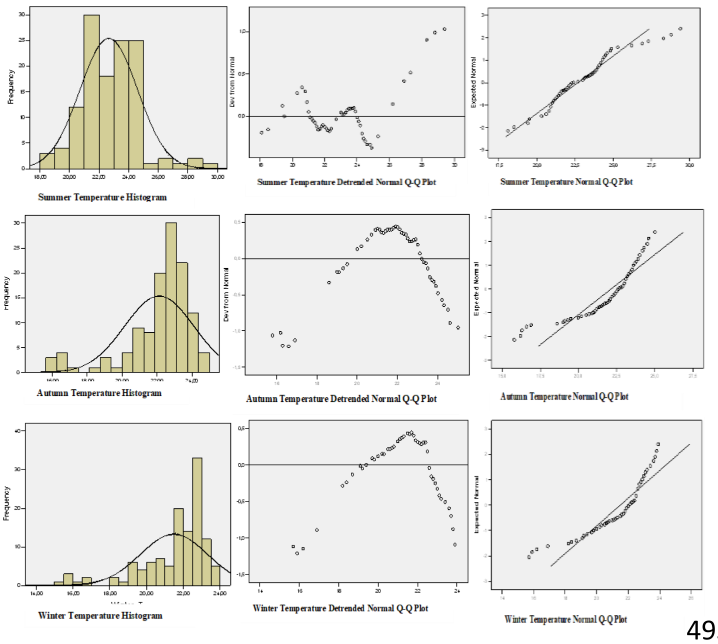
As observed in Graph. 1.a., 1e. and 1f., temperature changes increased towards the end of the month in July, and almost throughout the month in November and December. They were more stable in August, on the other hand they exhibited different pattern lowering below LL (18°C) and rising above UL

(22°C) towards the end of the month in September and October from morning to evening as seen in Graph. 1.d., ve 1e. Kolmogorov-Smirnov test was also performed to confirm whether this temperature measurement data-set had a normal distribution or not and test results are given in Table 2. in detail.

Table 2. Results for the test of Normality

Months	Time	Kolmogorov-Smirnov(a)		
		Statistic	df	Sig.
July – AM-T	AM	0,09	31	0,20
	PM	0,24	29	0,00
August-AM-T	AM	0,21	31	0,00
	PM	0,22	29	0,00
September-AM-T	AM	0,28	31	0,00
	PM	0,16	29	0,07
October-AM-T	AM	0,21	31	0,00
	PM	0,23	29	0,00
November-AM-T	AM	0,18	31	0,01
	PM	0,24	29	0,00
December-AM-T	AM	0,21	31	0,00
	PM	0,20	29	0,00
*	This is a lower bound of the true significance.			
a	Lilliefors Significance Correction			

Although Kolmogorov-Smirnov test reported that the temperature data-set has built a normal distribution patern, but skewness-kurtosis values, histogram and detrented graphs have revealed that they do not have a normal distribution character indeed, as it is illustrated in histogram and detrented graphs given in Graph 2. In terms of extreme values, 20 outliers (±2°C from UL and LL) were detected with a large number, and it was preferred not to remove these outliers, considering that both removing them would damage the data-set and preventing to reveal the real conditions of the existing heating, cooling and ventilation system of the storage plant.



Graph 2. Distribution characterization of the temperature data-set by seasons

By adding the monthly morning and evening temperature measurement values, July and August are defined as summer, September, October as autumn, November and December as winter, and the descriptive analysis are given in detail in Table 3.

Table 3. Descriptive analysis results of summer, autumn and winter temperature measurement data-set

Descriptive Statistics

	N	Range	Min.	Max.	Mean	Std. Dev.	Variance	Skew.	Std. Error	Skew./ Std.Err	Kurt.	Std. Error	Kurt. Std.Err.
Summer_T	124	11,4	18	29,4	22,66	1,95	3,80	0,60	0,22	2,73	1,46	0,43	3,39
Autumn_T	122	9,3	15,7	25	22,13	1,98	3,91	-1,74	0,22	-7,909	2,93	0,43	6,81
Winter_T	122	8,6	15,3	23,9	21,49	1,83	3,36	-1,60	0,22	-7,27	2,45	0,43	5,69
Valid N (listwise)	122												

As given in Table 3., Skewwness/Std.Error value was calculated as 2.73 for summer, as -7,909 for autumn and as -7.27 for winter. Moreover Kurtosis/Std. Error value was calculated as 3.39 for summer, as 6.81 for autumn, and as 5.69 for winter. Since these skewness and kurtosis values were greater than 1.5, it was determined that the temperature measurement data-set did not reval a normal distribution character during summer, autumn and winter [26]. Afterwards, a Levene's Test was performed to understand the homogeneity of the variances for the temperature data-set measured at AM and PM time and the results were given in Table 4.

Table 4. Test of Homogeneity of Variances

	Levene Statistics	*df1*	*df2*	*p*
July AM/PM-T	1,555	1/2	48	0,218
Agust AM/PM-T	,087	1/2	48	0,769
September AM/PM-T	0,573	1/2	48	0,453
October AM/PM-T	2,758	1/2	48	0,103
November AM/PM-T	0,505	1/2	48	0,481
December AM/PM-T	1,161	1/2	48	0,287

1= AM (measumements at 08:00 AM) 2=PM (measumements at 20:00 PM)

As observed in Table 4., it was found that the variances of the morning and evening temperature measurements had a homogeneous character in July, August, September, October, November and December (p>.05). In this case, the Mann-Whitney-U test was applied whether there is a significant difference between the temperature values measured in the morning and evening and in summer, winter and autumn under the operating conditions of this storage area. The results are presented in Table 5.

Table 5. Mann-Whitney-U test

	Summer-T	Autumn-T	Winter-T
Mann-Whitney U	1474,5	1376	1836,5
p (2-tailed)	0,0253	0,0131	0,9041

$p<0,05$

As seen in Table 5, under the operating conditions of the storage plant, there is a significant difference detected between the morning and evening temperature measurement values in summer and autumn months ($p<0.05$), but there is no significant difference detected in winter months ($p>0, 05$).

3.1 Effect of Temperature, %RH and WB on each other

%RH and WB values were measured in the storage area along with the temperature values during July, August, September, October, November and December. Pearson Product Moment Correlation Analysis was performed in order to determine the relationship between temperature, %RH and WB and to explain the effect of temperature, %RH and WB on each other. The results are given in detail in Table 6. A statistically non-significant negative correlation (at $p<0.01$ level) was determined between temperature and % RH values obtained by gathering monthly measurement values in summer, autumn and winter (respectively *r=-0,14772; p> 0.05;r=-0,03566; p>0 .05;r=-0,09376;p> 0.05).*On the other hand, a statistically significant positive correlation, at $p< 0.01$ level, was found between temperature and WB in summer, autumn and winter (respectively *r=0,.864727; p<0 .05 ;r= 0,865915; p<0.05;r=0,782454;p<0.05*).

Moreover, a statistically significant positive correlation, at $p< 0.01$ level, was found between %RH and WB in summer, autumn and winter (respectively, *r= 0,367618; p< .05 ;r= 0,4672; p< .05;r= 0,545235; p< .05*).

Table 6. Pearson Product Moment Correlation for T, %RH and WB in Summer, Autumn, Winter

		Summer_T	Summer_RH	Summer_WB
Summer_T	Pearson Correlation (*r*)	1	-0,14772	0,864727
	p (2-tailed)		0,101573	2,7E-38
	N	124	124	124
Summer_RH	Pearson Correlation (*r*)	-0,14772	1	0,367618
	p (2-tailed)	0,101573		2,67E-05
	N	124	124	124
Summer_WB	Pearson Correlation (*r*)	0,864727	0,367618	1
	p (2-tailed)	2,7E-38	2,67E-05	
	N	124	124	124
		Autumn_T	**Autumn_RH**	**Aumtumn_WB**
Autumn_T	Pearson Correlation (*r*)	1	-0,03566	0,865915
	p (2-tailed)		0,69658	6,59E-38
	N	122	122	122
Autumn_RH	Pearson Correlation (*r*)	-0,03566	1	0,4672
	p (2-tailed)	0,69658		5,8E-08
	N	122	122	122
Aumtumn_WB	Pearson Correlation (*r*)	0,865915	0,4672	1
	p (2-tailed)	6,59E-38	5,8E-08	
	N	122	122	122
		Winter_T	**Winter_RH**	**Winter_WB**
Winter_T	Pearson Correlation (*r*)	1	-0,09376	0,782454
	p (2-tailed)		0,304309	1,9E-26
	N	122	122	122
Winter_RH	Pearson Correlation (*r*)	-0,09376	1	0,545135
	p (2-tailed)	0,304309		8,47E-11
	N	122	122	122
Winter_WB	Pearson Correlation (*r*)	0,782454	0,545135	1
	p (2-tailed)	1,9E-26	8,47E-11	
	N	122	122	122

**. Correlation is significant at the 0.01 level (2-tailed).

3.2 Regression Analysis

A regression analysis was also carried out in order to explain the relationship between the temperature, %RH, WB values that are effective in the storage plant of the company. Normal distribution of the data, linearity, autocorrelation and the absence of multicollinearity problems between independent variables were the basic assumptions for multiple regression analysis. In line with these basic assumptions, the 'Enter' method of the regression test techniques, was used at a 95% confidence interval. The hypotheses and multiple linear regression equations formed within the framework of the research hypotheses in order to examine how much the T and WB values should increase in order to increase the %RH value under the existing system in the storage plant, that are as follows;

Hypothesis 1: T has an effect on %RH changes in the current system conditions in the storage plant.

Hypothesis 2: T and WB have a combined effect on %RH changes in the current system conditions in the storage plant.

(Hypothesis 1) $Y_1 = \beta_0 + \beta_1 X_1 + \varepsilon$,

(Hypothesis 2) $Y_2 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$ where,

Y = %RH values measured under the existing system in the warehouse in the enterprise

X1 = T - temperature value measured under the operating system of the storage plant.

X_2 = WB values measured under the operating system of the storage plant.

Multiple regression findings for testing hypotheses are presented in Table 7.

As a result of the analysis carried out, it has been seen that 2 different models have emerged for the summer months. Examining the significance of the models within the scope of regression analysis; Model 1 is not statistically significant with the value of $F(1,122)=2,722$ and $p= 0,102$ ($p>0,05$). On the other hand, Model 2. is statistically significant with the value of $F(2,121)=38256,275$ and $p= 0,000$ ($p<0,05$). Model 2. has the most powerful explanation of the effect of T and WB values on the increase in %RH values in summer. In other words, 99% of the variation in %RH values during the summer months is explained by the combination of the negative effect of the T values and the positive effect in the WB values. In line with the findings, it was determined that Hypothesis 1. was not fulfilled and was rejected, Hypothesis 2. was realized and the regression equation was $Y=\beta_0+\beta_1X_1+\beta_2X_2+\epsilon$ (Eq. 12).

For the summer mounths Model 2.; %RH= 71,773 + (-6,421x Summer T value) + (7,584 x Summer WB value)

Table 7. Regression analysis of T and WB values on the ability to increase %RH values

Summer							
Model 1	B	SE	*β*	Δ R^2	*t*	*F*	*p*
Constant	73,479	7,083					
Summer T	-0,513	0,311	–0,148	0,014	-1,650	*F*(1,122)=2,722	0,102
Model 2							
Constant	71,773	0,529					
Summer T	-6,421	0,046	–1,846		–138,599		0,000
Summer WB	7,584	0,051	1,964	0,994	147,453	*F*(2,121)=38256,275	0,000
Autumn							
Model 3							
Constant	66,612	7,877					
Autumn % RH	-0,139	0.355	-0,036	-0,07	-0,391	*F*(1,120)=0,153	0,697
Model 4							
Constant	75,068	0,673					
Autumn % RH	-6,839	0,060	-1,760		-113,460		0,000
Autumn WB	7,926	0,062	1,991	0,993	128,374	*F*(2,119)=8250,502	0,000
Winter							
Model 5							
Constant	61,356	9,938					
Winter % RH	-0,475	0,461	-0,094	0,001	-1,032	*F*(1,120)=1,064	0,304
Model 6							
Constant	69,086	0,688					
Winter % RH	-6,803	0,051	-1,342		-133,221		0,000
Winter WB	8,240	0,052	1,595	0,995	158,363	*F*(2,129)=12651,194	0,000

p<0,005

Also, 2 different models have emerged for the autumn months. Reviewing the significance of the models within the scope of regression analysis; Model 3. is not statistically significant with the value of $F(1,120)=0,153$ and $p= 0,697$ ($p>0,05$), but Model 4. is statistically significant with the value of $F(2,119)=38256,275$ and $p= 0,000$ ($p<0,05$).

Model 4. is the most powerful explanation of the effect of T and WB values on the increase in %RH values in autumn under the existing system in the storage plant. In other words, 99% of the variation in %RH values during the autumn months is explained by the combination of the negative effect of the T value and the positive effect in the WB values. In line with the findings, it was determined that Hypothesis 1. was not fulfilled and was rejected, Hypothesis 2 was realized and the regression equation was $Y=\beta_0+\beta_1X_1+\beta_2X_2+\epsilon$ (Eq. 12).

For the autumn mounths Model 4; %RH= 75,068 + (-6,839 x Autumn T value) + (7,926 x Autumn WB value)

Additionally, 2 different models have also emerged for the winter months. Among them, Model 5. is not statistically significant with the value of $F(1,120)=1,064$ and $p= 0,304$ ($p>0,05$), however Model 6. is statistically significant with the value of $F(2,119)=12651,194$ and $p= 0,000$ ($p<0,05$).

In summary, Model 6. is the most powerful explanation of the effect of T and WB values on the increase in %RH values in winter under the existing system in the storage plant. That means, 99% of the variation in %RH values during the winter months is explained by the combination of the negative effect of the T value and the positive effect in the WB values. In line with the findings, finally, it was determined that Hypothesis 1. was not fulfilled and was rejected, Hypothesis 2. was realized and the regression equation was $Y=\beta_0+\beta_1X_1+\beta_2X_2+\epsilon$ (Eq. 12).

For the winter months Model 6.; %RH= 69,086 + (-6,803 x Winter T value) + (8,240 x Winter WB value)

In this regression analysis, the VIF value less than 3 reveals that there is no linearity problem, and the tolerance statistics are calculated above 0.25.

Afterwards, in order to reach to the normal distribution for the validation, in this acceptance limits (UL=22ºC and LL=18ºC), the extreme values of the current temperature measurement data-set were calculated in SPSS and some of the extreme values such as 26- 28ºC and 15-16ºC were discarded, so that validation study was carried out. For this, the $F_{statistical}$ value was calculated to show the continuous distribution of the data-set after removing the outliers, and the results are given in Table 8.

Table 8. $F_{statistical}$ results for the summer, autumn and winter mounths

	F_{stat}	*df1*	*df2*	*F Critical two-tail*	*Mean*	
Summer-Autumn	1,16717	103	101	1,91262	22,15	22,36
Autumn-Winter	1,28967	101	114	1,92921	22,37	21,82
Summer-Winter	1,07058	103	114	1,98099	22,15	21,82

at %95 confidence interval

As it is given in Table 8. $F_{statistical} < F_{critical}$ was calculated for the summer, autumn and winter mounths and this means that the data-set consisting of temperature measurement values reveals continuous distribution and the temperature means values are within the acceptance limits at 95% confidence interval (respectively $F_{statistical}$ =1,16717< $F_{critical}$ =1,91262; $F_{statistical}$ =1,28967 < $F_{critical}$=1,92921;$F_{statistical}$=1,07058<$F_{critical}$=1,98099). It has been determined that even the removal of only a part of the outliers data, which are below and above ±2°C of the acceptance limits, from the data-set made it have a normal distribution. In addition, $t_{statistic}$ values were calculated in order to compare the measured temperature values with the mean value given at the determined acceptance limits (ULV=22ºC and LLV=18ºC) and to test whether they are within the acceptance limits at a certain confidence level, and the results are given in Table 9. collectively.

Table 9. Results for $t_{statistical}$ in summer, autumn and winter months

	t_{stat}	*df*	*t Critical two-tail*	*Mean*	
Summer-Autumn	1,13462	204	1,97166	22,15	22,36
Autumn-Winter	1,06710	215	1,97105	22,37	21,82
Summer-Winter	1,74957	217	1,97095	22,15	21,82

at %95 confidence interval

As seen in Table 9., since the extreme values were removed from the temperature measurement data-set, $t_{statistical}$ < $t_{critical}$ was calculated in summer, autumn and winter and which means that the difference between the mean values is not significant and that the temperature mean values are within the acceptance limits at the 95% confidence interval ($t_{statistical}$=1,13462<$t_{critical}$=1,97166; $t_{statistical}$=1,06710<$t_{critical}$=1,97105;$t_{statistical}$=1,74957<$t_{critical}$=1,97095). Afterwards, the standard deviation values or %Sr values of July, August, September, October, November, and December were compared with the %Sr values calculated from the standard deviations of the acceptance limits given in the MSDS of the mentioned chemicals. Mean, σ_X and %Sr values for the months of July, August, September, October, November, December are given in detail in Table 10.

For this, $\%S_r = (\sigma_X / \bar{X}).100$ Eq.(5) is used;

Table 10. $\bar{X}$, σ_X ve $\%S_r$ values

Months	$\bar{X}$	σ_X	$\%S_r$
Temmuz	21,65	1,5	6,92
Ağustos	22,45	1,2	5,34
Eylül	22,85	1,1	4,81
Ekim	22,01	1,0	4,54
Kasım	21,95	1,2	5,46
Aralık	21,70	1,3	5,99

$\%Sr_{crt}$ = 7,90569415

When $\%S_r$, value, calculated for the validation process, was compared with the $\%S_r$, value given in the standart method in MSDS, $\%Sr_{statistical}$ < $\%Sr_{crtical}$ determined for the months of July, August, September, October, November, December. It has been validated that the %Sr value of the temperature values measured by months is smaller than the %Sr values, calculated from the standard deviations of the acceptance limits given in the MSDSs' of the mentioned chemicals.

4. Conclusion

Hygroscopic materials react to changes in air humidity in the environment of the storage area. If appropriate air humidity levels are kept appropriately in storage areas, these products can be stored with less weight loss and longer shelf life. In this study, temperature, %RH and WB values in the storage area, with the existing heating, cooling and ventilation system, were measured in July, August, September, October, November and December. In the statistical descriptive analysis of the data-set consisting of these measurement values, it was understood that the data did not present normal distribution due to the fact that containing about 120 extreme values. It has been observed that these extreme values, which called as outliers, are much higher than acceptance limits of 18-22ºC (ULV-LLV), reaching 26, 27, 28ºC in July and August, and lowering to 15-16ºC in October, November and December. The Mann-Whitney-U test also revealed that the temperature values have not been set to the acceptance limits in the presence of the existing heating, cooling and ventilation system in the storage plant because there are significant differences in morning and evening temperatures. For validation, it was also determined that $T_{calculated}$ and $F_{calculated}$ values remained within the standard acceptance limits after removing the extreme values and found that the %Sr values of the temperature values were smaller than the %Sr values calculated from the standard deviations of the acceptance limits given in the MSDS of the mentioned chemicals. The fact that, the standard deviation of temperature mean values, in the measurement data-set, at the beginning of the summer months and the beginning of the winter months, are higher which indicates that the storage plant environment should be more cared and taken precautions during these months. Relative humidity can change when the water vapor content of the air or the air temperature changes. Provided the water vapor remains constant, cooling the air increases the relative humidity, and heating the air decreases the relative humidity. In this study developped models revealed that the increase in %RH values is affected by combining normal temperature and Wet Bulb temperature values. Modeling, with the measurement values taken in the current system, revealed that in order to increase the humidity of hygroscopic chemicals, the wet-bulb temperature should be increased by decreasing the normal thermometer temperature in the existing heating, cooling and ventilating system of the storage plant. However, it is already known that high temperature destabilizes these chemicals. In this case, it is needed obviously a climation system keeping the temperature within the acceptance limits, and giving humidity to the environment of the storage plant at the same time. For this reason, steam humidification systemsm, such as an adiabatic humidification systems, whose values will be adjusted according to the formulas in the models developped in this study, are recommended for the companies engaged in similar activities.

Acknowledgements

The technical support of Bengi Chemical Industry Trade. Inc. is gratefully acknowledged. The author and the company declares no conflict of interest for the work reported in this paper.

REFERENCES

[1] "Hygroscopic compounds". hygroscopiccycle.com. IBERGY. Archived from the original on April 8, 2017. Retrieved April 7, 2017.

[2] https://www.fda.gov/, 2023

[3] https://www.ema.europa.eu/en, 2023

[4] Beşeri Tıbbi Ürünler İmalathaneleri İyi İmalat Uygulamaları (GMP) Kılavuzu Versiyon: 2022/03

[5] ISO/IEC 17025 General requirements for the competence of testing and calibration laboratories standard, 2017

[6] https://eiti.org/sites/default/files/2022-07/EITI%20Validation%20Guide.pdf, 2022

[7] "Systems and software engineering - Vocabulary," ISO/IEC/IEEE std 24765:2010(E), 2010. | verification 6. comply with requirements (e.g., for correctness, completeness, consistency, and accuracy) for all life cycle activities during each life cycle process (acquisition, supply, development, operation, and maintenance), 2010

[8] https://en.wikipedia.org/wiki/Verification_and_validation, 2023

[9]https://www.dqsglobal.com/intl/blog/verification-and-validation-why-the-new-iso-iec-17029-standard-is-so-important, 2023

[10] Ryan M. Wheatcraft L.S., On the Use of the Terms Verification and Validation INCOSE International Symposium 27(1):1277-1290 DOI: 10.1002/j.2334-5837.2017.00427.x, 2017

[11]https://www.nweurope.eu/media/8126/etv-guidance-document-summary-final_forflyer.pdf, 2023

[12] https://byjus.com/statistics-formulas/, 2023

[13]https://acikders.ankara.edu.tr/pluginfile.php/12180/mod_resource/content/1/Verilerin%20%C3%A7%C3%B6z%C3%BCmlenmesi%203.pdf, 2023

[14]http://ist-fen.omu.edu.tr/tr/hakkimizda/ders-notlari/TEMEL%20%C4%B0ST%20II-2018.pdf, 23

[15]https://bookdown.org/ugurdar/dogrusalregresyon/r-ile-basit-do%C4%9Frusal-regresyon-uygulamas%C4%B1.html, 2023

[16] https://en.wikipedia.org/wiki/Glycerol, 2023

[17]"L-glucose". Biology Articles, Tutorials & Dictionary Online. 2019-10-07. Retrieved 2022-05-06.

[18]"Fructose". m-w.com. Merriam-Webster. Archived from the original on 5 June 2013. Retrieved 10 December 2014.

[19]"16.3 Important Hexoses | The Basics of General, Organic, and Biological Chemistry". courses.lumenlearning.com. Retrieved 2022-05-06., 2023

[20]"Riboflavin". Drugs.com. 22 July 2021. Retrieved 8 October 2021.

[21]"Niacin". Micronutrient Information Center, Linus Pauling Institute, Oregon State University, Corvallis, OR. 8 October 2018. Retrieved 16 September 2019.

[22]"Folic Acid". Drugs.com. American Society of Health-System Pharmacists. 1 January 2010. Archived from the original on 8 August 2017. Retrieved 1 September 2016.

[23]"Folic Acid". The PubChem Project. Archived from the original on 7 April 2014.

[24]"Pantothenic acid". Linus Pauling Institute at Oregon State University. Micronutrient Information Center. 1 July 2015. Retrieved 27 November 2020.

[25]Haynes, William M. (2014). CRC Handbook of Chemistry and Physics (95 ed.). CRC Press. pp. 4–92. ISBN 9781482208689. Retrieved 18 November 2018.

[26]Tabachnick, B. G., & Fidell, L. S., Using Multivariate Statistics (6th ed.). Boston, MA: Pearson. 2013