

## BAYESIAN APPROACH TO MODELLING THE RELATIONSHIP BETWEEN ORANGE ROT RATE AND TRANSPORTATION DISTANCE

By

**Aondoakaa Jege**

Dept. of Mathematics and Computer Science,  
Benue State University, Makurdi.

&

**Terlumun, I. T.**

School of Science Education, Department of Mathematics,  
Federal College of Education (Technical), Gombe.

E-mail: [imanditerlumun@yahoo.co.uk](mailto:imanditerlumun@yahoo.co.uk)

### Abstract

*A Bayesian Logistic Regression Model relating the rot rate of oranges and transportation distance using incomplete data was developed. The incompleteness of data limits the use of the Classical Logistic Regression Model and this limitation informed our development of a Bayesian Statistical Simulation Modelling Procedure. This work used the simulation procedure through the Markov Chain Monte Carlo (MCMC) algorithm executed in WINBUGS. It is found that transportation distance has a linear and positive relationship with orange rot rate. For both rainy and dry seasons, a kilometre increase in transportation distance will increase the odds that oranges transported from Benue State to Gombe Fruit Market will be rotten. It was also discovered that the mean rot rate of oranges transported from Benue State during the rainy season is 19% and 13% during the dry season. The researcher noted that the distribution of the estimated unknown parameter ( $\theta$ ) converges to the posterior distribution quickly indicating model adequacy. The model also predicted the best distance for maximizing profitability and minimising losses.*

**Keywords:** Bayesian Model, Markov Chain, Monte Carlo, Orange, Transportation.

### Introduction

In Nigeria, about 3,900,000 tons of citrus fruits were produced from an estimated hectareage of 800000 hectares of land in 2011 (FAO, 2012). There are two main markets for citrus fruit in Nigeria: the fresh fruit market and the processed citrus fruits market (mainly orange juice). Most citrus production is accounted for by oranges, but significant quantities of grape fruits, lemons and limes are also grown. Total production and consumption of citrus has grown strongly since the 1980's. The rise in citrus production is mainly due to the increase in cultivation areas, improvements in transportation and packaging, rising incomes and consumer preference for healthy foods (UNCTAD, 2007). Major citrus producing states in Nigeria include Benue, Nasarawa, Kogi, Ogun, Oyo, Osun, Ebonyi, Kaduna, Taraba, Ekiti, Imo, Kwara, Edo, and Delta.

According to Olife et al. (2015), farm inputs (nurseries, fertilizers and agro chemicals) are at the bottom of the citrus value chain with end points in large processing and exporting companies and the local domestic market. Major actors in the value chain for industrialized nations include farmers and commercial orchard owners, citrus pickers, local fruit marketers and exporters, citrus processing factories, industries that engage in the utilization of by-products for the production of specialty products. However, the citrus value chain in Nigeria is not developed. Farm inputs are at the bottom of the value chain with end points in fruit juice processing companies and domestic fresh fruit markets. Major actors include plantation owners, family orchard owners, transporters, local fruit marketers, fruit juice processing

companies, eateries that engage in direct crushing (fresh juice production and selling).

Consequently, Nigeria produces millions of tons of seasonal fresh fruits, however, high post-harvest losses (30-60%) due largely to inadequate infrastructure and high-cost of transportation continue to create demand for imports. The post-harvest sector includes all points in the value chain from production in the field to the food being placed on a plate for consumption. Post-harvest activities include harvesting handling, storage, processing, packaging, transportation and marketing (Mrema and Rolle, 2002). Losses of horticultural produce are a major problem in the post-harvest chain. Lack of value addition is also a major constraint in harnessing the citrus potentials in Nigeria. Other challenges include lack of infrastructure for storage and preservation, lack of improved varieties, and poor handling of the fruits during harvesting and transportation (Olife et al. 2015).

### Method of Data Analysis

There are several paradigms for approaching statistical inference, but the two dominant ones are *frequentist* (sometimes called classical or traditional) and *Bayesian*. According to the Bayesian paradigm the unobservable parameters in a statistical model are treated as random. When no data are available, a *prior distribution* is used to quantify our knowledge about the parameter (Paul & Brani, 2007). When data are available, we can update our prior knowledge using the conditional distribution of parameters given the data. The transition from the prior to the posterior is possible via the Bayes theorem.

This research adopts Bayesian inference for data analysis. In this section Bayesian methods applied will be explained and models will be formulated. Markov Chain Monte Carlo simulations to calculate the posterior are distribution carried out in WinBUGS.

### Bayesian Methods

Spiegelhalter et al (2004) define a Bayesian approach as the explicit use of external

evidence in the design, monitoring, analysis, interpretation and reporting of a scientific investigation. They argue that a Bayesian approach is more flexible in adapting to each unique situation, more efficient in using all available evidence, more useful in providing relevant quantitative summaries than traditional methods. In Bayesian analysis, inference is based on probability statements summarizing the posterior distribution of the treatment effects.

### Bayes' Theorem and its link with Bayesian Inference

#### *Bayes' theorem Provable from probability axioms*

Let  $A$  and  $B$  be events, then

$$p(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

If  $A_i$  is a set of mutually exclusive and exhaustive events (i.e.  $p(\cup_i A_i) = \sum_i p(A_i) = 1$ ), then

$$p(A_i|B) = \frac{p(B|A_i)p(A_i)}{\sum_j p(B|A_j)p(A_j)}$$

Bayesian inference makes fundamental distinction between Observable quantities  $x$ , i.e. the data and unknown quantities  $\theta$  can be statistical parameters, missing data, mismeasured data (Chawla, 2016). Here, parameters are treated as random variables. In the Bayesian framework, we make probability statements about model parameters while in the frequentist framework, parameters are fixed non-random quantities and the probability statements concerning the data (Paul & Brani, 2007).

As with any analysis, we start by positing a model,  $p(x|\theta)$ . This is the **likelihood**, which relates all variables into a **full probability model**. From a Bayesian point of view,  $\theta$  is unknown so should have a **probability distribution** reflecting our uncertainty about it before seeing the data. We therefore need to specify a **prior distribution**  $p(\theta)$  and  $x$  is known so we should condition on it and use Bayes theorem to obtain conditional probability distribution for unobserved quantities of interest given the data:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)d\theta} \propto p(\theta)p(x|\theta) \quad (2)$$

Where  $p(x)$  is a normalizing constant,

$$p(x) = \int p(x|\theta)p(\theta)d\theta \quad (3)$$

The equation (3) above is the **posterior distribution**. Once the data  $x$  are available,  $\theta$  is the only unknown quantity and the posterior distribution  $p(\theta|x)$  completely describes the uncertainty. There are two key advantages of Bayesian paradigm: (i) once the uncertainty is expressed via the probability distribution and the statistical inference can be automated, it follows a conceptually simple recipe, and (ii) available prior information is coherently incorporated into the statistical model.

### Ingredients for Bayesian Inference

A crucial application of Bayes' rule is to determine the probability of a model when given a set of data. What the model itself provides is the probability of the data, given specific parameter values and the model structure. We use Bayes' rule to get from the probability of the data, given the model, to the probability of the model, given the data (Kruschke, 2010).

$p(x|\theta) = L(\theta)$  is called the likelihood. The functional form of the term is fully specified up to a parameter  $\theta$ . According to the likelihood principle, all experimental information about the data must be contained in this likelihood function. The parameter  $\theta$ , with values in the parameter space  $\Theta$ , is considered a random variable. The random variable  $\theta$  has a distribution  $p(\theta)$  called the prior distribution. This prior describes uncertainty about the parameter before data are observed. If the prior  $\theta$  is specified up to a parameter  $\tau$ ,  $p(\theta|\tau)$ ,  $\tau$  is called the hyper parameter. Our goal is to start with this prior information and update it using the data to make the best possible estimator of  $\theta$ . We achieve this through the likelihood function to get  $p(\theta|x)$ , called the *posterior* distribution for  $\theta$ , given  $X = x$ .

To find  $p(\theta|x)$ , we use Bayes' rule to divide joint distribution for  $x$  and  $\theta$  ( $h(x, \theta) = p(x|\theta)p(\theta)$ ) by the marginal distribution  $p(x)$ , which can be obtained by integrating out the parameter  $\theta$  from the joint distribution  $h(x, \theta)$ ,

$$p(x) = \int_{\Theta} h(x, \theta)d\theta = \int_{\Theta} p(x|\theta)p(\theta)d\theta \quad (4)$$

The marginal distribution is also called the prior predictive distribution. Finally, we arrive at an expression for the posterior distribution  $p(\theta|x)$ :

$$p(\theta|x) = \frac{h(x, \theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int_{\Theta} p(x|\theta)p(\theta)d\theta} \quad (5)$$

Where;

$p(x \theta)$	= Likelihood
$p(\theta)$	= Prior Distribution
$h(x, \theta) = p(x \theta)p(\theta)$	= Joint Distribution
$p(x) = \int_{\Theta} p(x \theta)p(\theta)d\theta$	= Marginal Distribution
$p(\theta x) = p(x \theta)p(\theta)/p(x)$	= Posterior Distribution

### Estimating Binomial Proportions

Bayesian statistics has become increasingly popular today, and one reason for its increased application is that it allows

researchers to input expert opinion as a catalyst in the analysis (through the prior distribution). Expert opinion might consist of subjective inputs from experienced professionals, or perhaps a summary

judgement of past research that yielded similar results.

Suppose each of  $n$  independent trial reveal either a successful or unsuccessful outcome. If  $\theta$  represents the proportion of success, let  $y$  be the number of successful outcomes in  $n$  independent trials.  $Y$  is distributed binomial

with parameters  $n$  (known) and  $\theta$  (unknown). We won't expect an expert to quantify their uncertainty about  $\theta$  directly into a prior distribution  $p(\theta)$ . The researcher can elicit information such as the expected value and standard deviation of  $\theta$ .

## Formulating Bayesian Model

### Likelihood function (Binomial)

$$L(n|\theta) = \binom{n}{y} \theta^y (1 - \theta)^{n-y} \quad (6)$$

### Prior

If we suppose the prior distribution for  $\theta$  is Beta( $\alpha, \beta$ ), where the hyper parameters  $\alpha$  and  $\beta$  are known, then;

$$p(\theta; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1}, 0 \leq \theta \leq 1 \quad (7)$$

With  $Y/\theta \sim \text{Bin}(n, \theta)$ , the joint, marginal and posterior distributions are

$$h(y, \theta) = \frac{\binom{n}{y}}{B(\alpha, \beta)} \theta^{\alpha+y-1} (1 - \theta)^{n-y+\beta-1}, 0 \leq \theta \leq 1, y = 0, 1, \dots, n \quad (8)$$

$$p(y) = \frac{\binom{n}{y} B(y+\alpha, n-y+\beta)}{B(\alpha, \beta)}, y = 0, 1, \dots, n \quad (9)$$

### Posterior

$$p(\theta/y) = \frac{1}{B(y+\alpha, n-y+\beta)} \theta^{\alpha+y-1} (1 - \theta)^{n-y+\beta-1}, 0 \leq \theta \leq 1 \quad (10)$$

$$\Rightarrow p(\theta|y) \sim \text{Beta}(\alpha + y, n - y + \beta).$$

After formulating the posterior, the mean rotten orange bags can be computed. This is shown in equation 11.

### Mean Rotten Orange Bags

From the beta distribution:

$$E(\theta) = \frac{\alpha}{\alpha + \beta} \quad (11)$$

$$\text{Variance}(\theta) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \quad (12)$$

Meanwhile, we can solve for  $\alpha$  and  $\beta$  as a function of the expected value  $\mu$  and variance  $\sigma^2$  as shown below:

$$\alpha = \frac{\mu(\mu - \mu^2 - \sigma^2)}{\sigma^2} \quad (13)$$

$$\beta = \frac{(1-\mu)(\mu - \mu^2 - \sigma^2)}{\sigma^2} \quad (14)$$

### Credible Intervals

From the posterior we can obtain any type of interval estimate we desire, i.e. one-tailed, two-tailed and any %. For example an equal-tailed  $10(1 - \gamma)\%$  credible interval is  $(a, b)$  such that

$$\gamma/2 = \int_{-\infty}^a p(\theta|y)d\theta = \int_b^{\infty} p(\theta|y)d\theta \quad (15)$$

For the Beta posterior, we have

$$\gamma/2 = \int_0^a \frac{\theta^{(\alpha+y)-1}(1-\theta)^{(\beta+n-y)-1}}{B(\alpha+y, \beta+n-y)} d\theta = \int_b^1 \frac{\theta^{(\alpha+y)-1}(1-\theta)^{(\beta+n-y)-1}}{B(\alpha+y, \beta+n-y)} d\theta \quad (16)$$

which is usually obtained using statistical software, in the case of this research, WinBUGS.

### Logistic Regression Model Relating Orange Rot Rate to Transportation Distance

Given a logistic regression model with a single covariate for a binary response Y. We denote the probability of success ( $Y = 1$ ) with  $p_i$  and probability of failure ( $Y = 0$ ) with  $1 - p_i$

That implies;

$$p_i = P(Y = 1|x_i) \quad (17)$$

$$1 - p_i = P(Y = 0|x_i) \quad (18)$$

The covariate  $x_i$  is the transportation distance,  $i = 1, \dots, 5$  represents the different transportation distances.

$$\text{logit}(p_i) = \alpha_i + \beta_i x_i \quad (19)$$

And  $y_i|p_i \sim \text{Bin}(n_i, p_i)$ , for  $i = 1, \dots, 5$

Where  $\alpha$  and  $\beta$  are the logistic regression parameters,  $p_i$  (success rate) is the orange rot rate.

The logit model has as a dependent variable, a transformation of  $p_i$  that represents the odds of the unknown probabilities (Damisa et. al., 2017). This element can be expanded as the probability of rotten orange bags  $p_i$ , over the probability of non-rotten orange bags ( $1 - p_i$ ).

The logistic transformation is defined as the log odds.

$$\text{Odds} = \frac{\text{probability of rotten orange bags}}{\text{probability of non - rotten orange bags}} = \frac{p_i}{1 - p_i}$$

Thus;

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right) \quad (20)$$

From equation (20)

$$p_i = \frac{1}{1 + e^{-(\alpha_i + \beta_i x_i)}} \quad (21)$$

Also from (20)

$$x_i = \frac{\text{logit}(p_i) - \alpha_i}{\beta_i} \quad (22)$$

Where equation (22) predicts the transportation distance required for a given orange rot rate.

### Centering the Transportation Distance at the Mean ( $\bar{x}$ )

Centering the transportation distance ( $x_i$ ) at the mean distance ( $\bar{x}$ ), for non-centering means zero transportation distance which is unrealistic. Centering implies that  $x_i - \bar{x} = 0$

Therefore, equation 21 becomes

$$p_i = \frac{1}{1 + e^{-(\alpha_i + \beta_i(x_i - \bar{x}))}} \quad (23)$$

If  $\alpha < 0$  then;

$$p_i = \frac{1}{1 + e^{\alpha_i}} \quad (24)$$

And this has a negative impact on  $p_i$

If  $\alpha > 0$  then;

$$p_i = \frac{1}{1+e^{-\alpha_i}} \quad (25)$$

This has a positive impact on  $p_i$ .

### The above model in WinBUGS syntax:

In WINBUG syntax, we fit the Bayesian Logistic Regression Model with centered covariate as follows;

```
Model {For (i in 1: k) { y[i] ~ dbin(p[i],n)logit(p[i]) < - alpha[i] + beta[i] *
(x[i] - mean(x[]))alpha[i]~ dnorm(0,0.01) beta[i]~ dnorm(0,0.01) }}
```

Where the data list of number of rotten bags of oranges  $y[]$  and transportation distance (in kilometers)  $x[]$  as well as the initialization list for the model parameter arrays;  $\alpha[]$  and  $\beta[]$  are defined for each collection points of oranges  $i$ .  $k$  is set as the number of collection points for the oranges before being transported to Gombe while  $n$  the number of bags transported is set at 110 for the rainy season (scarce supply) and at 330 for the dry season (full supply).

Note that WINBUGS uses the equation (25) in computing the rate of rotten orange bags ( $p_i$ ) for a collection point of raw oranges  $i$ . This gives the simulated value of  $p_i$  (Ntzoufras, 2009).

### Estimation of an Unknown Transportation Distance with known probability of Success

An unknown transportation distance can be predicted using the model;

$$\text{logit}(p_i) = \alpha_i + \beta_i (ED - \bar{x})$$

$$ED = \frac{\text{logit}(p_i) - \alpha_i}{\beta_i + \bar{x}} \quad (26)$$

From the (22) above, the equation for predicting the transportation distance that will give us a maximum orange rot rate of 5% is given below;

$$ED5 = (\text{logit}(0.95) - \alpha) / \beta + \bar{x} \quad (27)$$

Where ED5 is the predicted distance that can give a maximum of 5% rate of rotten oranges.

In WinBUGS, simply add this code into the BUGS model code:

```
ED5 <- ((logit (0.95) - alpha[i])/beta[i]) + mean(x[])
```

### Bayesian Logistic Regression Model Parameters Estimation

The result of this study includes the actual values of the true orange rot rate  $p_i$  for each of the transportation distances from Benue State to the Gombe fruit main market in Gombe State. The results also include mean rate of rotten orange bags for the two seasons (rainy and dry). The results further capture the model parameter and measure values for each distance ( $x_i$ ). Estimation of ED5, a predicted distance that will keep maximum number of rotten orange bags at 5 %.

**Table 1:** Distribution of actual orange rot rate by transportation distance and seasons of purchase

Season						
Season		Rainy Season		Dry Season		
	Location	Distance (Km)	Rot rate		Distance (Km)	Rot rate
	Gboko	680	0.09	Gboko	680	0.05
	Tyowanye	695	0.10	Tyowanye	695	0.06
	Katsina Ala	720	0.13	Katsina Ala	720	0.06
	Vandeikya	768	0.17	Vandeikya	768	0.12
	Zaki-Biam	770	0.16	Zaki-Biam	770	0.10

**Table 2:** Parameter estimates, measure values and 95% credible regions for the logistic model fitted to the dry season data.

node	Mean	sd	MC error	2.5%	median	75.0%	97.5%	start	sample
alpha[1]	0.05722	9.999	0.1471	-19.58	0.1274	6.788	19.59	100000	100001
alpha[2]	-0.1763	9.896	0.1557	-19.52	-0.2724	6.44	19.43	100000	100001
alpha[3]	-0.3025	9.847	0.1673	-19.64	-0.2674	6.241	19.22	100000	100001
alpha[4]	-0.3926	10.06	0.1879	-19.99	-0.3984	6.356	19.24	100000	100001
alpha[5]	-0.1608	9.977	0.1619	-19.69	-0.1622	6.597	19.38	100000	100001
beta[1]	0.06243	0.2146	0.003156	-0.3591	0.06413	0.2068	0.4819	100000	100001
beta[2]	0.07954	0.3132	0.004929	-0.5328	0.07639	0.289	0.7001	100000	100001
beta[3]	0.2878	1.492	0.02534	-2.644	0.2914	1.28	3.245	100000	100001
beta[4]	-0.0392	0.243	0.004539	-0.5135	-0.03918	0.1247	0.4342	100000	100001
beta[5]	-0.04973	0.2299	0.003731	-0.5003	-0.04974	0.1062	0.4004	100000	100001
ed5[1]	657.4	11710.0	37.03	478.7	691.9	703.8	880.7	100000	100001
ed5[2]	704.5	1380.0	4.502	567.0	703.5	711.4	835.8	100000	100001
ed5[3]	721.5	796.3	2.54	699.7	723.9	725.3	746.9	100000	100001
ed5[4]	817.4	7329.0	23.12	639.0	766.0	782.1	901.4	100000	100001
ed5[5]	770.3	2759.0	8.756	613.0	766.7	786.3	929.8	100000	100001
p[1]	0.05502	0.007203	2.171E-5	0.04173	0.05476	0.05973	0.06983	100000	100001
p[2]	0.06401	0.007712	2.165E-5	0.04974	0.06374	0.06904	0.07998	100000	100001
p[3]	0.09999	0.009488	3.194E-5	0.08217	0.09971	0.1063	0.1194	100000	100001
p[4]	0.118	0.01017	3.105E-5	0.09875	0.1177	0.1247	0.1386	100000	100001
p[5]	0.08996	0.009027	3.047E-5	0.07308	0.08967	0.09588	0.1084	100000	100001

Note (i) The numbers 1 – 5 indicate the rot rate for each distance (estimated p)

(ii) Model parameters are alpha and beta.

**Table 3:** Parameter estimates, measure values and 95% credible regions for the logistic model fitted to the rainy season data.

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha[1]	0.0567	10.04	0.1677	-19.63	0.06931	19.73	100000	100001
alpha[2]	-0.09688	10.14	0.1893	-20.0	-0.07327	19.7	100000	100001
alpha[3]	-0.4401	9.843	0.2056	-20.26	-0.3396	18.71	100000	100001
alpha[4]	-0.3349	9.88	0.2365	-19.31	-0.4824	19.25	100000	100001
alpha[5]	-0.07126	9.959	0.2121	-19.56	-0.04652	19.43	100000	100001
beta[1]	0.05072	0.2154	0.003598	-0.3717	0.05089	0.4731	100000	100001
beta[2]	0.06661	0.3209	0.00599	-0.5631	0.06733	0.6935	100000	100001
beta[3]	0.2259	1.491	0.03116	-2.778	0.2416	3.128	100000	100001
beta[4]	-0.02977	0.2387	0.005712	-0.5027	-0.02607	0.4281	100000	100001
beta[5]	-0.03595	0.2295	0.004886	-0.4852	-0.0366	0.4136	100000	100001
ed5[1]	675.6	2079.0	6.464	519.1	689.8	850.4	100000	100001
ed5[2]	657.7	8214.0	25.9	590.3	702.2	801.4	100000	100001
ed5[3]	725.1	731.0	2.303	702.2	723.7	743.8	100000	100001
ed5[4]	782.6	2494.0	7.799	666.7	767.4	878.5	100000	100001
ed5[5]	737.8	8755.0	27.64	661.3	768.7	880.9	100000	100001
p[1]	0.09098	0.00908	3.026E-5	0.07404	0.09072	0.1095	100000	100001
p[2]	0.09999	0.009518	2.956E-5	0.08199	0.09972	0.1194	100000	100001
p[3]	0.127	0.0105	3.469E-5	0.1071	0.1267	0.1481	100000	100001
p[4]	0.1729	0.012	3.651E-5	0.1501	0.1727	0.1973	100000	100001
p[5]	0.164	0.01171	3.611E-5	0.1417	0.1638	0.1875	100000	100001

Note (i) The numbers 1 – 5 indicate the rot rate for each distance (estimated p)

(ii) Model parameters are alpha and beta.

**Table 4:** Estimation of the distance that will give minimal orange rot rate of at most 5%**STATISTICS FOR ED5****NODE**

Node	Mean	Sd	MC error	2.5%	Median	97.5%	Start	Sample
ED5	580.2	3546.0	15.21	105.2	600.3	672.6	10001	53000

Model parameters and measure values for each season are captured on tables 2 and 3. Details on these tables include the mean, standard deviation, Monte Carlo Simulation Error, median and 95 % credible interval. Estimates of the model parameters (alpha and beta) and their measures (orange rot rate ( $p$ )) and their respective mean values. The values of alpha and beta for each orange season helps to determine the relationship between orange rot rate and transportation distance. This is achieved when they are plugged into equation (27). Observe the close values of the actual orange rot rate ( $p$ ) on table 1 for the two seasons and their respective simulated values  $p[i]$  on tables 2 and 3. This also validates the model.

The actual values of the proportion of rotten orange bags (rot rate  $p$ ) across the seasons in each of the five (5) transportation distances are captured in table 1. As shown in this table, the seasons include the rainy – season, and the dry – season. These values were computed from actual data on the number of rotten orange bags per truck for each transportation

distance across the two orange seasons. The number of rotten orange bags per 110 bags (one truck) were determined for each distance in each season and used as number of trials for the Binomial distribution in the course of the Bayesian statistical modelling. The transportation distance (in Km) data sets also serve as input to the model.

Looking at the different transportation distances from where the orange marketers get their goods, the Bayesian logit equation for modelling the relationship between orange rot rate and transportation distance is given as;

$$p_1 = \frac{1}{1 + e^{-[0.0567 + 0.05072(x_1 - \bar{x})]}} \quad (28)$$

Where  $p_1, \dots, p_5$  are orange rot rates for each transportation distance from where the oranges were obtained. The above equation can be applied to all the distances and for both the dry and rainy seasons to predict orange rot rate. This model can provide profitable insights to the marketers on when and where to acquire their products in order to minimize losses.

## Reference

- Chawla, V. (2016). A Bayesian network approach to county-level corn yield prediction using historical data and expert knowledge. *Graduate Theses and Dissertations*. 15893. <https://lib.dr.iastate.edu/etd/15893>.
- Damisa, S.A., Musa, F. & Sani, S. (2017). On the Comparison of Some Link Functions in Binary Response Analysis. *Nigeria Statistical Society. Vol.1*.
- FAO (2012). Major Fruits and Vegetable Producing Countries in the World 2010 and 2011.
- Kruschke, J. K. (2009). Doing Bayesian data analysis- A Tutorial with R and WinBUGs Elsevier Academic Press.
- Mrema, C. G. & Rolle, S. R. (2002). Status of the postharvest sector and its contribution to agricultural development and economic growth. *9th JIRCAS International Symposium – Value Addition to Agricultural Product*, pp. 13-20.
- Ntzoufras, I. (2009). Bayesian Modelling Using WinBUGS. *John Wiley & Sons, Inc., Hoboken, N.J.* <https://doi.org/10.1002/9780470434567>.
- Olife, I.C., Ibeagha, O.A. & Onwualu, A.P. (2015). Citrus Value Chain Development in Nigeria. *Journal of Biology, Agriculture, and Healthcare. Vol. (5)4*.
- Spiegelhalter, D. J., Freedman, L. S., & Parmar, M. K. B. (1994). Bayesian approaches to randomized trials. *Journal of the Royal Statistical Society. Series A*, 157, 357–416.
- UNCTAD (2007). Market Information in the Commodities Area: Citrus Fruit.