



# **Regression-based Modeling of the Relationship between Weather and Tea Production in Sri Lanka**

**M.D.C.De.S.Jayatilake**  
(Reg. No.: MS20902216)

**M.Sc. in IT**

**Supervisor: Dr.Windhya Rankothge**  
**Assistant Professor**  
**Computer Systems Engineering Faculty of Computing**

**December, 2021**

**Faculty of Graduate Studies**  
**Sri Lanka Institute of Information Technology**

## Acknowledgements

Leading, I would like to convey my sincere gratitude to my supervisors Dr. Windhya Rankothge, Assistant Professor, Computer Systems Engineering Faculty of Computing, Sri Lanka Institute of Information Technology and Dr. Dharshana Kasthurirathna, Assistant Professor, Computer Science & Software Engineering Faculty of Computing, Sri Lanka Institute of Information Technology (SLIIT) for the continuous help of my MSc research and study. Supervisors endurance, motivation, keenness, and enormous of knowledge guidance helped me a lot in all the time of research.

My sincere thanks also go to the Mr. Samantha Rajapaksha, MSc Coordinator (IT), Senior Lecturer, Information Technology Faculty of Computing, Sri Lanka Institute of Information Technology. I would express my special great to Prof. Rahula Attalage, Dean - Faculty of Graduate Studies & Research.

I would like to thank to my loving parents and family members for giving birth to me at the first place and behind me spiritually throughout the life. Last but not the least; unfortunately, my Father was passed away 08<sup>th</sup> April 2021 during my research.

Further, I would like to greatly thankful to my loving wife, Thilina Pinidiya who had introduce me the first glance of SLIIT - MSc degree and my daughter, baby Mithuki Methara for encouraging me to achieve this journey success-fully.

# Table of Contents

Acknowledgements.....	ii
Table of Contents.....	3
List of Figures .....	6
List of Tables .....	7
Abstract.....	8
Chapter 1 Introduction .....	9
Tea Production Sri Lanka – January to September 2021 .....	9
Tea export of Sri Lanka (January-September 2021).....	10
1.1 Background of Research .....	11
1.2 Significant of the research study.....	11
Tea production sector - Main Tea Market Sri Lank -2020 .....	12
The percentage Tea contribution total merchandise expert in 2011-2021/August.....	12
1.3 UVA Province, Sri Lanka .....	13
Tea estates and gardens .....	15
Best UVA Teas of Sri Lanka.....	15
1.4 Challenges Tea Production.....	15
Improvements of the productivity.....	16
1.5 Present and Future Tea industry in Sri Lanka .....	16
1.6 Problem Definition and Research Question.....	18
1.7 Research Objectives and Scope .....	18
1.8 Summary of the introduction.....	19
Chapter 2 Literacy Survey .....	20
2.1 Exiting Research and Present Situation .....	20
2.2. Previous research finding with conclusions.....	20
2.2.1. Sri Lankan Tea Sector at Present.....	20
2.2.2 Small business holders in the remote areas in Sri Lanka and their issues and challenges .....	22
2.3.3 Impact of the climate variability on Tea Production in Sri Lanka .....	22
2.3 Previous research finding with other countries.....	23
2.3.1 Regression predict models for tea yield responses with climate parameters.....	23
2.3.6 Tea, moisture detection techniques on convolution neural network .....	27
2.3.7 UAV image data usage of the nitrogen application on tea quality monitor.....	27
2.3.8 Counterfeit climate effects .....	28
2.3.9 Sustainability assessment of tea farming with the effect of decisional and structural factors.	28

2.3.10 Sustainability assessment of tea farming with the effect of decisional and structural factors	29
Chapter 3 Research Methodology .....	30
3.1 Proposed Solution.....	30
3.1.1 Study Area- UVA Province, Sri Lanka .....	30
3.1.2 Random Forest.....	31
3.1.3 Working process of the Random Forest .....	31
3.1.4 Multiple Linear Regressions.....	32
Chapter 4 Methodology.....	34
4.1 Work flow for the Methodology .....	34
Chapter 5 Data Collection Methods of Research.....	35
5.1 Methods and the Materials of the research.....	35
5.1.1 Data collection of the Research .....	35
5.1.2 Sample data set- format-excel.csv file (i.e. training the prediction model) .....	36
5.1.3 Dependent Variable- Tea Yield data sample for the UVA province with eight Agro climate districts.....	37
Chapter 6 Machine Learning Implementation.....	38
6.1 Algorithm description of the prediction model .....	38
6.2 Results of the prediction model (RF, SVM and MLR) for region wise in UVA province Sri Lanka ....	38
Annex 1- A.....	39
6.2.1 Negative R square (score) value .....	39
6.3 Trainee accuracy of the main function implementation for the whole data set (Random Forest model).....	41
6.4 Final Results of the Tea-production prediction model .....	42
Appendix 1: Appendix B.....	44
Final Results of the Tea-production prediction variance training model accuracy .....	44
Final Results of the Tea-production variance accuracy of prediction model .....	45
Chapter 7 Evaluation of the Results.....	46
6.1 Regression Analysis.....	46
6.1.1 Random Forest application for the Haputhale region .....	46
Annex III-A- statistics analysis details attached .....	47
6.1.2 Variable importancy of the regression model .....	47
6.2 Relationship between climate parameters and the tea production .....	48
6.3 Regression Analysis with climate parameters .....	49
6.3.1 Liner regression model tea production variability of the climate parameters.....	49
Chapter 8 Regression Analysis for all 8 Areas of the UVA province .....	52
Appendix 3,4,5,6,7,8:- .....	71

Chapter 9 Conclusion .....	72
Chapter 10 Bibliography .....	74
Annex 1- A.....	78
Appendix 2: Appendix –D1- .....	84
Appendix 3: Haputhale .....	84
Appendix 4: Appendix-3 D3- Statistical analysis results with coding of the “Climate parameters Vs. Tea Production” graphs. (Bandarawela/ Poonagala agro climate distric) .....	90
Appendix 5: D4-Malwatte/Welimada agro climate distric (Climate parameters Vs. Tea Production)...	94
Appendix 6: Appendix3- D5- Statistical analysis results with coding of the Demodara/Haliells/Badulla agro climate distric (Climate parameters Vs. Tea Production) .....	97
Appendix3-D7- .....	103
Appendix3_D8- .....	106

## List of Figures

Figure 1-1 - Export Tea Performance 2011-2021-August[16].....	11
Figure 1-2: “percentage of tea contribution” [16].....	12
Figure 1-3-“Value of Exports Tea” .....	17
Figure 1-4 –‘World export tea’ .....	17
Figure 3-1Agro climate districts-Elevation.....	30
Figure 3-2 -UVA Province, Sri Lanka.....	31
Figure 3-3-RF working progress .....	32
Figuer 4-1- work flow of the methodology.....	34
Figure 5-1- training data sample.....	36
Figure 5-2 - Validation data sample .....	37
Figure 5-3 - Tea production data sample.....	37
Table 6-1- region wise prediction .....	38
Figure 6-2 -graphical representation .....	39
Figure 6-3 - liners regression repretation .....	40
Figuer 7-1 - Decision tree in RF .....	46
Figuer 7-2-graphical representation error .....	47
Figuer 7-3-tea production relationship with climate parameters .....	49
Figuer7-4 – Liner regresion analysis of the tea podution .....	51

## List of Tables

Table 1-1-January-September 2021 – MT.....	9
Table 1-2 Export figures from January to September 2021' .....	10

## Abstract

The research study shows the implementation of regression-based tea–weather prediction models for the tea production in Sri Lanka based on six weather parameters such as rainfall, , temperature minimum, temperature maximum, relative humidity, average wind speed, and sunshine hours. The statistics of the eight agro climate geographical regions of the UVA province, which contribute to the more productive contribution tea production of the country. They were used for this research and the significance of weather parameters on the tea production was explored by employing Random Forest algorithm and the variable importance of each of them was determined and further correlation in the middle of the climate parameters was examined. The results indicate that the Minimum Relative Humidity, Rainfall and the maximum Temperature throughout the tea plantation period are the most influential weather parameters. Furthermore, regression analysis implementations were applied for the Random Forest (RF), Linear Regression (LR), Multiple Linear Regression (MLR), and Support Vector Machine (SVM). According to the results, RF is the most reliable and accurate model for the prediction of tea production in Sri Lanka. UVA province prediction model accuracy (R square score) is 88.58% of the eight agro climate districts of the UVA province and region-wise prediction tea-production model accuracy is low parentage according to the results. Further RF, MLR and SVM, Machine Learning implementation already trained and validated for the same dataset. Although the results indicated that other models low percentage in comparison to the RF implantation model. The research regression analysis already applied for RF and LR for the region-wise of the UVA province Sri Lanka. Concluded that the final output of the research same process can be applied to the tea- weather prediction model of all the tea growing areas of the country.

Key words: Agro climate, Regression model, Random Forest, Tea production, Tea-weather