Laying the groundwork: A Preliminary General-Soil Based Model for Abaca Crop Productivity

UA&P
UNIVERSITY OF ASIA AND THE PACIFIC

Globe

Kevin Mikhail P. Atienza

A Legacy

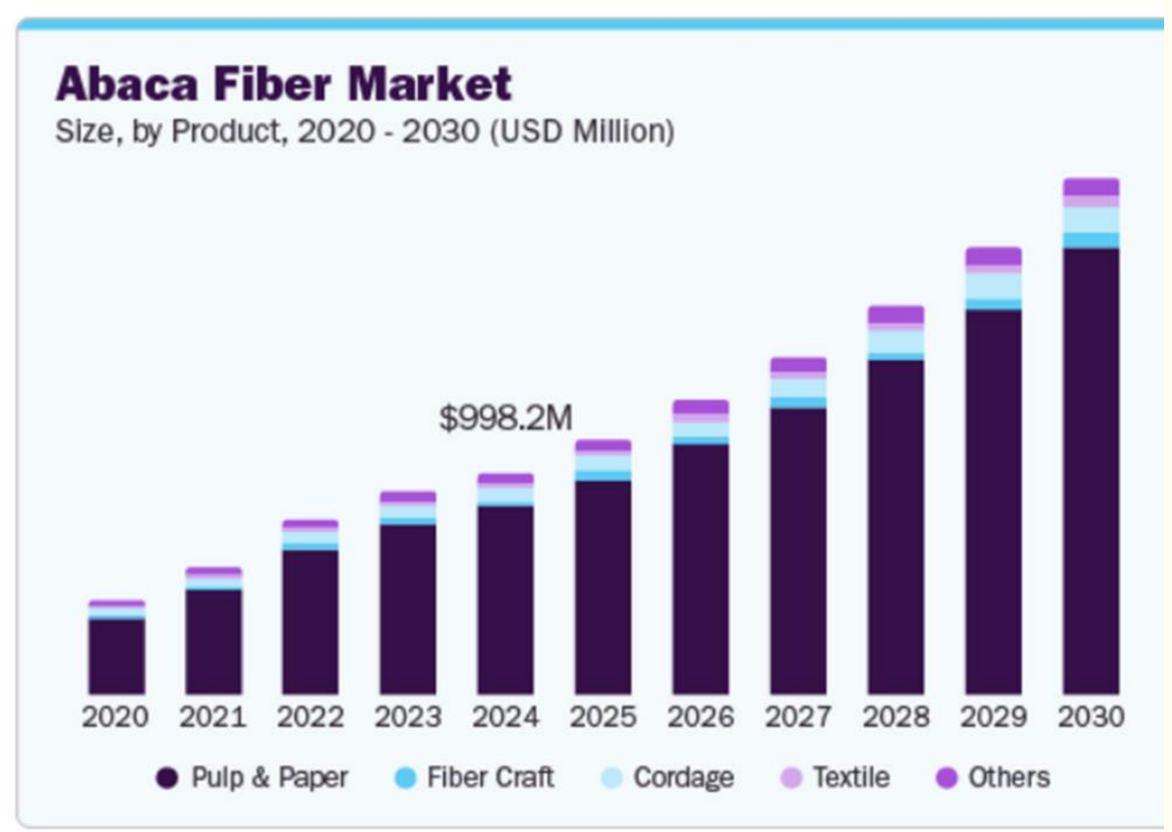
Philippine Fiber Industry Development Authority (PhilFIDA) is an attached agency of the Department of Agriculture. It's mandate is to promote the Philippine natural fiber industry through programs and policy guidelines.

- Research and Development
- Production and Extension Support
- Standards and Regularization
- Policy and Program Formulation
- Collaboration

Introduction - Market

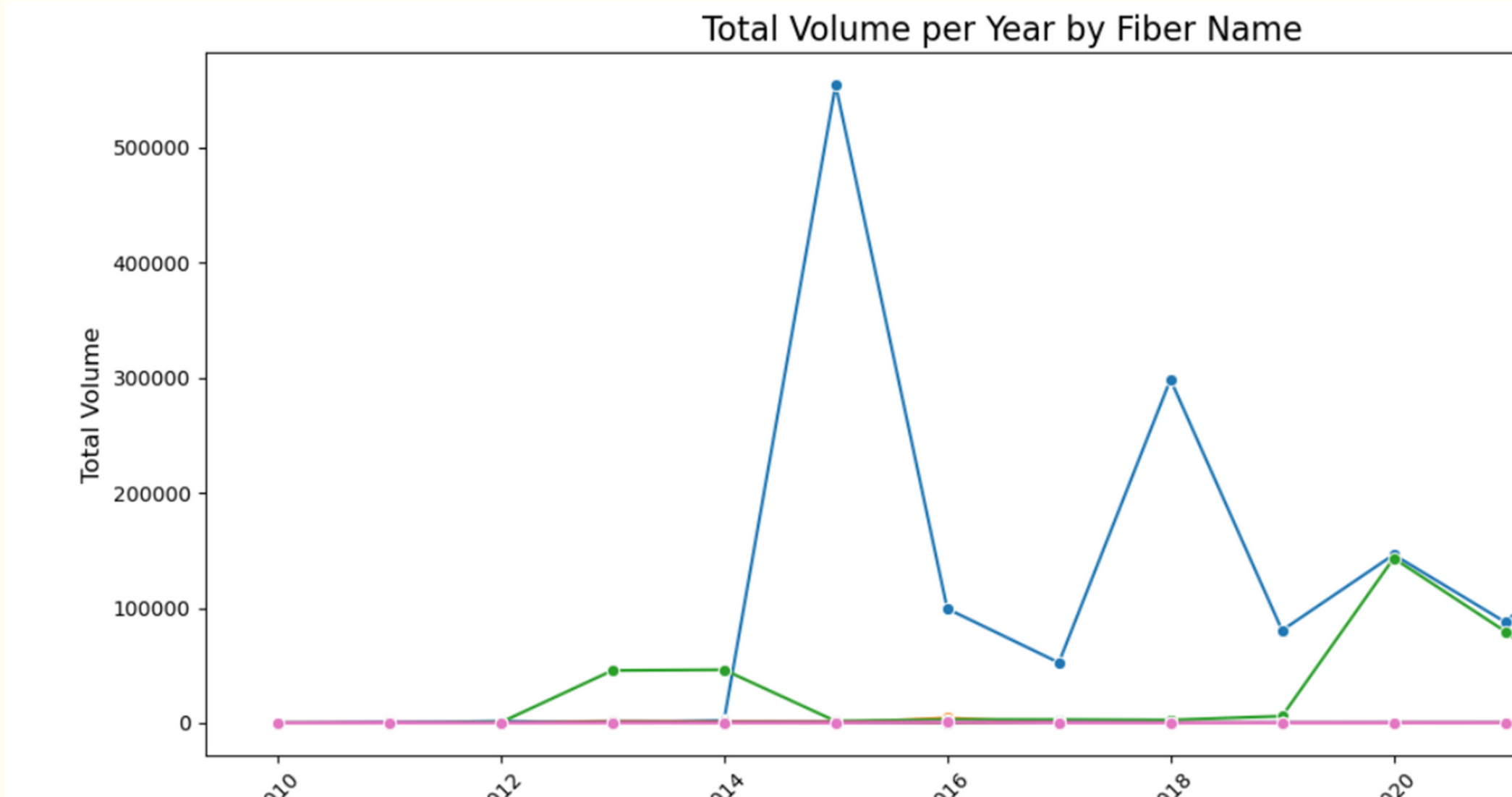
Grand View Research expects that the Abaca Fiber Market to expand at a 15% CAGR in the next 5 years amounting to $2.32B or Php 132.2B. With the country accounting for more than 72% of the global market share in 2024.

This growth was propelled by the consistent demand for specialty pulp and paper sector, and the increasing demand in natural fibers for sustainable practices. Additionally, applications that goes beyond the traditional markets like automotive industries and textiles.



Do we have sufficient research and studies that correlates Soil and Weather to Abaca Productivity?

Yes, but most of this studies are focused and done mostly as descriptive analytics, rather than predictive and prescriptive analytics. In addition, with the growing uncertainty of climate events such as typhoons, the lack of predictive analytics puts the industry at risk to lose and maintain its competitive dominance in terms of volume and/or quality to other countries such as Ecuador and Costa Rica.


Total Volume per Year by Fiber Name
Total Volume
500000
400000
300000
200000
100000
0

Objectives

The capstone project aimed to provide actual data analytics that could correlate soil and weather to the actual production of abaca.

It also aims to lay the groundwork for future predictive and prescriptive studies for the agency and other institutions as well as find potential areas in which private organizations can help the agency and the farmers with data and information that may or may not be available to them.

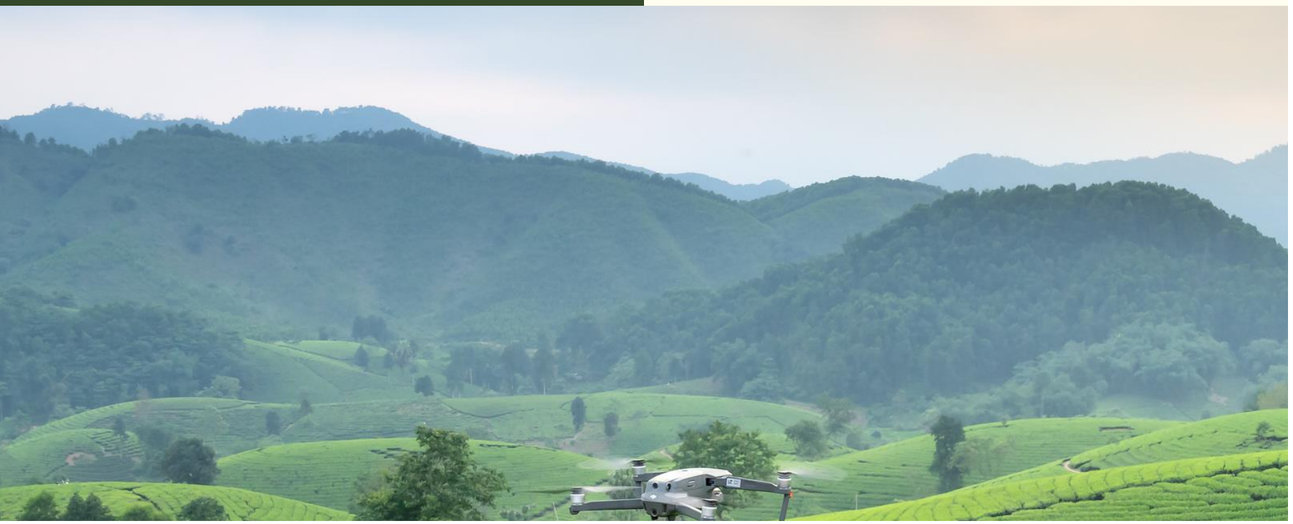
Data & Methodology

Data Sources

PAGASA

BSWM

PSA

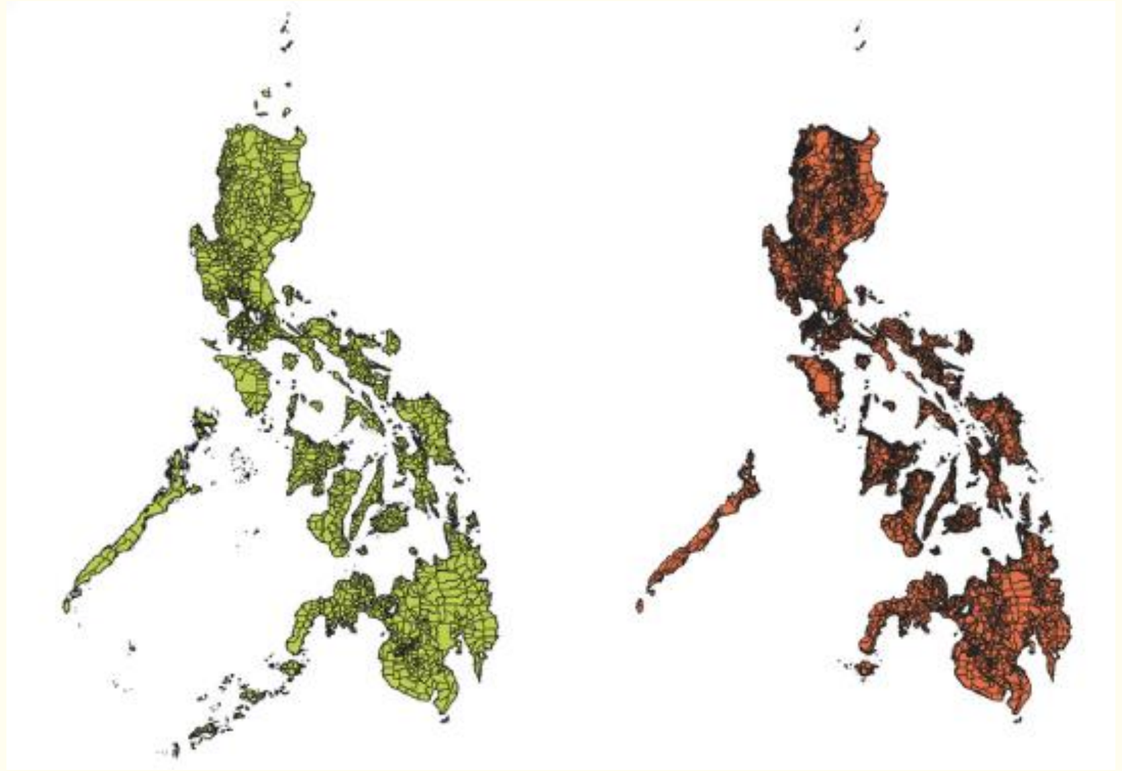


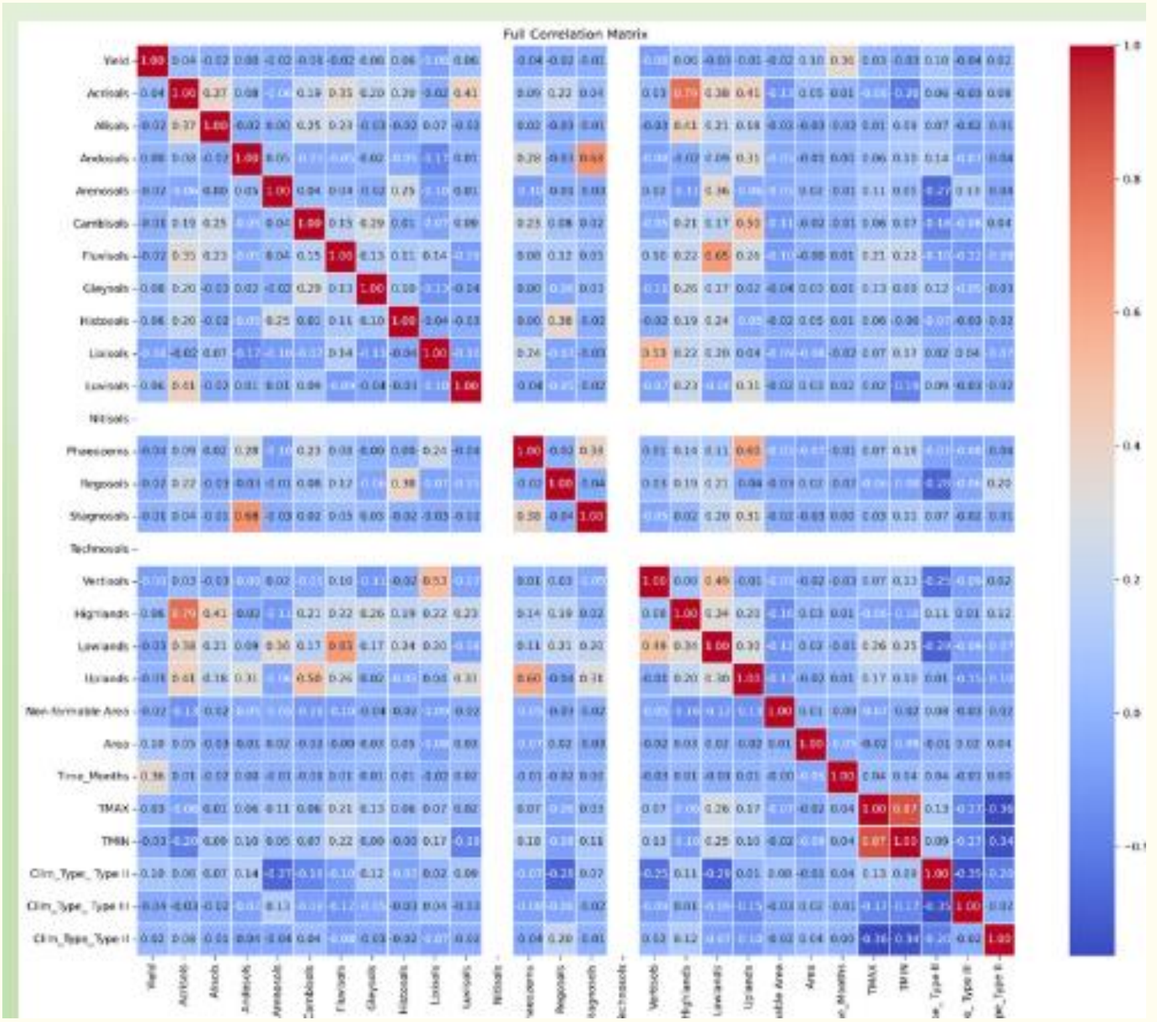
Challenges

01 Granularity - Provincial

02 Data Completeness

03 Lack of field validation




Full Correlation Matrix
Yield
Acrisols
Alisols
Andosols
Arenosols
Cambisols
Fluvisols
Gleysols
Histosols
Lixisols
Luvisols
Nitisols
Phaeozems
Regosols
Stagnosols
Technosols
Vertisols
Highlands
Lowlands
Uplands
Non-Arrivable Area
Area
Time_Months
TMAX
TMIN
Clim_Type_ Type II
Clim_Type_ Type III
Clim_Type_ Type IV
1.0
0.8
0.6
0.4
0.2
0.0
−0.1

01 Linear Regression

02 Decision Tree

03 Random Forest

04 XGBoost

Hanover Agriculture

Data Models used
Regression Models

Regression Models were used since we're trying to predict Yield

Based on other studies that aimed at predicting yield these models have had the best results in terms of metrics or scores.

Features

01 **Soil Data**

02 **Yield Data**

03 **Weather Data**

Soil data was from BSWM via shapefiles which were then used to interpolate with Administrative Boundaries to get the soil data distribution per province.

Yield data was from PSA which was based on export data from the Bureau of Customs.

Weather data had to be imputed as there were only less than 60 provinces that were available from PAGASA's database from 2010 to 2023

Results

Table 7.1 *Model Evaluation – With Climate TMIN and TMAX imputed*

	MAE	MSE	RMSE	R^2
Linear Regression	0.1317	0.0435	0.2085	0.1110
Decision Tree	0.0801	0.0222	0.1490	0.5459
Random Forest	0.0633	0.0149	0.1220	0.6955
XGBoost	0.0714	0.0187	0.1369	0.6170

Table 7.2 *Model evaluation – Without Climate TMIN and TMAX*

	MAE	MSE	RMSE	R2
Linear Regression	0.1405	0.0565	0.2377	0.0994
Decision Tree	0.0760	0.0206	0.1437	0.6710
Random Forest	0.0659	0.0152	0.1231	0.7584
XGBoost	0.0701	0.0174	0.1320	0.7221

- Random Forest

Feature	Importance
Time_Months	0.453205
Area	0.376882
Lowlands	0.025200
Arenosols	0.017914
Phaeozems	0.015049
Non-farmable Area	0.013485
Lixisols	0.012741
Gleysols	0.012583
Highlands	0.011588
Acrisols	0.011397

A good starting point for future studies

Narrowed down potential features to improve prediction

Insightful for policy makers as well as data scientists

Time_Months can be further broken down into events such as typhoons, policy changes, replanting among others.

THANK YOU!

kevin.atienza@globe.com.ph