

CLIMATE-DRIVEN VARIABILITY IN SUGARCANE PRODUCTIVITY: A DATA-DRIVEN ASSESSMENT FOR SÃO PAULO STATE (BRAZIL)

Giulia Cruz Lamas⁽¹⁾, Normando Perazzo Barbosa Sout⁽¹⁾, Igor F. Mendonça⁽¹⁾, Marcela Rodrigues Machado⁽¹⁾, Ana Filipa Ferreira⁽²⁾, Edgar Amaral Silveira⁽¹⁾

⁽¹⁾University of Brasília, Mechanical Sciences Graduate Program, Laboratory of Energy and Environment, Brasília-DF 70910-900, Brazil

⁽²⁾IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

Keywords: Data integration, Agricultural modeling, Sugarcane productivity, Climate variability, Bioenergy systems, Life Cycle Assessment

Abstract

Sugarcane is one of the most important bioenergy feedstocks worldwide, playing a central role in ethanol production and emerging sustainable aviation fuel (SAF) pathways. In Brazil, particularly in São Paulo State, sugarcane productivity is strongly influenced by climatic variability, which introduces significant uncertainty into biomass supply chains and life cycle assessment (LCA) results. Despite the relevance of climate–yield interactions, most LCA and process optimization studies rely on static or averaged productivity values, neglecting temporal and spatial variability. This simplification may lead to biased carbon intensity estimates and suboptimal decision-making in bioenergy systems.

This study proposes a data-driven framework to assess the influence of climatic variables on sugarcane productivity, focusing on municipal-level analysis in São Paulo State. A comprehensive panel dataset was constructed by integrating agricultural productivity data (tonnes per hectare) with meteorological variables, including precipitation, global horizontal irradiation, temperature, and wind speed. Climate data were obtained from meteorological station networks and satellite-derived datasets, while agricultural data were sourced from official statistics. The dataset was processed to ensure temporal consistency and spatial alignment across municipalities.

Exploratory statistical analyses are conducted to examine potential relationships between climatic variables and sugarcane yield. Particular attention is given to identifying patterns of association, potential nonlinear behaviors, and region-specific responses. The role of precipitation, global horizontal irradiation, photosynthetically active radiation, air temperature, and wind speed is being evaluated in order to better understand their combined influence on crop productivity.

In addition, Python-based data processing workflows are employed to integrate and harmonize heterogeneous data sources, including meteorological and agricultural datasets. The objective is to structure and standardize the dataset in a way that

enables its use in future machine learning applications. This preparation step includes data cleaning, temporal alignment, feature organization, and consistency checks across variables. From this perspective, the study establishes the foundation for subsequent application of machine learning techniques to capture complex interactions among climatic drivers and improve yield prediction.

From a systems perspective, the proposed framework enables the integration of climate-driven productivity variability into LCA and biofuel pathway assessments, mainly SAFs. By replacing static yield assumptions with dynamic, data-informed estimates, it becomes possible to better represent temporal fluctuations in biomass availability and associated environmental impacts. This approach is particularly relevant for regulatory frameworks such as RenovaBio and international programmes like CORSIA, where accurate carbon intensity calculations are essential.

The study contributes to advancing dynamic modeling approaches in bioenergy research and provides a foundation for integrating climate data, agricultural productivity, and life cycle modeling. The methodology is designed to be adaptable to other regions and feedstocks, supporting more robust sustainability assessments under varying climatic conditions.