Machine learning framework for evaluating climate change impacts on maple syrup yield in Quebec

Marco Alberto Nanfouet[1], Terence Epule Epule[1*], Cokou Patrice Kpadé[1], Simon Lafontaine[1], Liliana Perez[2], Changhui Peng[3]

[1] Institute for Research in Agriculture and Agri-Food (IRAA), Université du Québec en Abitibi-Témiscamingue (UQAT), Notre Dame du Nord, Québec, Canada

[2] Department of Geography, Université de Montréal, Montréal, Quebec, Canada

[3] Department of Biological Sciences, Institute of Environnemental Sciences, Université du Québec à Montréal (UQAM), Montréal, Québec, Canada

Terence Epule Epule (Corresponding author) *terenceepule.epule@uqat.ca

Abstract

Maple syrup production is an important cultural and economic activity in Quebec and contributes substantially to the livelihoods of maple producers. However, production has experienced considerable interannual variability in recent decades, partly associated with changing temperature and precipitation patterns. Improving the prediction of maple syrup yield and understanding its potential response to future climate change are therefore essential for supporting the resilience of Quebec's maple industry. This study developed and evaluated a machine learning (ML) framework for predicting maple syrup yield under historical and future climate conditions. Four ML algorithms—Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting Machine (GBM), and Extreme Gradient Boosting (XGBoost)—were evaluated using maple syrup yield data for 2005–2025 obtained from the Producteurs et Productrices Acéricoles du Québec (PPAQ) and corresponding temperature and precipitation data from the Coupled Model Intercomparison Project Phase 6 (CMIP6) obtained through Ouranos. The best-performing model was subsequently used to project maple syrup yield at the provincial scale and across five major maple-producing regions of Quebec: Centre-du-Québec, Estrie, Chaudière-Appalaches, Laurentides, and Montérégie, under the SSP2-4.5 and SSP3-7.0 climate scenarios. XGBoost demonstrated the highest predictive performance, achieving an RMSE of 0.04 kg tap^{-1} year^{-1}, an MAE of 0.03 kg tap^{-1} year^{-1}, and an R^2 of 0.99. Future projections indicated an overall decline in Quebec's maple syrup yield of 16.54% under SSP2-4.5 and 15.43% under SSP3-7.0 for 2040–2100 relative to the 2020–2024 baseline. However, substantial spatial heterogeneity was observed. In Centre-du-Québec, yield was

projected to decline by 13.26% and 16.57% under SSP2-4.5 and SSP3-7.0, respectively, whereas Montérégie experienced substantially larger projected declines of 30.11% and 34.79%. These findings demonstrate the potential of machine learning to capture complex climate–yield relationships and reveal important regional differences in the future vulnerability of Quebec's maple syrup sector. The resulting framework can support spatially targeted adaptation planning and inform strategies for maintaining the long-term resilience of maple syrup production under climate change.

Keywords: Maple syrup production; Machine learning; XGBoost; Climate change; Climate projections; Maple yield prediction; CMIP6; SSP2-4.5; SSP3-7.0; Quebec; Climate vulnerability; Agricultural adaptation.

1. Introduction

Climate change poses a major challenge to global food security and the achievement of Sustainable Development Goal 2 (SDG 2), which seeks to end hunger, achieve food security and improve nutrition (Campbell, 2022). Climate projections indicate that continued warming will increase the risks of food insecurity, malnutrition and diet-related mortality in many regions of the world (Intergovernmental Panel on Climate Change [IPCC], 2022). For example, Hultgren et al. (2025) estimated that a 1 °C increase in global temperature could reduce agricultural production by more than 5.5×10^{14} kcal annually, equivalent to approximately 120 kcal per person per day or 4.4% of recommended daily energy intake. Similarly, Li et al. (2025) projected that, in the absence of effective adaptation and mitigation, countries in South Asia, sub-Saharan Africa, North America and Oceania could face increasing difficulties in meeting national calorie requirements by the end of the twenty-first century. Although the magnitude and direction of climate impacts vary among regions and production systems, these findings highlight the need for improved approaches to quantify and predict climate-related changes in agricultural and agroforestry production.

Quebec is particularly important in this context because it is the world's leading producer of maple syrup. In 2024, Quebec accounted for approximately 73% of global maple syrup production (Government of Quebec, 2026). The province also hosts a large and economically important maple industry, with 7,838 maple syrup farms producing approximately 108.4 million kg of maple syrup in 2024 (Government of Quebec, 2026). Production is concentrated in several regions, with Chaudière-Appalaches accounting for approximately 48% of provincial production, followed by Estrie and Centre-du-Québec, which contributed approximately 13% and 12%, respectively (Producteurs et Productrices Acéricoles du Québec [PPAQ], 2025; Government of Quebec, 2026). Consequently, changes in the climatic conditions that regulate maple sap production could have important implications not only for individual producers but also for Quebec's agricultural economy, rural livelihoods and global maple syrup supply.

The economic importance of the maple industry further increases the significance of understanding its vulnerability to climate change. In 2024, Quebec's maple syrup sector generated more than $1.45 billion in economic activity and exported maple syrup to 71 countries (Government of Quebec, 2026). The United States was the principal export destination, accounting for approximately 65% of exports, followed by the European Union, the United Kingdom, Japan and Australia. Beyond its economic contribution, maple production provides important ecosystem services. Maple-

dominated forests contribute organic matter to forest soils, influence nutrient cycling and nitrogen mineralization, and can reduce nitrate leaching (Lovett et al., 2004; Lucash et al., 2012). They also contribute to the maintenance of plant and animal communities and broader forest biodiversity (Beier et al., 2012). The resilience of the maple industry is therefore relevant to both socioeconomic and ecological sustainability in Quebec.

Maple syrup production is strongly controlled by climatic and environmental conditions. The characteristic freeze–thaw cycle during early spring is particularly important because cold nighttime temperatures followed by sufficiently warm daytime temperatures generate the pressure gradients required for sap flow. Previous research has suggested that temperatures near 0 °C at night and above approximately 4 °C during the day provide favourable conditions for sap production (Pothier, 1995). Snow accumulation, winter and spring temperatures, precipitation, and the timing and duration of freeze–thaw events can therefore influence the onset, duration and intensity of the tapping season (Moss et al., 2024). Soil characteristics may also influence tree growth and productivity, with sugar maple generally performing well in slightly acidic soils (Moss et al., 2024). In addition, antecedent climatic conditions, including dry conditions during the preceding summer and precipitation during spring, can influence subsequent sap quantity and quality (Rapp et al., 2019; Moss et al., 2024). These interactions demonstrate that maple syrup production is governed by a complex combination of climatic, environmental and biological processes.

Climate change is consequently expected to alter the productivity, timing and geographic distribution of maple syrup production in Quebec and elsewhere in North America (Skinner et al., 2010; Moreau et al., 2020). Quebec has experienced significant climatic changes, including increasing temperatures, changes in precipitation patterns and increasing exposure to climate extremes (Epule et al., 2024). Such changes can modify the frequency and duration of the freeze–thaw cycles required for sap flow. Several studies have reported evidence of changes in the timing of the maple production season, with warming expected to result in earlier tapping and shorter production windows in some southern regions (Houle et al., 2015; Government of Canada, 2025). At broader spatial scales, suitable conditions for maple production are also expected to shift northward in response to warming (Skinner et al., 2010; Rapp et al., 2019). Within Quebec, recent analyses have identified northern or higher-latitude areas, including La Haute-Yamaska, Le Val-Saint-François, Saint-Joachim-de-Shefford and Lingwick, as potentially favourable future locations for maple syrup production (Moss et al., 2024).

Despite these findings, considerable uncertainty remains regarding the magnitude and spatial distribution of future changes in maple syrup yield. Previous studies have produced different estimates depending on the variables, spatial scales, climate scenarios and modelling approaches used. Houle et al. (2015), for example, used a logistic model to examine changes in the timing of maple syrup production in eastern Canada and projected a substantial shortening of the production season in southern Quebec under future climate conditions. Duchesne et al. (2009) applied multiple linear regression to estimate future climate impacts on maple syrup production in Quebec and projected yield reductions of approximately 15% by 2050 and 22% by 2090. Moreau et al. (2020) similarly reported that changes in climatic conditions could negatively affect maple syrup production in southern Quebec. In contrast, Rapp et al. (2019) found that changes in climatic suitability could create opportunities for maple production in parts of Quebec as favourable conditions shift northward. Filteau (2025) further showed that climate change may improve certain

aspects of maple syrup quality while simultaneously reducing the quantity produced. These contrasting findings demonstrate that climate change impacts on maple production are spatially heterogeneous and that assessments based solely on linear relationships may not adequately capture the complex interactions among climatic variables.

An additional limitation of the existing literature is the predominance of statistical and process-based approaches for predicting maple syrup production. Although these approaches have provided important insights into climate–production relationships, they may have limitations when relationships between predictors and yield are nonlinear, interactive or highly complex. Machine learning (ML) methods provide an alternative data-driven approach that can capture nonlinear relationships and higher-order interactions between climatic predictors and production outcomes (Crane-Droesch, 2018; Gharakhanlou & Perez, 2024). ML approaches have increasingly demonstrated strong predictive performance in agricultural yield modelling because of their capacity to identify complex patterns in high-dimensional datasets (von-Bloh et al., 2023; Dhillon et al., 2024). Algorithms such as Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting Machine (GBM) and Extreme Gradient Boosting (XGBoost) can accommodate nonlinear relationships without requiring the strict distributional assumptions commonly associated with conventional regression approaches.

However, ML approaches also present important limitations. Their performance can depend strongly on data availability, data quality, feature selection and model configuration, which can be challenging in agricultural and agroforestry systems where long-term observations are often limited (Parra et al., 2015; von Bloh et al., 2024). In addition, some ML models can be computationally intensive and difficult to interpret compared with conventional statistical models (Zhang & Ma, 2012). Consequently, rigorous model comparison and validation are necessary to determine whether ML methods provide reliable predictive improvements for maple syrup production. Importantly, despite the growing application of ML to agricultural yield prediction, there remains limited evidence on the application of ML to climate-driven maple syrup yield prediction in Quebec.

A further gap concerns the spatial scale of climate-impact assessments. Most previous studies have examined maple production at broad spatial scales or focused on specific aspects of production, such as production timing, climatic suitability or syrup quality. Yet climate impacts and adaptation opportunities are inherently spatially heterogeneous (Baker et al., 2012). A climate condition that reduces productivity in southern Quebec may simultaneously create more favourable conditions in northern regions. Provincial averages can therefore mask important regional differences in vulnerability and may provide limited guidance for region-specific adaptation planning. There is consequently a need for modelling frameworks capable of predicting future maple syrup yield at both provincial and regional scales.

To address these gaps, this study develops and evaluates a machine learning framework for predicting maple syrup yield under historical and future climate conditions in Quebec. Specifically, the predictive performance of four ML algorithms—Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Machine (GBM) and Extreme Gradient Boosting (XGBoost)—is evaluated using historical maple syrup yield, temperature and precipitation data. The best-performing model is then used to project future maple syrup yield under the SSP2-4.5 and SSP3-7.0 climate scenarios. Projections are conducted at both the provincial scale and across five major

maple-producing regions: Centre-du-Québec, Estrie, Chaudière-Appalaches, Laurentides and Montérégie.

This study makes three principal contributions. First, it provides a systematic comparison of four machine learning algorithms for climate-informed maple syrup yield prediction in Quebec, thereby assessing the suitability of ML approaches for this understudied production system. Second, it quantifies potential future changes in maple syrup yield under alternative climate scenarios and identifies the extent to which these impacts vary among major producing regions. Third, by providing regionally differentiated projections, the study offers information that can support climate adaptation planning, production management and long-term planning within Quebec's maple syrup sector. More broadly, the proposed framework demonstrates how machine learning can complement existing climate-impact modelling approaches and improve the spatial assessment of climate risks to climate-sensitive agricultural and agroforestry production systems.

2. Materials and Methods

2.1 Study Area

Quebec is the largest province in Canada, covering approximately 1.7 million km² and extending from approximately 45°N to 62°N latitude in eastern Canada (Government of Quebec, 2025). The province is administratively divided into 17 regions, 87 regional county municipalities (RCMs), and 1,123 local municipalities (Government of Quebec, 2025). Quebec City is the provincial capital, while Montréal is the largest metropolitan centre. The province had an estimated population of approximately 9.06 million inhabitants in 2025 (Institut de la statistique du Québec, 2026).

Quebec has a humid continental to subarctic climate, with substantial spatial and seasonal variations in temperature and precipitation. These climatic gradients are particularly important for maple syrup production because sap flow depends strongly on the occurrence and timing of spring freeze–thaw cycles. Maple production is concentrated primarily in the southern and central parts of the province, where climatic and ecological conditions are suitable for sugar maple (*Acer saccharum*) growth and spring sap production. The major maple-producing regions considered in this study are Centre-du-Québec, Estrie, Chaudière-Appalaches, Laurentides, and Montérégie. These regions collectively represent important components of Quebec's maple syrup production system and encompass substantial differences in climate and geographic conditions.

Quebec has experienced significant changes in temperature and precipitation over recent decades. Historical observations indicate that mean annual temperature increased by approximately 1.1 °C between 1948 and 2016, while winter and summer temperatures increased by approximately 1.4 °C and 1.5 °C, respectively (Zhang et al., 2019). Mean annual precipitation also increased by approximately 10.5% over the same period, with increases of approximately 5.3% during winter and 6.6% during summer (Zhang et al., 2019). These historical changes are particularly relevant to maple syrup production because changes in winter and early-spring temperatures can modify the frequency and duration of freeze–thaw events that regulate sap flow.

Future climate projections indicate that these trends are likely to continue. Under the SSP3-7.0 scenario, Quebec's mean annual temperature is projected to increase from approximately −1.4 °C in 2023 to 0.7 °C by 2060 and 3.0 °C by 2100 (Epule et al., 2024). Mean annual precipitation is similarly projected to increase from approximately 996 mm in 2023 to 1,052 mm by 2060 and 1,122 mm by 2100 under the same scenario (Epule et al., 2024). Although increased precipitation may partly offset some water-related stresses, increasing temperatures and changes in the timing and frequency of freeze–thaw events could substantially alter the climatic conditions required for maple sap production.

The combination of strong climatic gradients, a large and economically important maple syrup sector, and substantial projected changes in temperature and precipitation makes Quebec an appropriate study area for evaluating climate-related changes in maple syrup yield. The provincial-scale analysis provides an assessment of overall changes in maple syrup productivity, while the regional analysis allows differences in climate vulnerability among major maple-producing regions to be identified.

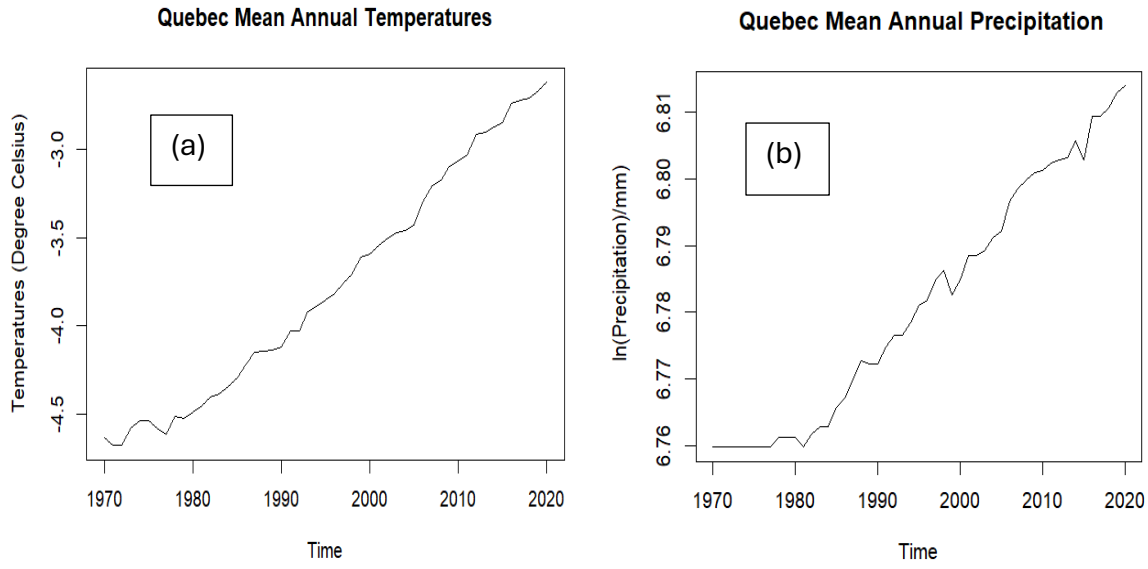


Figure 1. (a) Quebec's mean annual temperatures, (b) Quebec's mean annual precipitations

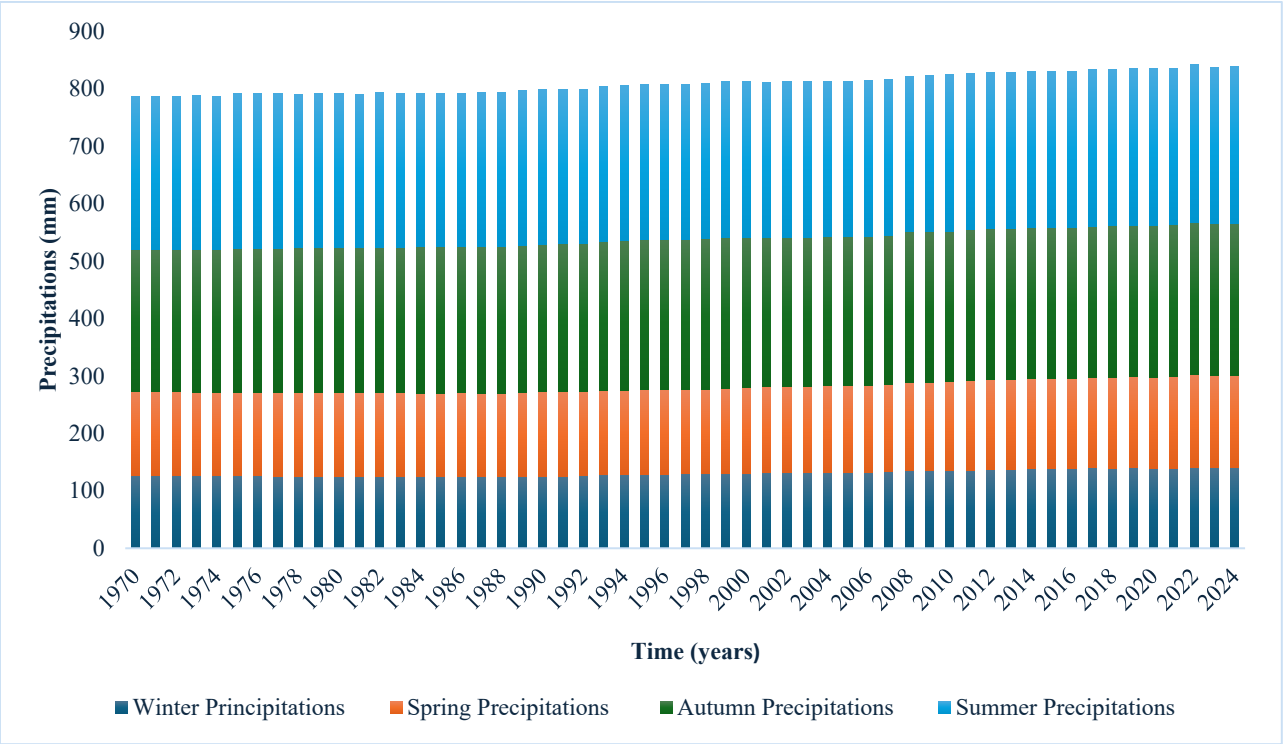


Figure 2. Quebec's seasonal precipitation trend over time (1970-2024)

Agriculture is an important component of Quebec's economy and contributes substantially to employment, food security and rural development. According to the 2021 Census of Agriculture, Quebec had more than 29,400 farms, a number that declined to approximately 28,200 farms by 2024 (Government of Canada, 2025). The agricultural sector employed approximately 51,600 people in 2024, representing about 1.2% of total employment in the province (Government of Canada, 2025). Montérégie, Chaudière-Appalaches and Centre-du-Québec are among the province's major agricultural regions, accounting for approximately 28.5%, 13.8% and 11.0%, respectively, of agricultural employment (Government of Canada, 2025). In 2020, Quebec's agricultural sector contributed more than $10 billion to the provincial economy and remains an important component of economic development and food security (Ouranos, 2026).

Quebec has a diversified agricultural production system that includes major field crops such as soybean, corn and barley, as well as maple syrup production. In 2024, Quebec produced approximately 1.4 million tonnes of soybean out of Canada's total production of approximately 7.6 million tonnes (Government of Canada, 2024). Quebec is also the world's dominant maple syrup-producing region, accounting for approximately 67% of global maple syrup production in 2023 (PPAQ, 2025). Maple syrup production is concentrated primarily in the southern and central regions of the province, including Chaudière-Appalaches, Estrie, Centre-du-Québec, Montérégie and Laurentides.

Quebec's agricultural and maple syrup production systems are strongly influenced by climatic conditions, particularly temperature and precipitation. Changes in these variables can affect crop growth, productivity and the timing of production activities, making the agricultural sector increasingly sensitive to climate change (Ouranos, 2026). For maple syrup specifically, temperature is particularly important because the occurrence and timing of spring freeze–thaw cycles directly regulate sap flow. Consequently, the observed and projected changes in Quebec's climate provide an important basis for assessing future changes in maple syrup yield at both provincial and regional scales.

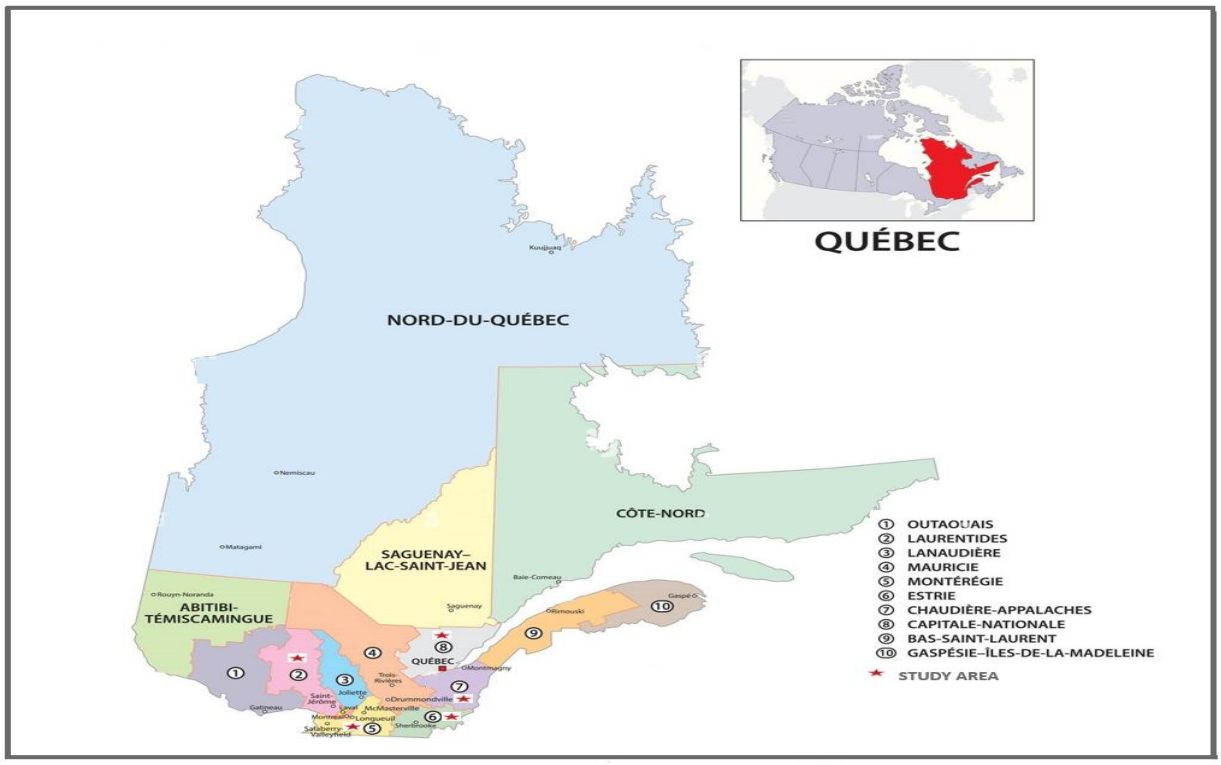


Figure 3. Map of Quebec showing the study areas

However, changing climatic conditions, characterized by rising temperatures and altered precipitation patterns, are expected to have substantial and heterogeneous effects on agricultural production in Quebec (Ouranos, 2014). Warmer temperatures may extend the growing season for some field crops, including maize and soybean, potentially creating more favourable conditions for crop development and increasing yields in some regions (Bootsma et al., 2005; St-Marseille et al., 2019). For maple syrup production, temperature changes have more complex effects because sap flow depends on the occurrence and timing of spring freeze–thaw cycles. Conditions characterized by nighttime temperatures near 0 °C followed by daytime temperatures above approximately 4 °C are considered favourable for sap flow (Pothier, 1995). Moderate warming may therefore improve sap-flow conditions in some regions, particularly where current temperatures are below optimal levels. However, excessive warming may reduce the frequency and duration of freeze–thaw events,

advance the onset of the tapping season, and shorten the overall production window, potentially reducing maple syrup yield (Duchesne et al., 2009; Rapp et al., 2019).

Beyond temperature changes, climate change is expected to create additional challenges for Quebec's agricultural sector, including the emergence and expansion of pests and plant diseases, increased risks of drought and flooding, and changes in the timing of sowing, harvesting and other agricultural activities (Ouranos, 2014). The combined effects of these climatic changes may therefore generate substantial spatial and temporal variability in crop productivity across Quebec. For maple syrup production in particular, the nonlinear relationship between temperature, precipitation, freeze–thaw dynamics and sap flow makes it important to account for these interactions when assessing future yield. Consequently, understanding and predicting the response of maple syrup production to changing climatic conditions is essential for developing effective regional adaptation strategies (Duchesne et al., 2009; Rapp et al., 2019).

2.2 Data Collection

This study used maple syrup yield and climatic data at both provincial and regional scales. Maple syrup yield data for Quebec and five major maple-producing regions—Centre-du-Québec, Estrie, Chaudière-Appalaches, Laurentides and Montérégie—were obtained from the Producteurs et Productrices Acéricoles du Québec (PPAQ) for the period 2005–2025. The dataset provided annual maple syrup production and was used to calculate maple syrup yield, expressed as kilograms of syrup per tap per year (kg tap^{-1} $year^{-1}$) (Table 1).

Climate data consisted of annual and seasonal temperature and precipitation variables corresponding to the maple production areas. Historical climate data for the study period were obtained from the Coupled Model Intercomparison Project Phase 6 (CMIP6) dataset through Ouranos. For future projections, climate data were obtained under two Shared Socioeconomic Pathways (SSPs): SSP2-4.5 and SSP3-7.0. SSP2-4.5 represents an intermediate emissions pathway in which socioeconomic and technological development follows broadly intermediate trajectories, while SSP3-7.0 represents a higher-emissions pathway characterized by regional rivalry, slower economic development, and comparatively limited international cooperation (Feng et al., 2020) (Table 1).

Temperature and precipitation variables were selected because of their established influence on maple sap flow and syrup production. The climate dataset included annual and seasonal temperature and precipitation measures to capture both long-term climatic conditions and seasonal variability relevant to maple production. The historical climate and yield data were used to train and validate the machine learning models, whereas projected climate data under SSP2-4.5 and SSP3-7.0 were subsequently used as inputs to the best-performing model to estimate future maple syrup yield.

For consistency between the climate and production datasets, climatic variables were spatially aggregated to the provincial and regional scales corresponding to the maple syrup yield data. The

resulting dataset therefore consisted of matched observations of maple syrup yield and climatic predictors for Quebec and the five selected maple-producing regions over the historical period, followed by projected climatic conditions for future periods.

Table 1. Data description

Data	***Units***	***Source***
Maple syrup yield	kg/tap/year	Producteurs et Productrices Acéricole du Québec (PPAQ)
Mean annual temperatures	^{0}C	Ouranos
Winter temperatures	^{0}C	Ouranos
Spring temperatures	^{0}C	Ouranos
Summer temperatures	^{0}C	Ouranos
Autumn temperatures	^{0}C	Ouranos
Mean annual precipitations	mm	Ouranos
Winter precipitations	mm	Ouranos
Spring precipitations	mm	Ouranos
Summer precipitations	mm	Ouranos
Autumn precipitations	mm	Ouranos

2.3 Machine Learning Models

Four supervised machine learning regression algorithms—Support Vector Regression (SVR), Random Forest Regression (RF), Gradient Boosting Machine (GBM), and Extreme Gradient Boosting (XGBoost)—were evaluated for predicting maple syrup yield in Quebec. These algorithms were selected because they can capture nonlinear relationships and interactions between climatic predictors and yield and have been widely applied to agricultural yield prediction (von Bloh et al., 2024; Dhillon et al., 2024; Gharakhanlou & Perez, 2024; Harraf et al., 2026). The four algorithms also represent different modelling approaches, including kernel-based learning (SVR), bagging-based ensemble learning (RF), and boosting-based ensemble learning (GBM and XGBoost), thereby allowing their predictive performance to be systematically compared.

The models were trained using historical maple syrup yield and climatic variables and subsequently evaluated using independent observations. The best-performing model was selected based on the predictive performance metrics described in Section 2.4 and was subsequently used with projected climate data to estimate future maple syrup yield under SSP2-4.5 and SSP3-7.0.

2.3.1 Random Forest Regression

Random Forest (RF) is a supervised ensemble learning algorithm that can be used for both classification and regression (Breiman, 2001). In regression applications, RF constructs a large number of decision trees using bootstrap samples of the training dataset and randomly selected

subsets of predictor variables. The random selection of observations and predictors reduces the correlation among individual trees and improves the stability and generalization of the ensemble (Dhillon et al., 2024; Gharakhanlou & Perez, 2024).

For each bootstrap sample, an individual decision tree is grown by recursively partitioning the predictor space. Each tree produces an independent prediction, and the final RF prediction is obtained by averaging the predictions from all trees. The RF regression prediction can be expressed as:

$$Y = 1/N \sum_{i=1}^{N} T(X) \quad \ldots\ldots\ldots\ldots\ldots\ldots 1$$

Where Y is the predicted output; N is the total number of decision trees in the forest; $T(X)$ is the prediction made by the i^{th} decision trees for input X.

RF is particularly suitable for this study because it can model nonlinear relationships between climatic variables and maple syrup yield without requiring predefined functional relationships. It is also relatively robust to noise and multicollinearity and can provide measures of predictor importance. However, RF predictions may be less effective when extrapolating beyond the range of conditions represented in the training data. This limitation is particularly relevant when applying historical models to future climate projections and was therefore considered when interpreting the results.

2.3.2 Support Vector Regression

Support Vector Regression (SVR) is the regression extension of the Support Vector Machine (SVM) framework. Unlike conventional regression approaches that attempt to minimize the prediction error for every observation, SVR seeks a function that approximates the target values within a predefined (\epsilon)-insensitive margin while controlling model complexity (Tian et al., 2018; Dhillon et al., 2024).

Given a set of observations ((X_i,Y_i)), where (X_i) represents the vector of climatic predictors and (Y_i) represents observed maple syrup yield, SVR estimates a function of the form:

Assume there is a sampling data

$$X = (x_1, x_2, \ldots\ldots x_i) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2$$

and the corresponding target is

$$Y = (y_1, y_2 \ldots\ldots y_i) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3$$

Where y_i belongs to a set of real numbers

SVR aims to find a function *f(x)* that has at most ϵ deviations from the obtained target y_i for all the training data.

The basic formular of a SVR model according to Gawdiya et al. (2024) is specified as

$$Y=\sum_{i=0}^{n} w ker(X_I, X_J)+b \ldots\ldots\ldots\ldots\ldots\ldots 4$$

Where *Y* is the dependent variable; *w* is the corresponding weight assigned to each predictor and $ker(X_I, X_J)$ is the kernel function.

The transformation (X_I, X_J) is defined through a kernel function, allowing SVR to model nonlinear relationships without explicitly calculating the transformed features. In this study, a nonlinear kernel can be used to capture complex relationships between temperature, precipitation and maple syrup yield. The commonly used radial basis function (RBF) kernel is expressed as:

K(X_i,X_j) =\exp(-\gamma|X_i-X_j|^2)\tag……*5*

where (K(X_i,X_j)) is the kernel similarity between observations (X_i) and (X_j), and (\gamma) controls the influence of individual training observations.

SVR was included because it can perform well with nonlinear relationships and relatively small datasets. However, unlike tree-based methods, SVR is sensitive to the scale of predictor variables. Therefore, all climatic predictors were standardized before model training. The principal hyperparameters, including the penalty parameter (C), the (\epsilon)-insensitive loss parameter, and the kernel parameter (\gamma), were optimized during model tuning.

2.3.3 Gradient Boosting Machine

Gradient Boosting Machine (GBM) is an ensemble learning method that sequentially combines weak prediction models, typically decision trees, to construct a strong predictive model (Friedman, 2001). Unlike Random Forest, where trees are constructed independently and their predictions are subsequently averaged, GBM builds trees sequentially, with each new tree focusing on reducing the errors generated by the existing ensemble.

For a regression problem, the GBM begins with an initial prediction and calculates the residuals between observed and predicted values. A new decision tree is then trained to approximate these residuals. The predictions are subsequently updated using a learning-rate parameter. The general form of the boosted model can be expressed as:

F_m(X) = F_{m-1}(X)+\eta h_m(X)\tag……*6*

where (F_m(X)) is the prediction after the addition of the (m)th tree, (F_{m-1}(X)) is the prediction from the previous iteration, (h_m(X)) is the newly fitted weak learner, and (\eta) is the learning rate.

The iterative nature of GBM allows the model to progressively reduce prediction errors and capture nonlinear relationships and interactions among climatic variables. GBM was therefore selected as a complementary boosting approach to RF and SVR. Key parameters, including the number of trees, tree depth, learning rate and minimum observations per terminal node, were optimized during model training to balance predictive performance and overfitting.

2.3.4 Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is an optimized implementation of gradient-boosted decision trees that incorporates regularization and computational improvements to increase predictive performance and reduce overfitting (Chen & Guestrin, 2016). Like conventional GBM, XGBoost sequentially builds decision trees, with each subsequent tree designed to improve the errors of the existing ensemble.

A major distinction between XGBoost and conventional GBM is the explicit regularization of the objective function. For a regression problem, the objective function can be expressed generally as:

\mathcal{L}=\sum_{i=1}^{n}l(y_i,\hat{y}i)+\sum{k=1}^{K}\Omega(f_k)\tag……7

where (l(y_i,\hat{y}_i)) is the loss function measuring the difference between observed and predicted values and (\Omega(f_k)) represents the regularization term controlling the complexity of the (k)th decision tree.

The regularization component penalizes overly complex trees and can improve model generalization. XGBoost also incorporates several algorithmic optimizations, including tree pruning, shrinkage, subsampling and efficient parallel computation (Chen & Guestrin, 2016). These characteristics make it particularly suitable for modelling nonlinear and interacting climate–yield relationships.

In this study, XGBoost was evaluated alongside conventional GBM to determine whether its additional regularization and optimization procedures resulted in improved prediction of maple syrup yield. Model hyperparameters, including the number of boosting rounds, maximum tree depth, learning rate, subsampling ratio, column-sampling ratio and regularization parameters, were optimized during model calibration.

2.3.5 Model Training and Hyperparameter Optimization

To ensure a fair comparison among the four algorithms, all models were trained using the same predictor variables and response variable. Because the study uses a relatively short annual time series, model validation was designed to preserve the temporal structure of the observations and avoid information leakage between the training and testing datasets. The data were divided into training and testing subsets (80% and 20%), with earlier observations used for model development and later observations retained for independent evaluation.

Hyperparameters were optimized using cross-validation within the training dataset. For SVR, the principal parameters included (C), (\epsilon), and (\gamma). For RF, the parameters included the number of trees, number of predictors considered at each split, maximum tree depth and minimum node size. For GBM and XGBoost, the parameters included the number of trees, learning rate, tree depth and subsampling parameters. The optimal parameter combination for each algorithm was determined based on its cross-validation performance.

To ensure comparability, model performance was evaluated using the same independent testing observations and the same performance metrics. The best-performing model was subsequently selected for future climate projections under SSP2-4.5 and SSP3-7.0. Because future climate projection involves extrapolation beyond the historical climate range, projected yield estimates were interpreted with particular attention to the range of climatic conditions represented in the model training dataset.

2.4 Data Analysis and Model Evaluation

The analysis was conducted in several sequential steps, including data transformation, assessment of predictor collinearity, model training and validation, comparison of machine learning performance, and future climate projections.

First, the distribution of maple syrup yield was examined to identify potential skewness and heteroscedasticity. Where appropriate, a natural logarithmic transformation was applied to the yield variable to reduce positive skewness and stabilize the variance:

$$Y'=\ln(Y)\tag{8}$$

where (Y) is the observed maple syrup yield and (Y') is the transformed yield. The transformation was applied consistently during model training, and predictions were back-transformed to the original yield scale for interpretation and reporting.

Because the dataset consists of annual observations, temporal characteristics of the yield series were also examined using the Augmented Dickey–Fuller (ADF) test. The ADF test evaluates the

null hypothesis that a time series contains a unit root and is therefore non-stationary. However, stationarity was not imposed as a strict requirement for the tree-based machine learning models because RF, GBM and XGBoost do not require stationary predictors in the same way as conventional time-series regression models. The ADF test was therefore used primarily as a diagnostic assessment of the temporal properties of the yield series rather than as a basis for automatically differencing the data.

Multicollinearity among climatic predictors was assessed using the Variance Inflation Factor (VIF). VIF values were calculated for all candidate predictors, and variables with VIF ≥ 5 were considered to exhibit substantial multicollinearity and were excluded from the predictor set. This threshold was selected to reduce redundancy among highly correlated climatic variables and improve the stability and interpretability of the modelling framework. Predictor selection was performed using the training data to avoid information leakage from the testing dataset.

Given the temporal nature of the data, the observations were divided chronologically into training and testing datasets. Approximately 80% of the earliest observations were used for model training and hyperparameter optimization, while the remaining 20% of the most recent observations were retained as an independent testing dataset. This temporal split provides a more realistic assessment of the ability of the models to predict future observations than a random split, particularly because the ultimate objective of the study is to project maple syrup yield under future climate conditions.

The four machine learning algorithms—Support Vector Regression (SVR), Random Forest Regression (RF), Gradient Boosting Machine (GBM), and Extreme Gradient Boosting (XGBoost)—were then trained using the selected climatic predictors. Hyperparameters for each model were optimized using cross-validation within the training dataset. For SVR, predictor variables were standardized before model fitting because of the algorithm's sensitivity to differences in variable scales. The same training observations and predictor variables were used for each algorithm to ensure a fair comparison.

The study employed three statistical metrics namely, Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Coefficient of Determination (R^2) to assess the predictive performance of the four ML algorithms employed. We then selected the ML algorithm with the best performance to predict Quebec's maple syrup yield by 2100. These statistical metrics have been used by Dhillon et al. (2024), Gawdiya et al. (2024) and Harraf et al. (2026) to assess the predictive performance of ML models for crop yield predictions.

RMSE was used to measure the magnitude of prediction errors while giving greater weight to larger errors. The equation deployed is given below:

RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2}\tag……..*9*

where (Y_i) is the observed yield, (\hat{Y}_i) is the predicted yield, and (n) is the number of observations.

MAE was calculated as:

MAE=\frac{1}{n}\sum_{i=1}^{n}|Y_i-\hat{Y}_i|\tag……*10*

where lower MAE values indicate greater predictive accuracy.

The coefficient of determination (R^2) was calculated as:

R^2=1-\frac{\sum_{i=1}^{n}(Y_i\hat{Y}*i)^2}{\sum*{i=1}^{n}(Y_i\bar{Y})^2}\tag…….11

where (\bar{Y}) represents the mean observed yield. Higher (R^2) values indicate that a greater proportion of the variability in observed yield is explained by the model.

The models were implemented in R using the **caret** package and associated machine learning packages. The algorithm producing the lowest RMSE and MAE and the highest (R^2) on the independent testing dataset was considered the best-performing model. XGBoost demonstrated the strongest predictive performance and was therefore selected for future maple syrup yield projections.

For future projections, the optimized XGBoost model was applied to climate projections under the SSP2-4.5 and SSP3-7.0 scenarios. Projected temperature and precipitation variables were processed using the same procedures and variable definitions applied to the historical data. Future yield predictions were generated for the provincial and regional study areas for the period 2040–2100. Changes in projected yield were calculated relative to the historical baseline period of 2020–2024 as:

\Delta Y(%)=\frac{Y_{future}-Y_{baseline}}{Y_{baseline}}\times100\tag…….12

where (Y_{future}) represents the mean projected maple syrup yield for the future period and (Y_{baseline}) represents the mean yield during 2020–2024. Negative values indicate projected yield reductions, whereas positive values indicate projected increases.

To facilitate interpretation of the regional impacts of climate change, projected percentage changes were calculated separately for Quebec as a whole and for Centre-du-Québec, Estrie, Chaudière-Appalaches, Laurentides and Montérégie. This approach allowed the analysis to identify spatial differences in projected climate impacts and to assess whether regions are likely to experience similar or contrasting changes in maple syrup productivity.

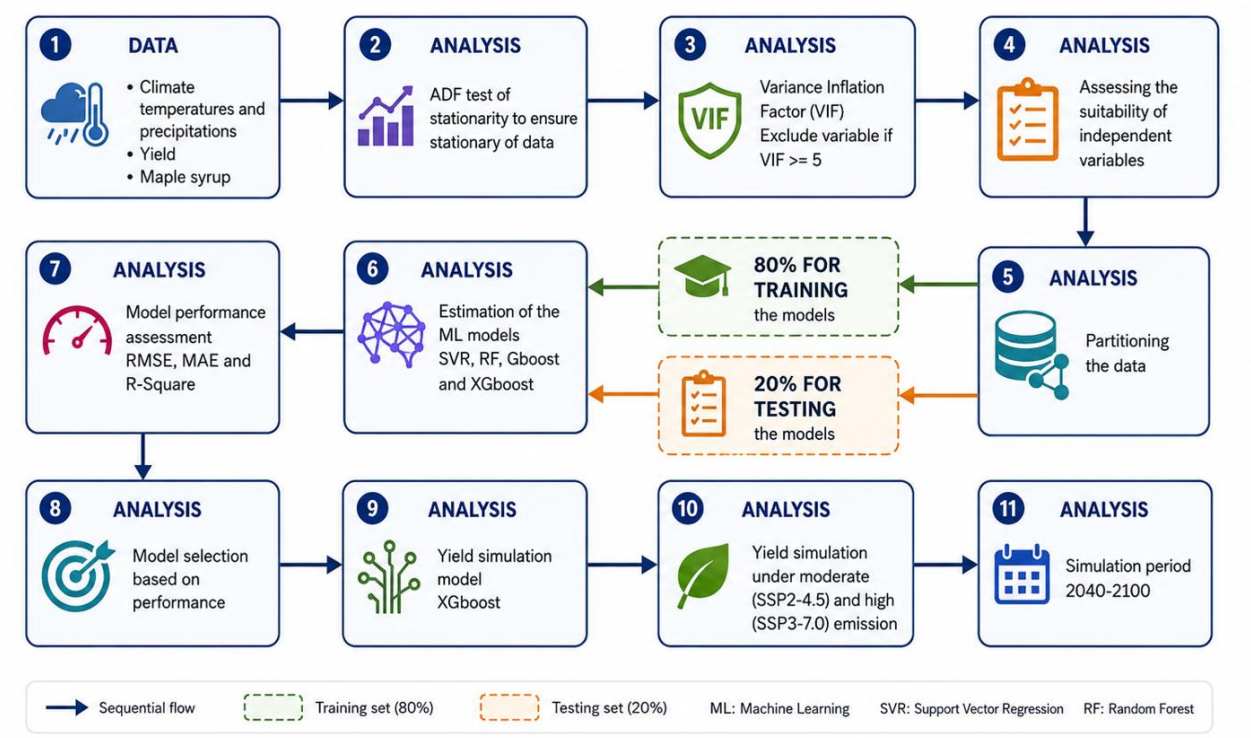


Figure 4. Flow diagram of the research methodology (Author's conceptualization, 2026)

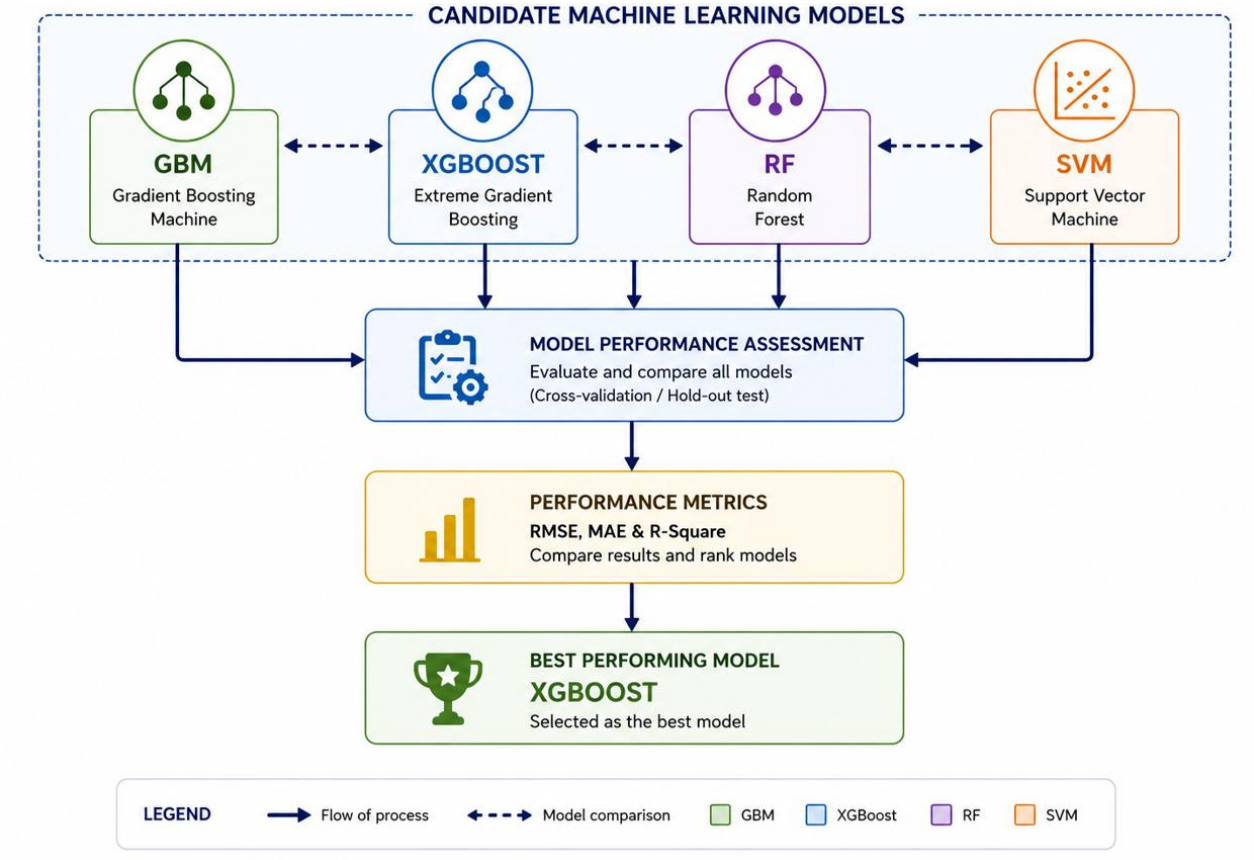


Figure 5. Flow diagram of the machine learning model performance assessment (Author's conceptualization, 2026

3. Results

3.1 Descriptive Results

The analysis indicates an overall increasing trend in Quebec's maple syrup yield over the study period, although substantial year-to-year fluctuations were observed. A marked increase in yield occurred in 2005, followed by a considerable decline beginning in 2006. From 2007 to 2011, maple syrup yield remained relatively stable before exhibiting greater interannual variability, with alternating periods of increases and decreases in subsequent years. Notably, maple syrup yield declined considerably in 2023, reaching levels comparable to those observed in 2007 (Figure 6).

The pronounced decline in yield in 2023 may be associated with the unusually warm conditions experienced in Quebec, particularly during the winter season. Changes in temperature and

precipitation patterns may influence the timing and duration of the maple sap flow season and, consequently, maple syrup production. These climatic variations may therefore contribute to the observed fluctuations in maple syrup yield over time.

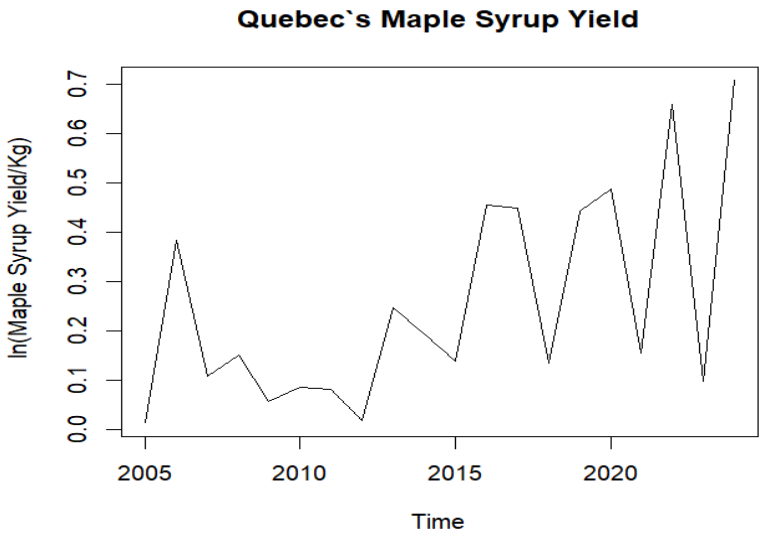


Figure 6. Trend of Quebec's maple syrup yield between 2005-2025

At the regional level, the analysis indicates a broadly similar pattern in maple syrup yield across the selected regions of Quebec, namely Centre-du-Québec, Montérégie, Laurentides, Estrie, and Chaudière-Appalaches. Overall, maple syrup yield exhibited an increasing trend over the study period, although substantial year-to-year fluctuations were observed across all regions (Figure 7).

Despite the similarities in temporal trends, notable differences in yield levels were observed among the regions. Relatively higher yield levels were recorded in Centre-du-Québec, Montérégie, and Chaudière-Appalaches (Figure 7a, b, and d), which are among Quebec's major maple syrup-producing regions. The relatively high yields observed in these regions may be associated with a combination of favourable climatic conditions, established maple production infrastructure, technological development, and other regional characteristics that support maple syrup production.

The regional trend analysis also shows a considerable decline in maple syrup yield in 2023 across most of the selected regions (Figure 7), consistent with the pattern observed at the provincial level. The widespread decline may be partly associated with the unusually warm temperatures recorded in Quebec during 2023, particularly during the winter and early spring periods. Changes in temperature and precipitation patterns, together with other environmental and production-related factors, may have contributed to the observed reduction in maple syrup yield. However, the relative contribution of these factors requires further investigation through the subsequent modelling and climate-impact analyses.

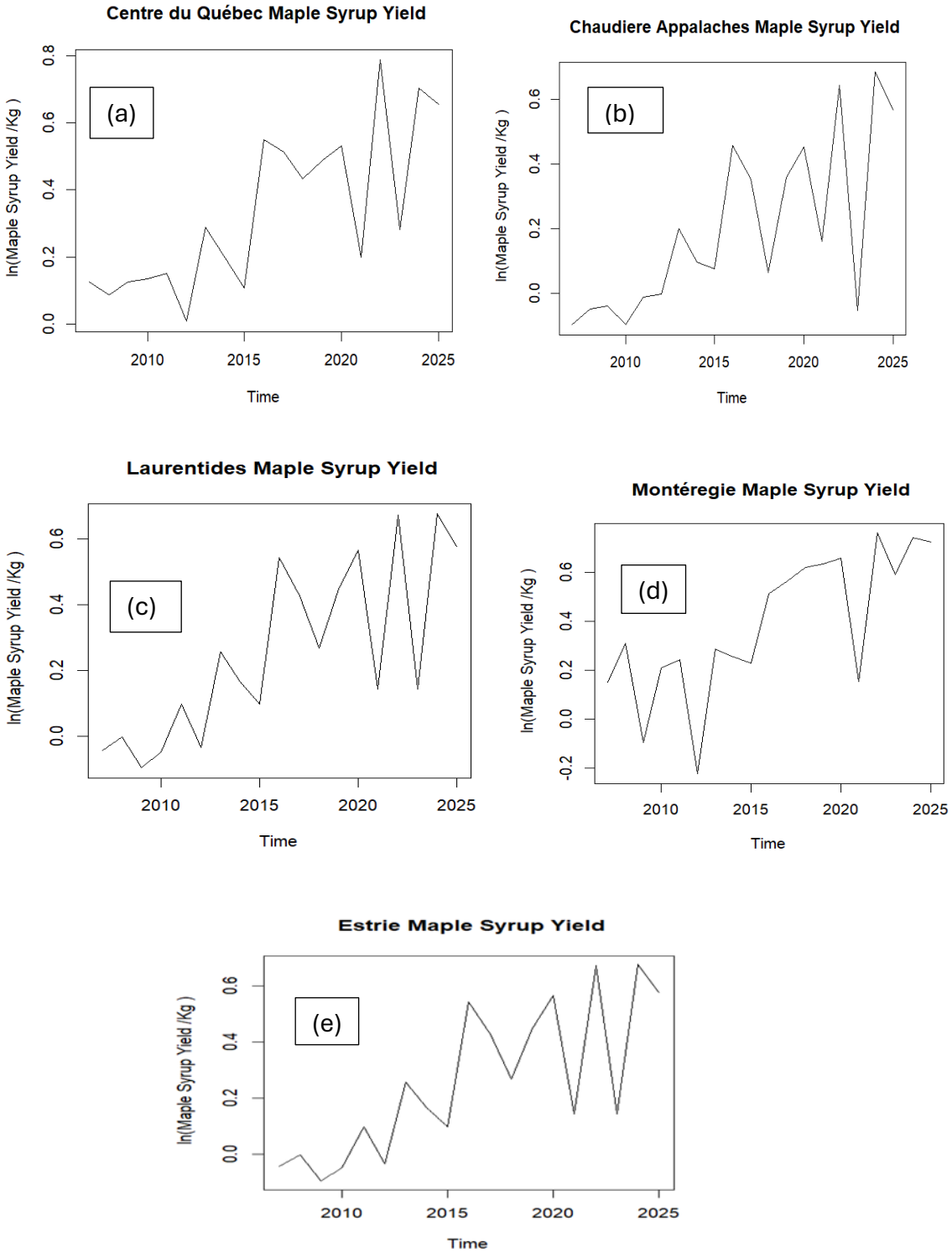


Figure 7. Regional trends of maple syrup yield in Quebec (a) Centre du Québec, (b) Chaudière-Apalachees, (c) Laurentides, (d) Montérégie, and (e) Estrie

3.2 Performance Assessment of Machine Learning Models for Predicting Maple Syrup Yield

The performance assessment indicates that the Extreme Gradient Boosting (XGBoost) model outperformed the other machine learning models evaluated in this study, achieving an RMSE of 0.04 kg/tap/year, an MAE of 0.03 kg/tap/year, and an R^2 of 0.99 (Table 2). These results demonstrate the model's high predictive accuracy, with the R^2 value indicating that XGBoost explained approximately 99% of the variability in maple syrup yield.

The superior performance of XGBoost may be attributed to its ensemble learning framework and built-in regularization mechanisms, which help reduce model complexity and minimize overfitting. In particular, XGBoost incorporates L1 (Lasso) and L2 (Ridge) regularization, which can improve model generalization and predictive performance when dealing with complex relationships between predictor variables and maple syrup yield. However, these mechanisms should be interpreted as potential explanations for the observed performance rather than as direct evidence unless their specific contribution has been evaluated experimentally.

Table 2. Performance assessment of the machine learning model for maple syrup yield

Models	Root Mean Square Error (RMSE)	Mean Absolute Error (MAE)	R-Square (R^2)
Random Forest Regression	0.09	0.25	0.87
Support Vector Regression	0.19	0.14	0.89
Extreme Gradient Boosting Regression	**0.04**	**0.03**	**0.99**
Gradient Boosting Regression	0.37	0.32	0.56

3.3 Plots of the predicted versus observed maple syrup yield (Quebec)

The predicted-versus-observed plots demonstrate that **XGBoost provided the strongest agreement between predicted and observed maple syrup yield**, followed by SVR and RF (Figure 8). This is consistent with their respective coefficients of determination (R^2), which were **0.99 for XGBoost, 0.89 for SVR, and 0.87 for RF**. The high R^2 values indicate a strong correspondence between the model predictions and the observed yield values, with XGBoost showing the closest fit.

In contrast, the Gradient Boosting (GBM/GBR) model exhibited a comparatively weaker relationship between predicted and observed maple syrup yield. This is consistent with its substantially lower R^2 value of **0.56**, indicating that the model explained a smaller proportion of the variability in the observed yield. Overall, the predicted-versus-observed plots corroborate the model performance results presented in Table 2 and further demonstrate the superior predictive capability of **XGBoost**, which was subsequently selected for maple syrup yield simulation.

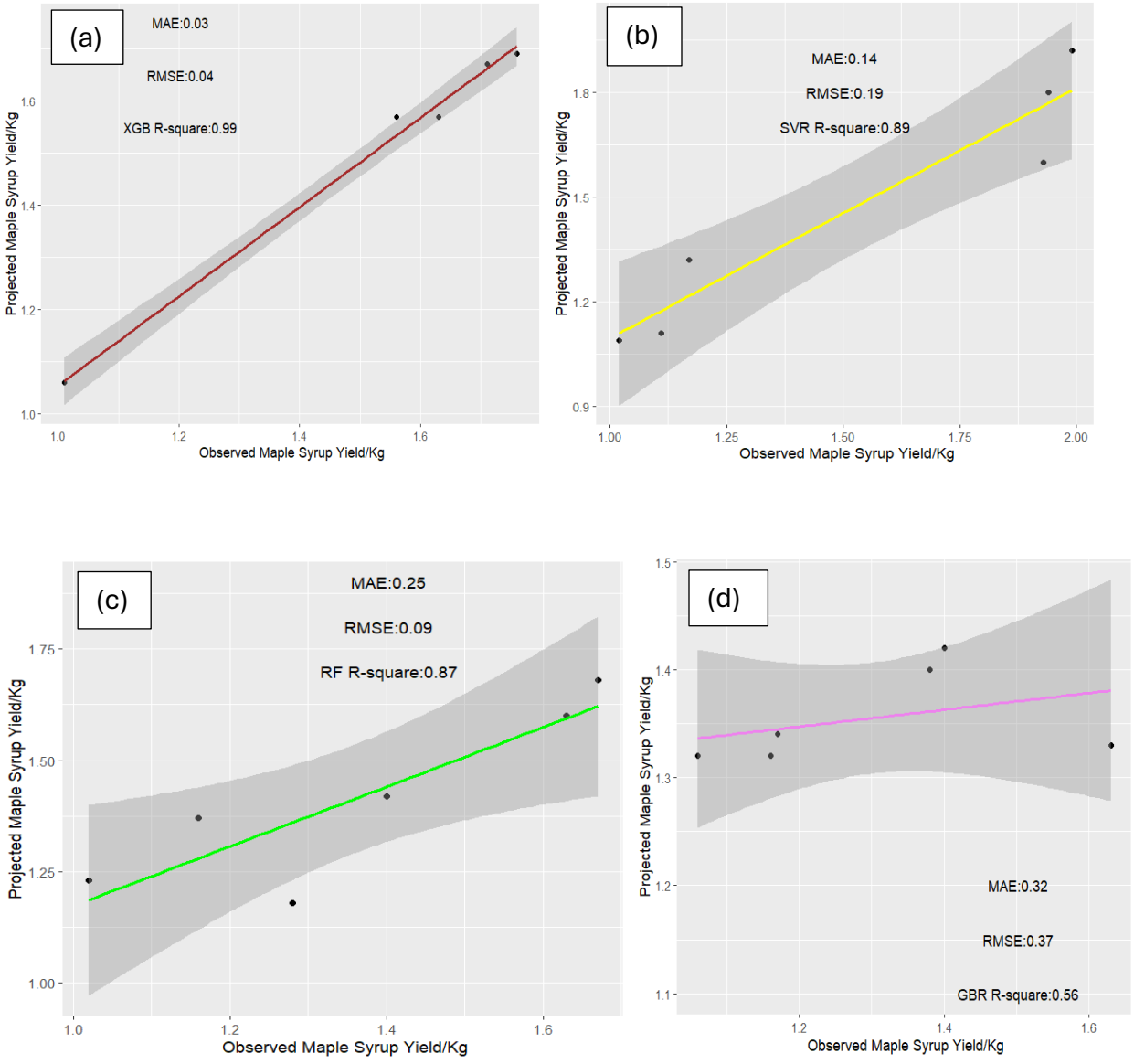


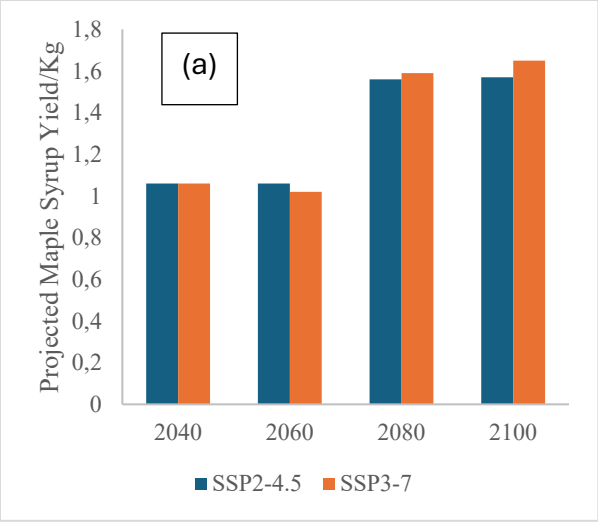
Figure 8. Plots of predicted versus observed yield (a) Extreme gradient boosting machine, (b) Support vector regression, (c) Random Forest, and (d) Gradient boosting machine

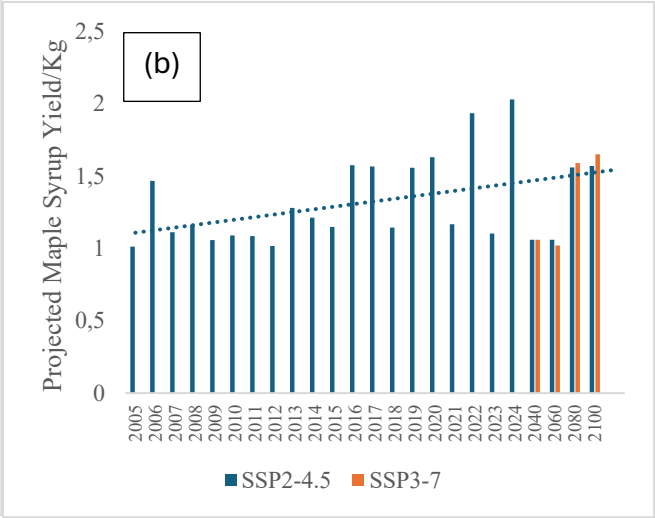
3.4 Projected Changes in Maple Syrup Yield Under Future Climate Scenarios

The analysis indicates that Quebec's maple syrup yield is projected to decline between 2040 and 2060, followed by a slight increase between 2080 and 2100 (Figure 9b). This pattern suggests that future changes in temperature and precipitation may have a nonlinear effect on maple syrup yield, resulting in fluctuations in projected production rather than a consistent decline throughout the projection period.

Despite the projected recovery toward the end of the century, maple syrup yield is expected to remain below the historical baseline. Relative to the historical period of 2020–2024, yield is projected to decline by 16.54% under the moderate-emissions scenario (SSP2-4.5) and 15.43% under the high-emissions scenario (SSP3-7.0) over the 2040–2100 projection period.

Interestingly, the SSP3-7.0 scenario is associated with a slightly smaller projected reduction in maple syrup yield than SSP2-4.5 (Figure 9). This result should be interpreted cautiously, as a higher-emissions scenario would not necessarily be expected to produce more favourable conditions for maple syrup production. One possible explanation is that the projected temperature trajectory under SSP3-7.0 may, during some periods, more closely approach the temperature range associated with optimal sap flow and maple syrup production. However, this interpretation requires further investigation of the projected temperature–yield relationship and should not be interpreted as evidence that higher emissions are beneficial to maple syrup production.





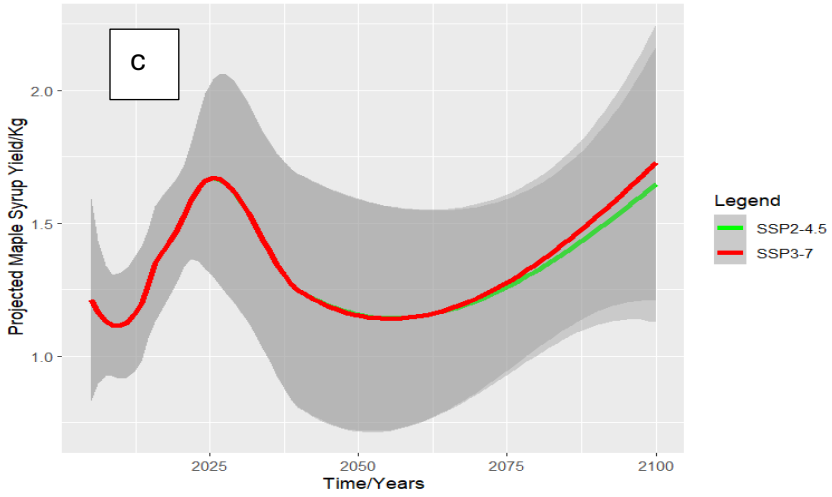


Figure 9. Projections of Quebec's maple syrup yield, (a) histogram showing the projections under the SSP2-4.5 and SSP3-7.0 scenarios, (b) histogram showing the observed and the projected yield, (c) curve of the observed and projected yield.

3.5 Plot of Projected Versus Observed Maple Syrup Yield in Centre-du-Québec

The projected versus observed maple syrup yield in Centre-du-Québec shows a strong agreement between the two variables (Figure 10). The model achieved a coefficient of determination (R^2) of 0.97, indicating that approximately 97% of the variability in the observed yield is explained by the model projections. The model also recorded a Root Mean Square Error (RMSE) of 0.09 kg/tap/year and a Mean Absolute Error (MAE) of 0.02 kg/tap/year, indicating relatively low prediction errors. Overall, the high R^2 value combined with the relatively low RMSE and MAE suggests that the model provides a strong representation of maple syrup yield in Centre-du-Québec and performs well in reproducing the observed yield patterns.

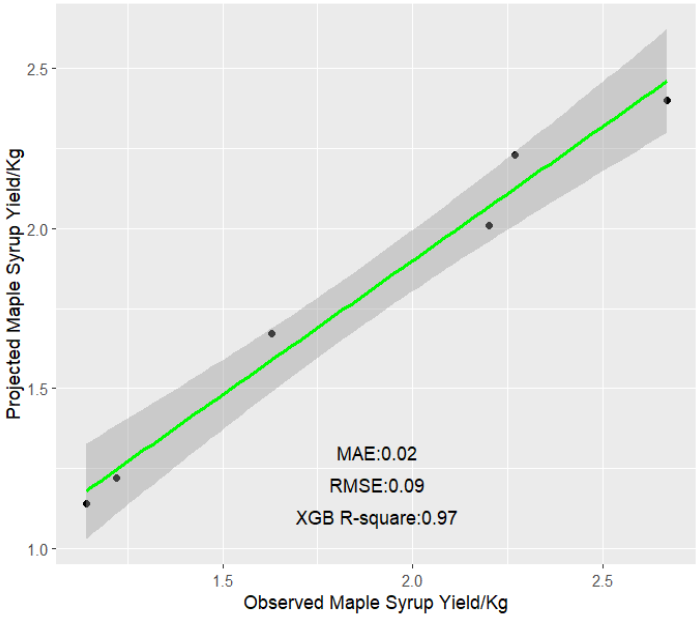


Figure 10. Projected versus observed maple syrup yield in Centre du Québec

3.6 Predictions of Maple Syrup Yield in Centre-du-Québec

The projections indicate that future changes in temperature and precipitation are likely to have varying effects on maple syrup yield in Centre-du-Québec by 2100 (Figure 11). During the 2040–2060 period, maple syrup yield is projected to decline under both climate scenarios, with yields remaining below 2 kg/tap/year. In contrast, a recovery in yield is projected during the 2080–2100 period, particularly under the SSP2-4.5 scenario, where yields are projected to exceed 2 kg/tap/year.

Despite this projected recovery, maple syrup yield over the 2040–2100 period is expected to decline by 13.26% under SSP2-4.5 and 16.57% under SSP3-7.0, relative to the historical baseline period of 2020–2024 (Figure 11b). These results suggest that the high-emissions scenario (SSP3-7.0) would have a greater negative impact on maple syrup yield in Centre-du-Québec than the moderate-emissions scenario (SSP2-4.5) (Figure 11a).

The greater projected decline under SSP3-7.0 may be associated with higher temperatures that increasingly exceed the temperature range favourable for maple sap flow and syrup production. The results therefore suggest that continued warming could reduce maple syrup productivity in Centre-du-Québec, although the projected fluctuations in yield indicate that the response to future climate conditions is likely to be nonlinear.

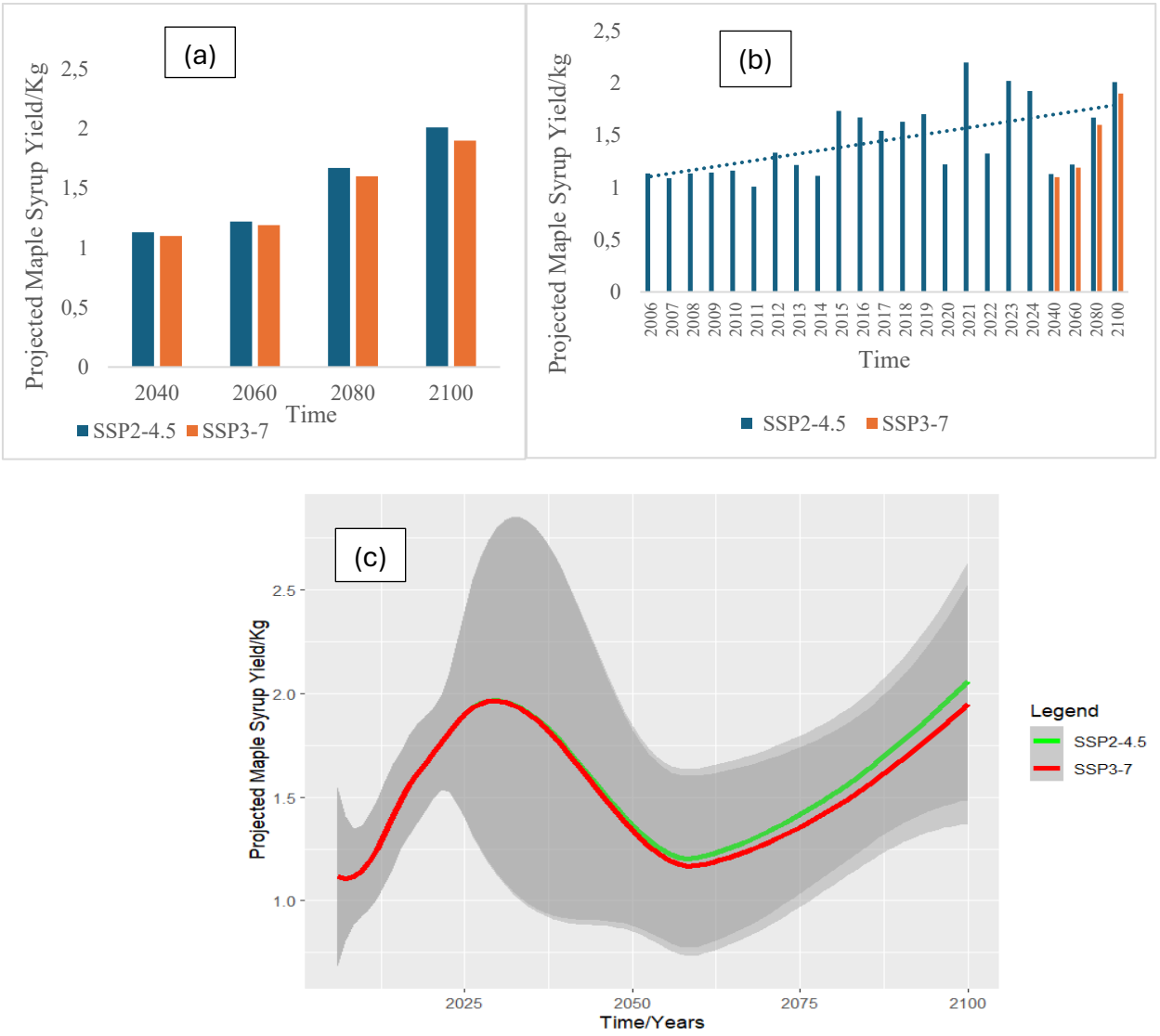


Figure 11. Projections of maple syrup yield in Centre du Quebec**,** (a) histogram comparing the projections under the SSP2-4.5 and the SSP3-7.0 scenarios, (b) histogram showing the observed and projected yield, (c) curve of the observed and projected yield

3.7 Plot of the projected versus observed maple syrup yield in Montérégie

The analysis shows a good fit of the projected yield against the observed yield in the Montérégie region of Quebec. The RMSE of the model is 0.02kg/tap/year and the MAE is 0.01kg/tap/year (Figure 11). Meanwhile, the R^2 of the model is 0.96 and implies that 96% of the change in the projected maple syrup yield is explained by the observed yield. The model's low value of RMSE and MAE and high value of R^2 indicates a high prediction accuracy between the observed and predicted maple syrup.

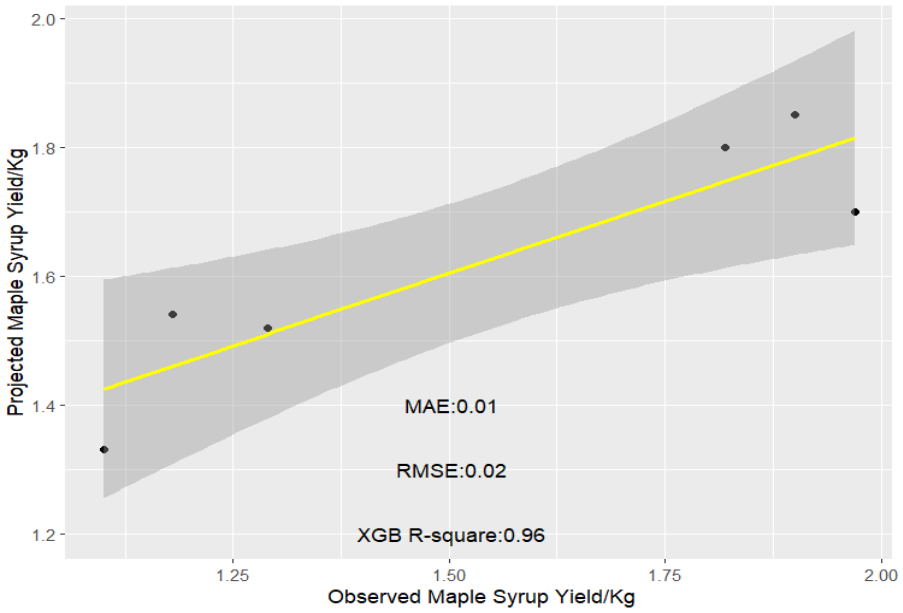


Figure 11. Projected versus observed maple syrup yield in Montérégie

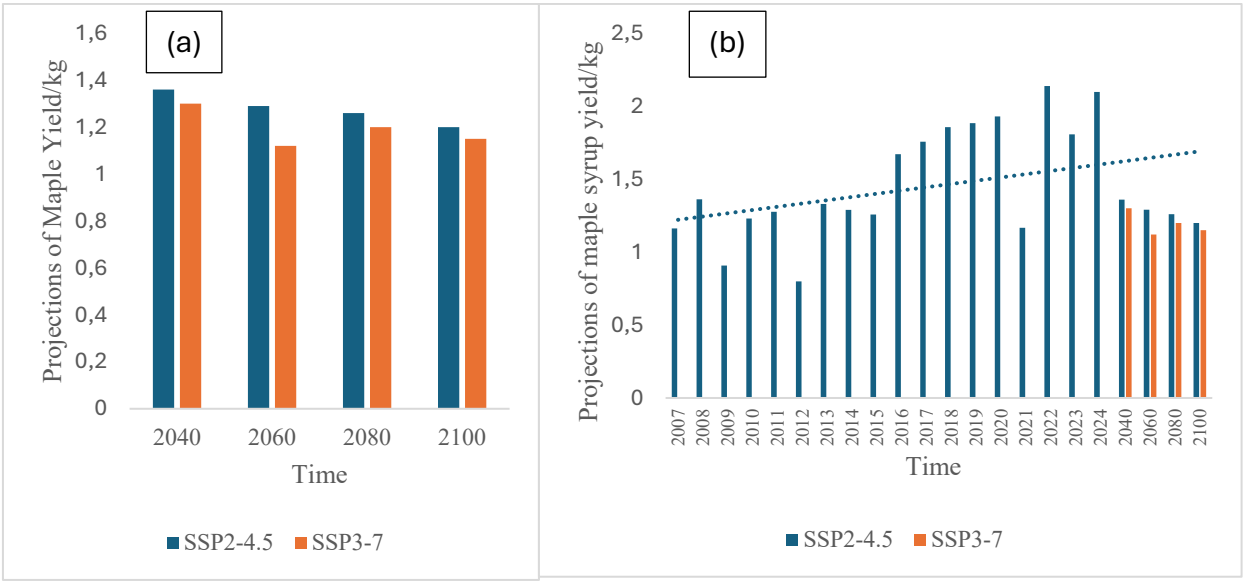
3.7 Projections of Maple Syrup Yield in Montérégie

The projections indicate that rising temperatures and changing precipitation patterns are likely to have a pronounced negative effect on maple syrup yield in the Montérégie region of Quebec by 2100. Unlike the other regions examined in this study, including Centre-du-Québec, Chaudière-Appalaches, and Laurentides, Montérégie exhibits a persistent decline in projected maple syrup yield throughout the 2040–2100 period (Figure 12b).

Relative to the historical baseline period of 2020–2024, maple syrup yield is projected to decline by 30.11% under the moderate-emissions scenario (SSP2-4.5) and 34.79% under the high-emissions scenario (SSP3-7.0) (Figure 12c). By 2100, projected yields are expected to fall below 2 kg/tap/year under both scenarios, indicating a substantial reduction in maple syrup productivity in the region.

Comparison with the projections for Centre-du-Québec suggests that Montérégie may experience a greater climate-related reduction in maple syrup yield over the projection period. The stronger response may be associated with the region's relatively warmer baseline conditions, which could make maple trees more vulnerable to additional warming. Under the SSP3-7.0 scenario, the larger projected decline of 34.79%, compared with 30.11% under SSP2-4.5, further suggests that greater warming could exacerbate the negative effects on maple syrup production in Montérégie.

Overall, the results identify Montérégie as the most climate-sensitive of the studied regions, based on the magnitude and persistence of the projected yield decline. However, attributing the regional differences solely to temperature requires caution, as precipitation variability and other climatic and environmental factors may also contribute to the projected changes in maple syrup yield.



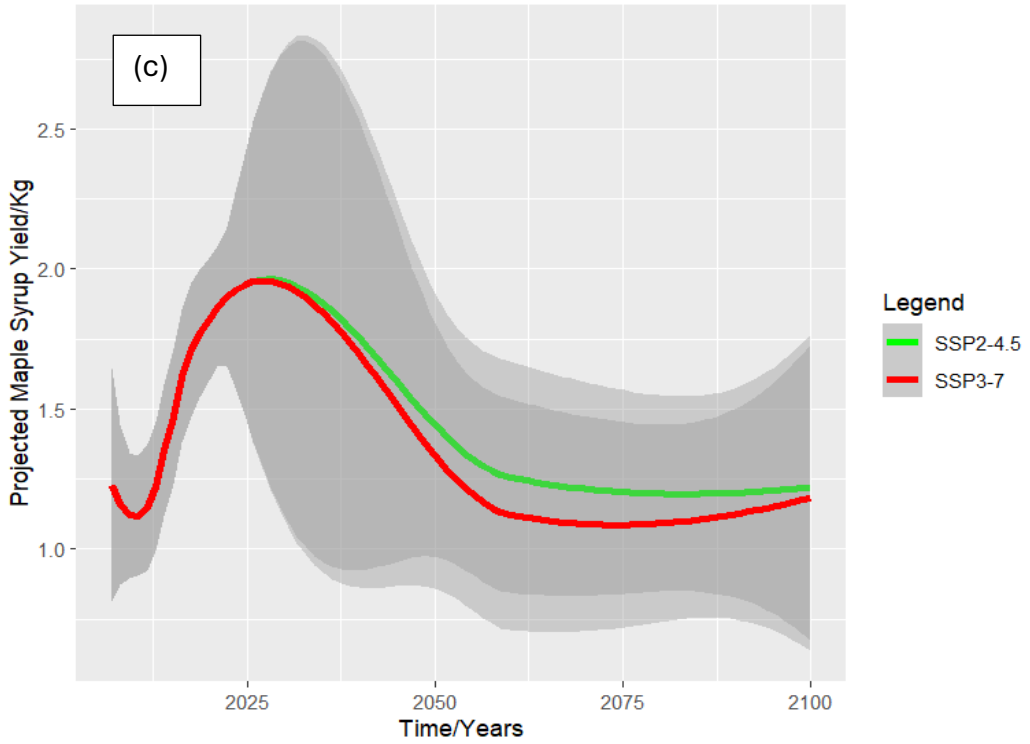


Figure 12. Projections of maple syrup yield in Montérégie**,** (a) histogram comparing the projections under the SSP2-4.5 and SSP3-7.0 scenario, (b) histogram showing the observed and projected yield, (c) curve of the observed and projected yield

3.8 Plot of projected versus observed yield in Estrie

The analysis shows a good fit between the projected yield against the observed yield in the Estrie region of Quebec. The RMSE of the model is 0.17kg/tap/year and the MAE is 0.02kg/tap/year (Figure 13). The R^2 value of the relationship is 0.97 and implies that 97% of the change in the projected yield is explained by observed yield. The values of these statistical metrics indicate the model's strong prediction accuracy between the observed and projected yield in the Estrie region of Quebec.

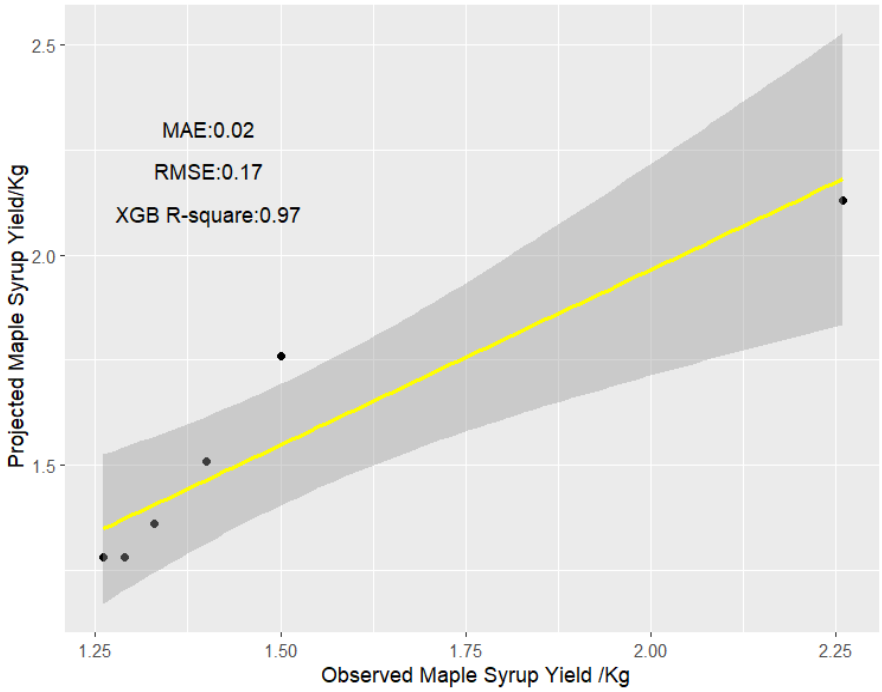
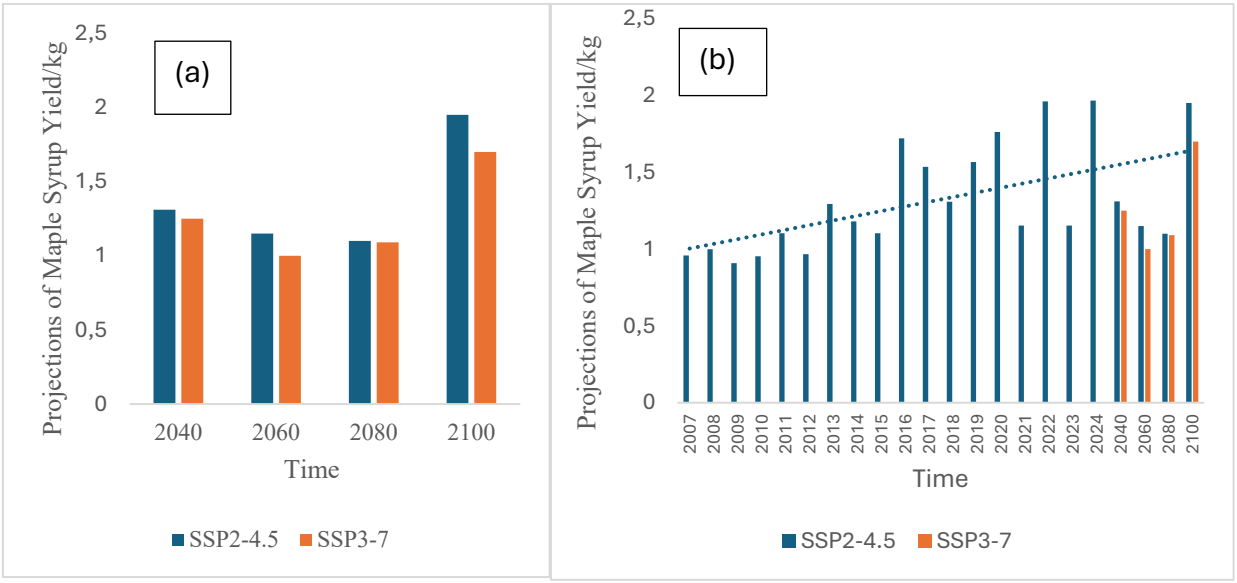


Figure 13. Projected versus observed maple syrup yield in the Estrie region of Quebec

3.9 Projections of Maple Syrup Yield in Estrie

The projections indicate varying impacts of rising temperatures and changing precipitation patterns on maple syrup yield in Estrie by 2100 (Figure 14c). Overall, the results show a persistent decline in yield between 2040 and 2080, followed by an increase by 2100 (Figure 14). The projected trend in maple syrup yield for Estrie is like that observed in Centre-du-Québec. Furthermore, the analysis indicates that maple syrup yield would decrease by 13.85% and 20.20% under the moderate (SSP2-4.5) and high (SSP3-7.0) emission scenarios, respectively, relative to the historical period of 2020–2024 (Figure 14b). These findings indicate that the high-emission scenario (SSP3-7.0) would have a greater negative impact on maple syrup yield in Estrie than the moderate-emission scenario (SSP2-4.5) (Figure 14a).



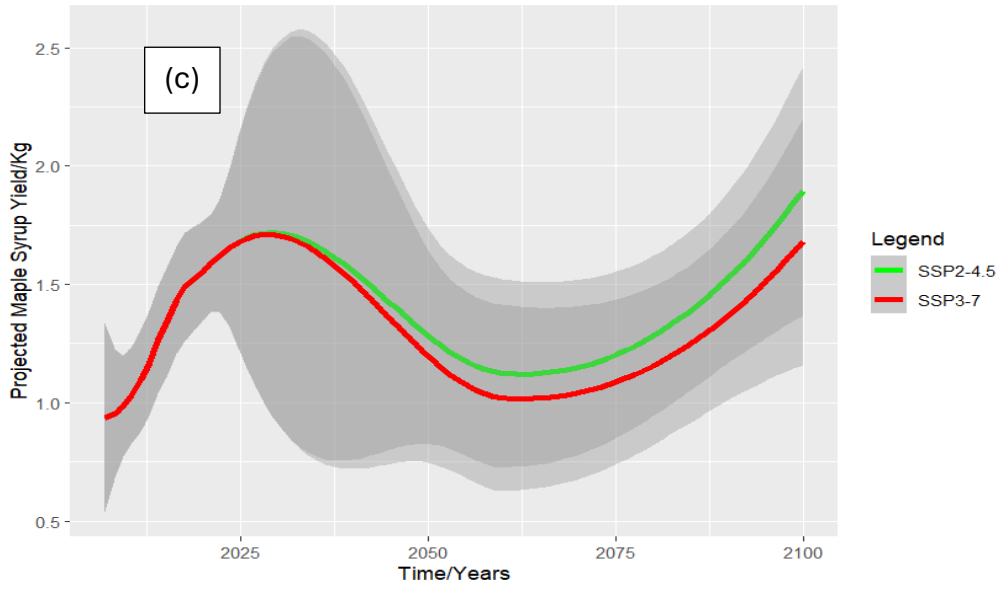


Figure 14. Projections of maple syrup yield in the Estrie region of Quebec, (a) histogram comparing the projections under the SSP2-4.5 and the SSP3-7.0 scenarios, (b) histogram showing the observed and the projected yield, (c) curve of the observed and projected yield

3.10 Plot of the projected versus observed maple syrup yield in Chaudière Appalaches

The plot shows a strong fit of the projected against the observed maple syrup yield in Chaudière Appalaches. The RMSE of the model is 0.08 kg/tap/year and the MAE is 0.05 kg/tap/year (Figure 15). Meanwhile, the R^2 of the model is 0.97 and indicates that the observed yield contributes to explain 97% of the change in the predicted maple syrup yield. These predictive performance indicators indicate a good model's precision and prediction accuracy for maple syrup yield in the Chaudiére Appalaches region of Quebec.

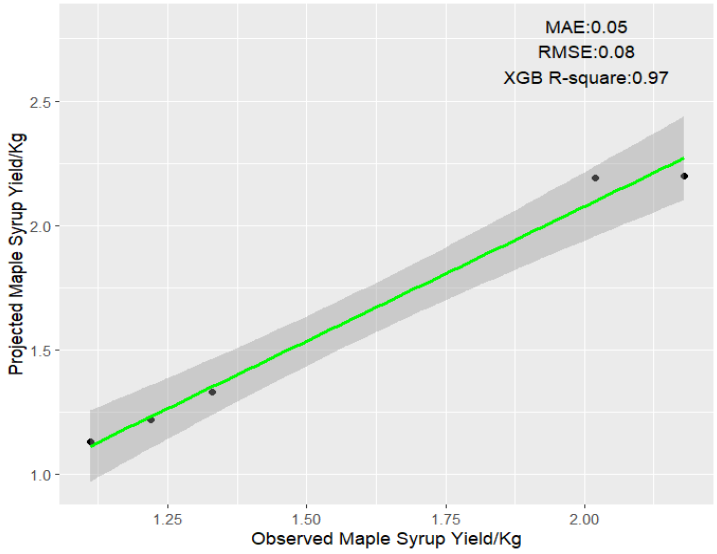


Figure 15. Projected Versus observed maple syrup yield in Chaudière Appalaches

3.11 Projected Maple Syrup Yield in Chaudière-Appalaches

The projections indicate that rising temperatures and changing precipitation patterns will have varying effects on maple syrup yield in Chaudière-Appalaches (Figure 16). The results show a substantial decline in yield between 2040 and 2080, followed by an increase by 2100; however, projected yield remains well below 2 kg/tap/year. The future projections for maple syrup yield in this region show a trend like those observed in Centre-du-Québec and Estrie. Overall, the analysis indicates that maple syrup yield would decrease by 13.93% and 18.07% under the moderate (SSP2-4.5) and high (SSP3-7.0) emission scenarios, respectively, relative to the historical period. The results further show that the high-emission scenario (SSP3-7.0) would have a greater negative impact on maple syrup yield in Chaudière-Appalaches than the moderate-emission scenario (SSP2-4.5) (Figure 16a). Moreover, under the SSP3-7.0 scenario, the negative impact of rising temperatures and changing precipitation patterns on maple syrup yield is greater in Chaudière-Appalaches than in Centre-du-Québec. This suggests that the impacts of climate change on maple syrup yield vary spatially across regions of Quebec.

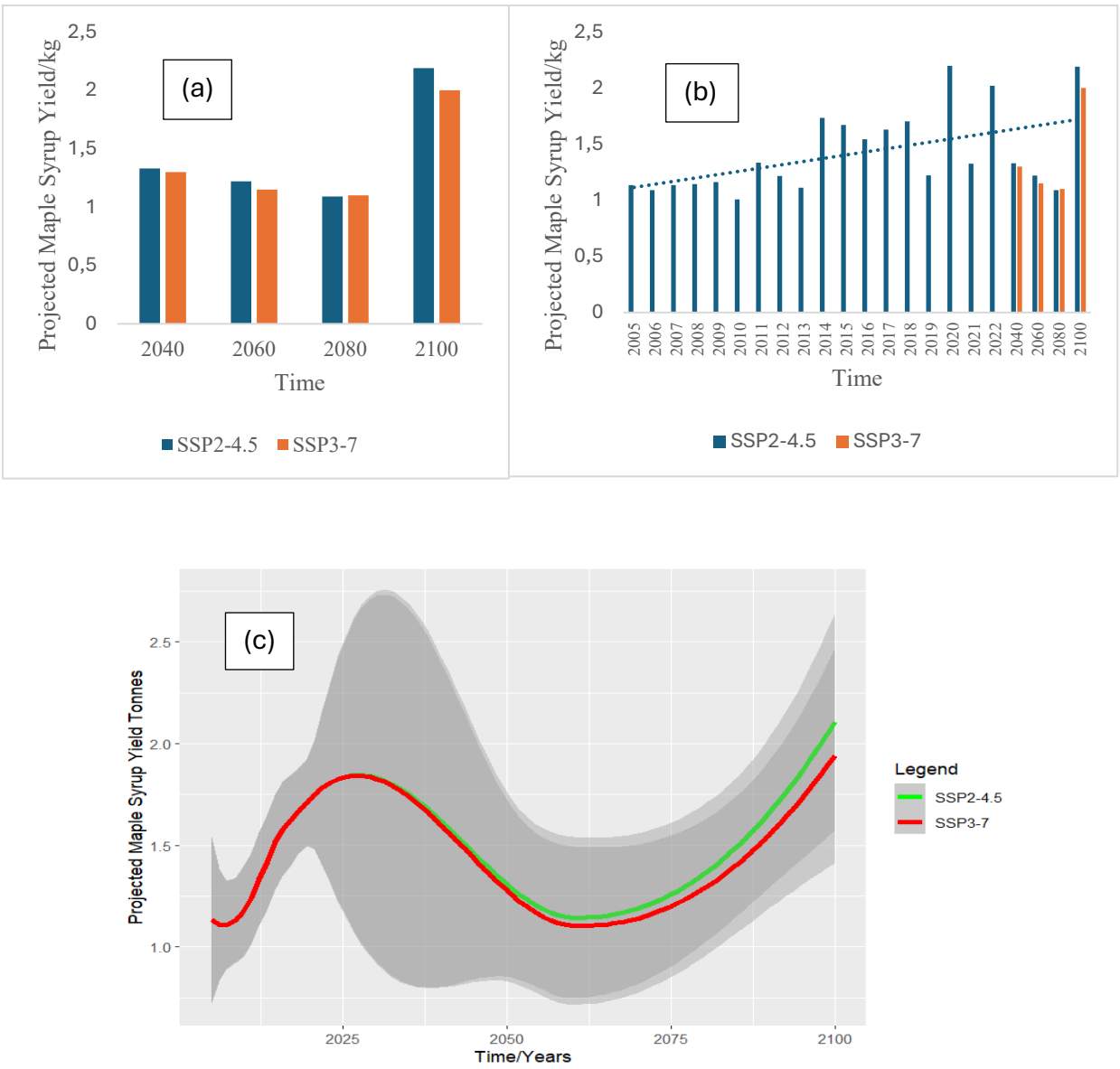


Figure 16. Projections of maple syrup yield in Chaudière Appalaches (a) histogram comparing the projections under the moderate (SSP2-4.5) and high emission (SSP3-7.0) scenarios, (b) histogram showing the observed and projected yield, (c) curve of the observed and projected yield.

3.12 Plot of the projected versus observed maple syrup yield in the Laurentian region of Quebec

The plot shows a strong fit of the projected yield against the observed yield in the Laurentian region of Quebec. The RMSE of the model is 0.03kg/tap/year and the MAE is 0.03kg/tap/year (Figure 17). The R^2 of the model is 0.97 and implies that 97% of the change in the projected maple syrup yield is explained by the observed yield. The values of these metrics indicate a high model prediction accuracy for maple syrup yield in the Laurentian region of Quebec

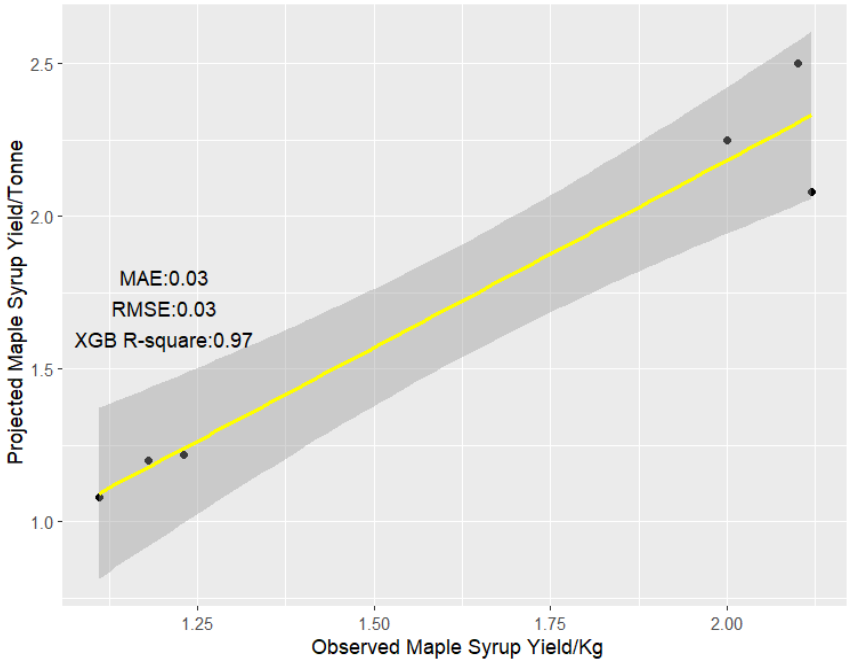


Figure 17. Projected versus observed maple syrup yield in Laurentian region of Quebec

3.13 Maple Syrup Projections in the Laurentian Region of Quebec

The projections indicate that rising temperatures and changing precipitation patterns will have varying impacts on maple syrup yield in the Laurentian region of Quebec (Figure 17). The results show a relatively lower projected yield by 2080 compared with the other projection periods (2040–2060) (Figure 17c). Furthermore, the analysis indicates that maple syrup yield in the Laurentian region would decrease by 18.36% and 23.22% under the moderate (SSP2-4.5) and high (SSP3-7.0) emission scenarios, respectively, relative to the historical period of 2018–2022 (Figure 17b). These findings suggest that the impacts of rising temperatures and changing precipitation patterns on maple syrup yield would be greater in the Laurentian region than in Centre-du-Québec and Chaudière-Appalaches. This greater impact may be associated with the higher projected future temperatures in the Laurentian region of Quebec.

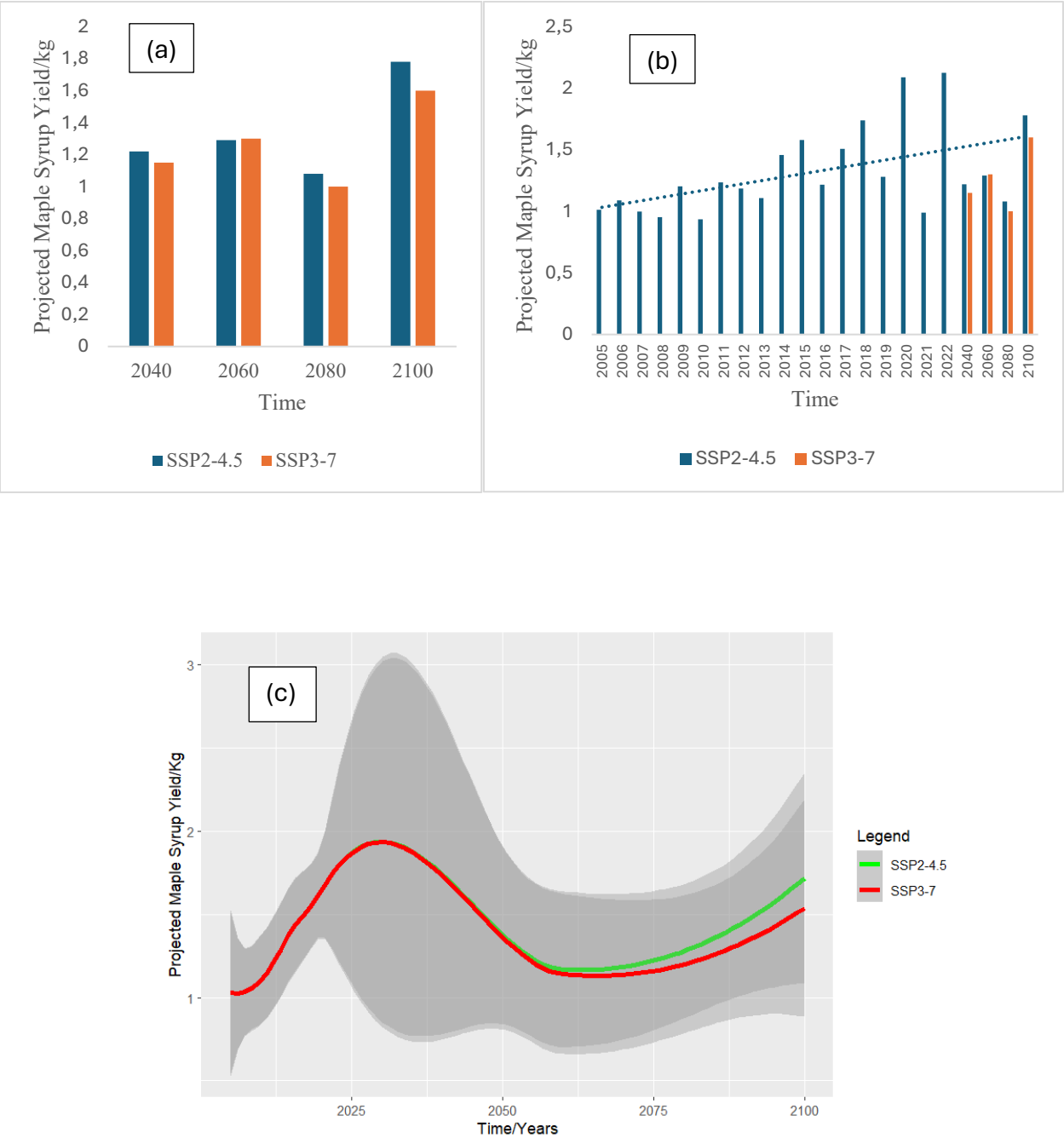


Figure 17. Projections of maple syrup yield in the Laurentian region of Quebec, (a) histogram comparing the projections under the SSP2-4.5 and SSP3-7.0 scenarios, (b) histogram showing the observed and projected yield, (c) curve of the observed and projected yield.

4. Discussion

This study employed four machine learning (ML) models—Random Forest (RF), Support Vector Regression (SVR), Gradient Boosting Regression (GBR), and Extreme Gradient Boosting Regression (XGBoost)—to assess the impacts of rising temperatures and changing precipitation patterns on maple syrup yield at both the provincial level (Quebec) and regional level (Centre-du-Québec, Estrie, Chaudière-Appalaches, Laurentides, and Montérégie). Among the four models, XGBoost demonstrated the highest predictive performance, with an RMSE of 0.04 kg/tap/year, an MAE of 0.03 kg/tap/year, and a coefficient of determination (R^2) of 0.99. The superior predictive performance of XGBoost may be attributed to its regularization techniques, including L1 (lasso) and L2 (ridge) regularization, which help reduce overfitting and improve model generalization and predictive accuracy.

The superior performance of the XGBoost over the other machine learning is consistent with recent application of machine learning models in crop yield prediction. For instance, Gharakhanlou and Perez (2024) employed five ML models—AdaBoost, XGBoost, LightGBM, GBM, and RF—for crop yield prediction across Canada and revealed that the XGBoost outperformed the other models, achieving an R^2 of 0.872, an RMSE of 119.48, and an MAE of 68.70 for canola, and an R^2 of 0.972, an RMSE of 102.39, and an MAE of 39.47 for soybean. Also, our findings corroborate those of Harraf et al. (2026), who assessed the impacts of climate change on crop production in North Africa using Random Forest, Extreme Gradient Boosting Regression, and Multiple Linear Regression (MLR). Their study identified XGBoost as the best-performing predictive model for crop yield. However, our findings differ from those of Dehillon et al. (2024), who employed four ML models to predict maize, and soybean yields in Ohio, United States, and found that Random Forest outperformed the other models. These differences highlight the importance of crop type, climatic conditions, geographical context, and model characteristics in determining the predictive performance of ML approaches.

The results further indicate that climate change may have a nonlinear effect on maple syrup yield in Quebec. The projections show a declining yield trend between 2040 and 2060, followed by an increasing trend between 2080 and 2100. The projected increase in yield during the later period may be associated with temperatures becoming closer to the optimal thermal conditions required for maple syrup production, which range from approximately 0°C at night to above 4°C during the day (Pothier, 1995). Our projections for 2080–2100 are consistent with the findings of Rapp et al. (2019), who examined shifting optimal temperatures for maple syrup production in North America and reported that climate change could have a positive effect on maple syrup production in Quebec. However, the nonlinear response observed in our projections also suggests that the relationship between temperature and maple syrup yield is complex and may vary across different future periods.

At the regional level, the study revealed considerable spatial variation in the impacts of rising temperatures and changing precipitation patterns on maple syrup yield. In Centre-du-Québec,

Estrie, Chaudière-Appalaches, and Laurentides, the projections indicate a decline in maple syrup yield in the short term (2040–2060), followed by an increase during the longer-term period (2080–2100). This regional pattern is broadly consistent with the provincial-level projections and suggests that the effects of climate change on maple syrup production may vary according to local climatic conditions.

Several factors, including technological improvements and the development and implementation of more effective adaptation strategies, may contribute to the projected increase in maple syrup yield over the long term in these regions. Improvements in tapping technologies, sap collection systems, forest management practices, and other adaptive measures could potentially offset some of the negative effects of changing climatic conditions. This interpretation is consistent with the findings of Rapp et al. (2019), who reported that rising temperatures could have an overall positive effect on maple syrup production in Quebec. Moreover, our findings are broadly consistent with Filteau et al. (2025), who found that increasing temperatures could improve the quality of maple syrup.

In contrast, the projections for Montérégie indicate a continued decline in maple syrup yield by 2100. The analysis further shows that the moderate-emission scenario (SSP2-4.5) would have a less negative impact on yield than the high-emission scenario (SSP3-7.0) in this region. This stronger negative response under SSP3-7.0 may be associated with the relatively higher projected temperatures in Montérégie by 2100, which could exceed the optimal thermal conditions for maple syrup production. Overall, these regional differences demonstrate that the impacts of climate change on maple syrup yield are spatially heterogeneous across Quebec and that the magnitude and direction of future changes may depend strongly on regional climatic conditions and the extent of future warming.

5. Limitations of the Study

This study has several limitations that should be considered when interpreting the results. First, the analysis primarily used annual and seasonal temperature and precipitation variables to predict maple syrup yield in Quebec through 2100. Other factors that may influence maple syrup production, including socioeconomic characteristics such as farmers' education, farming experience, access to technology, and management practices, were not incorporated into the models. These factors may influence producers' capacity to adapt to changing climatic conditions and, consequently, future maple syrup yields.

Second, the climate projections used in this study are based on CMIP6 climate model simulations rather than directly observed future temperature and precipitation data. Although CMIP6 provides scientifically robust climate projections, simulated climatic variables may differ from observed conditions because of model uncertainty, differences in spatial resolution, and the inherent uncertainty associated with future climate projections. Consequently, the projected maple syrup yields should be interpreted as estimates rather than precise predictions of future production.

Third, this study considered only the moderate (SSP2-4.5) and high (SSP3-7.0) emission scenarios. Although these scenarios represent plausible future climate pathways, excluding other SSP scenarios limits the range of potential climate outcomes considered in the analysis. Including additional scenarios in future studies could provide a more comprehensive assessment of the uncertainty surrounding future maple syrup production.

Finally, the findings are geographically limited to Quebec, even though maple syrup production is also important in other Canadian provinces. Therefore, the results may not be directly applicable to other maple syrup-producing regions with different climatic, ecological, and socioeconomic conditions. Despite these limitations, the study provides important insights into the potential future impacts of climate change on maple syrup production in Quebec and offers a basis for further research and adaptation planning.

6. Conclusion

This study assessed the predictive performance of four machine learning models—Extreme Gradient Boosting (XGBoost), Gradient Boosting (GB), Random Forest (RF), and Support Vector Machine (SVM)—and subsequently employed the best-performing model to project maple syrup yield in Quebec through 2100. XGBoost outperformed the other machine learning models, achieving an R^2 of 0.99, an RMSE of 0.04 kg/tap/year, and an MAE of 0.03 kg/tap/year. The high predictive performance of XGBoost demonstrates its potential for modelling the relationship between climatic variables and maple syrup yield.

The projections indicate that maple syrup yield in Quebec would decrease by 16.54% and 15.43% under the moderate (SSP2-4.5) and high (SSP3-7.0) emission scenarios, respectively, over the 2040–2100 projection period relative to the historical period of 2020–2024. However, the regional-scale projections reveal substantial spatial variation in the impacts of climate change on maple syrup yield across Quebec.

For example, in Centre-du-Québec, maple syrup yield is projected to decline by 13.26% and 16.57% under the SSP2-4.5 and SSP3-7.0 scenarios, respectively. In contrast, Montérégie is projected to experience substantially larger yield losses of 30.11% and 34.79% under the SSP2-4.5 and SSP3-7.0 scenarios, respectively. These differences highlight the spatially heterogeneous nature of climate change impacts on maple syrup production across Quebec.

Overall, the findings suggest that climate change will pose important challenges to the sustainability of maple syrup production in Quebec, although the magnitude and direction of these impacts will vary among regions. The study also demonstrates the potential value of machine learning approaches for assessing climate-related risks to agricultural and forest-based production systems. Future research should incorporate socioeconomic characteristics, forest management practices, technological advancements, and adaptation strategies to provide a more comprehensive assessment of future maple syrup production. Although the present analysis is limited to Quebec, its findings provide important insights that can support policy development, climate adaptation planning, and efforts to improve the sustainability and resilience of maple syrup production in the province.

References

Baker, I., Peterson, A., Brown, G., & McAlpine, C. (2012). Local government response to the impacts of climate change: An evaluation of local climate adaptation plans. *Landscape and urban planning*, *107*(2), 127-136.

Beier, C. M., Woods, A. M., Hotopp, K. P., Gibbs, J. P., Mitchell, M. J., Dovčiak, M., ... & Page, B. D. (2012). Changes in faunal and vegetation communities along a soil calcium gradient in northern hardwood forests. *Canadian Journal of Forest Research*, *42*(6), 1141-1152.

Biau, G., Cadre, B., & Rouvière, L. (2019). Accelerated gradient boosting. *Machine learning*, *108*(6), 971-992.

Bishop, D. A., Beier, C. M., Pederson, N., Lawrence, G. B., Stella, J. C., & Sullivan, T. J. (2015). Regional growth decline of sugar maple (Acer saccharum) and its potential causes. *Ecosphere*, *6*(10), 1-14.

Bootsma, A., Gameda, S., & McKenney, D. W. (2005). Potential impacts of climate change on corn, soybeans and barley yields in Atlantic Canada. *Canadian journal of soil science*, *85*(2), 345-357.

Bouras, E. H., Jarlan, L., Er-Raki, S., Balaghi, R., Amazirh, A., Richard, B., & Khabba, S. (2021). Cereal yield forecasting with satellite drought-based indices, weather data and regional climate indices using machine learning in Morocco. *Remote Sensing*, *13*(16), 3101.

Campbell, B. (2022). Climate change impacts and adaptation options in the agrifood system – A summary of recent IPCC Sixth Assessment Report findings. Rome, FAO. https://doi.org/10.4060/cc0425en

Caughron, A., Legault, S., Haut, C., Houle, D., & Reynolds, T. W. (2021). A changing climate in the maple syrup industry: Variation in Canadian and USA producers' climate risk perceptions and willingness to adapt across scales of production. *Small-Scale Forestry*, *20*(1), 73-95.

Cheng, K., Lu, Z., Zhou, Y., Shi, Y., & Wei, Y. (2017). Global sensitivity analysis using support vector regression. *Applied Mathematical Modelling*, *49*, 587-598.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Crane-Droesch, A. (2018). Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environmental Research Letters*, *13*(11), 114003.

Courville, S. (2009). *Quebec: a historical geography*. ubc Press.

Duchesne, L., Houle, D., Côté, M. A., & Logan, T. (2009). Modelling the effect of climate on maple syrup production in Québec, Canada. *Forest ecology and management*, *258*(12), 2683-2689.

Dhillon, R., Takoo, G., Sharma, V., & Nagel, M. (2024). Utilizing machine learning framework to evaluate the effect of climate change on maize and soybean yield. *Computers and Electronic in Agriculture, 224, 108982.*

Epule Epule, T., Poirier, V., Peng, C., & Etongo, D. (2024). Vulnerability of barley yields in Québec. *Environmental Monitoring and Assessment*, *196*(9), 850.

Filteau, M. (2025). Will climate change affect the quality of maple syrup?. *Heliyon*, *11*(12).

Feng, P., Wang, B., Li Liu, D., Waters, C., Xiao, D., Shi, L., & Yu, Q. (2020). Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. *Agricultural and Forest Meteorology*, *285*, 107922.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Gavin, D. G., Beckage, B., & Osborne, B. (2008). Forest dynamics and the growth decline of red spruce and sugar maple on Bolton Mountain, Vermont: a comparison of modeling methods. *Canadian Journal of Forest Research*, *38*(10), 2635-2649

Gawdiya, S., Kumar, D., Ahmed, B., Sharma, R. K., Das, P., Choudhary, M., & Mattar, M. A. (2024). Field scale wheat yield prediction using ensemble machine learning techniques. *Smart Agricultural Technology*, *9*, 100543.

Gharakhanlou, N. M., & Perez, L. (2024). From data to harvest: Leveraging ensemble machine learning for enhanced crop yield predictions across Canada amidst climate change. *Science of The Total Environment*, *951*, 175764.

Government of Quebec. (2026). Production de sirop d'érable. https://www.quebec.ca/agriculture-environnement-et-ressources-naturelles/agriculture/industrie-agricole-au-quebec/productions-agricoles/production-sirop-erable-acericulture

Government of Canada. (2025). Maple syrup production and climate change: does the future test as sweet? https://natural-resources.canada.ca/forest-forestry/state-canada-forests/maple-syrup-production-climate-change-does-future-taste-sweet

Government of Quebec. (2025). Aperçu du territoire Québécois. https://www.quebec.ca/gouvernement/portrait-quebec/geographie-territoire/apercu

Government of Canada. (2025). Agriculture, crop production, animal production and aquaculture, Quebec. https://www.jobbank.gc.ca/trend-analysis/job-market-reports/quebec/sectoral-profile-agriculture

Government of Canada. (2024). Principal production of field crops, November 2024. https://www150.statcan.gc.ca/n1/daily-quotidien/241205/dq241205b-eng.htm

Gujarati, D. N., & Porter, D.C (2009). Basic econometrics*: McGraw-Hill*.

Harraf, R. B., Mhada, M., & El Moçayd, N. (2026). Climate change and crop production in North Africa: Insights from machine learning models. *Journal of Environmental Management*, *398*, 128451.

Houle, D., Paquette, A., Côté, B., Logan, T., Power, H., Charron, I., & Duchesne, L. (2015). Impacts of climate change on the timing of the production season of maple syrup in Eastern Canada. *PLoS One*, *10*(12), e0144844.

Hultgren, A., Carleton, T., Delgado, M., Gergel, D. R., Greenstone, M., Houser, T., ... & Yuan, J. (2025). Impacts of climate change on global agriculture accounting for adaptation. *Nature*, *642*(8068), 644-652.

Li, C., Camac, J., Robinson, A., & Kompas, T. (2025). Predicting changes in agricultural yields under climate change scenarios and their implications for global food security. *Scientific Reports*, *15*(1), 2858.

Lovett, G. M., Weathers, K. C., Arthur, M. A., & Schultz, J. C. (2004). Nitrogen cycling in a northern hardwood forest: do species matter?. *Biogeochemistry*, *67*(3), 289-308.

Lucash, M. S., Yanai, R. D., Blum, J. D., & Park, B. B. (2012). Foliar nutrient concentrations related to soil sources across a range of sites in the Northeastern United States. *Soil Science Society of America Journal*, *76*(2), 674-683.

Moss, A., Raposo, L., Thorn, C. (2024). Evaluating optimal sites for maple syrup farms in Quebec' s Eastern Townships using ArcGlS Suitability Modeler. *Centre of Geographic Sciences.* https://scholars.esri.ca/wpcontent/uploads/2024/06/Esri2024_Report_Suitability Analysis.pdf

Moreau, G., Achim, A., & Pothier, D. (2020). An accumulation of climatic stress events has led to years of reduced growth for sugar maple in southern Quebec, Canada. *Ecosphere*, *11*(7), e03183.

Ouranos. (2014). Sommaire de la synthèse des connaissances sur les changements climatiques au Québec. Édition 14. Montréal, Québec, Ouranos.

Ouranos. (2022). What are the key takeaways of the IPPC report on climate change impacts and adaptation. https://www.ouranos.ca/en/news/2024-03-07/what-are-key-takeaways-ipcc-report-climate-change-impacts-and-adaptation

Ouranos. (2026). Agriculture. https://www.ouranos.ca/en/topics-of-interest/agriculture-background

Parra, E., Dimou, C., Llorens, J., Moreno, V., & Fraga, A. (2015). A methodology for the classification of quality of requirements using machine learning techniques. *Information and Software Technology*, *67*, 180-195.

Pothier, D. (1995). Effets des coupes d'éclaircie et des variations climatiques interannuelles sur la production et la teneur en sucre de la sève d'une érablière. *Canadian Journal of Forest Research*, *25*(11), 1815-1820.

Rapp, J. M., Lutz, D. A., Huish, R. D., Dufour, B., Ahmed, S., Morelli, T. L., & Stinson, K. A. (2019). Finding the sweet spot: Shifting optimal climate for maple syrup production in North America. *Forest Ecology and Management*, *448*, 187-197.

Sidhu, B. S., Mehrabi, Z., Ramankutty, N., & Kandlikar, M. (2023). How can machine learning help in understanding the impact of climate change on crop yields? *Environmental Research Letters*, *18*(2), 024008.

Skinner, C. B., DeGaetano, A. T., & Chabot, B. F. (2010). Implications of twenty-first century climate change on Northeastern United States maple syrup production: impacts and adaptations. *Climatic change*, *100*(3), 685-702.

St-Marseille, A. F. G., Bourgeois, G., Brodeur, J., & Mimee, B. (2019). Simulating the impacts of climate change on soybean cyst nematode and the distribution of soybean. *Agricultural and forest meteorology*, *264*, 178-187.

Tian, Y., Xu, Y. P., & Wang, G. (2018). Agricultural drought prediction using climate indices based on Support Vector Regression in Xiangjiang River basin. *Science of the total environment*, *622*, 710-720.

von Bloh, M., Lobell, D., & Asseng, S. (2024). Knowledge informed hybrid machine learning in agricultural yield prediction. *Computers and Electronics in Agriculture*, *227*, 109606.

von Bloh, M., Júnior, R. D. S. N., Wangerpohl, X., Saltık, A. O., Haller, V., Kaiser, L., & Asseng, S. (2023). Machine learning for soybean yield forecasting in Brazil. *Agricultural and Forest Meteorology*, *341*, 109670.

Wang, Q., Abah, E. O., & Cannella, M. (2025). The maple syrup industry in Canada and the United States: challenges and potential strategies towards a more sustainable development. *Frontiers in Sustainable Food Systems*, *9*, 1562645.

Zhang, X., Flato, G., Kirchmeler, M., Vincent, L., Wan, H., Wang, X., Rong, R., Fyfe, J., Li, G. & Kharin, V. V. (2019). Les changements de température et de précipitations au Canada. Dans Bush, E. & Lemmen, D. S. (eds.), *Rapport sur le climat changeant du Canada*, 113-194. Government of Canada. https://doi.org/10.4095/327812

Zhang, C., & Ma, Y. (2012). *Ensemble machine learning* (Vol. 144). New York: springer.