**1.Forecasting Global Developments and Challenges in Olive Oil Supply and Demand: A Delphi Survey from Spain**

**Discusses mainly the regions involved with olive oil production**

* Tunisia > Spain > Italy hold the majority share of export market of Olive oil. (Tunisia being 3rd largest in the world)
* Italy is worlds leading importer of Olive oil.
* Another defining feature of the international olive oil market is the absence of an official, universally recognized price benchmark. There are, instead, more or less representative regional markets.
* The key characteristics of the Delphi method are (i) anonymity of the participants in order to avoid the persuasive effect of dominant members of the group [32]; (ii) an iterative process: Delphi runs in two or more rounds of structured questionnaires to allow information exchange within the panel [27,31,33,34], and the process is repeated until the stability of the responses is achieved, irrespectively of whether or not a consensus is reached [35]; (iii) controlled feedback.
* In this study, we applied the Delphi method in four steps. First, the research objectives and the conceptual model were formulated. Then, a final questionnaire was prepared after several preliminary versions, and a rigorous selection of the participating experts was performed. Subsequently, the Delphi survey was implemented in two rounds during 2017. We used two rounds as a result of the satisfactory level of stability of responses attained in the second round, which makes further rounds fruitless according to the literature on ending the iterative process in Delphi studies (see Section 3.2). Finally, the major findings of the study were portrayed and analyzed.

**2. Impact of Climate Change on Olive Crop Production in Italy**

**Discusses mainly the factors affecting the Olive Crop Production**

* Numerous studies have pointed out the importance of olive yield forecasting of airborne pollen amounts and weather conditions during the months following olive pollination [4,6,7]. However, pollen monitoring of olive groves has other valuable purposes from an agronomic point of view besides predicting olive yield as it allows to study the olive reproductive cycle on large areas of olive orchards as well [8].
* Olive trees are influenced by human factors (e.g., agronomic techniques), environmental conditions (water deficit, extreme temperature) and phyto-pathological problems [6,9] during both the pre-flowering and post-pollination periods.
* Experimentation and Methodology involved - Identifying Olive cultivation areas, Olive Flowering Monitoring, Climate Change Diagnosis, Statistical Analysis and Future Projections
* Results - The most important olive cultivation areas in Central-South Italy belong to the province of Bari, which stands out from the rest with almost 120,000 ha of olive groves on average. Bari is followed by other provinces in the Southeast of the country, such as Lecce, Brindisi, and Foggia, all of them belonging to the Apulia region. Another important region is Calabria in the Southern limits of the Italian Peninsula, where the province of Reggio Calabria accounts for more than 50,000 ha of olive groves (Table 1).
* Olive production values showed high levels of variability over the years, but pollen emission values showed even higher coefficients of variation, which in most cases overcame 50% (Table 1). These results highlighted the great alternating behavior of pollen and fruit production in these Italian olive orchards.

**3. Flower and pollen production in the ‘Cornicabra’ olive (Olea europaea L.) cultivar and the influence of environmental factors**

*Taken from the citations of the above research paper*

* Key message The olive pollen production showed a variation related to temperature and rainfall during the dormancy period. A correlation was found between the number of olive tree inflorescences and airborne pollen counts.

**4. Key message The olive pollen production showed a variation related to temperature and rainfall during dormancy period. A correlation was found between the number of olive tree inflorescences and airborne pollen counts.**

* Shows the negative impact of abandonment and intensive regime of olive groves. Also examines the vulnerability of olive groves to climate change.

**5. Projected climate changes are expected to decrease the suitability and production of olive varieties in southern Spain**

* Used ‘biomod’ 2 R package. The algorithm implemented in this function randomly shuffles each variable and then model predictions are made with this 'shuffled' dataset. Then, Pearson’s correlation coefficient (Pc) is calculated for the relationship between the reference predictions (not shuffled) and the 'shuffled' ones, returning the importance score (I) equal to I = 1 - Pc. The higher the value, the greater the influence that the variable has on the model (a value of zero assumes no influence). We calculated variable importance scores for each selected model included in the average ensemble.
* The algorithm implemented in this function randomly shuffles each variable and then model predictions are made with this 'shuffled' dataset. Then, Pearson’s correlation coefficient (Pc) is calculated for the relationship between the reference predictions (not shuffled) and the 'shuffled' ones, returning the importance score (I) equal to I = 1 - Pc. The higher the value, the greater the influence that the variable has on the model (a value of zero assumes no influence). We calculated variable importance scores for each selected model included in the average ensemble.
* Using this model, we showed that olive production is related to the suitability of the area (for all varieties); therefore, future -suitability maps can be used to obtain projections of total fruit production.
* Further, we also obtained projections to assess the alteration in the trend of annual production due to changes in the environmental suitability and the distribution of the olive varieties as a consequence of climate change

**6. Assessing the long-term impact of climate change on olive crops and olive fly in Andalusia, Spain, through climate indices and return period analysis**

* The long-term climate change impacts on olive crops in Andalusia (South Spain) were examined using specially tailored climatic indicators along with the return period method, which is useful for risk description and quantification. The climatic indices used for this study have been identified in the framework of the Med-Gold Horizon 2020 project as the most appropriate for the olive sector. The return levels of interest for the calculation of the occurrence probabilities were estimated by the upper tail of the reference distribution. Both the climatic indices and the occurrence probabilities of the return period analysis were calculated using an ensemble of high-resolution, bias adjusted RCM-GCM model simulations from the EURO-CORDEX initiative.

**7. Mediterranean Olive Orchards under Climate Change: A Review of Future Impacts and Adaptation Strategies**

This might be a very useful paper for writing literature review

* The Mediterranean Basin is considered a climate change “hotspot” [123], since future projections hint at considerable warming trends and an increase of consecutive dry days [124], leading to an overall increase in aridity. In this context, climate change may be- come particularly challenging for olive growers [21]. The growing evidence for substantial climate change in the upcoming decades urges adaptation measures to be taken. As such, the development of future climate projections, based on feasible future socio-economic storylines, is of great value, as they provide objective information that can be used in developing suitable adaptation/mitigation managements to minimize climate change’s im- pacts on the environment and human activities. In effect, a single adaptation strategy may not be sufficient to counteract the negative impacts of climate change [99]. In a study by Lorite et al. [18], the authors indicated that the best adaptation measure was a combination of using cultivars with early flowering dates and regulated deficit irrigation. Nevertheless, to effectively cope with the projected changes, short and long-term strategies deserve much greater attention in future research [125]. Although the adaptation potential of the different strategies to cope with climate change impacts is still unclear [114], they can be highly beneficial for the agricultural sector as a whole [126].

**8. Climate change projections for olive yields in the Mediterranean Basin**

* Crop models were fed with an ensemble of EURO-CORDEX regional climate model data, along with soil and terrain data. For the recent-past, important differences between western and eastern olive growing areas are found. GSL presents a strong latitudinal gradient, with higher/lower values at lower/higher latitudes. Yields are lower in inner south Iberia and higher in Italy and Greece, which is corroborated by historical data. Southern Iberia shows higher GST and lower GSP, which contributes to a higher ETP, lower ETA and consequently stronger WD. Regarding WP, the recent-past values shows similar ranges across Europe. Future projections point to a general increase in GSL along with an increase in GST up to 3oC. GSP is projected to decrease in Western Europe, leading to enhanced WD and consequently a yield decrease (down to -45%). Over eastern European, GSP is projected to slightly increase, leading to lower WD and to a small yield increase (up to +15%). WP will remain mostly unchanged. We conclude that climate change may negatively impact the viability of olive orchards in southern Iberia and some parts of Italy. Thus, adequate and timely planning of suitable adaptation measures are needed to ensure the sustainability of the olive sector.

**9. Life cycle assessment of olive oil: A case study in southern Italy**

* The paper examines the life cycle of the production of a 0.75 L extra virgin olive oil glass bottle based on the results of field analysis in 50 different enterprises in the area of Reggio Calabria, in the southern part of Italy. The study aims to assess the energy and environmental impacts of different scenarios involving conventional and organic cultivations, plains and hills cultivations, and involving different operating techniques. The research has verified that pushing energy efficiency and innovative technologies in the post-agricultural stages has only limited effects in the whole life cycle from an energy and environmental point of view. The Evoline technique proves to be particularly efficient in largely reducing the energy use and environmental impacts within the transformation stages, having potential relevant impacts from the economical and technical side and from the point of view of the stakeholders operating in this segment of the life cycle. However, if a more appropriate holistic point of view is adopted extending the point of view to the whole life cycle, its potential is bound by the limited impact these stages have on the total of the life cycle energy uses and environmental impacts.
* The paper uses the Life Cycle Assessment (LCA) methodology to assess the energy and environmental impacts of different scenarios involving conventional and organic cultivations, plains and hills cultivations, and involving different operating techniques in the production of olive oil in the region of Calabria, in southern Italy. The LCA methodology is a useful tool to assess resource use, energy and environmental burdens related to the full life-cycle of products and services, widely used for analysis of the sustainability of the agro-food sector. The analysis was developed according to the LCA standards of the ISO 14040 series.
* The paper presents the results of a life cycle assessment (LCA) based analysis of the production of olive oil in the region of Calabria, in southern Italy. The study aims to assess the energy and environmental impacts of different scenarios involving conventional and organic cultivations, plains and hills cultivations, and involving different operating techniques. The results show that the first part of the life cycle, from the production, including the growth of the olive plant to the full production stage, is the most relevant for all indicators, with a variable between 80.6% share in the case of the particulate matter indicator to the 99.64% in the case of land use (Hill - Biological agriculture scenario). Relevant differences can be also traced for each specific indicator among all scenarios; high impacts are traced for the agricultural stages among all scenarios (70% −90% in all indicators) with high impacts caused by fertilizers. Among the transformation stages, the bottle production is one of the most relevant sources of life cycle energy uses and environmental impacts (80-90%).

**10. Comparison of five strategies for seasonal prediction of bioclimatic indicators in the olive sector**

* The paper is about assessing the forecast quality of five seasonal forecasting strategies applied to different bioclimatic indicators tailored to the olive sector. The cultivation of olive trees has been established in the Mediterranean region for centuries and is an essential staple in addition to being representative in the cultural, ecological, and economic aspects after its domestication. The EU accounts for a significant portion of the world's olive oil and table olives production, and Mediterranean Spain is one of the main pillars in the olive industry. Olive trees are sensitive to changes in the climate, and the paper aims to provide strategies for seasonal prediction of bioclimatic indicators to aid in the management of the olive orchard.
* The paper used five seasonal forecasting strategies to assess the forecast quality of different bioclimatic indicators tailored to the olive sector. The five indicators were selected based on their importance in the management of the olive orchard. The impact of the increasing share of actual observations included in the computation of the indicators was evaluated by examining the variabilities of correlation coefficients and fair rank probability skill scores for each initialization date. The paper used a methodology where the closest month(s) with available data were used to compute the indicator for each initialization date, and as time progressed through the indicator target period, the months from April to the month before the present were used.
* The paper's results show that blending either seasonal predictions or climatology with observations enhanced the capability of forecasting the tercile category for all the indicators when compared to the use of climatology or seasonal predictions alone. The combination of observations and SEAS5 predictions could outperform the other methods for most of the start months for Spring Maximum Temperature and Growing Season Temperature indicators. For threshold-defined indicators, namely Spring Heat Days and Summer Heat Stress Days, the end-users are highly encouraged to use climatology in the first month and combine it with observations as soon as the latter becomes available. The paper also assessed the differences between indicators, strategies, and start months. The remaining bias of each indicator could be found in the supplementary material.

**11. Performances of climatic indicators from seasonal forecasts for ecosystem management: The case of Central Europe and the Mediterranean**

* The paper evaluates the performance of seasonal forecasts in predicting key indicators for agriculture and forestry sectors in Central Europe and the Mediterranean region. The study suggests the potential usefulness of seasonal forecasts for decision making under different geographical and environmental contexts, considering sensitivity inherent to different processing chains.
* The paper uses authoritative climate data from Copernicus Climate Data Store, i.e., ERA5 reanalyses and hindcasts from CMCC SPSv3 and ECMWF SEAS5 seasonal prediction systems (SPSs), to evaluate the performance of seasonal forecasts for the summer period across Central Europe and the Mediterranean region in predicting some key indicators serving agriculture and forestry sectors, i.e., Potential EvapoTranspiration (PET), Potential Soil Moisture Deficit (PSMD), and Fire Weather Index (FWI). The study assesses the differences in performance across the two SPSs, two start dates (March and May) and four correction techniques applied to overcome modelling bias, namely bias correction (BC), calibration (CAL), quantile mapping (QM) and detrending (DET).
* The results of the paper are organized to represent -for the different SPSs and forecast start dates -the absolute bias, against the observations, in the indicators for the whole climatological summer season, the year-to-year correlation and the ability to detect out-of-norm and extreme events. The study found that seasonal predictions of PET perform better in Western and Eastern Europe and some areas of North Africa. PSMD predictions follow a similar spatial pattern as PET, except that for some areas in Central (Eastern) Europe in which the performance increases (decreases). FWI predictions reveal better results in some areas of the Iberian Peninsula, North-Western Africa, Balkan Peninsula, and Ukraine. Results also suggest that QM might be the most suitable technique for bias correction. Furthermore, the start date of the forecast might imply varying correlation significance, with the start date closest to the forecasted period not always being the best. Overall, the study suggests the potential usefulness of seasonal forecasts for decision making under different geographical and environmental contexts, considering sensitivity inherent to different processing chains.

**12. Projected climate changes are expected to decrease the suitability and production of olive varieties in southern Spain**

* The paper examines the effect of climate change on the area suitable for olive cultivation and production levels in Andalusia, one of the main areas of olive production. The authors used Species Distribution Models (SDMs) to predict the current and potential distribution of olive varieties and evaluated the expected alteration in annual olive production.
* The authors used Species Distribution Models (SDMs) to predict the current and potential distribution of olive varieties and evaluate the expected alteration in annual olive production. They used the 'biomod2' R package to ensemble the models and used expert knowledge, previous literature, and exploratory analysis to select predictors. They also derived future climate projections to assess the effect of climate change on the environmental suitability and productivity of each olive variety.
* The paper found that soil pH was the most important factor for most distribution models, while (bio-)climatic predictors such as continentality index, summer and autumn precipitation, and winter temperature provided important contributions. Projections based on regional climate change scenarios point to a decrease in the area suitable for olive crops in Andalusia, due to an increase in evapotranspiration and a decrease in precipitation. These changes in suitable area are also projected to decrease olive production for almost all the olive varieties.

**13. Determination of the quality and purity characteristics of olive oils obtained from different regions of Turkey, depending on climatic changes**

* The paper aims to determine the quality and purity characteristics of olive oils obtained from different regions of Turkey, depending on climatic changes. The study found that the delta-7-stigmastenol value is high when the annual average relative humidity is low and the annual average temperature is high.
* The methods used in this paper are as follows:
  + Five different olive varieties were harvested from four different regions of Turkey during the 2017/2018-2020/2021 harvest years.
  + Olive samples were collected from three orchards in 13 provinces/districts every year.
  + One hundred and fifty-six samples were subjected to purity, quality, and sensory analysis.
  + Basic climatic values (average, minimum, and maximum temperature, humidity, and precipitation) were examined for four consecutive years.
  + Multiple linear regression analysis was applied using a genetic algorithm-based inverse least squares method to determine whether there is a relationship between climate data and delta-7-stigmastenol values.
  + The FFA (in oleic acid %) and UV-spectrophotometric indices (K 232 and K 270 measurements) were measured according to the methods given by the International Olive Council (IOC).
* The paper found that all of the examined olive oil samples were determined within the legal limits in terms of fatty acid composition and fatty acid ethyl ester values. However, delta-7-stigmastenol value from the sterol composition was found to be above 0.5% in some samples in all the years studied (total 21 samples). Delta-7-stigmastenol values of olive oil samples varied between 0.16% and 1.14%. Multiple linear regression analysis was applied using a genetic algorithm-based inverse least squares method to determine whether there is a relationship between climate data and delta-7-stigmastenol values. According to this result, it has been determined that the delta-7-stigmastenol value is high when the annual average relative humidity is low and the annual average temperature is high.

**14. A new aerobiological indicator to optimize the prediction of the olive crop yield in intensive farming areas of southern Spain**

* The paper proposes a new forecasting model for predicting the olive crop yield in southern Spain using aerobiological and meteorological parameters. The model includes a new predictive variable, the number of days with pollen concentrations ≥400 pollen grains m −3, and weather-related variables such as cumulative precipitation and mean maximum temperature.
* The paper used partial least-squares regression to construct forecasting models for predicting the olive crop yield in southern Spain. The models were based on linear transformation of the original descriptors to a small number of orthogonal factors (latent variables), to maximize the covariance between the descriptors and the dependent variable. The annual olive yield was taken as the dependent variable, and several aerobiological and meteorological parameters were used as independent variables. The models were validated using a full cross-validation method, and a 23-year period was used for the study. Four regression models were built, two of which included the new aerobiological variables proposed in the study, and two of which were based on biometeorological forecasting models developed previously in the study area.
* The paper found that the number of days with pollen concentrations ≥400 pollen grains m −3 was a new predictive variable that accurately predicted the olive harvest in the study area, and was included in the forecasting model with the highest determination coefficient value (R 2 = 0.89). Weather-related variables such as the cumulative precipitation from October to December of the previous year or the mean maximum temperature from January to March were also factors of particular importance on crop production. The new model proposed provides early and effective olive crop forecasting by using independent variables which can be easily obtained towards the middle of June, also incorporating to the model the phenological variability associated with changes in the local weather.

**15. A Machine Learning Model for Early Prediction of Crop Yield, Nested in a Web Application in the Cloud A Case Study in an Olive Grove in Southern Spain**

* The paper presents a machine learning model for early prediction of crop yield in an olive orchard in southern Spain. The model uses spatio-temporal training data and meteorological parameters data to provide advance information on the profitability of the farm.
* The paper uses data mining techniques to generate predictive models of the amount of olive crop that will be harvested. The model is estimated using spatio-temporal training data, such as yield data from eight consecutive years, and more than twenty meteorological parameters data, automatically charged from public web services, belonging to a weather station located near the sample farm. The workflow requires selecting the parameters that influence the crop prediction and discarding those that introduce noise into the model.
* The paper reports that the developed machine learning model for predicting crop yields in advance has absolute errors better than 20%, which is crucial for making decisions on tillage investments and crop marketing. The model is accessible and easy to use by the farmer or farm manager from a web-based application.

**16. Evaluation of Three Feature Dimension Reduction Techniques for Machine Learning Based Crop Yield Prediction Models**

* The paper proposes a framework that uses feature selection (FS), feature extraction (FX), and a combination of both (FSX) to reduce the dimensions of features when developing machine learning-based models for crop yield prediction. The results show that FSX-based models perform the best, improving the accuracy of predicting crop yield by up to 60% in terms of RMSE.
* The paper uses machine learning-based rice yield crop models over the entirety of Vietnam based on vegetation condition index (VCI) and temperature condition index (TCI) data as a case study. It employs all common, widely used machine learning algorithms, namely linear, support vector machine (SVM), decision tree (Tree), artificial neural network (ANN), and Ensemble. The paper proposes a framework that uses feature selection (FS), feature extraction (FX), and a combination of both (FSX) to reduce the dimensions of features when developing machine learning-based models for crop yield prediction. The PCA method is used for feature extraction. The paper also uses an approach of using the same learner algorithm but training them on different subsets of the training set.
* The results of the paper show that the FSX-based models perform the best, improving the accuracy of predicting crop yield by up to 60% in terms of RMSE. The paper also highlights the significant role of FS, FX, and specially FSX coupled with a wide range of ML algorithms (especially Ensemble) for enhancing the accuracy of predicting crop yield. The study uses leave-one-out cross-validation for training all models. 21 of the best models are developed based on Ensemble (13 models), Tree (6 models), linear (1 model), and ANN (1 model).

**17. Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes**

* The paper proposes a Crop Yield Prediction Algorithm (CYPA) utilizing IoT techniques in precision agriculture to predict crop yields based on soil, climatic, environmental, and crop traits extracted via decision support algorithms. The algorithm incorporates climate, weather, agricultural yield, and chemical data to facilitate the anticipation of annual crop yields by policymakers and farmers in their country.
* The paper uses machine learning approaches to predict the top 10 crop yields eaten globally, including wheat, paddy, rice, sorghum, sweet potatoes, soy beans, plantains, cassava, and others. The study trains and verifies five models using optimal hyper-parameter settings for each machine learning technique. The DecisionTreeRegressor achieved a score of 0.9814, RandomForestRegressor scored 0.9903, and ExtraTreeRegressor scored 0.9933. Additionally, the paper introduces a new algorithm based on active learning, which can enhance CYPA's performance by reducing the number of labeled data needed for training.
* The paper presents the results of the Crop Yield Prediction Algorithm (CYPA) utilizing IoT techniques in precision agriculture to predict crop yields based on soil, climatic, environmental, and crop traits extracted via decision support algorithms. The proposed CYPA incorporates climate, weather, agricultural yield, and chemical data to facilitate the anticipation of annual crop yields by policymakers and farmers in their country. The study trains and verifies five models using optimal hyper-parameter settings for each machine learning technique. The DecisionTreeRegressor achieved a score of 0.9814, RandomForestRegressor scored 0.9903, and ExtraTreeRegressor scored 0.9933. Additionally, the paper introduces a new algorithm based on active learning, which can enhance CYPA's performance by reducing the number of labeled data needed for training. Incorporating active learning into CYPA can improve the efficiency and accuracy of crop yield prediction, thereby enhancing decision-making at international, regional, and local levels.

**18. Comparative Study for Classification Algorithms Performance in Crop Yields Prediction Systems**

* The paper focuses on using data mining classification algorithms to predict the impact of various parameters such as area, season, and production on crop yield quality. The performance of different algorithms is compared based on error values and accuracy.
* The paper uses data mining classification algorithms such as Naive Bayes, Decision Tree, SVM, KNN, and Random Forest to predict the impact of various parameters on crop yield quality. The performance of these algorithms is compared based on error values and accuracy. The datasets used in the experiment were collected from Kaggle publicly available datasets, which were pre-processed and organized in Microsoft Excel. CSV was the file format used.
* The paper compares the performance of different data mining classification algorithms in predicting the impact of various parameters on crop yield quality. The results are presented in tables and figures, which show the accuracy percentage, error values, and time taken to build the model for each algorithm. The SVM algorithm achieved the highest accuracy value with 76.82%, while the lowest is achieved by the KNN algorithm with 35.76%. The highest error value was 111.8855 for KNN. The paper concludes that data mining techniques can be used to predict crop yield quality, and SVM is the most accurate algorithm for this purpose.

**20. Crops yield prediction based on machine learning models: Case of West**

**African countries**

* The paper proposes a machine learning-based prediction system to forecast the yield of six crops (rice, maize, cassava, seed cotton, yams, and bananas) in West African countries based on climatic data, weather data, agricultural yields, and chemical data. The system aims to assist farmers and decision-makers in making quick decisions to combat climate change and ensure food security.
* The paper uses a combination of climatic data, weather data, agricultural yields, and chemical data to predict the yield of six crops in West African countries. The authors used three machine learning models, namely decision tree, multivariate logistic regression, and k-nearest neighbor, to build the prediction system. They applied hyper-parameter tuning techniques throughout cross-validation to get a better model that does not face overfitting. The authors also studied the correlation between the predicted results and the expected results.
* The paper achieved promising results with all three machine learning models. The decision tree model performed the best with a coefficient of determination (𝑅2) of 95.3%, while the K-Nearest Neighbor model and logistic regression performed with 𝑅2 = 93.15% and 𝑅2 = 89.78%, respectively. The prediction results of the decision tree model and the K-Nearest Neighbor model were found to be correlated with the expected data, which proves the efficacy of the model.

**21. A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches, With Special Emphasis on Palm Oil Yield Prediction**

* The paper provides a comprehensive review of the use of machine learning algorithms for crop yield prediction, with a special emphasis on palm oil yield prediction. It discusses the current status of palm oil yield, critical evaluation of state-of-the-art crop yield prediction approaches, and identifies the benefits and challenges associated with features and machine learning algorithms in the prediction of crop yield.
* The paper is a review article that critically evaluates the use of machine learning algorithms for crop yield prediction, with a special emphasis on palm oil yield prediction. The authors have reviewed a large number of articles related to crop yield prediction and proposed a forthcoming trend of palm oil yield prediction framework based on the critical assessment of related studies. They have also presented a prospective framework of palm oil yield prediction, which involves collecting a wide range of data including leave and fruit information, irrigation information, soil properties, climatic information, vegetation indices, cropland information, crop management data, historical yield data, fertilization information, and satellite data in the first step.
* The paper is a comprehensive review article that critically evaluates the use of machine learning algorithms for crop yield prediction, with a special emphasis on palm oil yield prediction. It does not present any specific results of an experiment or study. Instead, it provides a detailed analysis of the current status of palm oil yield, the fundamentals of crop yield prediction, critical evaluation of state-of-the-art crop yield prediction approaches, and identifies the benefits and challenges associated with features and machine learning algorithms in the prediction of crop yield. The paper also proposes a forthcoming trend of palm oil yield prediction framework based on the critical assessment of related studies.

**22. Agricultural decision system based on advanced machine learning models**

**for yield prediction: Case of East African countries**

* The paper proposes the development of a decision system using machine learning models to predict crop yield in East African countries. The system combines climate data, crop production data, and pesticides data to develop a decision system based on advanced machine learning models.
* The paper proposes three crop prediction models: Crop Random Forest, Crop Gradient Boosting Machine, and Crop Support Vector Machine. These models are based on an ensemble of decision trees and are used for both classification and regression of crop yield problems. The models take as input the number of features such as climate data, crop production data, and pesticides data to develop a decision system based on advanced machine learning models. The paper also discusses the data acquisition process, data prepossessing, and the data analysis results. Finally, the paper presents experimental results and discussion, followed by future directions.
* The paper presents experimental results that show the three proposed machine learning models fit well with the crop data with a high accuracy R2. The Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) associated with the models are very minimal because the agricultural prediction values are very close to reality. The proposed models are reliable and generalize well the agricultural predictions in East Africa. The experimentation was performed using a Linux computing system, consisting of Intel Core i5 CPU 16.0 and 8.0 GB of RAM. Anaconda environment was used to implement the models using Python 3.8.8. The training dataset was chosen as 70% of all datasets and was used to train the three proposed machine learning models.

**23. An intelligent decision support system for crop yield prediction using hybrid machine learning algorithms [version1; peer review: 2 approved, 1 approved with reservations]**

* The paper proposes a hybrid machine learning algorithm to predict crop yield with improved accuracy. The algorithm uses specialized ensembling methods such as stacked generalization, gradient boosting, random forest, and LASSO regression.
* The paper proposes hybrid machine learning algorithms that use specialized ensembling methods such as stacked generalization, gradient boosting, random forest, and least absolute shrinkage and selection operator (LASSO) regression. The proposed algorithm uses aerial-intel datasets from the github data science repository to demonstrate its applications. The performance of the individual algorithm and hybrid ML algorithms are compared using cross-validation to identify the most promising performers for the agricultural dataset.
* The experimental results done on the agricultural data show that the accuracy of random forest regressor, gradient boosted tree regression, and stacked generalization ensemble methods are 87.71%, 86.98%, and 88.89% respectively. The proposed stacked generalization ML algorithm statistically outperforms with an accuracy of 88.89% and hence demonstrates that the proposed approach is an effective algorithm for predicting crop yield.

**24. Deep learning for crop yield prediction: a systematic literature review**

* The paper is a systematic literature review that provides an overview of the state-of-the-art application of Deep Learning in crop yield prediction. The authors retrieved 456 relevant studies and selected 44 primary studies for further analysis. They observed that Convolutional Neural Network (CNN) is the most common algorithm and has the best performance in terms of Root Mean Square Error (RMSE). One of the most important challenges is the lack of a large training dataset and thus, the risk of overfitting and lower model performance in practice. The authors suggest that data augmentation and transfer learning can overcome the limitation of large training datasets. The paper concludes that crop yield prediction with Deep Learning methods depends on the kind of data used, as much as on the type of crop.