International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Developing An Integrated AI Model for Predictive Maintenance in Hydroponics

Sanay Dhote

Student, The Riverside School

Abstract:
Hydroponics have gained attention due to their ability to cope with future food needs. However, the farmers face the issue of maintenance: if done frequently, it is expensive, and if done rarely, it risks crop failure. This paper aims to ideate a setup of AI models for maintenance to be predictive, reducing crop failure risk while making maintenance more profitable. The model must be able to predict future readings of various conditions that can indicate a crop failure. For this reason, I found the various measurements that can indicate crop failure, and which sensors can detect them.

The factors were yield prediction, component malfunction, nutrition, electrical conductivity, power of hydrogen, and environmental factors (temperature, humidity, and light intensity). Next, I collected various studies done on how different models predict each of these factors. The results showed that for yield prediction, deep neural network; for component malfunction, random forest; nutrition, random forest and support vector regression; for electrical conductivity and power of hydrogen, nonlinear autoregressive with exogeneous inputs; and for environmental factors, extreme gradient boosting were the promising AI models. Also, deep neutral networks can be trained to mimic other model's decisions, which can lower economic investment of the various AI models.

INTRODUCTION:

The UN predicts that the world population will rise to 10 billion by 2050 (2025). To meet their needs, the production of food would also need to increase by 70% (United Nations, 2024). To achieve this ambitious goal, there is an urgent need to find an efficient, fast, and reliable food source. Over the years, environmentally harmful activities such as deforestation and pollution have caused the harvests to be less efficient year by year (Zongbo Shi, 2014). Pollution and climate change is a well known culprit to unpredictable, extreme weather conditions.

This scenario puts the agricultural sector at a tough spot with an urgent need to balance between their risk to reward (having a high initial investment that may not lead to profits) and providing a necessity to human society: food. There have been various innovations to solve this major issue, one of them is Hydroponics. Hydroponics, like other modern farming methods, is designed to be efficient, fast and easy to manage (Zainab Kakal, 2022).

Hydroponics is known to produce faster and safer fruits and vegetables (Zainab Kakal, 2022). Moreover, hydroponics is very scalable. For instance, they can be the size of a pencil holder or a green-house which allows hydroponics to be implemented in crowded cities, where growing output would be impossible. Hydroponics uses 90% less water and can deliver around 10-12 times more yield (Barbosa et al., 2015). The plants as well can be grown all year round, no matter the weather or season. These factors make hydroponics more efficient and continuous. Another advantage of hydroponics is the lack of pests and pesticides in their growth. Unlike traditional farming, plants are kept in a controlled environment limiting its interactions with the outside world which dramatically reduces pests as well as the pesticides needed, resulting in an organic and healthy harvest.

Despite its advantages, the current state of hydroponics faces the issue of maintenance. Maintaining hydroponic farms cost a lot, thus it is recommended not to be done often (Gomes, 2023) . Also, if not done properly, it can ruin the entire crop batch,generating a significant loss for the farmer. Hydroponics are mostly done with the machines to control various elements: temperature, power of hydrogen (pH), humidity, light intensity, and more. To the inexperienced, it will be almost impossible for them to distinguish between faulty mechanisms just by the readings, in fact it might not even be noticed. Another key factor is the substantial amount of energy hydroponic farms need. They are completely dependent on machinery, this means a single power outage can ruin a batch of harvest. The usage of electricity is also a big cost (Gomes, 2023).

The Internet of Things (IoT), the interconnectedness of devices via the internet allowing to send or receive data, has also helped hydroponics a lot. It makes managing hydroponics very handless, and accessible just through a device. It reduces the dependency of the farm to a person and also lessens the chances of human error causing major damage. Furthermore, if the farmer wants to take a break, everything can be managed just on a phone (Omolola Ogboluman, 2023).

While integrating IoT is known to help hydroponics, it can also increase the risk of failure by software/technological bugs. A single bug can ruin the flow of the hydroponic components.

This is where artificial intelligence (AI) can play an important role in overcoming these challenges. Using the AI algorithms, a farmer can anticipate failure in advance and solve the issue as soon as possible. Making hydroponics more economically profitable.

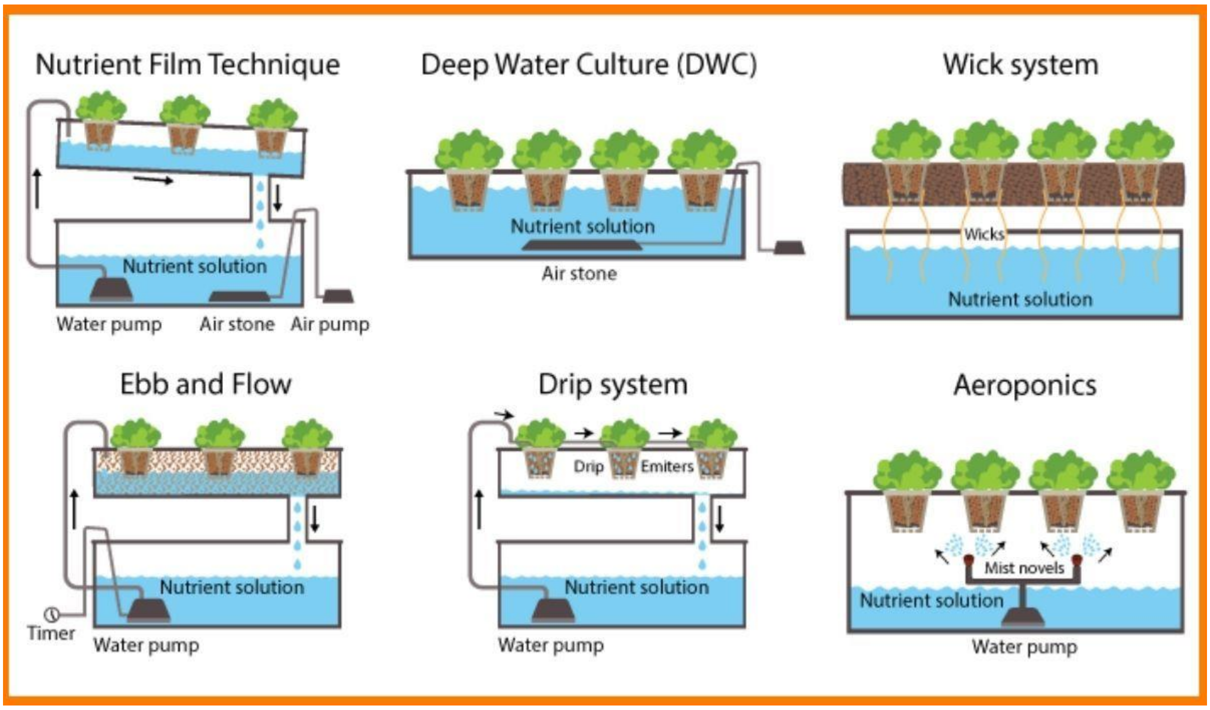
By relying on AI and IoT, a hydroponic farm can receive a lot of advantages.The AI algorithm can be tuned for a specific hardware. The hardware information can be fed into the AI and then make decisions based on the reviews of the hardware. For example, AI can be trained to detect anomalies in hardware performance, such as irregularities in nutrient delivery systems, based on historical and real-time data.

For the AI to work, it must receive numerical information to interpret in the first place. The question remains, whether there lies a number that will help predict a failure that can then be used to train AI. To train AI, we must give it various input sets of data, and label them, the result of the data (crop failure or not). This way the AI will fine tune by going through the training data sets. However there lies the issue of overfitting in AI, where too much training will lead to the AI pattern match. For this reason, there will be a test data set to give predictions. Then using the predictions and actual results, we can compare the accuracy and then set up the AI for a set amount of training repetitions (Boorla Arjun, 2025).

Predictive maintenance, a strategy of using real time data to predict any potential failure before they happen, is also a crucial part of managing hydroponic systems. It helps achieve efficiency in gaining output from hydroponics, while not bearing minimum costs in damage/maintenance of the system. Predictive maintenance allows the farmer to decide when to do checkups or maintenance on a system. While frequent checks might reduce output, no checks can cause crop failure. In sum, predictive maintenance helps a farmer to predict when a failure will occur in the system before it starts affecting the crops resulting in minimizing downtime and reducing the risk of crop failure.

This research paper will explore and evaluate various ideas from other research papers that integrate hydroponics with AI, and see the distinct advantages and shortcomings that each of the AI models tend to have. The end goal is to combine the knowledge and pattern seen in others' systems and ideate a solution that can be implemented as one fit for all models. At the same time ensuring any mechanical or technical failure can be predicted using AI with minimal false or missed alarms. In the future this model can then be implemented in hydroponic setups and tested for its accuracy and precision.

Different Hydroponic Setups



Methodology:

To reach an idealistic setup that solves predictive maintenance with the help of AI, smaller factors that can indicate maintenance would need to be found. AI needs numerical data on various factors in hydroponic setups to make calculations. Sensors will read and give out these numericals. It is then the AI's role to interpret, analyse, and predict trends that indicate maintenance. The best model for each factor should have the highest accuracy in its prediction for the assigned factor.

There are various sensors in a hydroponic setup: electrical conductivity (EC), power of hydrogen (pH), humidity, temperature, nutrition sensors, and many more. It is possible that for each of these factors, there lies a certain trend that can predict the possibility of a failure. To create a robust framework, the solution should incorporate multiple AI models for different factors, allowing farmers to cross-check with other parameters to avoid false or missed alarms.

As the name suggests, predictive maintenance needs to be able to predict a failure, not to inform after it has occurred. This means the model should accurately be able to predict the future readings of the setup based on the reading received. For this reason, using google scholar, I found 3 studies for each of the factors selected to predict by using keywords in the searches. The factors selected were environmental factors (light, humidity, and temperature), EC and pH, nutrition, yield prediction, component malfunction, and other factors (financial and scalability).

Each of the papers targeted predicting the value of the factor it has to. I found out the few best models from each study, its drawbacks, and tried to explain why a certain model worked best in that case. This helps to understand the scenario in which a certain model works best at. Using this knowledge, I can then create a solution that solves the issue of predictive maintenance by incorporating a model in each factor, keeping economical feasibility and scalability in mind. Each selection will not be made because it proved best in its study, but because its system of working prefers the nature of the specific data. This method allows that no bias is given to a certain model, and the solution has substantiate reasoning behind it.

Literature Review:

Predictive maintenance in hydroponics is a growing area of research. The failures in these systems usually don't happen instantly, but start with slight deviation in the environment. For example, a slight change in pH, EC or nutrient balance can eventually grow into a bigger issue that affects the entire batch of crops. Several researches suggest that IoT-enabled systems can provide a strong base for predictive maintenance. Rahman et al. created an AI and IoT integrated hydroponic system that would collect data from different sensors, and then use machine learning models to suggest recommendations and detect anomalies(2024). Another researcher developed a deep-water culture (DWC) system that uses ESP32 and the Blynk platform to automate pumps, while also giving alerts in case values start to deviate (Duangpakdee et al., 2024). These IoT integrations show that real-time monitoring reduces the chances of crop failure, and also creates a strong base for AI algorithms to be used in prediction.

Other than monitoring, machine learning can also be used for predicting yields. Mokhtar et al. tested various algorithms, such as SVR, Random Forest, and deep neural networks on lettuce grown with the help of hydroponics. They found that the models accurately predicted the fresh weight of the crop (2022). This finding can be very important as an indication of the actual yield being lower than the predicted yield, and can signal that something is wrong with the hydroponic system. This can help to be used as an early-warning mechanism for crop failure.

Nutrient management is also a critical factor where AI can help reduce crop failure. "A study on soybeans grown with hydroponics used machine learning to evaluate the relationship between nutrient concentrations—such as nitrogen, magnesium, and potassium—and plant uptake and growth (Dhal et al., 2024). The study showed how predictive models can notice deficiency or toxicity in nutrients before the plants show stress. A study also addressed improving the accuracy of ion-selective sensors, which are typically influenced by interference from other ions (Zhang et al., 2023)This helps predictive maintenance because it avoids faulty nutrient readings, one of the most common and unnoticed causes of crop failure in hydroponics.

Failures are not just from nutrient or environmental changes, but can also occur from biological causes such as plant diseases. A study by Musa et al. proposed a lightweight convolutional neural network that can detect diseases in hydroponic plants on low-power devices (2023). Pathogens can spread very fast in a nutrient solution that is shared across plants. For that reason, early detection using computer vision is an important form of predictive maintenance.

AI can also be used in controlling the hydroponic system in real-time. To optimize greenhouse environments, a deep learning system was developed to continuously balance temperature, humidity, and CO_2 (Viet et al., 2024). The system could predict when the values are about to deviate outside the safe range and automatically adjust the mechanism, and prevent crop stress before it can become irreversible.

Rathor et al. explained how IoT and AI solutions can be applied to vertical farming and hydroponics, with special focus on challenges such as sensor deviation, mechanism failures and pathogen detection (2024). Similarly, a compiled review of hydroponic systems highlighted the parameters that have the most influence on crop success, such as pH, EC and dissolved oxygen (Shareef et al, 2024). These reviews

provided clarity on which signals and failure points should researchers prioritize when building predictive maintenance systems.
To sum it up, AI helps predictive maintenance in three main ways. It can help predict crop yields and nutrients conditions before any problems show up. It can also detect anomalies in real-time. Finally, it can control systems in a manner to keep the environment stable and within the safe range. If IoT provides the constant flow of data, then AI algorithms can be trained to warn about these issues in their early stages, giving the user a better chance at avoiding complete crop failure. However, a noticeable gap in the reviewed literature is the lack of a complete and holistic approach. Most studies focus on individual aspects, like anomaly detection, yield prediction, or system control, and very few suggest an integrated framework that combines all these elements into a complete predictive maintenance solution.

Comparative Analysis:
This comparative analysis aims to explore and identify which AI model approach is the most effective at training AI systems to detect early signs of equipment malfunction in hydroponic environments. While prior research explores a range of objectives such as yield prediction and nutrient optimization, the ability of hydroponics to anticipate pump failures, sensor drift, or emitter blockages is critical to ensure its reliability and economic viability.By comparing anomaly detection frameworks such as Isolation Forests, autoencoders, and hybrid statistical, this study evaluates not only accuracy in fault detection but also robustness under noisy sensor data, scalability across system sizes, and suitability for real-time deployment.
With the final goal in mind, the analysis will include 3 research papers focusing on each of the various aspects that influence the hydroponics systems predictive maintenance.

Yield Prediction
Yield prediction can act as a very effective approach to decide the appropriate frequency of maintenance for hydroponic setups. Decision making becomes more numerical and planning based rather than instinctive. For instance, if the yield of corn in a farm is constantly reducing, the farmer may want to inspect the hydroponic system to identify if there is a fault and stop the problem before it amplifies. If implemented correctly and in some specific conditions, the model may predict yield reduction in advance, prompting the farmer to take action before any loss is generated.
The study by Duangpakdee et al. (2024) gives insights into this aspect of predictive maintenance with chinese celery being the plant in subject. The research focused on the effect of temperature and humidity on Chinese celery yield. The study involved 4 hydroponic setups, the first setup had controlled light intensity, second had controlled temperature, third had both, and the last was the controlled setup with natural conditions. Other sensors monitored parameters such as humidity, electrical conductivity, power of Hydrogen, and more. A ESP-32 microcontroller would monitor the environment and the actuators. The system was recalibrated every month to ensure reliability.The study showed a 14% economic profitability with a breakeven point of 27 months. However, the study also identified periodic calibration as a practical application issue. Also, it highlighted the importance of sensor accuracy for ensuring the yield consistency and long-term sustainability. Predictive maintenance can make this process way easier. Instead of calibration done at a frequent time period, it will be more calculated, which will maximize the yield.. This instance shows the scope of benefits a correctly implemented and planned predictive maintenance model can bring to the farmer.
The research by Mokhtar et al. (2022) looked into 4 different AI models, namely support vector regression

(SVR), extreme gradient boosting (XGB), random forest (RF), and deep neural network (DNN), to predict lettuce yield in a hydroponic setup. The models were tested under 3 different scenarios which showed merits of different models.For example the 3rd scenario showed lowest root mean square error (RMSE) for XGB, while the 1st scenario had the lowest RMSE for RF. This shows how there is no single answer to the problem, but it is dependent on the specific factors. However, it is important to note that DNN had promising predictions in all the 3 scenarios. Further on, its ability to take fewer input variables and deliver these results make the model easy to understand and feasible to scale-up for rapid decision-making.
The study by Kadam et al. (2025) looks into even more models, six to be precise. They include linear regression, ridge regression, lasso regression, support vector regression, random forest, and decision tree. The study revolves around predicting the growth of thai basil. The growth prediction can help determine if the plants are receiving enough nutrients and whether the predicted forecast matches reality. The more the deviation between the two, more the chances of a malfunction in the system. The conclusion of the research showed that the decision tree model performed the best at predicting growth rates. So if we had to compare the real growth to a forecast, it would be from the forecast of this model.

Component Malfunction

Component malfunction is one of the prevalent causes of crop failure, and one of the main reasons predictive maintenance needs to be done in systems. Component malfunction happens due to wear and tear in the sensors, actuators, or other critical features in the system. Unlike other issues, component malfunctions are relatively hard to detect. This is because these are the pieces that are the data inputs for detecting problems. Deviations in readings can either be from a bug in the system, or an actuator not deploying, or even the sensor itself taking wrong readings. As humans, we can’t discern the root cause in such cases. This is where AI models will help flag out if a system needs to stop complete operation due to some critical failure, or if it was just a small issue that can be resolved in seconds. This minimizes downtime of a setup and reduces the gambling nature of such faulty readings.
An example of this situation is presented by Rahman et al. (2024). Its primary aim was to provide crop recommendations to help achieve optimal parameters to make the yield be as efficient as possible. The methodology first collected various data (nutrition, pH, temperature, humidity) via IoT sensors, then preprocessed the information for consistency, and finally trained several machine learning models. The trained models were decision tree, random forests, k-nearest neighbours (KNN), support vector regression, and extreme gradient boosting. After training, these models were evaluated using accuracy, precision (the amount of true positives over all the positives given by the machine), recall (the amount of true positives over all the positives in the data set), and F1-score (which follows the formula $2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$). Random Forest achieved the best results (97.5% accuracy, ≈0.98 precision/recall). The multi-model training steps helped ensure that the most reliable model was found, and not assuming a single algorithm. The authors also did note the dependence of the models on the sensor quality and suggested increasing datasets and going deeper into anomaly detection methods.
The Random Forest model has many advantages for predictive maintenance in hydroponic systems. Its structure, which aggregates multiple decision trees, makes it highly reliable at capturing non-linear relationships in sensor data such as pH, nutrient balance, or pump flow rates. For the same, it helps reduce overfitting (when the model starts to memorize patterns) in AI models. This allows the model to detect small deviations that can signal early equipment malfunction. Moreover, its high precision and recall scores showcases its ability to minimize both false alarms and missed detections, making it a reliable

monitoring tool. In practice, predictive maintenance with Random Forest can be implemented by continuously feeding real-time sensor streams into the trained model, which classifies whether conditions fall within normal ranges or suggest an impending fault. Alerts generated from the model can then prompt farmers for interventions, such as recalibrating sensors or servicing pumps, before the issue leads to crop failure. However, the model also has limitations. It needs large, high-quality datasets to maintain accuracy, it can be computationally expensive when scaled for real-time monitoring, and its "black-box" nature makes it harder for farmers to interpret why a particular fault was alerted. Additionally, performance may degrade if sensors drift over time without proper recalibration, highlighting the need for more anomaly-detection methods. This makes Random Forest a very expensive model to implement on a large scale.
The study by Métwalli et al., (2025) takes into consideration anomaly detection in hydroponic setups, as well as looking into growth prediction at the same time. It was done by placing a camera and then various imaging processing techniques to feed the model, and then the model will detect anomalies. Random forest, linear regression, and neural network were tested. Random forest performed the best at predicting anomalies, with an accuracy of 94.55%.

Nutrition

Nutrition is a critical factor of plant growth in hydroponics. Unlike traditional farming, where nutrients are absorbed from the ground, hydroponics take nutrition directly from the water provided to them. Many times crop failure and predictive maintenance issues occur not only due to component malfunction, but because of the excess or lack of nutrients for a plant to grow. Nutrient levels continuously change as plants uptake nutrients and actuators input more nutrients. In fact, nutrient level drifts can also be a proof of component malfunction. For instance, a faulty timer that controls nutrient addition intervals could lead to such deviations.
A study by Dhal et. al. (2024) used machine learning to analyse how different compositions of nitrogen, magnesium, and potassium affect the water uptake of plants in hydroponic setups. It used various models and measured the most accurate model for each nutrient. To determine the accuracy, the study used RMSE on model's data sets. The results showed RF performed better for Nitrogen and Magnesium, while SVR performed better with potassium. This study helps gain insight into how crop failure can be reduced by models analysing water uptake, as nutrition imbalance is a major reason for crop failure. The model can directly catch minor rise or drop in nutrition and prompt farmers to check the hydroponic system for any malfunction in components. This will reduce frequency of inspection and reduce the loss due to downtime, without damaging crops. Nitrogen's stronger performance with Random Forest, compared to magnesium and potassium with Support Vector Regression, can be attributed to differences in their uptake patterns. Nitrogen happens to have a non-linear uptake, but trends in small pieces and chunks. Random Forest excels in discontinuous data and thresholds, making it better suited for nitrogen. On the other hand, potassium and magnesium follow a relatively linear uptake by water, creating a more linear relationship, which is the domain support vectorisation regression excels at.
This study uses a very robust methodology, and can be used in different hydroponic setups by just modifying the data with any other variable. The researchers used the SHAP analysis to explain the predictions of models, giving a clear picture of a certain output. This approach solves the problem of the previous study, where the root cause may be hard to interpret. For example, if the model thinks a pump will fail, the farmer can correlate it with rising motor temperature or failing water flow, which could not have been done with the previous study.

However, there are also a few limitations. The methodology relies on using interpolated data, potentially creating bias if there is a pattern noticed. It is further limited by the fact that the study was done in a completely controlled setup, which means the fewer external variables made it easy for AI to predict accurately. If applied in large scale models or commercial systems it may negatively impact AI's ability to compute accurately due to various uncontrolled factors. On the bottom line, the study's methodology and it used to analyse models make it a very useful piece of work, which can be implemented in any setup for various data sets.

The study done by Arjun et al. (2025) highlights other benefits and possible intervention in the automated AI system. The study focused on mixing nitrogen, potassium, and phosphate (NPK) salts using AI. The results showed a 25% reduction in nutrient wastage and also 18% increase in yield. The study used convolutional neural networks to gather and process data, and then used decision tree models to put actuators into play. After all this, the user can input feedback, which can enable the AI to learn from any mistake with the help of reinforcement learning. This allowed a continuous improvement and the operation to be more responsive.

Electrical Conductivity and power of Hydrogen

Similar to nutrition, electrical conductivity (EC) and power of hydrogen (pH) play a pivotal role in growth of plants.. Their fluctuation can also provide valuable insights into when predictive maintenance should be done, while also showing proof of component malfunction.

The study by Karimzadeh et al. (2025) evaluated 5 different models to detect and diagnose 5 different factors related to electrical conductivity and power of hydrogen. The models were random forest, artificial neural network, long short-term memory, k-nearest neighbours, and support vector regressor. The factors evaluated were bias, drift, precision degradation, spike, and stuck. In the results, random forest performed the best for fault detection in EC, with 100% in spike and 97.1% in bias. However, for pH there is no single model that detected and diagnosed best, but it varied based on the fault type. SVR excelled with 100% in drift, while random forest performed better in normal drift and precision degradation. Notably, artificial neural networks performed well to predict EC levels with variables and environmental factors and a fault detection accuracy of 93.2%.

The study by SIDIBE et al. (2023) showcases a model that can predict EC and pH levels 5 minutes prior. This prior prediction allows pumps and actuators to take appropriate actions to avoid a disaster before it can actually strike. The model was nonlinear autoregressive with exogenous inputs.The model used was computationally light, with only 2 hidden neurons, and was very efficient with micro-controllers. The results showed 0.99 regression fit, making the model a very strong fit. The model had a RMSE of 0.00268, showing that signs of underfitting and overfitting were not prevalent. For EC, the predictions were very accurate, with the predicted trend matching the real trend. For pH, however, there were small deviations of 0.1 pH, but still a notable achievement for the model. Since the model is highly compatible with micro-controllers, the cost of installing it is relatively low. There were a few limitations noted. Firstly, the reading was only for 15 days, a relatively short period of time. Secondly, uneven mixing of nutrients at times caused the data to be skewed.

Environmental Factors

In this context, environmental factors specifically refer to humidity and temperature. At a glance, these readings may not directly tell much about the hydroponic system directly as factors like EC, pH, solubility of nutritions, and many more factors are dependent on the surrounding temperature and humidity. Component malfunction is unlikely to go noticed with reading of temperature and humidity, but it can avoid false alarms. If a bad trend is observed in some of the sensors, which models are flagging out, there is a chance it can be linked to a variation in temperature or humidity and not really be a sign of predictive maintenance. This will allow farmers to have a greater economic yield and less wasted and unproductive efforts.

The study by Bouarroudj et al. (2025) evaluated the ability of different models to predict humidity, temperature, and light intensity in a hydroponic setup. Extreme gradient boosting showed the best results at the prediction with an F-1 score of 97.88%. The training was very fast and had robustness across all parameters. Fuzzy logic was also included in the system for actuators to take action based on the analysed data.

Other Factors

Models don't only need to be good at predicting readings and analyse them accurately, but also be economically scalable and easy to install. At times models can require too much physical space, needing larger controllers to feed in data. Other times models might require a lot of computational power, increasing the time for a model to make a prediction, unless a lot of money is invested.

The study by Khanh et. al. (2024) explores the usage of Nonlinear Model Predictive Control (NMPC) and Deep Neural Network (DNN) in a greenhouse environment. NMPC is known to work well with multiple inputs and outputs all at once, be dynamic with new data while also taking constraints of a system into account..

The model first creates a nonlinear model with various equations that describe cause and effect relationships between different factors. Then the model predicts the future of the system over a given time period. The model is then optimised using a cost function, for example the equation used can be $\Sigma(\text{error}^2 + \text{energy cost})$. The optimised sequence can then be converted to action. Then new sensor readings are taken and the entire process repeats. It acts as a continuous live-future predictor and can take steps accordingly. However, the computational power for this is very vast. For that reason, the study after using the NMPC, used the data to create a DNN model so that it accurately replicates an NMPC, with way lower computational costs.

DNN is good at managing large data sets by finding patterns that are very difficult to even notice. The DNN used in the study had 9 input layers, 5 hidden layers (each with 64 neurons), and 7 output layers. The NMPC input and output data was trained with the AI, so that it mimics NMPC behaviour. This example shows how models can be made easy to install, relative to other models, and not have a large initial cost at the same time.

Models-	**Abbreviations**
Artificial Neural Network -	ANN
Convolution Neural Network -	CNN
Decision Tree -	DT
Deep Neural Network -	DNN
Extreme Gradient Boosting -	XGB
K Nearest Neighbour -	KNN

Linear Regression -	LR
Long Short Term Memory -	LSTM
Nonlinear Autoregressive with Exogeneous Inputs -	NARX
Nonlinear Model Predictive Control -	NMPC
Random Forest -	RF
Ridge Regression -	RR
Support Vector Regression -	SVR

Study	**Year**	**Factor**	**Models**	**Key Findings**
Duangpakdee et al.	2024	Yield Prediction	Internet of Things	14% economic profitability with breakeven of 27 months. Needs periodic calibration Sensor accuracy
Mokhtar et al.	2022	Yield Prediction	SVR RF XGB DNN	XGB and RF performed best among the 4 in their specific scenarios. DNN has great predictions in all scenarios and took less computational power
Kadam et al.	2025	Yield Prediction	SVR LR RR DT RF Lasso Regression	DT gave the most accurate predictions of the growth rate
Rahman et al.	2024	Component Malfunction	DT RF KNN SVR XGB	RF had the greatest F-1 score, ensuring both accuracy and precision. Note that models are very dependent on the sensor quality, and more data sets would be required to arrive at a conclusion.
Métwalli et al.	2025	Component Malfunction	RF LR	RF performed best with an accuracy of 94.55%

			DNN	
Dhal et al.	2024	Nutrition	RF SVR KNN	RF performed best at predicting uptake of nitrogen and magnesium SVR performed best at predicting uptake of potassium Uses interpolated data, so possibility of biases Minimal background noise, may not replicate real world application
Arjun et al.	2025	Nutrition	CNN DT	25% reduction in nutrient wastage 18% increase in yield Used reinforced learning for increased responsiveness and continuous improvement
Karimzadeh et al.	2025	EC & pH	RF ANN LSTM KNN SVR	Random Forest performed best for EC anomaly detection. SVR excelled at drift anomaly detection for pH. RF performed better at normal drift detection and precision degradation for pH. ANN performed notably well to predict EC levels with variables and environmental factors.
SIDIBE et al.	2023	EC & pH	NARX	Accurately predicts EC levels 5 minutes prior. Predicts similar trend, with 0.1pH deviation, for pH 5 minutes prior. Easily compatible with micro-controllers, making it cost effective

				and easy to apply.
Bouarroudj et al.	2025	Environmental Factors	XGB	Had a F-1 score of 97.88% in predicting light intensity, humidity, and temperature. Easy to train, has robust parameters, and also integrated fuzzy logic.
Khanh et al.	2024	Other Factors	NMPC DNN	NMPC works well with multiple inputs and is very accurate. DNN can be trained to mimic NMPC decisions, giving notably accurate outputs. This makes it cost efficient and has lower computational cost.

Summary of Different Studies Analyzed

Results

The goal of this research project is to find a few viable models for each of the factors that can help predict a crop failure. Then, select final models, which will be used in the best case setup, based on economical factors. Let's delve into which

For yield prediction, the study by Mokhtar et al. (2022) does note deep neural networks to have consistent accuracy in various scenarios. It's no doubt a winner in this case. Consistency not only helps to be confident in the system, but also reduces missed alarms, which may be caused by certain conditions which the model struggles at. Having a model that is jack of all may be able to point out failures even at the most unimaginable conditions.

Looking into component malfunction, random forest was a clear go-to model. Both studies, by Rahman et al. (2024) and by Métwalli et al. (2025), show random forest to be more accurate than other models in this particular branch. It was because it detected anomalies and slight deviations better than any other model.

In the domain of nutrition, there are various models which could be picked. The study by Bhusan et al. (2024) highlighted how support vector regression worked better for certain nutrients, while random forest worked better for others. It would be ideal to use both in a system as they can then make decisions on linear and non-linear uptake of nutrients. This will allow different nutrients to be tracked accurately, and not just a few.

Electrical conductivity and power of hydrogen anomaly detection and diagnosis was best handled by random forest, with very strong results. However, another notable achievement was nonlinear autoregressive with exogenous inputs (NARX), which generally use neural networks. It precisely predicted the electrical conductivity values for 5 minutes, and had small deviations of 0.1 in pH predictions. This makes it effective to predict a failure 5 minutes before it occurs, allowing time for decision making and also preventing big damages to crops.

Environmental factors are a combination of multiple factors, and many studies came up with different results. However, the study by Bouarroudj et al. (2025), showed extreme gradient boosting to predict humidity, temperature, and light intensity with high accuracy.

Lastly, the financial capital of farmers is a very important factor. Models can cost a lot of money to implement in a setup or have high operational costs. For a model to be suitable, it must be affordable for the majority of the farmers. This is where the study by Khanh et al. (2024) is useful. It did not train DNN to predict or make decisions, but trained it to mimic results another model would give. DNN's are not that costly and can be operated via a simple micro-controller.

For the ideal models to be implemented, the go-to models are going to be random forest, for component malfunction; deep neural network, for yield prediction; and extreme gradient boosting, for environmental factors. For nutrition it may be more viable to use random forest, as it's suggested used for another factor, essentially cutting down costs of training support vectorisation regression. Both the models in electrical conductivity and power of hydrogen performed extremely well in their respective tasks. But, because NARX is slightly more complex, hard to set up, and costly, random forest would be preferred. In case costs are still very high for certain groups of farmers, it is always possible for them to train deep neural network models to try to replicate other models' ways of making decisions. Although it may not be as accurate as the originals, it is a good substitute in financial shortages. All these models will help farmers to make better, informed decisions by cross-checking across different factors readings, and allow maintenance to no longer be done on guesses and instinct, but predictively, making them profitable as well.

Chosen Models for Each Factor

Factor	**Chosen Model**
Yield Prediction	Deep Neural Network
Component Malfunction	Random Forest
Nutrition	Support Vector Regression & Random Forest
EC & pH	Random Forest or Nonlinear Autoregressor with Exogenous Inputs
Environmental Factors	Extreme Gradient Boosting
Financial Factor	Deep Neural Network (mimic other models)

Conclusion

This paper comes up with a list of AIs and their purpose in hydroponic setups, which will allow making predictive maintenance economically beneficially and reduce the risk of crop failure. The models chosen are random forest, extreme gradient boosting, support vector regression, nonlinear autoregression with exogenous inputs, and deep neural network for looking over the factors component malfunction and nutrition, environmental factors, nutrition, electrical conductivity and power of hydrogen, and yield and economical, respectively.

By providing the solution in the results section, it finds the perfect models to predict different factors in a hydroponic setup, using comparative analysis. This helps achieve the best possible prediction for crop failure, and indicate a need for maintenance. This helps the maintenance to be minimum and only at crucial times, reducing profit loss due to less produce and less exertion from the farmer. However, the study highly depends on the accuracy of other papers.

Since this research compares their results to find the best model, any data inaccuracy in the previous studies can put limitations on the result of this study.. Another issue is the lack of information on this topic. Since the hydroponic sector is still an upcoming sector, there is very little research done on neural networks integration in these setups. For this reason, very few studies are done on this topic, and, if practiced wrong, it can lead to partial/incomplete knowledge. There is still a vast future scope to continue research in this field. Since AI is continuously developing and generative AI is also improving, better models can be incorporated, including solutions that can take or give verbal explanations. Another scope would be to test the model suggested by this research in the results in a real setup and check its accuracy and precision in real-life application.

References

1. Arjun, B., Dhanushya, G., Pratussha, M., Srilaxmi, K., & Vinod Reddy, S. (2025, April). AI-driven automated hydroponic nutrient optimization system. Ijirt.Org. Retrieved October 27, 2025, from https://ijirt.org/publishedpaper/IJIRT176948_PAPER.pdf
2. Barbosa GL, Gadelha FD, Kublik N, et al. Comparison of Land, Water, and Energy Requirements of Lettuce Grown Using Hydroponic vs. Conventional Agricultural Methods. International Journal of Environmental Research and Public Health. 2015 Jun;12(6):6879-6891. DOI: 10.3390/ijerph120606879. PMID: 26086708; PMCID: PMC4483736.
3. Blazhevska, V. (2021, April 19). UN calls for urgent action to feed the world's growing population healthily, equitably and sustainably. United Nations Sustainable Development. https://www.un.org/sustainabledevelopment/blog/2021/04/un-calls-for-urgent-action-to-feed-the-worlds-growing-population-healthily-equitably-and-sustainably/
4. Bouarroudj, K., Babaa, F., & Touil, A. (2025, September). IoT-based monitoring and control for optimized plant growth in smart greenhouses using soil and hydroponic systems. Internet of Things (Amsterdam, Netherlands), 33(101710), 101710. https://doi.org/10.1016/j.iot.2025.101710
5. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020, July 22). Language Models are Few-Shot Learners. In arXiv [cs.CL]. http://arxiv.org/abs/2005.14165
6. Dhal, S. B., Mahanta, S., Moore, J. M., & Kalafatis, S. (2024, October 17). Machine learning-based analysis of nutrient and water uptake in hydroponically grown soybeans. Scientific Reports, 14(1), 24337. https://doi.org/10.1038/s41598-024-74376-7
7. Duangpakdee, K., Thananta, G., & Sukpancharoen, S. (2024,August 28). IoT enhanced deep water culture hydroponic system for optimizing Chinese celery yield and economic evaluation. Smart Agricultural Technology, 9(100545), 100545. https://doi.org/10.1016/j.atech.2024.100545
8. Kadam, S., Gohokar, V., & Kute, R. (2025, January). Machine learning and explainable AI for Thai basil growth prediction in hydroponics. IEEE Access: Practical Innovations, Open Solutions, 13(99), 99479–99489. https://doi.org/10.1109/access.2025.3576440

9. Kakal, Z. (2022, February 4). Growing a new leaf: learnings from using hydroponics in informal settlements in Fiji. UNDP. https://www.undp.org/pacific/blog/growing-new-leaf-learnings-using-hydroponics-informal-settlements-fiji
10. Karimzadeh, S., Li, Z., & Ahamed, M. S. (2025, October). Machine learning-based fault detection and diagnosis of electrical conductivity and pH sensors in hydroponic systems. Computers and Electronics in Agriculture, 237(110544), 110544. https://doi.org/10.1016/j.compag.2025.110544
11. Métwalli, A., Shalash, O., Elhefny, A., Rezk, N., Gohary, F. E., Hennawy, O. E., Akrab, F., Shawky, A., Mohamed, Z., Hassan, N., & Hassanen, M. (2025, October 1). Enhancing hydroponic farming with Machine Learning: Growth prediction and anomaly detection. Engineering Applications of Artificial Intelligence, 157(111214), 111214. https://doi.org/10.1016/j.engappai.2025.111214
12. Mokhtar, A., El-Ssawy, W., He, H., Al-Anasari, N., Sammen, S. S., Gyasi-Agyei, Y., & Abuarab, M. (2022, March). Using machine learning models to predict hydroponically grown lettuce yield. Frontiers in Plant Science, 13, 706042. https://doi.org/10.3389/fpls.2022.706042
13. Ogbolumani, O., & Mabaso, B. (2023, December 30). An IoT-based hydroponic monitoring and control system for sustainable food production. Journal of Digital Food Energy & Water Systems, 4(2). https://doi.org/10.36615/digital_food_energy_water_systems.v4i2.2873
14. Rahman, M. A., Chakraborty, N. R., Sufiun, A., Banshal, S. K., & Tajnin, F. R. (2024, August). An AIoT-based hydroponic system for crop recommendation and nutrient parameter monitorization. Smart Agricultural Technology, 8(100472), 100472. https://doi.org/10.1016/j.atech.2024.100472
15. Shi, Z. (2014, November 4). Why air pollution reduces crop yields. World Economic Forum. https://www.weforum.org/stories/2014/11/air-pollution-reduces-crop-yields/
16. Sidibe, A., Ndeda, R., Murimi, E., & Mutwiwa, U. N. (2023, August 30). Implementation of a microcontroller-based neural network for prediction of pH and EC levels in hydroponics. JOURNAL OF SUSTAINABLE RESEARCH IN ENGINEERING, 7(4), 106–117. https://jsre.jkuat.ac.ke/index.php/jsre/article/view/162
17. Van Khanh, P., Giang, L. T. T., & Loan, D. T. (2024). Real-time Greenhouse Climate Control with Deep Learning. Pakistan Journal of Life and Social Sciences, 22(2). https://doi.org/10.57239/PJLSS-2024-22.2.00855
18. Why air pollution reduces crop yields. (n.d.). World Economic Forum. Retrieved October 27, 2025, from https://www.weforum.org/stories/2014/11/air-pollution-reduces-crop-yields/
19. World Population Prospects. (n.d.). Population.Un.Org. Retrieved October 27, 2025, from https://population.un.org/wpp/graphs?loc=900&type=Demographic%20Profiles&category=Line%20Charts