

Article

A Machine-Learning Based IoT System for Optimizing Nutrient Supply in Commercial Aquaponic Operations

Sambandh Bhusan Dhal¹, Kyle Jungbluth², Raymond Lin², Pouyan Sabahi², Muthukumar Bagavathiannan³, Ulisses Braga-Neto⁴, Stavros Kalafatis^{5*}

¹ PhD student, Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA

² BS student, Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA

³ Associate Professor, Department of Soil and Crop Sciences, Texas A&M University, College Station, TX, USA

⁴ Professor, Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA

⁵* Professor of Practice, Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA

* Correspondence: skalafatis-tamu@tamu.edu

Abstract: Nutrient regulation in aquaponic environments has been the topic of research for many years. Most have focused on appropriate control of nutrients in an aquaponic set-up, but very little research has been done on commercial scale applications. In our model, the input data was sourced on a weekly basis from three commercial aquaponic farms in South-East Texas over the course of a year. Due to limited number of data points, dimensionality reduction techniques like pair-wise correlation matrix was used to remove the highly correlated predictors. Feature selection techniques like the XGBoost classifier and Recursive Feature Elimination with ExtraTreesClassifier were used to rank the features in order of their relative importance. Ammonium and calcium were found to be the top two nutrient predictors and based on the months in which lettuce was cultivated, the median of these nutrient values from the historical dataset served as the optimal concentrations to be maintained in the aquaponic solution. To accomplish this, Vernier sensors were used to measure the nutrient values and actuator systems were built to dispense the appropriate nutrient into the ecosystem via a closed loop.

Keywords: aquaponic; pair-wise correlation matrix; XGBoost; Recursive Feature Elimination; ExtraTreesClassifier; median; closed loop

1. Introduction

The system of aquaponics involves the adaptation of hydroponics and aquaculture in a combined set-up which can go on to eliminate the food scarcity and environmental crisis that the world faces today [1][2]. Due to rapid industrialization and large-scale globalization, the economic and social aspects of the varieties of food products emerging from aquaponics has been of great interest as research studies go on to show that it is possible to achieve a ten-fold increase in production without the use of harmful chemicals or pesticides, while using only 2-10% of the water used in traditional agricultural techniques [3] [4] [5]. Although there have been many studies over the last three decades on growing plants in aquaponic set-ups, none of them have focused specifically on determining the important nutrients which need to be monitored and controlled depending on the season in which the crop was cultivated.

There have been numerous studies in which different IoT systems have been implemented in small laboratory set-ups for optimal growth of plants in hydroponic and aquaponic environments. In [6], Valiente et. al. have constructed a system to monitor the pH level and water temperature for growing Nile Tilapia and Romaine Lettuce, and

constructed an actuation system to maintain the pH level between 6.2 to 7.5, and water temperature between 27°C and 30°C to ensure optimal growth. In [7], Yanes et. al. did a review study on IoT and smart systems for monitoring and regulating nitrite concentrations, electroconductivity, Dissolved Oxygen and hardness in the aquaponic solution to ensure a viable commercial model that can be implemented in small-scale semi-automated systems. In [8], Mahanta et. al. did a comprehensive comparative study on the growth of soybeans in hydroponic solution by applying different concentrations of plasma activated water at varying voltages and time intervals to the seeds to minimize heavy metal uptake and ensure better yield. In [9] [10] [11] and [12], the process of sensing and controlling the general parameters of the aquaponic solution like pH and water temperature for growing lettuce and other greens were carried out using different cloud-based approaches. However, very little has been done to formulate data-driven approaches to regulate the important nutrients for growing fish and plants which serves as the main motivation behind this paper.

One of the major problems that one faces while designing an intelligent system for nutrient regulation in aquaponic systems is lack of data. For this, different synthetic data generation techniques have been used. In [13], Soltana et al. used an iterative approach to generate data samples that meets desired statistical distribution, without taking into account the logical constraints, and later tweaking them to fix any violations. In [14], Anderson et al. proposed the generation of synthetic data for IoT devices which included extracting the structure of XML files and characterizing the values within the dataset. In [15], Alzantot et al. presented a deep learning based architecture for synthesizing sensory data named Sensegen which comprised of a LSTM network and a MDN in the first step followed by a LSTM network based discriminator to differentiate between real and generated synthetic data. Similarly, in [16], Dahmen et al. presented a ML-based synthetic data generation technique named SynSys which comprises of nested sequences using Markov models and regression models which are trained on real data. Similarly, in [17] [18] and [19], the concept of synthetic data generation using Monte-Carlo approaches have been stressed on by generating mean and covariance matrix between the classes.

The other problem which is generally faced while using small datasets is selection of relevant features from the dataset. In [20], Chandrashekhar et al. conducted a survey of the feature selection methods which included preprocessing of the training data using PCA and FFT, selection of appropriate features using different filter, wrapper and embedded methods and in the end, carrying out classification using RBF, SVM and ANN classifiers. In [21], Kira et al. focused on empirical test results in two artificial domains using LED Display domain and the Parity domain both with and without noise with a goal to optimize learning and improve quality. In [22], Li et al. proposed feature selection algorithms on conventional data by using similarity-based, information-theoretical-based, sparse-learning-based and statistical-based methods. In [23] and [24], the concept of correlation matrix to eliminate the highly correlated predictors was elaborated upon. In [25], an ensemble of various ML algorithms namely Decision Trees, Random Forest, ExtraTreesClassifier, MLP and SVM were used to generate Histogram of Oriented Gradients (HOG) features required for feature extraction. Similarly, in [26], Shafique et al. did a comparative study of different classification algorithms and ended up with a 90% accuracy score in detecting cardiovascular diseases using ExtraTreesClassifier. In [27] [28] [29] and [30], another important feature selection technique called as XGBoost classifier has been implemented in healthcare to generate feature importance and reduce the size of the datasets resulting in improved classification accuracy.

Finally, after selecting the appropriate nutrients to be regulated in an aquaponic system, sensing and maintaining these parameters at a desired level using an automated system is of utmost importance. In [31], Rau et al. proposed a smart irrigation system consisting of a DHT11 temperature and humidity sensor, and a Raspberry Pi to regularly monitor the weather conditions using an Android application for growing rice. Similarly, in [32], Zhang et al. proposed the use of sensors to measure air temperature, air humidity, soil temperature and soil humidity to monitor the growth of Chinese citrus in real-time

environments using Zigbee technology. In [33], Yimwadsana et al. proposed an IoT controlled system for plant growth consisting of sensors to measure air, light intensity and soil moisture, a dashboard to visualize the collected data and actuator modules consisting of relays, motor gear and water pump controlled by microcontrollers. In [34], Chaudhary et al. proposed a system consisting of a wireless sensor network to monitor and control the environmental parameters in a greenhouse. Similar to the above, in [35], Jaiswal et al. constructed a system to measure the parameters stated above in a greenhouse, controlling them through a Raspberry Pi module and displaying the data on a ThingSpeak cloud platform. Our research measures and regulates the important macronutrients which are determined from data-driven approaches rather than regulating the chemical properties of the aquaponic solution like pH, conductivity, or the environmental parameters of the greenhouse.

2. Methodology

The dataset used in our case was recorded over the course of a year from three aquaponic facilities in South-East Texas: Aquatic Greens Farm (Bryan), Wolff Family Farms (Caldwell) and Texas US Farms (Grimes). From the Aquatic Greens Farm, three samples were collected weekly (one from the tank which bred goldfish, other from the tank which bred tilapia and one from the greenhouse where lettuce was grown). From the Wolff Family Farms, two samples were collected weekly (one from the tank which bred Nile Tilapia and other from the main greenhouse where collard greens, jalapeno peppers and lettuce was grown). From the Texas US Farms, three samples were collected weekly (one from the main growth tank which was used to breed tilapia and shrimp, one from the greenhouse which was used to grow lettuce and kale greens, and the other from the tank which was used to breed seedlings before placing them in the main plant bed). After collecting the samples, these were sent to the Soil, Forage and Water Testing Laboratory at Texas A&M University, USA to determine the major nutrients in the aquaponic solution. A set up of the fish tank and the main plant bed which was used for growing lettuce has been shown in Figure 1.



1(a)



1(b)

Figure 1. a) and 1(b). A set-up of the plant bed which was used for producing lettuce and other greens, and a set-up of the fish tanks which served as a nutrient source for these plants in Texas US Farms, Grimes, TX.

2.1. Analysis of the Dataset

The dataset which was used for analysis in our case had a total of 211 observations and 12 predictors. The predictors consisted of the concentrations of calcium, magnesium, sodium, potassium, boron, carbonate, bicarbonate, sulfate, chloride, nitrate, ammonium and potassium (all of these measured in ppm). The month in which the observations were recorded was used as the class variable, i.e. the Winter and Spring months from November to March in which plants were cultivated were coded as 1, and the Summer and Fall months from April to October were coded as 0 respectively. The predictor which stored the carbonate concentration in ppm was dropped from the dataset before carrying out any preliminary analysis as the variance of the variable was zero throughout.

2.2. Generation of Synthetic Data

As the dataset was too small to make inferences on, it was an important measure to generate synthetic data before applying any data-driven approach. The dataset at hand had 107 observations belonging to Class 0 and 104 observations belonging to Class 1. Therefore, it was decided to use the concept of generation of mean and covariance matrix between and among the classes using two variations of Monte-Carlo (MC) approaches to generate synthetic data which was used for carrying out inferences. Both approaches have been highlighted in the next section of the paper.

2.3. Feature Selection

Before constructing an IoT system for sensing and regulating the nutrient parameters in a feedback loop, it is important to select the most important parameters to scale down the cost of the system. For this, three methods were used in a pipeline for reducing the dimensionality of the dataset and has been depicted in Figure 2 below. A pairwise correlation matrix was used to find out the correlation between the predictors, followed by XGBoost classifier to find out the F-scores for each of them, and finally, ExtraTreesClassifier with Recursive Feature Elimination to find out the two most important nutrients for which the IoT based control system was set up.

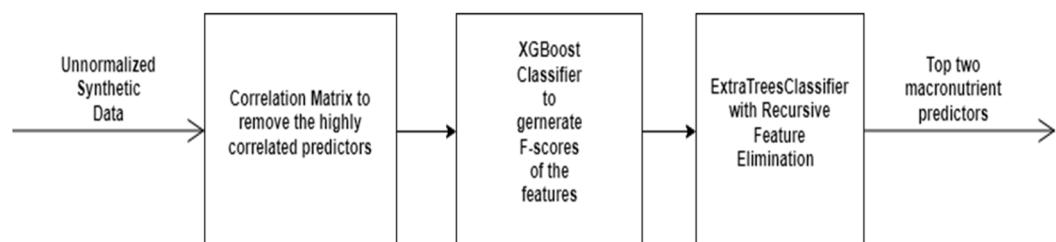


Figure 2. Feature Selection Pipeline for choice of appropriate nutrient predictors.

2.4. IoT System for Nutrient Regulation

In order to build the IoT system for regulating nutrients, we divided the entire system into three parts namely (a) the sensor subsystem, (b) the feedback loop and (c) the actuator system. The sensor subsystem consists of two sensors, manufactured by Vernier, to measure the important nutrients, and output the data to the Raspberry Pi through USB. Each sensor sampled the water 5 times before outputting an average of the results to the feedback loop. The data gathered will be in parts per million (ppm).

To properly measure and regulate the nutrient levels within the entire system, a feedback loop is used to continually monitor the environment. It connects to both the sensor system and actuator system to retrieve current nutrient amounts and be able to release more if needed. The feedback loop works by taking readings from the Vernier sensors and averaging them 10 times to use as the current nutrient level. If the measured levels remain below the targeted nutrient levels, the actuator system will receive a signal through the GPIO pins of the Raspberry Pi to run one cycle of nutrient release.

The actuator subsystem in our case consists of a PIC microcontroller, 12V tolerance motor modules, and two 12V stepper motors to dispense the nutrients. The stepper motors are connected to the motor modules which supply the motors with 12V as well as the signals to rotate the dispensing portion. The motor modules are then connected to the PIC microcontroller, which determines the rotational direction of the motor and the speed of rotation. The motors will only activate when the PIC microcontroller receives a bit signal from the Raspberry Pi to indicate when more nutrients are needed.

3. Results

After removing the carbonate values due to zero variance in the dataset, we proceeded with synthetic data generation using Monte-Carlo approaches. We did not standardize or normalize the data as all the predictors were distributed normally. A total of 2,675 data points were generated for each class using a variation of Monte-Carlo technique in which a separate mean and covariance matrix was produced for each class. Similarly, 5,350 synthetic data points were generated, sharing the mean and covariance matrix between the classes. This resulted in a total of 10,700 observations with 12 predictors, which were then used for further analysis.

As mentioned in the section above, a pipeline of feature selection techniques was applied to the dataset to select the top two nutrients. At first, a pair-wise correlation matrix was generated between the predictors and is mentioned in Figure 3.

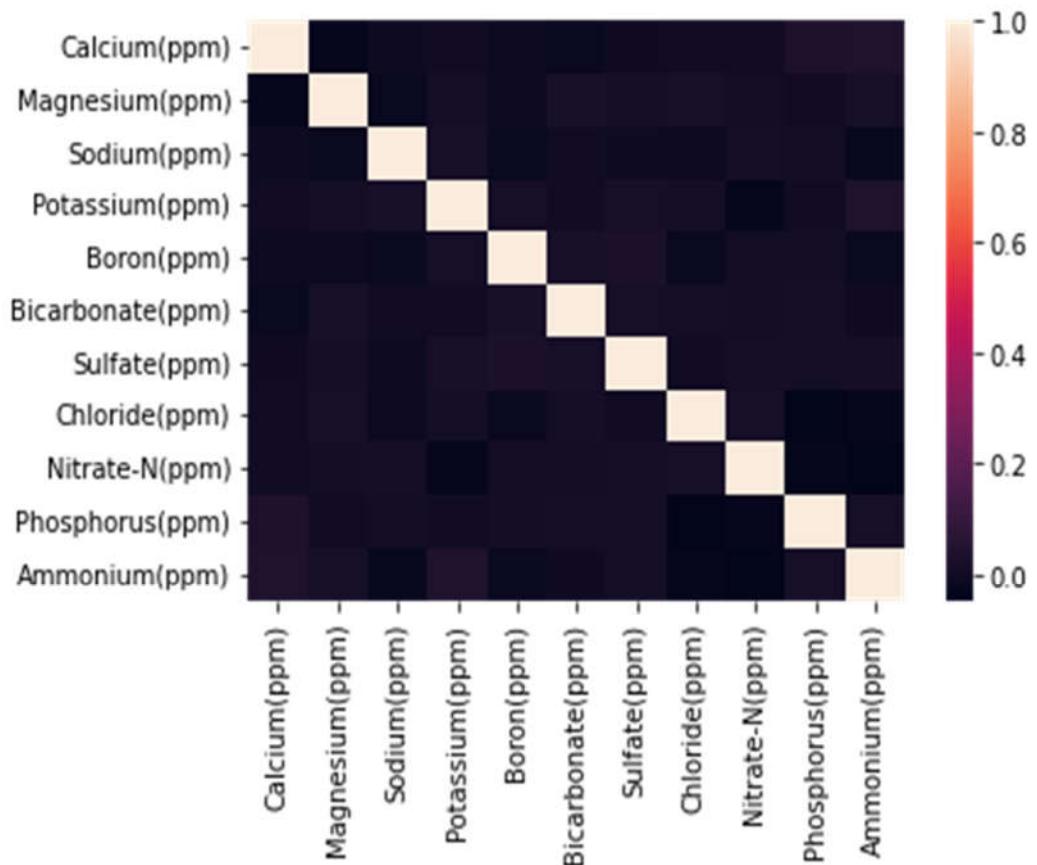


Figure 3. Pair-wise Correlation Matrix among all the nutrient predictors in the dataset.

From Figure 3, it can be inferred that the pair-wise correlation between the predictors is not greater than 0.05. Therefore, it was decided to proceed with the entire dataset without eliminating any predictors. Further, the XGBoost algorithm was used to generate F-scores for each predictor, shown in Figure 4.

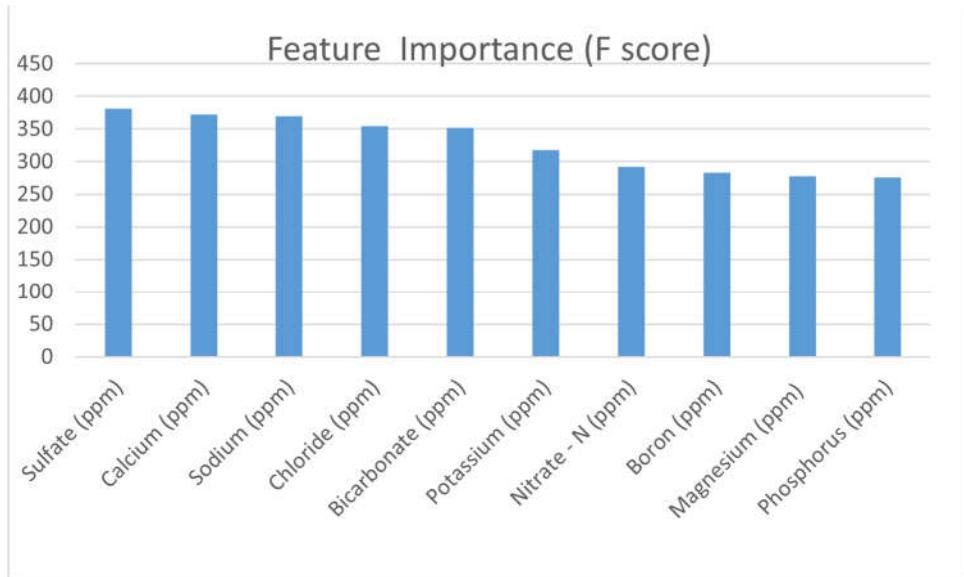


Figure 4. F-Score of each predictor generated by the XGBoost algorithm.

Figure 4 shows that the feature importance of sulfate is the highest as the F-score generated by the XGBoost algorithm is around 375. This was followed by the feature importance of calcium, sodium, chloride, bicarbonate and potassium with values ranging between 325 and 375. The rest of the nutrient predictors were excluded from our analysis before proceeding to the third step of the feature selection pipeline.

The final step in the feature selection process is applying the ExtraTreesClassifier with Recursive Feature Elimination on the trimmed dataset to generate feature importance, shown in Figure 5.

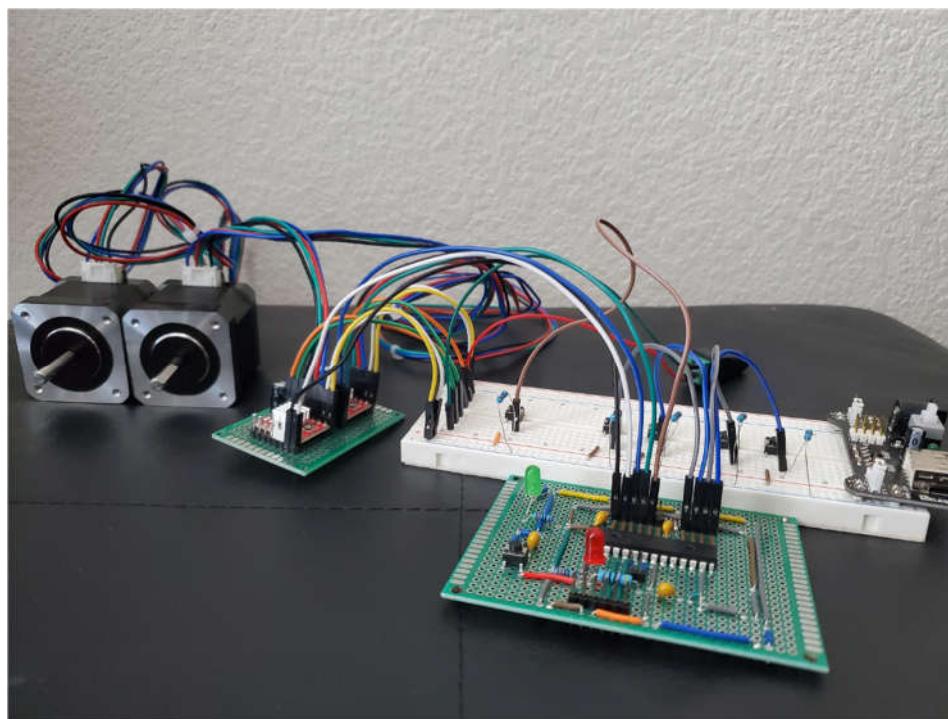


Figure 5. Feature Importance of each predictor generated by the ExtraTreesClassifier.

The ExtraTreesClassifier revealed that the feature importance values of ammonium and calcium were 52% and 16%, respectively, making up 68% of the total feature importances in the dataset. Next, the main focus was to automate the sensing and actuation of these nutrients using a feedback loop. As discussed above, two Vernier sensors were used to measure ammonium and calcium, and two actuator modules were designed to release these nutrients if they fall below the recommended concentrations. A model set-up of the sensing, actuation and feedback unit is shown in Figure 6.



6(a)



6(b)



6(c)

```

Machine Model Output: 1
Nutrient Level: 1.8 ML Level: 1.55 Don't release ammonium
Nutrient Level: 29.0 ML Level: 27 Don't release magnesium
Machine Model Output: 0
Nutrient Level: 1.95 ML Level: 1.82 Release ammonium
Nutrient Level: 29.15 ML Level: 32 Release magnesium
Machine Model Output: 1
Nutrient Level: 1.9 ML Level: 1.55 Don't release ammonium
Nutrient Level: 29.1 ML Level: 27 Don't release magnesium

```

6(d)

Figure 6. a), 6(b), 6(c), 6(d). A set-up of the Vernier sensors to measure calcium and ammonium, a set-up of the motor control units to control the actuators, a set up of the actuator modules to release the nutrients, and a feedback loop connecting the sensor and actuator modules.

From Figure 6, it can be inferred that the output of the ML model is 0 or 1 which states the month in which the observations were recorded and has a set of recommended levels of nutrients that need to be maintained in the aquaponic solution. For each

recommendation value stated in the table, the median of the nutrient values from the historical dataset per class was chosen as the concentration that should be maintained in the aquaponic solution. The recommendation system used here has been discussed in detail in Table 1.

Table 1. The Recommendation System for Nutrient Regulation in Aquaponic Environments.

Sl. No.	Class Name	Calcium Concentration (ppm)	Ammonium Concentration (ppm)
1	0 (April to October)	32	1.82
2	1 (November to March)	27	1.55

4. Discussion

The main motivation behind developing this approach is to build a recommendation system with real-time sensing and actuation units to regulate nutrient levels in the aquaponic solution for optimal growth of fish and crop plants in a single set-up. In order to achieve this, the pre-recorded observations from the historical dataset were used to design the approach and a set of rules were prescribed for both the classes.

This is the first of its kind where IoT systems were combined with Machine Learning for optimizing nutrient supply in aquaponic solutions. The major advantage of a data-driven IoT system compared to a traditional aquaponic system is cost savings through improved yield and quality of produce via the integration of AI. This in turn significantly increases the profits generated from the aquaponic operations. Another solution provided in this research is the reduction in the amount of nutrients supplied for plant growth. In many of the earlier aquaponic studies, excess phosphorus accumulation due to unregulated use of fertilizers has resulted in eutrophication, leading to algal growth and a steep decline in fish populations [36]. The nutrient accumulation problem has been addressed in this research through IoT-based nutrient optimization.

For the observations to which the month class is 0, cultivation occurs during the dry and hot summers of Texas. During these months, the fish in the aquaponic set-up grew from fries to the adult stage. As ammonium is a byproduct formed from protein metabolism, they are excreted in high amounts during these months [37] [38]. Since high ammonium concentrations are toxic to fish, it is important to restrict ammonia to a maximum level of 1.82 ppm. On the other hand, ammonium is a source of nitrogen for plants [39], resulting in higher yields, though greater than optimal levels can also be injurious to crop growth. For the observations to which the month class is 1, cultivation is done during the cold winter months. During this period, the fish growth is minimal, and in turn the amount of ammonium released in water is comparatively low. The concentration of ammonium during the winter months can be restricted to 1.55 ppm, which is sufficient for plant growth and is also a safe limit for sustainable fish growth in commercial set-ups.

Most of the fish grown during the summer months in aquaponic set-ups in Texas are warm-water fish species. As mentioned above, calcium is an important element required for normal growth and reproduction of fish species [40]. Calcium levels need to be maintained at 32 ppm during these months, which allows for maintenance of water hardness between 75 and 125 ppm, while still safe for plant growth. Calcium concentrations beyond this level can cause tip burn in lettuce plants. During the winter months, plant growth is considerably slower and excess levels of carbonates or bicarbonates can be toxic to plants

[41]. Therefore, the recommended level of calcium in the aquaponic solution was 27 ppm, which was sufficient to sustain fish growth without injuring the winter crops.

This study specifically addresses the optimization of the top two nutrients (ammonium and calcium), which should be regulated depending on the month in which the plants and fish are grown in a closed-loop control aquaponic set up. As ammonium is a key factor in controlling the pH and TDS of the solution, regulating the ammonium concentration addresses most of the issues which may occur due to TDS imbalance in the solution [42][43]. On the other hand, the concentration of calcium plays an important role in determining the hardness of the solution. Therefore, regulating both of these nutrients through an IoT based sensing and actuation system ensures that pH is adequately regulated for promoting optimal growth and yield of both fish and crop plants in a single set-up.

As the entire data on which our model was trained to build an AI-based approach was recorded from aquaponic farms located in Texas, the IoT system developed here cannot be used directly in places that are extremely harsh or cold. The important features determined by our dimensionality reduction techniques are bound to change in those conditions due to high variance expected in the dataset. However, this study provides a general approach that can be adapted to a range of environments and aquaponics settings with adequate modifications. Future research should investigate the effectiveness of the IoT approach for optimizing nutrient supply in extreme climatic conditions and diverse management scenarios.

5. Conclusions:

Using the dataset recorded from three aquaponic farms in South-East Texas, the two most important predictors identified from the existing design, after generating synthetic data and carrying out dimensionality reduction, were the concentrations of calcium and ammonium. These two nutrients were regulated using an IoT based sensing and actuation system in a closed loop set up. The tests were carried out for the cycle of 21 days to grow Romaine Lettuce and were verified experimentally as to how the proposed recommendations positively impacted plant and fish growth in a single system. The size and yield of the lettuce was compared to the ones grown in unregulated aquaponic solution and the yield showed significant increase in size with some of them being as large as 40 to 45 inches in diameter. The cost involved in regulating nutrient parameters compared to traditional aquaponic environments also decreased by more than 75%.

6. Future Work:

In the future, a cloud database can be set up to host the Machine Learning model and store data dynamically. The size of the sensing and actuation modules described above can be scaled up to regulate more macronutrients and heavy metals to build a comprehensive unit to be installed at commercial setups. The data collected in our case to generate the AI-based approach was collected from three aquaponic farms in Texas where the terrain is mostly plain. More data can be added in the future from diverse geographical locations with more variable weather conditions to improve the robustness of the model.

Author Contributions: Conceptualization, S.B.D; methodology, S.B.D; software, S.B.D, K.J; hardware, R.L., P.S., M.B.; validation, S.B.D, K.J., R.L., P.S.; formal analysis, S.B.D, M.B., S.K.; investigation, S.B.D., U.B.N., S.K.; resources, S.K.; data curation, S.B.D.; writing—original draft preparation, S.B.D.; writing—review and editing, S.B.D., M.B., S.K.; visualization, S.B.D.; supervision, S.K.; project administration, S.K.; funding acquisition, S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding, but this research was funded by the Departmental grants from Department of Electrical and Computer Engineering, Texas A&M University

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The dataset used in this case has been described in detail in the sections above. The researchers would be willing to provide more details if needed.

Acknowledgments: I am grateful to Ms. Sharon Wells, owner of Aquatic Greens Farm, Bryan; Mr. Robert Wolff, owner of Wolff Family Farms, Caldwell; and Mr. Joe Leveridge, owner of Texas US Farms, Grimes County for their cooperation and providing access to their hydroponic farms for experimentation. Their insights have been equally helpful in designing the approach that has been implemented here.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pillay, T. V. R., & Kutty, M. N. (2005). Aquaculture: principles and practices (No. Ed. 2). Blackwell publishing
2. Jensen, M. H. (1997). Hydroponics. HortScience, 32(6), 1018-1021.
3. Pillay, T. V. R. (2008). Aquaculture and the Environment. John Wiley & Sons.
4. Jones Jr, J. B. (2016). Hydroponics: a practical guide for the soilless grower. CRC press.
5. Roberto, K. (2003). How-to hydroponics. Futuregarden, Inc.
6. Valiente, Flordeliza L., et al. "Internet of things (IOT)-based mobile application for monitoring of automated aquaponics system." 2018 IEEE 10th international conference on humanoid, nanotechnology, information technology, communication and control, environment and management (HNICEM). IEEE, 2018.
7. Yanes, A. Reyes, P. Martinez, and R. Ahmad. "Towards automated aquaponics: A review on monitoring, IoT, and smart systems." Journal of Cleaner Production 263 (2020): 121571.
8. Mahanta, S.; Habib, M.R.; Moore, J.M. Effect of High-Voltage Atmospheric Cold Plasma Treatment on Germination and Heavy Metal Uptake by Soybeans (*Glycine max*). *Int. J. Mol. Sci.* 2022, 23, 1611. <https://doi.org/10.3390/ijms23031611>
9. Dutta, Abhay, et al. "IoT based aquaponics monitoring system." 1st KEC Conference Proceedings. Vol. 1. 2018.
10. Lee, Chien, and Yu-Jen Wang. "Development of a cloud-based IoT monitoring system for Fish metabolism and activity in aquaponics." Aquacultural Engineering 90 (2020): 102067.
11. Aishwarya, K. S., et al. "Survey on IoT based automated aquaponics gardening approaches." 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). IEEE, 2018.
12. Jacob, Nikhil Kurian. "IoT powered portable aquaponics system." Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing. 2017.
13. G. Soltana, M. Sabetzadeh and L. C. Briand, "Synthetic data generation for statistical testing," 2017 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE), 2017, pp. 872-882, doi: 10.1109/ASE.2017.8115698.
14. J. W. Anderson, K. E. Kennedy, L. B. Ngo, A. Luckow and A. W. Apon, "Synthetic data generation for the internet of things," 2014 IEEE International Conference on Big Data (Big Data), 2014, pp. 171-176, doi: 10.1109/BigData.2014.7004228.
15. Alzantot, Moustafa, Supriyo Chakraborty, and Mani Srivastava. "Sensegen: A deep learning architecture for synthetic sensor data generation." 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). IEEE, 2017.
16. Dahmen, Jessamyn, and Diane Cook. "SynSys: A synthetic data generation system for healthcare applications." Sensors 19.5 (2019): 1181.
17. Bonner, Courtney. "Markov chain Monte Carlo synthetic data generation from casino slot floor event data." (2022).
18. Socco, Laura Valentina, and Daniele Boiero. "Improved Monte Carlo inversion of surface wave data." Geophysical Prospecting 56.3 (2008): 357-371.
19. Kiviet, Jan F. "Monte Carlo simulation for econometricians." Foundations and Trends® in Econometrics 5.1–2 (2012): 1-181.
20. Chandrashekhar, Girish, and Ferat Sahin. "A survey on feature selection methods." Computers & Electrical Engineering 40.1 (2014): 16-28.
21. Kira, Kenji, and Larry A. Rendell. "A practical approach to feature selection." Machine learning proceedings 1992. Morgan Kauffmann, 1992. 249-256.
22. Li, Jundong, et al. "Feature selection: A data perspective." ACM computing surveys (CSUR) 50.6 (2017): 1-45.
23. Steiger, James H. "Tests for comparing elements of a correlation matrix." Psychological bulletin 87.2 (1980): 245.
24. Guyon, Isabelle. "Practical feature selection: from correlation to causality." Mining massive data sets for security: advances in data mining, search, social networks and text mining, and their applications to security (2008): 27-43.
25. Abhishek, L. "Optical character recognition using ensemble of SVM, MLP and extra trees classifier." 2020 International Conference for Emerging Technology (INCET). IEEE, 2020.
26. Shafique, Rahman, Arif Mehmood, and Gyu Sang Choi. "Cardiovascular disease prediction system using extra trees classifier." (2019).
27. Zhanshan, Li, and L. I. U. Zhaogeng. "Feature selection algorithm based on XGBoost." Journal on Communications 40.10 (2019): 101.
28. Chen, Cheng, et al. "Improving protein-protein interactions prediction accuracy using XGBoost feature selection and stacked ensemble classifier." Computers in Biology and Medicine 123 (2020): 103899

29. Ogunleye and Q. -G. Wang, "XGBoost Model for Chronic Kidney Disease Diagnosis," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 6, pp. 2131-2140, 1 Nov.-Dec. 2020, doi: 10.1109/TCBB.2019.2911071.
30. Guo, Junqi, et al. "An XGBoost-based physical fitness evaluation model using advanced feature selection and Bayesian hyper-parameter optimization for wearable running monitoring." *Computer Networks* 151 (2019): 166-180.
31. J. Rau, J. Sankar, A. R. Mohan, D. Das Krishna and J. Mathew, "IoT based smart irrigation system and nutrient detection with disease analysis," *2017 IEEE Region 10 Symposium (TENSYMP)*, 2017, pp. 1-4, doi: 10.1109/TENCONSpring.2017.8070100.
32. Zhang, Xueyan, et al. "Monitoring citrus soil moisture and nutrients using an IoT based system." *Sensors* 17.3 (2017): 447.
33. Yimwadsana, Boonsit, et al. "An IoT controlled system for plant growth." *2018 Seventh ICT International Student Project Conference (ICT-ISPC)*. IEEE, 2018.
34. Chaudhary, D. D., S. P. Nayse, and L. M. Waghmare. "Application of wireless sensor networks for greenhouse parameter control in precision agriculture." *International Journal of Wireless & Mobile Networks (IJWMN)* 3.1 (2011): 140-149.
35. Jaiswal, Himanshu, Ram Singuluri, and S. Abraham Sampson. "IoT and machine learning based approach for fully automated greenhouse." *2019 IEEE Bombay Section Signature Conference (IBSSC)*. IEEE, 2019.
36. Cohen, Abigail, et al. "Combined fish and lettuce cultivation: an aquaponics life cycle assessment." *Procedia Cirp* 69 (2018): 551-556.
37. Eshchar, Micha, et al. "Intensive fish culture at high ammonium and low pH." *Aquaculture* 255.1-4 (2006): 301-313.
38. Smutna, M., L. Vorlova, and Z. Svobodova. "Pathobiochemistry of ammonia in the internal environment of fish." *Acta Veterinaria Brno* 71.2 (2002): 169-181.
39. Bittsanszky, Andras, et al. "Nutrient supply of plants in aquaponic systems." *Ecocycles* 2.2 (2016): 17-20.
40. Wendelaar Bonga, S. E., and G. Flik. "Calcium regulation in fish." (1993).
41. Burstrom, H. G. "Calcium and plant growth." *Biological Reviews* 43.3 (1968): 287-316.
42. Arora, Amarpreet Singh, et al. "Energy saving anammox technology-based nitrogen removal and bioenergy recovery from wastewater: Inhibition mechanisms, state-of-the-art control strategies, and prospects." *Renewable and Sustainable Energy Reviews* 135 (2021): 110126.
43. Bu, Jianwei, et al. "Shallow groundwater quality and its controlling factors in the Su-Xi-Chang region, Eastern China." *International Journal of Environmental Research and Public Health* 17.4 (2020): 1267.