sustainability

MDPI

Article

A Prescriptive Intelligent System for an Industrial Wastewater Treatment Process: Analyzing pH as a First Approach

Luis Arismendy [1], Carlos Cárdenas [1], Diego Gómez [1], Aymer Maturana [2], Ricardo Mejía [2] and Christian G. Quintero M. [1,*]

[1] Department of Electrical and Electronics Engineering, Universidad del Norte, Barranquilla 081007, Colombia; arismendyl@uninorte.edu.co (L.A.); ccarlosa@uninorte.edu.co (C.C.); dgomez@uninorte.edu.co (D.G.)
[2] Department of Civil and Environmental Engineering, Universidad del Norte, Barranquilla 081007, Colombia; maturanaa@uninorte.edu.co (A.M.); marchenar@uninorte.edu.co (R.M.)
* Correspondence: christianq@uninorte.edu.co

Abstract: An important issue today for industries is optimizing their processes. Therefore, it is necessary to make the right decisions to carry out these activities, such as increasing the profit of businesses, improving the commercial strategies, and analyzing the industrial processes performance to produce better goods and services. This work proposes an intelligent system approach to prescribe actions and reduce the chemical oxygen demand (COD) in an equalizer tank of a wastewater treatment plant (WWTP) using machine learning models and genetic algorithms. There are three main objectives of this data-driven decision-making proposal. The first is to characterize and adapt a proper prediction model for the decision-making scheme. The second is to develop a prescriptive intelligent system based on expert's rules and the selected prediction model's outcomes. The last is to evaluate the system performance. As a novelty, this research proposes the use of long short-term memory (LSTM) artificial neural networks (ANN) with genetic algorithms (GA) for optimization in the WWTP area.

Keywords: artificial neural network (ANN); chemical oxygen demand (COD); data-driven decision making (DDDM); Industry 4.0; machine learning (ML); optimization; wastewater treatment plant (WWTP)

check for updates

Citation: Arismendy, L.; Cárdenas, C.; Gómez, D.; Maturana, A.; Mejía, R.; Quintero M., C.G. A Prescriptive Intelligent System for an Industrial Wastewater Treatment Process: Analyzing pH as a First Approach. *Sustainability* **2021**, *13*, 4311. https://doi.org/10.3390/su13084311

Academic Editor: Jae Kwang (Jim) Park

Received: 16 February 2021
Accepted: 16 March 2021
Published: 13 April 2021

1. Introduction

A brilliant explosion in the accessibility and availability of information through the data of many industrial processes has opened the data analysis to describe, predict, and prescribe to make better decisions in processes like iron extraction, food chains, medicine production, and energy generation [1]. However, there is not currently enough research in the prescription area. Therefore, there is plenty of topics to discuss how to interpret prediction events to make intelligent decisions [2]. In search of optimal achievement of the industries' goals, prescriptive analytics support their processes intelligently using inference and predictions to avoid future faults [3]. Through a timely prediction of out-of-range values, the authors in Reference [3] take precautions representing savings in operational costs. However, today "the deep relation between predictive and prescriptive analytics still is neither well understood nor fully exploited" [4]. Exploiting the research results in the area, the enterprises could analyze their processes through the continually growing data and the prescriptive analysis method, leading to controlling their activities efficiently [5]. In terms of competition, the industries empowered by prescriptive analysis techniques will lead to the evolution of the next industrial level with the ability to decide in real-time. This paper proposes research on the prescriptive analysis, showing the impact and potential in the industry. Specifically, this paper focuses on a case study for an industrial wastewater treatment plant (WWTP), a facility accustoming a mix of several processes (e.g., physical, chemical, and biological) to treat industrial wastewater and take away pollutants [6]. The

treated water in these plants can be classified according to its source. For instance, domestic wastewaters are liquids from residences, commercial, and institutional buildings. This municipal wastewater is the liquid waste transported by the sewer of a city or town and treated in a municipal treatment plant. On the other hand, industrial wastewater is the water from the discharges of manufacturing industries [7]. In WWTP, a biological treatment process known as activated sludge reduces the organic load in these waters thanks to the aerobic microorganisms' action. Figure 1 shows a general structure of biological treatment. The variety of microorganisms in this treatment makes the process highly nonlinear, which is a real process to investigate. However, not everything is lost yet since computational algorithms have high accuracy to cope with complex systems. For instance, artificial neural networks (ANNs) are part of a set of computational algorithms known for their ability to accurately approximate many processes' general behaviour, whether complex or straightforward. This technology is inspired by biological neural networks wherein input signals are processed by neurons that answer with an output signal [8]. For some years now, it has used neural networks in regression or classification problems. References [9,10] developed examples of regression problems using ANN technology. In classification, computer vision found an essential resource in ANNs to identify objects, animals, and people in images and videos. Some examples in classification can be seen in the latest works [11,12]. It is worth noting that there will be a type and a more suitable network architecture for each application type. Among the most popular networks are convolutional networks, recurrent networks, and multilayer perceptron networks.

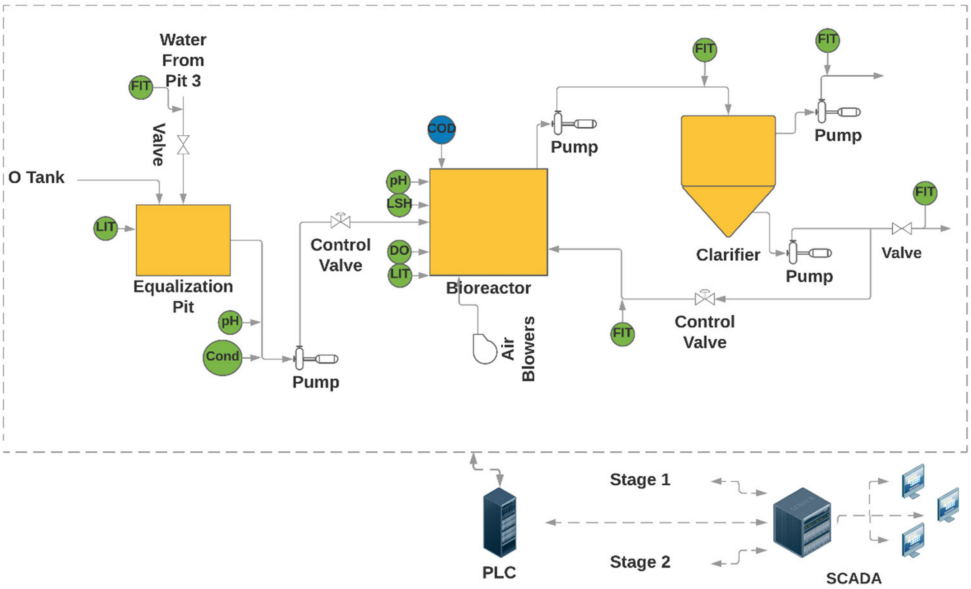

Figure 1. Biological treatment in the industrial wastewater treatment plant (WWTP) case study.

For processes where variables can be monitored and essential in their efficiency, the artificial neural networks used for regression applications could be quite helpful. One of the many benefits is the opportunity to predict the behaviour of a process [13]. Still, it would be much more advantageous if, beyond predictive analysis, a process applies prescriptive analysis techniques. Prescriptive analytics is about making intelligent decisions that favour the analyzing process based on the conclusions provided by predictive analytics. In other words, it optimizes the process [14]. Section 2 will detail works that take advantage of these techniques. One technique for prescriptive analysis purposes to highlight is genetic algorithms in evolutionary computation. The process of natural biological selection inspires this computational model, which, over time, finds the fittest species. This algorithm leads to find the best possible solution for a complex optimization problem with the conviction of not falling into local lows in the optimization zone [15]. Therefore, it is one of the most used algorithms in complex problems.

As mentioned above, the industrial WWTP process's biological treatment in this work is highly nonlinear. Consequently, the purpose of this work is to answer the question: can an intelligent computational model make a prescriptive decision based on available predictive data in an industrial WWTP?

2. Related Works

Currently, several works about descriptive and predictive analytics can be found in the literature. However, prescriptive analytics works are a bit behind in terms of research published. Recent interest in this topic has increased [16]. Reference [17] presents a systematic literature review on prescriptive analytics. According to this review, related works' prescriptive methods are classified into Probabilistic Models, Machine Learning/Data Mining, Mathematical Programming, Evolutionary Computation, Simulation, and Logic-based Models. The authors in Reference [17] make this classification based on the prescriptive techniques used in all the works found in their literature review. Then, it will discuss the intelligent systems approaches developed by the works. Figure 2 represents a review of the methods used. According to this, Table 1 presents the number of published articles in each category.

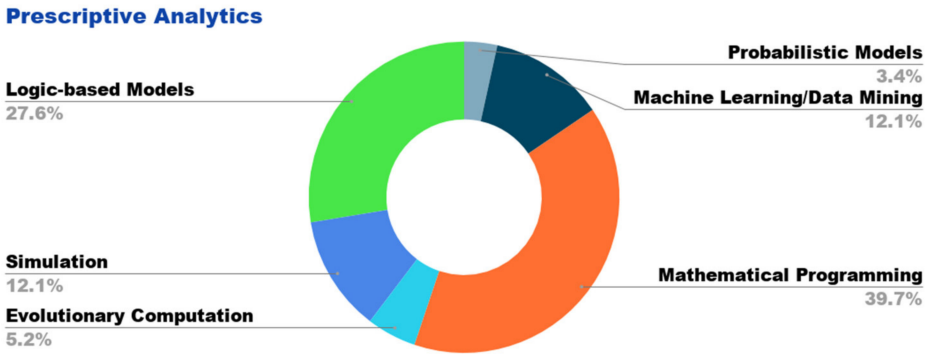

Figure 2. Prescriptive analytics literature review.

Table 1. Published articles according to a prescriptive analysis techniques classification in Reference [17].

Category	Number of Published Articles [1]
Probabilistic Models	2
Machine Learning/Data Mining	7
Mathematical Programming	23
Evolutionary Computation	3
Simulation	7
Logic-based Models	16

[1] Publications until 2020.

Below are the most relevant works. In Reference [18], prescriptive analytics help classify information into secure and insecure for avoiding security vulnerability in the Hadoop framework. Then the system decides the preferred location for writing the data. The authors implemented an unsupervised machine learning algorithm for clustering. In Reference [19], the authors developed two prescriptive methods based on Nadaraya-Watson and nearest-neighbours learning to prescribe an optimal decision applied to a small newsvendor problem. A case study in Reference [20] is tackled using the ANN model and a genetic algorithm to optimize product quality in complex industrial processes. The primary aim was to find the behaviour of the alloying elements in steel with the desired performance. Reference [21] presents a prescriptive system for determining hotel room prices to be published in price brochures. The work solved the trade-off between profit-maximizing and an easy-to-read price brochure. A human resource planning model in

Reference [22] is developed to make hiring decisions and maximize profit at firms that sell contract-based consulting projects.

In the lodging industry, prescriptive analytics let authors in Reference [23] plan well lodging capacity before the FIFA (French: Fédération Internationale de Football Association) World Cup. They tested alternative scenarios of 32 qualifying teams considering critical foreign spectator attendance factors and comparing it with the FIFA seat allocation mechanism.

Regarding health management for the electric power grid, IBM Research in Reference [24] developed a system to test asset health, suggesting an optimal maintenance strategy considering budgetary constraints for the electric power grid in an enterprise. The results improved the prioritization process based on both the risk and the impact of each budget allocation. In the simulation field, a project in Reference [25] named Predictive Analytics for Server Incident Reduction (PASIR) simulates the operation of some servers in real-time to classify them into problematic and non-problematic classes. For those problematic ones, the system recommends modernization actions analyzing the behaviour of the server. The research [26] conceives a new framework comprising random forest, Bayesian belief networks, and ARIMA (AutoRegressive Integrated Moving Average) models to overcome some difficulties, such as identifying key performance indicators (KPIs) and incorporating the KPIs' temporal effects into predictive analytics. Authors in Reference [27] combined both product portfolio configuration and prescriptive methods to satisfy volatile market demands and accomplish company objectives. This way, companies' product management is closer to making the right decision and catching more customers.

About logic-based models, EventAction is the first interface to give recommendations about temporal event sequences. Reference [28] shows event action in the context of student advising. The interface recommends temporal event sequences that might help students to accomplish their academic goals.

Finally, a recent case that uses the advantages of prescriptive analytics is the work carried out by the authors in the water treatment topic presented in Reference [3]. However, in the area, a lack of research related to prescriptive analytics is notorious, just like in other fields. In Reference [3], an intelligent decision-making system reduces the membrane fouling incidence. A self-organizing deep belief network compounds the first stage of the system. The second stage shows the strength of a multi-warning method based on independent component analysis and principal component analysis. Finally, authors in the third stage develop a multi-category diagnosis method on a kernel function.

In general, the researchers are just beginning to be interested in prescriptive analytics techniques in areas like information security, business, industry, health management, computing, and education, according to the works found mainly from 2017 to 2020. However, in the WWTP field, only one work with a system using computational techniques appears under the searching conditions mentioned before. As a novelty, the proposed work in our paper uses long short-term memory (LSTM) neural networks with genetic algorithms (GA) for prediction and optimization in the WWTP field, in which, as can be seen before, it has not found a publication using these techniques together.

3. Materials and Methods

Wastewater treatment processes usually monitor a set of variables to provide information on how the process is developing, such as chemical oxygen demand (COD), which gives information about efficiency. This set of variables detailed later classify whether the action of some other variable indirectly controls the variable or whether it can be manipulated directly with a controller based on the type of control used over each one [29]. Thus, changing the course of the efficiency of a process can be possible by manipulating the correct variables optimally.

A biological wastewater treatment process can be manipulated to affect its efficiency positively. Therefore, the proposed approach starts by characterizing the variables that lead to the improvement of this efficiency. In this sense, one important aspect is that the

objective will always be to optimize the process, so prescribing setpoints corresponding to each manipulated variable and including a set of components in the prescription's implementation system is necessary. A conceptual approach to develop a prescription system is shown in Figure 3 [20]. The proposed components of this system are:

(1) Prediction models;
(2) Desirability functions;
(3) Compound desirability;
(4) Optimization algorithm.

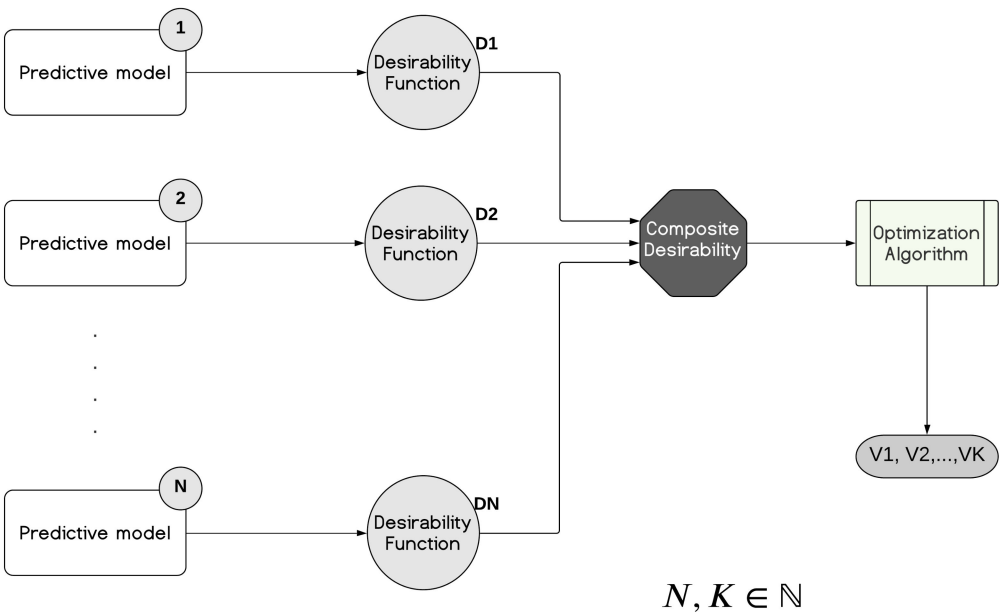

Figure 3. The conceptual approach for developing a prescription system.

Below, this paper states the justification for the use of these components. The first reason lies in taking advantage of the conclusions generated by N predictive models to analyze how the characterized variables will impact the future of the process. According to the literature review, these models are usually advanced computing techniques because of their accuracy [30]. As shown in the Related Works section, some advanced computing techniques are ANNs, Bayesian belief networks, ARIMA models, and self-organizing deep belief networks.

As mentioned before, the main objective is to modify variables intelligently in search of optimal improvement. Therefore, it is necessary to use optimization algorithms that find the best setpoint within an appropriate margin determined by experts in the process. It is appropriate to limit this margin to use functions that transform each variable's range into a convenient range. For this reason, desirability functions are essential. Using a compound desirability function is vital to condense the desirability functions when more than one variable is part of the system. Finally, the experts' recommendations and the biological process variables in the wastewater treatment plant lead to contextualizing the conceptual approach in Figure 3.

COD is one of the most critical variables in the process of a biological treatment [31] because of the information COD provides, leading to making important decisions. The objective of biological wastewater treatment is to remove the pollutant load in water [32]. In the bioreactor at this stage, a set of microorganisms of different natures break down a percentage of the water's organic matter. For studying the COD dynamics in the process, a dataset is received from a WWTP from Nantong, China, with a daily data frequency for 877 samples. In this dataset, it registered 22 variables from 1 December 2017 to 16 July 2020. Figure 4 describes the COD dynamics.

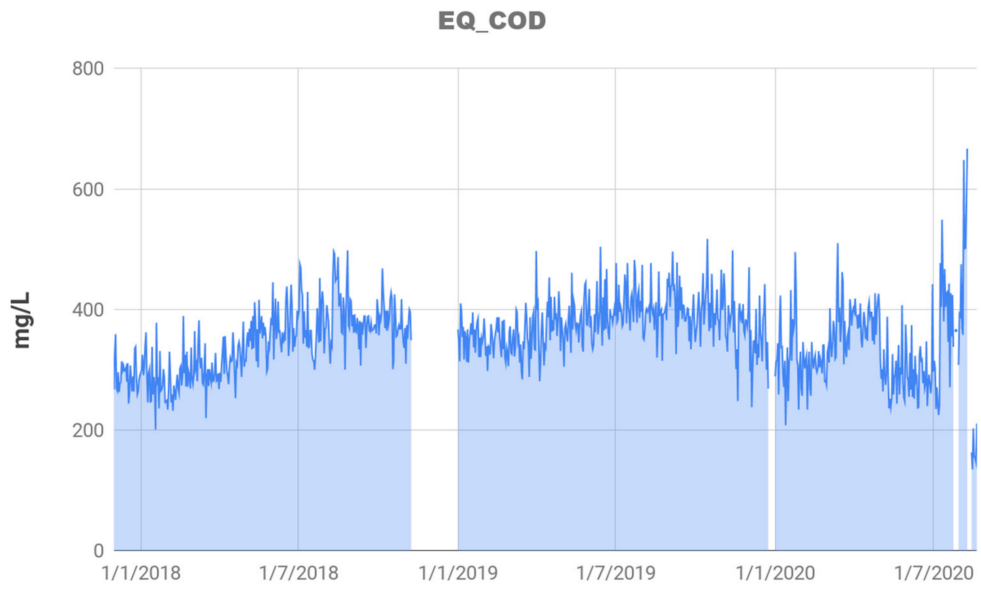

Figure 4. Chemical oxygen demand (COD) behaviour in the equalizer tank.

Within the dataset, the main variables of the process are:

- Flow;
- COD of influent water;
- Suspended solids on influent water (SS);
- Mixed Liquor suspended solids (MLSS);
- Mixed Liquor volatile suspended solids (MLVSS);
- Nitrogen (N);
- pH;
- Dissolved oxygen (DO);
- Food/Microorganism (F/M).

Below, it details the stages related to the biological treatment in which the variables can affect once or more:

- Equalizer (EQ);
- Bioreactor (BIO);
- Bioreactor Pit N (BT_N);
- Bioreactor Pit C (BT_C);
- Clarifier (Clari);
- Oxidation Tank (OxT);
- Discharge Pit (D).

The variables that can be directly manipulated based on experts' knowledge are pH, flow, and DO. These variables affect some other variables, such as MLVSS and COD. For example, the MLVSS could increase or decrease depending on the amount of DO. Thus, both have a directly proportional link. MLVSS is a variable that gives information about microorganism growth [33]. Depending on the microorganism growth, the efficiency of COD reduction will improve or worsen. Expressly, microorganisms break down COD. Hence, constant microorganism growth leads to a constant COD reduction theoretically. In general, the biological process's objective is to reduce the COD at the end of the discharge as much as possible. Thus, to reach the main objective, the optimization algorithms must determine the combination of setpoints for the manipulated variables (pH, flow, and DO).

The final proposed system considers an MLVSS predictive model and two COD predictive models in different stages each. The first stage is the equalizer at the beginning of the biological process and the second one is in the pit at the end of discharge. Figure 5 shows the final system.

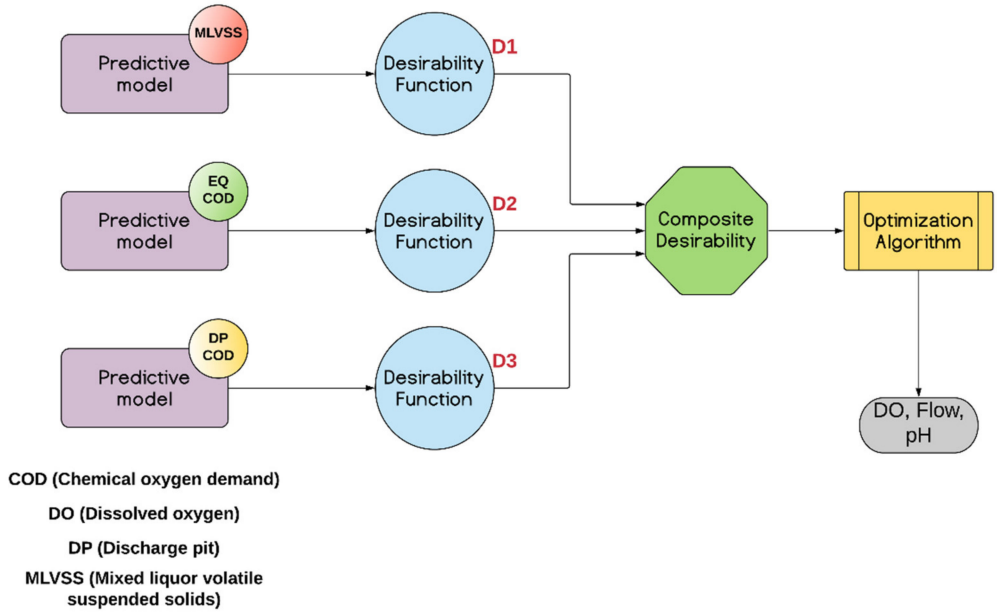

Figure 5. The proposed approach for the prescriptive system of the industrial WWTP case study.

However, this paper's approach focuses on the predictive model of COD in the equalizer, the optimization algorithm, and one variable to prescribe. The first stage of the system starts by using the COD model's predictions in the equalizer. This prediction model uses LSTM-ANN (Long Short-Term Memory Artificial Neural Network) developed with the methodology in Reference [30] to forecast the mean of the variable COD tomorrow. This neural network has 18 variables inputs to its model: MVLSS, SS, nitrogen, DO, F/M, and COD. The six variables are received as input more than once, as shown in Table 2, despite the repetitions that account for different stages of the process on different days.

Table 2. Inputs to the predictive model based on the long short-term memory (LSTM) artificial neural network.

N°	Variable	Delay in Days (t)
1	BT_C_MLVSS	2
2	D_SS	2
3	EQ_N	2
4	Clari_DO	2
5	F/M	2
6	EQ_COD	2
7	BT_C_MLVSS	1
8	D_SS	1
9	EQ_N	1
10	Clari_DO	1
11	F/M	1
12	EQ_COD	1
13	BT_C_MLVSS	0
14	D_SS	0
15	EQ_N	0
16	Clari_DO	0
17	F/M	0
18	EQ_COD	0

The LSTM-ANN architecture consists of one input layer of 18 neurons, two hidden layers with two neurons and 16 neurons, respectively, and one output layer of a neuron. A grid search that analyzes the neurons' distribution adjusts the number of neurons in the hidden layers for obtaining the network's best performance. The network's training

time does not exceed 1.5 min using hardware with the servers' characteristics provided by Google Colab. Due to the network structure, there are 1401 parameters trained with 80% of the previously mentioned dataset of 877 samples.

The correlation of each variable with the entire set of variables is analyzed to adapt the predictive model and build an intelligent system to optimize the process at each biological treatment stage. This is in contrast of which variables coincided with the variables, as recommended by the experts. In the correlation matrix presented in Figure 6, the relationship with each variable is detailed. Mainly, the objective is to analyze the variables inputs to the LSTM strongly correlated with the variables to manipulate. Results lead us to conclude that none of the inputs could be modified directly with a controller in this prediction model because the focus is on looking for variables that could be manipulated before the equalizer stage or in the equalizer itself. As a workaround, an indirect relationship could be found between the nitrogen variable in the equalizer (EQ_N) and the pH that could be modified in the oxidation tank, which is before the equalization stage. This variable is labelled as OxT_pH_PM and gives information about the water's pH in the oxidation tank. Figure 6 points this correlation with the white star. When choosing this variable, it is necessary to study which other variables are affected by the pH change since the pH prescription will change all the variables strongly correlated. Results show no variables have a considerable correlation other than nitrogen. Besides, the pH variable's operating margins in the process are studied to narrow down the search for the optimization algorithm and find the optimal setpoint. Figure 7 shows the probability density function (PDS) of the pH where the operating margin is between 3.0 and 5.0. Hence, it is the range taken as a study.

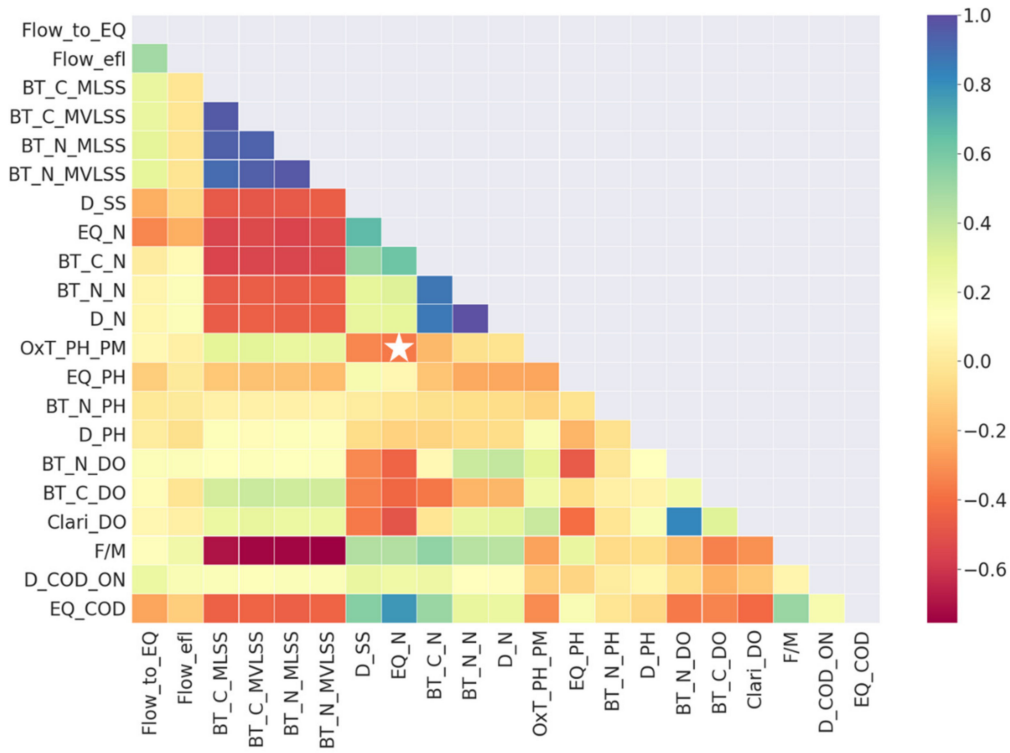

Figure 6. Pearson correlation matrix of variables at the biological treatment process.

A machine learning algorithm is implemented to estimate how the change in the pH variable affected the equalizer's nitrogen variable. Depending on the pH variable's behaviour and the nitrogen variable itself on days past, it could estimate the new EQ_N measurement. A decision tree is the chosen algorithm. This estimation algorithm has, in total, five inputs, which are: EQ_N (t-2), OxT_PH_PM (t-2), EQ_N (t-1), OxT_PH_PM (t-1), and OxT_PH_PM (t). Figure 8 shows the result of this estimation, comparing it with the actual variable. The estimation has a slight shift because, in some days, the algorithm

estimates the nitrogen remains equal to the day before. In this regard, using the Mean Absolute Percentage Error (MAPE) as a comparison metric between the actual and the estimated variables. A performance of the model equal to 7.09% is obtained. Because this model obtains a performance of less than 10% of MAPE, the prescriptive system uses it to estimate the EQ_N.

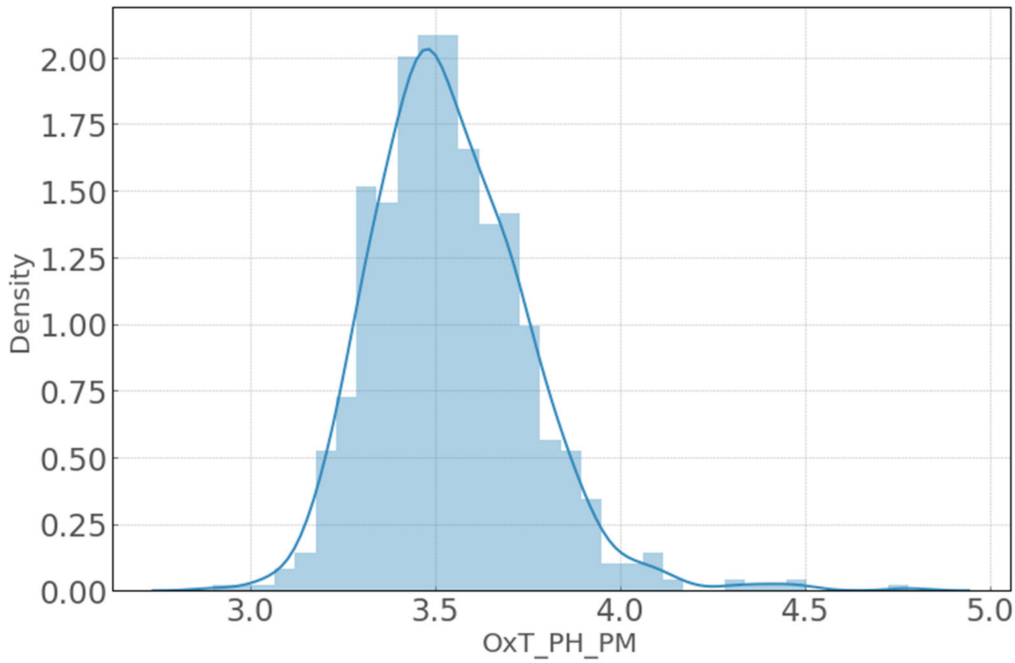

Figure 7. pH probability density function (PDS) of the oxidation tank before the biological treatment.

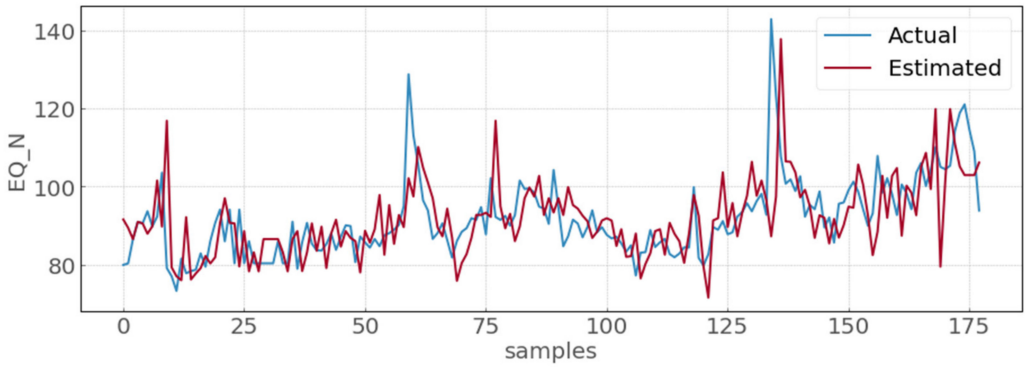

Figure 8. EQ_N estimation using a decision tree algorithm.

Consequently, the LSTM prediction model receives the estimate of the nitrogen by modifying the pH variable. Therefore, it is possible to analyze how pH changes directly influence the nitrogen EQ_N and study how the COD variable changes. The focus is on minimizing or stabilizing the COD variable over time. To pursue this, the optimization algorithm based on genetic algorithms is designed.

The design starts by generating an initial population from a uniform distribution. The initial population has a size of N = 20. The genes are various pH setpoints represented in binary in the range mentioned before (from 3.0 to 5.0). For creating new generations and making the crossover, the system selects 20% of the best chromosomes. In the case of mutations, the population generated as offspring has its genes randomly modified, conditioned by the effect of tossing a coin. If it is head ($x < 0.5$), the mutation will not be applied. Otherwise ($x \geq 0.5$), the genes are altered with a random value such as zero of either one. Finally, the number of generations will depend on three criteria:

(1) Maximum number of generations;
(2) Convergence;
(3) Ideal COD outcome.

The first criterion sets a maximum of 100 generations to save computing resources. Using the second criterion, the system states that if there is no improvement after a specific number of generations (10% of the maximum number of generations), the algorithm would stop and report the best value found. The last criterion states the algorithm would stop if it finds a COD less than or equal to 250 mg/L (value recommended by the experts). In addition, according to the experts' suggestions, small pH values are favoured. Figure 9 shows a general scheme about how the genetic optimization algorithm works.

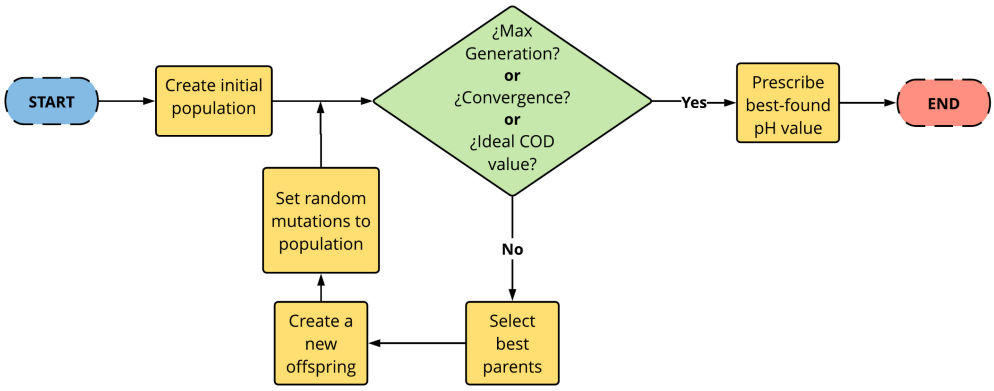

Figure 9. Genetic algorithm scheme for the chemical oxygen demand (COD) value optimization.

4. Results

A forecast is made with the LSTM predictive model day by day, taking 877 actual samples from the process to analyze the best pH setpoint that significantly decreased the COD variable. Figure 10 shows the pH points that are selected day by day based on the genetic algorithms (GA) optimization process. Using these setpoints, the predictive model forecasted how the course of the COD variable would change. Figure 11 compares the forecast with the actual values and the changes with the prescribed values for pH.

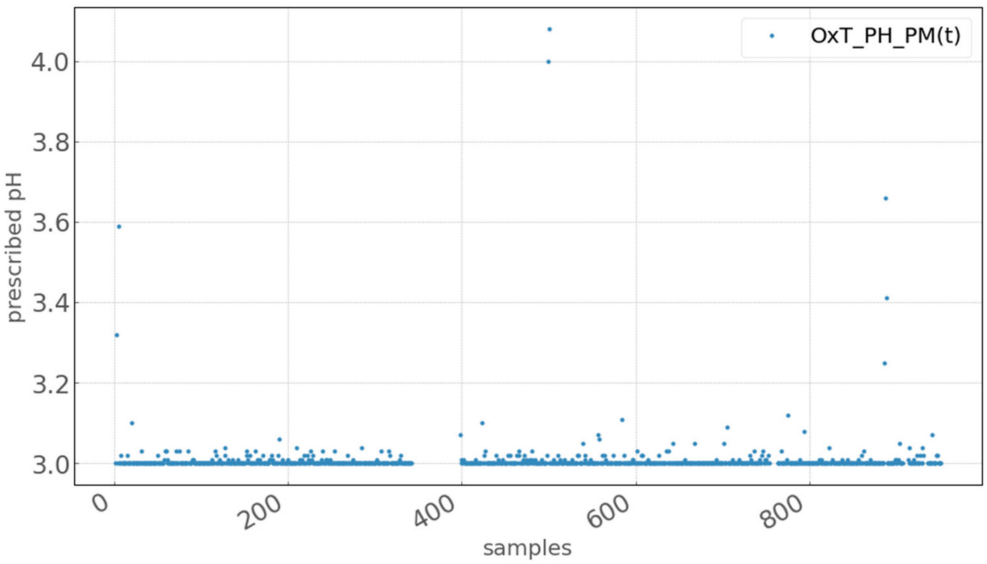

Figure 10. Prescribed pH day by day based on the optimization algorithm.

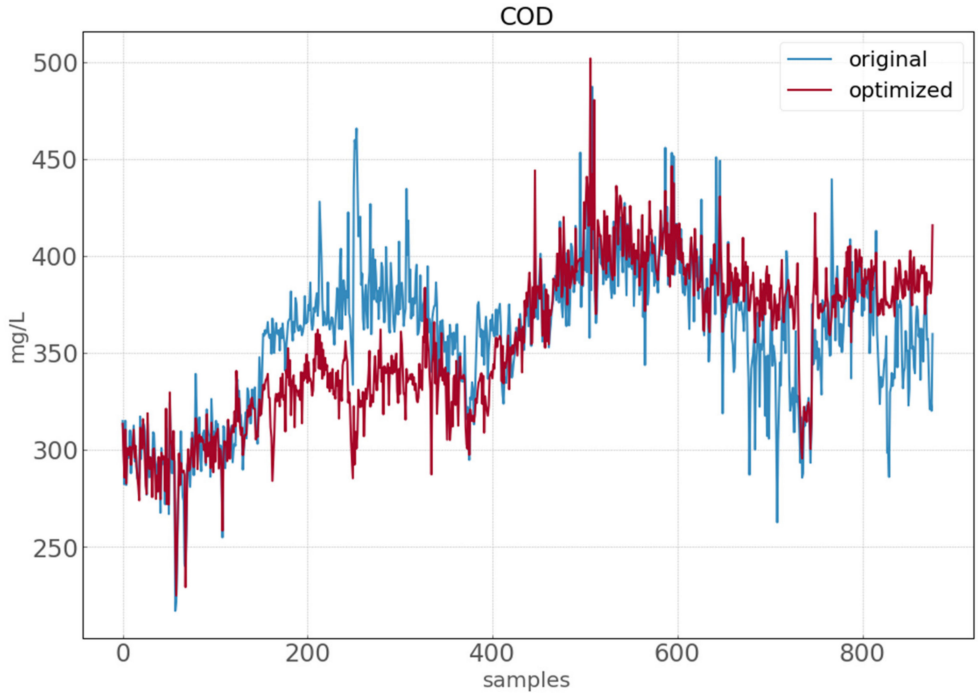

Figure 11. Comparison among original data vs. prescribed data for the COD variable.

For making the comparison of the two signals, it follows the procedures below:

(1) Calculation of mean and standard deviation;
(2) T-student test;
(3) F-Fisher test;
(4) Box-and-whisker plots comparison.

Table 3 shows the mean and standard deviation of each variable. As can be seen, the mean of the prescribed COD is lower than the original one even though the prescription increases measurements' variability. With this, the algorithm is optimizing the equalizer performance by decreasing overall COD values.

Table 3. Mean and standard deviation of the original and optimized time series through prescription.

Statistic	Original Time Series	Optimized Time Series
Mean (mg/L)	358.29	355.21
Standard deviation (mg/L)	38.04	39.87

In addition, to determine if the two-time series are statistically equal, a set of tests is performed to solve two null hypotheses. The first is that the two-time series have the same mean, and the second is if they have the same standard deviation. However, the p-value of each test (T-student test and F-fisher) reject the null hypothesis, so it leads us to conclude that the two-time series are statistically equal. The reason is that the two p-values are 0.108 and 0.811, respectively. Taking an alpha equal to 0.05, p-values that are more significant than alpha lead to rejecting the null hypothesis. Therefore, although the mean and standard deviation are close, it can be stated that, statistically, the two-time series are different.

For quantifying and evaluating the difference between the two signals, the analysis uses the metric in Equation (1). In this equation, x_{or_i} are the original signal's points and x_{op_i} are the prescribed signal's points. The equation result would be a day-to-day comparison

normalized by summing the original signal points to measure how much larger the original signal is than the prescribed signal.

$$\frac{\sum_{i=1}^{N} x_{or_i} - x_{op_i}}{\sum_{i=1}^{N} x_{or_i}} \times 100\%, \quad (1)$$

An important observation to highlight is that the original signal corresponding to COD is 0.86% higher than the optimized one. This result meets the objectives considering that the COD predictive model depends on 18 variables, and the prescriptive system only selects the EQ_N for the respective optimization by modifying pH. As mentioned before, Table 2 shows these 18 variables. From these variables, EQ_N is the variable controlled by OxT_pH_PM to optimize COD for the reasons stated before in the correlation analysis. Besides, box-and-whisker plots analysis supports the claim found using Equation (1). For further explanation, Figure 12 shows two boxes that account for each time series measurements' distribution.

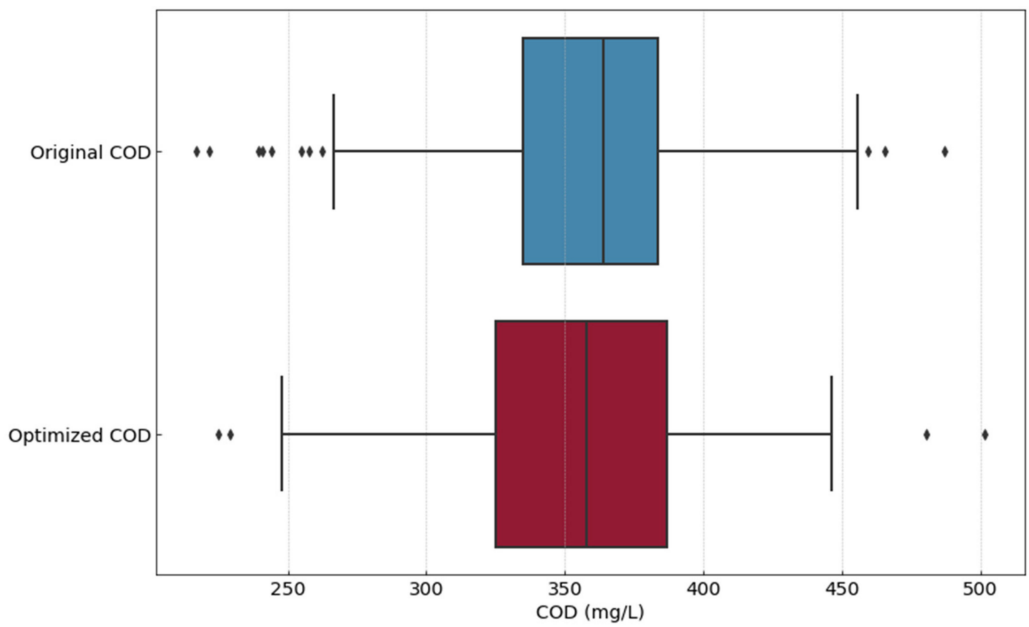

Figure 12. Box-and-whisker plots of the original and optimized time series through prescription.

A comparison between both boxes shows that the optimized COD box is more to the left than the original one, which means the optimized COD distribution is lower than the original COD. However, the interquartile range is more extensive for the red one, but it just reaffirms the increased variability in it, as shown in Table 3 in a standard deviation. Physically, this denotes less organic load in the equalizer tank when the process uses the prescription dictated by the optimizer. Giving solidity to the idea presented by the study of the mean and the metric in Equation (1) supports that the intelligent system can optimize the COD behaviour using the prescriptive strategies.

5. Discussion

This paper carries out the prediction model characterization and adaptation for the decision-making scheme successfully using an LSTM-ANN. Thus, a prescriptive intelligent system is developed based on rules made by expert knowledge, the outcomes provided by the selected prediction model, and a decision tree algorithm to estimate changes of the EQ_N based on the measurement of the OxT_pH_PM to optimize COD in the equalizer tank. Finally, the variables' mean, the calculus made with Equation (1), and the box-and-whisker plots comparison support the prescriptive system evaluation proving a decreased COD value. The optimization is notorious in analyzing the results in Table 3 and Figure 12,

which denotes that the equalizer's COD decreased. Comparing the means in Table 3, the mean for optimized COD is smaller than in the original performance, averaging the 877 samples for each variable. Figure 12 shows how the sample distribution in the optimized variable is further to the left than the original COD, which means the sample distribution in the optimized COD is decreased, according to the diagram analysis.

On the other hand, samples in Figure 11 for the optimized COD are worse than the original one, which are mostly the last samples. This phenomenon is the algorithm's performance in Figure 8 for the last samples that are not that close to the actual expected results. The prescription system will progress by improving the algorithm performance.

Consequently, this paper answers the question stated at the end of the introduction positively. An intelligent computational model can make a prescriptive decision based on available predictive data in an industrial WWTP. Future work will include more variables in the intelligent system that help improve the optimization outcomes. Therefore, the authors will complete the final proposal summarized in Figure 5. Finally, the work developed so far has a high probability of being scaled to other fields. These fields include the optimization in any manufacturing processes, whether of goods or services, electric power generation, or others. In general, these prescription techniques favour any process that can monitor, control, and store the information of the variables that define it. Other researchers can use the same methodology designed in this study.

Author Contributions: Conceptualization, A.M. and R.M. Methodology, C.G.Q.M. Software, L.A. Validation, L.A., C.C., and C.G.Q.M. Formal analysis, L.A., C.C., and C.G.Q.M. Investigation, L.A., C.C., and C.G.Q.M. Resources, D.G. Data curation, L.A. and C.C. Writing—Original draft preparation, L.A. Writing—Review and editing, L.A., C.C., D.G., A.M., R.M., and C.G.Q.M. Visualization, L.A. Supervision, C.G.Q.M. Project administration, D.G. Funding acquisition, D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Colombian Ministry of Science and Technology, MINCIENCIAS, Investment Tax Benefits, Call No. 786.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported by the Universidad del Norte, Barranquilla, Colombia.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding this paper's publication.

References

1. Soltanpoor, R.; Sellis, T. Prescriptive Analytics for Big Data. *Australas. Database Conf.* **2016**, 245–256. [CrossRef]
2. Yan, Z.; Ge, H.; Pan, C.; Mei, L. The Study on Face Detection Strategy Based on Deep Learning Mechanism. *Future Inf. Technol.* **2014**, 391–396. [CrossRef]
3. Han, H.G.; Zhang, H.J.; Liu, Z.; Qiao, J.F. Data-driven decision-making for wastewater treatment process. *Control. Eng. Pract.* **2019**, *96*, 104305. [CrossRef]
4. Hertog, D.D. Bridging the Gap between Predictive and Prescriptive Analytics—New Optimization Methodology Needed. 2011, 1–15. Available online: https://www.semanticscholar.org/paper/Bridging-the-gap-between-predictive-and-methodology-Hertog/98c09b4e4069ee044e71a1bebb5177a43b23cb45 (accessed on 24 January 2021).
5. Krumeich, J.; Werth, D.; Loos, P. Prescriptive Control of Business Processes: New Potentials Through Predictive Analytics of Big Data in the Process Manufacturing Industry. *Bus. Inf. Syst. Eng.* **2015**, *58*, 261–280. [CrossRef]
6. Anjum, M.; Al-Makishah, N.H.; Barakat, M. Wastewater sludge stabilization using pre-treatment methods. *Process. Saf. Environ. Prot.* **2016**, *102*, 615–632. [CrossRef]
7. Tchobanoglous, G.; Schroeder, E.D. *Water Quality: Characteristics, Modeling, Modification*; Addison-Wesley: Massachusetts, MA, USA, 1983.
8. Puri, M.; Solanki, A.; Padawer, T.; Tipparaju, S.M.; Moreno, W.A.; Pathak, Y. Introduction to Artificial Neural Network (ANN) as a Predictive Tool for Drug Design, Discovery, Delivery, and Disposition. *Artif. Neural Netw. Drug Des. Deliv. Dispos.* **2016**, 3–13. [CrossRef]

9. Kumar, B.R.; Vardhan, H.; Govindaraj, M.; Vijay, G. Regression analysis and ANN models to predict rock properties from sound levels produced during drilling. *Int. J. Rock Mech. Min. Sci.* **2013**, *58*, 61–72. [CrossRef]
10. Sharifi, S.S.; Rezaverdinejad, V.; Nourani, V. Estimation of daily global solar radiation using wavelet regression, ANN, GEP and empirical models: A comparative study of selected temperature-based approaches. *J. Atmos. Sol. Terr. Phys.* **2016**, *149*, 131–145. [CrossRef]
11. Hang, R.; Liu, Q.; Hong, D.; Ghamisi, P. Cascaded Recurrent Neural Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote. Sens.* **2019**, *57*, 5384–5394. [CrossRef]
12. Hong, D.; Gao, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J. Graph Convolutional Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote. Sens.* **2020**, 1–13. [CrossRef]
13. Kotu, V.; Deshpande, B. Time Series Forecasting. *Predict. Anal. Data Min.* **2015**, 305–327. [CrossRef]
14. Butka, P.; Bednar, P.; Ivancakova, J. Methodologies for Knowledge Discovery Processes in Context of AstroGeoInformatics. In *Knowledge Discovery in Big Data from Astronomy and Earth Observation*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 1–20.
15. Thompson, J.D. Statistical Alignment Approaches. *Stat. Bioinform.* 2016, pp. 43–51. Available online: http://alexeidrummond.org/assets/publications/2005-lunter-statistical.pdf (accessed on 12 January 2021).
16. Larson, D.; Chang, V. A review and future direction of agile, business intelligence, analytics and data science. *Int. J. Inf. Manag.* **2016**, *36*, 700–710. [CrossRef]
17. Lepenioti, K.; Bousdekis, A.; Apostolou, D.; Mentzas, G. Prescriptive analytics: Literature review and research challenges. *Int. J. Inf. Manag.* **2020**, *50*, 57–70. [CrossRef]
18. Revathy, P.; Mukesh, R.; Varadarajan, V.; Kommers, P.; Piuri, V.; Subramaniyaswamy, V. HadoopSec 2.0: Prescriptive analytics-based multi-model sensitivity-aware constraints centric block placement strategy for Hadoop. *J. Intell. Fuzzy Syst.* **2020**, *39*, 8477–8486. [CrossRef]
19. Bertsimas, D.; Van Parys, B. Bootstrap Robust Prescriptive Analytics. *arXiv* **2017**, arXiv:1711.09974.
20. Dey, S.; Gupta, N.; Pathak, S.; Kela, D.H.; Datta, S. Data-Driven Design Optimization for Industrial Products. *Manag. Ind. Eng.* **2018**, 253–267. [CrossRef]
21. Baur, A.; Klein, R.; Steinhardt, C. Model-based decision support for optimal brochure pricing: Applying advanced analytics in the tour operating industry. *OR Spectr.* **2013**, *36*, 557–584. [CrossRef]
22. Berk, L.; Bertsimas, D.; Weinstein, A.M.; Yan, J. Prescriptive analytics for human resource planning in the professional services industry. *Eur. J. Oper. Res.* **2019**, *272*, 636–641. [CrossRef]
23. Ghoniem, A.; Ali, A.I.; Salem, M.A.; Khallouli, W. Prescriptive analytics for FIFA World Cup lodging capacity planning. *J. Oper. Res. Soc.* **2017**, *68*, 1183–1194. [CrossRef]
24. Goyal, A.; Aprilia, E.; Janssen, G.; Kim, Y.; Kumar, T.; Mueller, R.; Phan, D.; Raman, A.; Schuddebeurs, J.; Xiong, J.; et al. Asset health management using predictive and prescriptive analytics for the electric power grid. *IBM J. Res. Dev.* **2016**, *60*, 4:1–4:14. [CrossRef]
25. Giurgiu, I.; Wiesmann, D.; Bogojeska, J.; Lanyi, D.; Stark, G.; Wallace, R.B.; Pereira, M.M.; Hidalgo, A.A. On the adoption and impact of predictive analytics for server incident reduction. *IBM J. Res. Dev.* **2017**, *61*, 9. [CrossRef]
26. Wang, C.-H.; Cheng, H.-Y.; Deng, Y.-T. Using Bayesian belief network and time-series model to conduct prescriptive and predictive analytics for computer industries. *Comput. Ind. Eng.* **2018**, *115*, 486–494. [CrossRef]
27. Jank, M.-H.; Dölle, C.; Schuh, G. Product Portfolio Design Using Prescriptive Analytics. *Adv. Prod. Res.* **2018**, 584–593. [CrossRef]
28. Du, F.; Plaisant, C.; Spring, N.; Shneiderman, B. EventAction: Visual analytics for temporal event sequence recommendation. In Proceedings of the 2016 IEEE Conference on Visual Analytics Science and Technology (VAST), Baltimore, MD, USA, 23–28 October 2016.
29. Lewis, H.W.; Iii, H.W.L. Classical Fuzzy Control Design and Implementation. *Found. Fuzzy Control* **1997**, 95–137. [CrossRef]
30. Arismendy, L.; Cárdenas, C.; Gómez, D.; Maturana, A.; Mejía, R.; Quintero, M.C.G. Intelligent System for the Predictive Analysis of an Industrial Wastewater Treatment Process. *Sustain. J. Rec.* **2020**, *12*, 6348. [CrossRef]
31. Hu, Z.; Grasso, D. Water Analysis | Chemical Oxygen Demand. *Encycl. Anal. Sci.* 2005, pp. 325–330. Available online: https://www.sciencedirect.com/topics/earth-and-planetary-sciences/chemical-oxygen-demand (accessed on 12 January 2021).
32. Inc, W.C. Methods for Treating Wastewaters from Industry. *Ind. Waste Treat. Handb.* 2006, pp. 149–334. Available online: https://www.iwapublishing.com/news/industrial-wastewater-treatment (accessed on 12 January 2021).
33. Deowan, S.; Bouhadjar, S.; Hoinkis, J. Membrane bioreactors for water treatment. *Adv. Membr. Technol. Water Treat.* **2015**, 155–184. [CrossRef]