environments

MDPI

Article

Statistical Characterization of Full-Scale Thermophilic Biological Systems to Inform Process Optimization

Maria Cristina Collivignarelli [1,2], Stefano Bellazzi [1,*], Francesca Maria Caccamo [1], Marco Sordi [3], Barbara Crotti [3], Alessandro Abbà [4] and Marco Baldi [5,*]

1 Department of Civil Engineering and Architecture, University of Pavia, Via Ferrata 3, 27100 Pavia, Italy; mcristina.collivignarelli@unipv.it (M.C.C.); francescamaria.caccamo01@universitadipavia.it (F.M.C.)
2 Interdepartmental Centre for Water Research, University of Pavia, 27100 Pavia, Italy
3 ASMia S.R.L., Mortara, Via Tiziano Vecellio 540, 27036 Mortara, Italy; m.sordi@asmortara.eu (M.S.); b.crotti@asmortara.eu (B.C.)
4 Department of Civil, Environmental, Architectural Engineering and Mathematics, University of Brescia, Via Branze 43, 25123 Brescia, Italy; alessandro.abba@unibs.it
5 Department of Chemistry, University of Pavia, Via Puccini 66, 27011 Belgioioso, Italy
* Correspondence: stefano.bellazzi01@universitadipavia.it (S.B.); marco.carlo.baldi@gmail.com (M.B.); Tel.: +39-0382-985314 (S.B.)

Abstract: This paper focuses on using a novel approach to assess the statistical variability of management data from an aerobic thermophilic biological plant (AWTP) utilizing a fluidized bed biological reactor. A proper statistical characterization of full-scale thermophilic biological systems, in fact, may inform process optimization in the light of a future automation of treatment plants. We present a case study that spans the period from 2018 to 2023 and encompasses various high-strength aqueous waste (AW) in continuous mode. Key aspects of the proposed analytical approach include: (i) utilizing advanced descriptive statistics, such as violin graphs, to depict the variability of monitored parameters over five years; (ii) conducting correlation analyses (Spearman and Pearson correlation matrices) specifically focusing on nitrogenous forms within the AW; (iii) applying multivariate statistical analysis to assess the correlation between pollutants released and the plant's energy and oxygen consumption; and (iv) reconstructing parameter trends by considering periodic and random components, thus enhancing the understanding of the system's behavior over time. The findings presented in this paper offer valuable insights into the performance and optimization of AWTPs, potentially leading to a proper planning of the loads and consequent feeding of the plants. If properly enacted, our approach may provide a significant contribution to the field of aqueous waste management.

Keywords: environmental engineering; WWTPs; thermophilic; aqueous waste

check for updates

Citation: Collivignarelli, M.C.; Bellazzi, S.; Caccamo, F.M.; Sordi, M.; Crotti, B.; Abbà, A.; Baldi, M. Statistical Characterization of Full-Scale Thermophilic Biological Systems to Inform Process Optimization. *Environments* **2024**, *11*, 36. https://doi.org/10.3390/environments11020036

Academic Editor: Naresh Singhal

Received: 1 December 2023
Revised: 26 January 2024
Accepted: 12 February 2024
Published: 17 February 2024

1. Introduction

The treatment of aqueous waste (AW) represents a critical issue due to the high variability of their physical-chemical characteristics and the possible presence of emerging conservative micro-pollutants, such as perfluoroalkyl substances [1], absorbable organic halogens (AOX) and clorides. Conventional biological activated sludge (CAS) processes are not able to achieve a sufficient performance to remove some types of conservative pollutants; in some cases the biota of the plant may even need to be purged since it is no longer active [2,3].

In today's landscape, numerous additional treatments are available to increase the acceptance of conventional biological processes, reducing the effects of acute or chronic toxicity due to the presence of highly refractory organic compounds. The treatments proposed are usually of a physico-chemical nature and therefore often involve the use of various chemical reagents [4,5].

Despite their widespread use, physical-chemical treatments present several drawbacks, such as high operating costs linked to chemical reagents, the need for highly specialized personnel in their management, and finally the high carbon dioxide emission factors for production processes and for the reactions of the reagents themselves [6,7]. For these reasons, biological processes are preferable given their lower environmental impact. They have these advantages with respect to physical-chemical treatments: easy of management, lower costs, and lower CO_2 emission factors [8].

Among biological treatments, the scientific community agrees that thermophilic ones can have high removal yields even in the case of recalcitrant pollutants [9,10]. Considering to the high performance and bio-resistance to some recalcitrant pollutants, the scientific community has identified thermophilic biomasses as those suitable for the treatment of aqueous waste [11]. The main issues with thermophilic treatment are poor biomass settling and the potential absence or inhibition of nitrification. These challenges can be addressed by using membranes for better solid separation and implementing post-treatment processes to encourage nitrification under more favorable conditions. Despite these drawbacks, thermophilic treatment offers advantages such as faster sludge digestion and pathogen destruction, requiring careful management for effective waste treatment [12].

Mesophilic bacteria are used in the conventional wastewater treatment plants (WWTPs); their working range is between 15 and 30 °C. Thermophilic biota are usually exploited for the treatment of AW; they provide optimal working conditions at temperatures between 50–60 °C [13]. For example, in previous works, a thermophilic biological process combined with a membrane ultrafiltration system (TMR—Thermophilic Membrane Reactor) has been used for both the treatment of aqueous wastes and the minimization of biological sludge produced by the WWTPs [14,15]. The great elasticity of this treatment makes it suitable to be applied to different substrates in terms of both chemical and rheological characteristics [16]. This technology has demonstrated resistance to some conservative parameters and micropollutants and it has been applied in two full-scale plants [17]. As a matter of fact, TMR combines the robustness of physico-chemical treatments with the advantages of biological ones, thus allowing a production of equivalent CO_2 lower than standard physico-chemical treatments [18].

So far, the statistical-based characterization of aerobic thermophilic biological plant (AWTP) has received relatively poor attention in the literature. Aguado et al. [19] applied a multivariate statistical approach to derive alerts conditions in the management of the plant. The same strategy has been presented in Liu et al. [20], after variable reduction based on the principal component analysis (PCA). Wang et al. [21] reported a statistical study based on the dynamic simulation to achieve model predictive control. O'Brien and Teather [22] developed a predictive model successfully applied to estimate the pollutant concentrations in a biological system (activated sludge).

In this paper an approach to evaluate the statistical variability of management data of an AWTP at full-scale is proposed. The main novelty stands in the increase of the number of environmental variables that we consider characterizing a newly developed plant, based on a thermophilic fluidized bed biological reactor. The approach is designed to inform process optimization in the light of a future automation of treatment plants, by potentially leading to a proper planning of the loads and consequent feeding of the plants.

The main points of interest of this analysis are:

- The application of advanced descriptive statistics (violin graphs) to describe the variability of each parameter monitored over the five years; this allowed reporting univariate probability distributions of the values observed based on the density profile kernel estimation;
- Correlation analysis (Spearman and Pearson correlation matrices) of the nitrogenous forms;
- The application of multivariate statistical analysis to evaluate the correlation between the different pollutants released and the energy and oxygen consumption of the plant to provide a useful tool to optimize TMR;

- The reconstruction of the trends of the parameters studied taking into account periodic and random components [23].
- In this work, we present a case-study that analyzes the data coming from monitoring a full-scale aerobic thermophilic biological plant in the period 2018 (last trimester)—2023 (first trimester); along this period the plant treated different high-strength AW in continuous mode.

2. Materials and Methods

2.1. TMR Configuration

The monitored TMR plant was usually fed with a wide range of AW such as landfill leachate, wastewater with high contents of organic content, salts, solvents, pharmaceutical compounds and with very different pH values (strong acid or alkaline wastewater). The water line is composed by both physico-chemical and biological treatments (Figure 1). This article focuses only on the thermophilic system composed of a thermophilic biological reactor (TBR) followed by ultrafiltration (UF) membranes placed upstream of a traditional CAS for municipal sewerage treatment.

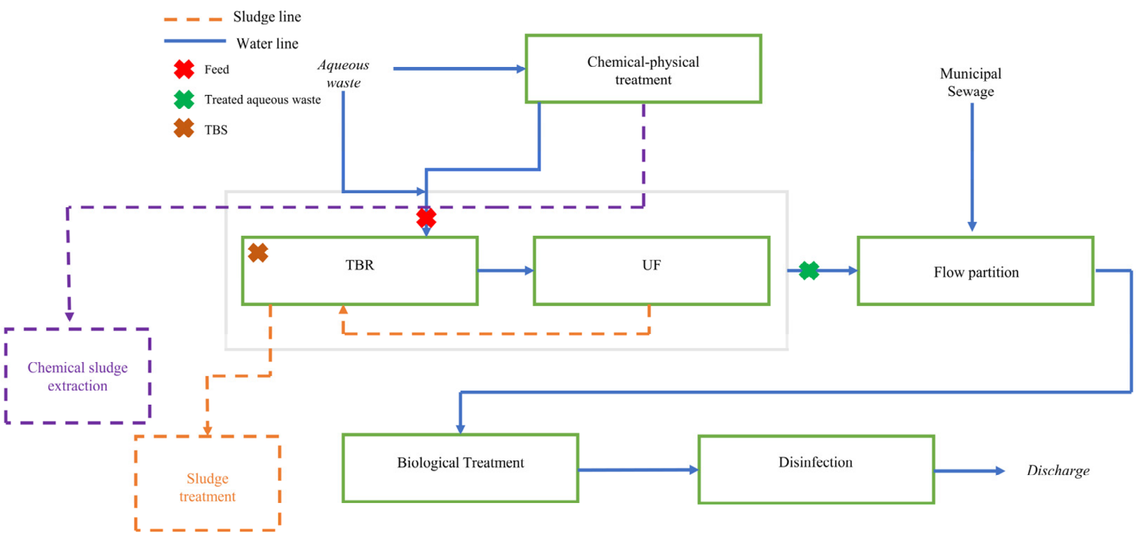

Figure 1. Scheme of the TMR and sampling points (X). TBR: Thermophilic biological reactor; UF: Ultrafiltration.

As seen from the plant scheme, AW have two treatment options based on their characteristics. Waste is pre-treated in batches in two physico-chemical compartments to safeguard the thermophilic biomass (T = 48 °C). The TBR has a volume of 1400 m^3 and is characterized by an average inlet flow rate Qin = 170 m^3d^{-1}, with a hydraulic retention time HRT of almost 8.0 d, as reported by Collivignarelli et al. the hydraulic retention time was studied in an experimental work using a pilot plant [9]. To improve the performance of the sludge and obtain a more compact biomass, a dissolved oxygen concentration of 1.5–2 mg L^{-1} is guaranteed in the reactor by injecting pure oxygen. Inside the biological tank, different areas have different concentrations of oxygen; this management strategy allows for both the aerobic removal of organic substance in the zone of high oxygen concentration (2 mgO_2 L^{-1}) and denitrification in the anoxic zone (<1 mgO_2 L^{-1}). As mentioned previously, the TBR is coupled to an ultrafiltration system due to the poor settleability of the thermophilic biomass probably caused by its carbohydrate content [24]. The platform is therefore equipped with two parallel UF units. The UF unit is composed of 6 vessels with 99 ceramic membranes (cut-off 300 kDa) with an operating pressure between 3 and 5 bar. As shown in Figure 1, the UF retentate is recirculated into the biological reservoir thus ensuring low excess sludge extraction.

2.2. Monitoring Plan

Monitoring activity focused on TMR from 2018 to 2023, sampling the inlet substrate and ultrafiltration permeate (Figure 1) once a week for five years. Analyzes of COD, nitrogen forms (TN–total nitrogen, N-NO_x–sum of nitrite nitrogen and nitrate nitrogen, N-NH_4^+–ammonia, and N_{org}–organic nitrogen), total phosphorus (TP), non-ionic surfactants (TAS) and anionic surfactants (MBAS) were measured according to standard methods [25]. The flow rates ingoing and outgoing the plant, the consumption of pure oxygen necessary for the degradation of the organic matter (O_2), and the energy consumption (kWh) were weekly monitored.

2.3. Data Processing

It was decided to evaluate the statistical variability through violin graphs for all of the input parameters of the plant. Violin graphs are an evolution of box plots, reporting the density profile of the observed values for a univariate distribution in the form of a kernel density plot. The kernel density plot [26] of the data is reported symmetrically by both sides of the distribution and this explains the typical shape of the graphs, which gives them their name [27].

To determine the removal efficiency of COD, nitrogen and total phosphorus, mass balances were performed weekly based on fed and extracted loads.

To obtain reliable modeling, the data were normalized by the maximum value of the series to obtain a series of homogeneous values between 0 and 1 [17]. Pearson and Spearman rank correlations were used to calculate the binary correlations between the nitrogenous forms monitored within the plant.

To evaluate energy and oxygen consumptions during the aerobic process, all input parameters were studied using a multiple linear regression analysis.

A stepwise regression method has been performed with the MATLAB function stepwisefit [28]; this method selects the variables that are most strongly correlated jointly with the reagent (pure oxygen input) and energy consumption of the system. Given the *i*-th parameter xi, it is described with the model:

$$x_i = \alpha_i + \sum_{j \neq i} \beta_{ij} x_j \tag{1}$$

If m parameters are measured, the stepwise procedure allows selecting the $m_i \leq m - 1$ parameters that are more statistically significant to predict each x_i.

To evaluate the presence of a random component and a periodic component in the parameters loaded to the aqueous waste treatment platform an algorithm called STL ("Seasonal and Trend decomposition using Loess") has been applied [29]. STL is a technique used for time series analysis and forecasting [30]. The objective was therefore to decompose a time series into easy-to-interpret components such as trends, seasonality, and noise. We will look for monthly seasonal components based on the variability of the influential loads and the trend in the five years of interest.

3. Results and Discussion

3.1. Monitoring of the Biological Thermophilic System-Violin Plot

To evaluate the distribution of the parameters entering the TBR industrial reactor, it was decided to create violin plot that show a boxplot divided for each parameter and the Kernel distribution reported symmetrically within it.

As can be seen in Figure 2, the association between the two numerical calculations shown (boxplot and Kernel distribution) effectively schematizes the great variability of the 8 pollutants in going to TMR. As well depicted in the trend of the distributions, the variability in terms of input parameters increased from 2018 to 2023 for polluting parameters such as N-NO_3^-, TAS and N_{org}. However, by observing the graph relating to TP it is possible to deduce the loading mode of the reactor with the presence of some peaks uniformly distributed over the various years as shown by the relative violin graphs; in the last quarter analyzed this aspect is less evident. However, with respect to the measurement of COD, an effort can be seen by the plant manager to homogenize the organic load entering the biological reactor.

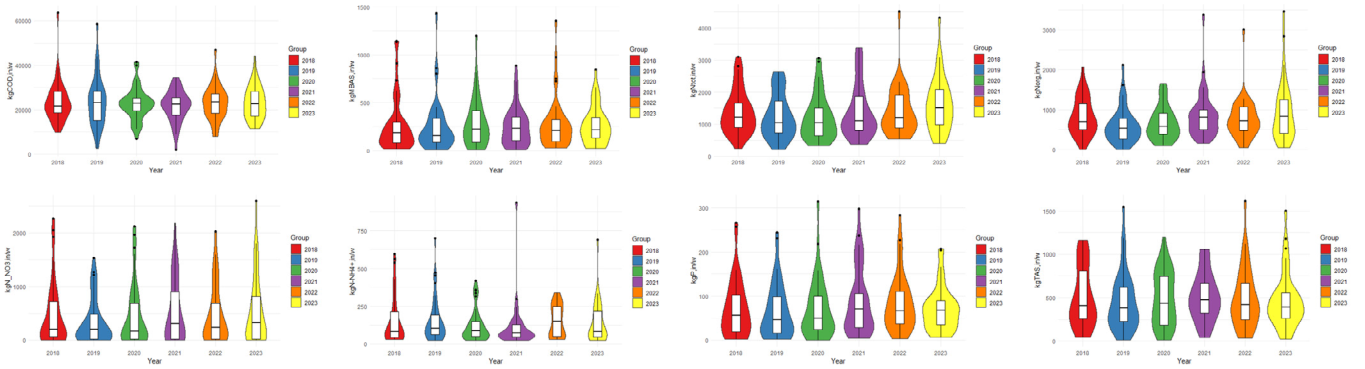

Figure 2. Violin Plot of the loads [kg/week] fed to the TMR relating to the following pollutants: COD, Ntot, N-NO_3^-, N-NH_4^+, N_{org}, TP, TAS, MBAS.

In Figure 3 it is possible to appreciate through the violin plots relating to the outputs from TMR the treatment effectiveness of this process which combines thermophilic biodegradation with an ultrafiltration treatment [31]. By observing the value in kg/week of each parameter output from the TMR it is possible to note the high removal yields shown subsequently in Table 1. The only parameter which has an output load value higher than the input is ammonia due of the strong ammonification process of organic nitrogen in thermophilic conditions [17]. The trend of the distribution of outgoing loads divided into the 6 years shown is instead useful for showing a second strong point of this technology [32]. The input loads presented a large variability at the input as demonstrated by the distributions in Figure 2, the outputs instead present a very low variability value, thus underlining on the one hand the great elasticity and resilience of the thermophilic system and on the other the very good applicability of this system placed in series with a CAS which requires homogeneity of the polluting parameters at input [33].

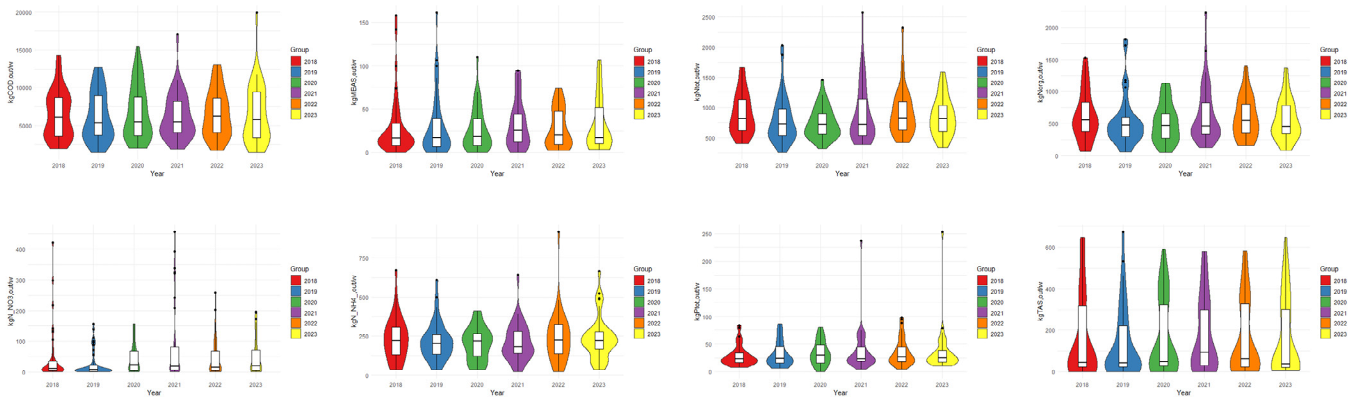

Figure 3. Violin Plot of the loads in output to the TBR relating to the parameters COD, Ntot, N-NO_x, N-NH_4^+, Norg, Ptot, TAS, MBAS [kg/week].

Table 1 shows the pollutant loads introduced and the performance of the thermophilic biological process. This analyzed biomass can treat high COD loads [13], the performance slightly dropped after 2019, reaching an average value over the 5 years of 71.8%. Nitrites and nitrates represent a particularly important parameter for the plant which aims to treat aqueous waste [34], the performance value obtained increased after an initial acclimatization phase reaching a value of 81.3%. The removal yield relating to ammonia is negative. This is due to the fact that, as already known by the scientific community, under thermophilic conditions a strong process of ammonification of organic nitrogen occurs. As already described by observing the graphs of the distributions at the entrance to the biological reactor, the TP presents a high variability value, despite this an average value of 55.3% was obtained over the five years, the accumulation of phosphorus in the sludge in inorganic form occurs mainly thanks to the formation of hydroxyapatite in pure oxygen aeration conditions [35]. The capacity of TMR therefore allows it to accumulate phosphorus within the sludge in the solid phase, facilitating possible recovery in agriculture depending on their content of persistent pollutants such as heavy-metals.

The effect on the removal of TAS and MBAS was also evaluated. The analyzed process allowed obtaining removal yields of 77.4% and 84.0%, respectively, demonstrating the good biodegradability of these types of surfactants in thermophilic conditions. Yields related to MBAS are generally higher due to their structure being more easily attacked by microorganisms [36]. The removal can be attributed to thermophilic biological degradation in the TAR reactor and not to filtration on the subsequent UF membrane, typical molecular weights of surfactants ranged from 0.2–0.4 kDa, depending on their molecular structure, and that the molecular weight limit of the UF membrane used in this study was 30 kDa, most of the surfactants (TAS and MBAS) could pass the pores of the UF membranes [37].

Table 1. Fed pollutants, loads and performance and other operational parameters of thermophilic system; *n* stands for the amount of data. The values are report plus or minus the standard deviation. The parameters read as follows: Chemical Oxygen Demand (COD), Nitric and Nitrous Nitrogen (N-NOX), Ammoniacal Nitrogen (N-NH_4^+), Organic Nitrogen (Norg), Total Nitrogen (TN), Total Phosphorus (TP), Non-ionic surfactants (TAS), Anionic surfactants (MBAS).

Parameter	2018 [*n* = 20]	2019 [*n* = 52]	2020 [*n* = 48]	2021 [*n* = 52]	2022 [*n* = 52]	2023 [*n* = 13]	2018–2023 [*n* = 237]
COD IN [kg week^{-1}] μ COD [%]	20,842 ± 3332 78.8 ± 4.3	19,174 ± 1756 80.9 ± 2.3	20,038 ± 1671 73.4 ± 2.9	25,735 ± 2508 66.2 ± 3.0	26,502 ± 2531 63.1 ± 3.1	25,484 ± 2971 74.1 ± 1.9	22,928 ± 1061 71.8
N-NO_X IN [kg week^{-1}] μ N-NO_X [%]	138.8 ± 72.8 73.2 ± 10.3	235.0 ± 64.3 79.4 ± 6.9	277.9 ± 141.7 76.6 ± 9.5	849.6 ± 157.6 89.1 ± 3.4	575.4 ± 187.7 81.0 ± 8.6	789.7 ± 330.8 79.3 ± 7.0	469 ± 73 81.3
N-NH_4^+ IN [kg week^{-1}] μ N-NH_4^+	208.3 ± 53.7 −51.3 ± 8.3	141.9 ± 34.6 −58.5 ± 7.3	111.7 ± 36.8 −66.3 ± 6.7	123.6 ± 47.8 −51.5 ± 9.3	132.9 ± 24.7 −52.5 ± 8.1	260.7 ± 63.7 −16.7 ± 3.2	140 ± 17.1 −55.8
Norg IN [kg week^{-1}] μ Norg [%]	688.5 ± 190.9 43.2 ± 13.6	564.5 ± 88.4 54.0 ± 5.9	749.6 ± 160.1 39.8 ± 8.6	894.8 ± 137.0 30.9 ± 6.7	989.3 ± 182.0 32.8 ± 5.7	989.3 ± 182.0 29.0 ± 4.4	776 ± 69 40.5
TN IN [kg week^{-1}] μ Ntot [%]	1020.1 ± 210.7 24.7 ± 10.7	950.8 ± 127.5 34.4.6 ± 6.2	1166.3 ± 222.9 42.7 ± 8.2	1834.6 ± 207.2 41.2 ± 6.0	1662.8 ± 205.0 39.1 ± 5.1	1624.5 ± 250.4 23.0 ± 9.2	1376 ± 96 37.5
TP IN [kg week^{-1}] μ Ptot [%]	62.1 ± 27.9 61.9 ± 12.1	103.2 ± 20.9 65.2 ± 6.7	65.7 ± 18.2 58.8 ± 8.1	72.9 ± 14.2 49.9 ± 8.2	76.9 ± 15.2 45.9 ± 6.8	17.5 ± 14.9 22.9 ± 9.2	75 ± 8 55.3
TAS IN [kg w^{-1}] μ TAS [%]	246.3 ± 66.9 85.2 ± 5.2	475.7 ± 81.3 91.2 ± 3.2	402.9 ± 60.5 84.6 ± 6.1	605.6 ± 82.7 69.4 ± 5.9	639.5 ± 100.5 54.9 ± 6.4	101.1 ± 40.3 40.8 ± 3.9	486 ± 41 77.4
MBAS IN [kg week^{-1}] μ MBAS [%]	115.0 ± 63.3 82.1 ± 9.3	252.0 ± 75.4 91.2 ± 3.0	311.9 ± 67.9 89.2 ± 3.9	325.0 ± 74.8 76.6 ± 6.0	279.3 ± 52.1 80.6 ± 4.4	52.8 ± 24.8 70.9 ± 9.8	267 ± 32 84.0

3.2. Data Analysis

The Pearson moment correlation coefficient and the Spearman rank correlation coefficient were calculated to evaluate a possible correlation between the monitored nitrogenous forms in thermophilic biological process input (Table 2).

This analysis allows us to have an indication of the composition of the total nitrogen loaded and the possible relationship between the various forms of nitrogen. The analysis of Pearson's correlation index highlighted a positive correlation between Norg introduced into the plant and TN (0.57). The correlation between the value of nitric and nitrous nitrogen and the total nitrogen entering the plant was 0.66 while the correlation between ammoniacal nitrogen and the monitored total was only 0.2. From the Person index it can therefore be deduced that the most significant amount of nitrogen in terms of total nitrogen was that attributed to nitric and nitrous nitrogen. It is also possible to note some negative correlations between the monitored nitrogenous forms (NH4-NO_x) and organic nitrogen, suggesting a treatment upstream of the biological process and a possible interference in the measurement method [38].

Table 2. Pearson correlation matrix.

	Norg IN	TN IN	N-NH_4^+ IN	N-NO_x IN
Norg IN	1.00	\	\	\
TN IN	0.57	1.00	\	\
N-NH_4^+ IN	−0.06	0.20	1.00	\
N-NO_x IN	−0.17	0.66	0.10	1.00

The Spearman correlation matrix shows very similar results (Table 3). A strong positive correlation can be observed between nitrates input into the system and total nitrogen input. This correlation can be explained by considering that nitrates constitute the largest share of the forms of nitrogen treated by the thermophilic biological system.

Table 3. Spearman correlation matrix.

	Norg IN	TN IN	N-NH_4^+ IN	N-NO_x IN
Norg IN	1.00	\	\	\
TN IN	0.45	1.00	\	\
N-NH_4^+ IN	−0.12	0.25	1.00	\
N-NO_x IN	−0.19	0.69	0.24	1.00

From the analysis of the two correlation matrices, a positive correlation emerges between all of the non-organic nitrogenous forms entering the biological reactor with the total nitrogen value. This analysis suggests the possibility of estimating the composition of the nitrogenous forms entering the plant based on the measurement of total nitrogen, usually simpler than that relating to nitrates.

According to Correlation Matrixes, Pearson (P) and Spearman (S) is interesting since if you have S > P, it means you have a monotonic but non-linear correlation [39]. The results could help in the future in the correct modeling of the observed parameters, providing information on the type of correlation between the data, whether linear or monotonic.

3.3. Multivariate Linear Regressions

Table 4 shows the matrix of the coefficients of the best multiple linear regressions able to describe the possible relationship between the parameters of the substrates entering the biological reactor and the chemical, pure oxygen, and energy consumption. The regressions are obtained by means of a stepwise search, as implemented in the MATLAB Statistics and Machine Learning toolbox [40].

As shown in Table 4, it was found that the value of oxygen needed by the system is linearly correlated with the input of organic nitrogen (B = 0.1382), i.e., around 7 kg of organic nitrogen correspond to 1 kg of oxygen consumption. As reported by Table 1, the analysis shows that the oxidation process of organic nitrogen corresponds to the highest rate of oxygen consumption, thus leading to an increase of the ammoniacal nitrogen. This regression has a low Pval value (0.0175) and a root mean square deviation (Rmse) value of 0.1226, i.e., well below the maximum value. This demonstrates a good applicability of the model as well as for forecasting purposes. The same approach was used to study the value of electricity consumption. The parameters selected by the model were total nitrogen and anionic surfactant. In this case the Pval value was very low (around 10^{-7}) showing excellent statistical significance of the model, with a Rmse value of 0.1042.

Table 4. Coefficients of multivariate linear regressions. N corresponds to the parameter selected by the model, being 1 equal to yes and 0 to no; Bare the linear regression coefficients; Pvals are the P-values, i.e., the probability of obtaining a result equal to or more extreme than the observed one, assuming that the null hypothesis is true (Pvals are computed using the F-statistic); finally Root Mean Square Error (Rmse) evaluates the performance of a forecasting model by comparing the prediction to the observed or actual data.

	COD, in [kg/week]	**N_{org}, in [kg/week]**	**TN, in [kg/week]**	**N−NH4, in [kg/week]**	**N−NO3, in [kg/week]**	**TP, in [kg/week]**	**TAS, in [kg/week]**	**MBAS, in [kg/week]**	**KgO_2 in [kgO_2/week]**
N	0	1	0	0	0	0	0	0	
B	\	0.1382	\	\	\	\	\	\	
Pval	0.0175								
Rmse	0.1226								
	COD, in [kg/week]	**N−Org, in [kg/week]**	**N−tot, in [kg/week]**	**N−NH4, in [kg/week]**	**N−NO3, in [kg/week]**	**P, in [kg/week]**	**TAS, in [kg/week]**	**MBAS, in [kg/week]**	**kWh in [kWh/week]**
N	0	0	1	0	0	0	0	1	
B	\	\	0.1297	\	\	\	\	0.2643	
Pval	71.44×10^{-7}								
Rmse	0.1042								

This type of analysis allows the AW utilities manager to be able to forecast the electrical and oxygen consumption values of the system based on the input parameters.

The proposal involves utilizing future linear regression analysis to establish a connection between oxygen consumption and removal efficiencies. This aims to set a minimum operational threshold for plant performance or adherence to discharge limits. Let us note that flow rate is an important determinant of the energy consumption of all MBR systems. In our experimental setting, however, the variability of the flow rate between output data was quite low, and therefore it was not included in the regression. In case of a different experimental setting, one may consider modeling as a first regression the relationships between energy consumption and flow rate, and then perform a new regression study, as we did, on the residuals between the first regression and the data. The suggestion to enlarge the number of regressors extends to exploring also other variables such as temperature, pH, or operational factors influencing oxygen consumption and pollutant removal. The suggested approach of correlating oxygen consumption with removal efficiencies offers a comprehensive perspective on plant performance, facilitating the establishment of minimum operational objectives.

3.4. Time Series Decomposition

Time series decomposition is a data analysis technique used to extract trend and seasonal components of a time series. Its main purpose is thus to highlight and separate long-term effects from periodical ones. This strategy is often used to break down, analyze, filter, and interpret complex data in a meaningful way for research, forecasting, or decision-making purposes. In detail, we have performed the following analysis:

1. Decomposition of time series: we have assumed an additive decomposition model, which assumes that a time series Y can be written as follows: Y = T + S + R, where T is a trend component, S is a seasonal component and R is a random component. For finding the components we have exploited the algorithm called STL ("Seasonal and Trend decomposition using Loess") [29] a robust approach able to deal with handle any type of seasonality, and capable of handling outliers and non-linear trends. In

our case, we looked for yearly seasonality (i.e., we wanted to highlight if there were regularities in the loadings of the same month over different years).

2. Data visualization: STL decomposition was used to visualize data by plotting long term trends and monthly seasonality.

The results helped to identify regularities that can be effectively exploited by the plant managers to plan the input loads to the biological reactor, forecasting and therefore minimizing fluctuations of the input loads [41].

To explain the practical impact of this analysis, we can concentrate our attention to the panel related to N-NO_3 in Figure 4. We can see that the original time series is quite noisy. However, the trend analysis clearly shows an increase starting on July 2020, with a peak in January 2022 and a following decrease in the last year. Moreover, the cyclic component highlights a low value in June (as it happens also for other pollutants) and a higher values in February and November. The combination of these aspects can be used to optimize the plant configuration at large, from the plants settings to the personnel shifts.

Looking at the results shown in Figure 4, it is possible to note that COD, TN, N-NO_3, N_{org} and TAS show a long-term increasing trend. Moreover, there is high long-term variability for N-NH_4^+, as well as for MBAS. The analysis of the seasonal components highlights that February and July are high loading months in terms of all input quantities, while August has low loadings for several inputs: COD, TN, MBAS, and N_{org}.

Figure 4. *Cont.*

Figure 4. Time series decomposition of the loads fed to the TBR relating to the parameters COD, Ntot, Ptot, MBAS, N−NH_4^+, Norg, N−NO3, TAS [kg/week]. Each subplot shows the original data (**upper plot**), the long-term trend (**middle plot**) and monthly boxplot over the years, where the blue line shows the boxplots and the red lines the average value (**lower plot**).

4. Future Outlooks and Conclusions

A full-scale TMR has a great elasticity in the examined AWTP throughout the five years of the research. An approach to evaluate the statistical variability of management data was investigated to derive alerts conditions in the management of the plant. Advanced descriptive statistics (violin graphs) highlight the treatment effectiveness of this process which combines thermophilic biodegradation with an ultrafiltration treatment; the loading mode of the reactor with the presence of some peaks uniformly distributed over the various years. Correlation matrixes (Pearson (P) and Spearman (S)) help in the future in the correct modeling of the observed parameters, providing information on the type of correlation between the data, whether linear or monotonic. Multiple linear regressions allow the AW utilities manager to be able to forecast the electrical and oxygen consumption values of the system based on the input parameters. Time series decomposition helped to identify regularities that can be effectively exploited by the plant managers to plan the input loads to the biological reactor, forecasting and therefore minimizing fluctuations of the input loads. Clearly, the approach presented in the paper needs to be further applied to other systems, to assess its potentials in different contexts. New analysis may provide insights about the methodological choices and suggest different statistical methods to highlight the main plant characteristics.

For future work, we will explore the different ways for effectively optimizing the treatment plant by looking both at the "loading" and "feeding" aspects. First, an important action is to try to reduce the variability of the loads in terms of physico-chemical properties and volumetric quantity. This may be possible by planning ex-ante the load based on the analysis of the historical loads of the plants. Second, a proper calibration of the feeding of pollutants may minimize oxygen and energy consumption during the oxidation process. Again, a proper statistical analysis of the monitoring data is essential to achieve the optimization goals.

Author Contributions: Conceptualization: M.C.C. and S.B.; methodology: M.C.C., F.M.C. and S.B.; validation: M.S., B.C. and M.B.; formal analysis: S.B.; investigation: A.A., S.B. and M.B.; resources: M.C.C. and S.B.; data curation: S.B. and F.M.C.; writing—original draft preparation: S.B., A.A. and B.C.; writing—review and editing: S.B. and F.M.C.; visualization: S.B., M.B., F.M.C. and A.A.; supervision: M.C.C. and S.B.; funding acquisition: M.C.C. All authors have read and agreed to the published version of the manuscript.

Funding: The authors thank Department of Civil Engineering and Architecture of Pavia for providing financial support to the reviews activities.

Data Availability Statement: Data are contained within the article.

Acknowledgments: We thank the Itelyum group and AsMortara's ownership of the plant for granting us the availability of the data. Finally, we appreciate the work of all volunteers who contributed to sample processing and data collection.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Ata, R.; Merdan, G.F.; Töre, G.Y. Activated Sludge Process for Refractory Pollutants Removal. In *Removal of Refractory Pollutants from Wastewater Treatment Plants*; CRC Press: Boca Raton, FL, USA, 2021.
2. Buttiglieri, G.; Knepper, T.P. Removal of Emerging Contaminants in Wastewater Treatment: Conventional Activated Sludge Treatment. In *Emerging Contaminants from Industrial and Municipal Waste: Removal Technologies*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 1–35.
3. Gurung, K.; Ncibi, M.C.; Fontmorin, J.M. Incorporating Submerged MBR in Conventional Activated Sludge Process for Municipal Wastewater Treatment: A Feasibility and Performance Assessment. *J. Membr. Sci. Technol.* **2016**, *6*, 1000158. [CrossRef]
4. Camargo, F.P.; Sérgio Tonello, P.; dos Santos, A.C.A.; Duarte, I.C.S. Removal of Toxic Metals from Sewage Sludge through Chemical, Physical, and Biological Treatments—A Review. *Water Air Soil Pollut.* **2016**, *227*, 433. [CrossRef]
5. Robinson, S.; Abdullah, S.Z.; Bérubé, P.; Le-Clech, P. Ageing of Membranes for Water Treatment: Linking Changes to Performance. *J. Membr. Sci.* **2016**, *503*, 177–187. [CrossRef]
6. Lackner, K.S.; Wendt, C.H.; Butt, D.P.; Joyce, E.L.; Sharp, D.H. Carbon Dioxide Disposal in Carbonate Minerals. *Energy* **1995**, *20*, 1153–1170. [CrossRef]
7. Nguyen, T.K.L.; Ngo, H.H.; Guo, W.; Chang, S.W.; Nguyen, D.D.; Nghiem, L.D.; Liu, Y.; Ni, B.; Hai, F.I. Insight into Greenhouse Gases Emissions from the Two Popular Treatment Technologies in Municipal Wastewater Treatment Processes. *Sci. Total Environ.* **2019**, *671*, 1302–1313. [CrossRef]
8. Hagos, K.; Zong, J.; Li, D.; Liu, C.; Lu, X. Anaerobic Co-Digestion Process for Biogas Production: Progress, Challenges and Perspectives. *Renew. Sustain. Energy Rev.* **2017**, *76*, 1485–1496. [CrossRef]
9. Collivignarelli, M.C.; Abbà, A.; Bertanza, G. Treatment of High Strength Pharmaceutical Wastewaters in a Thermophilic Aerobic Membrane Reactor (TAMR). *Water Res.* **2014**, *63*, 190–198. [CrossRef]
10. Simstich, B.; Beimfohr, C.; Horn, H. Lab Scale Experiments Using a Submerged MBR under Thermophilic Aerobic Conditions for the Treatment of Paper Mill Deinking Wastewater. *Bioresour. Technol.* **2012**, *122*, 11–16. [CrossRef]
11. Yang, S.-J.; Kataeva, I.; Hamilton-Brehm, S.D.; Engle, N.L.; Tschaplinski, T.J.; Doeppke, C.; Davis, M.; Westpheling, J.; Adams, M.W.W. Efficient Degradation of Lignocellulosic Plant Biomass, without Pretreatment, by the Thermophilic Anaerobe "*Anaerocellum thermophilum*" DSM 6725. *Appl. Environ. Microbiol.* **2009**, *75*, 4762–4769. [CrossRef]
12. Collivignarelli, M.C.; Abbà, A.; Frattarola, A.; Manenti, S.; Todeschini, S.; Bertanza, G.; Pedrazzani, R. Treatment of Aqueous Wastes by Means of Thermophilic Aerobic Membrane Reactor (TAMR) and Nanofiltration (NF): Process Auditing of a Full-Scale Plant. *Environ. Monit. Assess.* **2019**, *191*, 708. [CrossRef]
13. LaPara, T.M.; Nakatsu, C.H.; Pantea, L.; Alleman, J.E. Phylogenetic Analysis of Bacterial Communities in Mesophilic and Thermophilic Bioreactors Treating Pharmaceutical Wastewater. *Appl. Environ. Microbiol.* **2000**, *66*, 3951–3959. [CrossRef] [PubMed]
14. Collivignarelli, M.C.; Abbà, A.; Bertanza, G.; Setti, M.; Barbieri, G.; Frattarola, A. Integrating Novel (Thermophilic Aerobic Membrane Reactor-TAMR) and Conventional (Conventional Activated Sludge-CAS) Biological Processes for the Treatment of High Strength Aqueous Wastes. *Bioresour. Technol.* **2018**, *255*, 213–219. [CrossRef] [PubMed]
15. Collivignarelli, M.C.; Abbà, A.; Carnevale Miino, M.; Caccamo, F.M.; Argiolas, S.; Bellazzi, S.; Baldi, M.; Bertanza, G. Strong Minimization of Biological Sludge Production and Enhancement of Phosphorus Bioavailability with a Thermophilic Biological Fluidized Bed Reactor. *Process Saf. Environ. Prot.* **2021**, *155*, 262–276. [CrossRef]

16. Collivignarelli, M.C.; Todeschini, S.; Bellazzi, S.; Carnevale Miino, M.; Caccamo, F.M.; Calatroni, S.; Baldi, M.; Manenti, S. Understanding the Influence of Diverse Non-Volatile Media on Rheological Properties of Thermophilic Biological Sludge and Evaluation of Its Thixotropic Behaviour. *Appl. Sci.* **2022**, *12*, 5198. [CrossRef]
17. Collivignarelli, M.C.; Pedrazzani, R.; Bellazzi, S.; Carnevale Miino, M.; Caccamo, F.M.; Baldi, M.; Abbà, A.; Bertanza, G. Numerical Analysis of a Full-Scale Thermophilic Biological System and Investigation of Nitrate and Ammonia Fates. *Appl. Sci.* **2022**, *12*, 6952. [CrossRef]
18. Boldrin, A.; Neidel, T.L.; Damgaard, A.; Bhander, G.S.; Møller, J.; Christensen, T.H. Modelling of Environmental Impacts from Biological Treatment of Organic Municipal Waste in EASEWASTE. *Waste Manag.* **2011**, *31*, 619–630. [CrossRef]
19. Aguado, D.; Rosen, C. Multivariate Statistical Monitoring of Continuous Wastewater Treatment Plants. *Eng. Appl. Artif. Intell.* **2008**, *21*, 1080–1091. [CrossRef]
20. Liu, Y.; Pan, Y.; Sun, Z.; Huang, D. Statistical Monitoring of Wastewater Treatment Plants Using Variational Bayesian PCA. *Ind. Eng. Chem. Res.* **2014**, *53*, 3272–3282. [CrossRef]
21. Wang, X.; Ratnaweera, H.; Holm, J.A.; Olsbu, V. Statistical Monitoring and Dynamic Simulation of a Wastewater Treatment Plant: A Combined Approach to Achieve Model Predictive Control. *J. Environ. Manag.* **2017**, *193*, 1–7. [CrossRef]
22. O'Brien, G.J.; Teather, E.W. A Dynamic Model for Predicting Effluent Concentrations of Organic Priority Pollutants from an Industrial Wastewater Treatment Plant. *Water Environ. Res.* **1995**, *67*, 935–942. [CrossRef]
23. 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. *Sustainability* **2020**, *12*, 6348. [CrossRef]
24. Piterina, A.V.; Bartlett, J.; Pembroke, J.T. Morphological Characterisation of ATAD Thermophilic Sludge; Sludge Ecology and Settleability. *Water Res.* **2011**, *45*, 3427–3438. [CrossRef]
25. Rice, E.W.; Bridgewater, L.; American Public Health Association (Eds.) *Standard Methods for the Examination of Water and Wastewater*; American Public Health Association: Washington, DC, USA, 2012.
26. Aitchison, J.; Lauder, I.J. Kernel Density Estimation for Compositional Data. *Appl. Stat.* **1985**, *34*, 129. [CrossRef]
27. Azami, H.; Sarrafzadeh, M.H.; Mehrnia, M.R. Influence of Sludge Rheological Properties on the Membrane Fouling in Submerged Membrane Bioreactor. *Desalin. Water Treat.* **2011**, *34*, 117–122. [CrossRef]
28. Siwek, K.; Osowski, S. Data Mining Methods for Prediction of Air Pollution. *Int. J. Appl. Math. Comput. Sci.* **2016**, *26*, 467–478. [CrossRef]
29. Cleveland, R.B.; Cleveland, W.S.; McRae, J.E.; Terpenning, I. STL: A Seasonal-Trend Decomposition. *J. Off. Stat.* **1990**, *6*, 3–73.
30. Golyandina, N.; Zhigljavsky, A. *Singular Spectrum Analysis for Time Series*; SpringerBriefs in Statistics; Springer: Berlin/Heidelberg, Germany, 2013; ISBN 978-3-642-34912-6.
31. Lopetegui, J.; Sancho, L. Aerated Thermophilic Biological Treatment with Membrane Ultrafiltration: Alternative to Conventional Technologies Treating Paper Mill Effluents. *Water Supply* **2003**, *3*, 245–252. [CrossRef]
32. Pianosi, F.; Wagener, T. Distribution-Based Sensitivity Analysis from a Generic Input-Output Sample. *Environ. Model. Softw.* **2018**, *108*, 197–207. [CrossRef]
33. Fatone, F.; Di Fabio, S.; Bolzonella, D.; Cecchi, F. Fate of Aromatic Hydrocarbons in Italian Municipal Wastewater Systems: An Overview of Wastewater Treatment Using Conventional Activated-Sludge Processes (CASP) and Membrane Bioreactors (MBRs). *Water Res.* **2011**, *45*, 93–104. [CrossRef]
34. Siciliano, A.; Limonti, C.; Curcio, G.M.; Molinari, R. Advances in Struvite Precipitation Technologies for Nutrients Removal and Recovery from Aqueous Waste and Wastewater. *Sustainability* **2020**, *12*, 7538. [CrossRef]
35. Mañas, A.; Biscans, B.; Spérandio, M. Biologically Induced Phosphorus Precipitation in Aerobic Granular Sludge Process. *Water Res.* **2011**, *45*, 3776–3786. [CrossRef] [PubMed]
36. Zhu, F.-J.; Ma, W.-L.; Xu, T.-F.; Ding, Y.; Zhao, X.; Li, W.-L.; Liu, L.-Y.; Song, W.-W.; Li, Y.-F.; Zhang, Z.-F. Removal Characteristic of Surfactants in Typical Industrial and Domestic Wastewater Treatment Plants in Northeast China. *Ecotoxicol. Environ. Saf.* **2018**, *153*, 84–90. [CrossRef]
37. Chang, I.-S.; Chung, C.-M.; Han, S.-H. Treatment of Oily Wastewater by Ultrafiltration and Ozone. *Desalination* **2001**, *133*, 225–232. [CrossRef]
38. Arogo, J.; Westerman, P.W.; Heber, A.J.; Robarge, W.P.; Classen, J.J. Ammonia in Animal Production—A Review. In Proceedings of the 2001 Annual Meeting of the American Society of Association Executives, Philadelphia, PA, USA, 4–7 August 2001.
39. de Winter, J.C.F.; Gosling, S.D.; Potter, J. Comparing the Pearson and Spearman Correlation Coefficients across Distributions and Sample Sizes: A Tutorial Using Simulations and Empirical Data. *Psychol. Methods* **2016**, *21*, 273–290. [CrossRef] [PubMed]
40. Draper, N.R.; Smith, H. *Applied Regression Analysis*; John Wiley and Sons: Hoboken, NJ, USA, 1998.
41. Cheng, T.; Dairi, A.; Harrou, F.; Sun, Y.; Leiknes, T. Monitoring Influent Conditions of Wastewater Treatment Plants by Nonlinear Data-Based Techniques. *IEEE Access* **2019**, *7*, 108827–108837. [CrossRef]