

II-295 - DATA SCIENCE EM UMA ESTAÇÃO DE TRATAMENTO DE ESGOTO

Allan dos Anjos Pestana⁽¹⁾

Tecnólogo em Automação Industrial pela FATEC Itaquera e Tecnólogo em Gestão Ambiental pelo SENAC. Pós-graduado em Indústria 4.0 pelo SENAI e MBA em Gestão de Projetos pelo Centro Paula Souza. Encarregado em Automação/Instrumentação na Unidade de Tratamento de Esgoto da região metropolitana da Sabesp. Mestrando em Gestão e Tecnologia em Sistemas Produtivos no Centro Paula Souza, aluno do MBA em Data Science e Analytics da USP/ESALQ.

José Anselmo Pereira da Silva⁽²⁾

Tecnólogo em Informática com ênfase em Gestão de Negócios pela FATEC. Pós-graduando em Indústria 4.0 pelo SENAI. Técnico em Automação/Instrumentação na Unidade de Tratamento de Esgotos da região metropolitana da Sabesp.

Paulo André Caetano⁽³⁾

Técnico em Mecatrônica pelo Liceu Brás Cubas. Técnico em Automação Industrial pela ETEC Presidente Vargas. Aluno do curso técnico em Segurança e Saúde do Trabalho na ETEC Presidente Vargas. Técnico em Automação/Instrumentação na Unidade de Tratamento de Esgotos da região metropolitana da Sabesp.

Endereço⁽¹⁾: Rua João Lopes Maciel, 465 – Cidade Nitroquímica - SP - CEP: 08090-040 - Brasil - Tel: (11) 2037-4227 - e-mail: apestana@sabesp.com.br

Endereço⁽²⁾: Rua João Lopes Maciel, 465 – Cidade Nitroquímica - SP - CEP: 08090-040 - Brasil - Tel: (11) 2037-4270 - e-mail: janselmo@sabesp.com.br

Endereço⁽³⁾: Av. Major Pinheiro Fróes, 1560 – Vl. Maria de Maggi - Suzano - CEP: 08680-000- Brasil - Tel: (11) 4741-4528 - e-mail: pacaetano@sabesp.com.br

RESUMO

O objetivo é apresentar um piloto e os possíveis benefícios do uso das tecnologias habilitadoras da Indústria 4.0 no saneamento, mais especificamente em uma Estação de Tratamento de Esgotos (ETE), visando, por exemplo, melhoria no controle de Oxigênio Dissolvido (OD) nos tanques de aeração, controle dos produtos químicos utilizados na etapa de desidratação mecânica de lodo, controle do manto de lodo nos decantadores primários e secundários, entre muitos outros e consequentemente a estabilidade geral nos processos subsequentes. Visa também a eficiência energética nestas etapas do processo, considerando a Companhia de Saneamento, um dos maiores clientes da Companhia de Energia Elétrica da Região, e alguns de seus equipamentos podendo ser comparados ao consumo energético de pequenos municípios.

PALAVRAS-CHAVE: Data Science, Tratamento de Esgoto, Indústria 4.0, Digitalização, Transformação digital.

INTRODUÇÃO

A Sabesp é a maior companhia de saneamento da América Latina e considerada a quarta maior do mundo em população atendida, com mais de 28 milhões de pessoas abastecidas com água e 21 milhões de pessoas com coleta de esgoto. Detém a concessão dos serviços de saneamento básico de mais de 370 municípios do Estado de São Paulo, que é seu principal acionista.

A Unidade de Tratamento de Esgotos da região metropolitana de São Paulo, a MT, é uma das Superintendências da Sabesp. A Unidade em questão é dividida nos seguintes departamentos: Administrativo e Financeiro, Engenharia, Leste e Oeste, sendo esse penúltimo o objeto de estudo deste trabalho.

O Departamento Leste é composto por três grandes Estações de Tratamento de Esgoto e seus sistemas (Elevatórias e Sistemas isolados) são essas: ETE ABC, ETE São Miguel e ETE Suzano com média de vazão afluente de respectivamente, 2792 l/s, 1592 l/s e 977 l/s. As três plantas principais são estações de tratamento

por princípio de lodo ativado, onde resumidamente, é enviado oxigênio ao chamado licor misto nos tanques de aeração, com finalidade de que os microrganismos presentes no esgoto possam decompor a matéria orgânica. O fluxograma de uma dessas plantas, que evidencia também a separação dos processos entre fase líquida e sólida, e destaca os dois produtos finais do tratamento de esgoto, pode ser visto na figura 1:

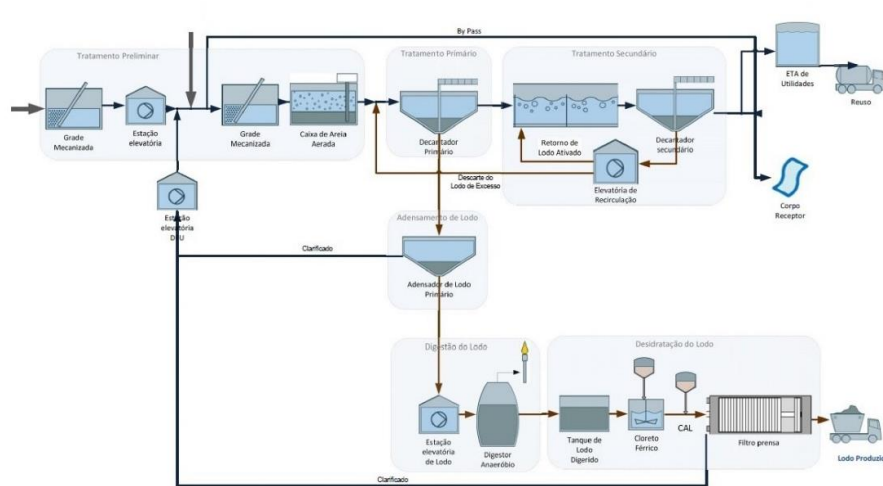


Figura 1 – Fluxograma do tratamento – ETE

Fonte: O autor (2022).

A Unidade de tratamento de esgoto, no ano de 2021 realizou uma contratação para dados analíticos de processo diferente do que se costumava. Ao invés da aquisição dos instrumentos, consumíveis, realizar calibração e limpeza, um contrato de performance agora é responsável por todas essas demandas, sendo remunerada enquanto o dado estiver disponível, chegando a um total de 255 pontos analisados em toda Superintendência. Entre eles pH, condutividade, amônia, oxigênio dissolvido (OD), sólidos suspensos totais (SST), turbidez, manto de lodo entre outros. Porém dados por si só não é informação, nem conhecimento, então trazer valor agregado para companhia com essas análises na forma de eficiência operacional e energética, enxergando a correlação entre esses dados é um dos objetivos desse trabalho.

Importante ressaltar que maiores níveis de Automação resultam em maior eficiência dos processos de tratamento e maior conformidade legal. Isso pode ser observado, por exemplo, num simples controle de nível em uma elevatória de esgoto, que implica em maior controle da vazão tratada, mitigação de eventos de extravasamento de esgoto e retorno em residências, e operação do equipamento de grande porte dentro de sua curva de projeto, resultando economia de energia e conservação do conjunto de bombeamento.

Objetivos

O objetivo geral desse trabalho é encontrar correlações entre as diferentes análises realizadas nas plantas, inclusive considerando o Tempo de Detenção Hidráulica (TDH) das etapas de tratamento. Para que em trabalhos futuros, as malhas de controle possam considerar essas correlações, visando inclusive antecipar ações necessárias.

Segundo Van Haandel e Marais (1999) para garantir que o tratamento de esgoto seja eficiente é preciso manter um valor mínimo de Oxigênio Dissolvido (OD) no licor misto (de 1 a 2 mg/L). A automação e controles simples podem auxiliar nesse controle, que é um exemplo, dentre os diversos controles existentes na planta.

Outro fato observado, é que a carga orgânica da estação pode ser muito variável, dependendo de diversos fatores, como por exemplo, horário do dia, período de chuvas, entre outros. A variação desses fatores, exige que a oxigenação nos tanques de aeração varie imediatamente. Segundo Metcalf e Eddy (2003) os aeradores devem permitir atender essas variações com flexibilidade e economizar energia, dado o custo energia nessa etapa do processo. Jordão e Pessoa (2014) afirmam que o oxigênio dissolvido é um parâmetro fundamental no tratamento por lodo ativado para decomposição da matéria orgânica.

REFERENCIAL TEÓRICO

Pirâmide de Automação

A literatura de automação, divide os elementos pertencentes a um sistema automático em cinco camadas, visando separá-los de acordo com sua funcionalidade hierárquica. A primeira camada, contempla todos os dispositivos de campo fundamentais para os controles, sejam atuadores, medidores de nível, vazão, pH e afins. Na segunda camada da pirâmide estão os controladores de processo, no caso das plantas de tratamento, são utilizados majoritariamente os Controladores Lógicos Programáveis (CLPs). Na terceira camada estão os sistemas de supervisão e as Interface Homem Máquina (IHMs) com a funcionalidade de realizar a primeira interface com usuário final. Na quarta camada, os gerenciadores da planta, e na Companhia de Saneamento em questão é utilizado o Plant Information Management System - Sistema de gerenciamento de informações da planta (PIMS), PI System da OSI Software. Por fim, na última camada estão os Enterprise Resource Planning - Sistema integrado de gestão empresarial (ERPs) e atualmente o utilizado na Companhia é o SAP (Goeking, 2010).

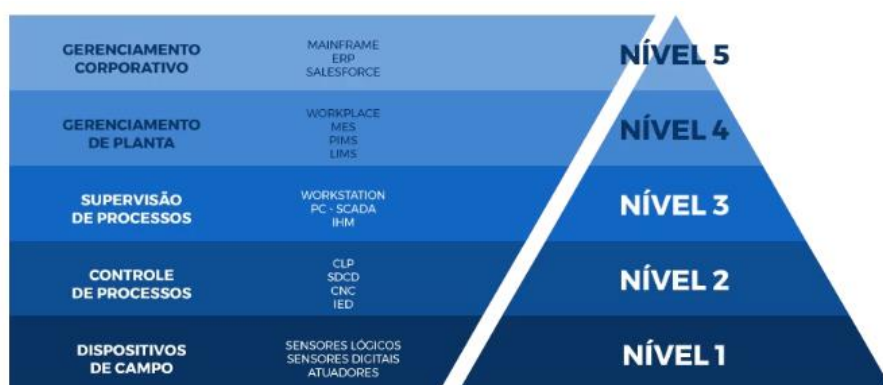


Figura 2 – Pirâmide de automação

Fonte: <https://www.altus.com.br/post/100/curso-de-introducao-a-automacao--5baula-01-5d>

IoT

Internet of things ou Internet das coisas, é um conceito fundamental para a indústria, onde o “caminho” de comunicação usualmente utilizado por seres humanos na internet, para realizar compras, se comunicar em redes sociais e afins, é nesse caso utilizado pelas máquinas (Sacomano et al., 2018).

Big Data

O termo ou conceito Big Data não é exatamente novo. De 2004 a 2010, por exemplo, a busca pelo tema se mantinha constante, foi então que o interesse pelo assunto começou a aumentar exponencialmente, chegando a ser dez vezes maior em 2018 (Google Trends, 2022). E como é possível afirmar isso sem muitas pesquisas? Utilizando a ferramenta do Google chamada Google Trends, que a partir de um número gigantesco de buscas em sua plataforma, converte esses dados em gráficos que relacionam as pesquisas em determinados locais e períodos do tempo, um bom exemplo de Big Data.

Um outro exemplo de Big Data pode ser visto nos mapas virtuais. Antes o mapa do GPS, era “egoísta”, encaminhava todos pelo melhor caminho, que a partir daí se abarrotava de motoristas e já não se mostrava ser exatamente o melhor. Hoje o Waze é colaborativo, utilizando-se dos dados dos usuários, indica o melhor caminho para cada indivíduo, e quando aquele caminho se tornar saturado de motoristas, o segundo melhor passará a ser a melhor rota, tudo isso de forma dinâmica. E isso só é possível com a utilização de Big Data. Em sua definição, Big Data, é uma grande quantidade de dados em volume, variedade e velocidade segundo Gartner (2022).

Indústria 4.0

Toda evolução humana e tecnológica que aconteceu gradualmente nos últimos séculos trouxeram benefícios para a indústria e meios de produção. A chamada primeira revolução industrial ocorre na segunda metade do século XVIII com a criação da máquina a vapor e a diminuição significativa do esforço braçal humano no trabalho. A segunda através da produção em escala e divisão do trabalho com o advento da eletricidade no começo do século XX. A terceira revolução é tida no início dos anos 70, com os robôs e controladores programáveis. (Kagermann; Wahlster; Helbig, 2013). A quarta revolução industrial é caracterizada pelos sistemas autônomos e capazes de tomar ou propor decisões até então impossíveis ao ser humano, haja vista a capacidade de processamento. Segundo Choi, Kang, Jung, Kulvatunyou e Morris (2016) é também caracterizado pela utilização de sistemas inteligentes e elevado grau de automação. O termo foi utilizado pela primeira vez em 2011 em Hanover na Alemanha segundo Drath e Horch, (2014).

Indústria 4.0 no Saneamento

Assim como a indústria, o saneamento demanda processos cada vez melhores e mais eficientes. Ainda levando em consideração o novo marco legal do saneamento, lei nº 14.026, de 15 de julho de 2020, que preconiza o setor de saneamento atender 99% da população com água potável e 90% com coleta e tratamento de esgoto até 31 de dezembro de 2033, o setor de saneamento tem um grande desafio pela frente.

O uso das tecnologias habilitadoras da Indústria 4.0 no saneamento se mostra como uma alternativa para o setor. Alguns dos pilares da indústria 4.0, tem a possibilidade de tornar, por exemplo, o tratamento de esgoto algo não mais reativo as mudanças de características do esgoto, do clima ou das bacias a que atende e sim podendo prever essas alterações com o uso de *Big Data e Machine Learning*.

Muitos controles necessários nas Estações de Tratamento de Esgoto, como por exemplo, controle de oxigênio dissolvido nos tanques de aeração, cloreto férrico e polímero na desidratação mecânica de lodo, ou hipoclorito na água de reuso já existem desde o início dessas Estações. Contudo, espera-se aplicar as tecnologias habilitadoras da Indústria 4.0 nesse setor, para alcançar um patamar que o controle humano não permite. Por exemplo, utilizar do Big Data, para minerar grande quantidade de dados relacionados as bacias pertencentes a determinada Estação, fatores climáticos e externos para prever e antecipar a atuação dos controles, tornando os processos muito mais eficientes.

Machine Learning

O *Machine Learning* pode ser definido como a capacidade de uma máquina de aprender sem ser previamente programado (Samuel, 1959). Posteriormente redefinido e categorizado em quatro possíveis grupos de algoritmos de *Machine Learning*: supervisionado, não supervisionado, semisupervisionado e aprendizado por reforço (Mitchell, 1997).

Inteligência Artificial

A inteligência artificial tem por objetivo fazer com que determinada máquina se comporte ou pense com inteligência segundo Gomes (2010). Por exemplo, em um dia de chuva, é comum que uma pessoa leve um guarda-chuva, em um dia de chuva e frio é esperado também que essa mesma pessoa leve um guarda-chuva e uma blusa. Agora imagine em uma indústria em que diversos fatores acarretam uma porção de implicações. É bem provável que a capacidade humana, não seja suficiente para dar a esses processos todas a resposta de que ele precise, e as máquinas, sua capacidade computacional e a inteligência artificial podem ser uma alternativa para tal.

Avaliação de Maturidade da Acatech

Para identificação de qual momento a empresa está na implementação das tecnologias habilitadoras da indústria 4.0, é possível utilizar uma avaliação de maturidade, como por exemplo, a da ACATECH (German Academy of Science and Engineering – Academia Alemã de Ciências e Engenharia) de Schumacher, Erol e Sihm (2016) e Schuh et al. (2017). Como apontado em algumas revisões bibliográficas, a Europa é o local, segundo suas publicações, com maior avanço no uso das tecnologias habilitadoras atualmente. No Brasil, infelizmente, boa parte da indústria ainda se encontra entre a segunda e terceira revolução industrial, sendo o setor automotivo o que apresenta melhor desempenho nesse sentido (FIRJAN, 2016)

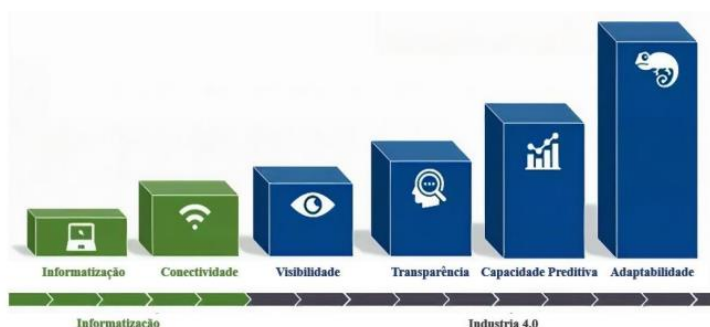


Figura 3 – Escala de maturidade – Indústria 4.0 - ACATECH

Fonte: https://www.gov.br/cti/pt-br/publicacoes/producao-cientifica/seminario-pci/x_seminario_pci-2020/pdf/seminario-2020_paper_25.pdf

MATERIAIS E MÉTODOS

A metodologia escolhida para o projeto, foi a experimental. Trabalhos pontuais executados anteriormente no departamento demonstram os resultados financeiros e operacionais obtidos com o investimento em automação nos processos. Contudo o potencial de benefícios projetados pode ir muito além. A Unidade de Tratamento em questão que, por exemplo, possuía poucos instrumentos analíticos para controle de seus processos, tem previsão de aumento exponencial dessas análises. Com os dados analíticos das plantas de tratamento de esgoto tabulados, estudamos a possibilidade de haver correlações entre as séries de dados. O Excel possui uma fórmula para o cálculo do coeficiente de correlação (correlação de Person), a fórmula CORREL. Abaixo a equação que a fórmula utiliza:

$$\text{Correl}(X, Y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Equação (1)

Ou seja, a partir desse ponto podemos enxergar possíveis correlações entre dados analíticos de diferentes etapas do processo. Como esses dados ainda estão em processo de registro segue exemplo da planilha criada no Excel, afim de criar um piloto para o estudo, inicialmente com dados aleatórios, com periodicidade de 5 minutos.

ABES - Associação Brasileira de Engenharia Sanitária e Ambiental

	Local	Comparação Real time																						
		Parâmetro	PHAT-101	PHAT-102	PHAT-103	PHAT-104	PHAT-105	PHAT-106	PHAT-107	PHAT-108	PHAT-109	PHAT-110	PHAT-111	PHAT-112	PHAT-113	PHAT-114	PHAT-115	PHAT-116	PHAT-117	PHAT-118	PHAT-119	PHAT-120		
	Esgoto bruto	pH		0,48	-0,06	0,24	-0,39	0,34	0,00	-0,40	0,25	-0,72	0,38	-0,15	-0,31	0,70	0,00	-0,30	-0,02	0,10	-0,40	0,26	0,51	
	Esgoto bruto	Condutividade			0,22	0,22	-0,29	0,15	0,55	-0,62	0,04	-0,79	0,20	0,26	-0,05	0,69	0,01	-0,57	0,27	0,00	0,02	-0,27	0,27	
	Esgoto desarenado	SST (imersão)				0,71	-0,44	0,31	0,41	0,04	0,38	0,15	-0,17	-0,26	-0,27	-0,34	0,33	-0,59	0,47	-0,61	-0,27	0,20	0,27	
	Esgoto desarenado	DOO					-0,24	0,16	0,12	0,12	0,12	0,14	0,48	0,09	-0,21	-0,03	0,70	-0,36	0,25	-0,50	-0,05	0,32	0,53	
	Captação Aquapolo	NH3						-0,95	0,02	0,37	-0,65	0,20	-0,02	0,35	0,32	0,15	0,07	0,05	0,17	0,12	0,54	0,12	-0,78	
	Efluente Primário	SST (imersão)							-0,14	-0,22	0,57	-0,15	0,15	-0,27	-0,38	-0,19	-0,17	0,18	-0,26	-0,23	-0,74	-0,25	0,82	
	Efluente Final	DOO							-0,09	0,24	-0,31	-0,40	0,26	-0,08	0,35	0,21	-0,51	0,19	-0,11	0,09	-0,40	-0,16		
	Licor misto	SST (imersão)								-0,08	0,63	0,09	0,35	0,34	-0,45	0,22	0,25	0,12	-0,40	0,04	-0,03	-0,28		
	Reatores Aeróbios	ORP (Médio) - Saída									-0,19	-0,57	-0,53	-0,18	-0,19	-0,34	-0,31	0,12	-0,08	-0,63	0,02	0,11		
	Reatores Aeróbios	Oxigênio Dissolvido (Médio) - Entrada									-0,12	0,02	0,26	-0,84	0,42	0,43	-0,21	-0,07	0,29	0,12	-0,16			
	Reatores Aeróbios	Oxigênio Dissolvido (Médio) - Saída												0,50	0,05	0,30	0,28	0,24	-0,08	-0,20	0,02	-0,03	0,56	
	Efluente Primário	NH3													0,36	0,30	0,36	0,22	-0,12	0,04	0,49	-0,61	-0,04	
	Efluente Final	NH3													-0,12	0,19	0,02	0,08	0,40	0,68	-0,12	-0,41		
	Efluente D5s	Turbidez (Média)														-0,01	-0,29	-0,01	0,26	0,09	-0,11	0,06		
	Tanque D5s	Manto de lodo (Médio)																-0,95	-0,23	0,10	0,49	0,15	0,24	
	Efluente Final	pH																	-0,78	0,23	-0,01	-0,41	0,19	
	Lodo Atm. Digestor	SST (imersão)																		-0,57	-0,06	0,36	-0,36	
	Lodo condicionado	SST (imersão)																				0,53	-0,02	-0,16
	Filtrado FPP 1+3	SST (imersão)																					0,02	-0,49
	Filtrado FPP 1+3	NH3																						-0,11
	Efluente Final	Condutividade																						

Figura 6 – Dados analíticos, correlações (ETE ABC)
Fonte: Do autor, 2022.

de Transferência			Fonte		Alinhamento																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
=CORREL(\$B\$5:\$B\$13;BD5:BD13)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
AT	AU	AV	AW	AX	AY	AZ	CORREL(matriz1; matriz2)										BF	BG	BH	BI	BJ	BK	BL	BM	BN	BO	BP																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
Loja At. 1.º andar	Loja cond. 1.º andar	Filial FPP 1 a 3	Filial FPP 1 a 3	Elav. Final																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				</

[illegible]

Figura 8 – Dados analíticos, correlações com delay 1 hora (ETE ABC)
Fonte: Do autor, 2022.

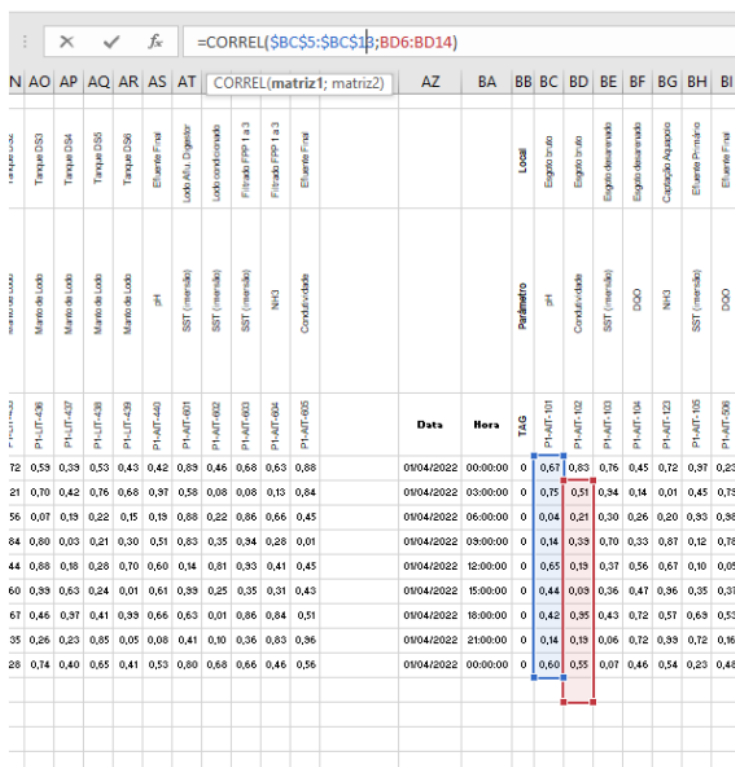


Figura 9 – Correlação com delay, seleção da série de dados (ETE ABC)

Fonte: Do autor, 2022.

Algumas etapas de fato possuem correlação com outras com determinado “delay”, porém, na maioria dos casos, esse delay não é conhecido. Portanto, a próxima etapa do trabalho foi realizar essa mesma análise de correlação, porém considerando diferentes atrasos horários, com a finalidade de encontrar quais são maiores para correlações que de fato existem. Segue:

HORAS	TAG	Parâmetro										Local										TAG									
		PI-AUT-102	PI-AUT-103	PI-AUT-104	PI-AUT-105	PI-AUT-106	PI-AUT-107	PI-AUT-108	PI-AUT-109	PI-AUT-110	PI-AUT-111	PI-AUT-112	PI-AUT-113	PI-AUT-114	PI-AUT-115	PI-AUT-116	PI-AUT-117	PI-AUT-118	PI-AUT-119	PI-AUT-120	PI-AUT-121	PI-AUT-122	PI-AUT-123	PI-AUT-124	PI-AUT-125	PI-AUT-126	PI-AUT-127	PI-AUT-128	PI-AUT-129	PI-AUT-130	
0		0,03	0,00	0,04	0,01	0,09	0,03	0,02	-0,05	0,00	0,00	0,14	-0,01	0,02	-0,01	-0,03	0,06	-0,01	-0,02	0,00	0,00	0,09	0,10	-0,01	-0,02	0,04	-0,01				
1		-0,05	-0,01	0,11	-0,05	0,03	-0,05	-0,03	-0,08	0,01	0,03	-0,07	-0,08	0,06	0,01	-0,01	-0,07	0,01	-0,08	-0,03	-0,02	0,06	-0,02	-0,05	-0,07	0,02	-0,05				
2		-0,01	-0,01	0,02	-0,04	0,06	0,09	0,02	-0,01	-0,04	0,03	0,03	-0,04	-0,03	0,01	0,01	0,04	0,04	0,03	0,04	0,01	-0,07	-0,09	0,03	0,03	0,02	-0,05				
3		0,00	0,02	0,02	0,04	-0,07	-0,03	-0,02	0,02	0,02	0,06	0,01	0,01	0,03	0,05	-0,01	0,06	-0,02	0,07	-0,01	0,09	0,04	-0,10	0,03	0,05	0,02	-0,03				
4		0,02	-0,01	-0,01	-0,03	-0,01	0,04	0,03	0,01	0,04	-0,01	0,03	-0,04	0,10	-0,07	0,07	0,04	-0,02	0,04	-0,05	0,03	-0,03	0,01	0,01	-0,03	-0,05	0,05				
5		-0,02	-0,05	0,02	0,02	0,02	0,02	0,05	-0,05	0,05	-0,16	0,01	0,06	0,01	-0,07	0,09	0,06	-0,01	-0,01	-0,04	-0,02	-0,06	-0,03	0,02	-0,02	0,00	-0,04				
6		0,12	0,00	0,02	-0,02	-0,02	0,02	-0,04	-0,01	-0,03	0,00	0,04	0,00	0,01	0,01	0,06	-0,03	0,04	0,07	-0,11	0,05	0,02	-0,06	-0,10	0,04	-0,01	0,03				
7		0,05	-0,02	-0,05	0,09	0,04	0,09	-0,04	0,07	-0,03	-0,01	0,03	-0,02	-0,07	0,06	0,02	-0,03	-0,02	-0,01	-0,03	0,00	-0,03	0,07	0,03	0,00	-0,02					
8		-0,03	0,07	0,04	-0,01	-0,07	-0,03	-0,01	-0,01	-0,01	0,03	-0,03	0,05	-0,04	-0,05	-0,02	-0,04	-0,01	-0,05	0,03	0,00	0,04	0,07	-0,05	0,00	-0,11	-0,01				
9		0,02	0,07	0,03	-0,09	0,05	0,05	0,02	-0,03	-0,03	0,04	-0,01	-0,07	0,03	0,04	0,00	-0,03	0,03	-0,02	0,01	-0,04	0,03	-0,01	0,00	-0,02	0,04	0,01				
10		-0,02	-0,05	0,04	0,01	0,01	0,00	-0,05	-0,03	0,04	0,03	0,04	0,02	-0,05	0,07	0,06	0,01	0,07	-0,05	-0,03	0,00	0,03	0,07	0,02	0,11	-0,04	0,01				

Figura 10 – Correlações com diversos delays

Fonte: Do autor, 2022.

Afim de identificar quais pares de analíticos possuem correlação e quais não, foi realizada uma soma em cada uma das colunas. Haja vista que a formula do Excel utilizada, retorna um número entre “-1” (altamente correlacionável, indiretamente proporcional) e “1” (altamente correlacionável, diretamente proporcional), valores maiores dessa soma, tanto positivo quanto negativo, nos mostrarão quais dados valem a atenção. Segue:

22	0,02	-0,05	0,07	0,01	-0,01	0,03	-0,04	0,07	-0,06	0,01	-0,05	0,12	-0,03	0,10	-0,05	0,04	0,06	-0,06	-0,01	-0,01
23	0,06	-0,02	0,05	-0,03	-0,09	-0,03	-0,11	0,07	0,02	-0,04	0,05	0,02	0,07	-0,14	-0,04	0,04	-0,02	0,05	0,04	-0,01
24	-0,03	0,04	-0,01	0,02	-0,01	-0,04	0,03	0,02	-0,01	-0,02	0,08	-0,04	-0,05	-0,09	-0,08	0,03	-0,09	0,04	0,03	-0,01
SOMA	0,19	-0,44	0,19	0,00	-0,21	-0,07	-0,43	0,06	0,06	-0,04	0,34	-0,11	-0,01	-0,30	-0,13	0,13	0,13	-0,09	0,06	0,00

Figura 11 – Soma das correlações

Fonte: Do autor, 2022.

Por fim, é possível visualizar os gráficos para cada uma das correlações, com base nessa “biblioteca” posteriormente criar padrões de operação e para os próprios controladores responsáveis por esses processos:

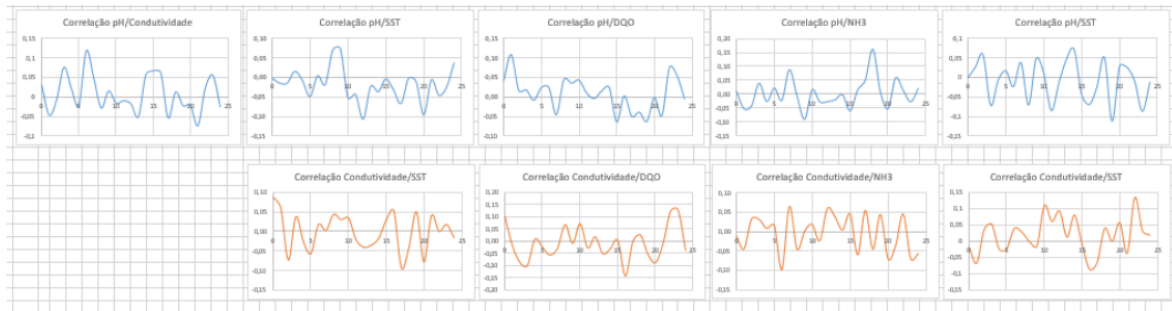


Figura 12 – Gráficos da modelagem
Fonte: Do autor, 2022.

Nessa etapa é possível já simular o uso da planilha, inicialmente de maneira manual, com dados reais, por exemplo, da ETE ABC, e foi possível enxergar algumas correlações inclusive considerando o tempo de detenção hidráulica das etapas envolvidas:

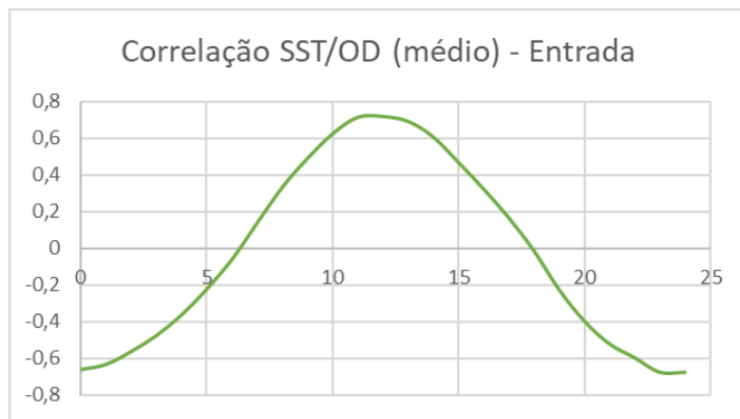


Figura 13 – Correlação do SST com OD em diversos delays
Fonte: Do autor, 2022.

Com esse gráfico podemos concluir que há uma forte correlação (0,72) entre os dados de Sólidos Suspensos Totais no esgoto desarenado com o Oxigênio Dissolvido nos Tanques de Aeração na ETE ABC. E não somente isso, mas conseguimos concluir que essa correlação é maior depois de exatamente 11 horas. Essa correlação, agora observando-se os dados propriamente ditos, podem ser vistos no gráfico abaixo:

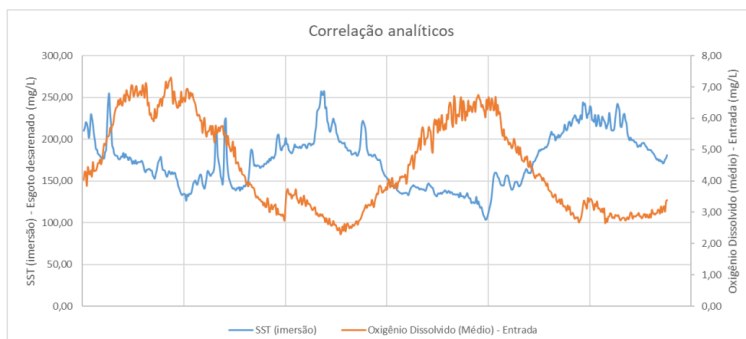


Figura 14 – Dados SST e OD – ETE ABC
Fonte: Do autor, 2022.

Outro exemplo pode ser visto, agora analisando-se a ETE São Miguel. Uma forte correlação (0,7) pode ser enxergada entre o SST do licor misto e os mantos de lodo dos DSs com um atraso de 12 horas:

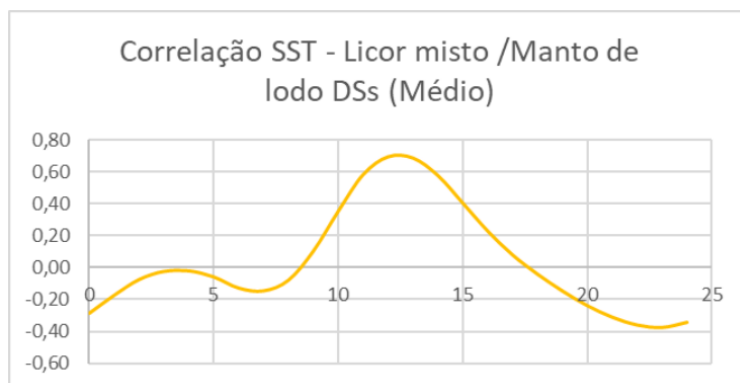


Figura 15 – Correlação do SST do licor misto com o manto de lodo dos DSs
Fonte: Do autor, 2022.

CONCLUSÕES

Muitas outras correlações podem ser obtidas com suas respectivas diferenças horárias, iniciando assim um banco de dados de como operar de maneira otimizada as Estações de Tratamento, prevendo falhas e antecipando atuações dos controles.

Etapas futuras do trabalho, devem fazer essa análise mostrada e outras de forma automática, haja vista que esses dados estarão chegando aos controladores, sendo enviados ao supervisão e armazenados no banco de dados das plantas.

A ideia nesse ponto do trabalho é utilizar o software da empresa BirminD que possui soluções para uso das tecnologias habilitadoras da Indústria 4.0. Espera-se utilizar dos módulos: “Relevância”, “Otimização”, “Simulação” e “Detecção de anomalias”. Abaixo podemos visualizar o módulo “simulação” no add-in para Excel utilizando o mesmo banco de dados que foi usado para obter as correlações observadas anteriormente. Observado o resultado da análise, temos que diversos dados analíticos possuem correlação com o pH nas porcentagens que foram apresentadas, ou seja, é possível entender quais dados possuem mais correlação em relação aos outros. Outro módulo interessante é de “Otimização” que determina como devem ser os outros parâmetros, afim de obter um set-point específico em determinada análise:

output	pH - Esgoto bruto	
inputs_relevance	SST (imersão) - Esgoto desarenado	13,05%
	ORP - Reator Aeróbio 1A	12,38%
	ORP - Reator Aeróbio 1B	10,41%
	ORP - Reator Aeróbio 2A	9,27%
	ORP - Reator Aeróbio 1B	8,65%
	Oxigênio Dissolvido - Reator Aeróbio 1B	6,87%
	Oxigênio Dissolvido - Reator Aeróbio 2A	6,63%
	Oxigênio Dissolvido - Reator Aeróbio 2B	5,37%
	Oxigênio Dissolvido - Reator Aeróbio 1A	4,89%
	Oxigênio Dissolvido - Reator Aeróbio 2B	4,63%
	Oxigênio Dissolvido - Reator Aeróbio 1B	3,63%
	SST (imersão) - Filtrado FPP 1 a 3	3,04%
	Manto de Lodo - Tanque DS5	2,59%
	Manto de Lodo - Tanque DS6	2,51%
	Manto de Lodo - Tanque DS1	2,11%
	Manto de Lodo - Tanque DS3	2,06%
	Manto de Lodo - Tanque DS2	1,92%

output	P1-AIT-101	
results	P1-AIT-603	3,35
	P1-LIT-439	1,04
	P1-LIT-438	1,04
	P1-LIT-436	1,31
	P1-LIT-435	0,5
	P1-LIT-434	0,92
	P1-AIT-421	5,43
	P1-AIT-419	1
	P1-AIT-418	5,27
	P1-AIT-417	4,86
	P1-AIT-415	0,66
	P1-AIT-414	4,8
	P1-AIT-411	108,94
	P1-AIT-410	282,16
	P1-AIT-409	245,43
	P1-AIT-408	234,09
	P1-AIT-103	182,94
entries_score	7	
optimal_point	7	

Figura 16 – Simulação de relevância e otimização add-in para Excel BirMind em dados analíticos do tratamento de esgoto
Fonte: Do autor, 2022.

Outro conceito importante é entender onde a companhia se encontra e onde pode se posicionar no médio prazo com o uso do Data Science, passando do Business Intelligence (BI) para o Business Analytics, com análises preditivas e prescritivas.

Para isso será fundamental a intersecção de algumas áreas do conhecimento afim de se chegar a esse Data Science, são essas: Matemática e Estatística, Ciência da Computação e o Domínio do conhecimento.

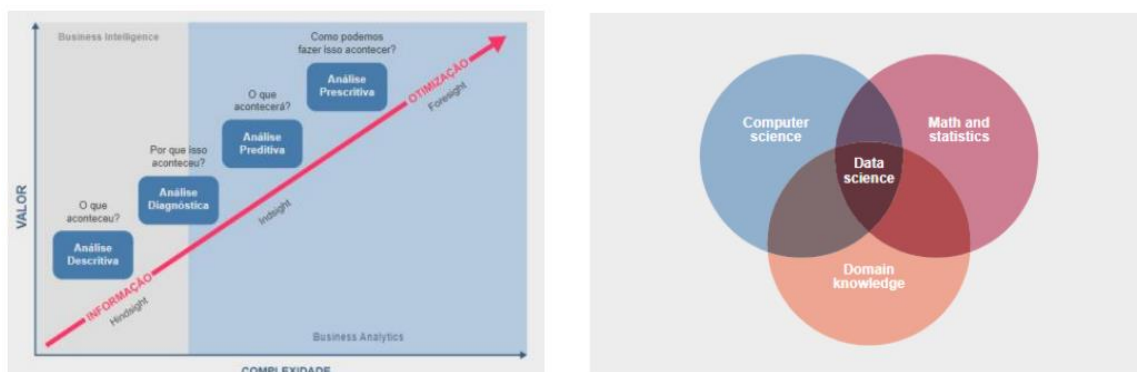


Figura 17 – Business Intelligence (BI) para o Business Analytics e Tripé do Data Science
Fonte: Paulucci & Rubim, 2022

Por fim a ideia é que possamos desenhar algo como um mapa das Estações de Tratamento de Esgoto, com base nos dados analíticos agora analisados, e elencar as correlações que foram enxergadas como os respectivos “delays”:

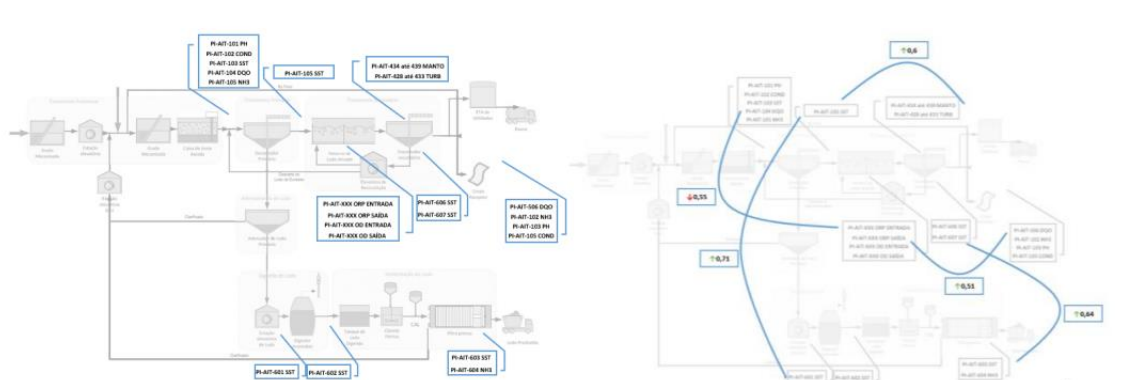


Figura 18 – Fluxograma do processo com localização dos dados analisados, elencando correlações (exemplo)
Fonte: Paulucci & Rubim, 2022

REFERÊNCIAS BIBLIOGRÁFICAS

1. BBC. Eles pararam de reclamar dos governos e estão usando o celular para melhorar a política. Disponível em: <https://www.bbc.com/portuguese/brasil-36660936>. Acesso em: 24 mar. 2022.
2. CHOI, S. S.; KANG, G.; JUNG, K.; KULVATUNYOU, B. & MORRIS, K.C. Applications of the factory design and improvement reference activity model. In: INTERNATIONAL CONFERENCE ON ADVANCES IN PRODUCTION MANAGEMENT SYSTEMS (APMS), 2016. Iguassu Falls, Brazil. Disponível em: <https://hal.inria.fr/hal-01615741/document>. Acesso em: 14 mar. 2022.
3. BRASIL. Presidência da República. Lei n. 14.026, de 15 de julho de 2020. Atualiza o marco legal do saneamento básico. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/lei/114026.htm. Acesso em: 14 mar. 2022.
4. DRATH, R.; HORCH, A. Industrie 4.0: Hit or hype? IEEE industrial electronics magazine, v. 8, n. 2, p. 56–58, 2014.
5. PAULUCCI, A.; RUBIM, L. FIAP - Nano courses. Big Data & Analytics. Disponível em: <https://on.fiap.com.br/local/nanocourses/index.php> Acesso em: 27 jun. 2022
6. JORDÃO, E. P.; PESSOA, C. A. Tratamento de esgotos domésticos. 7. ed. Rio de Janeiro: ABES, 2014.
7. GARTNER. Tecnologia da Informação. 2022. Disponível em: <https://www.gartner.com/en/information-technology/glossary/big-data>. Acesso em: 14 mar. 2022.

8. GOEKING, W., Da máquina a vapor aos softwares de automação, Revista O Setor Elétrico, Edição 52, Brasil, maio de 2010. Disponível em: https://www.voltimum.com.br/sites/www.voltimum.com.br/files/memoria_maio_10.pdf. Acesso em: 10 mar. 2022.
9. GOOGLE TRENDS. Big data. Disponível em: <https://trends.google.com.br/trends/explore?date=all&geo=BR&q=BIG%20DATA>. Acesso em: 08 mar. 2022.
10. KAGERMANN, H.; WAHLSTER, W.& HELBIG, J. Recommendations for implementing the strategic initiative industrie 4.0: final report of the Industrie 4.0, 2013. Disponível em: <https://www.din.de/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf>. Acesso em: 14 mar. 2022.
11. METCALF & EDDY. Wasterwater engineering: treatment and reuse. 4. ed. [s. l.]: McGraw-Hill, 2003.
12. MITCHELL, Tom. Machine learning. New York: McGraw-Hill, 1997.
13. ALTUS. Curso de introdução a automação. Disponível em: <https://www.altus.com.br/post/100/curso-de-introducao-a-automacao--5b9aula-01-5d>. Acesso em: 23 jun. 2022.
14. SACOMANO, J.B. et al. Indústria 4.0: conceitos e fundamentos. São Paulo: Blucher, 2018.
15. SAMUEL, Arthur L. Some studies in machine learning using the game of checkers. IBM Journal of Research and Development, v. 3, n. 3, pp. 210-229, July, 1959. DOI: 10.1147/rd.33.0210. Disponível em: <https://ieeexplore.ieee.org/document/5392560>. Acesso em: 23 jun. 2022.
16. VAN HAANDEL, A. C.; MARAIS, G. V. R. O Comportamento do Sistema de Lodo Ativado. Campina Grande: Epgraf, 1999.