

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CURSO DE TECNOLOGIA EM PROCESSOS QUÍMICOS
CAMPUS TOLEDO**

JANAINA BORGES FERREIRA

**MODELO ESTATÍSTICO PARA O CRESCIMENTO DA BACTÉRIA E. COLI
CONSIDERANDO INFLUÊNCIA DE PARÂMETROS FÍSICO-QUÍMICOS E ÍNDICE
PLUVIOMÉTRICO: UM ESTUDO DE CASO NO MANANCIAL DE CASCAVEL**

TRABALHO DE CONCLUSÃO DE CURSO

**TOLEDO
2018**

JANAINA BORGES FERREIRA

**MODELO ESTATÍSTICO PARA O CRESCIMENTO DA BACTÉRIA E. COLI
CONSIDERANDO INFLUÊNCIA DE PARÂMETROS FÍSICO-QUÍMICOS E ÍNDICE
PLUVIOMÉTRICO: UM ESTUDO DE CASO NO MANANCIAL DE CASCAVEL**

Trabalho de conclusão de Curso de graduação, apresentado à disciplina de TCCII do Curso Superior de Tecnologia em Processos Químicos da Universidade Tecnológica Federal do Paraná - UTFPR, Campus Toledo, como requisito parcial para obtenção do título de Tecnólogo em Processos Químicos.

Orientador: Prof.^o Dr.^o Ricardo Fiori Zara
Coorientadora: Prof.^a Dr.^a Suellen R. P Garcia

**TOLEDO
2018**

**TERMO DE APROVAÇÃO
DO TRABALHO DE CONCLUSÃO DE CURSO**

JANAINA BORGES FERREIRA

MODELO ESTATÍSTICO PARA O CRESCIMENTO DA BACTÉRIA E. COLI
CONSIDERANDO INFLUÊNCIA DE PARÂMETROS FÍSICO-QUÍMICOS E ÍNDICE
PLUVIOMÉTRICO: UM ESTUDO DE CASO NO MANANCIAL DE CASCAVEL

Trabalho apresentado como forma de avaliação para o Trabalho de Conclusão de
Curso do curso de Tecnologia em Processos Químicos da UTFPR,
Campus Toledo, e aprovado pela banca examinadora abaixo.

Prof.º Dr.º. Ricardo Fiori Zara
ORIENTADOR - UTFPR/TOLEDO

Prof.ª Dr.ª. Suellen Ribeiro Pardo Garcia
COORIENTADORA - UTFPR/TOLEDO

Prof.ª M.Sc. Michelle Maria Detoni Zanette
UTFPR/TOLEDO

Prof.ª Dr.ª. Regiane Slongo Fagundes
UTFPR/TOLEDO

“O Termo de Aprovação assinado encontra-se na Coordenação do Programa”.

AGRADECIMENTOS

Agradeço a Deus primeiramente por que sem ele não chegaria a lugar algum, também gostaria e agradecer a minha família, meu marido Marcio que muitas vezes deixou de fazer certas coisas por que eu tinha prova, trabalhos, etc. Aos meus filhos Bruno, Beatriz e Vitor que gostariam que eu estivesse em casa, mesmo assim me incentivaram a continuar nos momentos em que pensei em desistir. E agradecer em especial a minha filha Beatriz que muito me ajudou nos trabalhos, muitas vezes bravinha dizendo assim “o que mãe? Já te ensinei isso”.

Agradeço aos meus pais pelos ensinamentos do que é certo e errado na vida mesmo tendo saindo de casa muito cedo. “Que ninguém é melhor que ninguém”.

Agradecer as minhas colegas de trabalho de uma maneira geral que mesmo com muito serviço sempre me ajudaram e incentivaram que iria dar certo, e eu iria conseguir.

Também agradeço ao meu orientador Prof.^o Dr^o. Ricardo Zara, pelo incentivo durante o desenvolvimento do trabalho. Agradecer em especial a Prof.^a. Dr^a. Suellen Ribeiro Prado Garcia que muito me ajudou na elaboração dos modelos estatísticos, auxiliando com material e sugestões na elaboração do trabalho.

Agradeço a Companhia de Saneamento do Paraná pela oportunidade de realização da pesquisa.

Agradeço também ao Sistema meteorológico do Paraná pela disponibilidade dos dados com isso possibilitando o trabalho.

RESUMO

FERREIRA, Janaína Borges. **Modelo estatístico para o crescimento da bactéria *E.Coli* considerando influência de parâmetros físico-químicos e índice pluviométrico: um estudo de caso no manancial de Cascavel**, 2018. Trabalho de conclusão de curso – Universidade Tecnológica Federal do Paraná, Toledo – PR. Orientador: Prof.^o Dr.^o Ricardo Fiori Zara e Coorientadora Prof.^a Dr.^a. Suellen Ribeiro Pardo Garcia.

Estima-se um aumento nos estudos em que utilizam modelagem estatística como ferramenta para auxiliar no monitoramento da qualidade da água. O presente trabalho foi realizado para estudar a influência dos fatores físico-químicos e índice pluviométrico no crescimento microbiológico do manancial, que afeta diretamente a qualidade da água do rio Cascavel, indispensável para abastecimento da cidade, onde modelos de regressão linear foram ajustados para estimar a carga microbiana do rio, com dados da estação de tratamento de água de Cascavel entre os anos de 2014 a 2017. O estudo foi conduzido com base na carga microbiana indicadora de *E.coli*. As variáveis explicativas foram analisadas com base em dados hídricos e climáticos da bacia do rio Cascavel como precipitação acumulada, pH e turbidez. As primeiras análises estatísticas como o cálculo das estatísticas descritivas, correlações, transformações nas variáveis foram utilizadas para que o primeiro modelo fosse ajustado aos dados. Os resultados do primeiro modelo não foram satisfatórios então percebeu-se a necessidade de trabalhar com as séries diferenciadas, o que melhorou significativamente as estatísticas de teste do modelo. Por fim, tentou-se uma forma mais parcimoniosa, que mostrou resultados satisfatórios, portanto foi a escolhida. Os resultados da análise de regressão indicaram que as variáveis independentes explicam 70% da variação dos *E.coli*, valor também encontrado nos modelos da literatura (entre 50% e 80%), no entanto, o modelo de regressão linear desenvolvido não ficou tão simples como os modelos de referência, pois a diferenciação da série e os termos de correção de erro foram necessários. Os resultados obtidos no modelo adotado se mostraram eficientes visto que obtivemos uma correlação significativa entre a *E. coli* e a turbidez, onde a mesma é influenciada pela precipitação, mas a precipitação não influencia diretamente a *E. coli*. O modelo pode ser utilizado como ferramenta em monitoramento de mananciais.

Palavras-chave: Regressão, Previsão, Turbidez, Previsão e Modelagem.

ABSTRACT

FERREIRA, Janaína Borges. **Statistical model for the growth of E.Coli bacteria considering the influence of physico-chemical parameters and rainfall index: a case study in the Cascavel spring, 2018.** Undergraduate thesis – Federal University of Technology -Toledo, Paraná. Advisor: Professor Ricardo Fiori Zara, PhD, and Joint Advisor: Professor Suellen Ribeiro Pardo Garcia, PhD.

It has been reported that there is a raise in the number of studies that use the statistics model as a useful tool for water quality monitoring. This paper will focus on the influence of physical-chemical factors and pluviometric index on the microbiological growth of the wellspring, which affects instantly the water quality of the Cascavel river, necessary for the city water maintenance, where linear regression models were adjusted to estimate the microbial rate of the river, with data from the water treatment plant of Cascavel between the years of 2014 and 2017. The study was conducted based on the microbial rate of *E. coli*. The variables were analyzed according to water data and climate data from the Cascavel river basin, such as precipitation, pH and turbidity. The first statistical analysis, as the descriptive statistics evaluation, connections, changes in the used variables were used so the first model could be adjusted to the data. The results of the first model were not acceptable, so it was noticed the need of performing with differentiating series, which significantly improved the model trial statistics. Finally, it was tried out in a parsimoniously way, which showed satisfactory results, and therefore, was the chosen one. The results of the regression analysis indicated that the independent variables explain 70% of the *E. coli* variation, taking into account the numbers were superior to those found on the literature (between 40% and 60%), however, the developed linear regression model was not as simple as the referential ones, because the serie differentiation and the terms of error repair were necessary. The obtained results in the adopted model showed itself efficient, since it was obtained a significant correlation between *E. coli* and the turbidity, which is affected by precipitation. Nevertheless, the precipitation does not influence the *E. coli*, so the model can be used as an assistance for the wellspring monitoring.

Key words: Regression, precipitation, Turbidity, Forecast and Modeling.

LISTA DE SÍMBOLOS, SIGLAS E ABREVIACÕES

- ✓ ADF - Dickey-Fuller Aumentado
- ✓ APHA - American Public Health Association
- ✓ Colilert - Substrato enzimático
- ✓ E. coli - Escherichia coli
- ✓ FIB - Bactérias indicadoras fecais
- ✓ PRECCUM - Precipitação acumulada
- ✓ pH - Potencial Hidrogeniônico
- ✓ SIMEPAR - Sistema Meteorológico do Paraná
- ✓ SINGREH - Sistema Nacional de Gerenciamento de Recursos Hídricos
- ✓ TURB -Turbidez
- ✓ USGS - US Geological Survey
- ✓ VIF - Fator de Inflação de Variância

LISTA DE TABELAS

Tabela 1: Estatística descritiva das variáveis envolvidas no modelo.	27
Tabela 2: Correlação entre as variáveis E.coli, turbidez, precipitação acumulada, pH e termo de sazonalidade (cosseno).	28
Tabela 3: Resultados do modelo 1.....	29
Tabela 4: Resultados do modelo 1 modificado.....	30
Tabela 5: Resultados do modelo 2.....	31
Tabela 6: Resultados do modelo 3.....	32
Tabela 7: Resultado do teste de fator de inflação da variância para o modelo 3.	33
Tabela 8: Saída do teste de correlação serial nos resíduos do modelo 3.....	33

LISTA DE FIGURAS

Figura 1: Mapa das bacias hidrográficas de Cascavel	19
Figura 2: Localização das bacias Hidrográficas do município de Cascavel	20
Figura 3: Resíduos e valores observados e ajustados pelo modelo 3.	32
Figura 4: Valores previstos pelo modelo e limites inferiores e superiores para a previsão.....	34

SUMÁRIO

1 INTRODUÇÃO	11
1.1 OBJETIVOS	13
1.1.1 Objetivo Geral	13
1.1.2 Objetivos Específicos	13
2 REVISÃO BIBLIOGRÁFICA	14
2.1 IMPORTÂNCIA DA ÁGUA E POLUIÇÃO HÍDRICA	14
2.2 QUALIDADE DA ÁGUA.....	14
2.3 LEGISLAÇÃO BRASILEIRA	15
2.4 PARÂMETROS FÍSICO-QUÍMICO E MICROBIOLÓGICOS.....	16
2.4.1 Turbidez.....	16
2.4.2 Potencial Hidrogeniônico (pH)	16
2.4.3 Índice Pluviométrico.....	17
2.4.4. Coliformes	17
2.5 MODELAGEM PARA A PREVISÃO DA <i>E. COLI</i>	18
3 MATERIAL E MÉTODOS	19
3.1 BACIA HIDROGRÁFICA DE CASCAVEL	19
3.2 COLETA DE DADOS	21
3.2.1 Análises Microbiológicas.....	21
3.2.2 Análises de pH.....	21
3.2.3 Turbidez.....	22
3.2.4 Índice Pluviométrico.....	22
3.3 MODELO ESTATÍSTICO.....	23
4 RESULTADOS E DISCUSSÕES	27
4.1 MODELO 1	29
4.2 MODELO 2	30
4.3 MODELO 3	31
5. CONCLUSÃO	35
REFÊRENCIAS	36

1 INTRODUÇÃO

A água é indispensável na conservação do equilíbrio e biodiversidade, nas relações de seres vivos e ambientes naturais assumindo uma função importante no organismo, sendo indispensável para o metabolismo humano pois transportam nutrientes, vitaminas e sais minerais (BACCI, 2008)

Destinada ao consumo humano deve seguir um rigoroso padrão de potabilidade, com conjunto de valores permitidos e definidos pelos órgãos regulamentadores diminuindo os riscos de contaminações. Sendo submetidos a processos físicos, químicos e biológicos ou até mesmo combinados para obtermos um produto de qualidade à população (BRASIL, 2011).

Com isso sua qualidade pode ser definida pelo uso e ocupação das bacias hidrográficas sendo que a poluição ou contaminação pode ser natural, onde parâmetros utilizados são indicadores da qualidade e constituem impurezas, quando alcançam valores superiores determinados nas legislações (PORTAL TRATAMENTO DE ÁGUA, 2009).

A água provém de mananciais, lagos, rios e lençóis subterrâneos, onde corresponde a uma parte da água disponível. Esses rios são separados por classes em função de sua destinação.

A resolução do CONAMA nº 357/2005, art. 38 “determina o enquadramento dos corpos de água de acordo com normas e procedimentos definidos pelo conselho de recursos hídricos - CNRH e conselhos Estaduais de Recursos Hídricos - CERH, ” e o artigo 42 diz que” enquanto não for aprovado um novo enquadramento, as águas doces são consideradas classe dois” (BRASIL, 2005).

No controle de qualidade da água ao consumo humano são pesquisados microrganismos indicadores de contaminação, ou seja, as bactérias do grupo coliformes onde encontram-se um subgrupo, os coliformes totais e *E.coli*. (TORTORA, 2012).

A água proveniente de mananciais sem tratamento apresenta uma grande quantidade de parasitas e microrganismos causadores de doença, podendo entre eles citar a salmonella, shigella que são considerados patogênicas para o ser humano. As bactérias do grupo coliforme pertencem a família Enterobacteriaceae e são utilizadas como indicadores de contaminação da qualidade da água proveniente dos mananciais de abastecimento. Assim é fundamental a importância a verificação de presença de

coliformes termotolerantes mais especificamente a *E.coli* que constitui 95% da microbiota intestinal de humanos e animais de sangue quente (SILVA, et al. 2005)

Considera-se importante a relação velocidade do escoamento superficial e aumento na erosão podendo afetar diretamente a qualidade hidrodinâmica dos corpos hídricos, visto que partes das bacias hidrográficas perderam suas áreas vegetativas originais (ANDREOLI, et al. 2005).

A precipitação pode influenciar a vazão e impactos no solo, com isso desencadeando processos erosivos que aumentam a turbidez, pH, e teor do material orgânico, devido as entradas de água nos mananciais, mas dependendo da região é difícil estabelecer relação entre a precipitação e a qualidade da água (FRAGA et al. 2012)

Nesse caso, modelos estatísticos podem ser utilizados com dados de região próxima, onde foi possível realizar a coleta. Outra aplicabilidade seria quando o conhecimento da concentração microbiana é necessário para rápida tomada de decisões na estação de tratamento de água, ou seja, quando não há tempo hábil para a coleta e análise da quantidade microbiana em laboratório.

Os modelos estatísticos fornecem uma ferramenta para práticas de manejo de rios em escala regional ou mesmo em bacias hidrográficas, como a visualização da extensão e da tendência da carga de patógenos. Baseando-se nestas problemáticas é perceptível a relevância desta pesquisa, que tem como objetivo geral estudar a relação entre o crescimento microbiológico no manancial que abastece a cidade de Cascavel com as variações de pH, turbidez e índice pluviométricos acumulado.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Estudar a relação entre as variáveis físico-químicas e microbiológicas do manancial de abastecimento público do Rio Cascavel, utilizando dados de análises realizadas pela Sanepar.

1.1.2 Objetivos Específicos

- Identificar as correlações e propor um modelo estatístico para previsão da *E. coli*.
- Verificar a possível influência do índice pluviométrico no nível de contaminação do rio Cascavel com a bactéria.
- Examinar se existe relação entre o pH, turbidez e índices pluviométricos com o nível de contaminação da *E. coli*.

2 REVISÃO BIBLIOGRÁFICA

2.1 IMPORTÂNCIA DA ÁGUA E POLUIÇÃO HÍDRICA.

Segundo a Agência Nacional de Águas (2013), o estabelecimento de um parâmetro que indique a qualidade das águas é uma ferramenta importante para nortear ações de planejamento e gestão.

A água é um dos elementos mais importantes do mundo, sendo predominantemente essencial para vida humana. O Brasil é considerado um dos países mais ricos em água, porém pelo desmatamento e a falta de cuidados muitos rios e nascentes já secaram (DANTAS, 2008).

A Lei nº 6.938, de 31 de agosto de 1981 define poluição como a degradação da qualidade ambiental decorrente de atividades direta ou indiretamente, prejudicando a saúde, a segurança e o bem-estar da população. Com isso condições adversas às atividades sociais e econômicas, influenciem desfavoravelmente a biota e às condições estéticas, sanitárias do ambiente ou lancem matérias e energias em desacordo com os padrões ambientais estabelecidos (BRASIL, 1981).

Os recursos hídricos devem ser analisados, pois é por meio das poluições encontradas que são inseridas as substâncias nocivas, tóxicas ou patogênicas que modificam as características da água, fazendo com que a mesma necessite de monitoramento através de controles de qualidade (CORADI et al. 2009).

2.2 QUALIDADE DA ÁGUA

O Ministério da Saúde do Brasil, estabelece a Portaria de Consolidação nº5 de 28 de setembro de 2017, como “Consolidação das normas sobre as ações e os serviços de saúde do Sistema Único de Saúde”, originando-se da portaria nº 2.914 de 12 dezembro 2011, que rege os parâmetros de qualidade da água para consumo humano, auxiliando o controle de pontos críticos. De acordo com seu artigo 2º contido no anexo “Norma de Qualidade da Água para Consumo Humano”, toda água destinada ao consumo humano deve obedecer ao padrão de potabilidade e está sujeita à vigilância da qualidade da água.

Em decorrência de problemas relacionados a poluição criam-se séries de processos e equipamentos destinados a fornecer água potável a população. A captação é o início do processo de produção e retirada de água do manancial. Já a adução corresponde ao transporte da água, por meio de adutoras desde a captação até a estação de tratamento e reservatório de distribuição, considerado a parte mais importante do processo, removendo as impurezas e microrganismos presente. (BRAGA et al. 2002).

Estudos sistemáticos podem ser utilizados para modelar e prever a qualidade microbiológica da água bruta com base em dados físico-químicos disponíveis. Podendo ser os parâmetros para avaliar o nível de riscos para a saúde relacionados com a produção de água potável para melhorar as práticas de gestão de captação (EREGNO, 2013)

No Brasil os serviços públicos de abastecimento de água potável e de esgoto sanitário são prestados por agentes da administração pública direta ou indireta. Segundo dados do Sistema Nacional de Informações sobre o Saneamento (SNIS), até 2016, as companhias Estaduais de Saneamento Básico (CESBs) atendiam cerca de 71,1% dos municípios e 75,5% da população urbana no país (SNIS, 2016). No Paraná, a empresa responsável por fornecer água tratada para população é a companhia de Saneamento do Paraná (SANEPAR).

2.3 LEGISLAÇÃO BRASILEIRA

O Ministério da Saúde, através da portaria de consolidação nº 5 de setembro de 2017, dispõe sobre os procedimentos de controle e de vigilância da qualidade da água para consumo humano e seu padrão de potabilidade (BRASIL, 2017).

A água *in-natura* por meio da resolução nº 430, de 13 de maio de 2011 complementa e altera a resolução nº 357/2005. Dispõe sobre as condições e padrões de lançamento de efluentes (BRASIL, 2011).

A Lei nº 9.433/1997, instituiu a Política Nacional de Recursos Hídricos criou o Sistema Nacional de Gerenciamento de Recursos Hídricos (SINGREH). Onde a água é considerada um bem de domínio público e um recurso natural limitado, dotado de valor econômico. O instrumento legal prevê ainda que a gestão dos recursos hídricos

deve proporcionar os usos múltiplos das águas, de forma descentralizada e participativa, contando com a participação do poder público, dos usuários e das comunidades. A lei também prevê que em situações de escassez o uso prioritário da água é para o consumo humano e para a dessedentação de animais.

2.4 PARÂMETROS FÍSICO-QUÍMICO E MICROBIOLÓGICOS

2.4.1 Turbidez

A turbidez é um parâmetro baseado na comparação da intensidade da luz causada por matéria suspensa coloidal provocada por meio de rochas, argila, algas, de origem antrópica podendo estar relacionada à presença de compostos tóxicos e organismos patogênicos; diminuindo a penetração de luz, prejudicando a fotossíntese em corpos d'água (SPERLING, 2005).

Com isso, constatamos que a turbidez é uma variável extremamente importante em monitoramento de micro bacias hidrográficas, atuando como indicador de programas de manejo e conservação de solos, pois em sistemas de tratamento quando elevada formam flocos decantando rapidamente, dificultando o tratamento da água (FUNASA, 2013).

2.4.2 Potencial Hidrogeniônico (pH)

O pH representa a quantidade de íons H^+ presentes na água, parâmetro importante no tratamento, exercendo grande influência na solubilidade das substâncias afetando o processo de coagulação e também auxilia no controle da desinfecção. (SPERLING, 2005).

O pH de substância que pode variar de acordo com a composição, concentração de sais, metais, ácidos, bases, substâncias orgânicas e da temperatura. Tem grande influência na coagulação química, controle da erosão, abrandamento e desinfecção (SANEPAR, 2014).

No tratamento de água por processos químicos ou biológicos o pH deve ser mantido em faixas adequadas ao desenvolvimento de reações químicas do processo

podendo estar entre 6,0 a 9,0. Sendo um dos parâmetros mais importantes na determinação da maioria das espécies químicas tanto de água potável, como de água residuais (SANEPAR, 2014).

2.4.3 Índice Pluviométrico

O índice pluviométrico é uma medida que resulta na quantidade precipitada de água, podendo ser chuva, granizo ou neve, que influenciam na distribuição dos micro e macro nutrientes dos mananciais, elevando os níveis de turbidez, onde a mesma estando elevada pode comprometer a capacidade fotossintética dos microrganismos através da diminuição da luz solar (FRAGA. et al, 2012)

Conforme Richter (2009), as chuvas influenciam diretamente nos valores de matéria em suspensão em um corpo hídrico, devido o carregamento de material particulado, sendo a turbidez considerada uma medida indireta dos sólidos em suspensão.

2.4.4. Coliformes

O grupo coliforme é constituído por um número de bactérias que inclui os gêneros *Escherichia*, *Klebsiella*, *Citrobacter*, *Clebsiella* e *Enterobacter*. Originárias do intestino de animais ou humanos de sangue quente. Consideradas bacilos gram-negativos e aeróbios ou anaeróbios facultativos não formadores de esporos, oxidase negativos, capazes de se desenvolver na presença de sais biliares ou agentes tensoativos, fermenta a lactose com produção de ácidos, gás e aldeídos a 35° C mais ou menos 0,5° C em 24 horas, e que podem apresentar atividades da enzima Belta galactosidase. (CETESB, 2008).

Cerca de 95% dos coliformes existentes nas fezes humanas e animais são *E.coli*, sendo utilizada como um indicador de contaminação. Quando se fala em presença de coliformes fecais, possibilita a presença de bactérias patogênicas prejudicial à saúde, mas diante disto a água é tratada para essa e outras bactérias que possam estar presentes na mesma, visto que existem um padrão de potabilidade (SANEPAR, 2014).

De acordo com a Portaria de consolidação N°5 2017 do Ministério da Saúde (BRASIL 2017), quando forem detectadas amostras com resultados positivos para coliformes totais, mesmo que em ensaios presuntivos, devem ser coletadas novas amostras, em dias sucessivos que revelem resultados satisfatórios. Ainda, se em sistemas de distribuição, as recoletas devem apresentar três amostras simultâneas, uma sendo no mesmo ponto e as outras duas a montante e a jusante.

2.5 MODELAGEM PARA A PREVISÃO DA *E. COLI*

O modelo de regressão linear pode ser útil na identificação de fatores físico-químicos que afetam a qualidade da água e a equação de regressão são avaliadas por meio de medidas estatísticas como o R^2 ajustado. Os resultados da análise de regressão fornecem um meio para o monitoramento rápido da qualidade da água bruta microbiana com base nos parâmetros físico-químicos, como descarga fluvial, temperatura, precipitação, pH, turbidez, condutividade, cor e carbono orgânico total. (EREGNO, 2013)

No modelo de concentração *E.coli* desenvolvido por Vijayashanthar, et al. (2018), foi selecionado variáveis como a temperatura da água, turbidez, chuvas acumuladas, descarga de fluxo fluvial, distância entre a usina de recuperação a montante e número de emissários de esgoto combinados a montante.

Em Senior (2017) apresenta-se uma modelagem considerando amostras do fluxo da água para análise das concentrações de coliformes, durante sete meses. As equações de regressão incluíram possíveis variáveis explicativas de vazão de fluxo, turbidez, temperatura da água e fatores sazonais calculados e houve a necessidade de transformações logarítmicas nas variáveis selecionadas.

3 MATERIAL E MÉTODOS

3.1 BACIA HIDROGRÁFICA DE CASCAVEL

As águas que pertencem ao município de Cascavel são classificadas entre três bacias hidrográficas, sendo elas a do Rio Iguaçu, Bacia do Rio Piquiri e do Rio Paraná, tendo como latitude sul 24° 57' 21", e longitude oeste de 53° 27' 19", abrangendo a área de drenagem de 117,50 km² na bacia hidrográfica (OLIVEIRA, 2009).

Todas as bacias hidrográficas são consideradas áreas geográficas, local de drenagem de suas águas para um recurso hídrico, sendo relativa à qualidade da água de uma bacia para outra.

A Bacia Hidrográfica do Rio Cascavel tem sua principal nascente na região do Lago municipal e possui outras nas quais são em áreas rurais como o rio Saltinho e rio Peroba. Já as do perímetro urbano apresentam pontos negativos por serem localizadas próxima a rodovia federal BR 277 que em casos de acidentes comprometem o abastecimento de água da cidade, a captação de água para o abastecimento da cidade é feita no Rio Cascavel e seus afluentes (GLADIS, 2005)

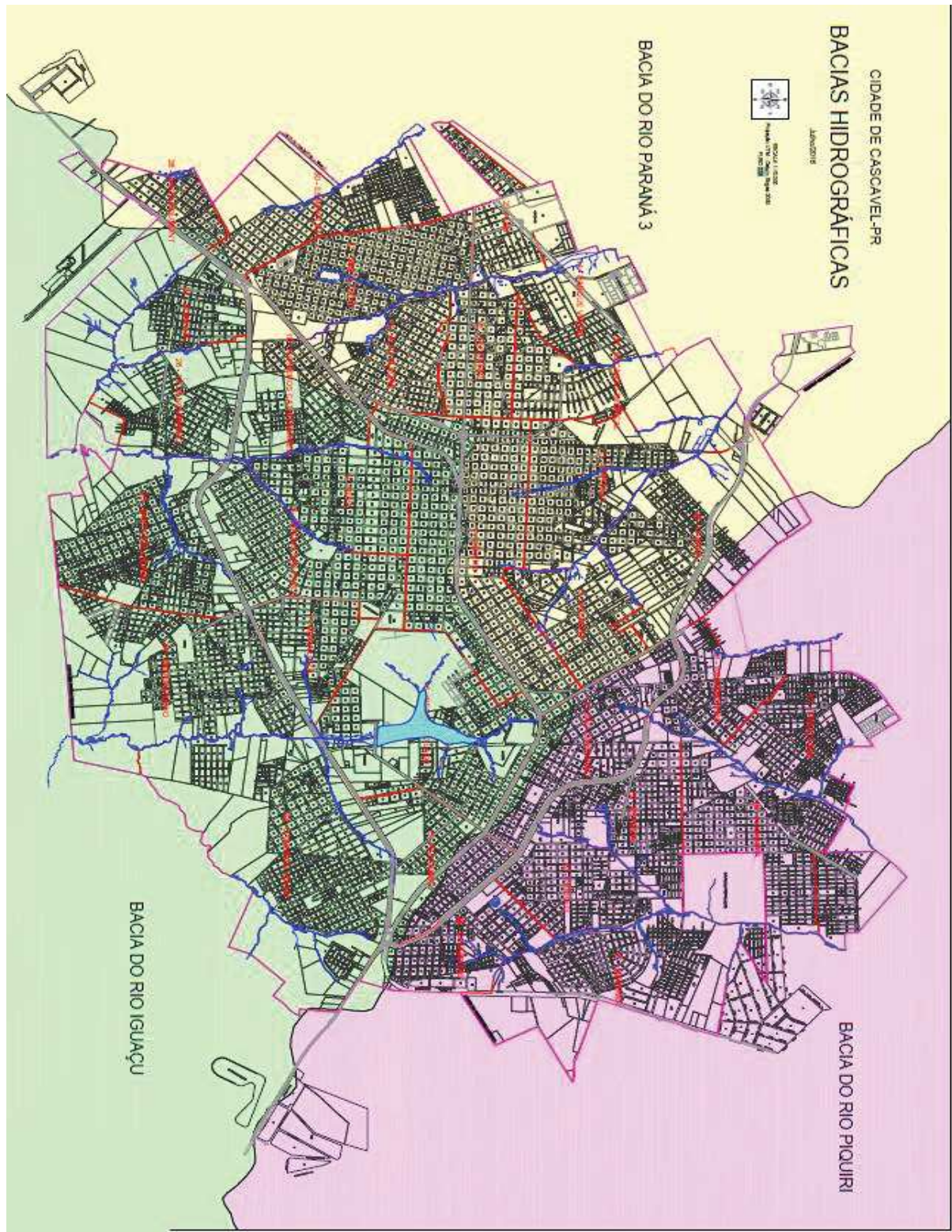
Figura 1: Mapa das bacias hidrográficas de Cascavel



Fonte: Revista Bacias Hidrográficas, 2010.

Já na Figura 2 mostra-se a localização das bacias hidrográficas de Cascavel, sendo o rio Cascavel pertencente ao rio Iguaçu, incluindo no mapa as bacias do rio Paraná e rio Piquiri

Figura 2: Localização das bacias Hidrográficas do município de Cascavel



Fonte: Portal do município de Cascavel (2018).

3.2 COLETA DE DADOS

As amostras foram coletadas nos mananciais de abastecimento de Cascavel, mais precisamente no rio Cascavel. As análises microbiológicas foram realizadas no laboratório central da companhia de saneamento do Paraná – Sanepar de Cascavel. E os demais parâmetros físico-químicos, turbidez, e pH foram realizados nos laboratórios descentralizados, localizados na estação de tratamento de água (ETA 3), situada na área rural da cidade.

As análises microbiológicas e físico-químicas foram realizadas utilizando a metodologia determinadas pelo American Public Health Association (APHA, 2005).

Foram considerados análises realizadas nos laboratórios descentralizados da Sanepar, junto a estação de tratamento de água da seguinte forma com observações mensais do manancial de Cascavel/PR, totalizando um ponto de amostragem com 48 observações no período de janeiro de 2014 a dezembro de 2017.

3.2.1 Análises Microbiológicas

As amostras microbiológicas foram coletadas em frascos autoclavados enviadas sobre refrigeração a temperatura máxima de 8 °C e foram analisadas utilizando a técnica de quantificação através de cartelas. Sendo elas diluídas conforme necessidade e características de cada fonte. Após este preparo inicial é adicionado o meio de cultura transferindo o conteúdo para a cartela Quanti-tray/2000 quando *in-natura*, levando então na seladora para selar a mesma. Após este processo deve-se incubá-la a 35°C ± 0,5°C. Sendo os resultados obtidos em 24 a 28 horas quando utilizado o substrato Colilert/ DEXX®.

3.2.2 Análises de pH

A determinação do pH foi realizada utilizando o método potenciométrico, onde dois eletrodos medem (eletrodos de referência e indicador) em função do volume do reagente adicionado. Após a chegada da amostra foram calibrados os pHmetro com

tampões de referência pH 4 e pH 7, logo após a calibração o eletrodo foi colocado na amostra e assim medido o potencial de Hidrogênio da água.

Os principais fatores que determinam o pH da água são: gás carbônico dissolvido e alcalinidade. Recomenda-se que o pH da água seja mantido na faixa de 6,0 a 9,5.

3.2.3 Turbidez

Para medir a turbidez utilizamos um nefelometro, o qual utiliza padrões secundário (ampolas), ou seja, padrões certificados para faixas específicas de turbidez, sendo os mesmos utilizados para padronizar o turbidímetro antes das leituras, e como controle para ajustes com padrão primário.

Se a turbidez menor que 40UT deve-se agitar a amostra suavemente até que as bolhas de ar desapareçam, fazendo a leitura da turbidez diretamente na escala do instrumento ou na curva de calibração apropriada. Mas caso a amostra da turbidez estejam acima de 40UT, deve-se diluir a amostra com 1 ou mais volume de água isenta de turbidez até que a amostra diluída fique entre 30UT a 40UT, após fazer a leitura e multiplicar pelo resultado de fator de diluição, obtêm-se o resultado (APHA, 2005).

3.2.4 Índice Pluviométrico

Foram considerados dados para a pesquisa os índices pluviométricos cedidos pelo SIMEPAR/PR, onde a precipitação determinada em (mm). Sendo utilizado instrumento para medida denominado de pluviômetro. Dados cedidos da estação nº 24535333 de Cascavel. Os resultados analisados a partir do método precipitação acumulada de sete dias de janeiro de 2014 a dezembro de 2017 que antecedem a coleta.

3.3 MODELO ESTATÍSTICO

O uso de estatísticas e modelagem envolvendo parâmetros de qualidade de água é considerada uma poderosa ferramenta analítica que tem crescido significativamente nos últimos anos.

Mais especificamente, as pesquisas que envolvem modelos estatísticos envolvendo a bactéria *E. coli* tem como variáveis explicativas (variáveis independentes) do modelo, propriedades da água e dados hidro climáticos como: descarga do rio, temperatura da água, precipitação, pH, turbidez, condutividade, cor, entre outros (ELERIA e VOGEL, 2005; EREGNO, 2013; MCCULLOCH, 2015; SENIOR, 2017).

A fim de encontrar um modelo estatístico que descreva o relacionamento entre a bactéria *E. coli* e algumas dessas variáveis explicativas precisa-se de um conhecimento do conjunto de dados. As estatísticas descritivas foram utilizadas para descrever as características básicas do conjunto de dados do estudo.

Realizando a análise de correlação, por meio do coeficiente de correlação linear de Pearson, para investigar as relações entre a bactéria *E. coli* e as variáveis ambientais e físico-químicas da qualidade da água. Quanto mais próximo de 1 ou -1 esse coeficiente se apresenta, maior é a relação entre as duas variáveis.

Toda análise estatística foi realizada por meio do *software EViews 9.5 StudentVersionLite*, que é uma versão gratuita. Esta versão possui algumas limitações quanto ao número de observações e variáveis envolvida, o que não impede que pesquisas no âmbito acadêmico sejam desenvolvidas.

A forma geral do modelo de regressão linear é dada por:

$$y_t = \hat{y}_t + e_t = b_0 + b_1x_{1t} + b_2x_{2t} + \dots + b_kx_{kt} + e_t, \quad t = 1, 2, \dots, n \quad (1)$$

onde y é a variável dependente, x_1, x_2, \dots, x_k são as variáveis independentes (explicativas), t são as observações da amostra e o termo e é o resíduo associado com a resposta y_t com o valor ajustado (WALPOLE, 2009, p. 285). Quando se usa o conceito de mínimos quadrados para chegar as estimativas dos coeficientes $b_i, i=0, \dots, n$ minimiza-se a expressão

$$SQE = \sum_{t=1}^n e_t = \sum_{t=1}^n (y_t - b_0 - b_1 x_{1t} - b_2 x_{2t} - \dots - b_k x_{kt})^2. \quad (2)$$

Para mais detalhes vide Walpole (2009) e Anderson; Sweeney; Williams, (2002).

Dois testes foram utilizados para avaliar a significância de uma relação no modelo de regressão, o teste F e o teste t. Para selecionar as variáveis independentes avaliamos a significância estatística dos coeficientes utilizando a estatística-t dada pela seguinte expressão.

$$t = \frac{b_i}{S_{b_i}}, \quad (3)$$

onde S_{b_i} é o desvio padrão de cada coeficiente.

Esta estatística verifica se determinada variável independente contribui significativamente para o modelo de regressão ou não (Anderson; Sweeney; Williams, 2002).

A hipótese nula do teste é que a variável não tem efeito sobre o modelo contra a hipótese alternativa de que a variável tem efeito sobre o modelo, ou seja, a hipótese nula é de que o coeficiente b_i da variável x_i seja nulo $H_0: b_i = 0$ e $H_a: b_i \neq 0$. Assim, o valor da estatística-t é calculado para cada estimativa do coeficiente, e se a probabilidade associada a cada valor da estatística-t ultrapassar o nível (α) de significância considerado, que foi o de 0,05, (valor-p), então a variável relacionada ao coeficiente em questão é não significativo, pode ser retirada do modelo.

O teste F verifica a significância da regressão. Este teste tem como hipótese nula que a equação de regressão não é significativa, ou seja, a equação de regressão não explica a variável dependente. A razão F é dada por:

$$F = \frac{MSR}{MSE}, \quad (4)$$

onde MSR indica a regressão média quadrática e MSE significa soma do quadrado médio (Anderson; Sweeney; Williams, 2002).

Para escolha do modelo ótimo utiliza-se algumas estatísticas que avaliam a qualidade do ajuste, como o coeficiente de determinação (R^2) e o R^2 ajustado, onde o

valor do R^2 é corrigido pelo número de coeficientes do modelo, assim em regressões múltiplas é preferível utilizar o R^2 ajustado.

Outras estatísticas utilizadas que são mais eficientes são os critérios de informação, que quanto menor o valor, melhor. Entre os critérios de informação disponíveis, optou-se pelo critério de Akaike cuja fórmula é dada por:

$$AIC = \frac{2}{n}(k - l), \quad (5)$$

onde n é o número de observações da amostra, k é o número de variáveis independentes do modelo e l é o log da verossimilhança que é uma estatística onde os erros são avaliados supondo-se uma distribuição normal.

As suposições necessárias para utilização do modelo de regressão escolhido são:

- i) O erro tem média zero e variância π^2 , desconhecida;
- ii) Os erros são não correlacionados;
- iii) Os erros têm distribuição normal;
- iv) As variáveis independentes assumem valores fixos.

As suposições (i)-(iii), podem ser representadas por $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$.

O termo multicolinearidade em uma regressão refere-se à correlação entre as variáveis independentes. Quando existe multicolinearidade na regressão, pesquisadores comprovaram que as estimativas dos parâmetros pelo método dos mínimos quadrados podem ter um sinal oposto ao do parâmetro estimado (Anderson; Sweeney; Williams, 2002). Para verificar a presença ou não de multicolinearidade utilizou-se o teste Fator de Inflação da Variância (VIF) dado por:

$$VIF = \frac{1}{1 - R_j^2}, \quad (6)$$

Em que R_j^2 é o coeficiente de determinação da regressão de umas sobre as outras variáveis independentes (EREGNO, 2013).

Utiliza-se o teste de Jarque-Bera para verificar se os erros têm distribuição normal. As hipóteses a serem testadas são: H_0 : o erro do modelo de regressão linear possui distribuição normal, contra H_1 : o erro do modelo de regressão linear possui

distribuição não-normal. O teste consiste em calcular os valores da assimetria e da curtose da amostra e utilizar a seguinte estatística de teste:

$$JB = \frac{N}{6} \left(S^2 + \frac{(K-3)^2}{4} \right) \quad (7)$$

onde S e K são, respectivamente, os coeficientes amostrais de assimetria e de curtose e N é o tamanho da amostra (FERREIRA, 2006).

Para avaliar os modelos, usamos medidas estatísticas de desempenho, que estão incluídas os Coeficiente de determinação (R^2 , R^2 ajustado (adj R^2), Estatística – F, critérios de Informação de Akaike (AIC), raiz quadrada de erro quadrático, média (RMSE).

Toda análise estatística foi realizada por meio do *software EViews 9.5 Student Version Lite*, que é uma versão gratuita. Esta versão possui algumas limitações quanto ao número de observações e variáveis envolvida, o que não impede que pesquisas no âmbito acadêmico sejam desenvolvidas.

4 RESULTADOS E DISCUSSÕES

O trabalho propõe a construção de um modelo estatístico que possa fornecer uma previsão da quantidade de bactéria *E.coli* na água do Rio Cascavel com base nos parâmetros de pH, turbidez e índice pluviométrico. Modelagens com esse tipo de dados são extremamente difíceis e complicadas porque o *E.coli* não pode ser cultivado em laboratório e não pode ser aferido após a exposição a condições extremas (EREGNO, 2013).

A estatística descritiva dos dados utilizados na modelagem é apresentada na Tabela 1, por meio da média, mediana, valor máximo e mínimo, desvio padrão, assimetria e curtose. O teste de Jarque-Bera foi utilizado para verificar se os dados da amostra têm a assimetria e a curtose correspondentes a uma distribuição normal. Como os valores das probabilidades foram menores que o nível de significância de 0,05 (p-valor < 0,05), exceto para a variável pH, entende-se as variáveis *E.coli*, Precipitação acumulada e turbidez não têm distribuição normal, logo precisa-se aplicar transformações nestas variáveis.

Tabela 1: Estatística descritiva das variáveis envolvidas no modelo.

	E.coli	Preccum	pH	Turbidez
Média	957,86	47,65	7,27	16,71
Mediana	738,50	35,40	7,25	11,25
Máximo	6488,00	315,40	7,60	84,40
Mínimo	131,00	0,00	6,90	4,98
Desvio Padrão	997,04	56,52	0,20	18,06
Assimetria	3,92	2,55	-0,09	2,80
Curtose	21,20	11,86	2,30	9,90
Jarque-Bera	785,05	209,35	1,07	157,78
p-valor	0,00	0,00	0,58	0,00

A matriz de correlação foi calculada para examinar as relações entre variáveis físico-químicas e o *E.coli*, apresentada na Tabela 2. O coeficiente de correlação linear de Pearson foi usado para verificar o grau de associação existente entre as variáveis. Correlação significativa positiva entre a variável dependente e as variáveis independentes é encontrada apenas entre *E. coli* e turbidez ($r = 0,41$, p-valor < 0,01).

A variável termo de sazonalidade consiste em $COSSENO = \frac{2\pi N}{12}$ onde n ($n=1, \dots, 12$) é o número correspondente ao mês da observação. Essa expressão descreve os efeitos sazonais e foi incluída no modelo como visto em Senior (2017).

Nota-se que a variável precipitação acumulada, pH e termo de sazonalidade (cosseno) têm correlações estatisticamente nulas com relação a variável *E.coli*, no entanto, um modelo de teste utilizando apenas a variável turbidez para descrever a variável *E.coli* não se mostrou interessante, assim as variáveis precipitação acumulada e pH foram mantidas. A variável termo de sazonalidade (cosseno) continuou a ser considerada, pois descreve a sazonalidade no modelo. Retirar esta variável não melhorou as estatísticas de avaliação do modelo.

Tabela 2: Correlação entre as variáveis *E.coli*, turbidez, precipitação acumulada, pH e termo de sazonalidade (cosseno).

Correlação/p-valor	ECOLI01C	TURBIDEZ	PRECCUM	pHC	COSSENO
E. coli01c	1,00				
Turbidez c	0,41 0,00	1,00 -----			
Preccum	0,14 0,33	0,46 0,00	1,00 -----		
pHc	0,06 0,68	-0,08 0,57	-0,07 0,26	1,00 -----	
Cosseno	-0,04 0,77	-0,27 0,06	-0,07 0,62	-0,21 0,16	1,00 -----

Foram aplicadas as seguintes transformações nas variáveis:

$$\begin{cases} LECOLI = \ln(ECOLI) \\ LPRECCUM = \ln(PRECCUM + 1) \\ LTURB = \ln(TURB + 0,1) \end{cases} \quad (8)$$

visto que o Teste Jarque-Bera (Tabela 1) mostrou que as variáveis não têm distribuição normal (p-valor < 0,05), exceto a variável pH. Para as variáveis transformadas o teste Jarque-Bera apresentou p-valor > 0,05 para as novas variáveis $|E.coli|$, $|$ precipitação acumulada e $|$ turbidez.

Foram realizadas algumas modelagens até se chegar em um modelo escolhido para determinar a concentração de *E.coli*. O modelo 1 é um modelo de regressão linear múltipla, onde as variáveis *E.coli*, precipitação acumulada, e turbidez foram transformadas por meio de logaritmos. O modelo 2 envolveu as variáveis na sua primeira diferença e termos de correção de erros, o que melhorou significativamente o

coeficiente de determinação de modelo. O modelo 3 é uma forma mais parcimoniosa do modelo 2, onde cada variável não significativa foi retirada uma por vez, todas as estatísticas recalculadas, até que se obteve uma forma simplificada do modelo, com um número menor de variáveis. Assim, o modelo 3 foi o escolhido para calcular a concentração de *E.coli*. Segue os resultados da modelagem e suas discussões.

4.1 MODELO 1

Na primeira tentativa de modelagem, ajustou-se o seguinte modelo:

$$LECOLI = b_0 + b_1LPRECCUM + b_2LTURB + b_3PH + b_4COSSENO + b_5n \quad (9)$$

Na Tabela 3 apresenta-se os coeficientes do modelo, os desvios-padrão de cada parâmetro, a estatística t para cada coeficiente e o p-valor. Observa-se que somente a variável LTURBIDEZ tem coeficiente estatisticamente não nulo ($p\text{-valor} = 0,0048 < 0,05$), logo rejeita-se a hipótese nula de que o coeficiente $b^2 = 0$. Como o coeficiente de determinação R^2 ajustado foi de 0,32 apenas, então a validação do modelo não foi concluída. E passou-se para outra simulação.

Tabela 3: Resultados do modelo 1.

Variável	Coefficiente	Desvio padrão	Teste t	p-valor
Constante	2,68	3,73	0,72	0,48
L preccum	0,08	0,06	1,35	0,18
L turbidez	0,46	0,15	2,98	0,00
pH	0,30	0,51	0,59	0,55
Cosseno	0,22	0,13	1,68	0,10
N	0,01	0,00	1,75	0,09
R^2	0,39	Prob (Estatística - F)		0,00
R^2 ajustado	0,32	Critério de Akaike		1,86
Estatística - F	5,47			

Optou-se por retirar a variável pH do modelo e o novo modelo voltou a apresentar I precipitação acumulada e termo de sazonalidade (cosseno) com coeficientes estatisticamente nulos (Tabela 4).

Tabela 4: Resultados do modelo 1 modificado

Variável	Coefficiente	Desvio padrão	T-estatística	Prob.
Constante	4,90	0,36	13,49	0,00
L preccum	0,09	0,06	1,51	0,14
L turbidez	0,44	0,15	2,95	0,00
Cosseno	0,20	0,12	1,59	0,12
N	0,01	0,00	2,09	0,04
R ²	0,39	Prob (Estatística - F)		0,00
R ² ajustado	0,33	Critério de Akaike		1,83
Estatística F				

Ao retirar as variáveis LPreccum e o cosseno do modelo, o coeficiente de determinação teve uma piora considerável, então voltou-se a investigar as variáveis e verificou-se a necessidade de trabalhar com as séries diferenciadas, pois os resultados do teste ADF (Dickey-Fuller Aumentado), apontaram que as séries são integradas de ordem 1. A hipótese nula do teste é que existe a presença de raiz unitária.

Assim, o novo modelo envolveu as séries em sua forma diferenciada, ou seja,

$$\begin{aligned}
 \Delta LECOLI_t &= LECOLI_t - LECOLI_{t-1} \\
 \Delta LTURB_t &= LTURB_t - LTURB_{t-1} \\
 \Delta PH_t &= PH_t - PH_{t-1} \\
 \Delta LPRECCUM_t &= LPRECCUM_t - LPRECCUM_{t-1}
 \end{aligned} \tag{10}$$

4.2 MODELO 2

Utilizando as variáveis diferenciadas no modelo, conseguiu-se melhorar o coeficiente de determinação, exceto para a variável do termo de sazonalidade, por ser considerada uma variável fixa. Incluímos termos de correção de erro ($LECOLI_{t-1}$, $LPRECCUM_{t-1}$, $LTURB_{t-1}$ e PH_{t-1}) para verificar se existia uma relação a longo prazo entre as variáveis e o resultado foi positivo, visto que o coeficiente da variável $LECOLI_{t-1}$ é estatisticamente não nulo (p -valor = 0,0). A fórmula do modelo 2 é dada por:

$$\begin{aligned}
 \Delta LECOLI_t &= 0,08\Delta PRECCUM_t + 0,48\Delta LTURB_t + 0,15\Delta pH_t + 0,14\text{COSSENO} + \\
 &5,55 + 0,01N - 0,06LECOLI_{t-1} + 0,11LPRECCUM_{t-1} + 0,29LTURB_{t-1} + 0,004 + pH_{t-1}
 \end{aligned} \tag{11}$$

Tabela 5: Resultados do modelo 2.

Variável	Coefficiente	Desvio Padrão	T-Estatística	Prob.
ΔL Preccum	0,08	0,07	1,13	0,27
ΔL turbidez	0,48	0,17	2,87	0,01
Δ pH	0,15	0,55	0,27	0,78
Cosseno	0,14	0,17	0,84	0,40
Constante	5,55	6,15	0,90	0,37
N	0,01	0,01	1,78	0,08
LEcoli 01 _(T-1)	-1,06	0,16	-6,52	0,00
LPreccum _(T-1)	0,11	0,11	0,99	0,33
LTurbidez _(T-1)	0,29	0,24	1,23	0,23
pH _(T-1)	0,00	0,83	0,01	0,99
R ²	0,73	Prob (Estatística F)		0,00
R ² ajustado	0,67	Crítérios de Akaike		2,01
Estatística F	11,26			

Na Tabela 5 temos o valor da estatística F de 11,26 com p-valor de 0,0, que indica a existência de uma relação entre a variável dependente ($\Delta lecoli_t$) e as variáveis independentes (fatores físico-químicos), ou seja, o modelo de regressão construído é bem determinado pelos fatores.

Como muitos coeficientes do modelo apresentaram p-valor $> 0,05$, ou seja, foram considerados estatisticamente nulos pelo teste t, passou-se a verificar uma forma mais parcimoniosa do mesmo.

4.3 MODELO 3

A fim de obter uma forma mais parcimoniosa do modelo anterior, passou-se a retirar uma variável de cada vez no modelo, que apresentava coeficiente estatisticamente nulo. A cada variável retirada, os valores das estatísticas t, F e R² ajustado eram novamente analisadas. A forma mais parcimoniosa do modelo sem prejuízo nos valores das estatísticas e coeficiente de determinação é dada por:

$$\Delta LECOLI_t = 5,755 + 0,015N + 0,528\Delta LTURB_t - 1,04LECOLI_{t-1} + 0,29LTURB_{t-1} \quad (12)$$

Na Tabela 6 são apresentados os resultados para o modelo 3. O termo de correção do erro $LTURB_{t-1}$ foi mantido, mesmo com p-valor $> 0,05$, pois retirá-lo impactava negativamente no coeficiente de determinação.

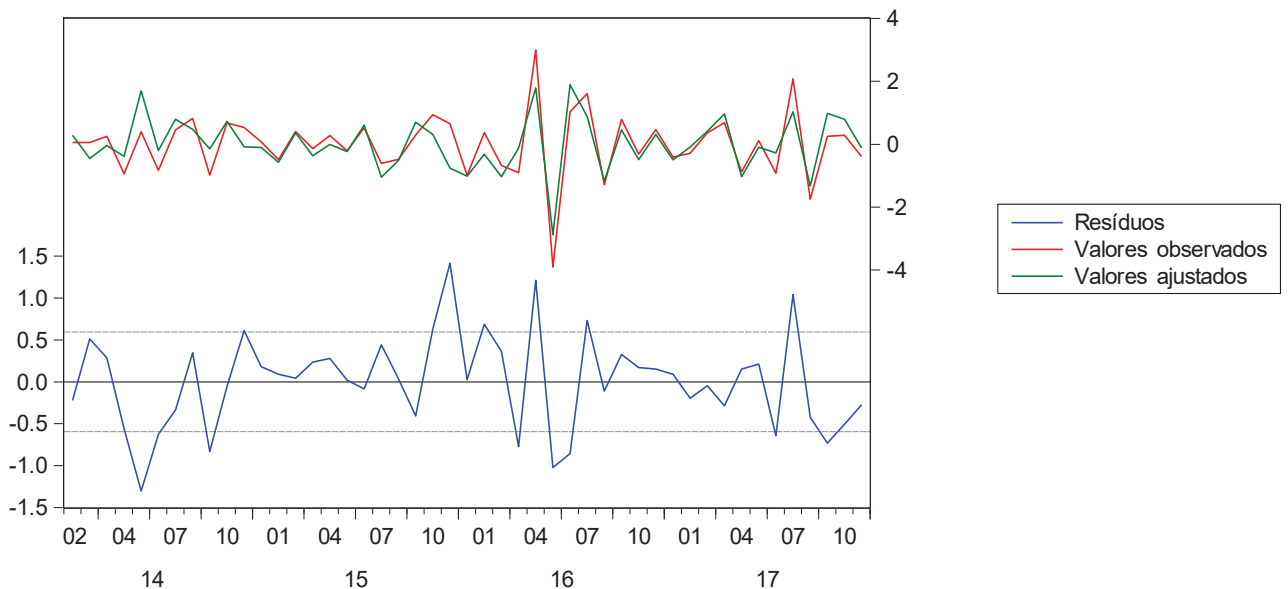
Tabela 6: Resultados do modelo 3.

Variável	Coefficiente	Desvio padrão	t-estatístico	Prob.
Δ LTurbidez C	0,52	0,13	3,98	0,00
Constante	5,75	0,91	6,31	0,00
N	0,01	0,01	2,26	0,03
LECOLI01 C (t-1)	-1,04	0,15	-6,99	0,00
LTurbidez C (t-1)	0,29	0,19	1,54	0,13
R ²	0,70	Prob (Estatística F)		0,00
R ² ajustado	0,67	Cr�terios de Akaike		1,90
Estat�stica F	24,81			

Para este modelo final, apresenta-se uma an lise completa dos resultados, bem como os valores de todos os testes de verifica o e valida o. Primeiramente, apresenta-se um gr fico com os valores observados (amostra), valores ajustados (previs es do modelo) e res duos referentes ao modelo 3.

Os res duos n o est o em torno de zero como seria o ideal, mas nos testes de aleatoriedade e normalidade os resultados foram satisfat rios como ser  apresentado no Figura 3 a seguir:

Figura 3: Res duos e valores observados e ajustados pelo modelo 3.



Primeiramente, apresenta-se na Tabela 7 o resultado do teste VIF que verifica a presen a de multicolinearidade. O modelo passou pelo teste de multicolinearidade entre as vari veis independentes, pois o VIF < 10. Se o VIF > 10 o modelo n o   confi vel, pois existe correla o entre as vari veis independente (multicolinearidade).

Tabela 7: Resultado do teste de fator de inflação da variância para o modelo 3.

Variável	Coeficiente	Centrado
	Variância	VIF
D (LTurb)	0,02	1,83
Constante	0,83	NA
N	0,00	1,12
LEcolil01 C _(T-1)	0,02	1,45
LTurbidez C _(T-1)	0,03	2,03

Quanto à análise dos resíduos, o modelo passou pelo teste de correlação serial ($p > 0,05$), ou seja, aceita-se a hipótese nula de não correlação até a defasagem de ordem 20 (Tabela 8).

Tabela 8: Saída do teste de correlação serial nos resíduos do modelo 3.

Autocorrelação	Correlação Parcial	Lag	AC	PAC	Estatística Q	p-valor
		1	-0,021	-0,021	0,0220	0,882
		2	-0,051	-0,052	0,1565	0,925
		3	0,219	0,217	2,6667	0,446
		4	-0,157	-0,161	3,9874	0,408
		5	0,085	0,117	4,3855	0,495
		6	-0,079	-0,164	4,7373	0,578
		7	-0,170	-0,083	6,3989	0,494
		8	0,128	0,055	7,3734	0,497
		9	0,119	0,199	8,2303	0,511
		10	-0,027	-0,014	8,2741	0,602
		11	-0,075	-0,131	8,6376	0,655
		12	0,187	0,190	10,940	0,534
		13	-0,083	-0,139	11,410	0,576
		14	-0,252	-0,243	15,826	0,324
		15	0,183	0,203	18,228	0,251
		16	-0,121	-0,021	19,322	0,252
		17	-0,031	-0,069	19,394	0,306
		18	-0,089	-0,283	20,022	0,332
		19	-0,285	-0,074	26,690	0,112
		20	0,144	0,020	28,469	0,099

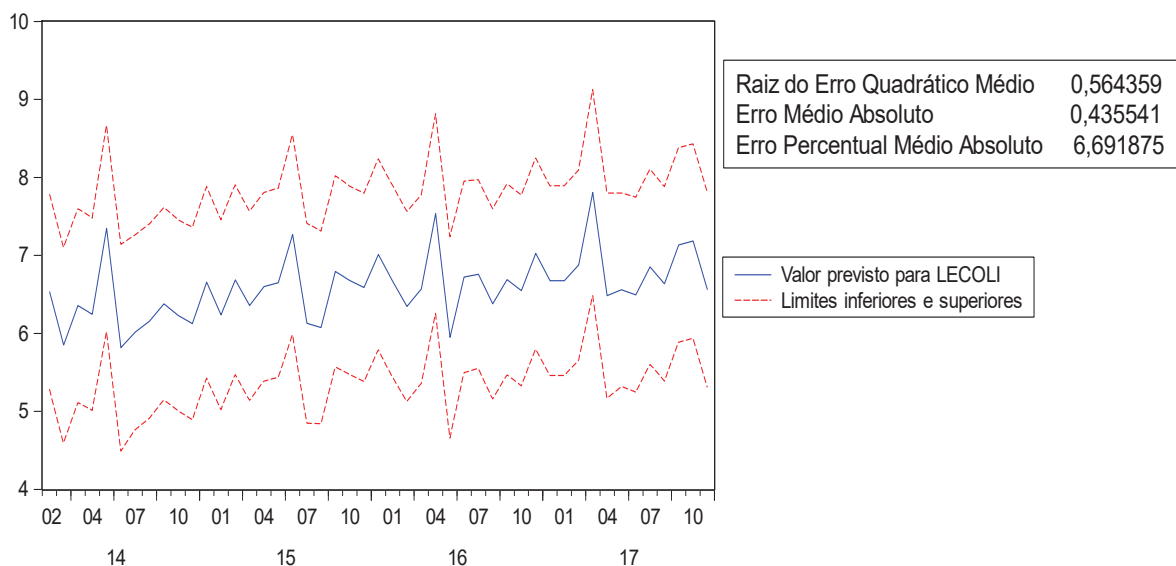
Quanto à normalidade dos resíduos, o resultado do teste de Jarque-Bera foi de 0,11, ou seja, a série dos resíduos tem distribuição normal ($p\text{-valor} = 0,95 > 0,05$).

Verificou-se também se a série dos resíduos é homocedástica. Homocedasticidade é o termo para designar variância constante dos erros experimentais para observações distintas (ANDERSON; SWEENEY; WILLIAMS, 2002). Caso a suposição de homocedasticidade não seja válida, o teste rejeitaria a hipótese nula com $p\text{-valor} < 0,05$, o que não ocorreu. O resultado do teste de Breusch-Pagan-Godfrey foi de estatística F igual a de 0,795 $p\text{-valor}$ de 0,53 $> 0,05$, ou seja, aceita-se a hipótese nula de que a série dos resíduos é homocedástica.

Por fim, utiliza-se o teste de verificação para correta especificação do modelo, denominado teste RESET. O resultado do teste aceita a hipótese nula de correta especificação do modelo ($p\text{-valor} = 0,23 > 0,05$).

Como o modelo passou por vários testes de verificação e validação, passou-se para o cálculo das previsões e dos limites inferiores e superiores ($LECOLI_t \pm 2\sigma$), bem como o cálculo dos erros de previsão no gráfico 2. Quanto as medidas de avaliação de erro, o RMSE foi de 0,57 e o MAPE foi de 6,7%. Como a medida MAPE é independente de escala, conclui-se que o modelo 3 tem um erro aceitável, mas essa é uma afirmação subjetiva, pois não houve comparação dessa medida entre modelos diferentes.

Figura 4: Valores previstos pelo modelo e limites inferiores e superiores para a previsão.



Comparado o modelo obtido escolhidos com os encontrados na literatura, o modelo R^2 ajustado foi de 0,77, contra 0,5 do modelo proposto por Eregno (2013). Já o modelo desenvolvido pelo Vijayashanthar, et al. (2018) apresentou um R^2 ajustado de 0,86, mas o mesmo também utilizou redes neurais como ferramenta além da estatística. Em Senior (2017) cujo os fatores sazonais foram incluídos e houve a necessidade de transformações logarítmicas nas variáveis selecionada, o R^2 ajustado variou de 0,61 a 0,76, selecionando variáveis diferentes no modelo. As referências citadas não apresentaram avaliação do erro então a comparação entre os valores RMSE e MAPE não foram possíveis

5. CONCLUSÃO

Quando aplicado modelo de regressão linear múltipla os resultados indicaram que as variáveis físico-químicas conseguiram explicar em torno de 70% da variação da concentração de *E.coli* para o manancial de abastecimento público de Cascavel.

Embora o modelo contenha variáveis independentes obteve-se um relacionamento entre elas o que não ocorreu de forma simples, como nas referências encontradas na literatura, mas forneceu um modelo com ajuste para o conjunto de dados quando avaliado aos vários testes estatísticos de diagnósticos e validação. Destaca-se então algumas possíveis falhas que podem ter ocorrido no processo de modelagem como a qualidade do conjunto de dados, e possivelmente falta de relação linear entre os fatores e as variáveis dependente, também a existência de fatores importantes que por ventura não foram contabilizados nos modelos. Contudo podemos concluir que existe correlação significativa entre a *E.coli* e a turbidez, onde a mesma é influenciada pela precipitação, mas a precipitação não influencia a *E.coli*.

Como trabalho futuro, pretende-se explorar outros dois conjuntos de dados provenientes de outros mananciais que abastecem a cidade de Cascavel, sendo eles o rio Saltinho e o rio Peroba.

Por fim, o trabalho proporcionou um estudo sobre a influência das variáveis, pH, turbidez e índice pluviométrico, no crescimento dos coliformes, o conhecimento de modelos de regressão e estatísticas, conceitos não vistos na graduação que podem ser utilizados em dados de outra natureza, ampliando assim as possibilidades de pesquisa.

REFERÊNCIAS

- ANDERSON, David R.; SWEENEY, Dennis J.; WILLIAMS, Thomas A. **Estatística aplicada à administração e economia**. São Paulo: Pioneira Thomson Learning, 2002.
- ANDREOLI, C.V.; WILLER, M. (Ed.) **Management of sanitation in planned communities**. Curitiba: AlphaVille Urbanismo, 2005.
- APHA; AWWA; WEF. **Standard methods for the examination of water and wastewater**. 21^o ed. Washington: APHA, 2005.
- BACCI, Denise de La Corte, et al. **Educação para a água**. Revista estudos avançados ed. 22^a, 2008.
- BRASIL, Conselho Nacional de Meio Ambiente. **Resolução 357**, de 17 de março de 2005. Dispõe sobre a classificação e enquadramento dos corpos de água. Disponível em http://pnqa.ana.gov.br/Publicacao/RESOLUCAO_CONAMA_n_357.pdf
- BRASIL, Portaria de Consolidação nº 5 de setembro de 2017: **Do controle e da vigilância da qualidade da água para consumo humano e seu padrão de potabilidade**: Ministério da Saúde, Brasília, 2017.
- BRASIL, Fundação nacional da Saúde. **Manual prático de análise de água/fundação nacional da saúde - 4^a Ed.** - Brasília: Funasa, 2013.
- BRAGA, B. et al. **Introdução a engenharia ambiental**. São Paulo: Prentice Hall/grupo Person Education, 2002.
- BRAGA, B.; HESPANHOL, I.; CONEJO, J. G. L. et al. **Introdução à Engenharia Ambiental. O desafio do desenvolvimento sustentável**. 2 ed. São Paulo: Prentice Hall, 2005.
- CETESB, Companhia Ambiental do Estado de São Paulo. **Relatório de Qualidade das Águas Interiores do Estado de São Paulo - Apêndice A - Significado ambiental e sanitário das variáveis de qualidade das águas e dos sedimentos e metodologias analíticas e de amostragem**, 2008. Disponível em: <<http://www.cetesb.sp.gov.br/Agua/rios/publicacoes.asp>>. Acesso em: 10 out. 2017.
- CHANG, Raymond. **Físico-Química para ciências químicas e biológicas**. Amgh editora, vol. 2, 3^a edição, 2010. Disponível em: <<http://www.R-project.org/>>. Acesso em: 20 de out. De 2017.
- CORADI, P.C.; FIA, R.; PEREIRA- RAMIREZ, O. **Avaliação da qualidade da água superficial dos cursos de água do município de Pelotas-RS**. Ambi-Água, Taubaté, v. 4, n. 2, p. 46-56, 2009. Disponível em: <www.ambi-agua.net/seer/index.php/ambi-agua/article/download/.../331>. Acesso em: 09 out. 2017.

COMPANHIA DE SANEAMENTO DO PARANÁ – SANEPAR. **Excelência na gestão e operação dos laboratórios descentralizados de água**. 5ª ed. Paraná. 2014.

DANTAS, T. N. P. **Avaliação da qualidade das águas da bacia hidrográfica do Rio Pirangi/RN**. Monografia (Curso de Tecnologia em Controle Ambiental) – Centro Federal de Educação Tecnológica do Rio Grande do Norte, Natal, 2008.

ELERIA, Anna; VOGEL, Richard M. **PREDICTING FECAL COLIFORM BACTERIA LEVELS IN THE CHARLES RIVER, MASSACHUSETTS, USA**. 1. JAWRA Journal of the American Water Resources Association, v. 41, n. 5, p. 1195-1209, 2005.

EREGNO, Fasil Ejigu. **Multiple linear regression models for estimating microbial load in a drinking water source case from the Glomma river, Norway**. Dissertação de Mestrado. Norwegian University of Life Sciences, Ås. 2013.

FERREIRA, M. Análise da sensibilidade dos testes de normalidade de jarque-bera e lilliefors em modelos de regressão linear. **Rev. Mat. Estat**, v. 24, n. 4, p. 89-98, 2006.

FRAGA, Rosimeri Silva da. **Influência da precipitação sobre parâmetros de qualidade da água utilizada para irrigação do morangueiro no município de turuçu-rs**. Revista brasileira de agrociência. Pelotas, v.18 n. 1-4, p.81-94, jan-mar, 2012. Disponível em: <<https://periodicos.ufpel.edu.br/ojs2/index.php/CAST/article/download/2491/2324>>. Acesso em: <13 de outubro de 2017>

KOBIYAMA, M. et al. **Recursos hídricos e saneamento**. Curitiba: Ed. Organic Trading, 2008.

LIBÂNIO, M. **Fundamentos de qualidade e tratamento de água**. Campinas: Áto-mo, 2005.

MCCULLOCH, Kaitlin J. **Analysis of Relationships Between Water Quality Parameters and Stream Sediment with Fecal Bacteria in Hidden Creek, Rock Hill, SC**. 2015.

NAVIDI, W. **Probabilidade e Estatística para Ciências Exatas**. Porto Alegre: AMGH, 2012.

OLIVEIRA, E. et al. **Caracterização morfometria da bacia hidrográfica do Rio Cascavel**. Guarapuava/PR, 2009. Disponível em: <revistas.unicentro.br/index.php/analecta/article/download/2098/1801>Acesso em: 14 de outubro de 2017.

PIVELI, R.P. **“Qualidade da Água”**. Apostila do Curso de Especialização em Engenharia em Saúde Pública e Ambiental da Fac. Saúde Pública – USP, 1996.

PORTAL TRATAMENTO DE ÁGUA, 2009. Disponível em: <<https://www.tratamento-deagua.com.br/artigo/qualidade-da-agua/>>. Acessado em: 04/12/2018.

R Development Core Team R: **A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2011.

SENIOR, Lisa A. **Estimated fecal coliform bacteria concentrations using near real-time continuous water-quality and streamflow data from five stream sites in Chester County**, Pennsylvania, 2007–16. US Geological Survey, 2017.

SECRETARIA DE ESTADO DO MEIO AMBIENTE E RECURSOS HÍDRICOS (SEMA). **Bacias Hidrográficas do Paraná**. Curitiba, 2010.

SILVA, Neusely et al. **Manual de métodos de análise microbiológica da água**. São Paulo: Livraria Varela, 2005.

SISTEMA NACIONAL DE INFORMAÇÕES SOBRE O SANEAMENTO, 2018. Disponível em: <<http://www.snis.gov.br/diagnostico-agua-e-esgotos/diagnostico-ae-2016>>. Acessado em: 04/12/2018.

SPERLING, Marcos Von. **Introdução à qualidade das águas e ao tratamento de esgotos**. 3ª Edição. Princípios do tratamento biológico de águas residuárias. Editora da Universidade Federal de Minas Gerais – UFMG – Belo Horizonte, 2005.

TORTORA, Gerard Jerry. **Microbiologia**. Editora Saraiva, 10ª edição, 2012.

TOSIN, G. **Caracterização Física Do Uso E Ocupação Da Bacia Hidrográfica Do Rio Cascavel**, 2005. Disponível em:<http://tede.unioeste.br/bitstream/tede/2842/1/Gladis_Tosin%202005.pdf> Acesso em: 14 de outubro de 2017.

WALPOLE, Ronald E. **Probability & statistics for engineering and science**. Pearson Prentice Hall, 2009.

VIJAYASHANTHAR, Vasikan et al. Modeling Fecal Indicator Bacteria in Urban Waterways Using Artificial Neural Networks. **Journal of Environmental Engineering**, v. 144, n. 6, 2018.