

PROPOSTA DE CALIBRAÇÃO DE UM MODELO MULTI-PARÂMETRO DE QUALIDADE DE ÁGUA UTILIZANDO O ALGORITMO GENÉTICO

Jonas Heitor Kondageski¹ & Cristóvão Vicente Scapulatempo Fernandes²

RESUMO --- No presente artigo foram desenvolvidas e empregadas rotinas de algoritmo genético para calibração de um modelo de qualidade da água para um rio hipotético dividido em trechos. Para isto, definiu-se uma função-objetivo, composta pela soma quadrática das diferenças entre as concentrações medidas e calculadas de DBO e OD. Foram definidos 5 cenários, onde empregou-se diferentes números de pontos de monitoramento para a calibração do modelo. Além disso, o critério para o término da execução do algoritmo genético foi diferente entre estes cenários. Para cada cenário foram realizadas 200 simulações, e em cada simulação obteve-se um conjunto diferente de constantes de qualidade da água que satisfizeram o problema de calibração. Observou-se que para aqueles cenários onde havia maior número de pontos de monitoramento as constantes de qualidade calculadas apresentaram forte relação linear entre si. Além disso, observou-se que as concentrações de DBO e OD calculadas com as médias das constantes aproximaram-se dos valores monitorados destas concentrações. A principal contribuição deste artigo é apresentar os resultados de calibração utilizando o algoritmo genético com um indicativo potencial de solução do problema de otimização ou de definição do valor dos parâmetros de Qualidade da Água de forma consistente e robusta e função do conjunto de informações de monitoramento disponíveis.

ABSTRACT --- Genetic algorithm routines were developed and applied to a water quality model calibration of a hypothetical river. An objective function was defined as the squared difference between gauged and calculated BOD and DO to achieve that purpose. There were 5 calibration scenarios, in each of them was used a particular number of field measurements for the calibration processes; besides, the ending criteria was different for the scenarios. In each scenario 200 simulations were carried out. For each simulation different water quality constant sets were found which were able to adjust the gauged concentration values to the calculated ones. It was presented that the BOD and DO concentrations calculated by an average constant set got near the gauged concentrations. The main contribution of this paper is to present the calibration results using GA technique as the potential real solution to the optimization Water Quality Parameters problem based upon the number of water quality data available.

Palavras-chave: Algoritmo genético, modelo de qualidade de água, calibração.

1) Engenheiro Ambiental, Mestrando - Departamento de Hidráulica e Saneamento - DHS - Centro Politecnico s/n, Bloco V - Jardim das Américas - Curitiba-PR-Brasil - CEP: 81531-990, Caixa Postal: 19011, Telefone: 0xx 41 361 3142 - FAX/Fone: 361-3143, E-mail: jonasheitor@yahoo.com.br
2) Professor Adjunto - Departamento de Hidráulica e Saneamento - DHS - Centro Politecnico s/n, Bloco V - Jardim das Américas - Curitiba-PR-Brasil - CEP: 81531-990, Caixa Postal: 19011, Telefone: 0xx 41 361 3142 - FAX/Fone: 361-3143, E-mail: cris.dhs@ufpr.br

1 - INTRODUÇÃO

A modelagem de qualidade de água constitui uma ferramenta que apresenta um papel muito importante na gestão de recursos hídricos. Ela pode ser aplicada, por exemplo, no auxílio à tomada de decisão de outorga de direitos de uso da água, além da outorga de lançamento de efluentes por indústrias. Pois através dos modelos matemáticos de qualidade da água é possível prever os impactos num sistema aquático, como um rio, causados pelo lançamento de efluentes industriais ou domésticos, pela descarga acidental de algum contaminante, ou mesmo o impacto relacionado a uma situação crítica de vazão (Nahom, 2006).

Porém, para que se aproveite todo o potencial de aplicação dos modelos matemáticos de qualidade da água, é necessário que eles descrevam da forma mais fiel possível a realidade do rio sob estudo. Esta realidade, que está relacionada às características físicas, químicas e biológicas do rio, é descrita pelas constantes de qualidade da água do modelo, que por sua vez devem ser ajustadas de forma a representar estas características do rio. Este processo de ajuste das constantes é denominado calibração do modelo de qualidade da água. Existem alguns métodos para a calibração como, por exemplo, a calibração manual. Este representa um método pouco prático e que consome muito tempo, principalmente se o modelo a ser calibrado apresentar grande número de constantes de qualidade de água.

Neste contexto, o Algoritmo Genético representa outro método que pode ser aplicado na calibração de modelos. O Algoritmo Genético constitui uma técnica computacional de otimização que baseia o seu funcionamento numa analogia da Teoria da evolução das espécies, declarada pelo fisiologista inglês Charles Darwin. Neste algoritmo os indivíduos (possíveis soluções do problema de otimização) evoluem através da aplicação dos operadores genéticos. Se o algoritmo foi bem elaborado, esta evolução conduzirá à formação de um indivíduo ótimo, ou seja, à resposta ótima do problema de otimização. Apesar do algoritmo genético representar uma técnica relativamente nova, vários autores já o empregaram em problemas de otimização: Wang (1991); Cho et al (1999); Wardlaw e Sharif (1999); Perera e Ng (2003).

O potencial de aplicação do Algoritmo Genético na calibração de modelos se deve ao fato da calibração poder ser tratada como um problema de otimização, pois se deseja minimizar a função soma dos resíduos quadráticos, ou seja, a soma das diferenças quadráticas entre a concentração de um poluente qualquer medida e aquela calculada pelo modelo. Desta forma empregou-se, no presente trabalho, o Algoritmo Genético para calibração de modelos de qualidade de água para rios. Um modelo de qualidade de água apresentado por Chapra (1997) foi programado de forma a funcionar acoplado às rotinas de algoritmo genético (módulo de calibração), que também foram programadas. Assim, o modelo de qualidade da água fornecia dados calculados de concentração

para o módulo de calibração, onde era realizado o processo de busca das constantes do modelo de qualidade de forma a ajustar as concentrações medidas às calculadas.

2 – REVISÃO BIBLIOGRÁFICA

2.1 - Utilização do Algoritmo Genético na literatura

Por ser uma técnica de otimização, o algoritmo genético apresenta um potencial de aplicação muito amplo em várias áreas da ciência e da engenharia, pois vários são os casos onde se está interessando em otimizar alguma função, como, por exemplo, minimizar o custo econômico ou a degradação ambiental. Um dos trabalhos desenvolvidos nesta linha de pesquisa foi o de Cho et al (1999) que acoplaram o algoritmo genético a um modelo de gerenciamento de qualidade da água, com o objetivo de maximizar a qualidade de água, adicionalmente à minimização dos custos de tratamento do esgoto lançado no rio Youngsan, um dos mais poluídos da Coréia do Sul, cuja bacia hidrográfica apresenta uma área total de 3374 km², e comporta uma população de 1,7 milhões. Após a execução deste modelo encontraram uma nova configuração das estações de tratamento que otimizava os custos e a qualidade da água. Em outro exemplo de aplicação, Wardlaw e Sharif (1999), aplicaram o algoritmo genético na otimização de operação de reservatórios em tempo real com previsão estocástica de vazões. Foi resolvido o problema dos 4 reservatórios, que apresenta solução conhecida através de técnicas de programação linear, e desta forma pode-se comparar com os resultados obtidos pelo algoritmo genético. Segundo os autores o algoritmo genético convergiu rapidamente para a solução conhecida, além de ser mais facilmente adaptável a problemas não lineares. Wang (1991) aplicou a metodologia do algoritmo genético na calibração de um modelo conceitual chuva-vazão com os dados da represa Bird Creek, localizada em Oklahoma, Japão. Havia 7 parâmetros no modelo que deveriam ser encontrados através desta calibração, que foi realizada pela minimização da função de mínimos quadrados. Resultados razoáveis foram obtidos na calibração pelo algoritmo genético. O autor concluiu que o algoritmo genético representou um método eficiente e robusto na calibração do modelo.

2.2 – Modelo de qualidade de água

O modelo de qualidade de água empregado foi apresentado por Chapra (1997). Este modelo é constituído por uma equação para o cálculo da demanda bioquímica de oxigênio (DBO) e outra equação para a concentração de oxigênio dissolvido na água do rio (OD). Estas equações estão apresentadas a seguir.

$$L = L_0 e^{-\frac{x}{u} k_r} + \frac{S_L}{k_r} \left(1 - e^{-\frac{x}{u} k_r} \right) \quad (1)$$

A constante k_r (T^{-1}) ($k_r = k_d + k_s$) representa a taxa de remoção total de matéria orgânica; k_d (T^{-1}) representa a taxa de remoção de matéria orgânica devido à decomposição pelos microorganismos; k_s (T^{-1}) representa a constante de remoção de matéria orgânica devido ao efeito da sedimentação, ou taxa de deposição; L_0 representa a concentração de DBO carbonácea inicial (ML^{-3}); u é a velocidade; x é a distância; S_L representa a carga difusa de DBO ($ML^{-3}T$).

$$D(x) = D_0 e^{\frac{x}{u} k_a} + \frac{L_0 k_d}{k_a - k_r} \left(e^{\frac{x}{u} k_r} - e^{\frac{x}{u} k_a} \right) + \frac{1}{k_a} \left(\frac{S_L k_d}{k_r} + \frac{S_b}{H} \right) \left(1 - e^{\frac{x}{u} k_a} \right) - \frac{S_L k_d}{k_r (k_a - k_r)} \left(e^{\frac{x}{u} k_r} - e^{\frac{x}{u} k_a} \right) \quad (2)$$

A variável D representa o déficit de oxigênio (ML^{-3}), isto é, a concentração de saturação menos a OD; D_0 o déficit inicial de oxigênio (ML^{-3}); K_a a constante de reeração (T^{-1}); S_b a demanda de oxigênio pelo sedimento (SOD) ($ML^{-2}T^{-1}$); e H a profundidade do rio (L). Para se chegar às equações apresentadas acima se considerou que a carga difusa insere massa de DBO no rio, mas não insere água.

3 – METODOLOGIA

3.1 – Calibração de modelos de qualidade de água

Observa-se nas equações (1) e (2) que a concentração de DBO e OD depende de algumas constantes de qualidade da água, como a taxa de sedimentação de matéria orgânica, taxa de reeração atmosférica. Porém, estas constantes são dependentes das características físicas, químicas, e biológicas do rio estudado; também dependem do tipo de poluente lançado e das condições atmosféricas. Desta forma, cada rio apresenta um conjunto particular de constantes de qualidade de água que o diferencia dos demais. Portanto, quando se pretende modelar algum rio, é necessário determinar, de alguma forma, estas constantes. A determinação das constantes é denominada calibração, e representa uma importante etapa no processo de desenvolvimento e implementação de um modelo de qualidade de água para um rio qualquer.

Como as características podem variar no tempo para um mesmo rio, o processo de calibração deve ser realizado com certa frequência, o que manterá atualizada as constantes do modelo que descrevem o comportamento do rio em relação à qualidade da água.

Existem algumas maneiras de se calibrar o modelo. Uma das formas mais conhecidas consiste no método de tentativa e erro, onde o modelador estima um valor que ele considera aceitável para a constante, roda o modelo e compara a resposta obtida com algum valor monitorado. Caso os valores calculados e monitorados estejam suficientemente próximos, considera-se o modelo calibrado; caso contrário, estipula-se outro valor para a constante e novamente testa-se a adequação desta rodando o modelo e comparando os valores calculados e medidos. Este processo iterativo prossegue até que o

modelo seja considerado calibrado. Obviamente este método apresenta dificuldade quando se deseja calibrar um modelo com grande número de constantes. Nestas situações o número de combinações possíveis entre os parâmetros cresce demasiadamente, o que torna o processo de calibração prolongado e cansativo.

A matemática pode auxiliar no processo de calibração. Por exemplo, defini-se a função: “soma quadrática dos resíduos”. Esta função é constituída pela soma quadrática das diferenças entre os valores de concentração medidas em campo e aqueles calculadas pelo modelo. Definida esta função, o objetivo consiste em minimizá-la, ou seja, encontrar as constantes tais que a concentração calculada pelo modelo torne-se mais próxima possível da concentração medida. Para minimizar esta função é possível lançar mão de alguns métodos de programação matemática não linear, como o método do gradiente, Newton, entre outros (Gonzaga, 2004). Os métodos heurísticos, onde se encaixa o algoritmo genético, constitui outra forma de minimizar a função.

A equação (3) representa a função “soma quadrática dos resíduos”, também denominada função-objetivo na linguagem de otimização, ou função-aptidão na linguagem dos algoritmos genéticos.

$$f = \sum_{i=1}^{nm} \left(w(i) [\hat{O}_D(i) - OD(i)]^2 + [1 - w(i)] [\hat{L}(i) - L(i)]^2 \right) \quad (3)$$

A variável $\hat{L}(i)$ representa a concentração medida de DBO; $L(i)$ a concentração calculada de DBO; $\hat{O}_D(i)$ a concentração medida de OD; $OD(i)$ a concentração calculada de OD; $w(i)$ o peso relativo da DBO e OD na função aptidão; e nm o número de pontos de concentração medida.

Assim, a aptidão do indivíduo é tanto melhor quanto mais próximo de zero for sua função aptidão, pois mais próximo estarão os dados calculados dos medidos em campo. No processo de calibração, o algoritmo genético será empregado para a minimização desta função.

3.2 – Estudo de Caso - Rio hipotético

Aplicaram-se as rotinas desenvolvidas de algoritmo genético na calibração de um rio hipotético onde as constantes do modelo de qualidade de água eram conhecidas. Assim, era possível compará-las com as constantes encontradas pelo algoritmo genético, e então testar o desempenho do algoritmo genético em relação ao processo de calibração ou otimização.

Este rio hipotético foi apresentado por Chapra (1997). Ele apresenta 100 km de comprimento, e está dividido em 3 trechos, onde as características hidráulicas, bioquímicas e morfológicas são constantes. Além disso, os trechos foram divididos em elementos computacionais com comprimento de 5 km cada. Um diagrama topológico deste rio está apresentado na Figura 1.

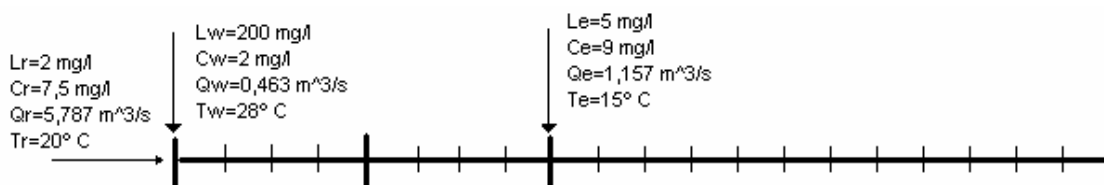


Figura 1- Diagrama topológico do Rio hipotético.

É possível observar uma descarga pontual de esgoto no início do trecho de rio considerado. Além disso, um rio afluente descarrega neste rio principal, na posição 40 km a partir do início do rio. Ao longo deste rio não existe carga difusa de matéria orgânica. A Tabela 1 apresenta as características as características do rio.

Tabela 1- Características do rio hipotético.

	1º Trecho	2º Trecho	3º Trecho
Comprimento (km)	20	20	60
Profundidade (m)	1.24	1.24	1.41
Área (m²)	15.5	15.5	18.05
Vazão (m³/s)	6.25	6.25	7.407
Velocidade (m/s)	0.403	0.403	0.410
Temperatura (°C)	20.59	20.59	19.72
C_{sat} saturação (mg/l)	8.987	8.987	9.143
K_a (dia⁻¹)	1.842	1.842	1.494
K_s (dia⁻¹)	0.25	0.0	0.0
K_d (dia⁻¹)	0.514	0.514	0.494
K_r (dia⁻¹)	0.764	0.514	0.494

Considerou-se um regime permanente, tanto em relação ao escoamento quanto em relação à qualidade da água. Esta aproximação não é necessariamente verdade; principalmente em relação às constantes de qualidade da água, que são afetadas fortemente pela temperatura que, por sua vez, apresenta variações significativas entre o dia e a noite. Além disso, outras condições atmosféricas podem afetar o valor destas constantes. A concentração de saturação que influencia a capacidade e velocidade da reaeração da água do rio, também é influenciada pela temperatura. Também é possível verificar que a remoção de matéria orgânica por efeito de sedimentação foi considerada importante somente no primeiro trecho do rio, ou seja, k_s é igual à zero no segundo e terceiro trechos. Além disso, não se considerou o termo da demanda bentônica de oxigênio ($S_b = 0$).

Possuindo as informações apresentadas na Tabela 1, utiliza-se as equações (1) e (2) para se obter os perfis de concentração de matéria orgânica (DBO) e oxigênio dissolvido (OD) ao longo do rio.

Para realizar a calibração foi percorrido o caminho inverso, ou seja, consideram-se as concentrações calculadas de DBO e OD como sendo “monitoradas”. Desta forma, calibra-se o modelo de qualidade da água lançando mão de algumas destas concentrações “monitoradas”. Como existem três constantes de qualidade de água a serem calibradas (k_d , k_a , k_s) e, além disso, estas

constantes podem variar para cada um dos três trechos do rio, a calibração corresponde a um problema de otimização de uma função com 9 variáveis (equação 3).

Para a calibração do modelo de qualidade de água se executou o algoritmo genético com os seguintes operadores genéticos: Representação real dos parâmetros; população inicial gerada aleatoriamente; seleção por torneio, onde 2 indivíduos são escolhidos aleatoriamente entre a população, e o melhor deles é selecionado para o cruzamento, ou o pior entre eles é eliminado da população; cruzamento aritmético; mutação uniforme. Por outro lado, os parâmetros genéticos utilizados foram os seguintes: População composta por 30 indivíduos; taxa de cruzamento de 100%; taxa de mutação é variável de acordo com o comprimento dos cromossomos; 20 indivíduos novos serão produzidos em cada geração; o programa termina sua execução após ter sido alcançado um resíduo mínimo estabelecido pelo usuário; substituição steady-state; o espaço de busca: $0 < k_d(\text{dia}^{-1}) < 1.6$; $0 < k_a(\text{dia}^{-1}) < 5.0$; $0 < k_s(\text{dia}^{-1}) < 1.6$ (Kondageski e Fernandes, 2005). Como critério para o encerramento da execução do algoritmo genético, ou critério de parada, foi definido que o valor da função aptidão deveria ser menor ou igual a algum valor definido.

A calibração foi realizada para 5 cenários que se diferenciavam entre si pelo número de pontos de monitoramento utilizados no processo de calibração. Além disso, o valor mínimo que a função aptidão (equação 3) deveria assumir, variou entre estes cenários. Para os cenários onde havia grande número de pontos de monitoramento (1 e 2) o critério de parada foi menos restritivo do que para aqueles cenários com poucos pontos. Pois nos cenários 1 e 2 o algoritmo genético apresentou dificuldade em convergir para um erro menor do que aqueles estabelecidos. As características de cada cenário de calibração estão apresentadas na Tabela 2.

Tabela 2 - Cenários de Calibração.

Cenário	Nº de Pontos de Monitoramento	Critério de Parada ($f \leq \dots$)
1	20	0.05
2	10	0.005
3	3	0.00001
4	2	0.00001
5	1	0.00001

No **cenário 1** procurou-se simular uma situação teórica onde haveria muitas informações de qualidade da água disponíveis: um ponto de monitoramento a cada 5 km do rio. Obviamente, este cenário representa uma situação puramente teórica, tendo o objetivo de avaliar o efeito do número de informações de qualidade de água sobre a calibração do modelo, uma vez que na prática o monitoramento custa caro, o que inviabiliza economicamente a obtenção de um número tão grande de informações. O **cenário 2** também apresentou muitos pontos de monitoramento (10 pontos), que estavam distribuídos entre os 3 trechos do rio. Os **cenários 3, 4 e 5** se apresentaram mais realistas em relação ao número de pontos de monitoramento existentes, ou seja, de informações de qualidade

de água disponíveis. Em relação ao **cenário 3**, os 3 pontos de monitoramento foram distribuídos ao longo dos trechos, um em cada trecho. No **cenário 4** os 2 pontos de monitoramento estavam localizados no segundo e terceiro trecho. Finalmente, em relação ao **cenário 5**, o mais desfavorável em relação à disponibilidade de informação de qualidade de água, o único ponto de monitoramento estava localizado no final do rio.

Muitos processos aleatórios estão envolvidos no método do algoritmo genético, como, por exemplo, a geração da população inicial, a escolha dos indivíduos para o cruzamento e morte, e a escolha dos genes que sofrerão mutação. Estes processos, por sua vez, dependem dos números aleatórios uniformemente distribuídos que são gerados pelo computador, os chamados números pseudo-aleatórios. Desta forma, a geração destes números representa um dos aspectos mais importantes na execução do algoritmo genético. A geração dos números pseudo-aleatórios, por sua vez, depende de uma semente inicial que é passada como parâmetro de entrada para a função de geração destes números; assim, ao empregar-se a mesma semente, obtém-se o mesmo conjunto de números pseudo-aleatórios uniformemente distribuídos. Para se evitar esta repetição, a semente utilizada é igual ao número de segundos - com precisão de duas casas decimais - decorridos desde zero hora do dia até o momento em que o programa é executado. Assim, garante-se que conjuntos diferentes de números pseudo-aleatórios uniformemente distribuídos são gerados, assegurando que as populações iniciais, os cromossomos selecionados para o cruzamento, mutação e morte não serão, necessariamente, os mesmos. Com isso, pode-se verificar a robustez do algoritmo, uma vez que partindo de populações iniciais diferentes, e apresentando diferentes cromossomos selecionados, o algoritmo genético deve convergir para a solução do problema. Porém, apesar do algoritmo genético convergir para qualquer execução do programa ou, de forma equivalente, para qualquer semente inicial, esta convergência não ocorrerá, obrigatoriamente, após o mesmo número de gerações.

Assim, para testar o desempenho do algoritmo genético em relação à semente inicial, foram realizadas, para cada um dos 5 cenários de calibração, 200 simulações do programa.

4 – RESULTADOS E DISCUSSÕES

4.1 – Resultados Gerais

Observou-se que para cada uma das 200 simulações, e para cada cenário de calibração, o algoritmo genético encontrou uma combinação diferente das constantes de qualidade de água que minimizaram a função objetivo (equação 3), ou seja, que calibra o modelo de qualidade de água. A Tabela 3 à Tabela 7 apresenta alguns parâmetros estatísticos das constantes de qualidade de água encontradas pelo algoritmo genético, separadas por trecho do rio e pelo número do cenário.

Tabela 3 - Resultados de Calibração para o cenário 1.

Trecho		Mínimo	média	Mediana	Máximo	Desv Pad	
Cenário 1	1	K _d	0,1981	0,4187	0,4137	0,6219	0,0881
		K _a	0,1873	1,3572	1,3366	2,3912	0,4503
		K _s	0,1267	0,3339	0,3383	0,5673	0,0908
		K _r	0,7369	0,7526	0,7490	0,7847	0,0111
	2	K _d	0,1685	0,3409	0,3317	0,5367	0,0845
		K _a	0,6125	1,2837	1,2694	1,9503	0,2807
		K _s	0,0172	0,1926	0,1902	0,3651	0,0806
		K _r	0,4626	0,5335	0,54203	0,5663	0,0243
	3	K _d	0,1076	0,3297	0,3352	0,5133	0,0796
		K _a	0,4999	1,0757	1,0824	1,5799	0,2086
		K _s	0,0020	0,1610	0,1539	0,3800	0,0763
		K _r	0,4703	0,49077	0,4881	0,5189	0,0114

Tabela 4 - Resultados de calibração para o cenário 2.

Trecho		Mínimo	média	mediana	Máximo	Desv Pad	
Cenário 2	1	K _d	0,3558	0,4477	0,4405	0,6370	0,0632
		K _a	0,9503	1,4701	1,4319	2,5124	0,3527
		K _s	0,1264	0,3138	0,3202	0,4087	0,0646
		K _r	0,7516	0,7615	0,7628	0,7755	0,0046
	2	K _d	0,1411	0,3562	0,3578	0,5290	0,0797
		K _a	0,7110	1,3630	1,3718	1,9445	0,2584
		K _s	0,0009	0,1619	0,1630	0,3709	0,0789
		K _r	0,4953	0,5181	0,5158	0,5347	0,0085
	3	K _d	0,2377	0,3437	0,3405	0,4873	0,0599
		K _a	0,8307	1,1038	1,0935	1,4661	0,1551
		K _s	0,0021	0,1498	0,1518	0,2575	0,0604
		K _r	0,4865	0,4935	0,4944	0,5039	0,0029

Tabela 5 - Resultados de calibração para o cenário 3.

Trecho		Mínimo	média	mediana	Máximo	Desv	
Cenário 3	1	K _d	0,2038	0,4325	0,4381	0,6852	0,0943
		K _a	0,1501	1,4044	1,4409	2,7156	0,5044
		K _s	0,0781	0,3310	0,3253	0,5596	0,0943
		K _r	0,7632	0,7635	0,7634	0,7648	0,0004
	2	K _d	0,1401	0,3603	0,3657	0,5147	0,0849
		K _a	0,6851	1,3736	1,3919	1,9302	0,2829
		K _s	0,0008	0,1547	0,1493	0,3746	0,0847
		K _r	0,5125	0,5149	0,5152	0,5156	0,0008
	3	K _d	0,1221	0,3116	0,3241	0,4608	0,0767
		K _a	0,5730	1,0506	1,0810	1,4163	0,1904
		K _s	0,0331	0,1821	0,1697	0,3716	0,0767
		K _r	0,4934	0,4938	0,4938	0,4946	0,0002

Tabela 6 - Resultados de calibração para o cenário 4.

Trecho		Mínimo	média	mediana	Máximo	Desv Pad	
Cenário 4	1	K _d	0,03556	0,31199	0,29251	0,696	0,12734
		K _a	0,33564	1,5978	1,50721	4,49234	0,70038
		K _s	0,04378	0,26108	0,25816	0,53582	0,0916
		K _r	0,19211	0,573077	0,57476	0,95837	0,13163
	2	K _d	0,13164	0,46654	0,45422	0,85758	0,146566
		K _a	0,28432	1,44167	1,39937	2,70449	0,46371
		K _s	0,05659	0,30171	0,28877	0,70185	0,12636
		K _r	0,25423	0,76824	0,76582	1,27726	0,17552
	3	K _d	0,09182	0,3032356	0,31453	0,46712	0,08125
		K _a	0,54562	1,0286	1,05012	1,43148	0,198989
		K _s	0,01798	0,17172	0,16226	0,40126	0,080695
		K _r	0,4363	0,47496	0,47497	0,51391	0,013311

Tabela 7 - Resultados de calibração para o cenário 5.

Trecho		Mínimo	média	mediana	Máximo	Desv Pad	
Cenário 5	1	K_d	0,0457	0,3817	0,3689	1,2011	0,1687
		K_a	0,0097	2,0461	2,0915	3,9385	0,8466
		K_s	0,0532	0,3404	0,3406	0,8425	0,1422
		K_r	0,1320	0,7222	0,7138	1,4088	0,2050
	2	K_d	0,1277	0,4183	0,4006	0,9517	0,1589
		K_a	0,0240	1,6885	1,5973	4,4131	0,8478
		K_s	0,0229	0,3620	0,3351	0,8984	0,1488
		K_r	0,3091	0,7802	0,7668	1,5831	0,2225
	3	K_d	0,0852	0,2523	0,2499	0,4451	0,0813
		K_a	0,2849	0,8284	0,8425	1,5176	0,2592
		K_s	0,0405	0,1747	0,1694	0,4546	0,0710
		K_r	0,1424	0,4270	0,4276	0,5928	0,0804

A média e mediana são parâmetros que representam a tendência central das constantes de qualidade de água, enquanto o desvio padrão, valor mínimo e máximo representam a dispersão destas constantes.

A constante de remoção total de matéria orgânica (k_r) não foi diretamente calculada, mas representa a soma das constantes k_d e k_s , obtidas pelo algoritmo genético.

Uma alternativa à visualização da tendência central e dispersão das constantes de qualidade da água é representada pelo gráfico “Boxplot”, onde são apresentados os quartis 25% e 75% da distribuição, os valores máximo e mínimo, além da mediana. A Figura 2 apresenta o Boxplot para as constantes de qualidade da água.

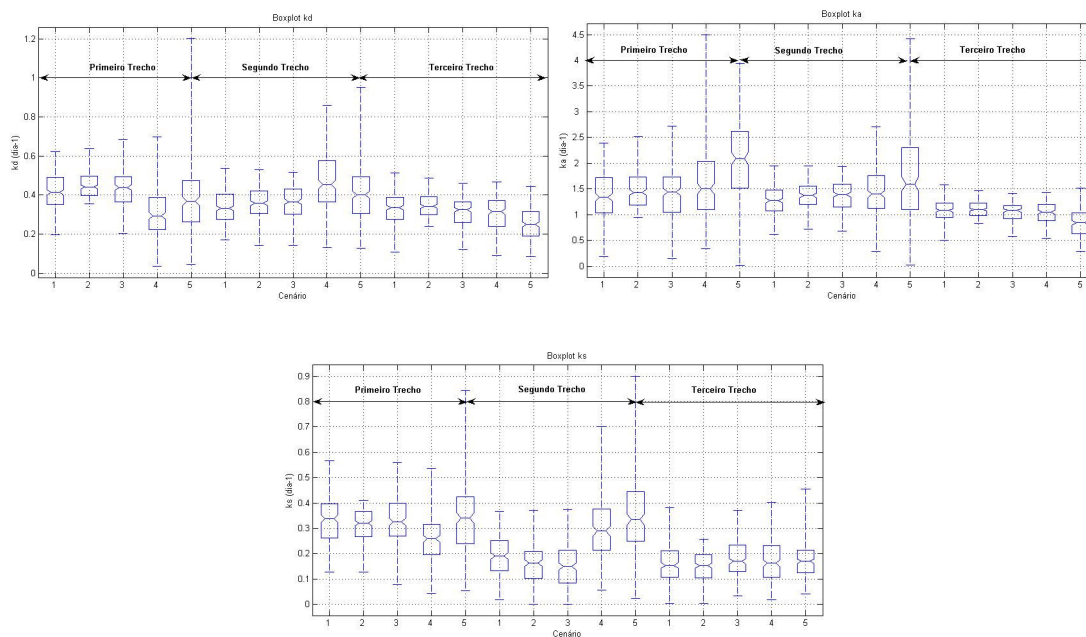


Figura 2 - Boxplots das constantes de qualidade de água.

Como foi mencionado anteriormente, o algoritmo genético robusto deve convergir para a solução do problema de otimização (calibração do modelo de qualidade de água) para qualquer simulação ou semente inicial. Verificou-se que esta convergência ocorreu em todas as simulações e

para cada um dos cenários; ou seja, o algoritmo genético encontrou, para cada cenário, 200 conjuntos de constantes de qualidade da água que produziam concentrações calculadas de DBO e OD iguais às “monitoradas”. Porém, como pode ser observado na Figura 2 e Tabela 3 à Tabela 7, estes conjuntos são diferentes entre si. Desta forma, este problema de calibração não tem solução única, pois qualquer um dos conjuntos de constantes de qualidade da água calibram o modelo. Porém, de um ponto de vista físico, existe apenas um conjunto de constantes que representa a condição real do rio modelado, ou seja, que representa as peculiaridades em relação à degradação microbiológica de matéria orgânica, a taxa de sedimentação e a reaeração atmosférica neste rio. Além disso, não se tem nenhuma garantia de que este único conjunto representativo do rio esteja entre os 200 conjuntos encontrados. Assim, apesar do algoritmo genético ter fornecido vários conjuntos de constantes de qualidade de água que calibram o modelo para o rio hipotético, não se conhece aquele conjunto de constantes que realmente representa as características física, química e biológica do rio.

Observando a Tabela 3 à Tabela 7, pode-se notar que, de forma geral, para os cenários onde havia poucas informações de monitoramento, a dispersão das constantes de qualidade de água foi maior do que para aqueles cenários onde havia um número maior informações de monitoramento. Por exemplo, comparando os cenários 2 (10 pontos de monitoramento) e 5 (1 ponto de monitoramento) observa-se que, em média, o desvio padrão das constantes de qualidade é aproximadamente 2 vezes maior para este cenário em relação àquele. Uma explicação para esta diferença pode estar relacionada ao fato de que no processo de calibração ajustam-se as concentrações de DBO e OD medidas às calculadas. Assim, ao se calibrar o modelo utilizando maior número de pontos de monitoramento, restringe-se mais as concentrações calculadas, pois uma quantidade maior destas concentrações deverá se ajustar às concentrações medidas. Por sua vez, se os valores das concentrações calculadas estão mais restritos, os valores das constantes de qualidade de água com as quais são calculadas estas concentrações também estarão mais restritas. Segue que, quando se tem apenas 1 ponto de monitoramento (cenário 5), a concentração de DBO e OD calculada deve se ajustar à monitorada neste único ponto, enquanto que em outros pontos do rio a concentração calculada pode assumir qualquer valor. Desde que no ponto de monitoramento a concentração calculada se ajuste à monitorada, entende-se que o modelo de qualidade de água está calibrado. Como as concentrações são pouco restritas, as constantes usadas para calcular estas concentrações também o serão, de forma que elas estarão sujeitas principalmente à restrição do espaço de busca. Por este motivo, as constantes de qualidade de água encontradas pelo algoritmo genético para este cenário apresentam maior variação, que pode ser percebida através do desvio padrão e do valor mínimo e máximo. Ainda dentro do cenário 5, este comportamento pode ser constatado pela observação de que nos trechos onde não havia pontos de monitoramento (trecho 1 e

2), o desvio padrão é aproximadamente o dobro do terceiro trecho, onde estava localizado 1 ponto de monitoramento. Semelhantemente, para o cenário 3, as constantes de qualidade de água apresentam menor desvio padrão em relação aos cenários 4 e 5, ou seja, maior restrição quanto aos valores das constantes. Isto ocorre principalmente no segundo e terceiro trecho.

Aceitando como verdadeira a explicação de que uma quantidade maior de pontos de monitoramento implica numa maior restrição em relação às constantes de qualidade da água, esperava-se que o desvio padrão das constantes para o cenário 1 fosse menor do que o desvio padrão das constantes para o cenário 2. Porém isso não ocorre; ao contrário, em média os desvios padrão das constantes para o cenário 1 são maiores. Isto se deve ao fato de que o critério de parada para o cenário 1 ($f \leq 0.05$) é menos restritivo do que para o cenário 2 ($f \leq 0.005$). Assim, se por um lado no cenário 1 as constantes de qualidade de água deveriam ser mais restritas em relação ao cenário 2 devido ao maior número de pontos de monitoramento, por outro lado estas constantes deveriam ser menos restritas no cenário 1 em relação ao cenário 2 devido à menor restrição do critério de parada. Desta forma, percebe-se que o efeito da restrição devido ao critério de parada sobrepuja o efeito da restrição devido ao número de pontos de monitoramento, o que implica na dispersão das constantes para o cenário 2 ser ligeiramente inferior à dispersão para o cenário 1.

4.2 – Comparação com a solução analítica

Das 200 simulações realizadas para cada cenário de calibração, foram tomadas aleatoriamente 20 destas simulações, e então as curvas de concentração de DBO e OD correspondentes foram inseridas num gráfico. O gráfico também apresenta o perfil de DBO e OD teórico, além dos pontos de monitoramento utilizados na calibração. Estes gráficos podem ser visualizados na Figura 3.

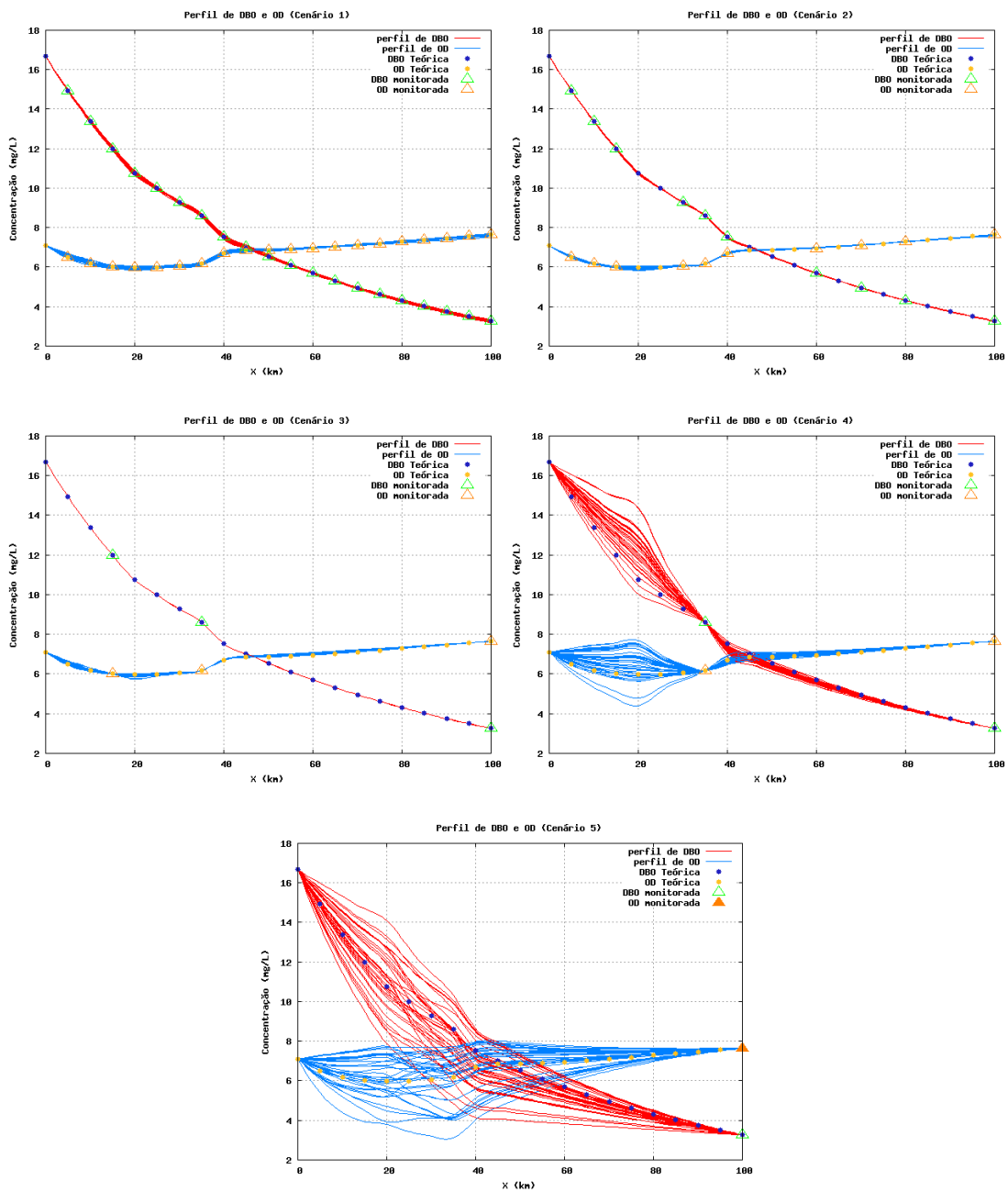


Figura 3 - Perfil de DBO e OD de 20 das 200 simulações para os cenários 1, 2, 3, 4 e 5.

Observa-se que os cenários de calibração para os quais as constantes de qualidade de água apresentaram dispersão mais elevada correspondem àqueles onde dispersão das concentrações de DBO e OD também é maior. Comparando os perfis de DBO e OD para o cenário 3 e 5 é possível visualizar claramente esta diferença de dispersão. Também é possível observar que, para os cenários 4 e 5, a dispersão de DBO e OD é alta quando se está longe dos pontos de monitoramento que foram empregados na calibração. Porém, todas as curvas de DBO e OD convergem exatamente para o valor monitorado nestes pontos. Por este motivo, para qualquer uma das simulações considera-se

que o modelo está calibrado, pois os conjuntos de constantes de qualidade de água minimizaram a função-objetivo (equação 3).

Para o cenário 1 os perfis de DBO e OD apresentados são praticamente iguais; além disso, estes perfis se ajustam bem aos pontos de DBO e OD teóricos, mesmo àqueles que não foram empregados na calibração. Apesar de praticamente não haver diferença entre os perfis apresentados, eles foram gerados a partir de conjuntos diferentes de constantes de qualidade da água. Semelhantemente, para os cenários 2 e 3 os perfis calculados se ajustam aos valores teóricos de concentração.

Observe que diferentemente do que ocorre nos cenários 1, 2 e 3 – onde se tem no mínimo um ponto de monitoramento por trecho do rio – nos cenários 4 e 5 os perfis de DBO e OD não são coincidentes para todas as simulações, ou seja, diferentes conjuntos de constantes de qualidade de água produzem diferentes perfis de DBO e OD.

4.3 – Critério representativo das constantes de qualidade da água

Devido ao fato de que em qualquer simulação o perfil de DBO e OD calculado se ajusta aos dados de monitoramento empregados na calibração, é possível tomar qualquer conjunto de constantes de qualidade da água e considerá-lo representativo do rio que está sendo estudado, empregando-o, desta forma, na modelagem de qualidade da água. Porém, também é possível utilizar a média das constantes obtidas nas 200 simulações, como está apresentado na Tabela 3 a Tabela 7. Desta forma, com o valor médio destas constantes plota-se o perfil de DBO e OD para cada cenário de calibração. Estes perfis estão apresentados na Figura 4.

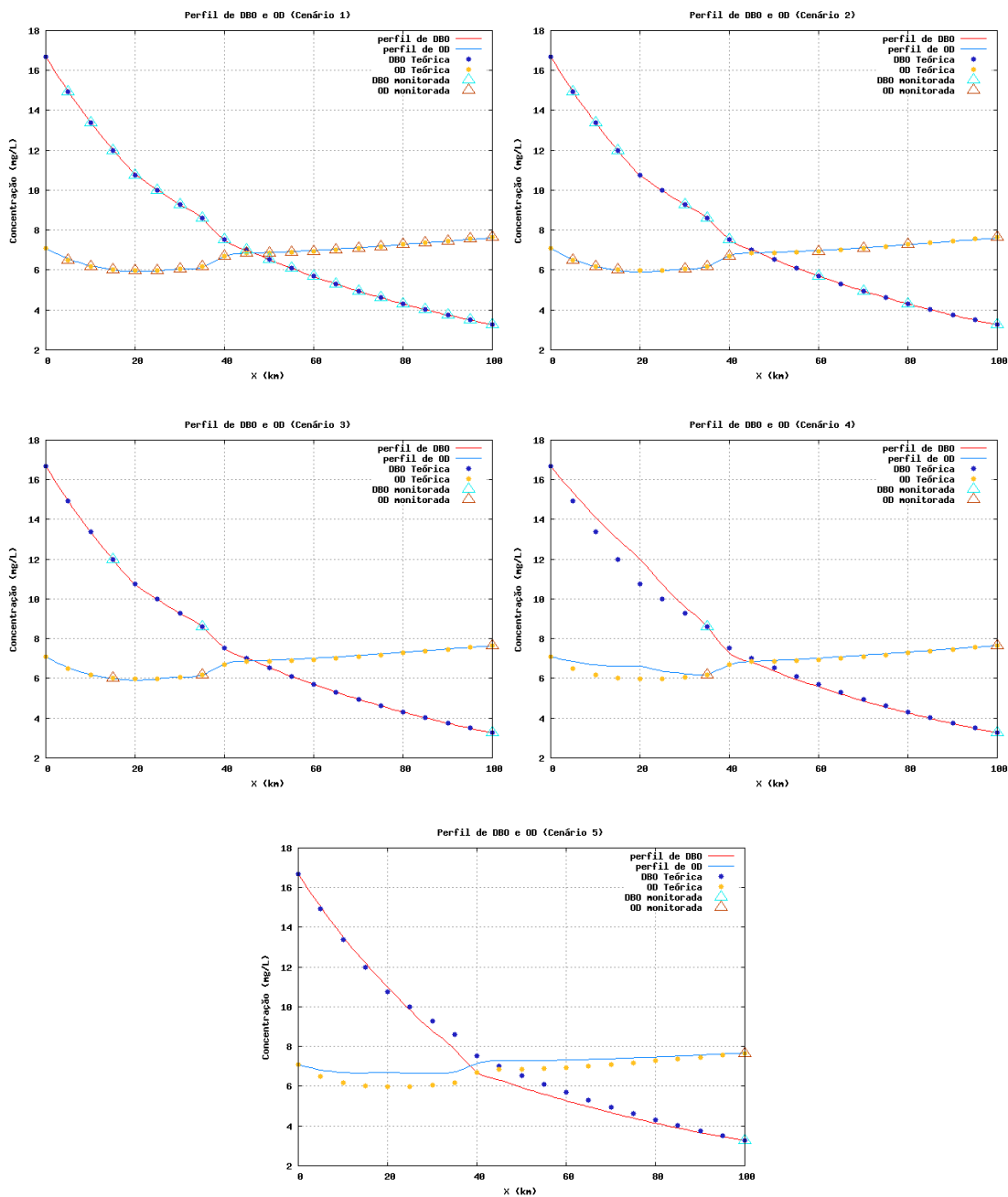


Figura 4 - Perfil de DBO e OD obtido com a média das constantes para os 5 cenários de calibração.

Observa-se que para os cenários 1 a 3 a média das constantes produziu perfis de DBO e OD praticamente iguais aos que tinham sido produzidos com cada conjunto particular de constantes. Estes perfis se ajustaram bem às concentrações monitoradas que empregadas na calibração e, também, aos valores teóricos destas concentrações. Para os cenários 4 e 5, o perfil também se ajustou bem às concentrações monitoradas, porém não às concentrações teóricas. Apesar disso, ao se utilizar a média das constantes de qualidade da água no modelo, resumem-se todos aqueles perfis da Figura 3 em apenas um perfil de DBO e OD que se aproximou razoavelmente bem do perfil teórico (Figura 4). Assim, o emprego da média das constantes pareceu ser um procedimento

adequado quando não se tem informação adicional para analisar qual dos conjuntos de constantes deve ser utilizado.

4.4 – Curvas das tendências das constantes de qualidade de água

Com os conjuntos de constantes de qualidade da água encontrados nas 200 simulações, foram plotadas, para cada trecho do rio, curvas que relacionam estas constantes entre si. Estas curvas podem ser visualizadas, para cada cenário de calibração, na Figura 5 à Figura 9. Na Tabela 8 estão apresentados os coeficiente de correlação linear ao quadrado (R^2) e as equação lineares .

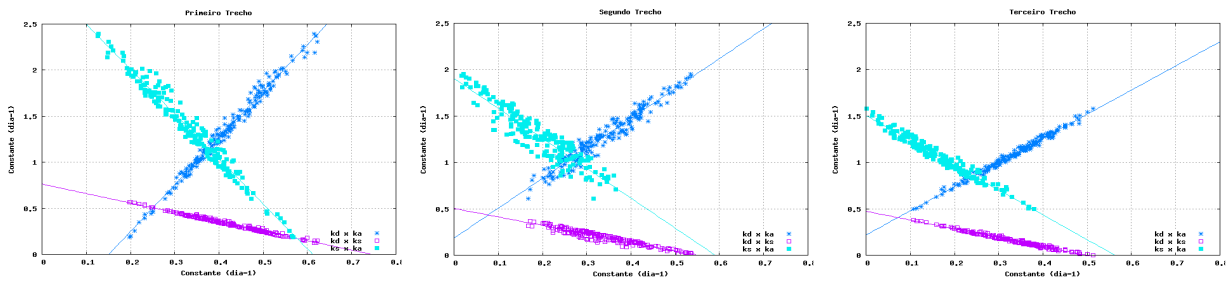


Figura 5 - Relação entre as constantes de qualidade de água para o cenário de calibração 1.

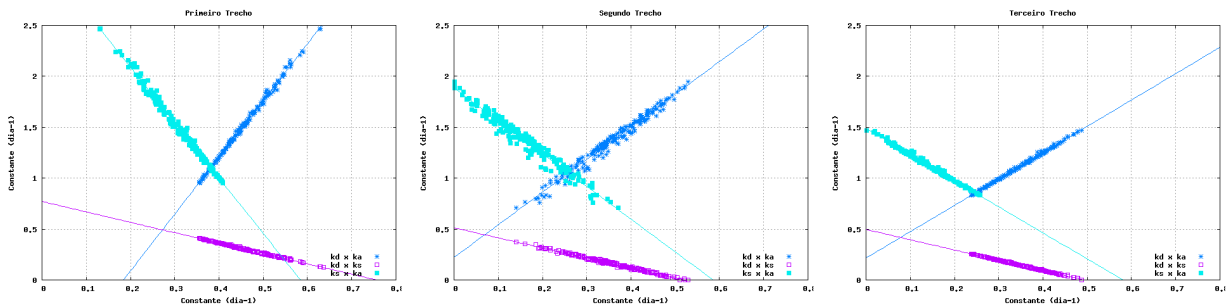


Figura 6 - Relação entre as constantes de qualidade de água para o cenário de calibração 2.

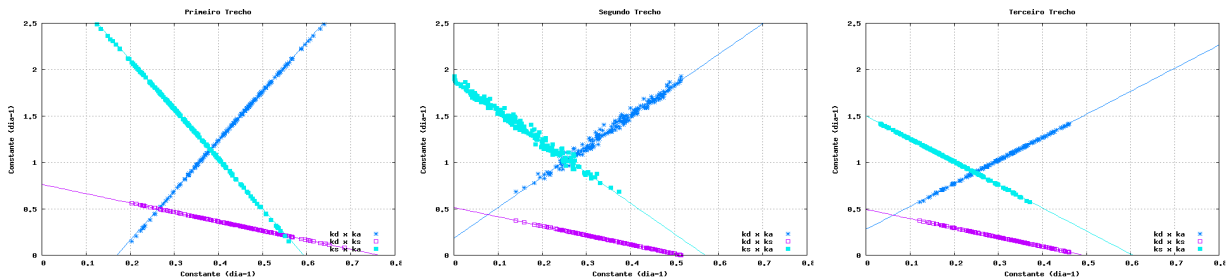


Figura 7 - Relação entre as constantes de qualidade de água para o cenário de calibração 3.

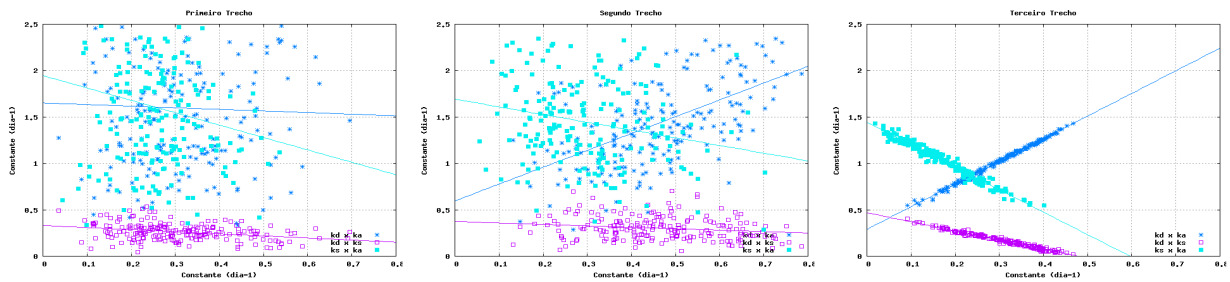


Figura 8 - Relação entre as constantes de qualidade de água para o cenário de calibração 4.

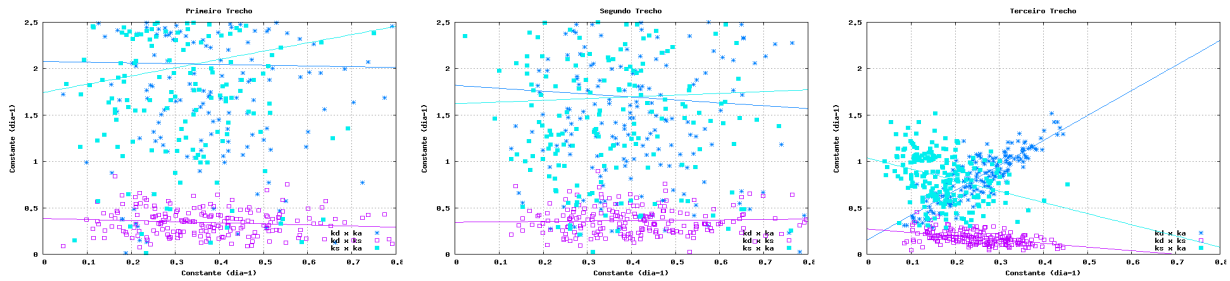


Figura 9 - Relação entre as constantes de qualidade de água para o cenário de calibração 5.

Tabela 8 - Equações que relacionam as constantes de qualidade da água.

Trecho		$K_s = f(k_d)$	$K_a = f(k_s)$	$K_a = f(k_d)$
Cenário 1	1	Função $k_s = -1.0228k_d + 0.7621$ $R^2 = 0.9856$	$k_a = -4.8792k_s + 2.9866$ 0.9684	$k_a = 5.0708k_d - 0.7657$ 0.9855
	2	Função $k_s = -0.9134k_d + 0.504$ $R^2 = 0.9176$	$k_a = -3.2267k_s + 1.9051$ 0.8574	$k_a = 3.2245k_d + 0.1843$ 0.9416
	3	Função $k_s = -0.9487k_d + 0.4738$ $R^2 = 0.9806$	$k_a = -2.6798k_s + 1.5072$ 0.9602	$k_a = 2.6029k_d + 0.2174$ 0.987
Cenário 2	1	Função $k_s = -1.0195k_d + 0.7702$ $R^2 = 0.9954$	$k_a = -5.4381k_s + 3.1766$ 0.9917	$k_a = 5.5741k_d - 1.0251$ 0.9978
	2	Função $k_s = -0.9839k_d + 0.5124$ $R^2 = 0.9888$	$k_a = -3.217k_s + 1.884$ 0.9644	$k_a = 3.2043k_d + 0.2217$ 0.9772
	3	Função $k_s = -1.0069k_d + 0.4959$ $R^2 = 0.9977$	$k_a = -2.561k_s + 1.4875$ 0.9937	$k_a = 2.5873k_d + 0.2147$ 0.9979
Cenário 3	1	Função $k_s = -1.0003k_d + 0.7636$ $R^2 = 1$	$k_a = -5.3476k_s + 0.3175$ 0.9997	$k_a = 5.3492k_d - 0.909$ 0.9997
	2	Função $k_s = -0.998k_d + 0.5142$ $R^2 = 0.9999$	$k_a = -3.3123k_s + 1.8859$ 0.9848	$k_a = 3.3058k_d + 0.1827$ 0.9847
	3	Função $k_s = -k_d + 0.4938$ $R^2 = 1$	$k_a = -2.483k_s + 1.5029$ 0.9998	$k_a = 2.4829k_d + 0.2769$ 0.9998
Cenário 4	1	Função $R^2 = 0.0974$	- 0.0305	- 0.001
	2	Função $R^2 = 0.0301$	- 0.0514	- 0.3293
	3	Função $k_s = -0.9798k_d + 0.4688$ $R^2 = 0.9732$	$k_a = -2.4218k_s + 1.445$ 0.9645	$k_a = 2.4418k_d + 0.2882$ 0.994
Cenário 5	1	Função $R^2 = 0.0193$	- 0.0224	- 0.0003
	2	Função $R^2 = 0.002$	- 0.001	- 0.003
	3	Função $R^2 = 0.2013$	- 0.1083	- 0.7128

Observe que para aqueles cenários onde havia ao menos 1 ponto de monitoramento em cada trecho do rio (cenário 1 a 3), as constantes de qualidade de água apresentaram uma relação linear bastante forte. Para o cenário 4, nos trechos 1 e 2 praticamente não existe relação nenhuma entre as constantes, enquanto que no último trecho esta relação existe. Para o cenário 5, em nenhum trecho do rio constata-se relação entre as constantes de qualidade de água.

Estas equações podem auxiliar na tarefa de encontrar as constantes representativas do rio. Assim, se por algum outro método qualquer - como experimentos de laboratório ou utilização de formulas empíricas - for determinada uma das constantes em cada trecho do rio, as outras duas podem ser obtidas através das relações apresentadas na Tabela 8. Por outro lado, esta informação adicional citada acima (valor de alguma constante de qualidade da água) também pode ser passada como dado de entrada ao programa antes da calibração. Desta forma, serão apenas 2 constantes de qualidade de água que deverão ser encontradas. Assim, como a função aptidão é composta por duas equações (resíduo em relação a DBO e OD), o algoritmo genético não encontrará vários conjuntos possíveis de constantes que calibrem o modelo, mas apenas um conjunto.

Inevitavelmente, para que o problema de calibração apresente apenas uma solução - aquela que representa as características reais do rio em questão - torna-se necessário que se forneça ao programa alguma informação adicional. Porém, esta informação pode ser fornecida antes ou depois da calibração. Caso ela seja fornecida antes, o algoritmo genético fornecerá, em qualquer uma das 200 simulações, o mesmo conjunto de constante de qualidade de água, ou seja, o único conjunto possível que satisfaz o problema de calibração. Porém, caso esta informação adicional seja fornecida depois da calibração, o algoritmo genético encontrará, em cada uma das 200 simulações, diferentes conjuntos de constantes que satisfaz o problema de calibração. Após isto, com estas constantes é possível encontrar relações lineares semelhantes às apresentadas na Tabela 8, e de posse destas equação, calcular as outras duas constantes em função daquela conhecida. Porém, a opção de inserir a informação adicional de qualidade de água após a calibração somente será efetiva para os cenários onde existe ao menos um ponto de monitoramento em cada trecho do rio. Estes cenários correspondem àqueles onde existe uma forte correlação entre as constantes de qualidade da água. Para os outros cenários, como não existe correlação entre as constantes, deve-se fornecer esta informação adicional ao programa de calibração.

5 – CONCLUSÕES

A calibração corresponde a um processo de otimização multi-objetivo. No presente estudo têm-se duas funções, uma relacionada à concentração de DBO e outra relacionada à concentração de OD. Porém existem 3 incógnitas a serem determinada (cada uma das constantes de qualidade da

água). Por este motivo, tem-se neste problema um grau de liberdade, o que implica no problema ter infinitas soluções. Assim, para que seja possível obter apenas uma solução para o problema de calibração, ou seja, apenas um conjunto de constantes de qualidade de água, seria necessário haver mais uma informação de qualidade de água ou equação.

Cada conjunto de constantes obtido pelo algoritmo genético ajustou as concentrações de DBO e OD calculadas às medidas. Como não havia informação que fundamentasse o emprego de um conjunto particular de constantes, uma metodologia aplicada foi usar a média das constantes obtidas nas 200 simulações. Agindo assim, percebeu-se que o perfil de DBO e OD obtido com esta média ajustou-se bem às concentrações monitoradas. Além disso, para os cenários 1 a 3, o perfil também se ajustou às concentrações teóricas.

O valor médio das constantes representou um critério para se obter alguma informação sobre as “verdadeiras” constantes de qualidade de água do rio em questão, porém não é o único. Por este motivo, é preciso procurar outros critérios que transforme todos aqueles conjuntos de constantes de qualidade de água em um único conjunto.

BIBLIOGRAFIA

- CHAPRA, C. S. **Surface Water Quality Modeling**. Colorado, USA: McGraw-Hill, 1997.
- CHO, J. H.; SUNG, K. S.; HA, S. R. A river water quality management model for optimizing regional wastewater treatment using a genetic algorithm. **Journal of Environmental Management**, v. 20, p. 1-14, jul. 2004.
- GONZAGA, C. C. Otimização irrestrita. In: _____. **Um curso de programação não linear**. Florianópolis, 2004. p. 13-49.
- KONDAGESKI, J. H.; FERNANDES, C.V.S. Calibração da equação de streeter – phelps através do método do algoritmo genético: um estudo de caso hipotético. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 15, 2005, João Pessoa.
- LACERDA, E. G. M. de; CARVALHO, A. C. P. L. F. de. Introdução aos algoritmos genéticos. In: Galvão, C. de O.; Valença, M.J.S. (orgs.). **Sistemas inteligentes: aplicações a recursos hídricos e ciências ambientais**: Porto Alegre: Ed. Universidade/UFRGS, 1999. p. 99-150.
- NAHON, I. M. **Sistema de apoio à análise de outorga de lançamento de efluentes para a variável demanda bioquímica de oxigênio**: Estudo de caso da bacia do Alto Iguaçu. Curitiba, 2006. 153 f. Dissertação (Mestrado em Engenharia de Recursos Hídricos e Ambiental) – Setor de Tecnologia, Universidade Federal do Paraná.
- PERERA, B. J. C.; NG A. W. M. Selection of genetic algorithm operators for river water quality model calibration. **Engineering Applications of Artificial Intelligence**, v. 16, p. 529-541, set. 2003.

TANOMARU, J. Motivação, fundamentos e aplicações de algoritmos genéticos. In: CONGRESSO BRASILEIRO DE REDES NEURAIAS, 2., 1995, Curitiba.

WANG, Q. J. The genetic algorithm and its application to calibrating conceptual rainfall-runoff models. **Water Resources Research**, v. 27, n. 9, p. 2467-2471, set. 1991.

WARDLAW, R.; SHARIF, M. Evaluation of genetic algorithms for optimal reservoir system operation. **Journal of Water Resources Planning and Management**, v. 125, n. 1, p. 25-33, jan./fev. 1999.