

Aula 3

8. Calibração e validação de modelos

8.1. Introdução

O desenvolvimento de um modelo obrigatoriamente passa pelas etapas de calibração e validação. A calibração de um modelo matemático se refere ao processo de ajuste de (alguns de) seus parâmetros para que ele represente bem o processo em estudo, coerente com a realidade observada. Nessa etapa necessita-se dispor de um conjunto de dados observados e as respectivas condições iniciais e de contorno.

A partir da calibração, realizada com um conjunto de dados observados, a validação consiste do teste ou avaliação do desempenho do modelo em outros cenários. Como na calibração, na fase de validação necessita-se também de um conjunto de dados observados e as condições iniciais e de contorno.

8.2. Calibração

Na etapa de calibração, alguns parâmetros do modelo são selecionados para realizar seu ajuste com base na comparação entre valores observados e preditos pelo modelo. Os parâmetros escolhidos são, normalmente, aqueles que possuem uma incerteza maior, ou seja, parâmetros de determinação mais difícil. Os valores desses parâmetros são então ajustados por um processo estatístico, objetivando-se a otimização de um indicador estatístico, muitas vezes o Raiz do Erro Quadrático Médio ou RMSE (veja adiante). Durante esse processo, a faixa de valores para os parâmetros pode ser restrita para garantir o significado físico do parâmetro.

8.3. Validação

Com o modelo calibrado, segue a etapa da validação, também chamada de teste ou avaliação. Essas últimas duas denominações têm a preferência da maioria dos modeladores, uma vez que a palavra “validação” tem a conotação de que o modelo seja “validado”, implicando na sua validade em qualquer circunstância. O que se faz, na verdade, é testar o desempenho do modelo para determinados cenários, avaliando seu desempenho.

A avaliação do desempenho do modelo na fase de validação é subjetiva. Como na fase de calibração, utilizam-se indicadores estatísticos. Uma forma de avaliar o modelo é pela comparação do RMSE obtido na calibração ($RMSE_c$) com o RMSE da validação ($RMSE_v$). Normalmente, o $RMSE_v$ será maior que o $RMSE_c$, mas se for muito maior pode indicar que a calibração não foi representativa. Modelos que apresentam $RMSE_v \gg RMSE_c$ normalmente possuem problemas conceituais ou mesmo erros de programação.

8.4. Indicadores de desempenho

Tanto na calibração e na validação necessita-se de um critério objetivo para otimizar o desempenho do modelo. Comparando-se um conjunto de n dados observados (O_i) ao conjunto correspondente de n dados simulados ou preditos (P_i), utilizam-se geralmente parâmetros estatísticos para quantificar a semelhança entre os dois. Evidentemente, o que se procura na calibração é ajustar parâmetros do modelo para se obter o maior grau de semelhança possível. Um excelente trabalho sobre os indicadores estatísticos mais comuns nessa avaliação foi publicado por Legates & McCabe (1999). Os próximos itens foram adaptados do trabalho deles, e vale a pena consultar o texto completo para maiores informações e detalhes.

Coeficiente de determinação (R^2)

O Coeficiente de determinação ou R^2 descreve a fração da variância total nos dados observados que pode ser explicada pelo modelo. Seu valor fica entre 0 e 1, e quanto maior o valor, melhor a concordância entre modelo e observação. A fórmula para seu cálculo é:

$$R^2 = \left\{ \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (P_i - \bar{P})^2}} \right\}^2$$

Coeficiente de eficiência (E)

Desenvolvido por Nash & Sutcliffe (1970), O Coeficiente de eficiência (também chamado de coeficiente de Nash-Sutcliffe) ou E é utilizado com muita frequência para a avaliação de modelos. Seu valor fica entre $-\infty$ e 1, e quanto maior o valor, melhor a concordância entre modelo e observação. A fórmula para seu cálculo é:

$$E = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}$$

A interpretação do valor de E pode ser feita da seguinte forma: E = 1 representa uma concordância total entre predito e observado, ou seja, o modelo desempenha com perfeição. E = 0 significa que o modelo desempenha tão bem quanto a simples média dos valores observados. Em outras palavras: tomando a média dos valores observados resulta numa predição tão boa quanto a do modelo o que, evidentemente, não é muito animador do ponto de vista do modelador. Valores negativos para E indicam que seria melhor tomar a simples média do que os valores preditos pelo modelo.

Índice de concordância de Willmott (d)

O Índice de concordância (d) foi desenvolvido por de Willmott (1981) e é definido como

$$d = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (|P_i - \bar{O}| + |O_i - \bar{O}|)^2}$$

O valor de d varia entre 0 e 1, e quanto maior seu valor, melhor o desempenho do modelo.

Raiz do Erro Quadrático Médio (RMSE)

O RMSE (sigla do inglês Root Mean Square Error) é definido como

$$RMSE = \frac{\sum_{i=1}^n (O_i - P_i)^2}{n}$$

O RMSE, ao contrário dos indicadores apresentados até aqui, possui uma unidade (dimensão), igual à dimensão dos valores observados e preditos. Interpreta-se seu valor como uma medida do desvio médio entre observado e predito, porém observe que as diferenças entre O e P são elevadas ao quadrado. A verdadeira média dos desvios é descrita pelo MAE, a seguir.

Erro médio absoluto (MAE)

O MAE (sigla do inglês Mean Absolute Error) é definido como

$$MAE = \frac{\sum_{i=1}^n |O_i - P_i|}{n}$$

Como o RMSE, o MAE possui dimensão igual à dimensão dos valores observados e preditos. Seu valor representa o desvio médio entre observado e predito. Comparando o RMSE e o MAE, o primeiro dá um peso maior para desvios grandes (pois são elevados ao quadrado), enquanto o MAE dá um peso igual a todos os desvios.

8.5. Referências

Legates, D.R., and G.J. McCabe Jr. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* 35:233-241, 1999.

Nash, J.E., and J.V. Sutcliffe. River flow forecasting through conceptual models, I, A discussion of principles, *J. Hydrol.*, 10, 282-290, 1970.

Willmott, C. J., On the validation of models, *Phys. Geogr.*, 2, 184-194, 1981.