

Cuidados na avaliação de modelos de simulação

Alfredo José B. Luiz^{1†}, Fernando Antônio M. Silva²

¹Embrapa Meio Ambiente.

²Embrapa Cerrados.

Resumo: O uso de modelos de simulação é cada vez mais frequente na pesquisa. A maior capacidade, tanto de memória quanto de processamento, assim como a grande disponibilidade, acessibilidade e quantidade de dados coletados e armazenados, induz e, por vezes, torna obrigatória a utilização dessa ferramenta. Por outro lado, quando calibrados, os modelos de simulação permitem a realização, em segundos e a custo quase zero, de inúmeras estimativas de comportamento futuro, nos mais diferentes cenários, o que, de outra maneira, demandaria muito tempo e recurso para serem obtidas. Modelos de crescimento de plantas e modelos climáticos estão entre os mais disseminados e utilizados na pesquisa agropecuária. Algumas métricas têm sido utilizadas com frequência pelos usuários desses modelos para avaliar sua eficiência. Entretanto, nem sempre são empregadas as melhores técnicas. Neste trabalho, foram analisados os resultados de simulação de dados de temperatura do ar e do solo, por dois modelos diferentes: ETA-HADGEM e STICS, respectivamente. As métricas usualmente indicadas EF (eficiência do modelo) e RMSE (raiz quadrada do erro médio), foram comparadas com resultados da análise de regressão e do teste de sinais para o viés. A introdução da correção do viés também foi aplicada e interpretada. Conclui-se que as métricas, se usadas de forma isolada e sem análise aprofundada das características das variáveis simuladas, podem resultar em uma interpretação equivocada da eficiência dos modelos.

Palavras-chave: STICS; ETA-HADGEM; viés; modelos de regressão; teste de sinais.

Care in the evaluation of simulation models

Abstract: The use of simulation models is increasingly frequent in research. The greater capacity, both in memory and processing, as well as the great availability, accessibility and amount of data collected and stored, induces and sometimes makes the use of this tool mandatory. On the other hand, when calibrated, the simulation models allow the realization, in seconds and at almost zero cost, of countless estimates of future behavior, in the most different scenarios, which, otherwise, would require a lot of time and resources to be obtained. Crop growth models and climate models are among the most widespread and used in agricultural research. Some metrics have been frequently used by users of these models to assess their efficiency. However, the best techniques are not always employed. In this work, the simulation results of air and soil temperature data were analyzed by two different models: ETA-HADGEM and STICS, respectively. The metrics usually indicated: EF (modelling efficiency) and RMSE (root mean square error), were compared with results of regression analysis and sign test for bias. The introduction of bias correction was also applied and interpreted. It is concluded that the metrics, if used in isolation and without in-depth analysis of the characteristics of the simulated variables, may result in a misinterpretation of the efficiency of the models.

Keywords: STICS; ETA-HADGEM; bias; regression models; sign test.

Introdução

Os modelos conceituais sempre estiveram presentes na pesquisa científica. Podemos encontrar uma tentativa de descrever o funcionamento da nossa circulação sanguínea por meio de um modelo conceitual num dos pilares fundamentais do atual Método Científico, o livro Discurso do Método, de Descartes, escrito provavelmente antes de 1637 (Descartes, 2001). Sabemos hoje que o modelo hidráulico utilizado por Descartes para tentar explicar o sistema cardiovascular humano, semelhante ao que descreve o funcionamento das máquinas a vapor, não está correto, apesar disso, o exemplo é útil para demonstrar que a Ciência sempre usou de modelos para tentar entender e reproduzir o funcionamento dos sistemas que estuda.

^{††} Autor correspondente: alfredo.luiz@embrapa.br.

Sigmae, Alfenas, v.12, n.3, p.213-223, 2023.

67ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria (RBras) e 20º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO)

As duas principais utilidades dos modelos de simulação são, sem dúvida, aprimorar o conhecimento sobre os sistemas que os modelos pretendem representar (Nash e Sutcliffe, 1970); e permitir a realização de experimentos virtuais, sem a necessidade da condução de experimentos físicos (Affin et al., 1992). Os modelos de simulação não precisam, necessariamente, serem matemáticos e nem mesmo estarem disponíveis na forma de programas computacionais (Vicente, 2005). Entretanto, a maior capacidade, tanto de memória quanto de processamento dos computadores, assim como a grande disponibilidade, acessibilidade e quantidade de dados coletados e armazenados digitalmente, induz e, por vezes, torna obrigatória a utilização dessa ferramenta.

Depois de calibrados, os modelos de simulação permitem realizar, em segundos e a custo quase zero, experimentos virtuais do comportamento futuro das variáveis de interesse nos mais diferentes cenários, o que, de outra maneira, demandaria muito tempo e recurso para ser feito se fossem utilizados experimentos físicos. Modelos de crescimento de plantas e modelos climáticos estão entre os mais disseminados e utilizados na pesquisa agropecuária (Andrade et al., 2009; Machado Filho et al., 2016).

No estabelecimento dos critérios recomendados pelo Programa Nacional de Zoneamento Agrícola de Risco Climático (ZARC), um instrumento de política agrícola e gestão de riscos na agricultura (MINITTI, 2020), estimativas de produtividade das culturas em função de características locais, principalmente do solo e do clima, são fundamentais. Os dados utilizados para cálculos de probabilidades de ocorrência das variáveis agrometeorológicas, nesses casos, são feitos baseados em séries históricas, com dados do passado recente. Entretanto, na iminência das mudanças climáticas, faz-se necessário estimar o comportamento futuro, caso as previsões dos modelos climáticos se verifiquem. Nesse aspecto, as projeções dos modelos climáticos globais ou regionais são úteis para o planejamento de medidas preventivas ou mitigatórias (MAIA et al., 2011). Como é impossível experimentar fisicamente o futuro, essa é uma aplicação por excelência dos modelos de simulação. Na realidade, é necessária a união dos dois tipos citados, os modelos climáticos para simular as variáveis agrometeorológicas futuras, que por sua vez servirão de entrada nos modelos de crescimento de plantas.

Para a utilização segura dos modelos de simulação é preciso que se tenha confiança que eles são capazes de reproduzir com consistência os dados reais. A avaliação da eficiência, também chamada de destreza, desses modelos é feita comparando suas projeções retrospectivas com dados originais, como os obtidos em estações climatológicas ou em experimentos de campo controlados. É aceitável que um modelo não reproduza exatamente a realidade e até apresente vieses passíveis de correção. Entretanto, carecemos de estudos que avaliem o desempenho desses modelos (Luiz e Maia, 2016).

Nesse trabalho, foi selecionado o modelo climático ETA-HADGEM para estimar a temperatura diária do ar, variável que serve de entrada para o modelo de crescimento de plantas STICS, que tem entre as variáveis que calcula, a temperatura do solo na camada superficial. As estimativas dessas temperaturas pelos modelos foram comparadas com dados observados na área experimental da Embrapa Cerrados, em Planaltina, DF. Os modelos utilizados permitem simular muitas outras variáveis e por períodos muito mais longos, entretanto, o objetivo presente é apenas o de apresentar algumas formas de estudar, genericamente, a eficiência desses modelos e discutir como diferentes ferramentas podem ser utilizadas para esse propósito.

Material e métodos

Foram utilizados dados experimentais e meteorológicos obtidos na Embrapa Cerrados, em Planaltina/DF, nos anos de 2013 e 2014. Os valores de temperatura do solo, de 0 a 5 cm de profundidade, de outubro de 2013 a setembro de 2014, foram obtidos de um experimento de longa duração localizado na latitude de 15°33'34"S, e na longitude de 47°44'12"O, numa altitude de 1.035 m, e comparados com dados simulados pelo STICS, um modelo genérico de simulação de culturas e dos seus balanços de água e nitrogênio (BRISSON et al., 1998 e 2002). O experimento possui diversos tratamentos, mas nesse trabalho foram utilizados os dados oriundos apenas do cultivo da soja sob plantio convencional. Já os dados medidos de temperatura do ar (mínima e máxima), de 01/01/2013 a 31/12/2014, foram comparados com os simulados pelo modelo ETA-HADGEM, que é derivado do downscaling dinâmico do modelo HADGEM2-ES (do relatório IPCC AR5) usando o modelo dinâmico ETA do Inpe, com resolução de 20x20 km de latitude e longitude (CHOU et al., 2014). Os dados meteorológicos são provenientes da estação climatológica principal da Embrapa Cerrados (disponíveis em <http://clima.cpac.embrapa.br/clima/dados-climaticos/#>), situada na altitude de 1.007 m, cujas coordenadas geográficas de latitude 15°35'30"S e de longitude 47°42'30"O (Silva et al., 2014), permitiram a extração dos dados correspondentes aos da sua quadricula envolvente, a partir da malha gerada pelo modelo ETA.

Uma das métricas frequentemente utilizada para verificar a eficiência de modelos de simulação é chamada de simplesmente de Eficiência (ou Efficiency – EF). Essa métrica foi proposta pela primeira vez por James Eamonn Nash, formado em engenharia civil, mas que deu importantes contribuições ao estudo da hidrologia, tendo sido pioneiro no uso da regressão ao relacionar os parâmetros dos modelos hidrológicos com as características da bacia hidrográfica. O artigo do *Journal of Hydrology* (Nash e Sutcliffe, 1970) introduziu a medida de eficiência de Nash-Sutcliffe para calibrar modelos hidrológicos e se tornou um dos artigos mais citados em hidrologia. Nele foi apresentada uma proposta muito simples para testar a eficiência de modelos. Da forma como é calculada, a métrica EF compara o afastamento absoluto médio entre os valores preditos e observados com o afastamento absoluto médio entre os valores observados e sua média. Outra estatística usada na avaliação dos modelos é a conhecida raiz do erro quadrático médio (*root mean square error* – RMSE). Também é calculada a média dos desvios (*mean deviation* – MD), que é simplesmente a média das diferenças entre o valor simulado e o valor observado, que é exatamente o mesmo que o viés do modelo.

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

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

$$MD = \frac{1}{n} \sum_{i=1}^n (S_i - O_i)$$

onde n é o número de observações, S_i e O_i são os valores simulados e observados, respectivamente, e \bar{O} é o valor médio dos dados observados.

Além dessas métricas, foram calculados, para cada variável, e no caso das temperaturas do ar, para cada ano, os valores médios observados (Mobs) e o coeficiente de correlação linear entre dados observados e simulados (R).

Depois de feitas as análises usualmente aplicadas pelos estudiosos de modelos de simulação, foram ajustados modelos de regressão linear simples entre os valores observados e simulados para cada tipo de dado e ano, de maneira similar ao feito para as métricas. Nesse caso, os valores do coeficiente de determinação (r^2), o valor da estatística F para o ajuste do modelo e seu respectivo valor $P > F$, foram determinados.

Por fim, foram aplicados os testes não-paramétricos de sinais (*sign test*) e de sinais ranqueado (*signed rank*) entre os valores observados e simulados, como uma forma de avaliar se o viés observado pode ser considerado significativo. Neste teste se define, em cada par, a diferença dos escores, baseando-se apenas no sentido das diferenças dos escores, registrados sobre a forma de sinais: (+) e (-). Nesse caso foram usadas para testar a semelhança entre os dois conjuntos os testes *t* de Student e *S* de Wilcoxon, com as respectivas probabilidades associadas ($P > t$ e $P > S$). O teste de sinais, medido pela estatística *t*, avalia se as diferenças positivas e negativas entre os valores observados e simulados se distribuem de forma uniforme, como seria esperado se o modelo simulasse corretamente os dados observados, errando as estimativas em torno do valor real, para mais ou para menos, apenas ao acaso. Já o teste de sinais ranqueado, medido pela estatística *S*, tem o objetivo de avaliar a simetria dos erros positivos ou negativos produzidos pelo modelo, ou seja, testa se a mediana dos erros é diferente de zero.

Todas as análises foram feitas para três diferentes situações, primeiro foram aplicadas sobre os dados originais, sem transformação. Em seguida, as mesmas análises foram feitas após a correção do viés calculado dos dados simulados (subtraído ou somado, conforme seu sinal). Para efeito de demonstração e discussão, essa operação foi realizada mesmo se o teste de sinais não identificou o viés como significativo. Por último, foi ajustada uma equação de regressão que considerou um componente sazonal, inserido por meio dos senos e cossenos dos dias do ano (dda).

A utilização de um fator sazonal no presente trabalho, além de permitir a comparação entre formas de análise de modelos, obedece a uma lógica baseada nas características fundamentais das variáveis analisadas. Como se sabe, pelo conhecimento astronômico disponível, a temperatura do ar e do solo é predominantemente influenciada pela energia solar incidente sobre o local, que varia segundo a órbita terrestre em torno do sol. Embora a órbita seja elíptica, ela pode ser incluída no modelo, de forma mais simples, se for considerada circular, o que permite medir o movimento diário em frações dos 360 graus do círculo. Dessa forma, a cada dia, a temperatura observada foi expressa como função do valor estimado pelos modelos, somado a esse valor multiplicado pelo seno ou cosseno do ângulo diário (em radianos).

Assim, para os ajustes temporais dessas análises, o valor da temperatura máxima do ar estimada pelo modelo ETA (TmaxEta) foi corrigido da seguinte forma:

$$T_{maxEta}' = T_{maxEta} + \{T_{maxEta} * \text{seno}_{Rad} [dda*(360/365)]\}$$

Já o valor da temperatura mínima do ar estimada pelo modelo ETA (TminEta) foi corrigido assim:

$$T_{minEta}' = T_{minEta} + \{T_{minEta} * \text{cosseno}_{Rad} [dda*(360/365)]\}$$

Da mesma maneira, foi corrigida a estimativa da temperatura do solo pelo modelo STICS (TsStics):

$$TsStics' = TsStics + \{TsStics * \cos\theta_{Rad} [dda*(360/365)]\}$$

Esses valores corrigidos foram usados, pareados aos valores observados, e repetidas as mesmas análises realizadas com os dados dos modelos e aqueles com viés incluído. Todas as análises foram realizadas no ambiente do programa SAS (SAS, 2013).

Resultados e discussão

Primeiramente, é preciso detalhar o significado dos valores de EF. Essa métrica pretende medir a eficiência do modelo, ou seja, o grau de concordância entre valores observados e simulados. Seu valor pode variar de 1, que significa predição perfeita, até menos infinito ($-\infty$). O valor negativo de EF indica que a média simples das observações é um preditor melhor que o modelo. Ou seja, se o valor de EF é negativo o modelo é pior que a média como estimador dos valores reais. E quanto mais EF se aproximar de 1, melhor a representação que o modelo faz das observações. Por esse critério, é possível observar na Tabela 1 que tanto o ETA quanto o STICS apresentam baixa eficiência, ou são ineficientes para a estimativa das temperaturas estudadas, pois EF sempre foi negativo.

Tabela 1. Critérios estatísticos que mostram o desempenho do modelo STICS em simular a temperatura do solo, no período de 22/10/2013 a 30/09/2014, e do modelo ETA-HADGEM em simular a temperatura do ar nos anos de 2013 e 2014, em Planaltina/DF, Brasil.

Tipo de dados	Variáveis	Métricas usadas na análise de modelos							Análise de regressão			Teste de sinais (viés=MD)				
		Ano	n	EF	MObs	MD	R	RMSE	r ²	F	P>F	viés	t	P>t	S	P>S
Modelo climático ETA	Temp. máxima	2013	365	-1,29	28,17	-0,75	0,06	3,67	0,01	1,49	0,2232	-0,753	-4,001	0,0001	-8.814	<0,0001
		2014	365	-0,85	28,66	-1,30	0,23	3,48	0,05	20,10	<0,0001	-1,300	-7,668	<0,0001	-15.307	<0,0001
	Temp. mínima	2013	365	-0,78	16,45	1,00	0,43	2,91	0,18	80,49	<0,0001	1,004	6,999	<0,0001	12.288	<0,0001
		2014	365	-0,77	15,87	2,07	0,47	3,26	0,22	102,09	<0,0001	2,072	15,660	<0,0001	25.301	<0,0001
Modelo de plantas STICS	Temp. do solo Prof. (cm) 0 a 5	82	-0,20	22,86	1,72	0,77	2,24	0,59	116,97	<0,0001	1,715	10,771	<0,0001	1.574	<0,0001	

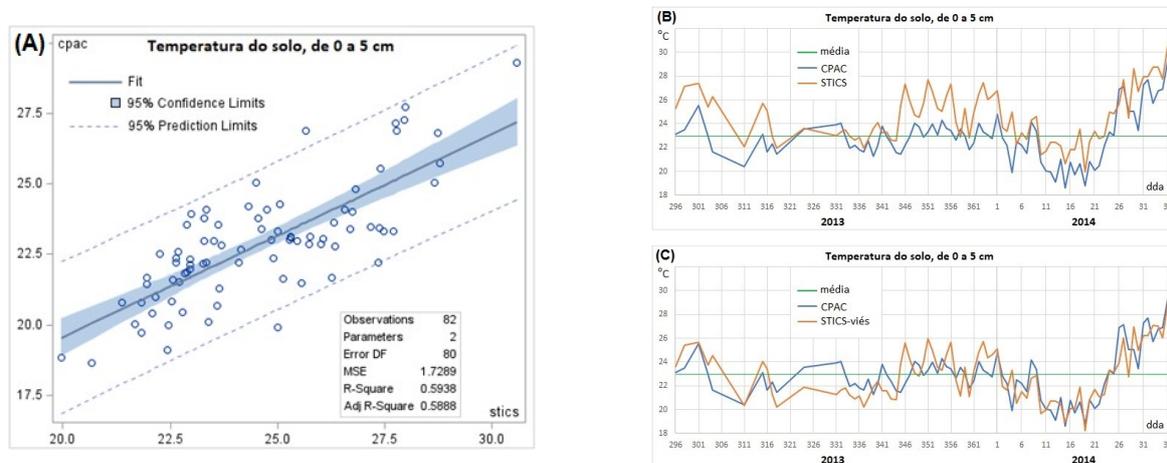
Fonte: Autores.

É importante ressaltar que a comparação de valores de MD e de RMSE não podem ser feitas como entre valores de EF. Isso porque a métrica proposta por Nash é invariante em relação à escala ou dimensão dos dados, enquanto MD e RMSE são dependentes desses valores. Ou seja, na tabela 1, só é possível comparar essas duas estatísticas entre anos, no caso das temperaturas do ar, máxima e mínima, pois as médias observadas (MObs) são, obviamente por se tratar de variáveis diferentes, bastante distintas em valores absolutos.

Apesar desse julgamento prejudicial aos modelos, proporcionado pelos resultados de EF, podemos observar que, com exceção da temperatura máxima em 2013, todos os ajustes de regressão linear simples foram significativos, apesar de baixos valores do coeficiente de determinação. No caso da temperatura do solo, entretanto, até por essa estatística podemos considerar que o STICS tem um desempenho razoável ($r^2 = 0,59$). Ainda na Tabela 1 é possível perceber que os testes indicam que o viés é altamente significativo, fazendo necessária uma maior investigação sobre seu comportamento.

Para investigar melhor essa questão, é apresentada abaixo a comparação gráfica entre dados observados e simulados para a temperatura do solo. A métrica EF (-0,20) nos diz que o modelo é ineficiente, ou seja, as previsões do modelo são piores que apenas a média dos valores observados. Em outras palavras, segundo a EF, a estimativa que indicasse que em todos os dias do ano a temperatura do solo será igual à média, acertaria mais do que o modelo de simulação com todos os seus complexos cálculos. Entretanto, assim como indicaram os resultados da análise de regressão linear, a realidade parece ser outra. Na Figura 1-A vemos os dados distribuídos conforme a escala de medida da temperatura do solo (graus centígrados) observada (eixo horizontal) e simulada (eixo vertical). Esse gráfico não considera a questão temporal, mas sim a associação entre valores segundo a escala de medida, ou seja, quando os valores observados são mais altos os valores simulados também são altos? Tanto essa figura quanto as medidas de correlação e regressão nos dizem que o modelo é razoável como estimador da temperatura do solo, conclusão diferente da que a EF indicou (Tabela 1).

Figura 1. Comparação gráfica entre valores observados e simulados pelo modelo STICS, da temperatura do solo, de out/2013 a set/2014, Planaltina/DF, Brasil.



Fonte: Autores.

E se olharmos a Figura 1-B, os dados observados (linha azul) e simulados (linha laranja) são expostos segundo a data em que foram medidos ou calculados. A linha verde, que representa a média dos valores observados, como se vê, não sofre nenhuma alteração ao longo do tempo. Fica evidente que o modelo é muito mais eficiente que a média na tentativa de estimar o valor observado ao longo do tempo, o que pode ser ainda melhor percebido ao observarmos que nos meses em que os valores observados ficaram bem abaixo da média, o modelo simulou valores baixos também, e fez o oposto nos meses em que a temperatura subiu. Por fim, uma outra questão a ser observada é a

que está exposta na Figura 1-C. Nesse gráfico, os dados observados e a média (linhas verde e azul) são os mesmos, mas os dados simulados (linha laranja) foram corrigidos pelo viés do modelo (MD = viés = 1,715, na tabela 1). É possível notar que a linha dos dados simulados acompanha muito bem os dados observados ao longo do tempo. É isso que faz a correlação e a regressão lineares resultarem em valores altos e significativos. Ou seja, o modelo está superestimando o valor da temperatura do solo em 1,72 graus de forma aproximadamente constante ao longo do ano. Esse é um bom indicativo de que o modelo pode ser aprimorado para essa variável. Como o STICS foi desenvolvido inicialmente para o clima temperado da França, provavelmente é necessário algum ajuste para ser aplicado ao clima tropical do Cerrado de Planaltina, DF.

Então, para efeito de discussão do comportamento das estatísticas e métricas, foi feita a correção do viés no valor simulado (somado ou subtraído, conforme seu sinal) e repetidas as análises, cujos resultados são apresentados na Tabela 2.

Tabela 2. Critérios estatísticos do desempenho do modelo STICS, em simular a temperatura do solo, no período de 22/10/2013 a 30/09/2014, e do modelo ETA-HADGEM em simular a temperatura do ar nos anos de 2013 e 2014, com correção de viés, em Planaltina/DF, Brasil.

Tipo de dados	Variáveis	Métricas usadas na análise de modelos com correção de viés							Análise de regressão com correção de viés			Teste de sinais com correção de viés (viés=MD)				
		Ano	n	EF	MObs	MD	R	RMSE	r ²	F	P>F	viés	T	P>T	S	P>S
Modelo climático ETA	Temp. máxima	2013	365	-1,19	28,17	0,00	0,06	3,59	0,01	1,49	0,2232	0,000	<0,0001	1,0000	958	0,6353
		2014	365	-0,59	28,66	0,00	0,23	3,23	0,05	20,10	<0,0001	0,000	<0,0001	1,0000	418	0,8360
	Temp. mínima	2013	365	-0,57	16,45	0,00	0,43	2,74	0,18	80,49	<0,0001	0,000	<0,0001	1,0000	780	0,6994
		2014	365	-0,06	15,87	0,00	0,47	2,52	0,22	102,09	<0,0001	0,000	<0,0001	1,0000	1.252	0,5357
Modelo de plantas STICS	Temp. do solo Prof. (cm)	0 a 5	82	0,51	22,86	0,00	0,77	1,43	0,59	116,97	<0,0001	0,000	-0,0004	0,9997	-61	0,7816

Fonte: Autores.

Como esperado, os resultados da análise de regressão são idênticos aos da tabela 1, posto que a transformação foi linear (soma ou subtração de uma constante). Também era previsível que o viés fosse eliminado (passou a ser igual a zero) e, portanto, não significativo. O mais importante nessa tabela é observar o comportamento das estatísticas EF e RMSE (já que MD é igual ao viés e também foi zerada). É notável a diminuição do valor negativo da EF e do valor de RMSE para todas as variáveis estimadas pelo modelo ETA, e um valor positivo de EF para a estimativa do modelo STICS (EF = 0,51) bem como a redução quase pela metade do valor de RMSE (de 2,24 para 1,43). Ora, a simples correção de um viés linear mudou completamente o julgamento sobre o modelo STICS para estimativa da temperatura do solo e melhorou o desempenho aparente do modelo ETA. É importante destacar que o modelo STICS é sítio específico, ou seja, ele é calibrado para cada situação específica de solo, clima, cultura e manejo, sendo originalmente destinado a estimativas pontuais. Por isso mesmo, era esperado que ele fosse mais eficiente para uma situação específica que o modelo ETA, que embora tenha o nome de regional, é regional em relação aos modelos

globais, que simulam o clima no mundo todo. Cada ponto (quadricula) do ETA corresponde a 400 km², ou 40.000 ha, e a malha completa cobre a América do Sul, da Patagônia ao Caribe e do oceano Atlântico ao Pacífico, com Andes e floresta amazônica no meio. Ele foi desenvolvido para estimativas sobre extensas áreas e para longo prazo, o presente trabalho apenas pretende demonstrar como se pode avaliar a destreza em situações específicas, que no caso não são o principal objetivo desse modelo.

Mas, sabendo da característica sazonal da temperatura, uma nova tentativa de correção dos dados simulados foi feita com a aplicação de uma componente senoidal ou cossenoidal, dependendo da variável. Os resultados podem ser vistos na Tabela 3.

Tabela 3. Critérios estatísticos do desempenho do modelo STICS, em simular a temperatura do solo, no período de 22/10/2013 a 30/09/2014, e do modelo ETA-HADGEM em simular a temperatura do ar nos anos de 2013 e 2014, com ajuste temporal¹, em Planaltina/DF, Brasil.

Tipo de dados	Variáveis	Métricas usadas na análise de modelos com ajuste temporal							Análise de regressão com ajuste temporal			Teste de sinais com ajuste temporal (viés=MD)					
		Ano	n	EF	MObs	MD	R	RMSE	r ²	F	P>F	viés	T	P>T	S	P>S	
Modelo climático ETA	Temp. máxima	2013	365	-62,05	28,17	-1,53	-0,11	19,26	0,01	4,21	0,0410	-1,530	-1,5204	0,1293	-3.604	0,0739	
		2014	365	-60,36	28,66	-2,07	-0,40	20,01	0,16	70,34	<0,0001	-2,072	-1,9870	0,0477	-4.105	0,0418	
	Temp. mínima	2013	365	-34,43	16,45	2,31	0,61	12,98	0,37	212,43	<0,0001	2,310	3,4489	0,0006	7.102	0,0004	
		2014	365	-25,54	15,87	2,99	0,67	12,64	0,44	290,52	<0,0001	2,989	4,6411	<0,0001	8.880	<0,0001	
Modelo de plantas STICS	Temp. do solo	Prof. (cm)	0 a 5	82	-137,37	22,88	23,57	0,53	23,95	0,28	31,84	<0,0001	23,570	49,986	<0,0001	1.743	<0,0001

¹O ajuste temporal foi feito em relação ao cosseno do dia do ano (dda) para as temperaturas mínima do ar e do solo, e pelo seno do dda para a temperatura máxima do ar.

Fonte: Autores.

Na Tabela 3 podemos ver que a aplicação de uma correção sazonal aos dados simulados pelo STICS, cuja entrada já contém dados diários e, portanto, sazonalizados, de temperatura e precipitação, entre outros, não resulta em melhoria da eficiência do modelo por nenhuma das métricas analisadas. Mesmo para as estimativas do modelo ETA, há apenas uma melhora quando observamos os ajustes das regressões, especialmente para as temperaturas mínimas e para a temperatura máxima de 2014. A partir desses resultados é possível afirmar que os dados simulados pelo modelo ETA-HADGEM devem ser usados com muito cuidado quando aplicados a situações sítio específicas ou para determinado ano.

Embora seja intuitivamente uma boa medida relativa da eficiência do modelo, a métrica EF desconsidera completamente a posição (neste caso, no tempo) em que são observados os afastamentos entre valores simulados e observados. Ou seja, sabemos que um bom modelo deve reproduzir o 'jeitão', a forma ou o caminho percorrido pela variável simulada, ao longo do tempo e da escala de medida; quando a variável cresce no mundo real o modelo deve resultar em valores simulados também crescentes, e vice versa, quando no mundo real a variável diminui de valor, o modelo deve seguir essa mesma tendência. Isso é muito mais importante para um modelo – apresentar um comportamento temporal e dimensional semelhante – do que acertar mais na média.

Sigmae, Alfenas, v.12, n.3, p.213-223, 2023.

67ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria (RBras) e 20º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO)

Conclusões

A métrica EF, apesar de ter um apelo intuitivo interessante, sozinha não é uma boa forma de julgar a eficiência de modelos de simulação de fenômenos naturais, biológicos, que sofrem grande influência tanto da fenologia quanto da meteorologia, ambas fortemente associadas ao passar do tempo, no sentido cronológico. Esses modelos, mesmo que não acertem exatamente os valores observados (afinal existem erros de medida tanto nos valores observados quanto nas variáveis de entrada do modelo), devem ser avaliados pela sua capacidade de predizer a tendência de mudança nas variáveis mais que seu valor exato (inalcançável).

As métricas, se usadas de forma isolada e sem análise aprofundada das características das variáveis simuladas, podem resultar em uma interpretação equivocada da eficiência dos modelos. Por fim, é bom lembrar as recomendações de Nash em seu artigo original sobre EF (Nash e Sutcliffe, 1970). Os modelos não são perfeitos nem devem ser considerados imutáveis, então, um procedimento sistemático de aprimoramento dos mesmos é desejável. Diante de um modelo, deve se medir sua eficiência (seja com EF ou outra estatística) e, caso necessário, modificar o modelo. Se ele for modificado, repetir a análise de eficiência e decidir sobre a aceitação ou rejeição da modificação. Escolher a próxima modificação. Usar gráficos entre valores simulados e observados pode indicar qual modificação é desejável.

Posfácio

Após a submissão deste trabalho ao evento da RBRAS/SEAGRO, uma cópia foi enviada a alguns membros da equipe de desenvolvimento do modelo STICS na França. Diante dos comentários sobre o viés encontrado nos dados simulados para a temperatura do solo, foi identificada a necessidade de alteração de uma informação de entrada utilizada pelos algoritmos do modelo de simulação, que diz respeito ao albedo do solo. Como havia sido suposto ao analisarmos o comportamento dos dados, o solo tropical apresenta um albedo diferente do que estava pré-indicado no modelo. Assim, foi alterado o valor do albedo e o modelo foi novamente aplicado para simular os dados de temperatura do solo de 0 a 5 cm de profundidade, com todas as demais informações mantidas.

O resultado das mesmas análises constantes da Tabela 1, agora para os dados corrigidos pelo novo valor de albedo, são apresentados na Tabela 4. Nesse caso apenas os dados de temperatura do solo de 0 a 5 centímetros foram analisados. Para facilitar a comparação, os mesmos resultados de temperatura do solo, já expostos na Tabela 1, são repetidos na Tabela 4.

Tabela 4. Critérios estatísticos que mostram o desempenho do modelo STICS em simular a temperatura do solo, antes e depois da correção do albedo, no período de 22/10/2013 a 30/09/2014, em Planaltina/DF, Brasil.

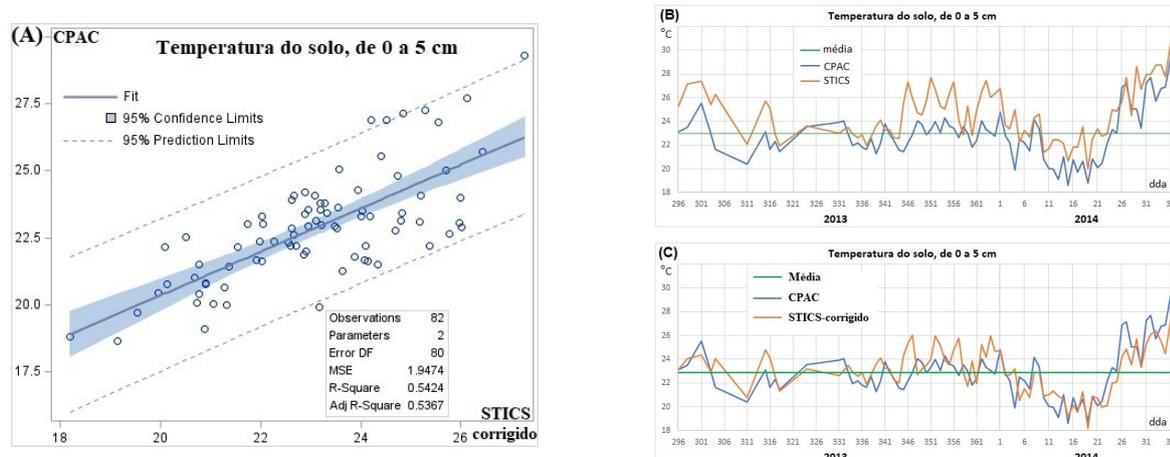
Modelo de plantas STICS	Métricas usadas na análise de modelos						Análise de regressão			Teste de sinais (viés=MD)				
	n	EF	MObs	MD	R	RMSE	r ²	F	P>F	viés	t	P>t	S	P>S
Temperatura do solo de 0 a 5 cm	82	-0,20	22,86	1,72	0,77	2,24	0,59	116,97	<0,0001	1,715	10,771	<0,0001	1.574	<0,0001
Temperatura corrigida do solo de 0 a 5 cm	82	0,50	22,86	0,22	0,74	1,44	0,54	94,82	<0,0001	0,223	1,4099	0,1624	240,5	0,2688

Fonte: Autores.

Sigmae, Alfenas, v.12, n.3, p.213-223, 2023.

67ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria (RBras) e 20º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO)

Figura 2. Comparação gráfica entre valores observados e simulados pelo modelo STICS, da temperatura do solo, depois da correção do albedo, de out/2013 a set/2014, Planaltina/DF, Brasil.



Depois de acertado o valor de entrada para o albedo do solo, os resultados para a análise de viés passam a indicar a ausência de erro sistemático, o que pode ser observado, em comparação com os resultados anteriores à correção, pela diminuição do valor do viés, de 1,715 para 0,223, mas especialmente pelo resultado dos dois testes, t e S, ambos com valores não significativos. Coerente com esse resultado, a análise de regressão apresentou muito pouca alteração, sem impacto na significância do modelo, ou seja, os dados simulados são efetivos no prognóstico dos dados observados. Por fim, as métricas usualmente aplicadas na análise da eficiência dos modelos de simulação sofreram importante modificação, o valor de EF, por exemplo, que era negativo (-0,20), indicando que o modelo seria pior que a simples média na estimativa da temperatura do solo, passou a um valor positivo (0,50), indicando agora que o modelo é eficiente na estimativa dessa variável. Além disso, o erro representado pelo valor de RMSE diminuiu de 2,24 para 1,44, outra evidência da qualidade do modelo.

Bibliografia

AFFIN, O. A. D.; BARCELLOS, J. M.; LUIZ, A. J. B.; SILVA, F. A. M.; CASTRO, L. H. R.; ZOBY, J. L. F. *Uso de modelos matemáticos de simulação para estabelecimento de prioridades de pesquisa*. Planaltina: EMBRAPA-CPAC, 1992. 43 p. (Embrapa CPAC. Documentos, 49).

ANDRADE, C. DE L. T. [et al.] *Modelagem do crescimento de culturas: aplicações à cultura do milho*. Sete Lagoas: Embrapa Milho e Sorgo, 2009. 65 p. (Embrapa Milho e Sorgo. Documentos, 91).

BRISSON, N.; MARY, B.; RIPOCHE, D.; JEUFRROY, M.H.; RUGET, F.; NICOUILLAUD, B.; GATE, P.; DEVIENNE-BARRET, F.; ANTONIOLETTI, R.; DURR, C.; RICHARD, G.; BEAUDOIN, N.; RECOUS, S.; TAYOT, X.; PLENET, D.; CELLIER, P.; MACHET, J.M.; MEYNARD, J.M.; DELECOLLE, R. STICS: a generic model for the simulation of crops and their water and nitrogen balances: I. Theory and parameterization applied to wheat and corn. *Agronomie*, v. 18, p. 311-346, 1998.

Sigmae, Alfenas, v.12, n.3, p.213-223, 2023.

67ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria (RBras) e 20º Simpósio de Estatística Aplicada à Experimentação Agrônômica (SEAGRO)

BRISSON, N.; RUGET, F.; GATE, P.; LORGEAU, J.; NICOUILLAUD, B.; TAYOT, X.; PLENET, D.; JEUFFROY, M. H.; BOUTHIER, A.; RIPOCHE, D.; MARY, B.; JUSTES, E. STICS: a generic model for simulating crops and their water and nitrogen balances: II. Model validation for wheat and maize. *Agronomie*, v. 22, p. 69-92, 2002.

CHOU, S. C.; LYRA, A.; MOURÃO, C.; DEREZYNSKI, C.; PILOTTO, I.; GOMES, J.; BUSTAMANTE, J.; TAVARES, P.; SILVA, A.; RODRIGUES, D.; CAMPOS, D.; CHAGAS, D.; SUEIRO, G.; SIQUEIRA, G.; MARENGO, J. Assessment of climate change over South America under RCP 4.5 and 8.5 downscaling scenarios. *American Journal of Climate Change*, v. 3, p. 512-527, 2014.

DESCARTES, R. *Discurso do método*. Martins Fontes, São Paulo, 2001. 102 p.

LUIZ, A. J. B.; MAIA, A. de H. N. Consistência entre dados originais, interpolações globais e projeções do modelo climático regional EtaHadGEM2-ES para a precipitação nas bacias dos rios Paracatu e São Marcos. In: SEMINÁRIO DA REDE AGROHIDRO, 4., 2016, Brasília, DF. Água e agricultura: incertezas e desafios para a sustentabilidade frente às mudanças do clima e do uso da terra: *Anais*. Brasília, DF: Embrapa, 2016. p. 152-159.

MACHADO FILHO, H.; MORAES, C.; BENNATI, P.; RODRIGUES, R. DE A.; GUILLES, M.; ROCHA, P.; LIMA, A.; VASCONCELOS, I. *Mudança do clima e os impactos na agricultura familiar no Norte e Nordeste do Brasil*. IPC-IG Working Paper, n.141, 2016. 68p. Disponível em:

https://ipcig.org/pub/port/Mudanca_no_clima_e_os_impactos_na_agricultura_familiar.pdf

MAIA, A. de H. N.; PAZIANOTTO, R. A. A.; COELHO, C. A. dos S. Avaliação da influência do ENOS sobre início da estação chuvosa via modelos semiparamétricos de sobrevivência. In: REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 56.; SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 14., 2011, Maringá. *Anais...* Maringá: Universidade Estadual de Maringá; Embrapa Pecuária Sudeste, 2011. 5 p.

MINITTI, A. F. *Relatório de avaliação dos impactos de tecnologias geradas pela Embrapa*. Campinas: Embrapa Informática Agropecuária, 2020. 57 p. (Embrapa Informática Agropecuária. Documentos, 170)

NASH, J. E.; SUTCLIFFE J. V. River flow forecasting through conceptual models: Part I - A discussion of principles. *Journal of Hydrology*, v. 10, n. 3, p. 282-290, 1970.

SAS Institute Inc. *SAS/STAT® 9.4 Users Guide*. Cary, NC: SAS Institute Inc. 2013.

SILVA, F. A. M. DA; EVANGELISTA, B. A.; MALAQUIAS, J. V. *Normal climatológica de 1974 a 2003 da Estação Principal da Embrapa Cerrados*. Planaltina, DF: Embrapa Cerrados, 2014. 98 p. (Embrapa Cerrados. Documentos, 321).

VICENTE, P. O uso de simulação como metodologia de pesquisa em ciências sociais. *Cadernos EBAPE.BR*, v. 3, n. 1, p. 1-9, 2005. Disponível em: <https://shorturl.at/bclMX>.

Sigmae, Alfenas, v.12, n.3, p.213-223, 2023.

67ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria (RBras) e 20º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO)