

# ANÁLISE DE TESTES ESTATÍSTICOS PARA COMPARAÇÃO DE DADOS DE ESTAÇÕES METEOROLÓGICAS E DE REANÁLISE PARA UM SEMESTRE EM PORTO ALEGRE

FRANCIELE DA COSTA TRASSANTE<sup>1</sup>; DOUGLAS DA SILVA LINDEMANN<sup>2</sup>; JAIRO VALÕES DE ALENCAR RAMALHO<sup>3</sup>; JOSEANE DA SILVA PORTO<sup>4</sup>; RUTH DA SILVA BRUM<sup>5</sup>

<sup>1</sup>Universidade Federal de Pelotas – [francieletrassante@gmail.com](mailto:francieletrassante@gmail.com)

<sup>2</sup>Universidade Federal de Pelotas – [douglasdasilva.lindemann@gmail.com](mailto:douglasdasilva.lindemann@gmail.com)

<sup>3</sup>Universidade Federal de Pelotas – [j.v.a.ramalho@gmail.com](mailto:j.v.a.ramalho@gmail.com)

<sup>4</sup>Universidade Federal de Pelotas – [joseaneclmd@gmail.com](mailto:joseaneclmd@gmail.com)

<sup>5</sup>Universidade Federal de Pelotas – [ruthdasilvabrum@gmail.com](mailto:ruthdasilvabrum@gmail.com)

## 1. INTRODUÇÃO

Atualmente, o acesso a banco de dados de estações meteorológicas automáticas está cada vez mais simples. Entretanto, a carência de estações em cidades menores e também as falhas apresentadas no armazenamento dos dados coletados são alguns dos problemas enfrentados, que ocorrem devido à diversas razões, causando ausência de registros que podem inclusive impossibilitar o seu uso.

Uma alternativa é o uso de dados de reanálise atmosféricas. Uma reanálise consiste em desenvolver e operar regularmente, dados meteorológicos globais e sistemas de assimilação de dados para construção de bancos de dados climáticos, com base na entrada de dados meteorológicos coletados por satélites e sistemas de observação da Terra, como estações automáticas, aeronaves, navios e balões meteorológicos. (Valeriano et al., 2019)

Sendo assim, o presente trabalho busca analisar através de testes estatísticos o desempenho da reanálise ERA5 com os dados observados de uma estação meteorológica na cidade de Porto Alegre, de modo que as informações obtidas possam ser incorporadas em estudos futuros relacionados à utilização de dados de reanálise.

## 2. METODOLOGIA

Neste estudo, foram utilizados dados horários que posteriormente foram transformados em diários de temperatura média (Tmed), máxima (Tmax) e mínima (Tmin) do ar (°C), no período entre 01/01/2007 à 30/06/2007, obtidos através da estação meteorológica automática do Instituto Nacional de Meteorologia (INMET), situada na cidade de Porto Alegre (30.05° S e 51.17° O). Foi utilizado também dados da reanálise ERA5, a quinta geração de reanálises atmosféricas pertencente ao *European Centre for Medium-Range Weather Forecasts* (ECMWF) do clima global (Hersbach et al., 2020).

O período de dados estudado servirá como base para futuros trabalhos que serão desenvolvidos no projeto de pesquisa intitulado "Trocadores de calor solar", onde serão analisados dados obtidos do ano de 2007 através da reanálise ERA5.

Realizou-se uma análise estatística semestral e mensal nos dados do ERA5 através dos testes de coeficiente de correlação (r), onde foi usado a classificação de Hopkins (2009), índice de concordância (d) proposto por Willmott et al., 1985, raiz do erro médio quadrático (RMSE, em inglês) (Loague; Green, 1991), que

representa a variação entre o dado estimado e o observado elevados ao quadrado, que resultam apenas valores positivos, sendo que quanto menor o valor, melhor a estimativa, e o erro do desvio médio (MBE, em inglês), que representa o desvio médio entre o dado observado e o predito, onde os valores abaixo de zero indicam subestimação e valores acima superestimação em relação aos dados observados. Os testes foram tratados e calculados no programa Microsoft Excel, com o propósito de avaliar o desempenho das estimativas encontradas na comparação de dados meteorológicos observados e os de reanálise. Em Lindemann (2012), encontra-se maiores detalhes sobre os testes estatísticos aqui aplicados.

### 3. RESULTADOS E DISCUSSÃO

De acordo com o teste de correlação ( $r$ ) ótimos desempenhos foram encontrados, em todas as análises as correlações entre o ERA5 e a estação automática para as temperaturas do ar ficaram acima de 0.9, exceto para fevereiro ( $T_{max}$  e  $T_{min}$ ) e março ( $T_{min}$ ). A  $T_{min}$  indicou resultado de correlação com o menor valor (0.76) em abril, enquanto as demais análises variaram entre 0.8 e 0.9.

A aplicação do índice de concordância ( $d$ ) apresentou comportamento semelhante à  $r$  como mostra a Tabela 1, com a maioria das variáveis indicando concordância com valores acima de 0.9. A menor concordância foi observada na  $T_{max}$  em abril (0.92), e as maiores concordâncias foram observadas nos meses de maio e junho referente à  $T_{min}$  e de junho referente à  $T_{med}$ , que obtiveram índices máximos. Os demais valores encontraram-se todos acima de 0.95 para as temperaturas, indicando que a reanálise apresenta uma boa condição de exatidão, em relação aos dados da estação meteorológica.

**Tabela 1.** Cálculo de índice de concordância semestral e mensal de janeiro a junho de 2007.

<b>d</b>	<b>Semestre</b>	<b>JAN</b>	<b>FEV</b>	<b>MAR</b>	<b>ABR</b>	<b>MAI</b>	<b>JUN</b>
<b>Tmax</b>	0,98	0,96	0,96	0,97	0,92	0,99	0,99
<b>Tmin</b>	0,99	0,99	0,96	0,99	0,98	1	1
<b>Tmed</b>	0,99	0,99	0,99	0,99	0,95	0,99	1

A raiz do erro médio quadrático médio (RMSE, em inglês) indicou valores próximos a 1 em todas as variáveis de temperatura do ar, sendo a  $T_{min}$  apresentando os valores mais baixos no semestre e nos meses de janeiro, abril, maio e junho (entre 0.86 e 1.34). Essa maior variabilidade dos resultados pode ser relacionado pela alta frequência de obtenção dos dados, que originalmente eram horários, portanto, pesos diferentes poderão ser atribuídos devido a esta variabilidade e quanto maior a variação do dado observado e o de reanálise, o índice RMSE aumenta consideravelmente, o que explica os resultados obtidos.

**Tabela 2.** Cálculo do erro da raiz do quadrático médio (RMSE) semestral e mensal de janeiro a junho de 2007.

<b>RMSE</b>	<b>Semestre</b>	<b>JAN</b>	<b>FEV</b>	<b>MAR</b>	<b>ABR</b>	<b>MAI</b>	<b>JUN</b>
<b>Tmax</b>	1,78	1,98	1,97	1,82	1,87	1,28	1,71
<b>Tmin</b>	1,02	0,86	1,34	1,04	0,97	0,99	0,90

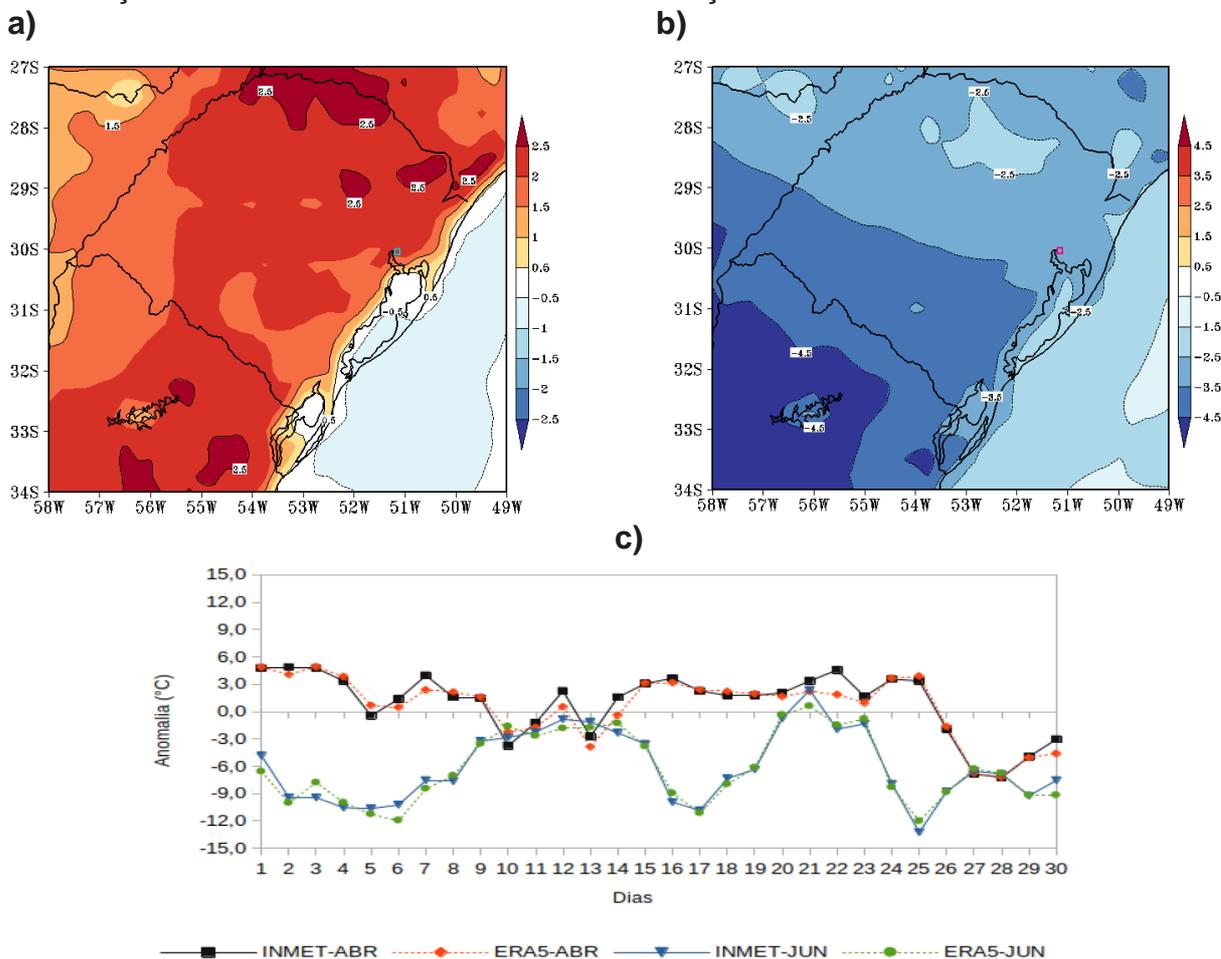
<b>Tmed</b>	1,08	0,94	1,01	1,02	1,38	1,13	0,92
-------------	------	------	------	------	------	------	------

O erro de desvio médio (MBE, em inglês) representado na Tabela 3., indica que há subestimação em todos os dados de temperatura máxima e média da reanálise. Os valores da *Tmin* foram os que apresentaram valores próximos a zero, exceto os meses de janeiro (0.37) e abril (-0.27).

**Tabela 3.** Cálculo do erro de desvio médio (MBE) semestral e mensal de janeiro a junho de 2007.

MBE	Semestre	JAN	FEV	MAR	ABR	MAI	JUN
<b>Tmax</b>	-0,99	-1,53	-1,12	-1,19	-1,46	-0,56	-0,03
<b>Tmin</b>	0,05	0,37	0,19	0,14	-0,27	-0,04	-0,08
<b>Tmed</b>	-0,66	-0,65	-0,55	-0,64	-1,10	-0,69	-0,30

Observou-se nas normais climatológicas do INMET que em 2007 foram registradas anomalias de  $\approx 2^{\circ}\text{C}$  acima da média nas temperaturas mínimas do mês de abril, e anomalias de  $\approx -2^{\circ}\text{C}$  abaixo da média nas temperaturas mínimas no mês de junho na cidade de Porto Alegre. Como uma forma de exemplificar essas variações espaciais, a Figura 1 a) e Figura 1 b) demonstram estas anomalias registradas no Rio Grande do Sul. Desse modo, observa-se com maior clareza na Figura 1. c) a série temporal com as anomalias indicando que mesmo em meses onde encontra-se padrões totalmente adversos de temperaturas, a reanálise consegue representar esses valores com um ótimo desempenho, com pequenas diferenças entre os extremos nos dados de estação e da reanálise.



**Figura 1** – a) Anomalia da *Tmin* para abril de 2007; b) Anomalia da *Tmin* para junho de 2007. Os retângulos nas Figuras a) e b) representam a posição da estação automática; c) Série temporal com as anomalias de *Tmin* da estação automática do INMET (linhas cheias) e ERA5 (linhas pontilhadas) para abril e junho.

#### 4. CONCLUSÕES

Conclui-se que na ausência de dados de estações meteorológicas é possível obter boas estimativas através dos dados de reanálise. Os objetivos foram alcançados, visto que os testes empregados apresentaram um bom desempenho geral, com correlações próximas a 0.9 e um ótimo índice de concordância, em que todos valores indicados foram acima de 0.95 provando sua exatidão.

Com base nas análises das figuras e gráfico, foi possível representar o desempenho da reanálise em relação aos dados da estação, provando que mesmo em condições meteorológicas totalmente diferentes seus resultados foram muito satisfatórios.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

Valeriano, T.T.B., de Souza Rolim, G., Bispo, R.C. et al. Evaluation of air temperature and rainfall from ECMWF and NASA gridded data for southeastern Brazil. **Theoretical and Applied Climatology**, v 137, 1925–1938 (2019).

HERSBACH, H. et al. The ERA5 global reanalysis. **Quarterly Journal of the Royal Meteorological Society**, p. 1999-2049, 2020.

INMET – Instituto Nacional de Meteorologia, BDMEP – Banco de dados Meteorológicos para o ensino e pesquisa. Disponível em: <https://portal.inmet.gov.br/>. Acesso em: 22/09/2020

LINDEMANN, D. **Variações na temperatura no Continente Antártico: observações e reanálises**. 2012. Dissertação de Mestrado em Meteorologia Agrícola – Programa de Pós Graduação em Meteorologia Agrícola, Universidade Federal de Viçosa (UFV).

HOPKINS, W. G. **A New View of Statistics**, 2009. Disponível em: <http://www.sportsci.org/resource/stats/effectmag.html> Acesso em: 22/09/2020

WILLMOTT, C.J.; AKLESON, G.S.; DAVIS, R.E. et al. Statistic for the evaluation and comparison of models. **Journal of Geophysical Research**, v. 90, p. 8995-9005, 1985.

LOAGUE, K.; GREEN, R.E. Statistical and graphical methods for evaluating solute transport models: overview and application, **Journal of Contaminant Hydrology**, v. 7, p. 51-73, 1991.