

# Processamento de dados de temperatura do solo para simulações computacionais de um agro-ecossistema

## Processing of soil temperature data for computational simulations of an agroecosystem

M.C. Paz<sup>1,\*</sup>, S.A.P. Santos<sup>1,2</sup> & R. Barreira<sup>3,4</sup>

<sup>1</sup> Instituto Politécnico de Setúbal, ESTBarreiro, CIQuiBio

<sup>2</sup> LEAF, Instituto Superior de Agronomia, Tapada da Ajuda, 1349-017 Lisboa, Portugal

<sup>3</sup> Instituto Politécnico de Setúbal, ESTBarreiro, INCITE

<sup>4</sup> CMAFcIO, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisbon, Portugal

(\*E-mail: catarina.paz@estbarreiro.ips.pt)

<https://doi.org/10.19084/rca.28405>

### RESUMO

Os serviços de ecossistema, como a limitação natural de pragas, incluem-se nas estratégias de gestão de agro-ecossistemas e o seu uso pode ser otimizado com base no conhecimento provindo da modelação computacional das pragas, dos seus predadores, e da paisagem. Neste artigo focamo-nos no processamento de dados de temperatura do solo, necessários para o funcionamento dos modelos de praga-predador *Bactrocera oleae* (mosca da azeitona) e *Haplodrassus rufipes* (aranha do solo) no olival. A metodologia de processamento permitiu-nos (1) completar as lacunas existentes na série de dados de temperatura do solo, recolhidos com periodicidade horária, (2) convertê-la numa série com periodicidade diária, e ainda (3) criar uma variável adicional, a temperatura média diária do solo durante o crepúsculo vespertino. Este último passo possibilita que não haja uma perda de informação no que diz respeito àquele período específico do dia, reduzindo-se o erro associado à desajustação da resolução temporal da variável climática média diária para exprimir um fenómeno que ocorre apenas durante um determinado número de horas do dia.

**Palavras-chave:** temperatura do solo, processamento de séries climáticas, conversão de resoluções temporais, linguagem R, modelação computacional de agro-ecossistemas

### ABSTRACT

Ecosystem services, such as natural pest control, are included in agro-ecosystem management strategies and their use can be optimized based on knowledge from computational modelling of pests, predators, and landscape. In this article we focus on the processing of soil temperature data, necessary for the operation of the predator-pest models *Bactrocera oleae* (olive fly) and *Haplodrassus rufipes* (soil spider) in the olive grove. The processing methodology allowed us to (1) complete the gaps in the series of soil temperature data, collected on an hourly basis, (2) convert it into a series with daily periodicity, and (3) create an additional variable, the average daily temperature of the soil during the twilight. This last step prevents the loss of information regarding that specific period of the day, reducing the error associated with the deajustation of the temporal resolution of the average daily climatic variable to express a phenomenon that occurs only during a certain number of hours of the day.

**Keywords:** soil temperature, climate series processing, conversion of temporal resolutions, R language, computational modelling of agroecosystems

## INTRODUÇÃO

A limitação natural de pragas é um serviço de ecossistema que consiste no controlo das espécies prejudiciais às culturas agrícolas através dos seus inimigos naturalmente presentes no ecossistema, e que pode ser promovido através do estabelecimento de componentes paisagísticas que forneçam alimento alternativo e refúgio a esses inimigos (Villa *et al.*, 2020, 2021). O uso deste serviço pode ser otimizado conhecendo o seu funcionamento, o que pode ser feito através da modelação computacional das pragas, dos seus predadores, e da paisagem (Corral & Calegari, 2011; Topping *et al.*, 2019; Ziołkowska *et al.*, 2021). Neste contexto, o projeto OLIVESIM está a construir um sistema de modelos para simular os comportamentos de espécies de pragas da oliveira e de seus inimigos, na região de Trás-os-Montes (Portugal), através do sistema ALMaSS (Animal, Landscape and Man Simulation System), que recebe séries de dados de ações de gestão agrícola e de variáveis climáticas, com periodicidade diária (Topping *et al.*, 2003). De facto, para alguns dos artrópodes modelados, a temperatura do solo é usada como dado de entrada na modelação computacional, uma vez que a temperatura desempenha um papel fundamental nas várias fases da sua vida e também porque interação espacialmente ao nível do solo.

Para que a modelação computacional possa funcionar, as séries de dados climáticos devem ter uma estrutura temporal regular. No entanto, muitas vezes não estão completas, o que exige a aplicação de metodologias fiáveis para o preenchimento das lacunas existentes. Por outro lado, uma série de dados climáticos é formada por dados recolhidos com uma dada periodicidade, que pode ser diferente da necessária para o sistema computacional, o que implica uma conversão de periodicidade. No caso da conversão de séries de dados com periodicidade horária em séries de dados com periodicidade diária, que neste caso são as usadas pelo sistema ALMaSS, há uma redução da resolução temporal e, portanto, uma perda de informação no que diz respeito a períodos específicos do dia. Isto pode ser um problema quando se tenta simular, por exemplo, um comportamento animal que ocorre durante um período específico do dia. Uma forma de segregar esse período específico é criar variáveis diárias que são calculadas usando apenas os dados horários compreendidos nesse período do dia.

Neste trabalho, debruçamo-nos sobre o processamento da série de temperatura do solo necessária para o funcionamento dos modelos de praga-predador *Bactrocera oleae* (mosca da azeitona) (Paz *et al.*, 2021) e *Haplodrassus rufipes* (aranha do solo) (Barreira *et al.*, 2021), ambos artrópodes dependentes da temperatura do solo em certas fases ou durante toda a sua vida. De facto, a aranha habita os primeiros centímetros do solo, e a mosca pupa na mesma zona durante os meses mais frios do ano, tornando-se um potencial alimento para a aranha. Completámos as lacunas existentes na série de dados recolhida, e convertimos a sua periodicidade de horária para diária, criando também uma variável adicional, a temperatura média diária do solo durante o crepúsculo vespertino, que é o momento em que a aranha normalmente se move para caçar (Benhadi-Marín, J., comunicação pessoal).

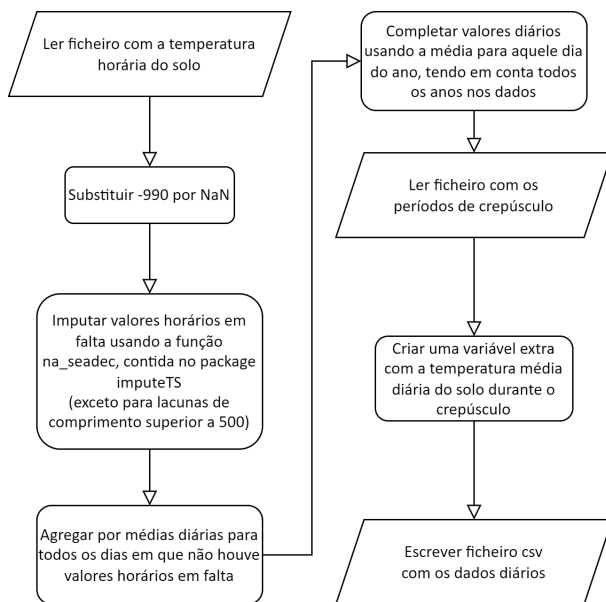
## MATERIAL E MÉTODOS

A temperatura do solo foi medida com periodicidade horária, de 2010 a 2020, a uma profundidade de 0.05 m, na estação meteorológica de Mirandela. Esta estação faz parte da rede de estações meteorológicas automáticas do Instituto do Mar e da Atmosfera (IPMA). O Quadro 1 resume as características da série de dados recolhida, que foi depois processada utilizando a linguagem R (R Core team, 2020), seguindo a sequência apresentada na Figura 1.

**Quadro 1** - Características da série de dados de temperatura do solo medidos com periodicidade horária na estação meteorológica de Mirandela. Entende-se por lacuna um grupo de valores em falta seguidos

Comprimento da série	96360
Número de valores em falta	9172
Percentagem de valores em falta	10%
Número de lacunas	88
Comprimento de lacuna mais frequente	1
Maior comprimento de lacuna	4285

Depois de completada a série de dados horários, calculou-se a temperatura média diária do solo (Ts), efetuando a média das 24 medições horárias de cada dia. Para o cálculo da temperatura média diária do solo durante o crepúsculo (Tsc) foi



**Figura 1** - Sequência de processamento dos dados de temperatura do solo. A série de dados de temperatura do solo, medidos com periodicidade horária no período 2010-2020, apresenta, como esperado, um padrão de sazonalidade, pelo que se qualifica para o uso da função `na_seadec`, contida no pacote `imputeTS`, na tarefa de imputação de valores nas lacunas com comprimento inferior a 500 valores consecutivos em falta. O pacote `imputeTS` foi especificamente desenvolvido para a imputação de valores em falta em séries temporais univariadas (Moritz & Bartz-Beielstein, 2017).

definido o período de crepúsculo como o período entre o pôr-do-sol e o fim do crepúsculo astronómico, isto é, quando o Sol está  $18^\circ$  abaixo do horizonte.  $T_{sc}$  foi então calculada efetuando a média ponderada dos dados de temperatura do solo medidos com periodicidade horária durante o período do crepúsculo.

## RESULTADOS E DISCUSSÃO

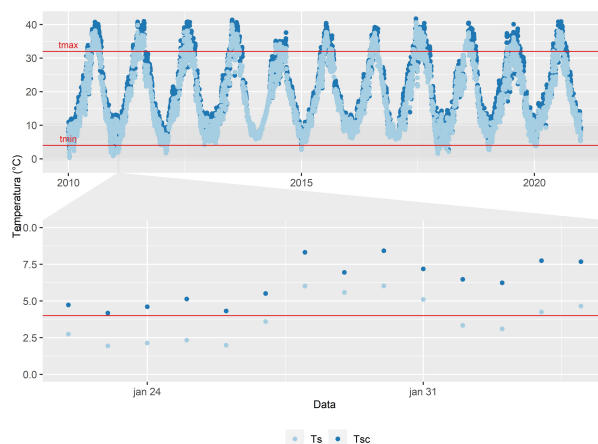
A Figura 2 mostra a série de dados de temperatura do solo, medidos com periodicidade horária, recolhida na estação meteorológica de Mirandela, e os dados imputados em lacunas inferiores a 500 valores consecutivos em falta. Podemos verificar que os valores imputados se enquadram perfeitamente na série.

A Figura 3 mostra as séries de  $T_s$  e  $T_{sc}$ , calculadas para o período 2010-2020, permitindo verificar



**Figura 2** - Série de dados de temperatura do solo, medidos com periodicidade horária (círculos azuis-claros), recolhida na estação meteorológica de Mirandela, no período 2010-2020, e dados imputados em lacunas inferiores a 500 valores consecutivos em falta (circunferências vermelhas), usando a função `na_seadec` contida no pacote `imputeTS`.

que  $T_{sc}$  é em geral superior a  $T_s$ . A Figura 3 mostra também, a título de exemplo, o período de 22 de janeiro a 4 de fevereiro de 2011, ano durante o qual se realizou uma campanha de amostragem de artrópodes e solos, e os limites do intervalo de valores de temperatura dentro do qual o movimento da aranha ocorre. Podemos ver que em alguns dias, por exemplo no dia 24 de janeiro de 2011,  $T_{sc}$  está acima do limite inferior ( $t_{min}$ ) e  $T_s$  está abaixo do mesmo, o que significa que se usássemos  $T_s$



**Figura 3** - Séries de temperatura média diária do solo ( $T_s$  –  $^\circ\text{C}$ ) e temperatura média diária do solo durante o crepúsculo ( $T_{sc}$  –  $^\circ\text{C}$ ), para o período 2010-2020, calculadas a partir da série de dados de temperatura do solo medidos com periodicidade horária, recolhida na estação meteorológica de Mirandela. As linhas vermelhas representam as temperaturas mínima ( $t_{min}$ ) e máxima ( $t_{max}$ ) entre as quais é desencadeado o movimento da aranha.

em vez de  $T_{sc}$  para simular o movimento da aranha, estaríamos a impedi-la de mover-se, quando deveria estar a fazê-lo. Isto poderá também acontecer em redor do limite superior, mas de forma inversa – a aranha estar em movimento, quando já deveria estar em repouso. Ou seja, o uso de  $T_s$ , que exprime a temperatura média das 24h do dia, em vez de  $T_{sc}$ , que exprime especificamente a temperatura durante o período do dia em que a aranha se move, resultaria numa simulação incorreta do seu movimento. Desta forma, podemos ver como a utilização de dados climáticos temporais de alta resolução oferece a oportunidade de modular aspetos que, de outro modo, não seriam perceptíveis (Afrifa-Yamoah *et al.*, 2020).

## CONCLUSÕES

A metodologia aplicada permitiu-nos (1) completar as lacunas existentes na série de dados de temperatura do solo, recolhidos com periodicidade horária, (2) convertê-la numa série com periodicidade diária, e ainda (3) criar uma variável adicional, a temperatura média diária do solo durante o crepúsculo

vespertino. Este último passo possibilita que não haja uma perda de informação no que diz respeito àquele período específico do dia, reduzindo-se o erro associado à desadequação da resolução temporal da variável climática média diária para exprimir um fenómeno que ocorre apenas durante um determinado número de horas do dia.

As séries criadas permitirão uma simulação mais correta das pragas e seus predadores, contribuindo para melhores definições de estratégias de gestão dos agro-ecossistemas.

## AGRADECIMENTOS

Este trabalho é financiado pela Fundação para a Ciência e Tecnologia (FCT) através do projeto PTDC/ASP-PLA/30003/2017 – OLIVESIM – Gestão dos serviços de ecossistema no olival utilizando modelos espaciais avançados. R.B. financiada pela FCT através do projeto UIDB/04561/2020.

As autoras agradecem ao IPMA a cedência dos dados de temperatura do solo.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Afrifa-Yamoah, E.; Mueller, U.; Taylor, S. & Fisher, A. (2020) - Missing data imputation of high-resolution temporal climate time series data. *Meteorological Applications*, vol. 27, n. 1, art. e1873. <https://doi.org/10.1002/met.1873>
- Barreira, R.; Paz, M.C.; Amaro, L.; Sousa, J.P.; Benhadi-Marín, J.; Rasko, M.; Alves da Silva, A.; Alves, J.; Chuhutin, A.; Topping, C.J. & Santos, S.A.P. (2021) - Developing an Agent-Based Model for *Haplodrassus rufipes* (Araneae: Gnaphosidae), a Generalist Predator Species of Olive Tree Pests: Conceptual Model Outline. *Biology and Life Sciences Forum*, vol. 4, n. 1, art. 45. <https://doi.org/10.3390/IECPS2020-08745>
- Corral, J. & Calegari, D. (2011) - Towards an Agent-Based Methodology for Developing Agro-Ecosystem Simulations. In: Barthe, G.; Pardo, A. & Schneider, G. (Eds.) - *SEFM 2011: Software Engineering and Formal Methods*, p. 431-446. Springer.
- Moritz, S. & Bartz-Beielstein, T. (2017) - imputeTS: Time series missing value imputation in R. *The R Journal*, vol. 9, n. 1, p. 207–218. <https://doi.org/10.32614/RJ-2017-009>
- Paz, M.C.; Santos S.A.P.; Barreira, R.; Rasko, M.; Duan, X.; Alves, J.; Alves da Silva, A.; Mina, R.; Topping, C.J. & Sousa, J.P. (2021) - *Developing a subpopulation-based model for the olive fruit fly Bactrocera oleae (Diptera: Tephritidae): conceptual model outline*. The 1st International Electronic Conference on Agronomy. *In press*.
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>
- Topping, C.J.; Hansen, T.S.; Jensen, T.S.; Jepsen, J.U.; Nikolajsen, F. & Odderskær, P. (2003) - ALMaSS, an agent-based model for animals in temperate European landscapes. *Ecological Modelling*, vol. 167, n. 1-2, p. 65–82. [https://doi.org/10.1016/S0304-3800\(03\)00173-X](https://doi.org/10.1016/S0304-3800(03)00173-X)
- Topping, C.J.; Dalby, L. & Valdez, J.W. (2019) - Landscape-scale simulations as a tool in multi-criteria decision making for agri-environment schemes. *Agricultural Systems*, vol. 176, art. 102671. <https://doi.org/10.1016/j.agsy.2019.102671>
- Villa, M.; Santos, S.A.P.; Aguiar, C. & Pereira, J. (2021) - Plants Biodiversity in Olive Orchards and Surrounding Landscapes from a Conservation Biological Control Approach. *Biology and Life Sciences Forum*, vol. 4, n. 1, art. 66. <https://doi.org/10.3390/IECPS2020-08604>
- Villa, M.; Santos, S.A.P.; Pascual, S. & Pereira, J. (2020) - Do non-crop areas and landscape structure influence dispersal and population densities of male olive moth? *Bulletin of Entomological Research*, vol. 111, n. 1, p. 1–9. <https://doi.org/10.1017/S0007485320000310>
- Ziółkowska, E.; Topping, C.J.; Bednarska, A.J. & Laskowski, R. (2021) - Supporting non-target arthropods in agroecosystems: Modelling effects of insecticides and landscape structure on carabids in agricultural landscapes. *Science of the Total Environment*, vol. 774, art. 145746. <https://doi.org/10.1016/j.scitotenv.2021.145746>