

ATIVIDADE 3 - Elaboração da seção de metodologia

Aluno: Wesley Lourenco Barbosa

NUSP: 10509976

Título do artigo: Problemas de Qualidade na Aquisição e no Processo de Análise de Dados Bioclimáticos Utilizados em Modelos de Distribuição de Espécies: Uma Revisão Sistemática da Literatura

Objetivo do artigo: Identificar os problemas de qualidade observados nos dados bioclimáticos utilizados em modelos de distribuição de espécies, tanto do ponto de vista da aquisição quanto do processo de análise, e investigar seus efeitos nos modelos.

Seleção do Veículo

O veículo escolhido para submissão do artigo é a PLOS ONE, revista científica multidisciplinar publicada pela Public Library of Science, que possui fator de impacto 3.752 e classificação Qualis A1 na área de Engenharias IV, Computação e Biodiversidade.

Abordagem Metodológica

Uma Revisão Sistemática da Literatura (RSL) permite criar uma base sólida para o avanço do conhecimento científico, fornecendo subsídios para descobrir lacunas de pesquisa (Webster & Watson, 2002). O processo empregado nesta RSL foi proposto por (Kitchenham & Chartes, 2007) que estabelece uma abordagem para revisões de literatura em computação. Este processo também foi utilizado por outras revisões sistemáticas como em (Rekik et al., 2018; Schneider et al., 2013; Sizo et al., 2019).

Na fase de planejamento da RSL, inclui-se as especificações das questões de pesquisa (RQ) e um conjunto de critérios de inclusão e exclusão criados para selecionar os artigos apropriados. A Tabela 1 mostra as 3 RQs que tentou-se responder nesta RSL, seguidas da motivação de estabelecer a respectiva pergunta.

Table 1: Research Questions

QUESTION	MOTIVATION
RQ1: What are the main DQ problems identified in data used to build SDM?	Conducting a survey into the main problems mentioned in the literature on three types of data used in SDM: occurrence, collected by CS projects, and climate.
RQ2: What are the strategies adopted to solve DQ problems?	Determining the main forms of correction adopted for the DQ problems identified.

As bases de dados de pesquisa utilizadas para selecionar os artigos foram: Web of Science (WoS), Scopus e ACM Digital Library. A escolha dessas bases de dados foi baseada em critérios de qualidade (Martín-Martín et al., 2018; Šubelj et al., 2015), relevância para o campo da computação e estudos de biodiversidade, e completude dos metadados retornados por esses repositórios.

Cada uma das bases escolhidas possui diferentes procedimentos de busca; consequentemente, 6 strings de busca semanticamente equivalentes foram criadas para cada RQ. Como exemplo, a Tabela 2 mostra a estrutura da consulta de pesquisa usada na WoS.

Table 2: Search Queries

RESEARCH QUESTION	QUERY
Q1.1	TS = (((“species distribution” OR “ecolog* niche”)) AND “data” AND (“quality” OR “error”)) NOT TS = ('marine' OR 'water' OR 'fish' OR 'sea' OR 'ocean*' OR 'river*' OR 'aqua*')
Q1.2	TS = ((“climat*” AND (“species distribution” OR “ecolog* niche”)) AND “data” AND (“quality” OR “error”) AND (“assess*” OR “control*” OR “measure*” OR “eval*” OR “estimat*” OR “analysis”)) NOT TS = ('marine' OR 'water' OR 'fish' OR 'sea' OR 'ocean*' OR 'river*' OR 'aqua*')
Q1.3	TS = (“citizen science” AND “data quality” AND (“error*” OR “problem*” OR “challenge”))
Q2	TS = (“species distribution*” AND “data” AND “error*” AND (“solution” OR “solv*” OR “fix” OR “correct*” OR “detect”))

A Tabela 3 apresenta os critérios de inclusão e exclusão definidos. O intervalo de tempo foi definido para garantir que os resultados descrevam o cenário mais atual.

Table 3: Inclusion and exclusion criteria

INCLUSION CRITERIA		EXCLUSION CRITERIA
1	Published between January 2013 and July 2022	Published outside our time span
2	Full text available	Not written in English
3	Published in journals or proceedings	Topic related to marine ecosystems
4	Queried database provided the abstract within metadata (bib file)	Books

A revisão centra-se na distribuição das espécies terrestres. A exclusão das espécies marinhas deve-se à existência de diferenças significativas nos processos ecológicos marinhos e às características intrínsecas deste tipo de ecossistema (Carr et al., 2003).

A Tabela 4 mostra o número de resultados retornados por banco de dados para cada consulta de pesquisa (Tabela 3), sem remover duplicatas. Os resultados da Tabela 4 são meramente informativos, não é objetivo desta RSL comparar as bases de dados de artigos científicos. Portanto, o detalhamento pelo banco de dados de pesquisa é apenas uma parte do processo de RSL.

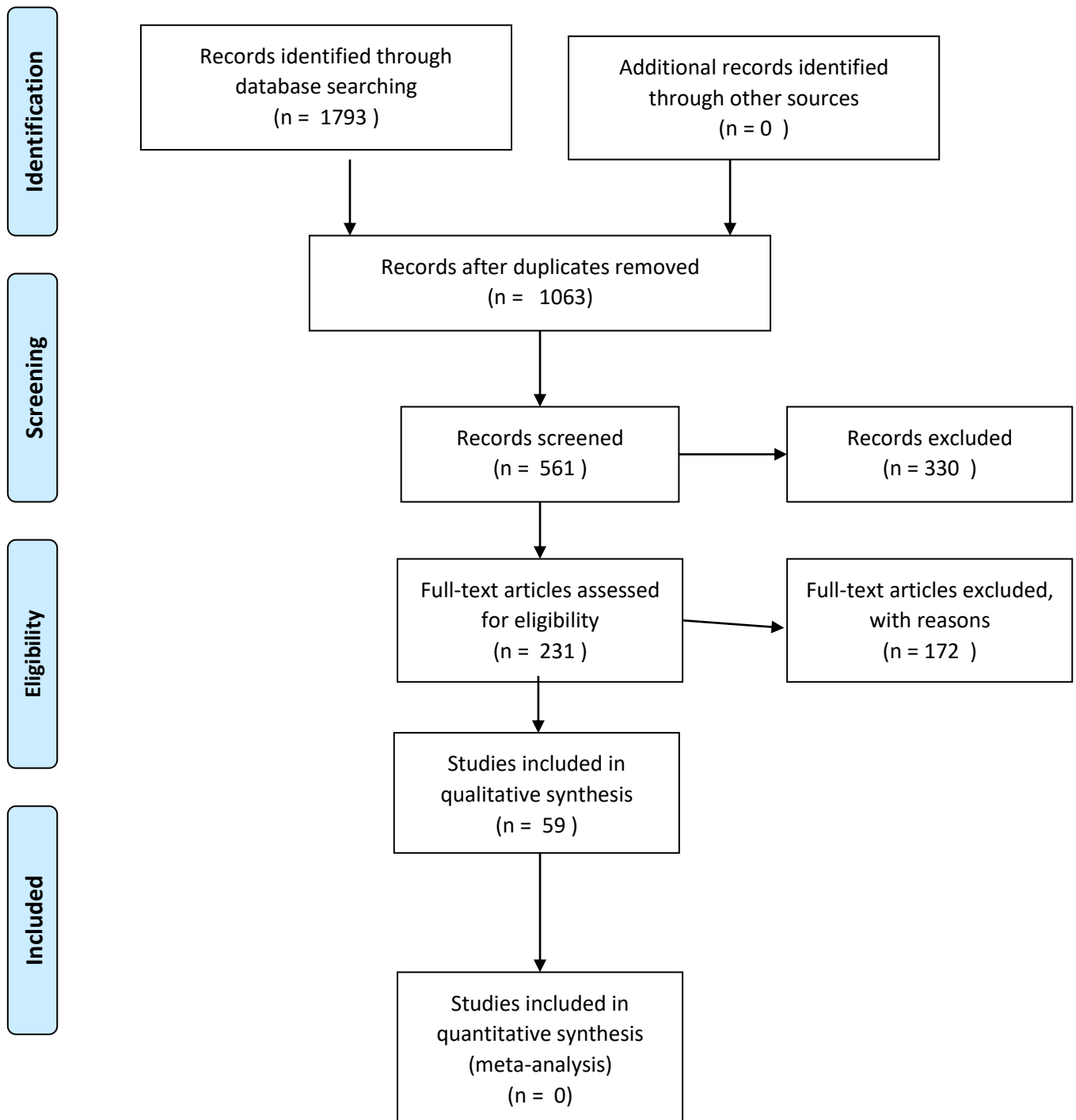
Table 4: Number of results per database – initial research

R Q	WOS	Scopus	ACM
RQ1.1	516	392	42
RQ1.2	175	183	0
RQ1.3	65	100	125
RQ2	76	81	38

Aplicando todos os critérios de exclusão, o número total de artigos foi de 1793. Removendo os artigos duplicados entre as bases de dados e entre as RQ, o número de artigos total foi de 1063 artigos.

A partir do volume inicial de artigos, os títulos e palavras-chaves dos artigos foram verificados para excluir aqueles cujo assunto indicava claramente que estavam fora do escopo da revisão pretendida. Essa etapa reduziu a amostra total de artigos para 526. Na etapa seguinte os resumos de todos os artigos foram avaliados. Atribuiu-se um grau de pertinência para os artigos de acordo com a pertinência percebida para cada questão de pesquisa. Os graus de relevância variaram de 1 a 5 (1 indicando relevância limitada e 5, relevância forte). Após a avaliação dos resumos, 231 artigos que foram classificados nos graus de pertinência 4 e 5 foram selecionados para a próxima etapa. A próxima etapa consistiu na análise das seções de introdução e conclusão dos artigos. Nesta fase os artigos foram selecionados de acordo com a relevância de cada artigo para os temas de interesse. Ao final, dos 231 artigos, 59 foram selecionados para leitura e análise na íntegra. O diagrama a seguir apresenta uma representação visual do diagrama do PRISMA (Page et al., 2021).

PRISMA Flow Diagram



Referências Bibliográficas (somente desta atividade)

- Carr, M. H., Neigel, J. E., Estes, J. A., Andelman, S., Warner, R. R., & Largier, J. L. (2003). COMPARING MARINE AND TERRESTRIAL ECOSYSTEMS: IMPLICATIONS FOR THE DESIGN OF COASTAL MARINE RESERVES. *Ecological Applications*, 13(sp1), 90–107. [https://doi.org/10.1890/1051-0761\(2003\)013\[0090:CMATEI\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2003)013[0090:CMATEI]2.0.CO;2)
- Kitchenham, B., & Chartes, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*.
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Scopus: a systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160–1177. <https://doi.org/10.1016/J.JOI.2018.09.002>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews*, 10(1), 1–11. <https://doi.org/10.1186/S13643-021-01626-4/FIGURES/1>
- Rekik, R., Kallel, I., Casillas, J., & Alimi, A. M. (2018). Assessing web sites quality: A systematic literature review by text and association rules mining. *International Journal of Information Management*, 38(1), 201–216. <https://doi.org/10.1016/J.IJINFOMGT.2017.06.007>
- Schneider, S., Torkar, R., & Gorschek, T. (2013). Solutions in global software engineering: A systematic literature review. *International Journal of Information Management*, 33(1), 119–132. <https://doi.org/10.1016/J.IJINFOMGT.2012.06.002>
- Sizo, A., Lino, A., Reis, L. P., & Rocha, Á. (2019). An overview of assessing the quality of peer review reports of scientific articles. *International Journal of*

Information Management, 46, 286–293.
<https://doi.org/10.1016/J.IJINFOMGT.2018.07.002>

Šubelj, L., Bajec, M., Mileva Boshkoska, B., Kastrin, A., & Levnajić, Z. (2015). *Quantifying the Consistency of Scientific Databases*.
<https://doi.org/10.1371/journal.pone.0127390>

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii–xxiii.
<http://www.jstor.org/stable/4132319>