data

MDPI

Data Descriptor

Data on 33 Years of Erroneous Usage of Rainfall Erosivity Equations

Nejc Bezak [1], Klaudija Lebar [1], Yu-Chieh Huang [2] and Walter Chen [2,*]

[1] Faculty of Civil and Geodetic Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia
[2] Department of Civil Engineering, National Taipei University of Technology, Taipei 10608, Taiwan
* Correspondence: waltchen@ntut.edu.tw; Tel.: +886-2-2771-2171 (ext. 2628)

Abstract: This paper describes the data gathered for a paper published in *Earth-Science Reviews* (DOI: 10.1016/j.earscirev.2023.104339) to address the problem of studies using incorrect equations to calculate rainfall erosivity (R factor), which can lead to issues related to land degradation, soil productivity loss, and biodiversity loss. The aim was to locate articles containing the incorrect equations and create a relational database that could be used to perform an in-depth analysis of the errors. Because the search target is an equation, it is impossible to directly query any literature database for the articles that contain the incorrect R equations. Therefore, a manual search of multiple databases was conducted. Subsequently, the literature search was broadened to identify the origin of the misuse of the R equations, and SQL (Structured Query Language) queries were formulated to understand why the errors continued to persist for a minimum of 33 years. The resulting entity-relationship-based Microsoft Access database was determined to be a valuable tool for performing in-depth analysis. It can be used to add incorrect studies and perform further analysis. It is suggested that further research should be conducted to determine the extent of the impact of these errors on soil erosion, ecosystems, and the environment.

Dataset: 10.17632/ryn6wdgckf.1.

Dataset License: CC BY 4.0

Keywords: relational database; empirical R equation; rainfall erosivity; land degradation; climate change; environmental education; soil erosion

check for updates

Citation: Bezak, N.; Lebar, K.; Huang, Y.-C.; Chen, W. Data on 33 Years of Erroneous Usage of Rainfall Erosivity Equations. *Data* **2023**, *8*, 75. https://doi.org/10.3390/data8050075

Academic Editor: Cidalia C Fonte

Received: 3 March 2023
Revised: 13 April 2023
Accepted: 19 April 2023
Published: 24 April 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Summary

One of the most critical concerns facing our planet is the impact of climate change on the environment. As such, it is essential to incorporate environmental education into soil erosion modeling, such as the Universal Soil Loss Equation (USLE) family of models. The rainfall erosivity factor (R) is a crucial parameter studied in these models, but it requires high-frequency precipitation data for accurate determination. However, in data-sparse regions, researchers have relied on empirical equations for decades to estimate the R factor. Unfortunately, one of these empirical equations, based on the Modified Fournier Index (MFI), has been widely misused, and the scientific community has not noticed the errors. The database presented here hold information on studies conducted since 1989 that employed incorrect equations to calculate rainfall erosivity (R). Due to the nature of the search criteria, it was impossible to directly query any literature database to locate articles containing the incorrect R equations. To locate these articles, we spent an extensive amount of time manually searching through scholarly sources such as Scopus, Google Scholar, and other Internet platforms using a combination of keywords such as "soil erosion", "ecosystem service", "erosivity", "1.735", "1.753", "0.08188", "0.8188", and their variations. Initially, the literature search was limited to English-language papers and

later expanded to include Chinese publications to establish the origin of the R equation misuse. We collected 125 articles, and their details were carefully recorded in an entity-relationship-based Microsoft Access database. Using SQL (Structured Query Language) queries, we were able to analyze the occurrence of the errors and better understand why they persisted for over 33 years without detection. The relational database is an invaluable tool for conducting detailed analyses, and it can be easily updated to add new incorrect studies to the database. Based on the information from the collected papers, we published a paper in *Earth-Science Reviews* [1] titled "A systematic review of the incorrect use of an empirical equation for the estimation of the rainfall erosivity around the globe" (DOI: 10.1016/j.earscirev.2023.104339). We hope that by raising researchers' awareness of the incorrect use of the R empirical equations, we can prevent similar errors from occurring in the future.

2. Data Description

Aside from the *Earth-Science Reviews* article's supplementary data, which was an unstructured Word document (https://doi.org/10.1016/j.earscirev.2023.104339), an entity-relationship model was implemented to structure the data for the analysis of the incorrect R equations. This analysis led to the development of the current database. The entity-relationship diagram, also known as an E-R diagram, is depicted in Figure 1, which explains the way in which all of the data in the database are linked to one another using the appropriate relations. The relational database, which was normalized to the third normal form, was implemented using Microsoft Access. SQL was used to implement common queries. To enhance the user experience, we also developed a custom-made greeting screen for the database's users, as shown in Figure 2. On the greeting screen, there are four buttons that are linked to four different SQL queries. These buttons can be used to accomplish the following actions:

- Detect the authors with more than one paper in the database and sort by the number of papers.
- Detect the most frequently cited articles and sort by the number of citations.
- Determine the number of times that Wischmeier and Smith (1978) [2] has been cited.
- Determine all types of incorrect R equations (ID and images).

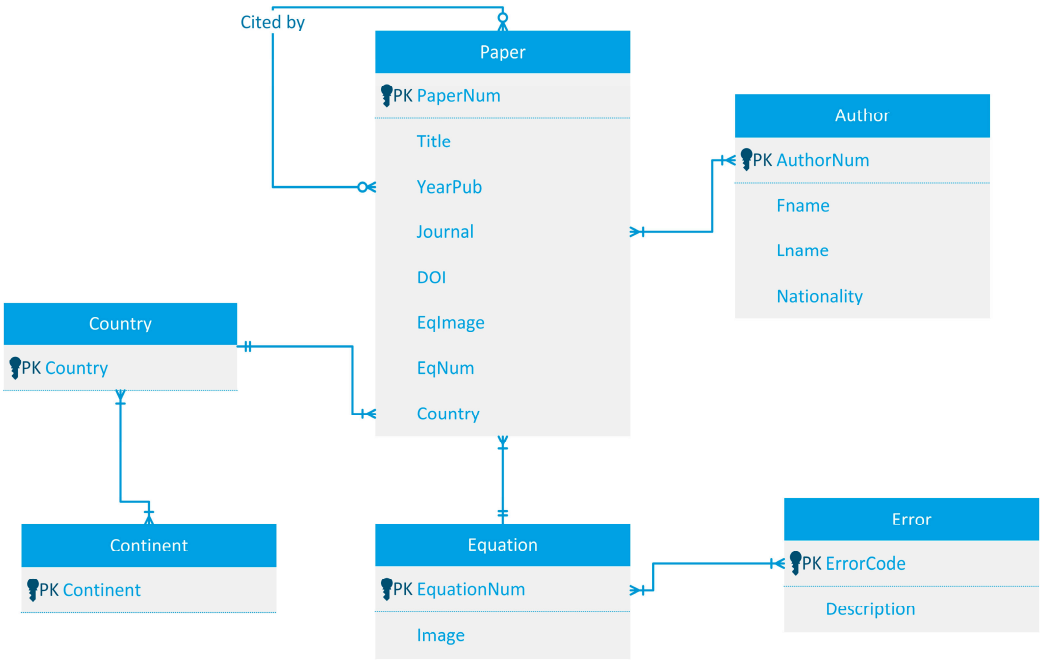

Figure 1. The entity-relationship diagram of this database.

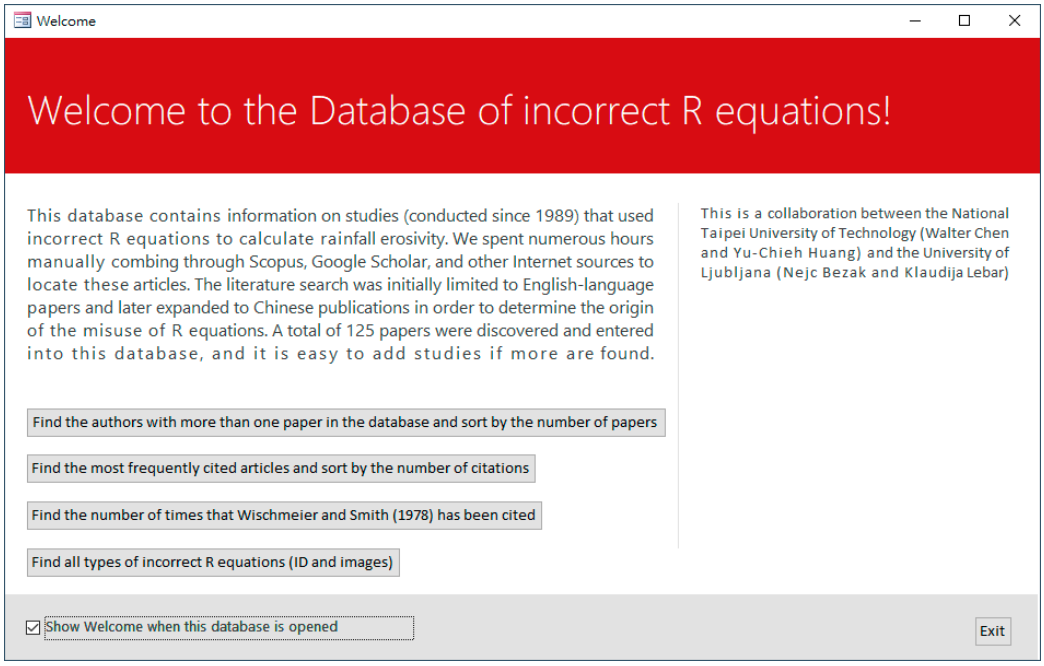

Figure 2. The welcome screen of this database.

3. Methods

The following procedures were utilized in order to collect the articles that had incorrect R equations:

- Scopus search terms such as "soil erosion," "ecosystem service," "erosivity," etc., were used to identify the probable list of papers, and then each of these papers was manually checked. At this stage, over 1300 articles were examined for possible inclusion.
- Search phrases such as "1.735," "1.753," "0.08188," and "0.8188" were used in Google Scholar, either singularly or in conjunction with one another. At this stage, more than one thousand articles were examined.
- A preliminary bibliometric examination was carried out on the papers that were identified in stages 1 and 2, totaling around sixty-five papers. At this point, some of the papers that had been published around 2010 and had received a significant number of citations were detected.
- A screening of any and all published works that mentioned any of the studies discovered in stage 3 was conducted. At this stage, we checked over five hundred different articles.
- A comprehensive inspection by hand of all the articles that were included in a review article on soil erosion published in 2021 was carried out. Approximately one hundred papers were reviewed at this stage.
- In addition, on the basis of the findings from stage 3, a selection of the early studies that were published between the years 1990 and 2000 was determined. Since those were not available online, we collected them with the assistance of the library at the National Taipei University of Technology. At this stage, about 20 papers were examined.
- The search for published material was initially restricted to articles written in English before being expanded later to include Chinese articles in order to locate the first study using an incorrectly cited R equation.
- In total, 125 papers were identified that had used incorrect R equations (along with three partially correct papers).

This database is valuable because it contains 125 studies spanning 33 years (from 1989) that have used incorrect R equations to calculate rainfall erosivity in soil erosion estimation. The fact that the error has been discovered for the first time has implications

for the accuracy of these 125 studies. Simple search commands cannot reconstruct this database's data. Only by manually examining publications in Scopus, Google Scholar, and other Internet sources and tracking citations was it possible to compile the data. The data were recorded in an easy-to-use Microsoft Access database to facilitate the addition of additional studies employing incorrect R equations.

4. User Notes

Other researchers may benefit from this database by skipping the laborious and time-consuming steps involved in obtaining these data and transforming them into a relational database.

It is possible to make use of the database to extend the currently conducted analysis of incorrect R equations [1]. Newly discovered articles from the past may be added to the database in order to conduct an in-depth analysis of the error pattern and provide insight into ways to prevent this kind of error from arising in future publications on soil erosion.

5. Patents

No patents have resulted from the work reported in this manuscript.

Author Contributions: N.B.: conceptualization, methodology, data curation, writing—review and editing; K.L.: data curation, writing—review and editing, W.C.: conceptualization, methodology, data curation, writing—original draft preparation; Y.-C.H.: data curation, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: N. Bezak and K. Lebar would like to acknowledge support from the Slovenian Research Agency through grants P2-0180 and J6-4628 and the Slovenian Ministry of Education, Science and Sport under UNESCO's Intergovernmental Hydrological Programme. Research activities were conducted within the scope of the UNESCO Chair on Water-related Disaster Risk Reduction. The Ministry of Science and Technology (Taiwan) Research Project (Grant Number MOST 111-2121-M-027-001) provided support for W. Chen and Y.-C. Huang.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The database has been deposited in Mendeley Data with the following DOI: 10.17632/ryn6wdgckf.1.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, W.; Huang, Y.C.; Lebar, K.; Bezak, N. A systematic review of the incorrect use of an empirical equation for the estimation of the rainfall erosivity around the globe. *Earth-Sci. Rev.* **2023**, *238*, 104339. [CrossRef]
2. Wischmeier, W.H.; Smith, D.D. *Predicting Rainfall Erosion Losses: A Guide to Conservation Planning (No. 537)*; Department of Agriculture, Science and Education Administration: Washington, DC, USA, 1978.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.