Site Planning for a Network of Weather Stations in the Dominican Republic Using Zonal Statistics from Geospatial Sources, Multi-Criteria Selection, and Neighborhood Analysis

Preprint

José-Ramón Martínez-Batlle
https://orcid.org/0000-0001-9924-0327
Facultad de Ciencias
Universidad Autónoma de Santo Domingo (UASD)
Santo Domingo, República Dominicana
joseramon@geografiafisica.org

Michela Izzo Gioiosa
https://orcid.org/0000-0003-4835-3967
Directora Ejecutiva
Guakia Ambiente
Santo Domingo, República Dominicana
michela.izzo@guakiambiente.org

December 24, 2024

Resumen

Abstract

Many weather station networks lack sufficient representativeness, and their station density is often inadequate to capture spatial and climatic variability effectively. Optimal site selection is therefore essential to enhance spatial coverage and improve data quality. This study proposes a methodology for identifying optimal sites for a meteorological station network in the Dominican Republic, utilizing a multi-criteria decision-making framework based on the Analytic Hierarchy Process (AHP) and neighborhood analysis. Using the H3 library as a spatial indexing tool, zonal statistics were derived from geospatial variables, including seasonality, habitat heterogeneity, proximity to water bodies, slope, solar radiation, and elevation. Expert-defined weights were assigned to each variable based on their relative importance. Areas with high topographic and climatic variability were prioritized to maximize spatial representativeness. Results highlight thermal and precipitation seasonality, elevation, and solar radiation as the most influential variables, emphasizing the need to collect data in elevated areas with marked seasonality. Sites were evenly distributed across three density scenarios, ensuring robust climatic and topographic coverage while avoiding redundancy through proximity constraints to existing stations. The proposed network would provide essential data for meteorological and climatic research in the region. Future studies should assess the accessibility and feasibility of the selected sites and incorporate additional environmental variables into the framework.

Keywords Weather stations networks · Optimal site selection · Spatial coverage · Multi-criteria decision-making · AHP

1 Introduction

Weather stations (WS) are essential for collecting accurate and up-to-date data on weather and climate in specific regions. The applications of the data collected by WS extend beyond meteorology and climatology, finding widespread use in fields such as engineering, agriculture, urban planning, and geography, among others (Chung et al., 2018; Marchi et al., 2019; Wilgen et al., 2016; World Meteorological Organization (WMO) & The International Association of Hydrological Sciences, 1976). The data provided by these stations are instrumental in predicting extreme weather events, such as tropical storms, hurricanes, tornadoes, and droughts, enabling communities to prepare and respond effectively. Furthermore, WS data underpin numerous scientific studies on climate and climate change, helping to better understand atmospheric dynamics and their impacts on the planet, ultimately contributing to more informed and effective planning strategies (World Meteorological Organization (WMO), 1996, 2017a, 2017b).

A robust WS network is crucial for making informed decisions across various domains and is fundamental for the well-being and safety of communities and the environment. Planning an adequate WS network is essential for effective land management. Previous studies, including those conducted in the Dominican Republic, reveal significant gaps in WS coverage in key areas and highlight the uneven spatial distribution and low density of existing networks, which likely affect the accuracy of collected data (Frei, 2003; Programa Mundial de Alimentos (PMA), 2019; Rojas Briceño et al., 2021; Theochari et al., 2021).

Many countries have evaluated the design of their WS networks, sometimes revisiting and improving them multiple times, often with successful implementations (Frei, 2003). Some have developed site selection protocols that align with general World Meteorological Organization (WMO) standards, adapting or extending them to meet the specific needs of their territories and intended applications (Rojas Briceño et al., 2021; Theochari et al., 2021).

The Dominican Republic is highly vulnerable to the impacts of climate change, and an insufficient WS network exacerbates this vulnerability (Izzo et al., 2010; Le, 2019; Lenderking et al., 2020; Lohmann, 2016; Mackay & Spencer, 2017; Roson, 2013). Improving and expanding the WS network requires investment in technology and infrastructure, as well as partnerships among government agencies, private entities, and research institutions (Programa Mundial de Alimentos (PMA), 2019). However, to optimize the use of limited resources, it is critical to design, evaluate, and select network alternatives using weighted criteria.

Research on the design of weather station networks consistently identifies multi-criteria decision-making (MCDM) methods as ideal for this purpose (Köksalan et al., 2011; Taherdoost & Madanchian, 2023; Thiriez & Zionts, 1975). These methods leverage geospatial data and include public input spatially integrated into decision-making using Geographic Information Systems (GIS) (Chakhar & Mousseau, 2008; Eastman et al., 1998; Malczewski, 2004; Rojas Briceño et al., 2021; Tekleyohannes et al., 2021; Theochari et al., 2021). Studies have demonstrated the effectiveness of traditional geostatistical techniques (Ali & Othman, 2018; Valipour et al., 2019), contemporary deep learning algorithms in combination with traditional methods (Safavi et al., 2021), and entropy-based approaches (Bertini et al., 2021). Combining geospatial data (e.g., GIS and remote sensing) with multi-criteria analysis (MCA) that assigns relative weights to geographical criteria is particularly efficient for analyzing diverse variables (Rojas Briceño et al., 2021).

The Analytic Hierarchy Process (AHP), a well-established multi-criteria decision-making (MCDM) method, is widely used due to its simplicity, its ability to provide insights into the analyzed attributes, and its structured framework for incorporating expert input (Rojas Briceño et al., 2021). Developed by Thomas Saaty in the 1970s (Saaty, 1977) and refined in subsequent decades (Saaty, 2001; Saaty & Tran, 2007), AHP is used to make decisions involving multiple criteria and alternatives. Traditionally applied in engineering, social sciences, economics, and business, AHP has recently been utilized effectively for selecting optimal WS sites in Peru (Rojas Briceño et al., 2021).

AHP involves breaking down a complex problem into a hierarchical structure of criteria and subcriteria, followed by pairwise comparisons to assign relative importance. The process includes identifying objectives and criteria, structuring them hierarchically, conducting pairwise comparisons, calculating priority values for criteria, and ranking alternatives based on aggregated priorities.

In this study, we integrate the Analytical Hierarchy Process (AHP) with geospatial and expert-driven data to systematically identify optimal sites for meteorological and climatic stations in the Dominican Republic. We prioritize key environmental and accessibility criteria to maximize spatial and resource efficiency while minimizing redundancy in existing networks. Additionally, we propose actionable scenarios for network expansion that align with international standards, offering solutions to address data gaps in poorly covered regions. Through this research, we advance geospatial methodologies and decision-support frameworks for meteorological infrastructure planning, with potential applications in broader climatological and environmental sciences.

2 Materials and Methods

We applied a sequence of four interdependent steps to develop alternative designs for weather station (WS) networks, emphasizing the multi-criteria selection of sites prioritized for their deployment. First, we gathered data on the existing WS network through consultations (via forms and visits) with government agencies, including the Dominican National Meteorological Office (ONAMET, now the Dominican Institute of Meteorology, INDOMET) and the National Institute of Hydraulic Resources (INDRHI). These forms were created and managed using the Open Data Kit (ODK) platform (Get ODK Inc., 2024; Hartung et al., 2010). We also consulted private entities managing WS networks. These efforts resulted in consolidated information on station locations, operational status, and other relevant attributes. This step ensured that the analysis was grounded in an accurate and comprehensive understanding of the current state of the WS network.

Subsequently, we implemented an Analytic Hierarchy Process (AHP) to select the optimal option among different alternatives using selection criteria weighted by individuals with expertise in the problem (Saaty, 2013). The selected criteria were distance to access points, thermal seasonality, rainfall seasonality, habitat heterogeneity, distance to water bodies, slope, hours of direct sunshine, elevation. These eight criteria were chosen based on their relevance to the problem, supported by our expertise as well as previous studies and recommendations from the World Meteorological Organization (Rojas Briceño et al., 2021; World Meteorological Organization (WMO) & The International Association of Hydrological Sciences, 1976).

We explicitly requested expert consultations, asking respondents to complete questionnaires electronically. After collecting the responses, we organized and recoded the data, then evaluated their consistency. We used only consistent responses to establish the criteria weights, which were subsequently applied to the available geographic information sources, including approximately 13,000 hexagons containing the corresponding multi-criteria information distributed across the Dominican Republic (Martínez-Batlle, 2022). Afterwards, we created these hexagons using the H3 spatial index library from Uber and computed zonal statistics with Google Earth Engine (GEE) (Gorelick et al., 2017; Uber Technologies, Inc., 2024). Specifically, we employed the GEE Python API to process the data programmatically, using packages like `geemap` for map visualization and data handling (Google Earth Engine Contributors, 2023; Wu, 2020). Finally, we assigned each hexagon an aggregated priority category, choosing from four possible options: marginally prioritized, moderately prioritized, prioritized and essential.

We designed the questionnaires, processed the responses, and weighted the criteria of geographic information sources using programming languages. These tasks were performed in the R statistical programming environment with the following packages: `ahpsurvey`, `sf`, `raster`, `terra`, `ggplot2`, `tidyverse`, `kableExtra`, `spdep`, `units`, `knitr`, and `rmarkdown` (Cho, 2019; Hijmans, 2023, 2024; Pebesma et al., 2016; Pebesma, 2018; Pebesma & Bivand, 2023; R Core Team, 2024; Wickham et al., 2019; Xie, 2014; Xie et al., 2020; Zhu, 2021). We also used Python to automate the design of questionnaires and their integration with Google Forms via its API.

Subsequently, we used the AHP results as input for a constraint-based exclusion process. In this step, we carefully analyzed the hexagons to identify those located in areas where accessibility was limited or where proximity to water bodies posed challenges. Hexagons situated near or within water bodies were deemed unsuitable for hosting meteorological stations and were excluded from further consideration. This process ensured that only feasible locations remained for the next steps of the analysis.

Finally, we conducted a neighborhood analysis between existing and proposed stations to ensure spatial homogeneity and avoid redundancy. First, we generated optimized station locations using a custom algorithm based on convex hulls and distance maximization. This method iteratively selected points that maximized their distance from previously chosen points, ensuring an optimal spatial distribution aligned with the station density criteria recommended by the World Meteorological Organization (World Meteorological Organization (WMO), 2020; World Meteorological Organization (WMO) & The International Association of Hydrological Sciences, 1976). Subsequently, we used continuous distance surfaces (e.g., rasters) as part of the neighborhood analysis to eliminate proposed stations that were too close to existing ones, thereby avoiding redundancy. Detailed methodological steps can be found in the Supplementary Information.

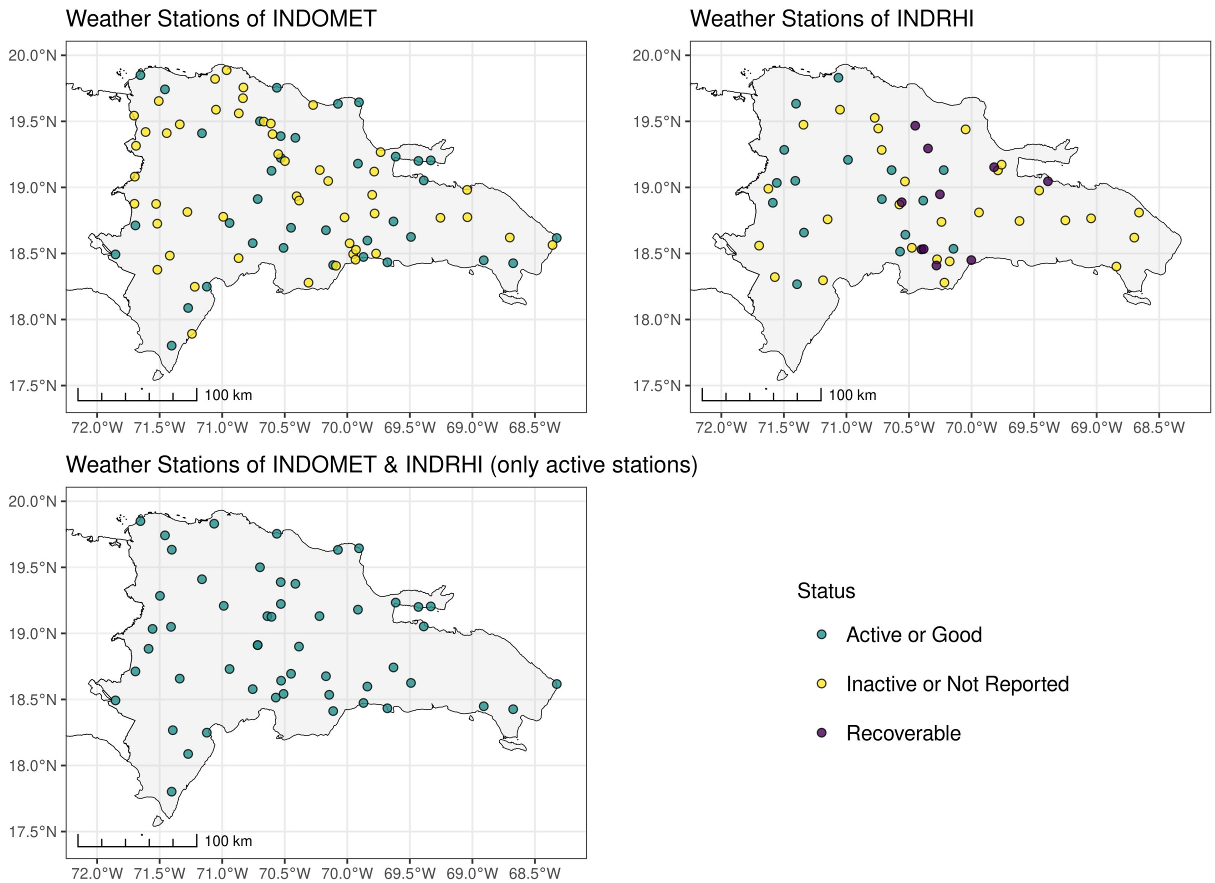

Figure 1: Weather station networks in the Dominican Republic for 2022, categorized by management entities: the Dominican Institute of Meteorology (INDOMET) and the National Institute of Hydraulic Resources (INDRHI). The maps illustrate the spatial distribution of stations by operational status, including "Active or Good," "Recoverable," and "Inactive or Not Reported." A combined map showcases only active stations from both entities to emphasize their geographic coverage

3 Results

(Table 1 and Figure 1)

Table 1: Summary of weather station status by owner (INDOMET and INDRHI) in the Dominican Republic for 2022, including the number of active or good, inactive or not reported, and recoverable stations, along with their total counts

Owner	Active or Good	Inactive or Not Reported	Recoverable	Total
INDOMET	36	51	0	87
INDRHI	16	28	10	54
Total	52	79	10	141

4 Discussion

Conflict of Interest Declaration

The authors declare that they have no conflict of interest related to the content of this article.

5 Author Contributions

JM and MI conceptualized and designed the study. JM was responsible for data collection. JM and MI established the methodology and conducted the research. JM developed the software, and supervision was carried out by MI. Both validated the work. JM and MI were in charge of visualization and drafted the original manuscript.

Data, Scripts, and Code Availability

The data supporting the findings of this study are openly available on Zenodo at []. The scripts used for data curation, analysis, and visualization are available in this section, as well as in the GitHub repository at and on Zenodo at [].

Supplementary Information

References

Ali, M. Z. M., & Othman, F. (2018). Raingauge network optimization in a tropical urban area by coupling cross-validation with the geostatistical technique. *Hydrological Sciences Journal*, *63*(3), 474–491. https://doi.org/10.1080/02626667.2018.1437271

Bertini, C., Ridolfi, E., de Padua, L. H. R., Russo, F., Napolitano, F., & Alfonso, L. (2021). An entropy-based approach for the optimization of rain gauge network using satellite and ground-based data. *Hydrology Research*, *52*(3), 620–635. https://doi.org/10.2166/nh.2021.113

Chakhar, S., & Mousseau, V. (2008). Spatial multicriteria decision making. *Encyclopedia of GIS*, *10*, 978–970.

Cho, F. (2019). *Ahpsurvey: Analytic hierarchy process for survey data*. https://CRAN.R-project.org/package=ahpsurvey

Chung, W., Abdel-Aty, M. A., & Lee, J. (2018). Spatial analysis of the effective coverage of land-based weather stations for traffic crashes. *Applied Geography*, *90*, 17–27.

Eastman, J. R., Jiang, H., & Toledano, J. (1998). Multi-criteria and multi-objective decision making for land allocation using GIS. *Multicriteria Analysis for Land-Use Management*, 227–251.

Frei, T. (2003). Designing meteorological networks for switzerland according to user requirements. *Meteorological Applications*, *10*(4), 313–317.

Get ODK Inc. (2024). *Get ODK*. https://getodk.org/.

Google Earth Engine Contributors. (2023). *Earth engine API*. https://github.com/google/earthengine-api

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/https://doi.org/10.1016/j.rse.2017.06.031

Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W., & Borriello, G. (2010). *Open data kit: Tools to build information services for developing regions*. https://doi.org/10.1145/2369220.2369236

Hijmans, R. J. (2023). *Raster: Geographic data analysis and modeling*. https://CRAN.R-project.org/package=raster

Hijmans, R. J. (2024). *Terra: Spatial data analysis*. https://CRAN.R-project.org/package=terra

Izzo, M., Rosskopf, C. M., Aucelli, P. P. C., Maratea, A., Méndez, R. E., Pérez, C., & Segura, H. (2010). A new climatic map of the dominican republic based on the thornthwaite classification. *Physical Geography*, *31*, 455–472.

Köksalan, M. M., Wallenius, J., & Zionts, S. (2011). *Multiple criteria decision making: From early history to the 21st century*. World Scientific. https://books.google.com.do/books?id=LqAw1539l/_cC

Le, T. D. N. (2019). Climate change adaptation in coastal cities of developing countries: Characterizing types of vulnerability and adaptation options. *Mitigation and Adaptation Strategies for Global Change*, *25*, 739–761.

Lenderking, H. L., Robinson, S., & Carlson, G. R. (2020). Climate change and food security in caribbean small island developing states: Challenges and strategies. *International Journal of Sustainable Development & World Ecology*, *28*, 238–245.

Lohmann, H. (2016). Comparing vulnerability and adaptive capacity to climate change in individuals of coastal dominican republic. *Ocean & Coastal Management*, *132*, 111–119.

Mackay, E. A., & Spencer, A. J. (2017). The future of caribbean tourism: Competition and climate change implications. *Worldwide Hospitality and Tourism Themes*, *9*, 44–59.

Malczewski, J. (2004). GIS-based land-use suitability analysis: A critical overview. *Progress in Planning*, *62*(1), 3–65.

Marchi, M., Sinjur, I., Bozzano, M., & Westergren, M. (2019). Evaluating WorldClim version 1 (1961–1990) as the baseline for sustainable use of forest and environmental resources in a changing climate. *Sustainability*.

Martínez-Batlle, J. R. (2022). *Estadística zonal multipropósito sobre información geoespacial de República Dominicana, usando Google Earth Engine, Python y R* (Version v0.0.0.9000) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.7367256

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446. https://doi.org/10.32614/RJ-2018-009

Pebesma, E., & Bivand, R. (2023). *Spatial Data Science: With applications in R*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429459016

Pebesma, E., Mailund, T., & Hiebert, J. (2016). Measurement units in R. *R Journal*, *8*(2), 486–494. https://doi.org/10.32614/RJ-2016-061

Programa Mundial de Alimentos (PMA). (2019). *Proyecto resiliencia a la sequía. Fortalecimiento de capacidades para mejorar la seguridad alimentaria y la resiliencia ante sequía en haití y la república dominicana. Proyecto de preparación ante emergencias basado en pronósticos de riesgos climáticos en república dominicana (FBF).* Programa Mundial de Alimentos (PMA).

R Core Team. (2024). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. `https://www.R-project.org/`

Rojas Briceño, N. B., Salas López, R., Silva López, J. O., Oliva-Cruz, M., Gómez Fernández, D., Terrones Murga, R. E., Iliquín Trigoso, D., Barrena Gurbillón, M., & Barboza, E. (2021). Site Selection for a Network of Weather Stations Using AHP and Near Analysis in a GIS Environment in Amazonas, NW Peru. *Climate*, *9*(12), 169. `https://doi.org/10.3390/cli9120169`

Roson, R. (2013). A modeling framework to assess the economic impact of climate change in the caribbean. *Cepal Review*, *111*, 23–36.

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, *15*(3), 234–281. `https://doi.org/10.1016/0022-2496(77)90033-5`

Saaty, T. L. (2001). *Fundamentals of the Analytic Hierarchy Process* (D. L. Schmoldt, J. Kangas, G. A. Mendoza, & M. Pesonen, Eds.; Vol. 3, pp. 15–35). Springer Netherlands. `http://link.springer.com/10.1007/978-94-015-9799-9_2`

Saaty, T. L. (2013). The Modern Science of Multicriteria Decision Making and Its Practical Applications: The AHP/ANP Approach. *Operations Research*, *61*(5), 1101–1118. `https://doi.org/10.1287/opre.2013.1197`

Saaty, T. L., & Tran, L. T. (2007). On the invalidity of fuzzifying numerical judgments in the Analytic Hierarchy Process. *Mathematical and Computer Modelling*, *46*(7-8), 962–975. `https://doi.org/10.1016/j.mcm.2007.03.022`

Safavi, M., Siuki, A. K., & Hashemi, S. R. (2021). New optimization methods for designing rain stations network using new neural network, election, and whale optimization algorithms by combining the Kriging method. *Environmental Monitoring and Assessment*, *193*(1), 4. `https://doi.org/10.1007/s10661-020-08726-z`

Taherdoost, H., & Madanchian, M. (2023). Multi-criteria decision making (MCDM) methods and concepts. *Encyclopedia*, *3*(1), 77–87. `https://doi.org/10.3390/encyclopedia3010006`

Tekleyohannes, M., Grum, B., Abebe, N., & Abebe, B. A. (2021). Optimization of rain gauge network using multi-criteria decision analysis and entropy approaches: case of Tekeze River basin, northwestern Ethiopia. *Theoretical and Applied Climatology*, *145*(1-2), 159–174. `https://doi.org/10.1007/s00704-021-03604-1`

Theochari, A.-P., Feloni, E., Bournas, A., & Baltas, E. (2021). Hydrometeorological-hydrometric station network design using multicriteria decision analysis and GIS techniques. *Environmental Processes*, *8*, 1099–1119.

Thiriez, H., & Zionts, S. (1975). *Multiple criteria decision making: Proceedings of a conference jouy-en-josas, france may 21–23, 1975* (Vol. 130). Springer Science & Business Media.

Uber Technologies, Inc. (2024). *H3. Hierarchical Hexagonal Geospatial Indexing System.* Available online: `https://h3geo.org/docs` (accessed December, 2024).

Valipour, E., Ghorbani, M. A., & Asadi, E. (2019). Evaluation and Optimization of Rain Gauge Network Based on the Geostatistic Methods and Firefly Algorithm. (Case study: Eastern Basin of Urmia Lake). *Irrigation Sciences and Engineering*, *42*(4), 153–166. `https://doi.org/10.22055/jise.2018.20549.1477`

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. `https://doi.org/10.21105/joss.01686`

Wilgen, N. J. van, Goodall, V. L., Holness, S. D., Chown, S. L., & McGeoch, M. A. (2016). Rising temperatures and changing rainfall patterns in south africa's national parks. *International Journal of Climatology*, *36*.

World Meteorological Organization (WMO). (1996). *Guía de instrumentos y métodos de observación meteorológicos.* Secretaría de la Organización Meteorológica Mundial.

World Meteorological Organization (WMO). (2017a). *Guía de instrumentos y métodos de observación meteorológicos.* World Meteorological Organization Geneva, Switzerland.

World Meteorological Organization (WMO). (2017b). *Guide to the WMO integrated global observing system. WMO-no. 1165.* World Meteorological Organization, Geneva, Switzerland.

World Meteorological Organization (WMO). (2020). *Guide to hydrological practices. Volume i: Hydrology—from measurement to hydrological information. WMO report no. 168* (p. 296). World Meteorological Organization, Geneva, Switzerland.

World Meteorological Organization (WMO), & The International Association of Hydrological Sciences. (1976). *Hydrological network design and information transfer*. Secretariat of the World Meteorological Organization.

Wu, Q. (2020). Geemap: A Python package for interactive mapping with Google Earth Engine. *Journal of Open Source Software*, *5*(51), 2305. https://doi.org/10.21105/joss.02305

Xie, Y. (2014). Knitr: A comprehensive tool for reproducible research in R. In V. Stodden, F. Leisch, & R. D. Peng (Eds.), *Implementing reproducible computational research*. Chapman; Hall/CRC.

Xie, Y., Dervieux, C., & Riederer, E. (2020). *R markdown cookbook*. Chapman; Hall/CRC. https://bookdown.org/yihui/rmarkdown-cookbook

Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and pipe syntax*. https://CRAN.R-project.org/package=kableExtra