[PD-02-025] Verification & Validation of Applied GIS Research

Abstract

Researchers and practitioners apply the concepts and tools of GIScience to answer questions about geographic phenomena and to inform policy and management decisions across a wide range of social and environmental domains. Verification and validation of applied GIS research is essential to the development and application of credible geographic knowledge. Attempting to verify and validate the claims researchers and practitioners make when they analyze phenomena using the concepts and tools of GIScience is essential to the development of geographic knowledge. Verification is the act of testing whether the concepts and methods used to make a research claim were implemented in a way that is appropriate for the question being investigated. Validation is the act of assessing whether the concepts, measurements, or conclusions of a study are logically sound and factually well-founded. Researchers can pursue the verification and validation of past studies by attempting to reproduce or replicate these earlier findings. During a reproduction, an independent researcher attempts to recreate the results of an initial study using the data and procedures of that study. During a replication, an independent researcher empirically tests the validity of the claims made in a study by selectively altering different aspects of the initial work when repeating the study. This entry outlines these processes and how they are used to verify and validate applied GIS research.

Keywords: replicability, reproducibility, validation

Author & citation

Kedron, P. and Holler, J. (2024). Validation and Verification of Applied GIS Research. The Geographic Information Science & Technology Body of Knowledge (2024 Edition), John P. Wilson (ed.). DOI: 10.22224/gistbok/2024.1.13.

Explanation

1. Introduction
2. Validity and Geographic Uncertainty
3. Verification of Geographic Research
4. Conclusion

1. Introduction

It is now common for researchers and practitioners alike to use the concepts and tools of GIScience to investigate social and environmental phenomena as spatial patterns and processes, and to plan interventions to change those phenomena. As researchers and practitioners study the world, they make claims about how processes function and the effects interventions will have on those processes. Assessing the credibility and reliability of those claims is essential to the development of geographic knowledge and decision-making systems. If future work or actions are based on an improper understanding of the world, then decisions and interventions may be inefficient, ineffective, or even potentially harmful. As a consequence, researchers and practitioners verify and validate their own work and the

work of others.

This process of continuous assessment and revision is a key characteristic of scientific investigations of geographic phenomena. As researchers produce studies, their results and interpretive claims become evidence for arguments that structure future investigations.

In this context, verification and validation can be defined as,

Verification: is the act of testing whether the concepts and methods used to make a research claim were implemented in a way that is suitable for the question being investigated. Verification can support the credibility, accuracy, or validity of a claim.

Validation: is the act of assessing whether a concept, measurement, or conclusion is logically sound and/or factually well-founded. The validity of research claims depends on both the quality of measurements, the execution of analyses, and the interpretation of results.

A common way to attempt to verify or validate a GIS study is by attempting to reproduce or replicate the results of that study. In a reproduction, researchers attempt to recreate the results of a prior study using the same data and procedures. During this process, the researchers verify the original work by evaluating whether concepts and methods were used appropriately and implemented properly. In a replication, researchers attempt to empirically test the validity of the claims made in a prior study by collecting new data and selectively altering different aspects of the initial work. Changing the data and procedures used in a study creates the opportunity to directly assess whether those elements affect results and the claims they are designed to support.

Recently, concern has arisen in the geographic (Konkol et al. 2019, Kedron et al. 2021, Goodchild and Li 2021) and general scientific literature (Baker 2016, Fanelli 2018, NASEM 2019) that researchers do not, or cannot, reproduce or replicate past research. Surveys of geographic researchers suggest that only a minority of researchers have attempted to replicate the work of their peers (2024a, 2024b), and that most studies are missing the components and information needed to even attempt a reproduction or replication (Ostermann and Granell 2017, Konkol et al. 2019). Extending this work, a growing literature on reproducibility and replicability highlights the important and complicated role of verification and validation in GIScience (Kedron and Holler 2022, Peng and Hicks 2021, Kedron et al. 2024a, 2024b).

Verifying and validating GIS studies is not simply a function of proper measurement, or the correct execution of an analysis. Rather, validity is tied to three elements: the current understanding of the phenomena, the location being studied, and the purpose of a study. Spatial data that is suitable and valid for one purpose in one location may not be suitable or valid for another purpose or another location. Moreover, as new evidence emerges and as theoretical explanations develop, the validity of past studies can be revised.

The primary focus of this entry is the verification and validation of the claims of researchers conducting studies of geographic phenomena using the concepts and tools of GIScience. From its inception, the discipline of GIScience has considered how well geographic phenomena are digitally represented and communicated, and how these representations contribute to uncertainty in geographic analyses.

Foundational elements of this literature are already present in the GIS&T BoK. The verification and validation of geospatial data and representational systems is discussed in BoK entries on Conceptual Models of Error and Uncertainty (Couclelis 2020), Error-based Uncertainty (Wechsler 2021), Spatial Data Uncertainty (Li 2017), and Representing Uncertainty (Kinkeldey and Senaratne 2018). These entries draw from a large and robust literature on sources of uncertainty in geographic data and means of validating measurements of geographic phenomena (see Fisher 1999, Couclelis 2003, Longley et al. 2015). Practical considerations are covered in entries on Ontology and Semantic Interoperability (Zhang 2019) and Usability Engineering and Evaluation (Ooms and Skarlatidou 2018). The simple feature access standard of the Open Geospatial Consortium provides a model and validation criteria for geometric representations of geographic features in a GIS. Relatedly, the ISO/TC 211 standards describe procedures for evaluating the quality of spatial data (ISO 19157-1) and for calibrating and validating remote sensing data and derivative products (ISO/TS 19124-1:2023).

Verification and validation also take on specialized meaning when examining GIS as software. In reference to software testing, verification is concerned with determining the quality of the software and whether it implements functions as expected. Successful verification typically reduces the chances of product failure. In contrast, GIS software validation is the process of determining whether the functions of the software meet its intended use. Software validation occurs after verification and is commonly used to detect errors in the software that were not anticipated or identified during verification.

In the following section, we outline four different types of validity and relate each to a conceptual model of uncertainty presented by Longley et al. (2015). This model is adopted because it synthesizes the wider literature on geographic uncertainty and because it is part of a commonly adopted GIScience textbook. In the final section, we examine verification and validation in the context of a theory-centered approach to inquiry that is facilitated by the spatial data management and computing capacity of GIS. In this section, we focus on reproduction and replication as means of verifying and validating prior empirical results and claims.

2. Validity and Geographic Uncertainty

In most GIS studies, validity is an inductive claim made based on empirical evidence, which cannot be definitively proven to be true or false. Validity can be assessed based on the strength or weakness of different forms of evidence presented by researchers. Evidence supporting four forms of validity are particularly important for establishing the credibility and reliability of GIS research.

Construct Validity relates to how well measured variables represent a concept that is not directly measured or observed. In applied GIS studies, do the spatial data inputs and constructed or modelled outputs truly represent the concepts in the research question? For example, a researcher interested in analyzing spatial variations in social vulnerability to environmental hazards could use an index composed of several variables (see Cutter et al. 2012) to represent this unobserved concept. The degree to which the index reflects the social vulnerability of people defines the construct validity of the index. In many cases, a researcher conducting a GIS study seeks construct validity by using definitions and measurement procedures that are already well-established in the existing literature. However, it may be the case that the same definitions and measurement procedures

cannot be used in all geographic contexts. Returning to the social vulnerability example, it is unlikely that the same set of social characteristics will contribute to vulnerability in the same way in different geographic or hazard settings. Construct validity may also change over time as the theory used to specify the meaning of measurements evolves, altering the relationship between theoretical concepts, empirical measurements, and study design. For example, geographic wildlife observations and habitat suitability models may lose construct validity once distinct species or sub-species are recognized. The Gunnison Sage Grouse was recognized as distinct from the Greater Sage Grouse and registered as “threatened” in 2014 (ECOS 2024).

Conclusion Validity is the extent to which the conclusions and claims in a GIS study are founded on an adequate analysis of the available data. Considering the research question and the characteristics of the input data, has the researcher generated valid results by choosing appropriate analytical methods and conforming to modelling assumptions? In many GIS studies, conclusion validity is closely associated with the assessment of the geographic models used in an analysis. The objective is to evaluate whether the analysis conducted provides adequate support for the conclusions drawn. Model validation typically involves scrutinizing the assumptions of the model in light of the input data characteristics, what is already known about the process under investigation, and if possible, how well the model fits or predicts outcome data. For example, problems of conclusion validity in GIScience could include bilinear interpolation of categorical data or modelling hydrological flows based on a digital surface model in a forested terrain. Valid models that meet assumptions and closely align with existing data support conclusion validity.

Internal Validity is the degree to which a study can identify true causal effects through unflawed study design and implementation of GIS analyses. In a narrow sense, have data collection and analysis methods adequately observed and isolated the geographic phenomena of interest, or could systematic bias or other factors also explain the results? Internal validity helps establish the soundness of the results of a study and whether those results can be trusted to answer the research question under investigation. A researcher can pursue internal validity by working to remove systematic errors (e.g., data processing errors) and biases (e.g., selection bias) from the design and implementation of their study. For example, a study designed to estimate the size of an animal population in a region will have less internal validity if it samples locations based on convenience and accessibility than if it samples stratified random locations without excluding remote and inaccessible places. Similarly, studies reliant on secondary data from volunteered geographic information like geocoded social media content or GPS tracks of cycling routes may suffer from internal validity problems rooted in data with uncontrolled systematic bias. However, there is no statistical test of internal validity. The internal validity of applied GIS research is a judgement based on study design.

External Validity is the degree to which the findings of a study can be generalized, or found to be true, in other contexts. Can the study design be applied to other regions or populations and yield similar results? External validity is nearly always a concern because GIS studies are conducted using data sampled from a region or population. If proper sampling procedures are followed (e.g., random sampling, sufficient sample size) and the procedures of the study support internal validity, the sample should be representative of the population, and the results may be generalized to the portions of that region or population that were unobserved. However, this information alone does not support claims of external validity.

Assessing the external validity of GIS studies can be particularly challenging because data are almost always drawn from a particular geographic region and period of time. While an internally valid GIS study can make reliable claims about the phenomena being studied in that location at that time, that evidence is insufficient to transfer those claims to new locations or new periods of time. Debate about the external validity of geographic studies and geographers' capacity to identify regular associations across locations animated the discipline's nomothetic-ideographic debate and motivated the development of popular local modelling frameworks like Local Indicators of Spatial Association and Geographically Weighted Regression.

Difficulties with external validity arise because many processes interact to produce the spatial data and patterns in complex regions, and because it is often unclear if and how processes may change over heterogenous regions and times. For example, a theorized relationship between minority language communities and social vulnerability may hold true in some regions, but may be counter-indicated in others where social capital within the minority language group contributes to resilience. Moreover, regions are open systems characterized by fuzzy boundaries, which makes it difficult to know how to best separate one region from another or represent those divisions in a GIS. Social vulnerability studies of the United States typically define regions as counties or as census tracts, neither of which correspond directly to minority-language enclaves in an urban region.

Each of these four types of validity can be linked to Longley et al.'s (2015) presentation of uncertainty in GIS studies (Figure 1). Longley et al. identify four distinct points in the research process where uncertainties emerge: 1) in the conception of geographic phenomena, 2) in the measurement and representation geographic phenomena, 3) in the analysis of geographic measurements, and 4) in the interpretation of analyses. Each of these filters distorts and transforms the representation of the real world as it is stored and analyzed in a GIS. As researchers make subsequent study design decisions with imperfect information and an incomplete understanding of the world, those distortion accumulate, which expands uncertainty about the results and claims. We add a fifth filter from Interpretation to Conception to illustrate this process of uncertainty cascading from one study to subsequent studies.

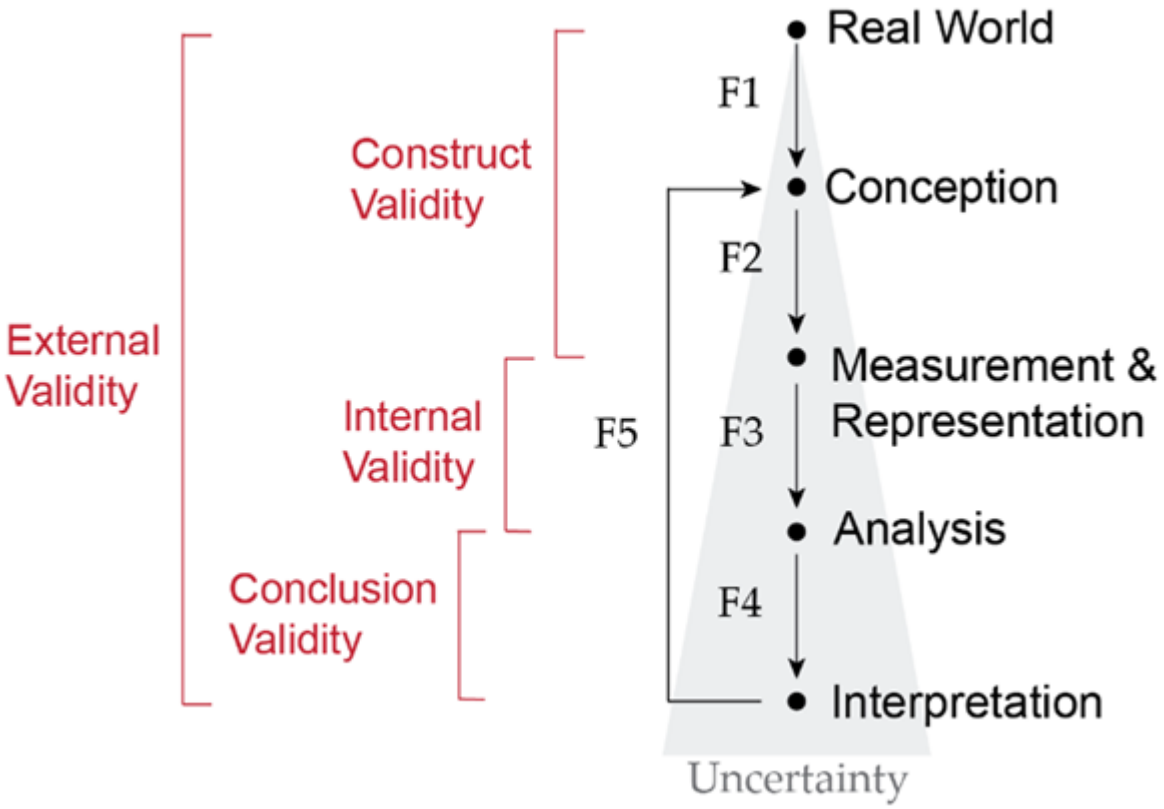


Figure 1. Conceptual model of uncertainty and validity. Different forms of uncertainty (F1-5) can filter and distort the study of the real world and expand uncertainty about results and claims. These filters impact different forms of validity. For example, construct validity is strongly impacted by uncertain conceptions and measurement of the real world. Source: authors.

Progressing sequentially along the research process, uncertainty in the conception and measurement of geographic phenomena (F1-2) may contribute to concerns about the construct validity of a GIS study. Researchers may attempt to address this uncertainty by measuring a construct in several different ways to assess how sensitive the representation is to different research design decisions. Extending the social vulnerability example above, researchers may attempt to increase construct validity by using several different vulnerability indices, by working with community members to adapt existing indices to local study context, or by validating the measurements with data on natural hazard outcomes.

Uncertainties in analysis (F3) impact conclusion and internal validity. Common threats to conclusion and internal validity are linked to the inherent fuzziness of many geographic entities, the large scale and complex nature of many geographic phenomena, and the well documented modifiable areal unit problem. For example, a researcher studying foraging patterns of American Robins might define foraging area polygons around nesting sites and then use an extract by mask operation to allocate resources stored in a set of raster files to those foraging areas. If that operation leads to the exclusion or misassignment of some raster cells to foraging areas because the operation uses raster centroids to allocate the resource, then the internal validity of the study may come into question. Similarly, if the rasters were developed to represent seasonal resources and the researcher inappropriately mixed together summer and winter rasters when modeling summer foraging behavior, then both the internal and conclusion validity of the study would be in question.

Threats to external validity also emerge across all four uncertainty filters. Claims of external validity rest on phenomena, regions, and populations being properly conceptualized within a study and across studies, and on conducting measurements and analyses in appropriate ways. Returning to the nesting example above, many of the study design and analysis decisions made when studying the American robin foraging may at first appear reliable across studies of different forested regions. However, robins are known to vary their foraging behavior with the distribution of food resources (Paszkowski 1982).

Therefore, a threat to external validity would arise if measurement decisions related to the extent of foraging areas and analytical decisions that define the parameterization of movement used in one study area were applied to another study area with different resource abundance. More broadly, analyses of the American robin, which nest in trees and typically forage locally for short periods, would make little sense when studying Laysan albatross, which nest on sandy beaches and forage across hundreds of miles of open ocean over the course of several days. Ultimately these issues raise questions about the proper interpretation of a study (F4) given the constructed of the evidence created during the course of a study.

A comprehensive graphical catalog of threats to each of the four types of validity in the context of epidemiological research is presented in Matthay and Glymour (2020). Further geographic treatments of these topics are included in several GIScience textbooks including Bolstad (2012), Montello and Sutton (2012), and Longley et al. (2015).

3. Verification of Theory-driven Geographic Research

Traditionally, studies of geographic phenomena that have used GIS to organize and analyze geospatial data have taken a theory-centered approach to inquiry. Within this approach, hypotheses are proposed based on the existing understanding of a phenomenon and then tested against empirical evidence with the aid of GIS tools. Hypotheses that are repeatedly supported by evidence produced during many independent studies gain in credibility and build support for the theories and explanations they are based on. Assessing the validity of individual studies as they are conducted is an important element of this process of evidence creation and evaluation. The validity of an individual study may be assessed by investigating how that study was conceived and executed, or by attempting to reproduce or replicate that work. As a counterpoint to what follows, Leszcynski (2017) provides a useful summary of epistemological critiques of this approach to GIScience, which includes a discussion of positivist and critical realist conceptions of the field that underlie the discussion that follows.

During a reproduction, an independent researcher attempts to recreate the results of an initial study using the data and procedures of that study. The epistemological goal of attempting to reproduce a GIS study is to verify that the results and claims of the initial study are supported by an adequate analysis of the data. What constitutes an adequate analysis can vary across researchers and studies, but generally consists of correctly processing spatial data and using methods that are logically capable of answering the research question. Reproductions therefore evaluate and offer the opportunity to test the conclusion validity of a study and may provide insight into other forms of validity. For example, Kedron et al.'s (2024a) reproduction of a county-level spatial hotspot analysis of COVID-19 identified concerns about the original authors implementation of spatial statistical procedures (conclusion validity), but also raised questions as to whether the

spatial weights matrix used in the study captured epidemiologically relevant interactions (construct validity).

In GIScience, attempts to formally reproduce prior studies often focus on the computational reproducibility of a study – whether the same results, figures, and maps can be recreated using the same data and procedures. For example, an ongoing effort by the Association of Geographic Information Laboratories in Europe annually attempts to reproduce the numerical results and figures of papers submitted to the association's annual conference. Similarly, much of the existing literature on reproducibility in the discipline focuses on the sharing of data and code, and providing sufficient documentation to reuse these research artifacts (see Konkol et al. 2019, Nust and Pebesma 2021, Tullis and Kar 2021). This privileging of the computational aspects of GIS analyses follows the computer science literature and standards presented by the ACM (2024) and NASEM (2019).

An emphasis on computational reproducibility is well founded because a researcher's ability to reproduce the computational aspects of a study is an important foundation for assessing or modifying prior work. However, recent work by Kedron et al. (2023b) argues for a wider view of reproducibility that recognizes that factors hindering the recreation of results and the evaluation of claims extend beyond computation into the conceptualization and design of geographic research. This approach acknowledges the importance of computation in GIS analysis, but aligns disciplinary practices surrounding reproducibility with those found across the physical and social sciences.

During a replication, a researcher empirically tests the validity of the claims made in a prior study by selectively altering different aspects of the initial work when repeating the study. A researcher attempting to replicate a GIS study will collect new data, but may also change the instruments and procedures used, the measurement of variables, the population being studied, and/or the location of the replication. For example, Kedron et al. (2022) used a replication to test the external validity of a regression model developed in New York City in the new context of Phoenix, AZ. For that replication, the authors collected new data, but attempted to match the disease environment and regression specification of the original study. This approach allowed the authors to evaluate the utility of the model form in a new context and the generalizability of key conclusions about associations between demographics and positive COVID-19 test rates.

The type of validity checks a specific replication provides depends on the combination of study characteristics a researcher changes (Sargent 1981, Schmidt 2009, Gomez et al. 2010, Radder 2003, 2012, Munafo et al. 2016, Plesser 2018), and may include changes to data, variable measurements, instruments and procedures, locations, and populations.

Changing Data: When a researcher attempting a replication keeps all aspects of a GIS study the same, but collects new data, the replication is designed to assess the internal validity of the initial study. The replication will control for sampling error and provide evidence as to whether the prior results were the product of chance variation. By recreating and closely examining the procedures and results of the initial GIS study, the independent researcher attempting the replication will work to determine whether the study was appropriately designed and executed and to identify systematic errors or alternative explanations of the original results and claims. Many researchers formally add this internal validity check to their own study designs with data sample partitioning and cross-validation.

Changing Variable Measurements: If the researcher changes how the variables used in a GIS study are measured when attempting to replicate that study, then the replication can act as a test of construct validity. For example, a researcher replicating a remote sensing analysis of forest health might shift from using the Normalized Difference Vegetation Index to the USGS Enhanced Vegetation Index or canopy structure measures derived from NASA's GEDI Mission. Observing consistent results and relationships across replications using alternative measures of forest health would typically raise researcher confidence that the operationalization of the construct used in the initial study did not affect the results or claims made. However, replications alone cannot provide definitive tests of construct validity because formation of a construct also depends on the theory that defines it.

Changing Instruments and Procedures: If a researcher changes the materials and procedures (e.g., data collection device, GIS analyses) used during the replication of a GIS study, then the replication can be used to establish if the initial results were the product of a particular device or procedure. This approach to replication can contribute to assessments of the conclusion validity of a GIS study. For example, while GIS software conform to many of the same OGC standards for fundamental features (see ESRI 2020, QGIS 2024), different software do use different algorithms for some analyses. As a result, changing GIS software alone may constitute a meaningful change in instrumentation. Moreover, the implementation of algorithms can change between versions of the same GIS software. In the majority of cases, GIS analysts do not provide extensive details about the full computational environment used in a study, which makes testing for the influence of instrumentation difficult in practice. Beyond data and code, the complete research compendium for a GIS study would include information about the software packages used and computational environment in which the analysis is executed (Nust et al. 2020, Konkol et al. 2020).

Emerging spatial research platforms such as CyberGISX or the Geospatial Analytics Extension for KNIME are in part developing infrastructure to automatically record and share this information to facilitate the reuse of analyses and replication.

Even when the GIS software and surrounding computational environment of a study are preserved in a replication, there remains a range of malleability in how those instruments are used in the procedures of the study. Many core functions (e.g., geometric and topological operations) are written directly into GIS software and cannot be adjusted without significant effort. Other components of a research procedure are easily accessible and are designed to be tailored to the context and data of any particular study. For example, spatial clustering tools included in most GIS software ask the user to specify key parameters such as the spatial weights matrix and multiple testing adjustments. As a researcher replicating a study adjusts these parameters, they are testing the sensitivity of the conclusion of the prior study to those changes.

Changing Locations: If a researcher replicating a GIS analysis collects new data from a new location, then the replication may be used to test the external validity of the claims of the initial study. Whether researchers should expect study results, and the theories that underlie them, to generalize across locations is at the heart of an ongoing discussion about nomothetic and ideographic approaches to geographic inquiry (Sui and Kedron 2021, Kedron and Holler 2022). On one hand, given geographers' acknowledgement of the uniqueness of places, we should expect challenges in replicating research in different locations. On the other hand, geographic researchers attempting to develop a generalizable

theory must demonstrate that the theory applies across a range of geographic contexts. Replication studies in different locations are therefore necessary to determine whether a geographic theory can be generalized and applied to other locations. If a theory can be generalized, replications are necessary to discover and describe the geographic contextual conditions in which the theory can be expected to apply.

As a practical matter, researchers conducting a GIS analysis often have limited or no ability to control factors that may affect the outcome of a study. This lack of control can make it difficult for the researcher conducting the replication to account for alternative explanations in their analysis. For these reasons, replication attempts in which locations are changed do not necessarily provide evidence that can be used to assess the claims of an initial GIS study. For example, if a replication of a GIS analysis of the effects of a conservation program on local biodiversity failed to find evidence of the biodiversity enhancing effects of the initial study, this does not mean that the effects identified by the initial study do not exist. It may simply be the case that other conditions needed for the effect to occur were not in place. If fencing restricts the movement of a newly introduced keystone species within a region, a researcher may not observe the biodiversity-enhancing effects introducing that species produced in another region without fences.

Changing Populations: As is the case with changing the location of a study, changing the population examined in a replication attempt tests the external validity of the claims of the initial study. Changing the location of the study usually also changes the population that the replication is sampling from, though this is not necessarily the case. A researcher may also attempt to replicate a GIS analysis in the same location, but from a different population. For example, replicating an analysis of healthcare accessibility and immigrant health in the same city could produce very different results, if the set of countries immigrants are arriving from has shifted between the times the two studies are taking place.

Planning, implementing, and evaluating replications of geographic research can be difficult for several reasons. Many of the processes studied using GIS can be expected to vary and interact differently across locations. This variability complicates the comparison of claims across locations. In practice, attempts to reproduce and replicate GIS studies often change several aspects of the initial study at the same time. For example, shifts in location may create unexpected changes in population, or require the use of new variables when the measures used in an initial study are not available. Instruments and procedures are often changed between studies because current publishing and reporting practices make it difficult to understand and identify how an initial study was conducted and even what GIS software and computational environments were used. Reproduction and replication studies are conducted after prior studies, by which time the geographic regions and phenomena being studied will have changed.

Geographers have little to no control over variables changing over space and time in the complex and open systems being studied. As a result, it is likely best to consider replications and reproductions along a continuum of change. Correspondingly, researchers should carefully read and consider what exactly was changed between studies and how the collective set of changes and results serve as evidence for the comparison of claims. Given the lack of reproducibility in published research, however, the greatest source of uncertainty in verifying and validating prior studies is a lack of knowledge about how prior studies were conducted.

4. Conclusion

The verification and validation of claims made by researchers using GIS to study the world is essential to the cumulative development and credible application of geographic knowledge. In this entry, we have outlined components of these processes and different types of validity essential to the reliability of research claims. We have presented a model of uncertainty filters in geographic research and conceptually mapped sources of uncertainty to types of validity. Reproduction and replication are discussed as avenues for verifying and validating different aspects of empirical GIScience research. While foundational to the research process, evidence of systematic verification and validation of past research is often missing in the published literature. To address this gap, there are needs to both 1) conduct and publish independent verifications and validations of prior research, and 2) conduct and publish original research using the concepts and tools of GIScience with greater reproducibility.

References

Association for Computing Machinery. (2024). Artifact Review and Badging Version 1.1 – Current. Association of Computing Machinery – Publications, Statements, Guidelines, and Procedures. Last accessed 4 April 2024.

Baker, M. (2016). Reproducibility crisis. Nature, 533(26), 353-66.

Bolstad, P. (2012). GIS Fundamentals (4th Edition). White Bear Lake, MN: Eider Press.

Couclelis, H. (2003). The Certainty of Uncertainty: GIS and the Limits of Geographic Knowledge. Transactions in GIS 7(2): 165-175.

Couclelis, H. (2020). Conceptual Models of Error and Uncertainty. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2021 Edition), John P. Wilson (Ed.).

Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2012). Social vulnerability to environmental hazards. In Cutter, S., (ed), Hazards Vulnerability and Environmental Justice (pp. 143-160). Routledge.

Environmental Conservation Online System (ECOS). (2024). Gunnison sage-grouse. US Fish and Wildlife.

Esri. (2020). Esri Support for Open Geospatial Standards. Esri White Paper. Last accessed 7 April 2024.

Fanelli, D. (2018). Is science really facing a reproducibility crisis, and do we need it to?. Proceedings of the National Academy of Sciences, 115(11), 2628-2631.

Fisher, P. F. (1999). Models of uncertainty in spatial data. Geographical Information Systems, 1, 191-205.

Gómez, O. S., Juristo, N., & Vegas, S. (2010, September). Replications types in experimental disciplines. In Proceedings of the 2010 ACM-IEEE international symposium on

empirical software engineering and measurement (pp. 1-10).

Goodchild, M. F., & Li, W. (2021). Replication across space and time must be weak in the social and environmental sciences. Proceedings of the National Academy of Sciences, 118(35), e2015759118.

Hey, T., Tansley, S., and Tolle, K. M. (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. Vol 1. Redmond, WA: Microsoft Research.

Kedron, P., & Holler, J. (2022). Replication and the search for the laws in the geographic sciences. Annals of GIS, 28(1), 45-56.

Kedron, P., Bardin, S., Hoffman, T. D., Sachdeva, M., Quick, M., & Holler, J. (2022). A Replication of DiMaggio et al.(2020) in Phoenix, AZ. Annals of Epidemiology, 74, 8-14.

Kedron, P., Bardin, S., Holler, J., Gilman, J., Grady, B., Seeley, M., ... & Yang, W. (2024c). A Framework for Moving Beyond Computational Reproducibility: Lessons from Three Reproductions of Geographical Analyses of COVID-19. Geographical Analysis, 56(1), 163-184.

Kedron, P., Frazier, A., Trgovac, A., Nelson, T., Fotheringham, S. (2021). Reproducibility and Replicability in Geographical Analysis. Geographical Analysis, 53(1), 135-147.

Kedron, P., Holler, J., & Bardin, S. (2024a). Reproducible Research Practices and Barriers to Reproducible Research in Geography: Insights from a Survey. Annals of the American Association of Geographers, 114(2), 369-386.

Kedron, P., Holler, J., & Bardin, S. (2024b). A Survey of Researcher Perceptions of Replication in Geography.

Kinkeldey, C., & Senaratne, H. (2018). Representing Uncertainty. The Geographic Information Science & Technology Body of Knowledge (2nd Quarter 2018 Edition), John P. Wilson (ed.).

Konkol, M., Kray, C., & Pfeiffer, M. (2019). Computational reproducibility in geoscientific papers: Insights from a series of studies with geoscientists and a reproduction study. International Journal of Geographical Information Science, 33(2), 408-429.

Konkol, M., Nüst, D., & Goulier, L. (2020). Publishing computational research-a review of infrastructures for reproducible and transparent scholarly communication. Research Integrity and Peer Review, 5, 1-8.

Leszczynski, A. (2017). Epistemological critiques. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2017 Edition), John P. Wilson (ed.).

Li, L. (2017). Spatial Data Uncertainty. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2017 Edition), John P. Wilson (ed).

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). Geographic

Information Science and Systems (4th Edition). England: John Wiley & Sons.

Matthay, E. C., & Glymour, M. M. (2020). A graphical catalog of threats to validity: Linking social science with epidemiology. Epidemiology, 31(3), 376-384.

Montello, D., & Sutton, P. (2012). An introduction to scientific research methods in geography and environmental studies. Sage.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie du Sert, N., ... & Ioannidis, J. (2017). A manifesto for reproducible science. Nature Human Behaviour, 1(1), 1-9.

National Academies of Sciences, Policy, Global Affairs, Board on Research Data, Information, Division on Engineering, ... & Replicability in Science. (2019). Reproducibility and replicability in science. National Academies Press.

Nüst, D., & Pebesma, E. (2021). Practical reproducibility in geography and geosciences. Annals of the American Association of Geographers, 111(5), 1300-1310.

Ooms, K. and Skarlatidou, A. (2018). Usability Engineering and Evaluation. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2018 Edition), John P. Wilson (ed).

Ostermann, F. O. and Granell, C. (2017). Advancing science with VGI: Reproducibility and replicability of recent studies using VGI. Transactions in GIS, 21 (2), 224–237.

Peng, R. D., & Hicks, S. C. (2021). Reproducible research: a retrospective. Annual Review of Public Health, 42(1), 79-93.

Plesser, H. E. (2018). Reproducibility vs. replicability: a brief history of a confused terminology. Frontiers in Neuroinformatics, 11, 76.

QGIS. (2024) Working with OGC / ISO Protocols. in QGIS Server Guide/Manual. Last accessed 7 April 2024.

Radder, H. (2003). Technology and theory in experimental science. In The philosophy of scientific experimentation (pp. 152-173). University of Pittsburgh Press.

Radder, H. (2012). The material realization of science: From Habermas to experimentation and referential realism (Vol. 294). Springer Science & Business Media.

Sargent, C. L. (1981). The repeatability of significance and the significance of repeatability. European Journal of Parapsychology, 3(4), 423-443.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. Review of General Psychology, 13(2), 90-100.

Sui, D. Z., & Kedron, P. (2021). Reproducibility and Replicability in the Context of the Contested Identities of Geography. Annals of the American Association of Geographers, 111(5), 1275-1283.

Symons, J., & Horner, J. (2014). Software intensive science. Philosophy & Technology, 27, 461-477.

Tullis, J. A., & Kar, B. (2021). Where is the provenance? Ethical replicability and reproducibility in GIScience and its critical applications. Annals of the American Association of Geographers, 111(5), 1318-1328.

Wechsler, S. (2021). Error-based Uncertainty. The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2021 Edition), John P. Wilson (ed.).

Zhang, C. (2019). Ontology for Geospatial Semantic Interoperability. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2019 Edition), John P. Wilson (ed.).

Zhu, A. X., Lu, G., Liu, J., Qin, C. J., and Zhou, C. (2018). Spatial prediction based on the Third Law of Geography. Annals of GIS, 24(4):225–240.