

R or QGIS? A Case Study on Raster Analysis Efficiency in Agricultural Engineering Education

Alberto Macancela-Herrera^{1*}; Eduardo Tacuri-Espinoza²;
Lucía Lupercio-Novillo³; Mateo López-Espinoza⁴

Abstract

Geographic Information Systems (GIS) are essential in agricultural and environmental sciences, yet handling complex spatial data requires efficient, reproducible workflows. This study evaluated the operational efficiency and student perception of R (via R Markdown) versus QGIS for raster analysis. Fifty-two agronomic engineering students from the University of Cuenca completed a raster temperature extraction exercise using both platforms. Survey responses and completion times were analyzed using McNemar's tests, exact binomial tests, and generalized linear models. Results showed that exercise completion was significantly higher in R (98%) than in QGIS (87%; $P=0.031$). R demonstrated superior time efficiency, with 65% of students finishing within three minutes compared to only 29% in QGIS. Additionally, students reported significantly greater self-perceived understanding of the workflow when using R (71% vs. 29%; $P<0.01$). While preference for a primary GIS teaching platform was evenly split (QGIS 52%, R 48%; $P>0.05$), an overwhelming majority (96%; $P<0.001$) indicated they would adopt R as a complementary tool due to execution speed and workflow transparency. Although instructional format asymmetry presents a potential confounding factor, these findings demonstrate that integrating reproducible scripting environments into GIS education enhances computational efficiency, user confidence, and procedural clarity in raster processing.

Keywords: Rstudio, Rmarkdown, GIS, spatial analysis, learning, raster

¿R o QGIS? Un estudio de caso sobre la eficiencia del análisis de imágenes ráster en la enseñanza de ingeniería agronómica

Resumen

Los Sistemas de Información Geográfica (SIG) son esenciales en las ciencias agrícolas y ambientales; sin embargo, el manejo de datos espaciales complejos requiere flujos de trabajo eficientes y reproducibles. Este estudio evaluó la eficiencia operativa y la percepción estudiantil de R (mediante R Markdown) frente a QGIS para el análisis ráster. Cincuenta y dos estudiantes de ingeniería agronómica de la Universidad de Cuenca completaron un ejercicio de extracción de temperatura ráster en ambas plataformas. Las respuestas de la encuesta y los tiempos de ejecución se analizaron mediante la prueba de McNemar, pruebas binomiales exactas y modelos lineales generalizados. Los resultados mostraron que la finalización del ejercicio fue significativamente mayor en R (98%) que en QGIS (87%; $P = 0.031$). R demostró mayor eficiencia temporal: el 65% de los estudiantes terminó en tres minutos o menos, frente a solo un 29% en QGIS. Además, los participantes reportaron una comprensión autopercebida del flujo de trabajo significativamente mayor al usar R (71% frente a 29%; $P < 0.01$). Aunque la preferencia por la plataforma principal de enseñanza estuvo dividida (QGIS 52%, R 48%; $P > 0.05$), una abrumadora mayoría (96%; $P < 0.001$) indicó que adoptaría R como herramienta complementaria debido a su rapidez de ejecución y transparencia metodológica. Aunque la asimetría del formato instruccional representa un potencial factor de confusión, estos hallazgos demuestran que integrar entornos de programación reproducibles en la educación SIG fortalece la eficiencia computacional, la confianza del usuario y la claridad procedimental.

Palabras clave: Rstudio, Rmarkdown, SIG, análisis espacial, aprendizaje, ráster

Recibido: 05 de mayo de 2026
Aceptado: 24 de agosto de 2026

¹ Laboratorio de Geomática, Facultad de Ciencias Agropecuarias, Universidad de Cuenca, Cuenca, Ecuador; Geoinformática, Agrociencias y Soluciones Sostenibles (G.R.A.S.S); Universidad de Cuenca, Cuenca, Ecuador; alberto.macancelah@ucuenca.edu.ec; <https://orcid.org/0000-0003-1461-4364>

² Laboratorio de Geomática, Facultad de Ciencias Agropecuarias, Universidad de Cuenca, Cuenca, Ecuador; Geoinformática, Agrociencias y Soluciones Sostenibles (G.R.A.S.S); Universidad de Cuenca, Cuenca, Ecuador; eduardo.tacuri@ucuenca.edu.ec; <https://orcid.org/0000-0002-4094-209X>

³ Laboratorio de Geomática, Facultad de Ciencias Agropecuarias, Universidad de Cuenca, Cuenca, Ecuador; Geoinformática, Agrociencias y Soluciones Sostenibles (G.R.A.S.S), Universidad de Cuenca, Cuenca, Ecuador; lucia.lupercio@ucuenca.edu.ec; <https://orcid.org/0000-0002-4798-6108>

⁴ Laboratorio de Geomática, Facultad de Ciencias Agropecuarias, Universidad de Cuenca, Cuenca, Ecuador; Geoinformática, Agrociencias y Soluciones Sostenibles (G.R.A.S.S), Universidad de Cuenca, Cuenca, Ecuador; mateo.lopeze@ucuenca.edu.ec; <https://orcid.org/0000-0002-6996-8701>

* Autor de correspondencia

I. INTRODUCTION

A Geographic Information System (GIS) comprises integrated hardware and software for storing, analyzing and displaying geographically referenced data. These systems enable the identification of spatial patterns, trends and relationships through validated spatial analysis (Scholten, 1990). GIS has become increasingly relevant in agriculture and environmental sciences, where advanced geospatial techniques are employed to monitor crop performance, assess vegetation dynamics and evaluate ecological parameters. Beyond scientific applications, GIS methodologies have also been adopted in marketing and business analytics to optimize spatial decision making (Turk et al., 2014; Li et al., 2024).

Specialized programs now include tools to download, visualize, analyze and graph spatial information (Zhou, 2025). Among these, QGIS (Quantum GIS) is a widely used open-source desktop GIS. It benefits from a global community that continuously contributes to its development, documentation and technical support. Its integrated interface enhances usability, making it accessible to both novice and advanced users. The extensive collection of functions, tools and plugins makes QGIS a powerful platform (Li et al., 2024). Nevertheless, several users have reported problems with layer alignment, software performance and interface glitches, which can be counterproductive when managing large raster datasets (Wu et al., 2024).

The incorporation of programming languages into GIS education has become increasingly important as spatial datasets continue to grow in size, complexity and diversity. Traditional graphical GIS interfaces are highly valuable for visualization and exploratory analysis; however, they may become inefficient when repetitive operations, large raster collections or automated workflows are required. In this context, learning programming languages such as R provides students and professionals with the ability to create reproducible, transparent and scalable geospatial analyses (Pebesma and Bivand, 2023). Through scripting, users can document every analytical step, reducing human error and facilitating replication of results by other researchers.

R is a high-level programming language originally developed for statistical analysis. Its application in

GIS and spatial analysis has expanded considerably due to the development of specialized packages that integrate geospatial functionalities (Saqr and López Pernas, 2024). R is now regarded as one of the most important tools for statistical and spatial data analysis and has become essential in research and general data processing workflows (Parra et al., 2023; Graser et al., 2025; Tucker et al., 2023). Compared with integrated GIS platforms such as QGIS or ArcGIS Pro, R often requires fewer computational resources, making it a more efficient alternative for large scale raster processing and spatial modelling (Nguyen Tien et al., 2019). In addition, R integrates data acquisition, preprocessing, spatial analysis, statistical modelling and visualization within a single environment. Packages such as *terra*, *sf*, *tidyverse* and *ggplot2* allow efficient manipulation of raster and vector data while maintaining compatibility with modern data science workflows (Hijmans, 2025; Pebesma, 2018; Wickham et al., 2019). This integration is particularly advantageous in agricultural engineering, where spatial information is frequently combined with experimental, climatic and socioeconomic datasets. Furthermore, learning R promotes coding skills that can be transferred to other programming environments, thereby strengthening students' computational thinking, analytical reasoning and problem-solving abilities in GIS applications.

Raster data are digital images composed of grids of cells or pixels that collectively form an overall image. Raster resolution depends on pixel size; therefore, the smaller the pixel, the higher the resolution. Each pixel possesses a value determined by the information contained in the image (Poynton, 2003). Contemporary raster datasets originate from diverse sources, including satellite imagery, orthophotographs, unmanned aerial vehicles (UAVs) and digital elevation models (DEMs), all of which are fundamental inputs for geospatial analysis and environmental monitoring (Singla et al., 2021). The use of raster data in agriculture and ecology has increased because of the wide availability of raster sources (Steiniger et al., 2017). One specific application is the calculation and analysis of vegetation indices, which help improve crop management by identifying stress, disease or nutrient deficiency (Rasul et al., 2020). Another important use is in climate studies,

where raster data model climatic conditions for agriculture, biodiversity and climate change research (Martin Gomez and Bartolome Muñoz de Luna, 2023).

From an educational perspective, the use of programming languages encourages active learning. Instead of following only predefined graphical procedures, students must understand the logical sequence of the analysis, the structure of spatial objects and the relationships between inputs and outputs. Previous studies have shown that programming-based instruction can improve analytical reasoning, autonomy and long-term retention of methodological concepts (Tucker et al., 2023). In GIS, this means that students not only learn how to generate maps but also how to design workflows capable of processing hundreds of spatial layers automatically. Another relevant advantage is that R is free and open source, which facilitates its adoption in universities with limited economic resources and promotes equitable access to advanced geospatial tools. Consequently, integrating R and other programming languages into GIS curricula may strengthen both technical competencies and research capacity in future agricultural and environmental professionals.

In this context, R offers the facility to download raster data from multiple repositories using, in most cases, only one line of code (Giofandi et al., 2023). We hypothesized that the R Markdown workflow would enable students to complete the raster extraction exercise more efficiently (higher completion rates and shorter times) and with greater perceived understanding than the QGIS handbook-based workflow. We interpret efficiency and comprehension as outcomes associated with the scripted, scaffolded nature of the R environment, rather than as inherent properties of the R software itself. Based on our findings, we propose that incorporating R and similar scripted tools as complementary resources in GIS curricula can enhance reproducibility and analytical clarity, while supplementary coding instruction is needed to address syntax-related challenges.

II. MATERIALS AND METHODS

Data Collection

Students of agronomical engineering at the University of Cuenca performed the exercise in QGIS

and R programs. Subsequently, a questionnaire was administered during August 2025. Participation required students to meet three criteria: (i) completion or current enrollment in courses related to GIS, topography, or precision farming, and (ii) prior experience with at least one geospatial software package (QGIS, ArcMap, ArcGIS Pro) and (iii) understand the logic behind code generation to R programming language. A methodological handbook for QGIS was developed to guide students through the exercise. In parallel, an R Markdown was developed to replicate the same process, detailing each line of code. This approach was intended to help students associate new GIS code with statistical code learned previously (statistics and experimental design). We received survey results from 52 students to perform results. The order in which students performed the exercises was not predetermined; each participant independently decided whether to begin with R or QGIS. Consequently, the sequence of execution varied across students and was dependent on individual choice.

R and QGIS exercise

R and QGIS were software used to carry out the extraction of pixel value from raster. It was necessary to use a raster temperature average image from Ecuador, which was downloaded from the database WorldClim with 0.5 min arc pixel size (~ 1 km). In general, the exercise consisted of: 1) download raster data, 2) save raster data (step 1 and 2 were not evaluated for the questionnaire), 3) open raster file, 4) visualize raster, 5) georeferenced, 6) extract raster values using points from a data frame, 7) create a database in .csv file. For QGIS exercise, the raster was downloaded directly from WorldClim (www.worldclim.org) using a search engine. The remaining procedure basic function from QGIS was executed. On the other hand, for R program the raster download was carried out in the same script, although, last process mentioned was not evaluated. The students must identify, understand and execute every step for both programs. We recommend using the last version of RStudio and QGIS LTR 3.40.

Exercise survey

The questionnaire elaborated to validate the use of R included 8 questions. All these have a relation

with QGIS and R programs because of the raster exercise. With the results of the questionnaire, we searched for validating students' use of R. Below are the questions that were part of the survey:

1. Are you familiar with the software?
2. Were you able to complete the exercise?
3. Was it complicated for you to use the software?
4. If your last answer was yes, what is the reason?
5. How long did it take you to complete the exercise?
6. In which software did you best understand the procedure to extract raster image data?
7. What software would you use as the main tool for GIS?
8. Would you use R as complementary software and why?

Data analysis

A survey consisting of eight questions was designed through Google Forms. From the analysis of the questionnaire results, we calculated the percentages for each question. Of the eight survey questions, only Questions 6 and 7 were subjected to inferential analysis using generalized linear models (GLMs) with a binomial distribution, because these questions directly addressed the primary comparative objectives of the study: students' perceived understanding of the raster extraction procedure and their preferred software for GIS applications. The remaining questions (1–5 and 8) were analyzed descriptively because they were intended primarily to characterize previous familiarity, task completion, perceived difficulty, completion time, and willingness to use R as a complementary tool. Thus, percentages were considered sufficient to describe these aspects of the participants' experience, whereas Questions 6 and 7 allowed direct comparison of categorical responses between R and QGIS. exercise completion

rates (Question 2) were additionally compared between R and QGIS using McNemar's test for paired binary outcomes, given that each student completed the exercise under both conditions. Willingness to adopt R as a complementary tool (Question 8) was evaluated against a null expectation of 50% using an exact binomial test. In addition, bar plots were generated to visually represent the percentage-based responses. All data management and graphics were obtained through R with the ggplot2 library (R Core Team, 2026).

III. RESULTS

According to the initial survey question, all students reported familiarity with QGIS, whereas only two respondents indicated that they were not familiar with R (Figure 1a). After the execution of the exercise, we found an interesting result: 98% of participants successfully completed the raster extraction exercise in R using the R Markdown script, a percentage significantly higher than the 87% obtained for QGIS (McNemar's test, $P = 0.031$; Figure 1b). Likewise, 15% of the students indicated that they experienced difficulties completing the extraction of raster values with both software packages (Figure 1c). The principal reasons are presented in Figure 1d. For QGIS, the main limitation was related to the graphical interface, followed by difficulties associated with the coordinate reference system (CRS), software lag, and the identification of appropriate tools for the procedure. In contrast, for R, the reported difficulties were mainly associated with coding syntax and compatibility issues related to the RStudio version. Despite these challenges, the higher completion rate observed in R suggests that the scripted and sequential workflow provided by R Markdown facilitated a more standardized and reproducible execution of the raster analysis among participants.

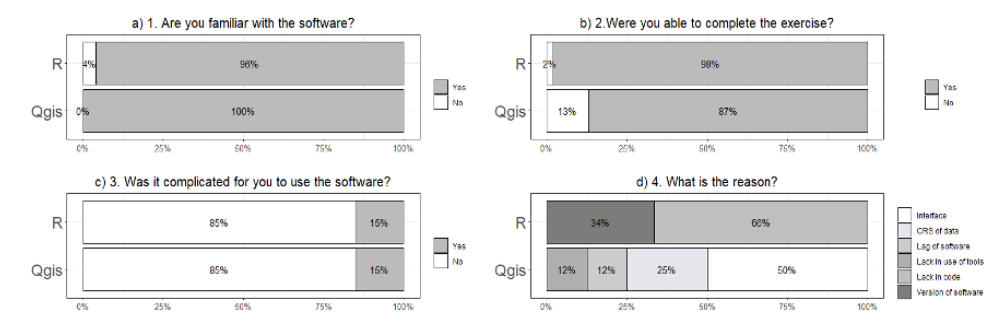


Figure 1. Software familiarity, exercise completion, perceived difficulty, and reported operational obstacles for R and QGIS (n = 52). a) Prior software familiarity. b) Workflow completion rates. c) Perceived operational difficulty. d) Categorization of primary obstacles encountered in each platform.

The time required to complete the raster extraction exercise differed between the two software environments and was lower when using the R Markdown script than when using QGIS (Figure 2a). For the R/R Markdown workflow, 15% of students completed the exercise in approximately one minute, 50% in approximately three minutes, 31% in approximately five minutes, and 4% required more than five minutes. In contrast, students using QGIS generally required approximately five minutes or more to complete the exercise, with a greater proportion of participants requiring more than five minutes compared with the R workflow. These results indicate that the R was associated with shorter and more homogeneous completion times

under the instructional conditions evaluated in this study. However, this difference should be interpreted considering the different levels of instructional guidance provided for each software environment. Furthermore, Figure 2b shows that a significantly greater proportion of students reported a better understanding of the extraction process when using R than when using QGIS ($P < 0.01$). These findings suggest that the sequential and reproducible structure provided by R Markdown was associated with lower execution times and higher self-reported understanding of the analytical workflow and the relationship between each processing step and the final raster output.

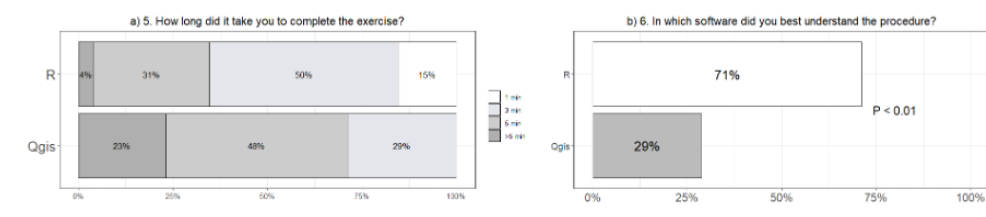


Figure 2. Operational efficiency and self-perceived understanding of raster processing in R and QGIS (n = 52). a) Distribution of exercise completion times across four time intervals (1, 3, 5, and >5 min). b) Proportion of students reporting higher self-perceived comprehension of the analytical workflow by software platform ($P < 0.01$).

Responses to Question 7 (Figure 3a) revealed that the comparison between QGIS and R was not statistically significant ($P > 0.05$), indicating that both software packages showed similar levels of acceptance among the evaluated students. Specifically, 52% of participants preferred QGIS as the primary software for GIS classes, whereas 48% selected R. This result is particularly noteworthy considering that, for most students, this was the

first experience using R for geospatial applications, suggesting a rapid adaptation to the programming-based workflow. The relatively balanced distribution of preferences indicates that students recognized the potential of R despite their limited prior exposure to spatial coding.

According to the final survey question, most respondents (96%) stated that they would use R as a complementary tool for GIS analysis (exact binomial

test, $P < 0.001$). The principal reason was the rapid execution of code, which represented 78% of the responses, followed by the simplicity of generating plots and presenting results (10%), and the greater clarity in understanding the analytical process (6%). Additional comments indicated that students valued the possibility of reproducing the entire workflow and modifying the analysis with minimal manual intervention. In contrast, the small proportion

of students who were not inclined to adopt R as complementary software mainly attributed their decision to difficulties related to coding logic, syntax, and limited experience with programming environments (Figure 3b). These results suggest that, although QGIS remains slightly preferred as the principal teaching platform, students perceive substantial advantages in incorporating R into GIS-based analytical workflows.

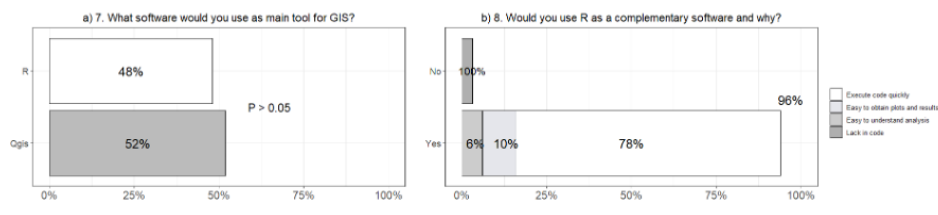


Figure 3. Student platform selection for primary GIS instruction and complementary tool adoption. a) Preferred primary GIS software selection. b) Willingness to adopt R as a complementary GIS tool partitioned by underlying motivations.

In addition to the completion time, the distribution of responses suggests a more homogeneous performance among students when using R. Most participants completed the exercise within a narrow time range, indicating lower variability in execution compared with QGIS, where completion times were more dispersed. This pattern may reflect the standardized workflow provided by the R Markdown script, which guided students through each analytical step in a sequential and reproducible manner. Furthermore, the higher success rate observed in R was achieved even though most students had no previous experience applying R to geospatial analyses, suggesting that the scripted environment reduced operational uncertainty during raster extraction. The questionnaire also revealed that difficulties associated with QGIS were primarily related to interface management and coordinate reference system handling, whereas challenges in R were mainly associated with coding syntax and software version compatibility. These results indicate that the obstacles encountered in QGIS were predominantly procedural, while those in R were technical and potentially easier to overcome through additional programming practice. Overall, the findings support the interpretation that R not only improved efficiency but also promoted a more consistent and reproducible execution of the raster analysis workflow among participants.

IV. DISCUSSION

A critical methodological consideration must be addressed upfront when interpreting the findings of this study. The comparison between R and QGIS was confounded by a structural asymmetry in the instructional materials provided for each software environment. Specifically, the R Markdown script guided students through each analytical step in a fully sequential, executable format, whereas the QGIS handbook provided written instructions that required students to independently locate and execute each tool within the graphical interface. This asymmetry in the level of instructional scaffolding likely contributed to the higher completion rates, shorter execution times, and greater self-perceived understanding observed in the R workflow, independent of any inherent advantage of the software itself.

Therefore, we frame our conclusions not in terms of R versus QGIS as software platforms, but rather in terms of scripted, reproducible workflows versus manual GUI-based workflows. The scripted environment reduced operational uncertainty, provided immediate feedback, and explicitly documented each analytical step pedagogical features that can improve student performance regardless of the underlying software. This distinction is essential for correctly interpreting our results and for designing future comparative studies.

Future research should consider standardizing the instructional format across platforms (e.g., equally detailed step-by-step guides for both QGIS and R) or using a counterbalanced design to isolate the effect of the software from the effect of the guidance material.

Previous research

Bearing this limitation in mind, our findings indicate that the R Markdown workflow was associated with greater efficiency than the QGIS handbook-based workflow for extracting pixel values in basic raster exercises. This supports the growing body of literature that highlights the advantages of programming-based workflows for spatial analysis (Pebesma and Bivand, 2023). The higher completion rate and shorter execution time observed in our study suggest that scripted environments can reduce operational complexity and improve consistency during geospatial procedures. In addition, the use of R Markdown enhanced students' comprehension by providing a transparent and reproducible workflow in which every analytical step was explicitly documented.

Reproducibility is increasingly recognized as a fundamental principle in spatial and environmental research, and scripting languages such as R facilitate the replication, verification, and extension of analyses (Pebesma and Bivand, 2023). The ability to download, preprocess, analyze, and visualize raster data within a single computational environment also represents a substantial advantage for agricultural and environmental applications, where large datasets are frequently used (Nguyen Tien et al., 2019). Previous studies have recommended the use of programming languages for handling extensive raster information because they allow automation of repetitive tasks and reduce the limitations associated with manual graphical workflows (Giofandi et al., 2023; Zhou et al., 2018).

Our results are consistent with previous studies reporting that programming-based approaches provide greater efficiency and reproducibility for spatial analysis than conventional graphical GIS workflows. Nguyen Tien et al. (2019) demonstrated that R offers high versatility for processing large geospatial datasets, particularly in environmental and agricultural applications. Similarly, Pebesma and Bivand (2023) emphasised that scripting-based

workflows improve transparency, scalability, and reproducibility compared with manual operations in desktop GIS software. In educational contexts, Tucker et al. (2023) showed that learning R strengthens analytical reasoning and data management skills.

Moreover, the availability of specialized packages has considerably expanded the capacity of R for efficient raster and vector processing. Although none of the participants had prior experience using R for GIS applications, nearly half selected it as their preferred tool, suggesting that well-designed coding environments may facilitate the adoption of programming approaches even among novice geospatial users. R is a flexible and powerful statistical software environment widely applied to GIS analysis, and it has been increasingly adopted by researchers, professionals, and students across the social sciences, humanities, and STEM disciplines (Tucker et al., 2023). R has been specialized in spatial analysis due to the development of libraries (e.g., raster or terra), facilitating the operation of raster images (Hijmans, 2020). Several important research projects have used R and these libraries to obtain results for agriculture, climate, and vegetation health (Nguyen Tien et al., 2019). In Peru, the introduction of programming languages in public universities revealed efficiency for topics that involve mathematics (Paucar-Curasma et al., 2023). According to our survey results, the R Markdown workflow was associated with shorter execution times, and participants reported higher self-perceived comprehension of the extraction procedure compared with QGIS.

Broader Educational and Research Context

Earlier research has shown that students who engage with programming languages achieve higher academic performance in lectures, largely due to their enhanced problem-solving capacity and skill development (Cuevas et al., 2025). This finding is particularly relevant in the context of our study, where the scripted R workflow appears to have facilitated not only task completion but also a clearer understanding of the analytical process. Furthermore, from 1996 to 2022, scientific production in Europe increased considerably; this phenomenon responds mainly to reforms that occurred 20 years ago within the higher education system of the European Union

(Knutas et al., 2021). Notably, a key finding of that trend was the higher rate of scientific publication among researchers who developed the capability to utilize a programming language. In this regard, our study specifically highlights the use of R (Martin Gomez and Bartolome Muñoz de Luna, 2023), suggesting that this programming language has the potential to develop inherent abilities for future research among students who employ R as a resource for data and spatial analysis. Thus, integrating R into GIS education may not only improve immediate task performance but also foster long-term research capacities and scholarly output.

Pedagogical Challenges and Implications for GIS Education

Despite the growing relevance of programming languages in geospatial sciences, the incorporation of R into GIS education continues to present important pedagogical challenges. Previous studies have reported that students frequently experience difficulties when transitioning from graphical interfaces to code-based analytical environments, particularly because programming requires the simultaneous understanding of syntax, logical structures, data organization, and analytical reasoning (Pavlenko et al., 2022). These cognitive demands are often intensified in spatial analysis, where students must also comprehend coordinate systems, raster and vector structures, and the sequential nature of geoprocessing workflows. Consequently, high levels of frustration, reduced self-confidence, and lower persistence have been documented in introductory programming courses (Mladenović et al., 2017). Educational research suggests that these difficulties can be mitigated through scaffolded instruction, guided practice, and interactive learning activities that progressively connect theoretical concepts with practical applications (Portella-Cleves and Rodríguez-Hernández, 2024).

In our study, the use of an R Markdown script functioned as a structured learning support, allowing students to visualize each command and immediately observe its effect on the raster analysis. This step-by-step approach likely reduced cognitive overload and facilitated the comprehension of geospatial procedures, reinforcing the value of reproducible and guided programming

environments in GIS education. Furthermore, the challenges identified by participants regarding R syntax and coding should not be interpreted solely as limitations of the software, but rather as indicators of insufficient exposure to computational thinking within traditional GIS curricula. Son et al. (2021) emphasized that meaningful learning in quantitative disciplines is strengthened when students repeatedly apply concepts through practical exercises, feedback, and reflection. Similarly, Tucker et al. (2023) demonstrated that teaching with R promotes deeper analytical understanding and long-term skill development when programming activities are integrated progressively into coursework.

In the context, students were already familiar with R for statistical purposes, yet they had limited experience applying it to geospatial problems. This distinction is important because spatial programming involves additional competencies, including the manipulation of spatial objects, coordinate reference systems, and raster operations. Therefore, reinforcing R training through complementary workshops, interactive tutorials, and project-based GIS exercises may improve both technical proficiency and conceptual understanding. Such strategies could help students overcome initial barriers to coding while simultaneously strengthening their capacity to conduct efficient, reproducible, and scientifically rigorous spatial analyses, which are increasingly required in agricultural, environmental, and climate-related research (Pebesma and Bivand, 2023).

A major limitation of R and other programming languages in GIS education is the steep learning curve faced by university students, particularly those in agronomical engineering programs. Many students have limited prior exposure to coding, algorithmic thinking, and data structures, which can generate anxiety, frustration, and reduced motivation during spatial analysis exercises (Pavlenko et al., 2022). In addition, agronomical students often prefer field-based and applied activities, making abstract programming concepts more difficult to assimilate. Challenges related to syntax, package management, debugging, and coordinate reference systems may further hinder learning, emphasizing the need for gradual, practice-oriented, and interdisciplinary teaching strategies (Tucker et al., 2023; Son et al.,

2021).

Additional Limitations and Future Research

Beyond the instructional confound discussed above, several other methodological limitations must be acknowledged when interpreting the findings. First, the sample size was relatively small ($n = 52$) and restricted to agronomic engineering students at a single institution, which may limit the generalizability of the results to other academic programs, skill levels, or institutional contexts. Second, the evaluation was confined to a single raster temperature extraction exercise; thus, performance and software acceptance may differ when executing other geospatial workflows, such as vector processing, spatial interpolation, or complex multi-criteria analyses. Third, the exercise execution order was not randomized or counterbalanced; participants independently decided whether to start with R or QGIS, introducing potential sequence, fatigue, or cross-platform learning transfer effects. Finally, key outcome variables specifically process understanding and operational difficulty were measured strictly through self-reported survey responses rather than objective pre- and post-instructional knowledge assessments. While self-perceived clarity provides valuable insights into user confidence and workflow transparency, it does not directly capture verified conceptual retention or skill acquisition, as lower operational friction in scripted environments may artificially elevate perceived comprehension.

Consequently, future research should implement randomized, counterbalanced designs across multi-institutional cohorts, employ standardized instructional materials, and incorporate objective pre- and post-exercise conceptual evaluations to isolate intrinsic software capabilities from pedagogical formatting and rigorously quantify actual learning gains in geospatial education. Longitudinal studies would also be valuable to assess long-term retention and skill transfer.

Practical Recommendations for Educators

Based on our findings and their limitations, we recommend that GIS educators adopt R Markdown or equivalent scripted tools to create transparent, reproducible, and modifiable workflows, while providing scaffolded scripts that students can

progressively modify as their skills develop. A blended instructional approach combining GUI based QGIS for intuitive visualization with code based R for automation and reproducibility is advocated to leverage the strengths of both environments. Supplementary coding support through workshops, tutorials, and office hours should address syntax and debugging challenges, which students identified as primary barriers. Finally, we encourage using R for reproducibility demanding assignments, such as research projects or theses, to reinforce best practices in scientific computing and better prepare students for data driven careers in agricultural and environmental sciences.

V. CONCLUSION

This study demonstrates that the R Markdown workflow was associated with higher completion rates, shorter execution times, and greater perceived understanding compared with the QGIS handbook-based approach. However, these advantages are substantially confounded by instructional asymmetry an executable script versus a static manual precluding any claim of inherent software superiority. Instead, our findings underscore the pedagogical value of scripted, reproducible workflows in GIS education. While 96% of students would adopt R as a complementary tool, primary preferences remained balanced (52% QGIS vs. 48% R), supporting a blended instructional model that combines intuitive visualization with code-based automation.

From an educational perspective, we recommend scaffolded R Markdown scripts to enhance transparency and reproducibility, complemented by workshops to address syntax and coding challenges identified as key barriers. Critically, the main obstacle was insufficient prior programming training, not R's geospatial functionality itself. Future studies should employ counterbalanced, multi-institutional designs with standardized materials and objective assessments to disentangle software effects from pedagogical formatting, thereby rigorously quantifying actual learning gains in agricultural and environmental spatial analysis education.

Acknowledgments

The authors are grateful to the students of

agronomical engineering of the University of Cuenca who participated in completing the exercise and questionnaire.

Conflict of interests

There is no type of conflict of interest related to the subject of the work.

Ethical consideration

In accordance with the ethical principles set forth in the Declaration of Helsinki, electronic informed consent was secured from all participants before administering the practical exercise and questionnaire. Students received comprehensive information regarding the study objectives, the voluntary character of their participation, and the complete anonymization of their responses. Crucially, participants were notified that opting in or out of the study would not influence their academic performance or standing in any manner.

Availability of deposited data:

The R markdown and data questionnaires are available from the corresponding author upon reasonable request.

VI. REFERENCES

- Cuevas, E. H., Zaldivar, D., and Perez, M. (2025). Impact of Programming Languages on Learning Performance. *International Journal of Information and Communication Technology Education*, 21(1), 1–17. <https://doi.org/10.4018/IJICTE.371419>
- Giofandi, E., Munibah, K., Kraugusteeliana, K., Novalinda, A., and Sekarrini, C. (2023). The Comparison of Vector and Raster Data for The Calculation of Landscape Environment Using a Geographic Information System Approach. *IT Journal Research and Development*, 7(2), 209–219. <https://doi.org/10.25299/itjrd.2023.10878>
- Graser, A., Sutton, T., y Bernasocchi, M. (2025). The QGIS project: Spatial without compromise. *Patterns*, 6(7), 101265. <https://doi.org/10.1016/j.patter.2025.101265>
- Hijmans, R. J. (2020). raster: Geographic Data Analysis and Modeling. R package version 3.4-5. <https://cran.r-project.org/package=raster>
- Hijmans, R. J. (2025). terra: Spatial data analysis (R package version 1.8-60). CRAN. <https://cran.r-project.org/web/packages/terra/index.html>
- Knutas, A., Hynninen, T., and Hujala, M. (2021). To Get Good Student Ratings should you only Teach Programming Courses? Investigation and Implications of Student Evaluations of Teaching in a Software Engineering Context. *Proceedings - International Conference on Software Engineering*, 253–260. <https://doi.org/10.1109/ICSE-SEET52601.2021.00035>
- Li, X., Yue, J., Wang, S., Luo, Y., Su, C., Zhou, J., Xu, D., and Lu, H. (2024). Development of Geographic Information System Architecture Feature Analysis and Evolution Trend Research. *Sustainability (Switzerland)*, 16(1). <https://doi.org/10.3390/su16010137>
- Martin Gomez, S., and Bartolome Muñoz de Luna, A. (2023). Systemic Review through Bibliometric Analysis with RStudio of Skills Learning to Favor the Employability of Its Graduates. *Trends in Higher Education*, 2(1), 101–122. <https://doi.org/10.3390/higheredu2010007>
- Mladenović, M., Krpan, D., and Mladenović, S. (2017). Learning programming from scratch. *Turkish Online Journal of Educational Technology*, 2017(November Special Issue INTE), 419–427. <https://www.semanticscholar.org/paper/Learning-programming-from-Scratch-Mladenović-Krpan/7fc1de582f2fa9230b2c505ba07770adf347d5ef>
- Nguyen Tien, H., Nguyen Thi, H., and KOIKE, K. (2019). High versatility and potential of spatial data analysis with R programming. *Geoinformatics*, 30(1), 3–14. https://doi.org/10.6010/geoinformatics.30.1_3
- Parra, M. I., Sanjuán, E. L., Robustillo, M. C., and Pizarro, M. M. (2023). Using R for teaching and research. <http://arxiv.org/abs/2306.12200>
- Paucar-Curasma, R., Silveira, I. F., Rondon, D., Robles, Z. M. L., and Porras-Ccance, L. E. (2023). Analysis of the Teaching of Programming and Evaluation of Computational Thinking in Recently Admitted Students at a Public University in the Andean Region of Peru. *CEUR Workshop Proceedings*, 3353, 101–110.

<https://dspace.mackenzie.br/items/5b561733-5c90-44c6-8385-49bd7ace46b9>

Pavlenko, L., Pavlenko, M., Khomenko, V., and Mezhyuev, V. (2022). Application of R Programming Language in Learning Statistics. January, 62–72. <https://doi.org/10.5220/0010928500003364>

Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, 10(1), 439–446.

Pebesma, E., and Bivand, R. (2023). Spatial data science: With applications in R. Chapman and Hall/CRC.

Portella-Cleves, J. E., and Rodríguez-Hernández, A. A. (2024). Enhancing Programming Education with an Active Learning Plan and Artificial Intelligence Integration. *Revista Facultad de Ingeniería (Rev. Fac. Ing.)*, 33(67), 2024. <https://doi.org/10.19053/01211129.v33.n67.2024.16328>

Poynton, C. (2003). 1 - Raster images. In C. B. T.-D. V. and H. Poynton (Ed.), *The Morgan Kaufmann Series in Computer Graphics* (pp. 1–16). Morgan Kaufmann. <https://doi.org/https://doi.org/10.1016/B978-155860792-7/50067-7>

R Core Team. (2026). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.r-project.org/>

Rasul, A., Ibrahim, S., Onojeghuo, A. R., y Balzter, H. (2020). A trend analysis of leaf area index and land surface temperature and their relationship from global to local scale. *Land*, 9(10), 1–17. <https://doi.org/10.3390/land9100388>

Sagr, M., and López-Pernas, S. (2024). Learning Analytics Methods and Tutorials: A Practical Guide Using R. *Learning Analytics Methods and Tutorials: A Practical Guide Using R*, June, 1–736. <https://doi.org/10.1007/978-3-031-54464-4>

Scholten, H. J. (1990). Application of geographical information systems in the Netherlands. *Kartografisch Tijdschrift*, 16(4), 27–36. <https://pubmed.ncbi.nlm.nih.gov/1949884/>

Singla, S., Eldawy, A., Diao, T., Mukhopadhyay, A., and Scudiero, E. (2021). Experimental study of big raster and vector database systems. *Proceedings - International Conference on Data Engineering*, 2021-April, 2243–2248. <https://doi.org/10.1109/ICDE51399.2021.00231>

Son, J. Y., Blake, A. B., Fries, L., and Stigler, J. W. (2021). Modeling First: Applying Learning Science to the Teaching of Introductory Statistics. *Journal of Statistics and Data Science Education*, 29(1), 4–21. <https://doi.org/10.1080/10691898.2020.1844106>

Steiniger, S., De La Fuente, H., Fuentes, C., Barton, J., and Muñoz, J. C. (2017). Building a geographic data repository for urban research with free software - Learning from Observatorio.Cedeus.Cl. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(4W2), 147–153. <https://doi.org/10.5194/isprs-archives-XLII-4-W2-147-2017>

Tucker, M. C., Shaw, S. T., Son, J. Y., and Stigler, J. W. (2023). Teaching Statistics and Data Analysis with R. *Journal of Statistics and Data Science Education*, 31(1), 18–32. <https://doi.org/10.1080/26939169.2022.2089410>

Turk, T., Kitapci, O., and Dortyol, I. T. (2014). The Usage of Geographical Information Systems (GIS) in the Marketing Decision Making Process: A Case Study for Determining Supermarket Locations. *Procedia - Social and Behavioral Sciences*, 148(August), 227–235. <https://doi.org/10.1016/j.sbspro.2014.07.038>

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., and Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.

Wu, J., Gan, W., Chao, H. C., and Yu, P. S. (2024). Geospatial Big Data: Survey and Challenges. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 17007–17020.

<https://doi.org/10.1109/JSTARS.2024.3438376>

Zhou, C. (2025). Exploring future GIS visions in the era of the scientific and technological revolution. *Information Geography*, 1(1), 100007. <https://doi.org/10.1016/j.infgeo.2025.100007>

Zhou, G., Pan, Q., Yue, T., Wang, Q., Sha, H., Huang, S., and Liu, X. (2018). Vector and raster data storage based on morton code. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 42(3), 2523–2526. <https://doi.org/10.5194/isprs-archives->