

Review of

Doctoral Dissertation written by Khansa Gulshad

From Flood Dynamics to Nature-Based Solutions: A Geospatial AI Framework for Urban Resilience

*Od dynamiki powodzi do rozwiązań opartych na przyrodzie:
geoprzestrzenne ramy sztucznej inteligencji dla odporności miast*

1. Formal and legal basis

The formal basis for preparing the review is the invitation letter written by prof. dr hab. inż. Ewa Wojciechowska, the Dean of the Faculty of Civil and Environmental Engineering at Gdansk University of Technology (l.dz. 605/WILiŚ/2026). The evaluation was conducted based on the doctoral dissertation prepared in English by the PhD Candidate, Ms. Khansa Gulshad. Prof. dr hab. inż. Michał Szydlowski, from the Department of Geotechnical and Hydraulic Engineering at Gdansk University of Technology, is the supervisor of the dissertation. The dissertation was submitted in both paper and digital forms to the Council of the Discipline of Civil Engineering, Geodesy, and Transport at Gdansk University of Technology. The review was conducted in accordance with the requirements outlined in Article No. 187 of the Act on Higher Education and Science of July 20, 2018 (POL-2018-R-109017), as amended by subsequent regulations (POL-2018-R-109019).

2. Overview of the conducted research

The main subject of the study is the modeling of flood hazard, with special emphasis on urban flooding. The research is focused on implementing modern technologies such as geoprocessing, hydrological modeling, and especially Machine Learning/Deep Learning (ML/DL) methods. In her dissertation, the PhD Candidate correctly identifies the greatest gap that exists in this field currently, i.e. the implementation of ML/DL methods to generate flood susceptibility maps without properly accounting for the physical principles governing water flow in natural environments. The second interest is related to the development of Nature-based Solutions (NbS) and planning Green Infrastructure (GI) in

cities. In this area, the investigations are focused on connecting existing computer vision methods with effective planting strategies, which is also one of today's challenges. The main concept of this doctoral dissertation is to integrate the methods currently used in flood risk analysis and urban water management and to develop a coherent methodology that enables efficient management of urban infrastructure. This concept goes beyond the standalone application of ML/DL methods—which is currently the most common practice—and instead seeks to implement them while giving due consideration to hydrological processes and hydraulic interactions within the system under study.

To illustrate the importance of a integrated approach, this thesis uses the city of Gdansk as a case study. Taking into account the historical flood experiences in Gdansk, the choice is reasonable. As indicated in the thesis, the flood hazard in the city is composed of three different factors: (1) surface runoff from post-glacial moraine hills, (2) river outflow during intense precipitation events, and (3) rising sea levels in the Gulf of Gdańsk. The PhD Candidate also noticed that *"[...] no study applied to a coastal city has yet integrated these [mentioned previously] components [...]."*

Hence, the importance of the presented research is unquestionable. The goals are properly defined, and the novelty is clearly indicated. The applied methods are advanced and consistent with current practices and research. The procedure ensures that the results are valid and that the defined purposes are properly met.

3. Analysis of the dissertation

The submitted dissertation comprises 9 regular chapters, 3 appendices, the list of abbreviations (82 items), the abstract, and references. At the beginning of the dissertation, there are also the graphical abstract and the acknowledgments, with a statement about the lack of substantial funding from external sources. The formal statements are added before the title page. The dissertation includes 51 figures, 30 tables, and 57 equations, most of which appear in the main chapters. Adding list of figures and list of tables would improve readability. There are 288 references, including 218 papers in scientific journals. Some repetitions in the list of references were detected. The problem is described in the detailed remarks.

Chapter 1, entitled *"Introduction,"* presents the problem and illustrates all its different facets. The PhD Candidate starts with a general overview of the modern challenges related to climate change and increasing flood risk, which is supported by examples of past urban floods in Gdansk. Next, the description is focused on the first major concern in this research: the variability of methods for modeling hydrological hazards. The PhD Candidate leads discussion from more physically oriented models like HEC-HMS to modern data-driven methods, including AI algorithms. The second concern is the spatial planning of Nature-based Solutions (NbS) and Green Infrastructure (GI) in the city. The author precisely identifies the main gaps in these two areas and briefly explains the approach applied in the presented research. Then the purpose of the research is formulated:

“[...] to demonstrate that GeoAI, [...], can directly support transparent, equitable, and evidence-based urban adaptation.”

Additionally, the PhD Candidate indicates that this is not only a methodological advancement in urban hydrology and GI planning, and that the key to success is applying GeoAI in a structured, integrated pipeline rather than as a collection of tools. The scientific tasks are specified as follows:

1. *“To quantify the historical and projected effects of LULC dynamics and climate change on surface runoff at basin and city scale, [...]”*
2. *“To develop an interpretable, ensemble machine learning framework, [...], to map urban flood susceptibility and identify the dominant spatial drivers of flood.”*
3. *“To diagnose the spatial inequity of green infrastructure distribution across Gdańsk [...]”*
4. *“To quantify urban greening potential at rooftop and street level [...], and to assess the cumulative stormwater retention benefits of integrated GI interventions across design storm scenarios.”*

The above statements are sufficient for precise definition of the scientific problem. However, to align with scientific methodology, the research hypothesis could also be stated.

The next chapter, entitled *“State of the art”*, describes the current state of knowledge in four specific fields important to the dissertation. The discussion of climate change impacts, changes in land use/land cover (LULC), and applications of hydrological modeling is followed by a discussion of methods for flood susceptibility assessment. This survey also includes specific reflections formulated by the PhD Candidate. Although arguments about time-consuming and calibration problems with hydrologic and hydraulic models are repeated following the scientific literature, the criticism about the black-box nature of AI models is also provided. The latter is totally correct, but it is not usually mentioned in papers on flood susceptibility maps. The next sections illustrate the concept of green infrastructure in cities and its driving factors with examples. The next part discusses the implementation of NbS and GeoAI tools for urban planning.

Chapter 3 is entitled *“Study area and data framework”*. In the first passages, the author clearly explains the motivation for choosing Gdansk, the largest harbor in the Baltic Sea, as a case study. The chapter includes details about this area and an interesting discussion of GI implementation in Polish cities. Table 3.1 lists the source data, with references to the subsequent chapters where the data are used. A short discussion with a classification of the different data types could highlight the differences between spatially distributed data and point data, as well as static data and time-variable data. It would be helpful to specify which data come from Polish sources, which come from European services, and which are global.

Chapters 4 to 7 discuss specific methods used in the research. The internal structure of each such chapter is similar. First, each chapter introduces the topic, then describes the methodology, presents specific results, discusses the particular subject, and concludes briefly. The first two chapters (Chapters 4 and 5) are focused on flood modeling using two

approaches: standard hydrological routing and flood susceptibility mapping with AI methods. The next two chapters (Chapters 6 and 7) are focused on green infrastructure and Nature-based Solutions. In the next paragraphs the methodological part is described with more details.

Chapter 4, entitled “*Land Use Dynamics and Hydrological Response*”, is focused on changes in land use and land cover in the cities, with special emphasis on Gdansk. The investigations use satellite data from Landsat and Sentinel-2. Google Earth Engine (GEE) was used for data processing. The impact of LULC changes on urban hydrology is then modeled using HEC-HMS. Figure 4.1 illustrates the overall concept, while Figure 4.2 shows the catchment. There are also two equations discussing core elements of the method applied for calculating abstractions, namely NRCS-CN (previously known as SCS-CN). According to the text, the SCS-CN version of the unit hydrograph models the transformation of excess rainfall into watershed outflow, and the Muskingum model is applied for flow routing in the channels. Because these methods are not the default choice in the applied software, it would be good to justify choice of them and provide their parameters briefly. It is also unclear what baseflow model was chosen. Table 4.1 summarizes the simulation configuration, while Table 4.2 briefly explains the climate change scenarios. The PhD Candidate focused on projections related to RCP4.5 and RCP8.5 scenarios recommended by the IPCC (Intergovernmental Panel on Climate Change). The results are discussed using 3 tables and 10 figures. In my opinion, the maps showing historical LULC changes and trends may have high practical importance. The results of the hydrological section are mainly focused on changes in the runoff coefficient, which could help design water management facilities.

The next chapter, “*Flood Susceptibility Mapping with Interpretable ML*”, describes ML/DL methods applied to determine flood-prone areas in the city. Figure 5.1 presents the overall approach. The author trained models of three different types: Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Networks (ANN). The data about interventions during floods in the period of 12 years (2010-2022), obtained from the State Fire Station and Gdansk Water Company, are the basis for the learning process. The driving factors are classified as physical and anthropogenic. The SHapley Additive exPlanations (SHAP) is applied for specific sensitivity analysis. In the chapter, there are 21 equations explaining key elements of the models and defining the statistical metrics. Table 5.1 summarizes the model parameters selected in this research. The results are shown in 4 tables and 9 figures. Figure 5.9 presents the final susceptibility maps, while Figure 5.10 shows only the areas prone to flash floods.

The focus on concept of green infrastructure and Nature-based Solutions begins in Chapter 6, “*Spatial Distribution and Determinants of Green Infrastructure*.” Figure 6.1 illustrates the methodology elements, and 9 pivotal equations support their description. In general, the stages include data collection, three steps of data processing, and analysis of results presented at the end. The three-step data processing consists of preprocessing, including the choice of resolution, spatial autocorrelation, and regression between GI

density and “*explanatory factors*”. The grid with 1 km resolution is used to represent spatial distribution of GI density across the city. The methods for calculating autocorrelation are used to check whether GI density values are spatially clustered, dispersed, or randomly distributed. The regression between GI density and “*explanatory factors*” is processed in the Python Colab environment. The methodology used is advanced and ensures accurate results with practical significance. The examples are presented in 6 figures and 2 tables. Appendix B includes the rest of results. It is worth to highlight that chapter 6 includes results of particular theoretical importance, e.g., graphs presenting correlations between explanatory factors (Figure 6.4) and a comparison of regression model performance (Table 6.2). However, there are also results of high utility for practitioners, e.g., a map of GI distribution (Figure 6.2), a map of GI priority areas (Figure 6.7).

Chapter 7, “*GeoAI Framework for NBS Suitability & Stormwater Analysis*”, describes the application of GeoAI tools for robust planning of Nature-based Solutions such as blue-green roofs (BGR), street-level greenery, and street trees, and then presents the estimation of retention and interception in the city. The methodology applied combines geoprocessing procedures, image processing and recognition, and ML/DL methods. The entire procedure is presented with a diagram in Figure 7.1. Different kinds of models are implemented at this stage of investigations, e.g., Mask2Former, XGBoost, and RESILO. The description of the methodology is supported by 8 equations and one table (Table 7.1), which includes a short characteristic of tested rainfall scenarios. In general, the methodology implemented in the models could be described more extensively. The results are presented and illustrated with 5 figures and 2 tables. The figures include examples of maps and city images (ground-level photos) that illustrate the implemented concepts. There are more results in Appendix C. The tables show the characteristics of retention reduction and interception impact. Two publicly available web services, called *GreenRoofs* and *GreenStreets*, are the effect of the analyses, too.

The integration of the above concepts is located in Chapter 8, “*Synthesis, Implications & Transferability*”. It starts with a graphical summary of the previously explained ideas (Figure 8.1). The chapter is a broad discussion, but conclusions come in the following Chapter 9. The PhD Candidate leads this discussion from a synthesis of the methodological chapters, through reflections on the methodological contribution made by the GeoAI framework, and considerations about its uncertainty and robustness. The recommendations for the city of Gdansk are presented in three tables, including policy recommendations, implementation pathways, and spatial conflicts.

In the chapter 9, “*Conclusion and Future Research*”, in the first steps it is proved that the purpose of the research was achieved and explains how it was done. Then the focus is on the scientific problems defined in the Introduction. Chapter 4 presents the first, quantification of LULC dynamics and the effects of climate change on surface runoff, what is the scientific problem no. 1. The next chapter illustrates flood susceptibility assessment in the urban environment, which is problem no. 2. Chapter 6 is focused on problem no. 3,

i.e., green infrastructure distribution across the city of Gdansk and its relationship to hydrological characteristics. Finally, Chapter 7 determines urban greening potential, taking into account the benefits of cumulative stormwater retention, which is problem no. 4. Additionally, the author emphasized some specific innovations achieved during research, e.g., the application of “*a scenario-based GPS (greening potential score) framework that links green roof technical feasibility to planning constraints.*” Next, the limitations of the study are summarized and briefly discussed. The author presents observations and suggestions that go beyond current practice and the state of knowledge, e.g., she suggests that flood type is not taken into account in flood susceptibility mapping, but the obtained results indicate significance of this factor. Finally, the author presents the conclusions and highlights the “most important finding”: combining the four more-or-less distinct methods into a single procedure. According to the reviewer’s knowledge, this is exactly the current concern in the discipline and the purpose to achieve for many scientists working in the field of flood risk management. Hence, innovative character of the proposed methodology is obvious.

It is worth highlighting that describing the methodology in four separate chapters has a definitely better effect than following the typical scheme: *materials-methods-results-discussion*. The presented description is clear and understandable. The connections between elements are well defined and finally synthesized in Chapter 8.

4. Remarks and questions

Major remarks and questions

- A. The hydrological components are present in the hydrodynamic models now. Software such as HEC-RAS or Iber can simulate 2D flow over specific areas, including rainfall and infiltration. Is it possible that a model based on the full mass and momentum balance equations, e.g., shallow water equations, give more interesting results? How such a model could be verified on the data used in the current research?
- B. The bimodality of the graph representing dependence of the frequency on distance from the river channel (Figure 5.2) suggests that there are fluvial and non-fluvial floods in the dataset. Supposedly, inundations at long distances from the river were caused by rainfall and surface runoff or failure of the collectors. In the Conclusions, there is a comment about the role of the flood type (in section “*9.1. Limitations and future research*”). How could this aspect impact the flood susceptibility maps presented? Is it possible that the distribution of flood susceptibility areas the city is different if the flood type is taken into account?
- C. Is it possible to use satellite images of the historical floods to calibrate ML/DL models? Would such a model still produce a flood-susceptibility map, or could it simulate the flood process and produce a full flood hazard map?

- D. A new class of methods appears in the field of hydrological modeling, namely physics-informed neural networks (PINN). It integrates classical deep learning based on artificial neural networks with physically based modeling based on classical balance principles (mass, momentum, etc.). Is it possible to implement such concepts in the framework presented here?

Additional remarks and questions

- (1) It would be good if the hypothesis behind the conducted research is formulated and the conclusions refer to it.
- (2) Why is the river network taken from the global service OpenStreetMap? Was it not possible to obtain such data from the Polish MPHP (Polish: *Mapa Podziału Hydrograficznego Polski*)? Which of the sources is more accurate?
- (3) The analysis is done for the city, which requires spatial data at a finer resolution or highly accurate vector layers. However, some raster data describing LULC and vegetation indices (NDVI, NDWI) have a rather coarse resolution (30 m, 100 m). Has it impacted the results?
- (4) What are the parameters of the particular models applied in HEC-HMS, namely the unit hydrograph based on SCS-CN and the Muskingum model? What model was chosen for baseflow? What kind of formula and definition is adopted for widely used runoff coefficient?
- (5) Different measures are mentioned for evaluation of the ML/DL models' performance (RMSE, MAE, ACC, etc.). Which were used during training (70% of cases), and which were used for verification (30% of cases)? What kind of training method was implemented, e.g., gradient methods, simple search methods, global optimization, trial-and-error?
- (6) Is the application of Google Colab as a Python environment motivated by the modules available there? Is it possible to make the same analysis in a standard Python environment with `pip` or `conda` for package management?
- (7) Generally, it would be good if the different models used in Chapter 7, i.e., Mark2Former, XGBoost, and RESILO, were better described.

Detected problems and technical inaccuracies

- (1) Some elements should be mentioned in the Contents:
 - a. the graphical abstract (page 6)
 - b. the acknowledgments (page 7)
 - c. the References (page 150)
 - d. the Appendices (start on page 172)
- (2) In the Introduction (page 17), one and the same concept is formulated twice as follows:

“Machine learning and explainable artificial intelligence offer a pathway beyond these modelling limitations, but their transformative potential lies not in applying them individually, but in chaining them with remote sensing and hydrological analysis into a coherent analytical pipeline.”

And the following sentence:

“Machine learning models offer a pathway beyond these modelling limitations, but their transformative potential lies in chaining them with remote sensing and hydrological analysis into a coherent pipeline rather than applying them individually [...]”

- (3) Figure 3.1: The scale of the DEM in the top could be rounded, e.g., from -20 to 180 m a.s.l.
- (4) It's not quite clear what kind of unit [ms] is used in the sentence *“Storm surges, [...], can raise sea levels by over 1.5 ms, with projections indicating potential surges reaching up to 2.5 ms.”* (page 32).
- (5) Table 3.1: The description in the second column (Resolution & Year) is a little bit confusing. The scale of the map (e.g., soil map) represents the level of detail, not resolution. Additionally, the vertical accuracy could be provided in the case of DEMs.
- (6) Figure 4.1: The fonts used in the descriptions should be larger. These are not clearly visible.
- (7) Figure 4.2: The scale of the DEM could be rounded, e.g., from 0 to 160 m a.s.l.
- (8) Wrong figure number in Chapter 4 on page 46. There is “Figure 2.3” instead of “Figure 4.3”.
- (9) Figure 5.1: The flowchart is a little bit complex. It's not so clear what the direction of information is in this chart.
- (10) Figure 5.2: The graphs for Soil and LULC should be presented and interpreted in a different way than the graphs for other elements. Soil and LULC represent discontinuous, unordered characteristics. Additionally, the vertical axis representing frequency should have the same scale. Maybe it would be better to use relative frequency to keep the vertical values between 0 and 1.
- (11) Figure 6.4: It's not clear what PR1, PR2, PR3, and PR4 are in the list of dependent variables. Are these projected precipitations related to scenarios RCP4.5 and RCP8.5 determined for the periods 2021-2050 and 2071-2100?
- (12) In section 6.1.4, the term *“explanatory factors”* is used. This term should be explained here. The *“factors”* are mentioned in the Introduction (page 26), but these are not called *“explanatory”*. For the reader not familiar with terminology used in regression methods, this statement might be confusing. In my opinion, it should be clarified in the text.
- (13) In the description of equation 7.3, the term *“proportion”* is used. However, the context suggests that *“ratio”* would fit better.

- (14) Wrong figure number in Appendix A on page 174. There is “Figure 3” instead of “Figure A.2”.
- (15) Wrong number in Figure C.2 in Appendix C on page 189. There is “Figure C4” instead of “Figure C.2”
- (16) There are some repetitions in the list of references, e.g.:
- 44 Cheng, B., Misra, I., Schwing, A.G., Kirillov, A., Girdhar, R., 2022a. Masked-attention mask transformer for universal image segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1290–1299.
 - 45 Cheng, B., Misra, I., Schwing, A.G., Kirillov, A., Girdhar, R., 2022b. Masked-attention mask transformer for universal image segmentation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1290–1299.
- (17) There are also some formatting problems in the list of references, e.g.:
- 73 Ewelina, C., Liliana, J., Katarzyna, K., Mateusz, K., Wojciech, Ł., Aleksandra, Ł., Piotr, O., Karolina, P., Paweł, P., Agnieszka, R., Agnieszka, S., Magdalena, Z., 2021a. Raport o stanie polskich miast. Środowisko i adaptacja do zmian klimatu / Report on the State of Polish Cities: Environment and Adaptation to Climate Change. Warszawa – Kraków.

5. Final evaluation of the dissertation

The research problem has been correctly defined. Its significance is beyond doubt, and the methods used are advanced and consistent with the current state of knowledge. No serious errors or shortcomings were detected. The problem addressed is highly complex, and the scope of the analyses conducted is very broad. The choice of an advanced and diverse methodology is fully justified. All these elements together make the submitted doctoral dissertation original and highly innovative. The results represent a valuable contribution to the state of the art and hold enormous practical potential. The presented methodology can be directly applied in Gdańsk and easily adapted to other locations. Some of the results are already available to potential stakeholders and decision-makers through professionally designed websites. This also serves as a strong example of how to organize public communication.

The research topic is a critical issue in Civil Engineering, Geodesy, and Transport. The dissertation emphasizes the importance of accurate hydrological assessment and identifying flood-prone areas in planning and designing urban infrastructure. Consequently, it has a significant impact on other fields of scientific research and practice, such as hydrology, flood protection, and land-use planning. The usefulness of green infrastructure and nature-based solutions has been clearly demonstrated. Furthermore, the study illustrates the potential of GeoAI tools in land-use planning and management. The analyses conducted contribute to a better understanding of processes, impacts, and management practices in the urban environment. They may also influence the further development of hydrological and hydrodynamic modeling, as well as the application of Machine Learning and Deep Learning methods in hydrology.

6. Conclusions

Overall, the doctoral dissertation is a valuable contribution to research on urban flood modeling, flood prevention, and spatial planning of flood protection infrastructure. The dissertation also demonstrated the potential application of the results. The PhD Candidate has demonstrated substantive knowledge and a deep understanding of the professional literature in formulating the goals of the work, conducting research and analyzing results. In my opinion, the PhD dissertation deserves distinction due to its broad scope and advancement of the presented research, its importance to science, and its practical usefulness.

The PhD thesis submitted by Ms. Khansa Gulshad meets all the requirements for doctoral dissertations by the Act of 20 July 2018 Law on Higher Education and Science. Based on my positive assessment, I recommend to the Chairman of the Scientific Council of the Discipline of Civil Engineering, Geodesy and Transport to admit Ms. Khansa Gulshad to the next stages of the doctoral dissertation process, including public defense of the dissertation.



Tomasz Dysarz