RESPOND: Realistic Environment Simulation of Population and Natural Disasters with LLM-Driven Agents

Roman Sultimov[1,2], Mikhail Mozikov[2,3], Dmitrii Abramov[1], Maria Kovalchuk[1,2], Maksim Malykh[1,2], Ilya Makarov[4], Andrei Osiptsov[5], Aleksandr Volkov[5], Yury Maximov[5]

[1] Lomonosov Moscow State University, Lomonosovsky prospekt, 27/1, Moscow 119192 Russia
[2] Moscow Independent Research Institute of Artificial Intelligence, Moscow Russia
[3] AI Research Institute, 105064 Moscow, Russia
[4] ISP RAS, Yerevan Armenia
[5] Interdata LLC, Astana 010013 Kazakhstan

Abstract

Climate change is driving more frequent and severe disasters, putting people and infrastructure at risk. Protecting communities requires models that capture both natural disasters dynamics and how people behave under extreme conditions. This demo presents RESPOND, a multi-agent LLM-enhanced platform that jointly simulates natural hazards and human response. RESPOND couples high-fidelity flood AI forecasting an agent-based model of human behavior. LLM modules improve each agent decision-making, enabling context-aware reasoning over alerts, road closures, social signals, and changing water levels. The system simulates evacuation flows, resource seeking, and communication patterns producing actionable outputs for emergency management, urban planning, and policy. In the live demo one can run what-if or predicted scenarios, adjust assumptions, and observe emergent population behavior and risk hot spots in real time. By tightly coupling dynamic hazards with LLM-driven multi-agent behavior, RESPOND moves beyond fragmented tools and offers a practical, integrated platform for disaster preparedness and response.

Introduction

The frequency and intensity of natural disasters are rising, demanding robust and adaptive management strategies to protect people and infrastructure (Alexander 2018; Krichen et al. 2024; Liu et al. 2024; Taniushkina et al. 2024; Shevchenko et al. 2024). Effective preparedness depends on accurately predicting impacts on at-risk people and realistically simulating their responses as conditions change. Yet many simulators simplify human behavior and ignore real-time environmental data, which limits their value in critical situations.

Current simulation approaches often limited as using simplified models of human behavior and lacking real-time environmental data integration, which restrict their predictive power in complex, evolving scenarios (Martin et al. 2016; Kattenstroth et al. 2024; Törnberg and Larooij 2025; Gao 2024). Multi-Agent Systems (MAS) can model complex disaster settings by representing individuals (agents) with distinct attributes, decisions, and interactions — capturing emergent behaviors like evacuation dynamics and communication under stress (Dorri, Kanhere, and Jurdak 2018; Keykhaei et al. 2024). Unfortunately, even advanced MAS struggle to incorporate nuanced human cognitive processes and adapt to rapidly changing environmental conditions or unstructured information inherent during disasters (Xue et al. 2024). Large Language Models (LLMs) enable missing capabilities with their power in natural language understanding, generation, and complex reasoning.

This paper presents RESPOND, a demo system that combines MAS, LLMs, and advanced flood forecasting to jointly simulate hazards and human response. RESPOND provides high-fidelity, near real-time flood dynamics prediction and uses LLM-driven agent cognition to produce realistic, context-aware decisions at population scale. It advances the state-of-the-art by

- *dynamic hazard integration*, i.e. near-real-time high-fidelity forecasts of flooding dynamics;
- *LLM-driven agent cognition* enabling interpretation of alerts, road closures, and other unstructured inputs and context-aware decision making beyond rule-based logic;
- *complex population response simulation* capturing evacuation, resource access, and communication across scenarios to inform planning and policy.

Related Work

Research in multi-agent systems has long supported disaster management by simulating evacuation dynamics, coordination, and communication during crises (Massaguer, Balasubramanian, and Mehrotra 2006). More recently, the integration of LLMs into agent-based modeling has enabled agents to reason and communicate in more human-like ways (Gao 2024; Larooij 2025). Emerging work showcases LLM-driven agents in large-scale simulations for understanding behavior and social dynamics (Piao 2025; Chen et al. 2025; Grobelnik et al. 2025), as well as for decision support in disaster contexts (Xu 2025; Xie 2025; Bao et al. 2025; Owen, Maximov, and Chertkov 2019). These systems advance realism by allowing agents to demonstrate sophisticated cognition and emergent collective behaviors (Piao et al. 2025; Zhang et al. 2025).

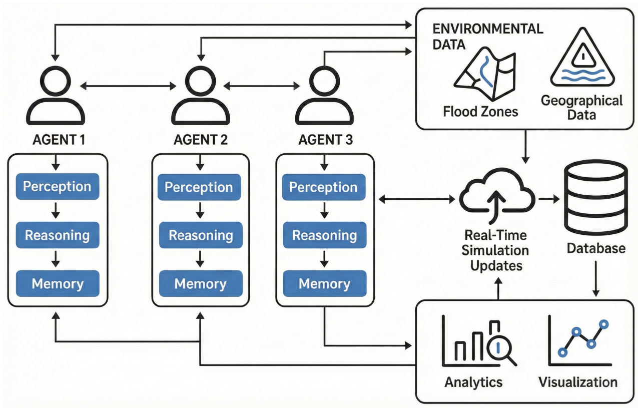

Figure 1: RESPOND architecture: interacting agents enhanced by LLMs and real-time flood data, produce population-scale simulations and predictions

In parallel, natural disasters and flood forecasting models have advanced dramatically, producing high-resolution, real-time flood depth maps critical for preparedness and response (Okeukwu-Ogbonnaya et al. 2025; Floodbase 2025; Delrieu et al. 2021). Recent innovations include Large Multimodal Models (LMMs) for flood-depth estimation from images (Lyu 2025; DeepFlood 2025) and LLM–GIS integration for geospatial analysis (Zhu 2024).

Unfortunately, a key state-of-the-art gap remains: while MAS and LLM research improves agent realism, it rarely integrates dynamic, high-fidelity hazard forecasts into simulations. Conversely, natural disasters forecasting systems generate rich predictions but lack mechanisms to translate them into real-time, agent-level behavioral responses that capture complex human decision-making under stress (Alomari et al. 2023; Ma et al. 2025).

RESPOND closes this gap by combining real-time flood maps with LLM-driven agents to simulate physical, social, and logistical disaster impacts in a unified testbed.

System Description

RESPOND is a modular platform that simulates the interaction between floods dynamics and human responses through three core modules.

Flood forecasting and mapping module generates high-resolution flood-depth maps as the main environmental input. It combines real-time meteorological, river gauge, and topographic data with hydrological and machine learning models (Mitrovic et al. 2023) to simulate water flow and inundation (Okeukwu-Ogbonnaya et al. 2025). The maps are continuously updated, refined with multimodal models that integrate visual data (Lyu 2025), and organized within a geospatial information framework for spatial reasoning.

Multi-agent simulation module represents individuals as agents, each with unique attributes, goals, and decision-making processes. Agents sense their environment, including local flood depths, and interact through a simulated communication network.

Figure 2: RESPOND user interface, simultaneously demonstrating agents behavior and natural disaster impact.

LLM-enhanced agent cognition module equips agents with advanced reasoning capabilities using LLMs. Unlike rule-based systems, LLMs let agents interpret unstructured inputs such as road-closure reports, social media updates, or official alerts. Agents can reason about risks, evaluate evacuation routes, adjust priorities (e.g., family safety over property protection), and communicate in natural language. This creates more realistic simulations of evacuation, resource seeking, and information sharing (Gao 2024; Chen et al. 2025; Grobelnik et al. 2025).

Together, these modules create a powerful simulation where real-time environmental dynamics directly shape adaptive, human-like population behaviors. RESPOND provides new insights into how collective responses emerge during disasters, helping researchers and practitioners better prepare for and manage crises.

Discussion and Conclusion

Our platform substantially advances disaster preparedness by integrating dynamic environmental models with realistic human behavior. Unlike traditional simulators, it establishes a continuous feedback loop between flood forecasting, multi-agent simulation, and LLM-driven cognition, enabling adaptive responses to evolving threats (Delrieu et al. 2021; Gao 2024). This interplay is critical for understanding cascading disaster effects.

LLM-enhanced cognition allow agents to interpret unstructured information, reason in context, and adapt decisions beyond fixed rules (Gao 2024; Chen et al. 2025; Grobelnik et al. 2025). This allows exploration of scenarios with ambiguity and social dynamics that simpler models cannot capture (Piao 2025; Zhang et al. 2025).

RESPOND delivers actionable insights for emergency managers, urban planners, and policymakers. It can improve evacuation planning, resource allocation, and communication strategies, while also supporting resilience planning and policy evaluation (Alomari et al. 2023; Ma et al. 2025).

Future directions optimizing scalability and LLM performance, expanding multimodal data integration, and exploring human-in-the-loop simulation (Lyu 2025; Chen et al. 2024; Zhu 2024). Thanks to its modular design, RESPOND can also be extended to other disaster types.

Acknowledgments

This work of R.S, M.K, and M.M. was supported by the Ministry of Economic Development of the Russian Federation (agreement No. 139-15-2025-013, dated June 20, 2025, IGK 000000C313925P4B0002).

References

Alexander, D. 2018. *Natural disasters*. Routledge.

Alomari, A.; et al. 2023. A situation-aware emergency evacuation (SAEE) model using multi-agent-based simulation for crisis management after earthquake warning. *International Journal of Disaster Risk Reduction*.

Bao, D.; Xue, Z. G.; Hiatt, M.; Xu, K.; Harris, C. K.; and Trepanier, J. C. 2025. A machine learning-based prediction-to-map framework for rapid and accurate spatial flood prediction. *npj Natural Hazards*, 2(71).

Chen, Q.; et al. 2024. Multimodal social sensing for the spatio-temporal evolution and assessment of nature disasters. *Sensors*, 24(18): 5889.

Chen, Y.; et al. 2025. Simulating Human Behavior with AI Agents. Stanford HAI.

DeepFlood. 2025. DeepFlood for Inundated Vegetation High-Resolution Dataset for Accurate Flood Mapping and Segmentation. *Scientific Data*. Nature Scientific Data.

Delrieu, G.; et al. 2021. Probabilistic Flood Hazard Maps from Monte Carlo Derived Peak Flow Values—An Application to Flood Risk Management in Zamora City (Spain). *Applied Sciences*, 11(14): 6629.

Dorri, A.; Kanhere, S. S.; and Jurdak, R. 2018. Multi-agent systems: A survey. *Ieee Access*, 6: 28573–28593.

Floodbase. 2025. Global Flood Database. https://global-flood-database.cloudtostreet.ai/.

Gao, C. 2024. Large language models empowered agent-based modeling and simulation: a survey and perspectives. *Nature*.

Grobelnik, A. M.; et al. 2025. Emergent Behaviors from LLM-Agent Simulations. In *SiKDD*.

Kattenstroth, F.; Disselkamp, J.-P.; Lick, J.; and Dumitrescu, R. 2024. Challenges in the implementation of simulation models for the digital factory twin - a systematic literature review. *Procedia CIRP*, 128: 442–447.

Keykhaei, M.; Samany, N. N.; Jelokhani-Niaraki, M.; and Zlatanova, S. 2024. Multi-agent-based human cognition simulation of Situation-aware earthquake emergency evacuation. *International Journal of Disaster Risk Reduction*.

Krichen, M.; Abdalzaher, M. S.; Elwekeil, M.; and Fouda, M. M. 2024. Managing natural disasters: An analysis of technological advancements, opportunities, and challenges. *Internet of Things and Cyber-Physical Systems*, 4: 99–109.

Larooij, M. 2025. Do Large Language Models Solve the Problems of Agent-Based Modeling? arXiv:2504.03274.

Liu, Z.; Coleman, N.; Patrascu, F. I.; Yin, K.; Li, X.; and Mostafavi, A. 2024. Artificial intelligence for flood risk management: A comprehensive state-of-the-art review and future directions. *International Journal of Disaster Risk Reduction*, 117: 105110.

Lyu, H. 2025. Assessing Large Multimodal Models for Urban Floodwater Depth Estimation. *Water Resources Research*, 61(2): e2024WR039494.

Ma, J.; et al. 2025. Large-language-model-driven agents for fire evacuation simulation. *Safety Science*.

Martin, T.; Hofman, J. M.; Sharma, A.; Anderson, A.; and Watts, D. J. 2016. Exploring limits to prediction in complex social systems. *Proceedings of the 25th ACM International World Wide Web Conference (WWW)*. ArXiv:1602.01013.

Massaguer, D.; Balasubramanian, V.; and Mehrotra, S. 2006. Multi-agent simulation of disaster response. In *Proceedings of the ATDM Workshop at AAAI*.

Mitrovic, M.; Lukashevich, A.; Vorobev, P.; Terzija, V.; Budennyy, S.; Maximov, Y.; and Panov, A. I. 2023. Data-driven stochastic AC-OPF using Gaussian process regression. *International Journal of Electrical Power & Energy Systems*.

Okeukwu-Ogbonnaya, A.; et al. 2025. LLM-Based Community Surveys for Operational Decision Making in Interconnected Utility Infrastructures. ArXiv:2507.13577v1.

Owen, A. B.; Maximov, Y.; and Chertkov, M. 2019. Importance sampling the union of rare events with an application to power systems analysis. *Electronic Journal of Statistics*.

Piao, J. 2025. AgentSociety: Large-Scale Simulation of LLM-Driven Generative Agents Advances Understanding of Human Behaviors and Society. arXiv:2502.08691.

Piao, J.; et al. 2025. Large-Scale Simulation of LLM-Driven Generative Agents Advances. arXiv:2502.08691.

Shevchenko, V.; et al. 2024. Climate change impact on agricultural land suitability: An interpretable machine learning-based Eurasia case study. *IEEE Access*, 12: 15748–15763.

Taniushkina, D.; et al. 2024. Case study on climate change effects and food security in Southeast Asia. *Scientific Reports*, 14(1): 16150.

Törnberg, P.; and Larooij, M. 2025. Do Large Language Models Solve the Problems of Agent-Based Modeling? A Critical Review of Generative Social Simulations.

Xie, Y. 2025. MARSHA: multi-agent RAG system for hazard adaptation. *Nature*.

Xu, F. 2025. Large language model applications in disaster management. *International Journal of Disaster Risk Reduction*, 105642.

Xue, J.; Park, S.; Mondal, W. U.; Reia, S. M.; Yao, T.; and Ukkusuri, S. V. 2024. An agent-based model of post-disaster recovery in multilayer socio-physical networks. *Sustainable Cities and Society*, 115: 105863.

Zhang, Y.; et al. 2025. Beyond Static Responses: Multi-Agent LLM Systems as a New Paradigm for Social Science Research. arXiv:2506.01839.

Zhu, J. 2024. A flood knowledge-constrained large language model for flood analysis. *International Journal of Geographical Information Science*, 38(2): 297–318.