

Towards AI-Assisted Process-Based Modeling by Combining LLM-Based Reasoning with Executable Knowledge Graphs

Long Jiang^{1, 2, *}, Yang Yang³, Huanfeng Duan¹, Hoshin Vijai Gupta⁴

¹*State Key Laboratory of Climate Resilience for Coastal Cities, Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China*

²*Puget Sound Institute, University of Washington, Tacoma, Washington, USA*

³*School for the Environment, University of Massachusetts Boston, Boston, Massachusetts, USA*

⁴*Department of Hydrology and Atmospheric Sciences, The University of Arizona, Tucson, Arizona, USA*

*Email: jlong@connect.hku.hk jlon@uw.edu

ABSTRACT (199 words)

Process-based models encode decades of scientific knowledge, yet much remains fragmented across publications, documentation and model-specific code, limiting comparison, transfer and reuse. Here we introduce the Process-Model Knowledge Graph (PMKG), an executable knowledge graph that represents variables, parameters, equations, processes, models and hydrological systems as linked, machine-interpretable entities with explicit provenance, assumptions, applicability conditions and validation evidence.

Implemented through a suite of executable large language model (LLM) skills, PMKG converts source-backed process formulations into standardized specifications, verified micro-modules and versioned graph records, and assembles them through constraint-aware reasoning. The illustrative hydrological implementation, PMKG v1.0, contains 530 entities, 1,661 typed relationships and 89 contract-verified executable micro-modules.

Four analyses demonstrate reconstruction, adaptation, comparison and transfer: PMKG reconstructs two SUPERFLEX structures and reproduces their reference hydrographs in the Thur catchment; adapts model pathways as evidence and objectives change in the Río Grande catchment; isolates the effects of alternative routing formulations across 20 Swiss catchments; and tests transfer by applying declared profile-to-process rules consistently across 60 basins spanning multiple regions. PMKG thus bridges scientific process knowledge and operational model execution while preserving component semantics and provenance, providing a traceable foundation for model reconstruction, evidence-conditioned assembly, controlled process comparison and AI-assisted hydrological reasoning.