Darwin Gödel Machine Agentic AI Framework for Earth System Observation and Prediction Applications

Jakub F. Juranek

Institute of Psychology, Polish Academy of Sciences, Warsaw, Poland

✉ jakub.juranek@sd.psych.pan.pl

This work investigates the potential for the use of novel agentic AI framework, the Darwin Gödel Machine (DGM), for Earth System Observation and Prediction (ESOP) applications. DGM is grounded in the idea of a self-improving system capable of recursively rewriting and bettering itself and the predictive solution it produces. We explore DGM capabilities in the context of weather prediction, by applying the framework with the use of varied time frame and data sources (single, few and multi weather station networks). The proposed agent operates by iteratively refining its predictive solutions by both monitoring its past performance and the accuracy forecasts produced by its solutions, within predefined constraint parameters. It maintains a population of its "children" variants evolved through internal meta-learning. Beyond the use cases of archived weather data, we additionally test and report on the use case testing feasibility of applying the framework in an online, real-time, real-world data update mode, assessing potential for autonomous, online self-improving prediction agent engines for continuous deployment. The reported work and results provide a foundation upon which similar solutions can be developed for a broader range of ESOP challenges, extending well beyond the cases explored here.

1. Background: Iterative and Self-Referential Self-Improvement of Problem Solvers

In 2007, Schmidhuber (2007) proposed the concept of mathematically rigorous, general, fully self-referential, self-improving, and optimally efficient problem solvers inspired by Kurt Gödel's (1931) self-referential formulas. These types of problem solvers are theoretically able to rewrite their own code as soon as they find proof that a rewrite is useful, through the use of axioms encoded within an initial proof searcher. Building upon and further developing this concept—while addressing the computational bottleneck of requiring strict axiomatic proofs in complex environments and the practical impossibility of implementing the machine's original formulation due to the challenge of proving the impact of most self-modifications—Zhang and colleagues (2025) introduced the Darwin Gödel Machine (DGM). The DGM concept married the idea of Gödelian self-reference with inspirations from Darwinian evolution, utilizing the mechanics of meta-learning to iteratively evolve a population of self-improving "children" variants of itself. This framework, developed in the context of contemporary AI agentic systems, has proven effective in coding tasks, inspiring applications in other domains such as pharmacology (Darwin Gödel Drug Discovery Machine - DGDM; Wu et al., 2025) and finding further advancement in concepts such as the Huxley Gödel Machine (Wang et al., 2025).

2. The Darwin Gödel Weather Prediction Machine (DGWPM)

Building upon above mentioned self-improving frameworks we propose the Darwin Gödel Weather Prediction Machine (DGWPM), an artificial intelligence agentic system for the domain of earth system observation and prediction (ESOP). A system able to iteratively refine its weather forecasting capabilities without manual retraining, able to learn on archived data as well as potentially in an online, real-time environment, having also ability to improve its own code base.

3. Architectural Topology of the DGWPM: Nested Hierarchical Optimization Loops

The DGWPM builds upon the basic premises of DGM, drawing also inspirations from the DGDM (Wu et al., 2025). It introduces a tripartite nested loop architecture governed by the principle of least-cost intervention, ensuring the system remains "graceful on resources" while pursuing long-term metaproductivity.

- **The Minor Inner Loop (Parameter Refinement):** Executing at the highest frequency, this loop facilitates continuous optimization of active hyperparameters (η) and threshold refinements. It responds to immediate fluctuations in predictive accuracy, attempting to maintain stability without altering the model's underlying structure.
- **The Major Inner Loop (Evolutionary Adaptation):** Triggered only upon sustained stagnation ($\Delta < \epsilon$) or iterative budget exhaustion ($T_{\{max\}}$), this loop initiates the evolutionary process. The agent, acting as a "Mutable Mutator," generates a population of variant "Solutions" ($\{s_0\}$), exploring structural permutations and hyperparameter spaces to identify more robust predictive architectures.
- **The Gödelian Outer Loop (Recursive Self-Improvement):** This meta-level governs the agent's own codebase. It is invoked only when stagnation persists across both inner loops or when developmental efficiency T_g—the ratio of performance gain to resource consumption (e.g., time or tokens)—falls below optimal levels. This loop executes fundamental self-modifications, allowing the agent to refine its code.

4. Current Implementation

The DGWPM is currently instantiated as a Python-based agentic pipeline, evaluating its recursive self-improvement capabilities against a decade of hourly meteorological data (2016–2026). To test the framework's spatial scaling and data-integration limits, the agent's predictive targets—comprising localized temperature, pressure gradients, humidity, wind kinematics, and precipitation—are evaluated across three escalating network topologies in the Emilia-Romagna region. These range from a baseline Single-Node stream (Bologna), to a Sparse Network (a synoptic cross-pattern integrating Parma, Ferrara, Ravenna, and Florence), and culminating in a Dense Network (a dynamically generated 5x5 mesoscale grid).

Figure 1: Operational Algorithm of the Hierarchical DGWPM

Algorithm

Input: Initial agent g_0, initial solution s_0, dataset D, threshold ϵ, weights w, stagnation limit K, max iterations T
Output: Evolved pair (g^*, s^*), Archive $\mathcal{R}$

Initialize $\mathcal{R} \leftarrow \emptyset$, stag_count $\leftarrow 0$
for $t = 1$ **to** T **do**
  `// Step I: Inner Minor Loop`
  Execute s_t on $D \rightarrow \text{logs}_s$
  Calculate
   $\mathcal{T}_s = w_1 \cdot \text{Loss} + w_2 \cdot \text{Compute} + \lambda \cdot \text{MDL}$
  if $\Delta \mathcal{T}_s < \epsilon$ **then**
    stag_count $\leftarrow$ stag_count $+ 1$
  end
  else
    stag_count $\leftarrow 0$
  end
  `// Step II: Inner Major Loop`
  if $1 \leq$ *stag_count* $\leq K$ **then**
    $g_t \xrightarrow{\text{mutate}} s_{t+1}$ `// Architectural Mutation of s`
    $\mathcal{R} \leftarrow \mathcal{R} \cup \{(g_t, s_{t+1}, \text{type='Major'})\}$
  end
  `// Step III: Gödelian Outer Loop (Meta-Evolution)`
  if *stag_count* $> K$ **or** $\mathcal{T}_g > \theta$ **then**
    $g_t \xrightarrow{\text{self-mutate}} g_{t+1}$ `// Recursive Modification of Agent Logic`
    $\mathcal{R} \leftarrow \mathcal{R} \cup \{(g_{t+1}, s_t, \text{type='Gödelian'})\}$
    stag_count $\leftarrow 0$
  end
end
return $\textit{argmax}_{\mathcal{R}}(\textit{Combined Score})$

CONTACT

WWW.LINKEDIN.COM/IN/JAKUB-F-JURANEK

ORCID.ORG/0000-0003-2438-5235

WWW.RESEARCHGATE.NET/PROFILE/JAKUB-JURANEK

References:

Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. Monatshefte für Mathematik und Physik, 38(1), 173–198.

Schmidhuber, J. (2007). Gödel machines: Fully self-referential optimal universal self-improvers. In Artificial General Intelligence (pp. 199-226). Springer.

Wang, W., et al. (2025). Huxley-Gödel Machine: Human-Level Coding Agent Development by an Approximation of the Optimal Self-Improving Machine. arXiv preprint arXiv:2510.21614.

Wu, X., et al. (2025). The Darwin-Gödel Drug Discovery Machine (DGDM): A Self-Improving AI Framework. bioRxiv.Zhang, J., et al. (2025). Darwin Gödel Machine: Open-Ended Evolution of Self-Improving Agents. arXiv preprint arXiv:2505.22954.

Zhang, J., et al. (2025). Darwin Gödel Machine: Open-Ended Evolution of Self-Improving Agents. arXiv preprint arXiv:2505.22954.