A Global, Multi-Factor Geospatial Framework for Predicting Species Encounter Probability in Recreational Hunting and Foraging

Petr Salnikov
Independent Researcher
Email: pravda121312@gmail.com

Abstract

Traditional species distribution models (SDMs) are often computationally heavy and inaccessible to recreational users such as hunters and foragers. This paper presents a novel, precomputed geospatial framework (WildTracks) designed to calculate real-time species encounter probability. The system utilizes a 150 GB PostgreSQL/PostGIS database containing vectorized species ranges, ecoregions, Köppen climate zones, soil data, and elevation models. By precomputing spatial intersections using GiST indexes and centroid-based ST_Intersects queries, the framework reduces real-time query latency to milliseconds. The final probability is calculated using a multi-factor algorithm that weighs habitat suitability, seasonality, time of day, lunar phases, and terrain. Additionally, the system incorporates a safety module that identifies dangerous look-alike species (e.g., poisonous mushrooms) and hazardous taxa. This approach demonstrates how heavy geospatial data can be optimized for web-based, user-facing applications, providing a scalable model for citizen science and wildlife management.

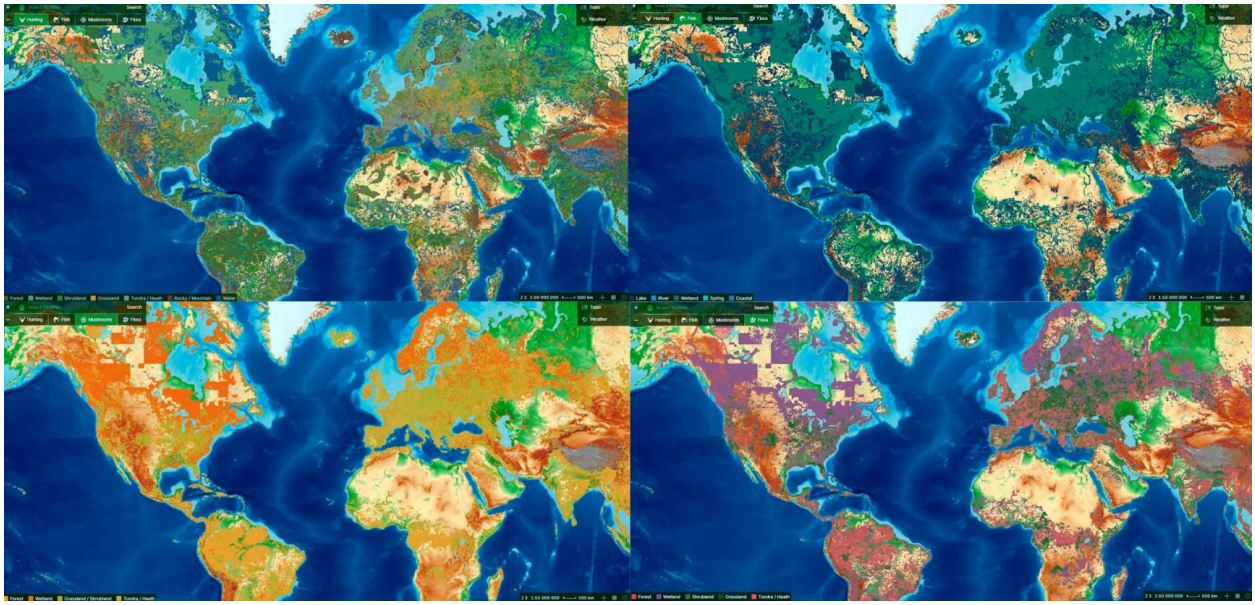


Figure 1. Global distribution of four primary activity layers in the WildTracks platform at zoom level 3.
(A) Hunting layer, showing base land cover (forests, grasslands, wetlands) used for terrestrial species prediction.
(B) Fishing layer, highlighting water bodies, rivers, and coastal zones.
(C) Mushroom picking layer, displaying forest and grassland habitats used for fungal distribution modeling.
(D) Flora/berry foraging layer, indicating vegetation and habitat types suitable for plant species.
Color-coding represents the underlying GIS data (e.g., OpenStreetMap land use) that serves as a foundation for the multi-factor probability algorithm.

1. Introduction

Predicting where a species is likely to be found is a fundamental task in ecology and wildlife management. While academic Species Distribution Models (SDMs) are highly accurate, they often require extensive computational resources and specialized software (e.g., R, MaxEnt), making them impractical for real-time web applications. Recreational hunters and foragers require instant, accessible predictions. This project addresses this gap by developing a scalable, open-source GIS framework that precomputes complex spatial intersections to deliver instantaneous, multi-factor probability scores based on user coordinates.

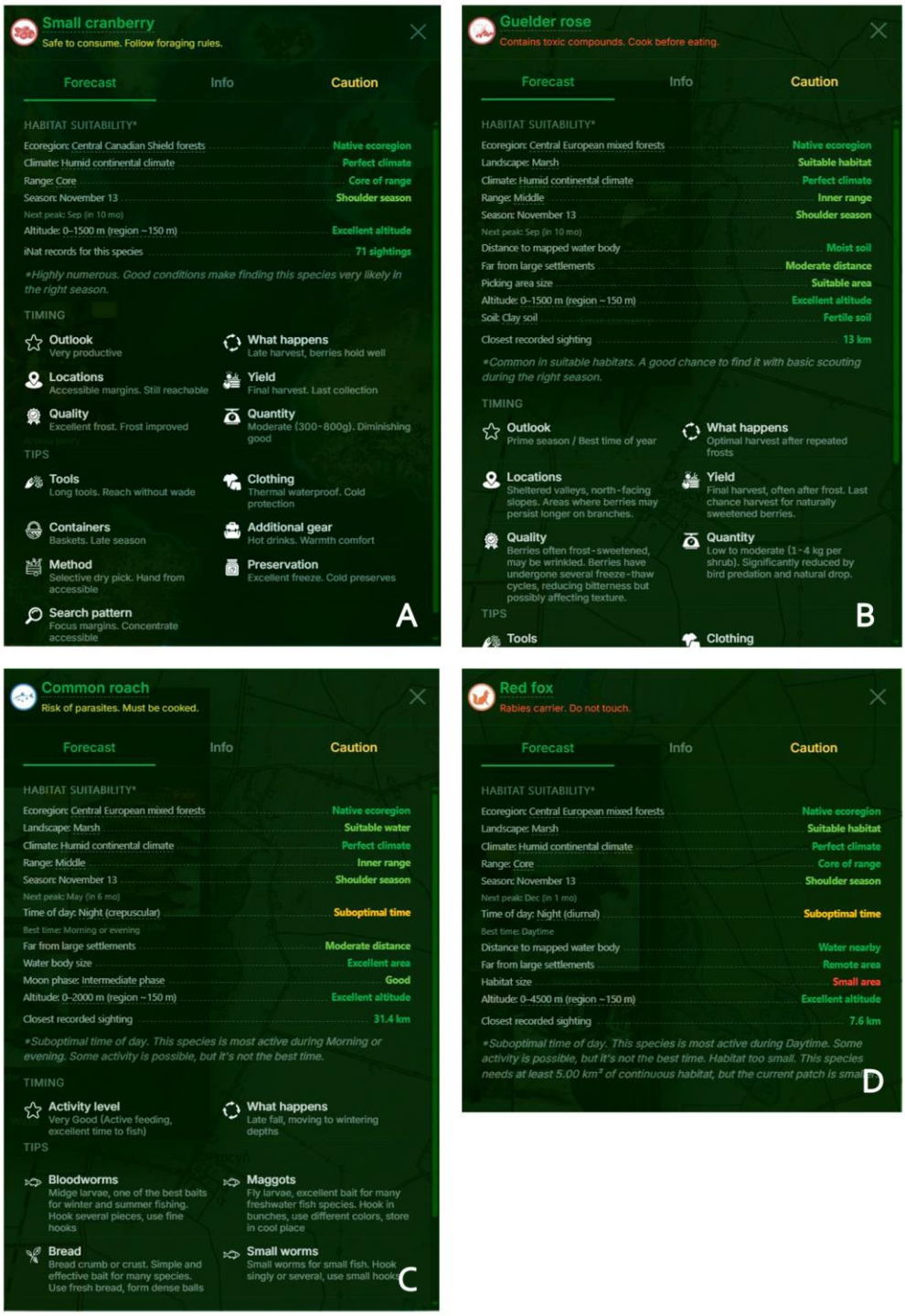


Figure 2. User interface examples of the WildTracks framework for different activity types and safety warnings. **(A)** Small cranberry (flora/foraging): shows habitat suitability, seasonality, and foraging tips. The system confirms the species is safe to consume. **(B)** Guelder rose (flora/foraging): provides a safety warning ("Contains toxic compounds. Cook before eating"), demonstrating the system's look-alike and toxicity detection. **(C)** Common roach (fishing): displays multi-factor fishing predictions, including optimal time of day (crepuscular), moon phase, and recommended bait. **(D)** Red fox (hunting): shows a safety warning regarding rabies ("Rabies carrier. Do not touch"), highlighting the "Danger" module for hunters. All panels provide explainable text outputs (e.g., "Suboptimal time," "Small area") alongside the probability score, ensuring transparency in algorithmic predictions.

2. Methods

2.1. Data Sources and Database Architecture

The system is built on PostgreSQL 15 with the PostGIS extension. We integrated vectorized species ranges from global biodiversity databases, WWF Ecoregions, Köppen-Geiger climate classification maps, OpenStreetMap (OSM) derived land cover data (forests, wetlands, water bodies, soil types), and elevation models.

2.2. Precomputation and Spatial Indexing

To avoid computationally expensive real-time spatial joins on a 150 GB dataset, we utilized a precomputation strategy. This involves centroid generation for all land cover polygons and bulk spatial intersection algorithms (ST_Intersects) using GiST indexes. The results (which species intersect with which habitat polygon) are stored in a materialized table (polygon_taxon_sets). This reduces the query time from minutes to milliseconds.

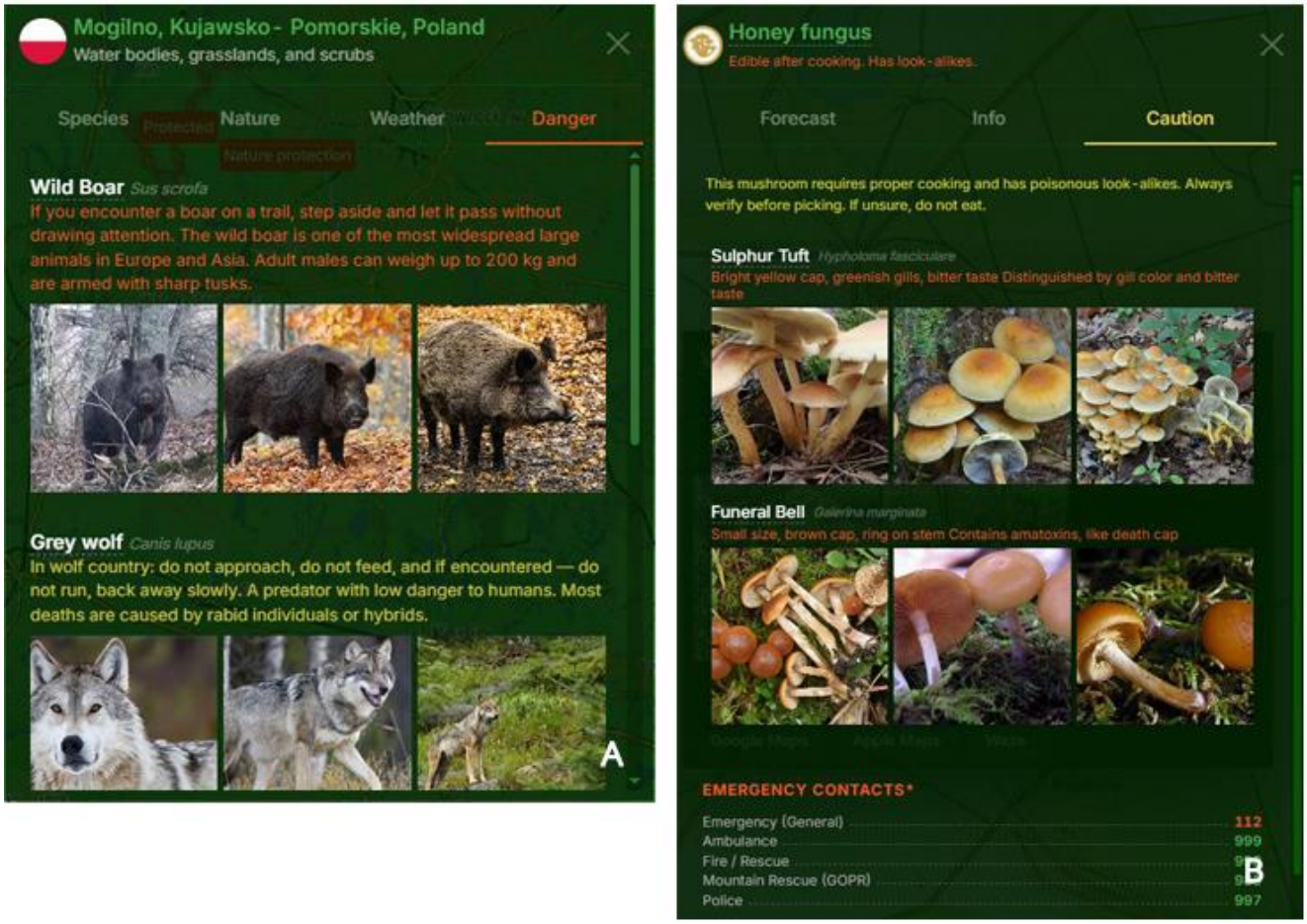


Figure 3. Safety and risk assessment modules in the WildTracks platform.
(A) The "Danger" tab for hunters: displays spatial warnings for potentially aggressive wildlife, such as the Wild Boar (Sus scrofa) and Grey Wolf (Canis lupus), including behavioral advice.
(B) The "Caution" tab for foragers: displays a safety warning for the Honey fungus (Armillaria mellea), including its poisonous look-alikes (Sulphur Tuft and Funeral Bell) and a direct link to emergency contacts (ambulance, police, mountain rescue) based on the user's current location.

2.3. Multi-Factor Probability Algorithm

The final probability score (P) is calculated as a weighted product of several factors:

- F_range: Position within the species range (Core, Edge, Outside).
- F_climate: Match between species preference and local Köppen zone.
- F_habitat: Overlap with specific land cover (e.g., forest, marsh, grassland).
- F_soil: Match with specific soil types (critical for fungi and flora).
- F_season: Seasonal activity coefficient (e.g., mating season, migration).
- F_time: Diurnal/nocturnal/crepuscular activity match.
- F_moon: Lunar phase (applicable only to specific fish species).
- F_elevation: Altitude suitability based on species limits.

2.4. Safety and Look-alike Species Module

A critical feature of the framework is the safety module. For foraging (mushrooms and berries), the system identifies "dangerous look-alikes" (e.g., poisonous mushrooms that resemble edible ones). Using the same spatial intersection logic, the system flags if a dangerous taxon shares the same habitat. Additionally, the interface provides a dedicated "Danger" tab for hunters, warning about aggressive or hazardous species (e.g., bears, snakes) whose range overlaps with the user's location.

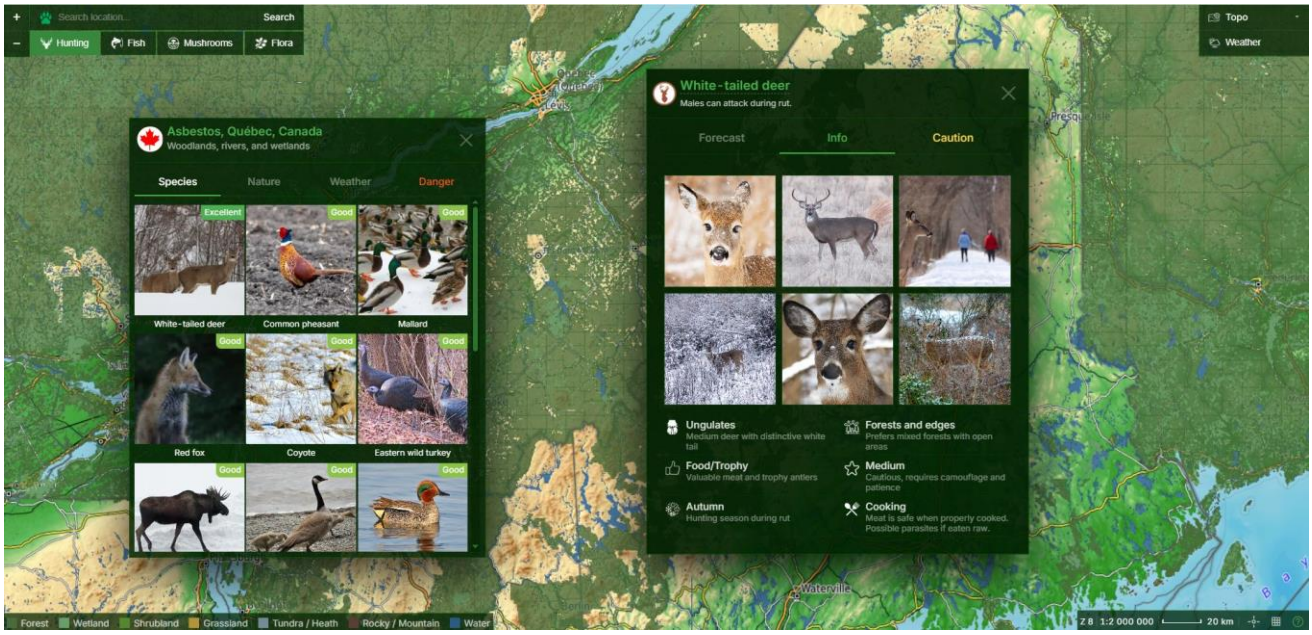


Figure 4. Regional species prediction and detailed information module for hunting in Quebec, Canada. The left panel lists predicted species (e.g., White-tailed deer, Wild turkey, Moose) with their suitability ratings. The right panel displays the "Info" tab for the White-tailed deer (Odocoileus virginianus), providing educational content, including behavioral warnings during the rut ("Males can attack during rut"), habitat preferences, and culinary safety notes ("Meat is safe when properly cooked").

3. Results

The framework was tested on a dataset of 5.8 million habitat polygons globally. Precomputation of spatial intersections for 235 species reduced the database query time from ~10 seconds per 500 records to ~1.5 seconds. The real-time API returns a probabilistic score and an explainable text output (e.g., "Poor altitude," "Suboptimal time of day") in under 100 milliseconds. The system successfully operates across both hemispheres, accounting for seasonal inversions. The user interface with a detailed breakdown of factors for the European perch is shown in Figure 2C.

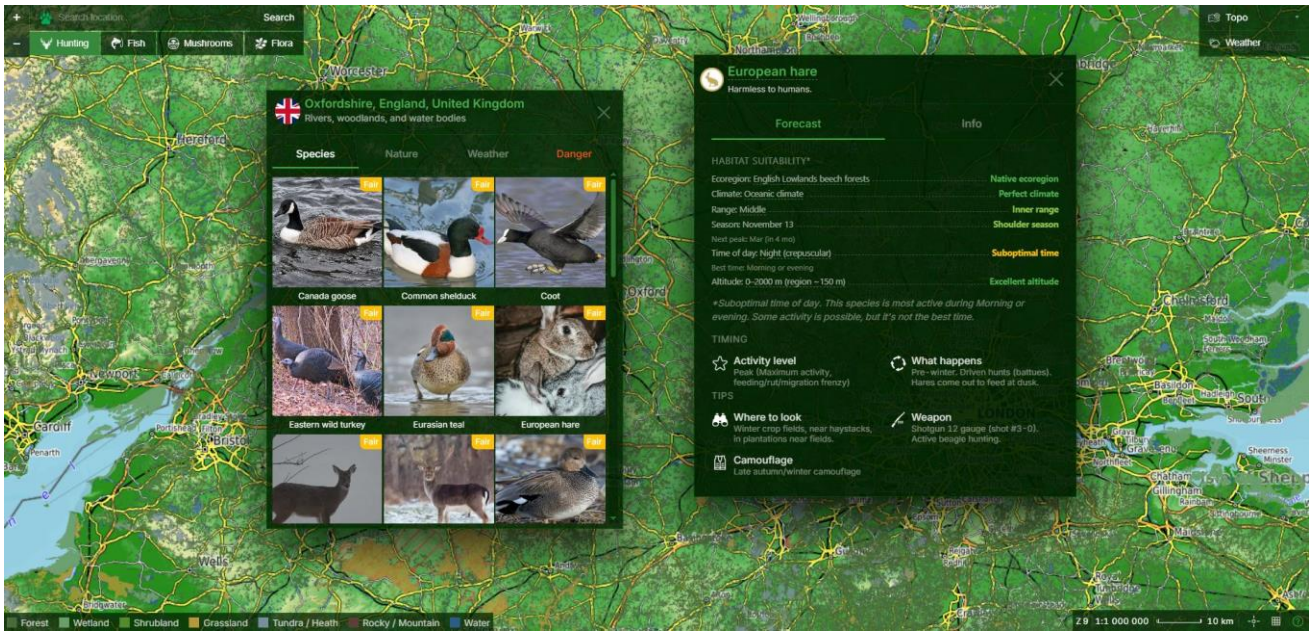


Figure 5. Example of a regional species prediction interface (Oxfordshire, UK) for the hunting activity. The left panel displays a list of species predicted to be present in the selected area, including the European hare (Lepus europaeus). The right panel shows the detailed multi-factor forecast for the European hare, highlighting a "Peak" activity level due to pre-winter driven hunts and an "Excellent altitude" suitability. The system provides practical hunting tips (where to look, weapon choice, camouflage) based on the current season and habitat.

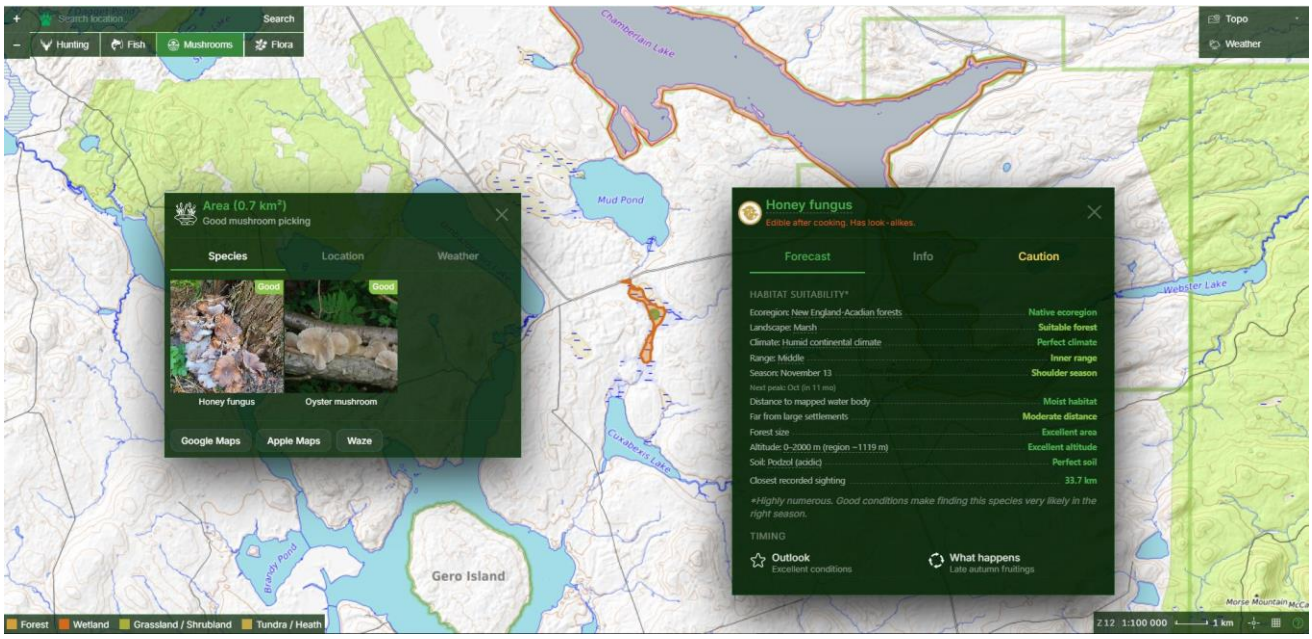


Figure 6. Local-scale habitat analysis for mushroom foraging (New England, USA). The left panel displays a specific habitat polygon (0.7 km²) classified as "Good mushroom picking," listing predicted species such as Honey fungus and Oyster mushroom. The right panel shows the detailed forecast for the Honey fungus (Armillaria mellea), including native ecoregion data, specific soil type (Podzol), altitude, and distance to water, demonstrating the system's ability to utilize fine-grained environmental variables for precise local predictions.

4. Discussion

This framework demonstrates that complex ecological modeling does not require heavy real-time computation. By shifting the computational load to precomputed spatial joins, we can deliver academic-grade predictions to end-users on low-end hardware. Limitations include the reliance on global polygon data, which may lack local precision. Future work will involve integrating real-time weather data and validating predictions against citizen science observations (iNaturalist).

References

Elith, J., et al. (2006). Novel methods for predicting species distributions from occurrence data. Ecography.

Hijmans, R. J., et al. (2005). Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology.

PostGIS Documentation: Spatial Queries and Indexing.