THE OUTDOOR ADVENTURE RISK INDEX: FORECASTING BACKCOUNTRY EVACUATION LIKELIHOOD FROM WEATHER AND VISITOR EXPOSURE

Matthew Sheahan[1]*, Peter Tsigaris[1], Anne Terwiel[1], and Yaou Hu[1]

[1] *Thompson Rivers University, BC, Canada*

ABSTRACT: Backcountry risk assessment requires practitioners and operators to interpret continually changing environmental conditions alongside uneven patterns of visitor exposure. Although professional judgment remains as a requirement, the likelihood component of risk assessment is often represented categorically rather than estimated from observed outcomes. The Outdoor Adventure Risk Index (OARI) was developed as a probability-based decision support model for estimating evacuation likelihood across monitored backcountry locations and hours.

OARI was developed using 511,328 location-hour observations from thirty recreation monitoring locations in Mount Revelstoke and Glacier National Parks, British Columbia, representing 771,212 recorded visitor passes between December 2016 and May 2024. Each location-hour integrates visitor exposure, weather, temporal information, location, and evacuation-relevant incident occurrence. Incidents were spatially and temporally linked to a monitored location before model fitting. An operationally constrained genetic algorithm selected ten forecast-compatible weather variables, which were incorporated into a binomial generalized linear model with a complementary log-log link. Visitor exposure entered the model through a log-exposure offset.

The final model achieved a precision-recall area under the curve of 0.010679, ROC-AUC of 0.787008, and Brier score of 0.002216 on held-out historical observations. More importantly for operational prioritization, the highest-ranked 1% of location-hours contained incidents at 10.0 times the rate expected under random ranking, while the highest-ranked 5% produced a lift of 5.5.

OARI is not an avalanche-hazard model and does not predict that an evacuation will occur within a particular time or place. It is a dynamic probability signal that is intended to aid in the identification of location-hours that warrant additional operational attention. The results support the feasibility of integrating environmental forecasts and visitor exposure into probabilistic decision support for backcountry safety management.

KEYWORDS: backcountry risk, evacuation likelihood, rare-event prediction, visitor exposure, decision support, search and rescue

1. INTRODUCTION

Backcountry risk management depends on decisions made under changing environmental conditions, uneven visitor exposure, and incomplete information. Professional judgement remains central, but one component of risk assessment is difficult to quantify consistently: the likelihood that an incident will occur under a particular combination of conditions.

Applied risk assessments often represent likelihood through categories such as unlikely, possible, or likely. These categories can organize professional judgement, but they are not necessarily empirical estimates of event probability. Risk-matrix research has shown that categorical systems can compress quantitatively different risks into the same rating and produce rankings influenced by category design (Cox, 2008; Duijm, 2015).

Exposure adds another source of variation. Similar environmental conditions may imply different incident likelihoods when visitor numbers differ substantially. Protected-area research has shown that recreation use varies across space and time, while search and rescue occurrence has also been associated with visitation patterns (Miller et al., 2017; Lu et al., 2021). Estimating incident likelihood therefore requires information on both environmental conditions and the number of people exposed to them.

Probabilistic methods have been used in wilderness search and rescue to estimate missing-person locations and spatial patterns of

* Corresponding author address:
Matthew Sheahan, Thompson Rivers University, Kamloops, BC, Canada, V2C 0C8;
tel: +1 613-894-0300;
email: matthewsheahan17@gmail.com

incident occurrence (Lin and Goodrich, 2010; Doherty et al., 2014). The Outdoor Adventure Risk Index (OARI) employs a similar predictive logic before an incident occurs. It estimates the probability that at least one evacuation-relevant incident will occur at a monitored recreation location during a specified hour.

OARI was developed using visitor-use, weather, location, and incident data from Mount Revelstoke and Glacier National Parks, British Columbia. Each observation represents one monitored location during one hour. The resulting probability is intended as an operational decision-support signal for comparing locations and forecast periods.

OARI does not estimate avalanche hazard and is not intended to replace avalanche forecasts or professional judgement. Its purpose is to identify location-hours where evacuation likelihood is elevated relative to background conditions. Performance is evaluated as a rare-event prediction problem, with emphasis on whether observed incidents are concentrated within the model's highest-ranked forecasts.

2. STUDY AREA AND DATA

2.1 Study Area

OARI was developed using data from Mount Revelstoke and Glacier National Parks in British Columbia, Canada, with monitoring locations distributed through the Revelstoke-Rogers Pass-Golden corridor. The area combines high-use mountain recreation with substantial spatial and seasonal variation in weather, access, and visitor activity.

2.2 Observation Unit and Data

The analysis uses a location-hour as its observation unit: one monitored recreation location during one hourly interval. This resolution allows visitor counts, weather observations, location information, and incident occurrence to be aligned within the same spatial and temporal structure.

The dataset contains 511,328 location-hours across 30 monitored locations, representing 771,212 recorded visitor passes between December 2016 and May 2024. Visitor passage was measured using automated trail counters and treated as an exposure measure rather than a count of unique recreationists.

Four data sources were combined:

- hourly visitor counts from trail counters;
- timestamped and geocoded Parks Canada incident records;
- hourly weather data assigned to each monitoring location; and
- monitoring-location coordinates and identifiers.

After spatial and temporal matching, each eligible location-hour was assigned a binary outcome indicating whether at least one evacuation-relevant incident was associated with that observation.

3. METHODS

3.1 Incident Matching

Parks Canada incident records were linked to the location-hour dataset using their timestamp and geographic coordinates. Duplicate incident records were removed before matching. Each incident was assigned to the nearest monitoring location when that location was within 10 km.

Incidents were then matched to the most recent eligible hour with recorded visitor activity, using a maximum 36-hour lookback. This window allowed for differences between visitor passage, incident occurrence, reporting, and recorded response time. Multiple incidents assigned to the same location-hour were collapsed into a single positive observation because the model estimates the probability of at least one incident.

3.2 Model Development

Weather predictors were selected using a genetic algorithm (GA). Candidate variables were restricted to weather fields available in both the historical dataset and operational forecasts. Candidate feature sets were evaluated using the same model structure later used in production, with precision-recall area under the curve (PR-AUC) as the primary selection criterion. Models producing unstable or saturated forecast probabilities were excluded from promotion.

The production model is a binomial generalized linear model (GLM) with a complementary log-log link:

$$\log\left[-\log\left(1-p_{s,t}\right)\right] = \alpha + X_{s,t}\beta + \log\left(E_{s,t}\right)(1) \quad (1)$$

where $p_{s,t}$ is the probability of at least one incident at location s during hour t, $X_{s,t}$ contains the selected weather variables, hour of day, weekend status, and location effects, and $E_{s,t}$ represents visitor exposure. Exposure enters as a log offset, allowing estimated incident probability to vary with the number of people exposed without treating visitor count as an ordinary predictor.

Operational forecasts replace observed exposure and historical weather with expected

visitor exposure and forecast weather while retaining the fitted model structure.

3.3 Validation

Model performance was evaluated using a sensor-wise chronological holdout. Observations were ordered by time within each monitoring location; the first 80% were used for model fitting and the final 20% were reserved for testing.

The held-out set contained 18,004 location-hours, including 40 incident-positive observations, giving an event prevalence of 0.222%. Evaluation therefore emphasized rare-event performance rather than overall classification accuracy. Precision-recall analysis is particularly informative when the positive class is rare (Saito and Rehmsmeier, 2015).

PR-AUC measured the concentration of incident-positive observations among higher model scores, ROC-AUC assessed ranking discrimination, and Brier score measured probabilistic error. Operational performance was also evaluated using top-k lift, which compares incident capture within the highest-ranked proportion of location-hours with the capture expected under random ranking.

4. RESULTS

4.1 Model Performance

The held-out test set contained 18,004 location-hours, including 40 incident-positive observations (0.222%). The final production model achieved a PR-AUC of 0.010679, ROC-AUC of 0.787008, and Brier score of 0.002216.

Ranking performance was strongest among the highest-scored location-hours. Top-1% and top-2% lift were both 10.0×, declining to 5.5× at the top 5% and 4.25× at the top 10%. The model therefore concentrated incident-positive observations within a relatively small portion of the ranked test set.

4.2 Selected Predictors

The constrained GA selected ten weather variables: air temperature, mean sea-level pressure, total and mid-level cloud cover, direct radiation, direct normal irradiance, wind direction at 10 and 100 m, weather code, and deep soil temperature. These were retained alongside hour of day, weekend status, location effects, and the visitor-exposure offset.

Coefficient diagnostics identified the strongest non-location signals for deep soil temperature, direct normal irradiance, and direct radiation. These coefficients describe associations within the fitted predictive model rather than causal effects. Should the model receive new training data, the model would undergo a new GA feature selection process.

5. DISCUSSION

5.1 Interpretation

OARI performed most clearly as a ranking model. The 10.0× lift within the highest-ranked 1% of location-hours shows that incident-positive observations were concentrated well above the random baseline. The comparatively low PR-AUC should be read against the rarity of the outcome rather than as a conventional classification score.

The model is therefore more useful for identifying periods of elevated likelihood than for determining whether an evacuation will occur. Forecast-stability requirements also prevented feature sets with strong historical performance but unstable operational probabilities from being promoted.

5.2 Operational Balance

A location-hour forecast allows many locations and time periods to be compared using a consistent probability estimate. In practice, higher-ranked periods could inform monitoring, visitor communication, staffing, or readiness while leaving the final decision with operational personnel.

Exposure is an important part of this interpretation. Similar environmental conditions do not necessarily imply similar evacuation likelihood when visitor use differs. The premise regarding the incorporation of expected visitor volume allows OARI to account for this variation rather than treating environmental conditions alone as the source of incident likelihood.

5.3 Limitations and Future Development

The main limitation is the small number of incident-positive observations. Performance estimates are therefore sensitive to relatively few events and should not be treated as evidence of equivalent performance across other regions or future seasons.

The predictor set is also incomplete. In-depth terrain features, snowpack, surface conditions, route exposure, visitor experience, and real-time visitor use are not represented in the current production model. Forecast performance also depends on the quality of weather and exposure inputs. It can be assumed that these variables would offer more meaningful directions for model

development than simply increasing the number of weather predictors.

6. CONCLUSION

OARI demonstrates that backcountry evacuation likelihood can be estimated at an hourly, location-specific scale using environmental conditions and visitor exposure. The final model concentrated observed incidents within its highest-ranked predictions, including a 10.0× lift in the top 1% of held-out location-hours.

The model is intended to support, rather than replace, operational judgement. Its main contribution is a reproducible way to convert changing weather and exposure conditions into comparable estimates of evacuation likelihood. Further development should focus on richer exposure, terrain, snowpack, and surface-condition data and on validation across additional seasons and locations.

REFERENCES

Cox, L. A. T., 2008. What's wrong with risk matrices? *Risk Analysis*, 28(2), 497–512. https://doi.org/10.1111/j.1539-6924.2008.01030.x.

Doherty, P. J., Guo, Q., Li, W., and Doke, J., 2014. Space-time analyses for forecasting future incident occurrence: A case study from Yosemite National Park using the presence and background learning algorithm. *International Journal of Geographical Information Science*, 28(5), 910–927. https://doi.org/10.1080/13658816.2014.890202.

Duijm, N. J., 2015. Recommendations on the use and design of risk matrices. *Safety Science*, 76, 21–31. https://doi.org/10.1016/j.ssci.2015.02.014.

Lin, L. and Goodrich, M. A., 2010. A Bayesian approach to modeling lost person behaviors based on terrain features in Wilderness Search and Rescue. *Computational and Mathematical Organization Theory*, 16, 300–323. https://doi.org/10.1007/s10588-010-9066-2.

Lu, Z. N., Briggs, A., Saadat, S., and Algaze, I. M., 2021. The associations between visitation, social media use, and search and rescue in United States National Parks. *Wilderness & Environmental Medicine*, 33(1), 77–84. https://doi.org/10.1016/j.wem.2021.08.005.

Miller, A. B., Leung, Y.-F., and Kays, R., 2017. Coupling visitor and wildlife monitoring in protected areas using camera traps. *Journal of Outdoor Recreation and Tourism*, 17, 44–53. https://doi.org/10.1016/j.jort.2016.09.007.

Saito, T. and Rehmsmeier, M., 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3),e0118432. https://doi.org/10.1371/journal.pone.0118432.