CZECH POLAR REPORTS 15 (1): 128-145, 2025

Developing explainable AI models for predicting snowpack variability in Polar regions

Sidra Sidra*, Shah Mehmood Wagan

Sichuan University, Chengdu, Sichuan, China

Abstract

Accurate prediction of variability of polar snowpack is central to understanding global climate change and its impacts on ecosystems, sea-level rise, and weather patterns. Traditional, physically based snowpack models commonly fail to capture full range of complexity that realistic snow dynamics can display, especially under extreme climate events. On the contrary, machine learning (ML) models can enhance predictive accuracy but lack interpretability and thus are challenging to apply in scientific contexts. This research proposes construction of explainable AI models for prediction of polar snowpack, emphasizing how to interpret domain knowledge coupled with state-of-the-art techniques in AI. Physics-informed neural networks are used herein to embed physical laws, Graph Neural Networks (GNNs) to represent spatial dependencies, and Recurent Neural Network (RNNs) for temporal sequences. Several post-hoc explanation techniques, such as SHAP and saliency maps, together with the development of causal inference models, are applied in such a way that the transparency and scientific basis of the models are preserved.

Key words: snowpack, variability, polar regions, explainable AI, neural networks, climate change

DOI: 10.5817/CPR2025-1-8

Introduction

Snowpack variability predictions are crucial in polar regions as understanding such conditions ultimately helps in several global climate patterns, further affecting sea-level rise, ecosystems, and human activities. Polar areas, which include regions like Antarctica and Greenland, involve a very complex interplay of climatic variables, including temperature, wind patterns, precipitation, and solar radiation (Pletzer et al. 2024, Roy et al. 2018). These factors affect not only the volume of snow mass but also its physical properties, further influencing energy balance, albedo, and regional water cycles (Ru and Ren 2024). Accurate modeling and high-resolution prediction of the snowpack variability is urgent for climate policy and preparation regarding extreme weather events driven by changes in polar environments (Singh et

Received April 8, 2025, *accepted* August 14, 2025.

*Corresponding author: S. Sidra <sidra_scu@outlook.com>

al. 2024, Sobota et al. 2020). Traditional models are usually based on physical simulation but can rarely capture the full complexity of snow dynamics, given the very high degree of spatial and temporal variability associated with polar regions (Spolaor et al. 2018, 2021). Machine learning (ML) models have revealed huge promise in improving predictive capabilities, yet lack transparency that prevents climate scientists from interpreting or trusting their output (Spolaor et al. 2014). This gap underlines the need for explainable AI models, which balance predictive accuracy with interpretability (Thapa et al. 2024). Explainable AI models enable users-like scientists and policy makers-to understand not only what the model foretells but also why it does so in a particular respect (Stroeve et al. 2024).

Literature Review

Traditionally, snowpack modeling has been controlled by physical models that simulate snow dynamics built on known principles of mass and energy balance (Arp et al. 2018). These include the Snow Water Equivalent (SWE) model, that evaluates the accumulation and density of snow, and the Crocus model, that simulates snowpack evolution by explaining complex thermodynamic equations (Blau et al. 2021). These models, while built on well-known physical laws, very frequently fail to duplicate the complete variability of the snowpack dynamics and especially its response to extreme climatic events-sudden climatic changes triggered in a time range of hours to several days including a sudden rise in temperature or a burst of heavy precipitation (Bonsoms et al. 2024). For the same reasons, the computational complication of these models, maintained as such, sets yet another significant limit to their scalability, particularly when lengthy to large spatial domains similar to polar regions (Celli et al. 2023).

More recently, methods of ML have been tried for the prediction of snowpack. Works investigating deep learning models for snowpack forecasting, such as those by Collow et al. (2022) and/or Davesne et al. (2022), have shown that neural networks are capable of learning complex patterns in large datasets to improve predictive accuracy (Valt and Salvatori 2016, van Dijk et al. 2014). Most of these models are "black boxes" like physics-based models that offer little insight into the relationships between input variables and outputs (Docherty et al. 2018). In particular, temperature and wind are input variables, while snow depth and melt rates are output variables (Ebrahimi and Marshall 2016). This lack of interpretability poses a big challenge in fields like climate science, where the processes responsible for driving variability are as important to understand as are the predictions themselves (Ernakovich et al. 2014, Owczarek and Opala 2016).

It was for such challenges that a field now identified as Explainable AI came into being (Varty et al. 2021, Wahl et al. 2024). Explainable AI-developed methods that makes ML algorithms more interpretable (Weinhart et al. 2021). Some feature attribution methods applied to provide insight into model decision-making are SHapley Additive exPlanations, or SHAP, and Local Interpretable Model-agnostic Explanations, LIME (Thapa et al. 2024). A work by Homan and Kane (2015), showed how the mentioned techniques help in enhancing the interpretability for the deep learning models (Irannezhad et al. 2016). However, XAI applications for snowpack modeling are still in their early days and have so far been explored only by a few studies, focusing on how ML is linked to physical laws controlling snow dynamics (Jennings and Jones 2015).

The most promising avenue seems to be that using Physics-Informed Neural Networks (PINNs), which encode physical laws as part of the learning itself (Li et al. 2018) developed PINNs, which constrain neural networks with known physical constraints that guarantee the predictions to be consistent with established scientific theories (Li et al. 2020).

Research Methodology

This section portrays the research methodology for the development of explainable AI models that predicts snowpack variability within polar regions. This is focused on model architecture and model interpretation techniques. For that reason, this research methodology has been designed to ensure scientific rigor, accuracy, and completeness so that stakeholders may have faith and understand the predictions made by the AI.

Model Development

Identification of Core Variables and Features

The first step in the development of AI models for the prediction of snowpack variability is the identification of those physical and environmental variables that influence the dynamics of snowpack. Key variables that control processes of accumulation and snow melting were identified (*see* the list below in Table 1). These typically include those listed below.

Variable	Impact on Snowpack
Temperature	Increased temperature enhances the rate of snowmelt, which lowers the depth of the snowpack.
Wind Patterns	Wind redistributes the snow by drifting and blowing; this redevelops the accumulation and ablation rates.
Albedo	The albedo, a factor impacting solar radiation amount absorbed by snow, evaluates the rate at where snowmelt happens.
Precipitation	Directly impacts snowfall and thus the amount of accumulation.
Atmospheric Conditions	If atmospheric pressure, humidity, and storms differ, then accumulation and melting change.

Table 1. Core variables and features.

Once these variables have been determined through the literature review as seen in Table 1, domain expertise of climate scientists can be included in order to validate that the choice of variables matches known climatic processes and snowpack dynamics. Without the input of such an expert, the complexity introduced into the problem by the behavior of the snowpack-especially at polar latitudes where local influences may be highly relevant, the modeling cannot be done correctly.

Snowpack Prediction Model Selection

We do have a number of ML and deep learning models that are used to predict snowpack variability. Since our focus is on developing explainable models, we identify those showing scientific grounds for interpretability.

Physics-Informed Neural Networks (PINNs)

Physics-Informed Neural Networks incorporate physical laws into the learning process directly. In the context of snowpack modeling, this translates to embedding the basic principles of the conservation of snow mass and energy fluxes into the architecture of the network as displayed in Figs. 1 and 2. These sets of governing equations for the conservation in snow mass and energy can be represented as:

$$\frac{\partial S}{\partial t} = P - M + E \qquad \textit{Eqn. 1}$$

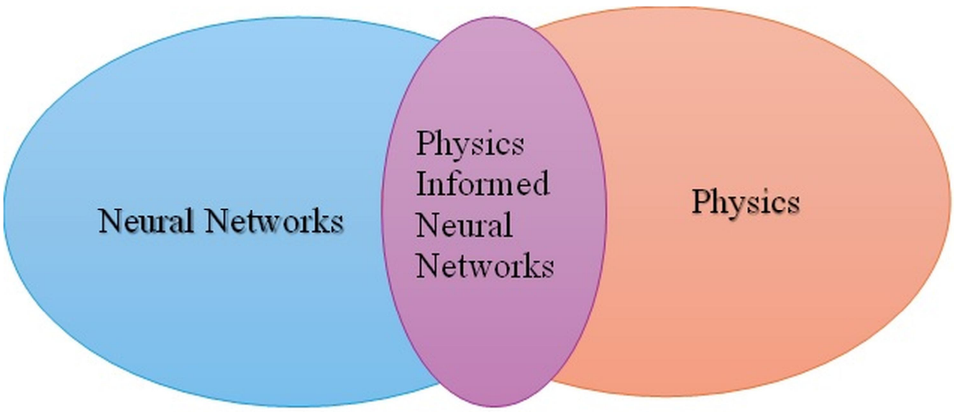

Fig. 1. Physics-informed neural network combination of nueral networks & physics.

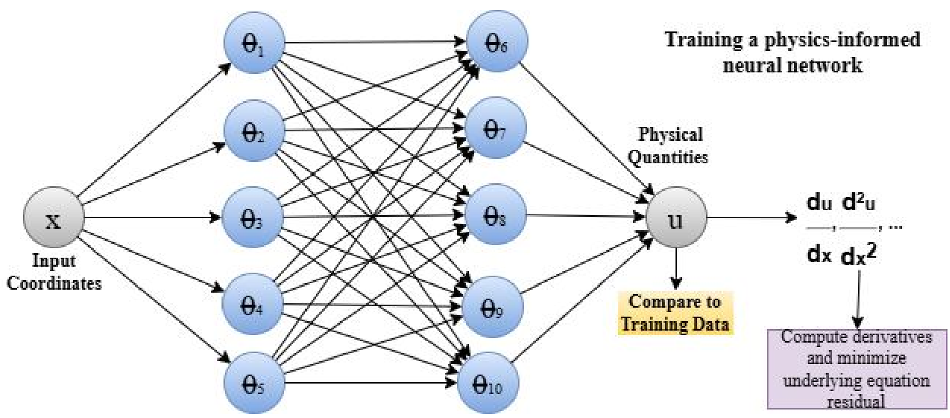

Fig. 2. Schematic of physics-informed neural network.

In equation 1 where S is a snowpack storage, P is precipitation (snowfall), M is snowmelt, dependent on temperature and radiation, E shows evaporation and sublimation losses. While training, PINNs force the model by hard-wiring these physical equations into the model architecture to ensure that the model behaves as expected from known scientific laws. The result of this is better generalization, especially in regions that might be poorly captured by the data.

Graph Neural Networks

With spatially dependent snowpack across polar regions, Graph Neural Networks (GNNs) can be applied to learn relational information from one geographical region to another. Many a time, snow variability can be very different depending on other surrounding regions due to common climatic patterns the regions share. This kind of structure in data is what GNNs are designed for; hence, it is eminently suitable for modeling spatial correlations in snow variability. A typical GNN may be represented as in Fig. 3.

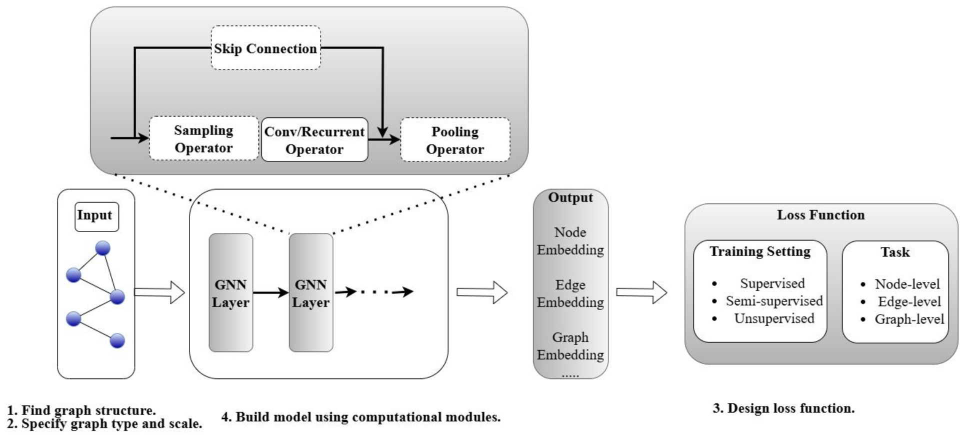

Fig. 3. GNN model design.

$$h_v^{(k+1)} = \sigma\left(W^{(k)} \cdot \sum_{u \in N(v)} h_u^{(k)} + b^{(k)}\right) \qquad \textit{Eqn. 2}$$

In equation 2 where $h_u^{(k)}$ is the hidden state of node v representing a geographic location at layer k, N(v) is the set of neighboring nodes connected to node v, $W^{(k)}$ and $b^{(k)}$ are the weight matrix and bias at layer k, σ is an activation function such as ReLU. GNNs can, therefore, represent how the snowpack at any given point in the polar region is linked to neighbouring regions, affording a more holistic and connected forecast snowpack variability in polar regions.

Recurrent Neural Networks (RNNs) / Long Short-Term Memory (LSTM) Networks

These dynamics are not only spatial but also temporal. For capturing this time-dependent nature, models based on RNNs or LSTMs are appropriate as displayed in Figs. 4 and 5. These models have been designed for processing sequential data, and, therefore, are perfectly adapted for modeling snowpack variability in time conditioned on prior snow measurements and climate variables. The LSTM cell is controlled by the following set of equations:

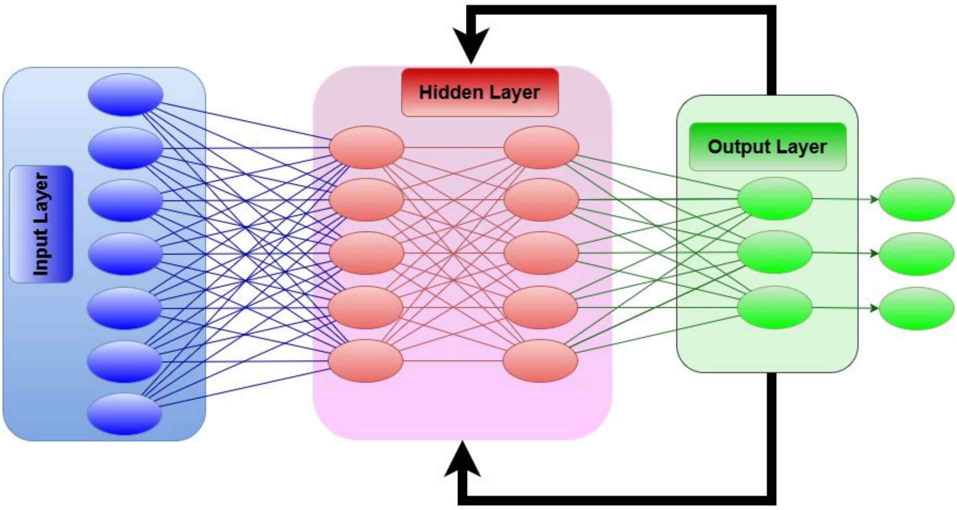

Fig. 4. Recurrent Neural Networks (RNNs) model design.

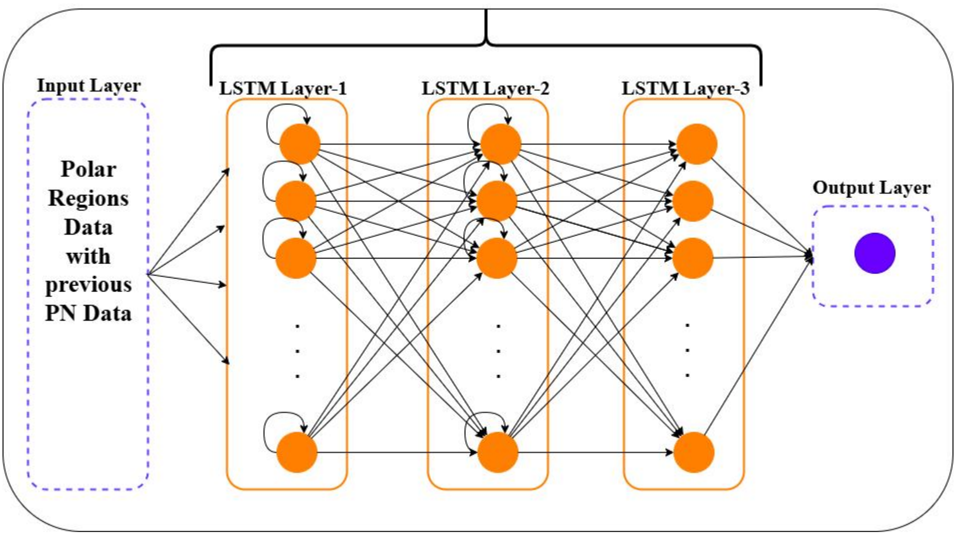

Fig. 5. Long Short-Term Memory (LSTM) Networks model design.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \qquad \textit{Eqn. 3}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \qquad \textit{Eqn. 4}$$

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \qquad \textit{Eqn. 5}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \qquad \textit{Eqn. 6}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \qquad \textit{Eqn. 7}$$

$$h_t = o_t * \tanh(C_t) \qquad \textit{Eqn. 8}$$

In the equations 3, 4, 5, 6, 7 and 8 where f_t, i_t, and o_t represent the forget, input, and output gates, denote the forget, input, and output gates, C_t is the cell state, while h_t is the hidden state at time t. This would enable the AI to memorize crucial temporal patterns in snowpack, including seasonable accumulation and melt cycles.

Explainability-first model design

While developing and training our AI models, the explainability needs to be one of the high-priority methods intrinsic to interpretability. These are decision trees and rule-based models where decision paths are easy to trace hence providing better explainability for the reason behind the model making a particular prediction. Hybrid models are applied when more complicated models are needed. In these models, simple and interpretable components (for instance, linear regression) are combined with more complicated ones that involve deep learning architecture (Thapa et al. 2024). In particular, in cases where there is a need to explain the relationship between temperature and snowmelt, a linear model can do that, while spatial dependencies are handled by more complex structures involving GNNs.

Explainability Methods

Post-Hoc Explanation Techniques

Post-hoc techniques provide explanations of predictions after the model is trained. These are SHAP and Lime. SHAP (SHapley Additive exPlanations) offers an additive feature attribution method using cooperative game theory. It assigns a "Shapley value" to each feature, which indicates its contribution to the prediction in Fig. 6. For a feature x_i, the Shapley value is calculated as in equation 9:

$$\phi_i(f) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \qquad \textit{Eqn. 9}$$

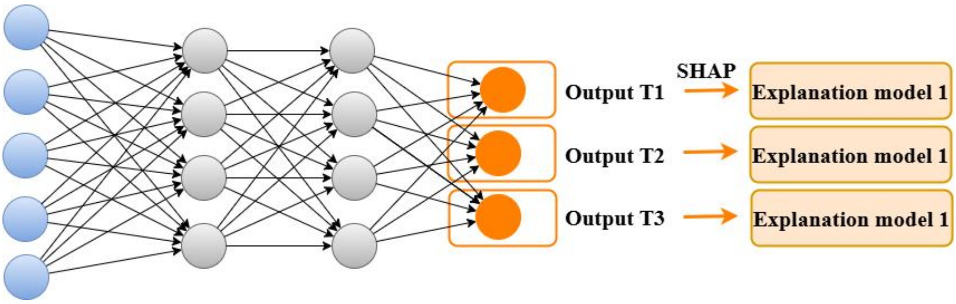

Fig. 6. SHAP (SHapley Additive exPlanations) design.

By doing this, it is well known what contribution every variable makes. In particular, temperature or wind gives to the model in making such a prediction about the snowpack. LIME (this means a Local Interpretable Model-agnostic Explanations) generate a simpler local model around a particular prediction. As a result, it is made rather easy to understand which features were important for that particular prediction. Saliency maps show which regions of the input data were most influential to the model's predictions for CNN/RNN and highlight spatial and temporal hotspots in snowpack variability.

Intrinsic Explainability Approaches

Intrinsic explainability: design models to be transparent by nature. Attention mechanisms in RNNs or transformer models allow a model to "attend" to the most informative parts of the input sequence, hence providing insight into which variables-or points in time-were most relevant for a prediction. The obtained attention scores can then be visualized, pointing out critical moments, such as critical temperature shifts. Monotonic neural networks ensure that known physical relationships, such as the inverse relationship between temperature and snowpack, are respected within the model. This ensures that valid scientific outputs outside the model is interpretable from the outset.

Causal Inference Models

Causal inference approaches, including SEMs and Bayesian networks, explicitly model cause-and-effect relationships among variables. By pointing out exactly how changes in one variable, say temperature, directly influence another variable, say snowpack, these models yield more insightful analysis than typical statistical analyses. CIMs have the capability of revealing the direct effects of cause-effect (*e.g.*, temperature affecting snowmelt) in a more insightful way as compared to analyses based on correlations. Causal inference models yield a more intuitive and transparent characterization of the climate processes driving snowpack variability. Given the probability for snowpack depleting in a Bayesian network in equation 10:

$$P(Snowpack = low \mid Temperature = high, Precipitation = low) \qquad \text{Eqn. 10}$$

The latter allows for more actionable insights, pointing not just at the correlations but also the causal pathways driving the changes in the snowpack.

Model Validation and Testing Approaches without Datasets

The goal of this section is to validate the performance, robustness, and scientific soundness of explainable AI models at predicting snowpack variability in polar regions without using any real-world datasets. This section introduces some theoretical frameworks, synthetic environments, and stress-testing methodologies that could be used for the estimation of model performance under different conditions for predicting snowpack changes.

Theoretical and Synthetic Validation

Theoretical validation becomes critical for AI models. The primary technique of validation is the development of synthetic environments based on established physical models of snow dynamics. These physical models, such as energy balance models, provide a first-principles-based simulation of accumulation, compaction, and melting processes in snow. The general equation of energy balance in the snowpack can be written as:

$$Q_{net} = Q_{SW} - Q_{LW} + Q_H + Q_L + Q_G \qquad \textit{Eqn. 11}$$

In equation 11, Q_{net} is the net energy flux affecting the snowpack, Q_{SW} is the incoming shortwave radiation, Q_{LW} is the outgoing longwave radiation, Q_H is the sensible heat flux, Q_L is the latent heat flux, Q_G is the ground heat flux. The aforementioned energy fluxes are important in determining snowmelt rates. The AI model should be trained and validated by the generation of synthetic data in order to produce predictions consistent with the physical laws governing snow dynamics. For instance, the AI model prediction of snowpack depletion subsequent to a temperature rise can be checked against synthetic snowmelt curves generated by an energy balance model for similar conditions. The good match between the AI model's results and the synthetic model serves as a good indicator of accuracy and reliability.

Other theoretical validation may come from climate model simulations developed from GCMs or regional snow models as shown in Table 2.

Validation Technique	Description	Application
Energy Balance Models	Employ physically based energy flux principles to simulate snow accumulation and ablation.	Predict snowmelt based on rising temperatures and validate it with snowmelt curves generated using energy balance equations.
Global Climate Models (GCMs)	Use global climate models to drive a generation of synthetic scenarios of snowpack with various climate forcings.	Compare AI predictions with GCM simulations of depletion in snowpack due to increased CO_2 levels or El Niño events.
Regional Snow Models	Focus on the localized behavior of synthetic snowpack in regional climatic patterns.	Compare the predictions of snowpack changes given by AI in specific polar regions to regional snow models, which simulate snowfall and ablation using prevailing conditions.

Table 2. Validation technique and description.

These models simulate snowpack variability in various climate scenarios, including historical and future projections. It may also include SWE predictions by the GCMs, for instance, using some climate forcing parameters like CO_2 concentrations and temperature increases.

Model Robustness Testing

Model robustness is one of the important paradigms in the validation process of an AI model, especially in such extremist events as sudden snowmelt or rapid accumulation. In the absence of real-world data, adversarial testing can be performed to appraise the model regarding its robustness under extreme and unanticipated situations. Perturbation of input variables beyond their typical range-for instance, temperature, wind speed, or precipitation-and sees how the model responds.

This could include testing the model for its response under a hypothetical scenario of a sudden extreme temperature spike during an Antarctic winter-such as might happen in a warming climate as seen in Table 3. The AI model's response, in terms of predicting sudden snowmelt, can then be evaluated both for accuracy and interpretability. It follows that this kind of stress-testing research ensures the model is capable not only of operating under average conditions but also of being robust at extremes. Under these perturbed conditions, the model predictions should be scientifically consistent with the known physical laws governing the dynamics of snow.

$$R_{ext} = \frac{1}{n}\sum_{i=1}^{n}\left(P_{AI}(x_i) - P_{theory}(x_i)\right)^2 \qquad \text{Eqn. 12}$$

In equation 12, P_{AI} (x_i) is the AI model prediction under the perturbed input x_i, P_{theory} (x_i) is the theoretical model prediction under the same perturbed input, and n is the total number of test scenarios. Moreover, this ensures that the model has fidelity and robustness under extreme, rare, or novel climatic conditions.

Test Scenario	Description	Application
Sudden Snowmelt	Adversarial testing of the rapid snowmelt prediction under extreme increases in temperature in the case of winter.	Run a scenario with a temperature rise of +10°C during polar night to explore whether the model predicts that snow is lost as would be dictated by physical principles.
Extreme Precipitation Events	Stress-testing under hypothetical heavy precipitation periods to evaluate model responses.	Apply a +50% increase in snowfall over one month and see if the model appropriately predicts large accumulation of snow without snow density limits being violated.
Prolonged Dry and Warm Periods	Testing of model response to longer periods with no snowfall together with increasing temperatures.	Introduce a 6-month period with no snowfall and rising temperatures; the model should continue with progressive snowmelt, as dictated by energy balance calculations.

Table 3. Test scenario and description.

Interpretability Evaluation

Interpretability evaluation also forms the critical basis of the model validation so that not only does it forecast accurately, but the explanation also makes sense to support already known scientific principles. Qualitative evaluations by domain experts, such as climate scientists, especially play an important role in assessing whether explanations provided by the model make any climatological sense as discussed in Table 4.

Lastly, it is possible for analysts to estimate the interpretability of the model results, meaning, the degree by which they can make sense out of the model's reason behind a prediction. In particular, in a case where the model predicts a steep drop in snowpack, the explanation, perhaps with respect to an atmospheric pressure pattern, should be intuitive and make easy sense to climate experts. Here, coherence means that the model's explanations are compatible with the scientific understanding already laid down. If the model foresees an increase in snowpack for a rise in temperatures, then this would become a flag. The given explanation would be cross-checked by experts whether or not is in agreement with so-called physical laws of snow dynamics, such as a negative relationship between temperature and snow accumulation.

Metric	Description	Application
Simplicity	The ease of the explanation, in that the domain expert can easily understand the model's explanation for a prediction.	It should easily explain from a linear relationship with precipitation why it predicted snow accumulation for the expert.
Coherence	The degree to which the model's explanations are consistent with prior knowledge and fundamental physical laws.	The model should explain that, because of known principles of energy balance, an increase in temperature results in the melting of snow.
Fidelity	The degree to which the interpretable model is a good approximation to the behavior of the complex model.	This must be made clear in the explanations of the simplified decision tree or rule-based model, while being in agreement with predictions made by the deep learning model.

Table 4. Evaluation metrics for interpretability.

$$Fidelity = \frac{1}{n}\sum_{n=1}^{n} 1\left(P_{simplified}(x_i) = P_{complex}(x_i)\right) \quad \textit{Eqn. 13}$$

In equation 13 for fidelity, $P_{simplified}$ (x_i) is the prediction from the simplified, interpretable model, $P_{complex}$ (x_i) is the prediction from the deep learning, complex model. 1 is an indicator function: equals 1 in the case where the predictions are the same and 0 otherwise. The above makes sure that the interpretable model keeps consistency with the behavior of the more complex model, and consequently, provides explanations that can truly represent the processes underlying the snowpack variability predictions.

Integration of AI with Domain Knowledge

This is required for enhancing interpretability and reliability in the snowpack predictions on the basis of domain knowledge emanating from polar climate science. In this regard, one possible effective research could be knowledge distillation-a process utilized in transferring awareness from complicated, high-performance models (of normally termed "teacher models") into simpler, more interpretable models (known as "student models"). The process aims at retaining the most critical decision-making abilities of a complex AI model with further enhancements for better transparency and interpretability.

Knowledge Distillation Process

The process of knowledge distillation generally involves a complex model, such as deep learning-based PINNs or GNNs, which characterizes snow dynamics with complex patterns both in spatial dependencies and temporal variability. Most of the time, this model is a much lightweight model, typically a simple decision tree or linear regression, trained on the task of emulating the teacher model by minimizing the difference in their predictions as seen in Table 5. The objective function for knowledge distillation can be written as:

$$L_{distill} = \alpha L_{student}(y, \hat{y}) + (1-\alpha) L_{transfer}\left(y_{complex}, \hat{y}_{simple}\right) \qquad \textit{Eqn. 14}$$

In equation 14, $L_{student}$ $(y, \hat{y})$ is the loss between the student model's prediction ($\hat{y}$) correct results (y), $L_{transfer}$ (y_{comlex}, $\hat{y}_{simple}$) is the transfer loss, or the difference between the complex model's output ($y_{complex}$), and the output of the student model output ($\hat{y}_{simple}$), α is the balance hyperparameter that decides the importance of learning directly from the dataset and emulating the complex model.

The process provides a simple model with the ability to learn domain knowledge and learned representations directly from the complex AI model in balance with interpretability and performance. Another integration method can be the inclusion of heuristic-based corrections based on polar climate knowledge. The thresholds which are very well understood can be used to drive model predictions. Snow density (ρ_s) is one of the most fundamental parameters in the snowpack's response to many controlling climatic factors. Domain expertise might give hard constraints, such as in equation 15:

$$\rho_s \leq \frac{600 \text{ kg}}{m^3} \qquad \textit{Eqn. 15}$$

In polar regions, snow density hardly ever exceeds this value. Expert-driven rules of this nature can serve to refine the predictions of the AI model, keeping the output values within physically plausible bounds. These kinds of constraints prevent the models from generating outputs that clearly violate well-known climatic principles, thereby increase their reliability and interpretability.

Iterative Feedback with domain experts

This iterative involvement of climate scientists in the process develops their AI models for snowpack variability. Such a feedback loop can enable the domain expert to assess predictions and explanations coming from a model and provide insight into necessary adjustments in the model to have the outputs of the AI much better aligned with physical laws and expert understanding. It provides for an iterative process aimed at gradual improvement in both accuracy and interpretability of the model.

Integration Method	Description	Application
Knowledge Distillation	The knowledge transfer is done using complex models and transferred to simpler, more interpretable models.	Have a deep learning model act as a teacher and a decision tree acting as the student model while performing predictions on snowpack accumulation.
Heuristic-Based Adjustments	Apply domain knowledge-based rules in order to constrain model predictions within scientifically accepted limits of the outputs.	Thresholds for snow density below which unrealistic snowpack density predictions cannot occur during times of rapid accumulation.

Table 5. Integration methods.

Improvement Process for Explanations

Polar climate experts iteratively improved the explanations based on the evaluation of model performance. Assuming an initial iteration of the model yield predictions regarding the variability of snowpack, the experts reviewed the coherence of the prediction with established climatic phenomena. These phenomena include the relation of the albedo of the snow to the solar radiation. According to the model prediction, the snow accumulation increased when the temperature was high, contrary to the expectations of the expert. In that case, the experts may propose fine-tuning of the model regarding a better representation of energy balance equations or its time features. This can be formalized as a feedback loop below:
Model Prediction: The AI model makes a prediction about snowpack variability from given atmospheric and environmental inputs.
Expert Evaluation: Climate experts review the predictions for their consistency with well-known physical processes, such as energy flux balance and laws of snow compaction.
Feedback and Changes: Based on the feedback given by the experts on instances where the model fails to behave as expected, modification is done either in the model architecture or in the considered features.
Model Refining: The model is revised with the contribution of experts, giving way to better levels of physical realism and interpretability.

Model Deployment and Use Case Testing

After refinement and validation, the explainable AI models should be subjected to realistic use-case scenarios. Those relevant to polar regions such as Antarctica and Greenland are the key regions of interest in terms of snowpack modeling because of the high climatic specifics, the large volume of ice, and the outstanding impact over global sea-level rise. These are involved with complex climatic interactions that interact significantly with the snowpack behavior. Prototype development of decision-support tools has to be done in order to make the model integrations realistic in applications involving scientific and policy stakeholders by assisting them in making informed decisions.

Use-Case Scenarios

It might consider, for instance, the snowpack variability under various climate forcing scenarios in Antarctica. One such use-case could be the determination of snowmelt rates in response to forecasted temperature rises induced by global warming. The AI model embedded in a decision-support tool would provide not only the predictions but also explainable insights, such as the main causes of snowmelt, including increased solar radiation or stronger wind patterns driving sublimation.

In another use-case over Greenland, the model predicts how sudden warming events or anomalous precipitation would influence snowpack. The decision-support tool enables users to view the predictions and also the spatial-temporal distribution of key features controlling snow accumulation and melting. This can be visualized as an interactive visualization in which the user is allowed to navigate the ranking of important variables, changing the climatic scenarios, and receive real-time information on the snowpack variability predictions.

Model Deployment Process

This is completed through the following stages of the deployment process: Integration into decision-support tools: AI models are integrated into interactive tools that allow scientists and policymakers to explore the predictions, explanations, and underlying features.
User Testing: Polar researchers and decision-makers test the system under a range of scenarios, providing feedback on usability and clarity of explanation as shown in Table 6.
Iterative Improvements: This tool is iteratively improved based on the users' feedback. It helps in settling users' practical needs, so that the explanations given by the model can be clear and actionable.

Use-Case	Region	Scenario	Decision-Support Application
Predicting Snowmelt	Antarctica	Predict snowmelt in response to an increase in temperature and changes in solar radiation.	Modular visual display of drivers of snowmelt with respect to temperature trends and albedo change.
Sudden Warming Event	Greenland	Assess how the sudden warming events would impact the stability of snowpack.	Hands-on experience-a workstation where users simulate different warming scenarios and visualize their effect on the snowpack through interpretable model outputs.

Table 6. Use-Cases for regions and scenarios.

In either case, the models can be refined once new scientific understanding becomes available with the effect of keeping them open towards future insights into polar climate dynamics.

Equations and Interpretability in Deployment

To help ensure that the results of these models are transparent, equations that drive such key snow dynamics must be provided within the decision-support tools. Specifically, users can get explanations referring them to basic physical principles, such as explaining how temperature relates to snowmelt:

$$\dot{S} = -\alpha \cdot T \; for \; T > T_{freeze} \qquad \textit{Eqn. 16}$$

In equation 16, $\dot{S}$ is the rate of snowmelt, α is a proportionality constant related to the energy balance, T is the temperature, T_{freeze} is the freezing temperature. Generally, interpretability is enhanced by providing users both model predictions and relevant physical equations, since AI outputs relates to scientifically established laws. In this context, the research makes sure that models are scientifically valid and interpretable, embedding domain knowledge in AI models, while keeping a feedback loop with experts. Second, the development of decision-support tools for real-world applications have a result in making such models theoretically sound while practically useful for those policymakers and scientists concerned with polar snowpack dynamics.

Discussion

Development of explainable AI models for predicting variability in polar snowpack is a very unique and cross-disciplinary approach; it merges ML with domain knowledge from climate science. Interpretable application has developed a clear focus that in such complex dynamics of the behavior of the snowpack there is no explicit dependence on an extensive dataset. This has also aligned with the current demand for more transparency in AI systems, especially in critical fields related to climate science, where model decisions need to be understandable and scientifically valid.

Some of the important challenges being addressed in this research paper involves applying physical principles to the AI models, including but not limited to methods such as Physics-Informed Neural Networks. The embedding of fundamental laws related to snow dynamics, like mass and energy balance into models, has the implication of not just being informed by empirical data but actually constrained by scientific theory. This hybrid approach makes up the gap from purely data-driven methods to physics-based models by keeping the AI system physically realistic.

GNNs and Recurrent Neural Networks have been used to give spatial and temporal predictions of snowpack, respectively, showing the need to capture the interdependence of different regions and time. Such vast and connected polar regions are efficiently mapped by GNNs for the variability of the snowpack, while the RNNs is very important in handling sequential climatic events. The challenge is how far these complex deep learning models should go while balancing with interpretability, and it is there that post-hoc explanation techniques such as SHAP and saliency maps play an important role. This research is further strengthened by the feedback loop of climate scientists. It thus creates a sort of iteration that helps the continuous refinement by experts' insights into the models for better tuning of the AI predictions against real-world physical phenomena. As these models continue to evolve, they start to integrate heuristic-based adjustments and expert-driven rules, therefore enhancing their accuracy and relevance for the particular use case, such as predicting snowmelt in Antarctica or sudden warming events in Greenland.

Conclusion

This research paper focuses on some of the solid frameworks that exist to date in order to build explainable AI models that suit the desired task of snowpack variability prediction in polar regions. The state-of-the-art architectures bring together PINNs, GNNs, and RNNs with knowledge distillation and heuristic adjustments to balance predictive power with interpretability. This makes them even more interpretable by incorporating techniques from both causal inference and intrinsic explainability, hence addressing one of the most significant concerns in applying AI to climate science.

In our paper, the approach is proposed wherein AI models are developed to be surrounded by the basic principles of physical science while revised iteratively in collaboration with domain experts to further refine models and their explanations. The conception of decision-support tools renders the practicality of these AI models, providing actionable insights into snowpack variability and their larger implications upon polar climates for policymakers and scientists. Of course, beyond the successful incorporation of AI into domain knowledge, lie further opportunities for such expansion in this approach to other critical aspects of climate science, such as sea ice prediction or permafrost thaw modeling. However, if these explainable AI models are to provide useful support to climate policy and related decision-making in an ever-changing global environment such models require further validation with real data and ongoing interaction with the climate science community.

References

ARP, C. D., JONES, B. M., ENGRAM, M., ALEXEEV, V. A., CAI, L., PARSEKIAN, A., HINKEL, K., BONDURANT, A. C. and CREIGHTON, A. (2018): Contrasting lake ice responses to winter climate indicate future variability and trends on the Alaskan Arctic coastal plain. *Environmental Research Letters*, 13(12): 125001. doi: 10.1088/1748-9326/aae994

BLAU, M. T., TURTON, J. V., SAUTER, T. and MOLG, T. (2021): Surface mass balance and energy balance of the 79N Glacier (Nioghalvfjerdsfjorden, NE Greenland) modeled by linking COSIPY and Polar WRF. *Journal of Glaciology*, 67(266): 1093-1107. doi: 10.1017/jog.2021.56

BONSOMS, J., OLIVA, M., ALONSO-GONZALEZ, E., REVUELTO, J. and LOPEZ-MORENO, J. I. (2024): Impact of climate change on snowpack dynamics in coastal Central-Western Greenland. *Science of the Total Environment*, 913: 169616. doi: 10.1016/j.scitotenv.2023.169616

CELLI, G., CAIRNS, W. R. L., SCARCHILLI, C., CUEVAS, C. A., SAIZ-LOPEZ, A., SAVARINO, J., STENNI, B., FREZZOTTI, M., BECAGLI, S., DELMONTE, B., ANGOT, H., FERNANDEZ, R. P. and SPOLAOR, A. (2023): Bromine, iodine and sodium along the EAIIST traverse: Bulk and surface snow latitudinal variability. *Environmental Research*, 239: 117344. doi: 10.1016/j.envres.2023.117344

COLLOW, A. B. M., THOMAS, N. P., BOSILOVICH, M. G., LIM, Y.-K., SCHUBERT, S. D. and KOSTER, R. D. (2022): Seasonal variability in the mechanisms behind the 2020 Siberian heatwaves. *Journal of Climate*, 35(10): 3075-3090. doi: 10.1175/jcli-d-21-0432.1

DAVESNE, G., DOMINE, F. and FORTIER, D. (2022): Effects of meteorology and soil moisture on the spatio-temporal evolution of the depth hoar layer in the polar desert snowpack. *Journal of Glaciology*, 68(269): 457-472. doi: 10.1017/jog.2021.105

DOCHERTY, C. L., RIIS, T., MILNER, A. M., CHRISTOFFERSON, K. S. and HANNAH, D. M. (2018): Controls on stream hydrochemistry dynamics in a high Arctic snow-covered watershed. *Hydrological Processes*, 32(22): 3327-3340. doi: 10.1002/hyp.13256

EBRAHIMI, S., MARSHALL, S. J. (2016): Surface energy balance sensitivity to meteorological variability on Haig Glacier, Canadian Rocky Mountains. *Cryosphere*, 10(6): 2799-2819. doi: 10.5194/tc-10-2799-2016

ERNAKOVICH, J. G., HOPPING, K. A., BERDANIER, A. B., SIMPSON, R. T., KACHERGIS, E. J., STELTZER, H. and WALLENSTEIN, M. D. (2014): Predicted responses of arctic and alpine ecosystems to altered seasonality under climate change. *Global Change Biology*, 20(10): 3256-3269. doi: 10.1111/gcb.12568

HOMAN, J. W., KANE, D. L. (2015): Arctic snow distribution patterns at the watershed scale. *Hydrology Research*, 46(4): 507-520. doi: 10.2166/nh.2014.024

IRANNEZHAD, M., RONKANEN, A. K. and KLOVE, B. (2016): Wintertime climate factors controlling snow resource decline in Finland. *International Journal of Climatology*, 36(1): 110-131. doi: 10.1002/joc.4332

JENNINGS, K., JONES, J. A. (2015): Precipitation-snowmelt timing and snowmelt augmentation of large peak flow events, western Cascades, Oregon. *Water Resources Research*, 51(9): 7649-7661. doi: 10.1002/2014wr016877

LI, F., ORSOLINI, Y. J., WANG, H., GAO, Y. and HE, S. (2018): Atlantic multidecadal oscillation modulates the impacts of Arctic sea ice decline. *Geophysical Research Letters*, 45(5): 2497-2506. doi: 10.1002/2017gl076210

LI, J., LI, F. and WANG, H. (2020): Subseasonal prediction of winter precipitation in southern China using the early November snowpack over the Urals. *Atmospheric and Oceanic Science Letters*, 13(6): 534-541. doi: 10.1080/16742834.2020.1824547

OWCZAREK, P., OPALA, M. (2016): Dendrochronology and extreme pointer years in the tree-ring record (ad 1951-2011) of Polar willow from Southwestern Spitsbergen (Svalbard, Norway). *Geochronometria*, 43(1): 84-95. doi: 10.1515/geochr-2015-0035

PLETZER, T., CONWAY, J. P., CULLEN, N. J., EIDHAMMER, T. and KATURJI, M. (2024): The application and modification of WRF-Hydro/Glacier to a cold-based Antarctic glacier. *Hydrology and Earth System Sciences*, 28(3): 459-478. doi: 10.5194/hess-28-459-2024

ROY, A., LEDUC-LEBALLEUR, M., PICARD, G., ROYER, A., TOOSE, P., DERKSEN, C., LEMMETYINEN, J., BERG, A., ROWLANDSON, T. and SCHWANK, M. (2018): Modelling the L-Band Snow-Covered surface emission in a winter Canadian prairie environment. *Remote Sensing*, 10(9): 1451. doi: 10.3390/rs10091451

RU, Y., REN, X. (2024): Subseasonal variability of sea level pressure and its influence on snowpack over mid-high-latitude Eurasia during boreal winter. *Climate Dynamics*, 62(8): 8299-8318. doi: 10.1007/s00382-024-07345-5

SINGH, S. K., THAKUR, P., TRIPATHI, N., JOSHI, P., GARG, V., DUBEY, A., GHETIYA, S., SRIGYAN, M., DAS, J., JAYAPRASAD, P., SHUKLA, A. and OZA, S. R. (2024): Harnessing the potential of EOS-04 SAR data for Himalayan and polar cryospheric studies. *Current Science*, 126(9): 1077-1087. doi: 10.18520/cs/v126/i9/1077-1087

SOBOTA, I., WECKWERTH, P. and GRAJEWSKI, T. (2020): Rain-On-Snow (ROS) events and their relations to snowpack and ice layer changes on small glaciers in Svalbard, the high Arctic. *Journal of Hydrology*, 590: 125279. doi: 10.1016/j.jhydrol.2020.125279

SPOLAOR, A., ANGOT, H., ROMAN, M., DOMMERGUE, A., SCARCHILLI, C., VARDE, M., DEL GUASTA, M., PEDELI, X., VARIN, C., SPROVIERI, F., MAGAND, O., LEGRAND, M., BARBANTE, C. and CAIRNS, W. R. L. (2018): Feedback mechanisms between snow and atmospheric mercury:

Results and observations from field campaigns on the Antarctic plateau. *Chemosphere*, 197: 306-317. doi: 10.1016/j.chemosphere.2017.12.180

SPOLAOR, A., BURGAY, F., FERNANDEZ, R. P., TURETTA, C., CUEVAS, C. A., KIM, K., KINNISON, D. E., LAMARQUE, J.-F., DE BLASI, F., BARBARO, E., CORELLA, J. P., VALLELONGA, P., FREZZOTTI, M., BARBANTE, C. and SAIZ-LOPEZ, A. (2021): Antarctic ozone hole modifies iodine geochemistry on the Antarctic Plateau. *Nature Communications*, 12(1): 5836. doi: 10.1038/s41467-021-26109-x

SPOLAOR, A., VALLELONGA, P., GABRIELI, J., MARTMA, T., BJORKMAN, M. P., ISAKSSON, E., COZZI, G., TURETTA, C., KJAER, H. A., CURRAN, M. A. J., MOY, A. D., SCHOENHARDT, A., BLECHSCHMIDT, A. M., BURROWS, J. P., PLANE, J. M. C. and BARBANTE, C. (2014): Seasonality of halogen deposition in polar snow and ice. *Atmospheric Chemistry and Physics*, 14(18): 9613-9622. doi: 10.5194/acp-14-9613-2014

STROEVE, J. C., VEYSSIERE, G., NAB, C., LIGHT, B., PEROVICH, D., LALIBERTE, J., CAMPBELL, K., LANDY, J., MALLETT, R., BARRETT, A., LISTON, G. E., HADDON, A. and WILKINSON, J. (2024): Mapping potential timing of ice algal blooms from satellite. *Geophysical Research Letters*, 51(8): e2023GL106486. doi: 10.1029/2023gl106486

THAPA, K. K., SINGH, B., SAVALKAR, S., FERN, A., RAJAGOPALAN, K. and KALYANARAMAN, A. (2024): Attention-based models for snow-water equivalent prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(21): 22969-22975. doi: 10.1609/aaai.v38i21.30337

VALT, M., SALVATORI, R. (2016): Snowpack characteristics of Broggerhalvoya, Svalbard Islands. *Rendiconti Lincei-Scienze Fisiche E Naturali*, 27: 129-136. doi: 10.1007/s12210-016-0509-7

VAN DIJK, A. I. J. M., RENZULLO, L. J., WADA, Y. and TREGONING, P. (2014): A global water cycle reanalysis (2003-2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble. *Hydrology and Earth System Sciences*, 18(8): 2955-2973. doi: 10.5194/hess-18-2955-2014

VARTY, S., LEHNHERR, I., ST PIERRE, K., KIRK, J. and WISNIEWSKI, V. (2021): Methylmercury transport and fate shows strong seasonal and spatial variability along a High Arctic freshwater hydrologic continuum. *Environmental Science & Technology*, 55(1): 331-340. doi: 10.1021/acs.est.0c05051

WAHL, S., WALTER, B., AEMISEGGER, F., BIANCHI, L. and LEHNING, M. (2024): Identifying airborne snow metamorphism with stable water isotopes. *Cryosphere*, 18(9): 4493-4515. doi: 10.5194/tc-18-4493-2024

WEINHART, A. H., KIPFSTUHL, S., HOERHOLD, M., EISEN, O. and FREITAG, J. (2021): Spatial distribution of crusts in Antarctic and Greenland snowpacks and implications for snow and firn studies. *Frontiers in Earth Science*, 9: 630070. doi: 10.3389/feart.2021.630070