Probabilistic Machine Learning in Polar Earth and Climate Science: A Review of Applications and Opportunities

Kim Bente,[1] **Judy Kay,** [1] **Roman Marchant** [2]

[1] School of Computer Science, The University of Sydney, Australia
[2] CSIRO's Data61, Australia
kim.bente@sydney.edu.au, judy.kay@sydney.edu.au, roman.marchant@data61.csiro.au

Abstract

Our world's climate future is on thin ice. The study of long-term weather patterns in the polar regions is an important building block in tackling Climate Change. Our understanding of the past, the present and the future of the earth system, and the inherent uncertainty, informs planning, mitigation, and adaptation strategies. In this work we review previous applications of machine learning and statistical computing to polar climate research, and we highlight promising probabilistic machine learning methods that address the modelling needs of climate-related research in the Arctic and the Antarctic. We discuss common challenges in this interdisciplinary field and provide an overview of opportunities for future work in this novel area of research.

Introduction and Background

This section introduces and defines the Polar Earth and Climate Science domain and substantiates the urgent need for continuing research in this field, most importantly to inform policy and decision-making based on a scientific, **uncertainty-aware** foundation. Next, we give a brief background on the recent growth in machine learning (ML) to address Climate Change, propelled by an increase in data availability, the simultaneous leaps in computing power, and advances in artificial intelligence (AI) and machine learning methods. We motivate the emphasis on using a probabilistic framework to convey uncertainty modelling needs.

With this uncertainty-aware perspective, this paper contributes a review of important machine learning applications in the polar parts of the Earth and Climate Science domain, thus building on the wider work of Rolnick et al. (2019). We discuss methodological aspects and the types of domain problems addressed in previous work to then synthesise common challenges. We introduce suitable probabilistic machine learning methods, particularly Bayesian Optimisation and causal methods, and highlight novel research from these areas where we recognise strong opportunities for future work in polar climate applications.

Polar Earth and Climate Science

Climate Change is one of the greatest challenges humanity is facing today. While on average, our globe is warming, temperatures in the Arctic have increased by more than double the global mean over the last two decades (IPCC 2019). Projections from a framework of state-of-the-art physics-based climate simulation models, CMIP6 (the acronym for Coupled Model Intercomparison Project phase 6) (Eyring et al. 2016), predict that the Arctic ocean will become largely free of sea ice during summer months by 2050, even under optimistic anthropogenic emission scenarios (Notz and Community 2020). Contributions from ice sheets and glaciers, notably the Greenland Ice sheet and the West Antarctic Ice Sheet, are understood to be the dominant source of the rise in sea level (IPCC 2019). This poses a direct threat to the livelihoods of a large number of people who live on low-lying islands, in coastal regions but also in inland, flood prone areas. These and other concerning changes like ocean acidification resulting from absorption of anthropogenic CO2 emissions, ocean warming (IPCC 2019), or the acceleration in the Antarctic Circumpolar Current (Shi et al. 2021) highlight the critical role the polar regions hold in the context of the climate system: The cryosphere, describing all frozen water part of the Earth system, as well as the neighboring oceans, are strongly linked to other components of the global climate system through the exchange of carbon, water and energy (IPCC 2019).

Accelerated by the recent pace of change and the looming threats to livelihoods and ecosystems, there are strong academic efforts in further growing our understanding of the field. Earth Science and Climate Science are both well-established research areas. The Earth Sciences are traditionally decomposed into the five interacting systems of earth, namely the atmosphere, the hydrosphere, the biosphere, the geosphere, and the cryosphere. Climate Science is the study of long-term weather patterns, primarily investigating atmospheric properties, but also building on the other subsystems of the Earth Sciences by studying interactions with, for example, the ocean, or, over longer timescales, the geosphere (Springer Nature 2022). To illustrate the interconnection of these disciplines, ice cores from Antarctica for instance, enable paleoclimatology researchers to determine past concentrations of greenhouse gases in the atmosphere. To do so, they analyse air bubbles which were trapped in the ice up to a million years ago. Thus, discoveries in the Earth Sciences often seed new insights for Climate. The problem of mapping the bedrock topography of Antarctica further show-

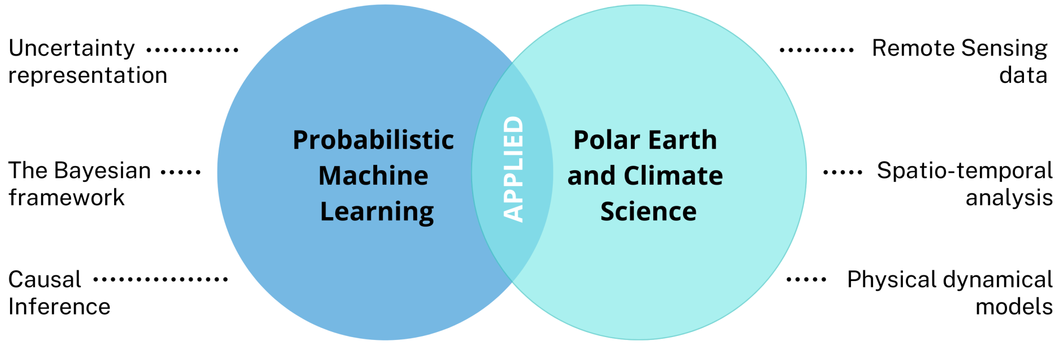

Figure 1: Schematic diagram of selected aspects relevant to the intersection of the Probabilistic Machine Learning field and the Polar Earth and Climate Science domain discussed in this paper.

cases how these fields are intertwined. This geology and Earth Sciences problem is directly related to the estimation of the ice topography and ice mass - quantities climate scientists are highly concerned with (Lythe and Vaughan 2001; Fretwell et al. 2013). Both examples fall under the umbrella of the Polar Sciences, a term that generally denotes scientific research from different disciplines relating to the polar regions (Elsevier 2022).

Because of the urgency imposed by the rapidly changing climate and its transnational scope, global organisations have formed and governments have committed to direct research resources, investments, and policy changes at this pressing issue. CSIRO, Australia's national science agency, identified adapting to climate change as a global megatrend, with particular concern about natural disasters, climate-driven migration and impacts on water quality, infrastructure and also public health (Naughtin et al. 2022). The Intergovernmental Panel on Climate Change (IPCC), a body of the United Nations and paramount international platform, was created to assess the scientific foundations of Climate Change and to inform policy makers about their findings. Leading researchers from the various interconnected fields, including those studying earth and climate, contribute to the IPCC assessment reports. The most recent, sixth assessment cycle includes a 'Special Report on the Ocean and Cryosphere in a Changing Climate' (IPCC 2019), emphasising the need to deepen understanding in this dedicated domain.

Machine Learning for Climate Research

Whilst Earth and Climate Science are well-established areas of research, and specifically research concerning the Arctic and Antarctic is advancing, the intersection of Artificial Intelligence/Machine Learning/Data Science and the Climate Sciences is a fairly novel field. Within this community the first of the annual Climate Informatics conference series, referred to as Climate Informatics, was held in 2011. Monteleoni, Schmidt, and McQuade (2013) provide an overview of new opportunities in this field, and in 2022 a new journal, named Environmental Data Science, positioned at the interface of Data Science and the environment, was established by leaders from the Climate Informatics community. Figure 1 presents a schematic overview of the two intersecting research areas and highlighted concepts reviewed in this paper. These concepts, discussed throughout this section, are signaled with bold font. The data-focused and Earth Science research communities differ in their modelling paradigms, publishing norms, and research priorities. Despite these research silos, a growing community of researchers has acted on the great opportunity of truly interdisciplinary research and has established research organisations with the aim of combining powerful machine learning and statistical methods with the deep understanding of climate and earth system processes and the high-impact questions driving research in related fields. The organisation, Climate Change AI, emerged in 2019 from a series of workshops on 'Tackling Climate Change with Machine Learning' at leading machine learning conferences, as well as side events at the 2019 and 2021 United Nations Climate Change Conferences (COP25 and COP26 respectively) (Climate Change AI 2022). The eponymous paper (Rolnick et al. 2019) gives a big picture overview of problems associated with climate change where machine learning can be applied with impact. Rolnick et al. (2019) allocate areas of machine learning to suitable climate change solution domains, spanning mitigation and adaptation strategies. Within climate prediction to inform adaptation strategies, Rolnick et al. outline important sub problems like data assimilation or the incorporation of ice sheet dynamics into climate models to improve projections. In this paper, we aim to build on this overview, by deepening the review of machine learning applications to Polar Climate and Earth Science problems, and by outlining opportunities suitable to this specific domain.

Remote sensing measurements (see Figure 1) from satellites and aircrafts, data from fixed monitoring stations, and field measurements from ice cores, roaming UAVs, or oceanographic research vessels and floats (Shi et al. 2021) are all contributing to an increase in earth observation data available today. Continuous earth observations by satellite only started with Landsat 1 in 1972 (NASA 2021) so it can-not support the study of long-term climate patterns. Fortunately, indirect measurements of ice cores, rocks and corals, can provide data that goes multiple glacial periods back. Advances in remote sensing technology allow a wide variety of properties to be directly measured or inferred, including altimetry, seismic activity, gravimetry, surface albedo, sea surface wind speeds or atmospheric properties. Data in this domain commonly have **spatial and temporal** dimensions (see Figure 1) and thus exhibit varying resolutions. These special characteristics can incur challenges with the **data fusion** and modelling process. Shirmard et al. (2022) provide a review of how machine learning and specifically deep learning is utilised to process various remote sensing data for mapping geological features - a use cases which is closely related to Climate Science applications. Overall, the data surge is a momentous opportunity to increase our understanding of the least explored and less understood parts of the Earth, such as the oceans, the Arctic and Antarctic. Together with the simultaneous increase in computing power (hardware and algorithms) and the rise of machine learning and statistical computing, particularly in deep learning and causal inference methods (see Figure 1), this is creating vast opportunities to harness data-centric methods for scientific discovery.

Uncertainty is an essential aspect of climate change data and its analysis. Predictions from climate models, together with their associated uncertainty, need to be interpreted to inform sensible decision-making. The uncertainty materialised in predictions arises from multiple sources and can be classified into measurement and model related. Some of these source of uncertainty arise from: physical limitations on sensors that place an upper bound on accuracy, data sets which can present biases, models with limited complexity which are imperfect representations of natural phenomena and inaccurate assumptions. We therefore believe that quantifying model uncertainty with probabilistic machine learning methods, is important, especially in this domain. Many probabilistic machine learning methods are rooted in the **Bayesian framework** (see Figure 1), where model parameters are represented with random variables, whose probability distributions are used as a central tool to represent uncertainty on different layers of abstraction in the model. Furthermore, a fully Bayesian approach incorporates domain expert knowledge through prior distributions, which after careful elicitation are combined with data and model assumptions to provide logically consistent and uncertainty aware estimations. We will therefore emphasise the perspective of uncertainty quantification throughout this paper.

Review of applications

This section reviews machine learning and statistical computing applications for Polar Climate and Earth Science. Applications are grouped into climate model emulators, sea level rise prediction, topography mapping, sea ice forecasting, and lastly climate feedbacks and teleconnection. Table 1 gives an overview of selected applications discussed, the geographical region addressed, the methods used, and the high-level discipline of the publishing venue.

Climate model emulators

State-of-the-art climate models, also known as Earth System Models (ESMs), simulate the interactions between the main climate drivers (atmosphere, land, ocean and ice) through physics-based coupled dynamics, to study the processes based on simulated data and to make predictions about future climate (Rolnick et al. 2019; Balaji et al. 2017). The latest state-of-the-art CMIP model, CMIP6 (Eyring et al. 2016), is highly computationally expensive and data intensive (Balaji et al. 2017). This complexity arises because the model simulates a large set of different processes and subprocesses within and between the climate drivers, which take place on different time and spatial scales. Furthermore, CMIP6 is a multi-model ensemble of around 100 models which were developed by over 50 different modelling groups (Copernicus 2021), scaling computational demands. One weakness of climate models is their sensitivity to small changes in initial conditions or other inputs (Balaji et al. 2017), known as the butterfly effect from early chaos theory literature (Abraham and Ueda 2000). The characteristics of this challenge, i.e. to learn complex and often spatially distant interactions within an uncertain environment, matches the potential of machine learning which can help with model estimation from fusing large amounts of multi-modal and disparate sources of data.

To combat the computational and robustness issues of climate models, deep learning can be used to create emulation models, which do not sacrifice accuracy but are computationally highly efficient once trained (Reichstein et al. 2019). While climate models remain the benchmark for most general climate prediction tasks today, the use of machine learning models to replace, complement or improve traditional first principle models is gaining momentum: Reichstein et al. propose to combine the strengths of theory-driven and data-driven modelling in a hybrid approach. Physical models are usually interpretable and deeply rooted in theoretical understanding of the phenomenon, while machine learning models are highly flexible and can adapt to data. Based on these different strengths of either paradigm, Reichstein et al. suggest that suitable domain problems replace physical sub-model components which are less well described by physical theory, with machine learning models, which may even be able to learn unexpected patterns unknown to experts. Because the cryosphere is a component of the earth system that is challenging to simulate (Gagné, Gillett, and Fyfe 2015), this could be a great opportunity to apply deep learning emulation models. The authors of (Reichstein et al. 2019) further identify that machine learning models could also be used as a calibration layer on top of traditional models, to correct error patterns of the model. In addition Reichstein et al. emphasise the need to quantify models' credibility and confidence, specifically in the case of extrapolation. This could be achieved by using Bayesian Deep Learning Models, which bridge exactly this gap within deep learning (Chandra, Azizi, and Cripps 2017). On a meta-level, decreasing the computational load for climate modelling will both speed up the process, and benefit the footprint of research in this field.

Application	Region	Method	Reference	Venue category
Emulation of climate models	Global	Deep Learning	Reichstein et al. (2019)	Interdisciplinary
Sea level rise prediction	Antarctic	Hybrid probabilistic modelling [Statistics]	Kopp et al. (2017)	Earth & Climate
Sea level rise prediction	Antarctic	Bayesian Hierarchical Models [Statistics]	Zammit-Mangion et al. (2014, 2015)	ML & Statistics
Bedrock and ice topography mapping	Antarctic	Convolutional Neural Networks (CNN) [Deep Learning]	Leong and Horgan (2020)	Earth & Climate
Sub-seasonal sea ice forecasting	Arctic	Attention-based Ensemble Model (EA-LSTM) [Deep Learning]	Ali et al. (2022)	ML & Statistics
Seasonal sea ice forecasting	Arctic	U-Nets [Deep Learning]	Andersson et al. (2021)	Interdisciplinary
Determining causal climate drivers	Arctic	Causal Effect Networks (CEN)[Causal Inference]	Kretschmer et al. (2016)	Earth & Climate
Determining causal climate feedbacks	Antarctic	Convergent cross-mapping (CCM) [Causal Inference]	van Nes et al. (2015)	Earth & Climate

Table 1: Overview of selected applications of machine learning (ML) and statistical computing methods to problems from the Polar Earth and Climate Science domain. The 'Venue category' reflects the broad research community and is based on the subject area of the journal which the cited work is published in.

Sea level rise predictions

The prediction of sea level rise is an important problem due to its far reaching implications on human habitat. Because the mass balance (the sum of ice losses and gains) from the Greenland ice sheet, the Antarctic ice sheet and glaciers are the primary drivers of sea level rise (IPCC 2019), these modelling tasks are directly related to each other and consequently also to the dynamics of climate models (Rolnick et al. 2019). Government agencies like the United States' NOAA, Australia's CSIRO, dedicated research groups like the Sea Level Research Group from CIRES at the University of Colorado Boulder, or IMBIE, an international collaboration of scientist led by the University of Leeds, all work in this field. The emission sensitivity in the predictions of the IPCC (2019) for mass loss is eminent. Especially in the high-emission scenario the accumulating uncertainty in predicted global mean sea level rise is visible through the wide range of predicted increase at low confidence. In addition, sea level rise is not distributed uniformally around the globe (IPCC 2019). Particularly the modelling of ice loss in the Antarctic is recognised to be challenging. A recent mechanistic understanding of accelerating effects from ice-shelf hydro-fracturing and collapsing of ice cliffs on mass loss, produces non-linear trends that far exceed established predictions (Kopp et al. 2017). In this work Kopp et al. incorporate an ensemble of Antarctic ice-sheet (AIS) simulations with a probabilistic framework. Kopp et al. argue strongly for the use of fully Bayesian models, and recommend for future work to identify domain-imposed constraints and well-informed prior beliefs over parameters.

Aligning with the emphasis on probabilistic methods to address this highly uncertain task, is the work of Gopalan, Zammit-Mangion, and McCormack. This predicts the Antarctica's contribution to sea-level rise using a **Bayesian Hierarchical Model** (Zammit-Mangion et al. 2014, 2015). On a high level, the different hierarchical layers constitute of the parameter model, the process model (modelling latent dynamical processes), and the observation model (Gopalan, Zammit-Mangion, and McCormack 2021). Altimetry, gravimetry and GPS observations are used. Knowledge about multiple relevant physical processes is incorporated into the statistical model as prior distributions and dependence structures, informed by traditional numerical ice dynamics models. A strong advantage of this technique is that all estimated quantities, not just predicted sea-level rise, have an associated credible interval reflecting uncertainty. Estimates, e.g. gravimetry parameter estimates, can be interpreted, offering insights for domain experts. Further, Gopalan, Zammit-Mangion, and McCormack (2021) used approximation methods to improve computational efficiency. They provide an overview of Bayesian modelling and inference in glaciology, showcasing two projects, one being the above work by Zammit-Mangion et al. (2014).

Topography mapping

An understanding of the topography underneath the ice forms the basis for ice sheet modelling. The series of BedMap models, BedMap and the updated BedMap2, comprise of gridded digital topographical models of the surface elevation, subglacial bed rock elevation, sea floor elevation, and also ice thickness for the continent of Antarctica (Lythe and Vaughan 2001; Fretwell et al. 2013). Data from various surveys, at different spatial scales, were assimilated to construct state-of-the-art mappings. The BedMap2 data set lays the foundation for many other researchers in this field. The dependence on up-stream estimates of quantities like subglacial bed rock elevation, which can not be directly measured, exemplify the role of uncertainty within polar research. Building on top of BedMap2, Leong and Horgan (2020) introduce DeepBedMap to address the problem of imputing high spatial resolution bed elevation grids for areas in Antarctica where no data at high resolution is available. A variant of Deep Convolutional Neural Networks, adapted from Enhanced Super-Resolution Generative Adversarial Network, is used to generate high-resolution maps. Additional gridded data on ice surface elevation, velocity and snow accumulation, all available at high spatial resolutions, are used as inputs. To capture the spatial interaction of the different properties, the neural network was trained on ground truth data. Resulting surface roughness was evaluated as an indicator for realistic topography maps. Other recent work uses topographic satellite data to map supraglacial lakes in regions of Antarctica using Random Forest classifiers (Dirscherl et al. 2020). Despite the black-box character of such models, this showcases how machine learning can be used for assimilation and imputation purposes, as a vital element within the process of polar climate research.

Sea ice forecasting

The prediction of sea ice extent is an important task that informs safe shipping routes, hazard alerts, and climate prediction models (Wang et al. 2016). Predictions can even be used to issue warnings prior to events like massive haulouts of walruses, providing the opportunity to prevent high mortality of the species (Andersson et al. 2021). Interannual variability makes sea ice forecasting a challenging task (Gagné, Gillett, and Fyfe 2015; Andersson et al. 2021). Gagné, Gillett, and Fyfe (2015) investigate the contrary resulting trends of simulated and actually observed sea ice data in the Antarctic by extending the historic records with recovered satellite based estimates from 35 to 50 years. The additional data further highlight the presence of high historic variability in the phenomenon, but emphasizes the view that existing climate simulations do not holistically describe the behaviour of sea ice extent. An application at the opposite end of the globe, the Beaufort Sea in the Arctic, uses convolutional neural networks (CNNs) to estimate high-resolution ice concentration maps directly from satellite synthetic aperture radar (SAR) data (Wang et al. 2016). SAR remote sensing is not impaired by cloud cover or the absence of daylight and is therefore a robust input. Although the regional scale of this application is constricted and sea ice concentration is not predicted for the future, the resulting performance, ranking close to the human expert benchmark, is a promising outcome. Since then, various researchers have applied deep learning models to predict sea ice concentrations, however for short, sub-seasonal lead times: Chi and Kim (2017) use deep learning and Kim et al. (2020) later use Convolutional Neural Networks (CNNs), a variant of deep learning, to predict Artic sea ice concentrations. Ali et al. (2022) propose an attention-based Long Short Term Memory (LSTM) ensemble method, combining the strength of attention-based methods to learn distant connections and the ability of LSTMs to remember previous states, analog to previous weather conditions. Ali et al.'s model outperforms previous state-of-the-art models. However, these applications only evaluate 1-month ahead predictions.

In more recent work Andersson et al. (2021) present a machine learning model to predict monthly averaged sea ice probability classes across the entire Arctic region at lead times of 1 to 6 months. They use a range of different input data, including climate variables from the atmosphere and ocean. The model is constructed as an ensemble of U-Nets, a variant of CNNs. U-Nets were originally developed for biomedical image segmentation, a conceptually similar Computer Vision task, mapping from gridded inputs (e.g. images) to gridded outputs. Andersson et al.'s IceNet model outperforms the state-of the art physics-based model at longer prediction lead times. The deep learning ensemble performs particularly well on predicting extreme sea ice conditions. Andersson et al.'s work is exemplary in integrating domain knowledge and machine learning: It not only displays a high level of understanding for the domain, but it also extracts **interpretable results** from the model, that may in turn provide new insights to domain experts and their models. A variable importance analysis is used to understand what inputs are contributing most to yield the predictive results for different months and lead times. The findings are compared to expectations from sea ice forecasting experts, and are mostly found to match domain knowledge. Nonetheless some new discoveries were also made from this data-driven approach. One interesting result is that extensively pre-training the model on CMIP6 climate simulation data barely increased the predictive performance. This supports the recognition that relatively small amounts of observational data, rather than large amounts of simulated data, can be highly indicative of future phenomena when used within suitable modelling settings. Andersson et al. (2021) suggest extending their work by using inputs at higher temporal resolution, with the intention of improving predictive ability at short, 1-month, lead times, where the model is currently under-performing. Furthermore, the authors suggest incorporating ice thickness as a model input to further improve forecasts. While the classifier predicts a discrete probability distribution over the possible sea ice probability classes as an output, there are opportunities to expand on the methodological approach by incorporating the probabilistic framework.

Climate feedbacks and teleconnections

Teleconnections are persistent patterns of climate anomalies that span large geographical areas. Such patterns and their causal structures are hard to detect but they influence climate processes at the global scale. The work by Kretschmer et al. (2016) demonstrates an application of **causal hypothesis testing** to understand Arctic teleconnection patterns: Causal effect networks (CEN), a type of graphical model, are used on time series data to identify autumn Barents and Kara sea ice concentrations as an important driver for mid-latitude winter circulation, which can show as extreme winter conditions in North America and Euroasia. Artic teleconnections are currently not very well understood, and as identified by Rolnick et al. (2019) incorporating them into climate models is likely to improve climate projections at global and regional resolutions. Work by van Nes et al. (2015) uses another type of technique, convergent cross-mapping (CCM), a **non-linear state-space method**, to investigate causal feedback structures in the field of paleoclimatology (van Nes et al. 2015). They use more than 400,000 years of temperature data and greenhouse gas concentrations reconstructed from the Vostok Ice core from Antarctica as a proxy time series. Their results demonstrate that orbital forcing (e.g. insolation) have no significant causal association with either temperature or greenhouse gas concentrations. However, they found a strong feedback effect of temperature variability on greenhouse gases, indicating that warming in itself may drive an increase in greenhouse gas concentrations. This constitutes an important finding on the level of cause and effect structures associated with climate change.

Discussion, Opportunities and next steps

In the following we discuss the applications reviewed in the previous section, and we examine their methods, and distill common challenges. Based on this we introduce and motivate opportunities for probabilistic Machine Learning methods, in particular we introduce Bayesian Optimisation and Causal methods and what use cases for future work they provide.

Discussion

Based on our review of recent works applying machine learning to the Polar Climate and Earth Science domain (see Table 1) we can observe that particular high-impact applications, i.e. sea ice forecasting and sea level rise prediction, have received more attention than others. Deep learning methods are often used in conjunction with satellite data, likely motivated by the success of deep learning for prominent Computer Vision problems, as well as recent achievements in the cryosphere domain (Andersson et al. 2021; Ali et al. 2022). Currently, distributed sub-communities contribute to this new field and relevant work is published across research venues in Earth & Climate, Machine Learning & Statistics or interdisciplinary venues, where terminology, contribution emphasis, and reproducibiity standards vary.

Challenges repeatedly discussed in the literature include:

- combining data from various sources (data fusion);
- dealing with varying spatial and temporal resolutions;
- increasing computational efficiency;
- interpretability of models and model outputs;
- modelling natural variability of phenomena;
- modelling systemic sources of uncertainty related to data and models.

Andersson et al. (2021) showcase how variable importance analysis can be used for deep learning models to make sense of the mechanism behind the black-box-model to address interpretability. However, this work incorporates probabilistic representation only for predicted outputs. Zammit-Mangion et al. (2014) is one of the few to have used Bayesian statistics to model uncertainty throughout the hierarchical model; an endeavour calling for a high degree of domain expertise to inform prior distributions, parameterisation and model structure. As uncertainty is inseparable from Climate research, there are major opportunities to use Probabilistic Machine Learning methods to solve the challanges faced.

Opportunities for Probabilistic Machine Learning

Probabilistic Machine Learning describes those methods that utilise a probabilistic framework to represent uncertainty. The probabilistic modelling framework is rooted in principled theoretical and highly practical approaches that are concerned with “representing and manipulating uncertainty” (Ghahramani 2015). Uncertainty arises from incorrect or biased measurements, from decisions about model structure, from model parameters and from the stochastic nature of the world. Therefore, uncertainty should be propagated through the model and included in model predictions. A review paper (Ghahramani 2015) provides an excellent introduction to Bayesian inference, the core of Bayesian statistics, and an overview of recent advances, specifically, **Bayesian Optimisation**, probabilistic programming, probabilistic data compression, and automatic model discovery. Ghahramani highlights the importance of the probabilistic modelling framework for problems where uncertainty is a “key ingredient”. The paper also discusses a common computational challenge among these probabilistic methods - inference - and how approximate integration methods like Markov Chain Monte Carlo (MCMC) (refer to Andrieu et al. (2003); Brooks et al. (2011) for more detail) or Variational Inference (refer to Blei, Kucukelbir, and McAuliffe (2017)) are related research fields addressing this challenge.

Various methods featured in Ghahramani (2015), including the aforementioned Bayesian Optimisation and its most common underlying surrogate model, Gaussian Processes, originated from the **spatial and spatio-temporal modelling** literature. Gaussian Process regression, which is also known as Kriging in geostatistics, is a class of flexible non-parametric models that has been particularly successful in modelling spatial correlation structures (Marchant and Ramos 2012). Gaussian Process models are discussed in great detail in the textbook (Rasmussen and Williams 2006). In the area of spatio-temporal modelling, the seminal textbooks by Cressie (2011, 1993) combine classical statistical methods and modern computational algorithms and are therefore influential across theoretical and applied fields. Other methods which have been gaining scholarly popularity and are thus worth mentioning are Bayesian Neural Networks (Chandra, Azizi, and Cripps 2017), which combine standard Neural Networks with Bayesian Inference, and **Causal Inference** (refer to Pearl (2009)), which is concerned with ascertaining causal relationships using probabilistic tools. As reviewed in the previous section, Kretschmer et al. (2016) and van Nes et al. (2015) demonstrate the use of causal methods in climate studies. Next, we will discuss Bayesian Optimisation and Causal Inference, as well as opportunities for applying these to polar climate research, in more detail.

Bayesian Optimisation. Bayesian Optimisation is a tool for global optimisation. It is particularly suitable when the objective function is unknown and complex, and when evaluations of the objective function are noisy and costly to obtain (Marchant and Ramos 2012; Archetti and Candelieri 2019; Shahriari et al. 2016). Over iterations, each new query point, where the objectively function is then evaluated, will be chosen carefully and efficiently. While some applications focus on finding the global optimum, other application focus on the iterative determination of the next optimal query point, known as active learning (Shahriari et al. 2016). Well-known use cases for Bayesian Optimisation exist in the design of exploration strategies for mining and geology in environmental applications, where Bayesian Optimisation can inform the design of sensing networks (Shahriari et al. 2016). Exploration drilling is a way of evaluating the unknown objective function, which describes the distribution of sub-surface minerals across space. Exploration drilling is very costly. Hence, data-efficient Bayesian Optimisation is well suited to inform decision making about the selection of promising drilling sites. In environmental monitoring Bayesian Optimisation is used to inform optimal sequential decisions which result in efficient data acquisition of environmental variables of concern (Marchant, Ramos, and Sanner 2014). Bayesian Optimisation takes into account the expected value based on the global model (for the 'optimisation' in Bayesian Optimisation) and also the degree of uncertainty the model has with regard to the expected value, based on the data, the underlying model assumptions and the prior. This is the trade-off between exploitation and exploration. To take into account the added desiderata of minimising sensor travel distance (Marchant and Ramos 2012) propose a new acquisition function, the Distance Based Upper Confidence Bound. They demonstrate considerably reduced travel distance in a real world and a simulated experiment without sacrificing accuracy. Use cases in polar research have strong parallels to this work, with limited sensing resources available, a vast space to explore, and high-uncertainty models. Because Bayesian Optimisation simultaneously updates the probabilistic model of the unknown function and sequentially suggests sampling locations (active learning), the method has dual utility. Therefore, methods building on top of these ideas, for example reflecting geographic or other asymmetric constraints in the acquisition strategy, may be a possible extension of previous work with high practical relevance for Polar Climate research.

Recent work has applied Bayesian Optimisation to actively monitor urban air pollution in London using Hierarchical Bayesian modelling as the surrogate model (Hellan, Lucas, and Goddard 2022). Further work is suggested to explore the use of other kernel families and kernel variations that can capture correlations appearing at different time scales. Another application of Bayesian Optimisation to the environmental domain is the localisation of a contamination source (Pirot et al. 2019). This work provides a good example for integrating hydrology domain knowledge into the objective function. Within the Machine Learning community, Bayesian Optimisation has attracted a lot of attention for its use in optimising hyperparameters of Machine Learning models (Snoek et al. 2014). Open-source software packages like Dragonfly (Kandasamy et al. 2020) enable a ready-to-use implementation of these ideas. Potentially this use case can be transferred to the optimisation of climate models, or to Machine Learning models of the Earth's sub-systems.

Expanding on existing research, future work could apply Bayesian Optimisation to optimise sensor networks for climate monitoring in polar regions, or as an active learning strategy to determine drilling locations for ice cores. An extension to the work of Marchant and Ramos (2012) could propose new acquisition functions that uses a non-stationary cost function which reflects the physical characteristics of the environment. Another challenging problem, shared across the reviewed literature (Gopalan, Zammit-Mangion, and McCormack 2021; Leong and Horgan 2020) is data fusion. Combining data from different remote sensing technologies as well as in situ measurements, demands a principled way of fusing varying uncertainty distributions, interpolating missing data or unifying scales. Whilst this is a sub-problem of applied research generally, the Bayesian framework may offer an elegant way to address this and therefore could benefit other applications of Data Science to climate-related domain problems in the Arctic and Antarctic.

Causal Inference. Climate Modelling is predominately associated with prediction through the implementation of deterministic physical systems which are highly interpretable. With the rise in machine learning methods, a sizeable component of the research community has focused on developing predictive black-box models that can be deployed as flexible and accurate regression (e.g. neural networks, computer vision, recommender systems). These methods, under the Rolnick et al. (2019) framework, are attributed to informing adaptation strategies in response to consequences of Climate Change. However, these methods present serious limitations from a scientific perspective since they: i) do not provide interpretability, thus limiting the capacity for climate scientists to learn from model predictions; ii) show a lack of transparency into the underlying working of the models, which may lead to a lack of trust; and iii) capture correlations and not causation which may result in misleading and incorrect recommendations.

In contrast, modelling the causal mechanisms of Climate Change, thereby discerning anthropogenic and natural causes of warming, will provide insights that inform mitigation strategies with stronger and interpretable evidence. Understanding the causes of phenomena we observe lies at the very heart of scientific discovery (Runge et al. 2019). Many domains, like medicine, use controlled experiments to establish causal links. However, in a large and complex field like Earth Science, where controlled experiments are impossible or unethical, Causal Inference methods based on observational data are a promising new research direction. Runge et al. provide an overview of Causal Inference frameworks for dealing with observational time series data and they suggest suitable applications in the Earth System Sciences. Computer simulation experiments, the prior standard for causal discovery in the Earth Sciences, are computationally expensive and constrained to assumptions made about the systems. Concurrent with the rise in Machine Learning, data availability and increased computing power paved the way for these new causal methods, which rely only on observational data. Research in Bayesian Networks (Pearl 2009) dates back a few decades but forms the foundation for many causal models. An important framework reviewed in Runge et al. (2019) is **Structural Causal Models (SCMs)** (refer to Peters, Janzing, and Schölkopf (2017) for more detail). These are closely related to Bayesian Networks. Both are graphical models, where the nodes of the graph represent variables of interest and the links between nodes represent causal relationships. SCMs are a particularly appealing framework, because various strong assumptions (e.g. about the noise structure) that were previously unavoidable, can be relaxed. SCMs can be viewed as a complement to black box ML models, to increase understanding of the mechanisms of the system (Runge et al. 2019). This understanding of causal relationships is not just a means to an end, but has also been recognised to increase robustness, particularly for out-of-distribution predictions (Runge et al. 2019).

In the context of probabilistic machine learning and uncertainty quantification, the recent rise in fully probabilistic Bayesian network inference has the power of incorporating uncertainty about causal structures by providing posterior distributions over graph structures (Kuipers and Moffa 2017). Furthermore, if causal inference and causal effects are also treated in a fully probabilistic framework, they have the capacity to quantify uncertainty and guide sequential decision making. Causal inference can also be connected with Bayesian Optimisation (Aglietti et al. 2020), which can be generalised to active sampling and intervention strategies that acquire data in order to find the most valuable actions.

Novel developments in Causal Inference frameworks including **Bayesian Networks** and **Structural Causal Models** enable us to gain understanding of causal structures of underlying systems from observational data. These offer great opportunities for future work, for instance, to build more robust climate models, to further understand causal feedbacks in climate change as demonstrated by van Nes et al. (2015), or to distinguish anthropogenic from natural drivers of Climate Change.

Conclusion

Opportunities for applying Machine Learning to solve problems from the Climate Sciences and the Polar Climate Sciences more specifically, are widely recognised and have the potential to be highly impactful (Rolnick et al. 2019). However, because this interdisciplinary research area is still novel, and remote sensing data has only become more accessible and more meaningful with increased sensing coverage and accompanying computing power in recent years, there are more research opportunities than existing work. A large body of work exists on the use of deep learning for remote sensing application (Ma et al. 2019) as well as the Earth Sciences (Reichstein et al. 2019). Aligned with this, many reviewed applications of machine learning to polar climate research use deep learning in combination with satellite data. Some of these applications outperform state-of-the art physics-based models (Andersson et al. 2021), suggesting further promising advances in this direction of research in the future. Common challenges across reviewed literature include the need for data fusion, assimilating multi-resolution data, increasing computational efficiency, enhancing interpretability, and modelling uncertainty. Addressing these challenges is another opportunity for future work and will benefit research down-stream.

Although probabilistic modelling is inevitable for making sensible and informed decisions, methods applied to problems in this fields often lack a framework for uncertainty quantification. To address this need, the class of probabilistic Machine Learning offers a toolbox of methods which are well-suited to reflect real-life uncertainty. We particularly highlight Bayesian Optimisation and Causal Inference methods which are well suited to problems from the Polar Climate and Earth Science domain. Bayesian Optimisation may be used to inform drilling site selection of ice cores, sequential selection of monitoring locations for autonomous sensors, or to optimise stationary sensor networks across the polar regions. Other non-spatial applications include the global optimisation of hyperparameters for machine learning and traditional climate models. Improved experimental design may help in reducing the computational footprint of this computationally intensive field of research. Advances in Causal Inference techniques provide another great opportunity for future work: Quantifying causal drivers of climate change or building more robust prediction models, by resembling the underlying causal structures of the system, could strengthen the uncertainty-aware, scientific foundation for global decision making in stewarding human impact on climate, thereby supporting climate change mitigation and adaption efforts.

References

Abraham, R.; and Ueda, Y. 2000. *The Chaos Avant-garde: Memories of the Early Days of Chaos Theory*. World Scientific. ISBN 978-981-238-647-2.

Aglietti, V.; Lu, X.; Paleyes, A.; and González, J. 2020. Causal Bayesian Optimization. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, 3155–3164. PMLR. ISSN: 2640-3498.

Ali, S.; Huang, Y.; Huang, X.; and Wang, J. 2022. Sea Ice Forecasting using Attention-based Ensemble LSTM. *arXiv:2108.00853*. ArXiv:2108.00853.

Andersson, T. R.; Hosking, J. S.; Pérez-Ortiz, M.; Paige, B.; Elliott, A.; Russell, C.; Law, S.; Jones, D. C.; Wilkinson, J.; Phillips, T.; Byrne, J.; Tietsche, S.; Sarojini, B. B.; Blanchard-Wrigglesworth, E.; Aksenov, Y.; Downie, R.; and Shuckburgh, E. 2021. Seasonal Arctic sea ice forecasting with probabilistic deep learning. *Nature Communications*, 12(1): 5124.

Andrieu, C.; de Freitas, N.; Doucet, A.; and Jordan, M. I. 2003. An Introduction to MCMC for Machine Learning. *Machine Learning*, 50(1): 5–43.

Archetti, F.; and Candelieri, A. 2019. *Bayesian Optimization and Data Science*. SpringerBriefs in Optimization. Cham:

Springer International Publishing. ISBN 978-3-030-24493-4 978-3-030-24494-1.

Balaji, V.; Maisonnave, E.; Zadeh, N.; Lawrence, B. N.; Biercamp, J.; Fladrich, U.; Aloisio, G.; Benson, R.; Caubel, A.; Durachta, J.; Foujols, M.-A.; Lister, G.; Mocavero, S.; Underwood, S.; and Wright, G. 2017. CPMIP: measurements of real computational performance of Earth system models in CMIP6. *Geoscientific Model Development*, 10(1): 19–34. Publisher: Copernicus GmbH.

Blei, D. M.; Kucukelbir, A.; and McAuliffe, J. D. 2017. Variational Inference: A Review for Statisticians. *Journal of the American Statistical Association*, 112(518): 859–877.

Brooks, S.; Gelman, A.; Jones, G.; and Meng, X.-L. 2011. *Handbook of Markov Chain Monte Carlo*. London, United Kingdom: CRC Press LLC. ISBN 978-1-4200-7942-5.

Chandra, R.; Azizi, L.; and Cripps, S. 2017. Bayesian Neural Learning via Langevin Dynamics for Chaotic Time Series Prediction. In Liu, D.; Xie, S.; Li, Y.; Zhao, D.; and El-Alfy, E.-S. M., eds., *Neural Information Processing*, Lecture Notes in Computer Science, 564–573. Cham: Springer International Publishing. ISBN 978-3-319-70139-4.

Chi, J.; and Kim, H.-c. 2017. Prediction of Arctic Sea Ice Concentration Using a Fully Data Driven Deep Neural Network. *Remote Sensing*, 9(12): 1305.

CIRES. 2022. Sea Level Research Group. Https://sealevel.colorado.edu/. Accessed: 2022-07-20.

Climate Change AI. 2022. Climate Change AI - About. Https://www.climatechange.ai/. Accessed: 2022-07-20.

Copernicus. 2021. Latest projections of future climate now available. Https://climate.copernicus.eu/latest-projections-future-climate-now-available. Accessed: 2022-07-20.

Cressie, N. 1993. *Statistics for Spatial Data*. New York, United States: John Wiley & Sons, Incorporated. ISBN 978-1-119-11517-5.

Cressie, N. A. C. 2011. *Statistics for spatio-temporal data*. Wiley series in probability and statistics. Hoboken, N.J: Wiley. ISBN 978-0-471-69274-4.

CSIRO. 2016. Sea-level Rise: CSIRO & ACE-CRC. Https://www.cmar.csiro.au/sealevel/sl_about_us.html. Accessed: 2022-07-20.

Dirscherl, M.; Dietz, A. J.; Kneisel, C.; and Kuenzer, C. 2020. Automated Mapping of Antarctic Supraglacial Lakes Using a Machine Learning Approach. *Remote Sensing*, 12(7): 1203.

Elsevier. 2022. Polar Science Aims and scope. Https://www.sciencedirect.com/journal/polar-science/about/aims-and-scope. Accessed: 2022-07-20.

Eyring, V.; Bony, S.; Meehl, G. A.; Senior, C. A.; Stevens, B.; Stouffer, R. J.; and Taylor, K. E. 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5): 1937–1958. Publisher: Copernicus GmbH.

Fretwell, P.; Pritchard, H. D.; Vaughan, D. G.; Bamber, J. L.; Barrand, N. E.; Bell, R.; Bianchi, C.; Bingham, R. G.; Blankenship, D. D.; Casassa, G.; Catania, G.; Callens, D.; Conway, H.; Cook, A. J.; Corr, H. F. J.; Damaske, D.; Damm, V.; Ferraccioli, F.; Forsberg, R.; Fujita, S.; Gim, Y.; Gogineni, P.; Griggs, J. A.; Hindmarsh, R. C. A.; Holmlund, P.; Holt, J. W.; Jacobel, R. W.; Jenkins, A.; Jokat, W.; Jordan, T.; King, E. C.; Kohler, J.; Krabill, W.; Riger-Kusk, M.; Langley, K. A.; Leitchenkov, G.; Leuschen, C.; Luyendyk, B. P.; Matsuoka, K.; Mouginot, J.; Nitsche, F. O.; Nogi, Y.; Nost, O. A.; Popov, S. V.; Rignot, E.; Rippin, D. M.; Rivera, A.; Roberts, J.; Ross, N.; Siegert, M. J.; Smith, A. M.; Steinhage, D.; Studinger, M.; Sun, B.; Tinto, B. K.; Welch, B. C.; Wilson, D.; Young, D. A.; Xiangbin, C.; and Zirizzotti, A. 2013. Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. *The Cryosphere*, 7(1): 375–393. Publisher: Copernicus GmbH.

Gagné, M.-E.; Gillett, N. P.; and Fyfe, J. C. 2015. Observed and simulated changes in Antarctic sea ice extent over the past 50 years. *Geophysical Research Letters*, 42(1): 90–95.

Ghahramani, Z. 2015. Probabilistic machine learning and artificial intelligence. *Nature*, 521(7553): 452–459. Number: 7553 Publisher: Nature Publishing Group.

Gopalan, G.; Zammit-Mangion, A.; and McCormack, F. 2021. A Review of Bayesian Modelling in Glaciology. *arXiv:2112.13663 [stat]*.

Hellan, S. P.; Lucas, C. G.; and Goddard, N. H. 2022. Bayesian Optimisation for Active Monitoring of Air Pollution. *arXiv:2202.07595 [physics]*. ArXiv: 2202.07595.

IMBIE. 2022. IMBIE. Http://imbie.org/. Accessed: 2022-07-20.

IPCC. 2019. Summary for Policymakers. In Pörtner, H.-O.; D.C. Roberts; V. Masson-Delmotte; P. Zhai; M. Tignor; E. Poloczanska; K. Mintenbeck; A. Alegría; M. Nicolai; A. Okem; J. Petzold; B. Rama; and N.M. Weyer, eds., *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, pp. 3–35. Cambridge, UK and New York, NY, USA: Cambridge University Press.

Kandasamy, K.; Vysyaraju, K. R.; Neiswanger, W.; Paria, B.; Collins, C. R.; Schneider, J.; Póczos, B.; and Xing, E. P. 2020. Tuning hyperparameters without grad students: scalable and robust Bayesian optimisation with dragonfly. *The Journal of Machine Learning Research*, 21(1): 81:3098–81:3124.

Kim, Y. J.; Kim, H.-C.; Han, D.; Lee, S.; and Im, J. 2020. Prediction of monthly Arctic sea ice concentrations using satellite and reanalysis data based on convolutional neural networks. *The Cryosphere*, 14(3): 1083–1104. Publisher: Copernicus GmbH.

Kopp, R. E.; DeConto, R. M.; Bader, D. A.; Hay, C. C.; Horton, R. M.; Kulp, S.; Oppenheimer, M.; Pollard, D.; and Strauss, B. H. 2017. Evolving Understanding of Antarctic Ice-Sheet Physics and Ambiguity in Probabilistic Sea-Level Projections. *Earth's Future*, 5(12): 1217–1233.

Kretschmer, M.; Coumou, D.; Donges, J. F.; and Runge, J. 2016. Using Causal Effect Networks to Analyze Different Arctic Drivers of Midlatitude Winter Circulation. *Journal of Climate*, 29(11): 4069–4081. Publisher: American Meteorological Society Section: Journal of Climate.

Kuipers, J.; and Moffa, G. 2017. Partition MCMC for Inference on Acyclic Digraphs. *Journal of the American Statistical Association*, 112(517): 282–299. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/01621459.2015.1133426.

Leong, W.; and Horgan, H. 2020. DeepBedMap: A deep neural network for resolving the bed topography of Antarctica. *Cryosphere*, 14(11): 3687–3705.

Lythe, M. B.; and Vaughan, D. G. 2001. BEDMAP: A new ice thickness and subglacial topographic model of Antarctica. *Journal of Geophysical Research: Solid Earth*, 106(B6): 11335–11351.

Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; and Johnson, B. A. 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152: 166–177.

Marchant, R.; and Ramos, F. 2012. Bayesian optimisation for Intelligent Environmental Monitoring. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2242–2249. ISSN: 2153-0866.

Marchant, R.; Ramos, F.; and Sanner, S. 2014. Sequential Bayesian optimisation for spatial-temporal monitoring. In *Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence*, UAI'14, 553–562. Arlington, Virginia, USA: AUAI Press. ISBN 978-0-9749039-1-0.

Monteleoni, C.; Schmidt, G. A.; and McQuade, S. 2013. Climate Informatics: Accelerating Discovering in Climate Science with Machine Learning. *Computing in Science & Engineering*, 15(5): 32–40.

NASA. 2021. Landsat 1 | Landsat Science. Https://landsat.gsfc.nasa.gov/satellites/landsat-1/. Accessed: 2022-07-20.

Naughtin, C.; Hajkowicz, S.; Schleiger, E.; Bratanova, A.; Cameron, A.; Zamin, T.; and Dutta, A. 2022. Our Future World: Global megatrends impacting the way we live over coming decades. Technical report, CSIRO, Brisbane, Australia. Publisher: CSIRO.

NOAA. 2022. Explore Sea Level Rise Tools, Services, and Educational Material. Https://oceanservice.noaa.gov/hazards/sealevelrise/. Accessed: 2022-07-20.

Notz, D.; and Community, S. 2020. Arctic Sea Ice in CMIP6. *Geophysical Research Letters*, 47(10): e2019GL086749.

Pearl, J. 2009. *Causality: models, reasoning, and inference*. Cambridge ;: Cambridge University Press, 2nd ed. edition. ISBN 978-0-521-89560-6.

Peters, J.; Janzing, D.; and Schölkopf, B. 2017. *Elements of Causal Inference: Foundations and Learning Algorithms*. Adaptive Computation and Machine Learning series. Cambridge, MA, USA: MIT Press. ISBN 978-0-262-03731-0.

Pirot, G.; Krityakierne, T.; Ginsbourger, D.; and Renard, P. 2019. Contaminant source localization via Bayesian global optimization. *Hydrology and Earth System Sciences*, 23(1): 351–369. Publisher: Copernicus GmbH.

Rasmussen, C. E.; and Williams, C. K. I. 2006. *Gaussian processes for machine learning*. Adaptive computation and machine learning. Cambridge, Mass: MIT Press. ISBN 978-0-262-18253-9.

Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N.; and Prabhat. 2019. Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743): 195–204.

Rolnick, D.; Donti, P. L.; Kaack, L. H.; Kochanski, K.; Lacoste, A.; Sankaran, K.; Ross, A. S.; Milojevic-Dupont, N.; Jaques, N.; Waldman-Brown, A.; Luccioni, A.; Maharaj, T.; Sherwin, E. D.; Mukkavilli, S. K.; Kording, K. P.; Gomes, C.; Ng, A. Y.; Hassabis, D.; Platt, J. C.; Creutzig, F.; Chayes, J.; and Bengio, Y. 2019. Tackling Climate Change with Machine Learning. *arXiv:1906.05433 [cs, stat]*. ArXiv: 1906.05433.

Runge, J.; Bathiany, S.; Bollt, E.; Camps-Valls, G.; Coumou, D.; Deyle, E.; Glymour, C.; Kretschmer, M.; Mahecha, M. D.; Muñoz-Marí, J.; van Nes, E. H.; Peters, J.; Quax, R.; Reichstein, M.; Scheffer, M.; Schölkopf, B.; Spirtes, P.; Sugihara, G.; Sun, J.; Zhang, K.; and Zscheischler, J. 2019. Inferring causation from time series in Earth system sciences. *Nature Communications*, 10(1): 2553.

Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R. P.; and de Freitas, N. 2016. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1): 148–175. Conference Name: Proceedings of the IEEE.

Shi, J.-R.; Talley, L. D.; Xie, S.-P.; Peng, Q.; and Liu, W. 2021. Ocean warming and accelerating Southern Ocean zonal flow. *Nature Climate Change*, 11(12): 1090–1097.

Shirmard, H.; Farahbakhsh, E.; Müller, R. D.; and Chandra, R. 2022. A review of machine learning in processing remote sensing data for mineral exploration. *Remote Sensing of Environment*, 268.

Snoek, J.; Swersky, K.; Zemel, R.; and Adams, R. 2014. Input Warping for Bayesian Optimization of Non-Stationary Functions. In *Proceedings of the 31st International Conference on Machine Learning*, 1674–1682. PMLR. ISSN: 1938-7228.

Springer Nature. 2022. Climate sciences. Https://www.nature.com/subjects/climate-sciences. Accessed: 2022-07-20.

van Nes, E. H.; Scheffer, M.; Brovkin, V.; Lenton, T. M.; Ye, H.; Deyle, E.; and Sugihara, G. 2015. Causal feedbacks in climate change. *Nature Climate Change*, 5(5): 445–448.

Wang, L.; Scott, K. A.; Xu, L.; and Clausi, D. A. 2016. Sea Ice Concentration Estimation During Melt From Dual-Pol SAR Scenes Using Deep Convolutional Neural Networks: A Case Study. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8): 4524–4533.

Zammit-Mangion, A.; Rougier, J.; Bamber, J.; and Schön, N. 2014. Resolving the Antarctic contribution to sea-level rise: a hierarchical modelling framework. *Environmetrics*, 25(4): 245–264.

Zammit-Mangion, A.; Rougier, J.; Schön, N.; Lindgren, F.; and Bamber, J. 2015. Multivariate spatio-temporal modelling for assessing Antarctica's present-day contribution to sea-level rise. *Environmetrics*, 26(3): 159–177.

Acknowledgments

This research was supported by the Australian Government through the Australian Research Council's Industrial Transformation Training Centre in Data Analytics for Resources and Environments (DARE) (project IC190100031). Furthermore, this research was supported by an Australian Government Research Training Program (RTP) Scholarship.