Monitoring Seasonal Snow Density from Satellite Based Passive Microwave Remote Sensing and Automatic Weather Stations

by

Jeffrey James Welch

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Science
in
Geography

Waterloo, Ontario, Canada, 2023

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis contains a manuscript journal article exploring the potential for satellite based passive microwave remote sensing and operational automatic weather stations to estimate bulk snow density in the Canadian tundra. The manuscript will be submitted to an appropriate journal after the defence of this thesis and will be co-authored with my advisor Dr. Richard Kelly. Dr. Kelly suggested introducing an inertia force into the model in order to constrain the evolution of the snowpack over the course of the season, and also advised on the configuration of the algorithm (described in the manuscript) and editing/word choice suggestions to ensure the proper communication of the relevant information. My contributions being the conceptualization and development of the machine learning model, experimental design and performance analysis, and the creation of the original draft and subsequent versions of this work.

Abstract

Seasonal snow plays an important role in Earth's systems and for hydrological applications one of the most important properties is the quantity of liquid water stored in the snowpack, referred to as snow water equivalent (SWE). SWE is related to the depth and density of a snowpack, so accurate estimates of both those properties are necessary to estimate SWE. However, the current understanding of snow density is limited to sparsely distributed *in situ* samples, which is especially limiting in an environment with restricted access like the Canadian tundra. Models can be used to estimate snow density in lieu of *in situ* sampling and there are a variety of such models available. However, it was determined that none of the available snow density models were entirely suitable for an environment like the Canadian tundra, each for their own reasons.

A new remote sensing algorithm was proposed to estimate snow density from satellite based passive microwave observations and operational automatic weather station (AWS) networks. In this research, an experiment was designed to evaluate the potential for the remote sensing algorithm to monitor snow density in the Canadian Tundra. AWS data were used parametrize a two-layer snowpack model (representing a depth hoar layer underlying a wind slab) and 3D gradient descent machine learning was used to isolate the volume scattering contributions of each layer density independently. New components were added to the machine learning cost function to incorporate prior knowledge and constrain the model's behaviour. The model was trained at the AWS site in Eureka, Nunavut and was then applied to AWS sites distributed across the Canadian tundra. Model performance was quite consistent at high arctic sites but began to degrade across the subarctic with increased distance from the training site, suggesting the need for more robust model training and forcing in the future. Estimation skill consistently improved over the course of algorithm runs and snow density estimates were often close to the ±10% uncertainty range of the *in situ* samples by the end of the season – showing good promise for estimating snow density at peak SWE accumulation, which could be useful for applications where total water storage in the snowpack is of concern.

Acknowledgements

First, I would like to thank my supervisor Dr. Richard Kelly for his incredible help throughout my master's program. One hears stories of supervisors who go above beyond for their students, and I can honestly say that Richard possesses all the outstanding qualities that I was told to look out for. Everyone should to be so lucky to find such a thoughtful and kind academic supervisor.

I would like to thank my parents Dr. Patricia Newcombe and Dr. William Welch for their unwavering support throughout my academic career. My parents have always offered sound advice, supported my decisions, and celebrated my successes, and I could not have accomplished what I have done without them. I hope one day to offer the same qualities to my own children (should my mother's wish be granted and I find Mrs. Right).

To my best friend Duncan Wey, thank you for always being there for me. We've been inseparable since the age of 10 and I sincerely hope that never changes.

Last, I would like to offer a special acknowledgement to my dog Alfie who was faithful by myside for the duration of my master's journey. When we weren't out exploring the parks and trails of Beechwood together to take a break from research, Alfie was curled up under my desk truly earning the title *man's best friend*. In all honesty, Alfie could be listed as a co-author on this work as he was present for every step of the process and made for a great [silent] partner for discussion.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Snow covers most northern landscapes during winter months and as such plays a role in many of Earth's systems. The presence of snow, and its properties, is of concern for a number of natural and anthropogenic applications. Snow affects land-atmosphere interactions in the energy cycle through albedo enhancement and thermal insulation (Cohen & Rind, 1991). Snow is a highly reflective surface so changes to its extent and properties can affect the local energy budget and in turn the response of hydrological systems (Malmros et al., 2018). Snow also has significant ecological implications for flora and fauna (Jones, 1999). Many species in northern landscapes rely on the presence of snow in their habits (Boelman et al., 2019) and changes in snow properties in a changing climate could threaten sensitive ecosystems (Sullender et al., 2023). Snow is also a critical water resource because most winter precipitation held in a frozen state within the snowpack, and has massive financial implications for anthropogenic activities in industries like agriculture and tourism that rely on presence and amount of snow (Sturm et al., 2017). Furthermore, snow cover has become a key indicator of climate change, by the World Meteorological Organization, as the extent of snow-covered areas will almost certainly decrease under a warming climate (Fox-Kemper *et al.* 2021). Thus, it has become increasingly important to characterize seasonal snow to understand how it has changed in recent history and how it may continue to do so in the future. These themes become even more pressing in the arctic

environment where temperatures are increasing at an alarming rate due to arctic amplification (Serreze & Barry, 2011).

Seasonal snow contributes significantly to water storage resources across most of the Northern Hemisphere. Accordingly, large efforts are applied to monitoring seasonal snow conditions for estimating future water availability, and in an historic context for climate analyses. The National Snow and Ice Data Center, at the University of Colorado, defines seasonal snow as "snow that accumulates during one season, [or] snow that lasts for only one season"; the latter definition is more applicable in this regard as most snow in the Northern Hemisphere does not last over the summer (save for that found on ice sheets and glaciers). From a hydrological perspective, one of the most import factors about seasonal snow is the quantity of water stored in the snowpack, referred to as snow water equivalent (SWE). The water stored in the snowpack is released back into the environment through snowmelt and impacts the water supply for the coming season. Snowmelt is a predominant surface water supply for nearly all the Northern Hemisphere above 45° N and the warming global climate threatens a large portion of the global population who rely on it for their livelihood (Barnett et al., 2005). Thus, accurate estimates of SWE are of concern for applications involving water storage/availability, such as, hydroelectric power generation, agricultural irrigation, and flood/drought forecasting. Specifically, those applications rely on accurate spatiotemporal estimates of SWE (Farmer et al., 2009), which can prove difficult to produce due to a lack of data at the necessary scales. The total amount of water stored in a seasonal snowpack can be estimated by calculating SWE at the time of peak accumulation (i.e. April 1st), to provide context on water resource availability (Bohr & Aguado, 2001).

SWE stored in a snowpack can be quantified by its depth multiplied by bulk density (i.e. average vertical density profile), so accurate estimates of both those properties are required to make informed decisions about water resources from seasonal snow. An issue arises when one considers the potential sources for snow depth and density data available to estimate SWE; there exist a variety of methods to directly measure snow depth, but snow density is not directly measurable so it must be estimated (Kinar & Pomeroy, 2015). Instead, a standard volume of snow can be collected and weighed to estimate snow density – a process typically limited to *in situ* manual observations. The process of manually weighing snow samples can be very laboursome and takes considerably longer than measuring snow depth. As a result, the total recorded snow measurements are heavily skewed towards snow depth, and far fewer snow density data are available (Sturm et al., 2010), and so too are many SWE analyses (e.g. Pulliainen et al., 2020). Because of these factors, the current understanding of the variability in snow density is limited to sparsely distributed *in situ* samples. Consequently, many studies regarding the estimation of SWE have focused on snow depth because of the lack of information available about snow density.

One avenue for estimating SWE on large scales is through satellite based passive microwave remote sensing. Passive microwave observations exhibit many qualities that offer a unique perspective for monitoring SWE; however, more research is required to realize its full capabilities (Saberi et al., 2020). The prevalence of snow depth over snow density carries over to passive microwave remote sensing of SWE. To date, essentially all passive microwave algorithms concerning snow mass estimate snow depth, and, in turn, convert to SWE with ancillary snow density data which were not considered in the passive microwave retrieval (e.g. Foster et al., 1997; Luojus et al., 2021; Tedesco & Jeyaratnam, 2016; Xu et al., 2010).

Conversely, estimates of snow density have only been retrieved with passive microwave remote sensing in two ways. Lemmetyinen et al. (2016) used a tower-based L-band (1-2 GHz) radiometer to estimate snow density at the ground-snow interface (i.e. bottom ~10cm of snow) through a refraction-based model. This refraction-based approach was developed for use in parameterizing the effects of overlying snow in soil moisture retrievals (see Houtz et al., 2019; Naderpour et al., 2017; Schwank & Naderpour, 2018). Oppositely, Champollion et al. (2019) estimated the near surface snow density (top ~3 cm) in Antarctica by using the 37 GHz polarization ratio, for surface mass balance modelling of ice sheets. Importantly, neither approach concerns *bulk* snow density which is required for estimating SWE. So, it appears there is a gap in the current understanding of passive microwave remote sensing of snow around density.

This research proposes a novel method to monitor snow density using passive microwave remote sensing. The effective grain size optimization procedure from the European Space Agency's GlobSnow algorithm (described in Takala et al., 2011) was modified to estimate snow density rather than grain size. This approach leverages automatic weather station networks to parameterize a microwave snowpack model so that it can be applied in remote areas using existing infrastructure. In this way, the algorithm has the potential to produce large scale estimates of snow density in areas which were not previously observable. Should the algorithm prove effective, the estimates of snow density could be useful for a range of applications whereas effective grainsize is only relevant for microwave modelling. In general, density estimates from this approach could improve SWE estimates by providing data that were not previously available. In that way, the snow density estimates from independent observations (i.e. passive microwave) could provide better context on total SWE accumulation to provide better

information for hydrological monitoring applications. On the other hand, there are some very specific uses for snow density estimates on their own. For example, snow density estimates from this approach could be used in caribou herd monitoring in the Canadian Arctic, where those data are typically limited. There is a critical snow density threshold (~350 kg/m^3) which impedes foraging and leads to population declines and improved estimates of density could aid in caribou conservation efforts (Martineau et al., 2022).

1.2 Goals and Objectives

The goal of this research is to demonstrate a sensitivity to bulk snow density in satellite based passive microwave observations. Should passive microwave remote sensing prove useful for monitoring snow density, estimates from this approach could provide some understanding of the temporal and spatial variability in snow density at scales not previously observable. To that end, this work presents a novel method to observe bulk snow density, so two objectives are set:

1. Determine if local changes in bulk snow density (i.e. at the site level) can be monitored by satellite based passive microwave remote sensing.
2. Produce synoptic estimates of bulk snow density at scales that were not previously observable.

1.3 Thesis Structure

This thesis is structured as follows. Chapter 2 provides a background review on passive microwave remote sensing of SWE. That review concerns the mechanics of satellite based passive microwave retrievals of snow characteristics and specifics for the study area (i.e. the tundra biome). Chapter 3 is written in the style of a manuscript. It describes the rationale, methodology, and results of the experiment testing the effectiveness of passive microwave

remote sensing for monitoring snow density, structured as a standalone paper – "Tundra Snow Density Estimates from Satellite Passive Microwave Remote Sensing and Automatic Weather Station Measurements" – to be submitted to an appropriate journal after the defence of this thesis. Chapter 4 concludes with a summary of the research findings, with a brief discussion of the broader implications and potential future research on this topic.

Chapter 2

Background

2.1 Passive Microwave Remote Sensing

Satellite based remote sensing systems present an interesting opportunity to monitor snowpack properties, compared to the limited scope of *in situ* sampling. *In situ* samples can provide very detailed information about snow properties but traditional manual methods are difficult to scale over space or time. Conversely, sensors onboard satellites offer a far greater potential for data acquisition in terms of their spatial coverage and continuous observation routines. However, the amount of detail in those satellite data are limited to the spatial resolution of the sensor and the repeat frequency of the orbit. There are a variety of sensor types that can be used to monitor snow characteristics from space, but large-scale estimates of SWE are usual retrieved from spaceborne passive microwave radiometers. A microwave radiometer measures energy emitted by an object as brightness temperature (Tb) – a function of the physical temperature and emissivity of the target – in kelvin. In this context, *passive* refers to the source of radiation being naturally emitted by the target, rather than an active sensor which emits pulses of radiation. As such, passive microwave remote sensing of snow properties concerns measuring radiation emitted by the Earth that interacts with snow present on its surface effecting observed Tb.

2.1.1 Microwave Snow Principles

A number of the characteristics of satellite based passive microwave remote sensing make it aptly suit for providing large-scale estimates of SWE. First, microwaves are mostly uninhibited by Earth's atmosphere, allowing for passive microwave observations under almost all-weather conditions. This quality is frequency dependent where lower frequency microwave channels have a nearly transparent atmospheric window, but higher frequency observations can interact more with the atmosphere attenuating observed Tb (Pulliainen et al., 1993). This quality is very appealing for cyrospheric research because of winter cloud cover, which can prevail over northern latitudes and restrict the effective use of other forms of remote sensing of land surface processes. Nor are passive microwave retrievals limited by prolonged periods of polar darkness, as are other passive sensors such as visible-infrared radiometers. Second, spaceborne passive microwave radiometers benefit from near-total, daily coverage of the Northern Hemisphere, with increasing observation frequency at higher latitudes. However, that wide spatial coverage comes at a cost and microwave radiometers are limited to coarse spatial resolutions because of the relative low energy radiation emitted by the Earth. For example, passive microwave observations are not well suited for SWE retrievals in mountainous areas because of uncertainties associated with the complex terrain variations (and corresponding variations in SWE) within the relatively large observation footprint (Smith & Bookhagen, 2016). Third, the relatively long wavelengths of microwave radiation can penetrate through a snowpack. Most other sensor types are limited to surface interactions with the snowpack, so they are only suitable for discerning snow height (e.g. LiDAR; Deems et al., 2013). On the other hand, penetration by microwaves through the snowpack allows for inferences to be made about its internal composition, thus, permitting SWE retrievals. Finally, the spaceborne passive microwave data record extends back ~45 years in time,

allowing for analyses of relative long-term trends in Northern Hemisphere SWE (e.g. Gan et al., 2013; Goodison & Walker, 1993; Pulliainen et al., 2020).

During early passive microwave experiments, focused on sea ice, unexpected differences in Tb over snow covered scenes were observed, and the presence of snow and its composition were proposed to explain the differences in observed emissivity (Gloersen et al., 1974; Rango et al., 1979). Subsequently, a great effort was put into understanding the contributions of different snow properties in observed microwave signatures (e.g. Chang et al., 1979; Mätzler et al., 1982; Hofer & Mätzler, 1980). Chang et al. (1987) identified the Tb difference (ΔTb) relationship between a high frequency scattering channel (i.e. 37 GHz), sensitive to the accumulation of SWE, and a low frequency background channel (i.e. 19 GHz), mostly unimpeded by the snowpack, to account for the effects of microwave volume scattering within the snowpack. Other approaches using different combinations of microwave channels have been proposed since (e.g. polarized brightness temperature ratio; Goodison & Walker, 1995), but the ΔTb method has become the basis for most current SWE retrievals. In the ΔTb approach, Tb of the scattering channel is attenuated by microwave interactions within the snowpack, while the background channel is mostly unaffected by the snowpack so to provide context on the underlying conditions. Thus, it can be inferred that increases in ΔTb are the result of SWE accumulation and uncertainties associated with other factors are reduced by comparing the relative, rather than absolute, Tb values. It is important to effectively parameterize snowpack properties affecting microwave interactions to make use of the ΔTb approach for estimating SWE.

2.1.2 Microwave Emission Modelling

Microwave radiation, emitted by the Earth's land surface, will interact with the overlying snowpack to different extents depending on its various physical properties. At higher frequencies

(i.e. 37 GHz), the primary interaction between microwave radiation and a dry snowpack is through volume scattering (Chang et al., 1982). Beyond total SWE accumulation, the volume scattering response in a dry snowpack is largely governed by differences in snowpack microstructure and stratigraphy (i.e. snow grain geometry and layering, respectively) (Hall, 1987; Markus et al., 2006). As a result, snowpacks with the same accumulation of SWE but different physical properties can produce different microwave signatures. Therefore, prior knowledge of snow conditions in the observed scene, and how they may affect observed Tb, is essential to isolate the scattering contribution of SWE sperate from other physical properties.

An understanding of the physical interactions between microwave radiation and various media present in the observation scene allows for those process to be simplified and represented through physically based models. Physically based models attempt to simulate known processes within a system, as opposed to empirical models that are deduced from observations. Physically based microwave modelling concerns the use of radiative transfer models (RTMs) to simulate microwave interactions with media. The relevant snow properties need be quantified for use in the RTM and are prescribed as parameters to the model to produce synthetic microwave emissions. The synthetic microwave signature from the RTM can then be compared to satellite observations to draw inferences about snow conditions on the ground. In this way, the RTM can be used to retrieve snowpack properties through inverse modelling. In an inverse modelling approach, all the relevant parameters are prescribed to the RTM except for the variable of interest and an optimization procedure applied on that variable to minimize the difference between synthetic and observed microwave emissions. The solution that best simulates passive microwave observations is the most likely representation of conditions on the ground.

Given the challenges in parameterizing physically based RTMs, increasingly machine learning approaches have been employed to retrieve snow characteristics, either in combination with or in lieu of physically based modelling. Machine learning models do not require the same level of prior knowledge as physical based models because of their ability to identify non-linear relationships in data, allowing for a simpler parametrization process (Yuan et al., 2020). In a very general way, machine learning iteratively applies mathematical models to a set of training data while changing a set of input parameters – called hyperparameters – to identify the best solution to a complex system. Unlike standard model parameters, hyperparameters are intrinsic to the machine learning implementation and cannot necessarily be conceptualized. For this reason, machine learning models are typically seen as black boxes and it can be difficult to explain why certain decisions are made by the model. There have been some successes in pure machine learning implementations of passive microwave SWE retrievals but more interesting are the machine learning approaches that are complemented with physical knowledge (e.g. King et al., 2020; Xue et al., 2018; Yang et al., 2021). These synergistic approaches present an opportunity to exploit the emerging benefits of machine learning while making informed decisions by retaining physically based knowledge of geophysical systems in the model.

2.1.3 Snow Electromagnetic Properties

It is important to understand the effects of the relevant snow properties on passive microwave observations to properly utilize the RTM. In the next chapter, the Dense Media Radiative Transfer model for multi-layered snowpacks (DMRT-ML; (Picard et al., 2013)) is used to simulate microwave emissions. As such, the relevant snowpack parameters in DMRT-ML are discussed for their contributions to synthetic microwave signatures produced from the RTM.

Snow Depth

Snow depth – the most ubiquitously measured snow property – significantly impacts microwave interaction within the snowpack. The depth of a snowpack is analogous to the number of grains in the energy's path (Kelly, 2009), where deeper snow conditions increase the possible number of interactions with individual snow grains and exhibit greater volume scattering. On one hand, shallow, dry snow is essentially transparent to microwave radiation and, in turn, a minimum of 5-10 cm of is generally required to detect snow with spaceborne microwave observations (Hall et al., 2002). Conversely, there is also a maximum retrievable snow depth from passive microwave observations because less radiation from the underlying surface will be transmitted through the snowpack as it deepens and exhibits a greater volume scattering response. The positive correlation between ΔTb and SWE holds up to a threshold (generally considered to be ~150mm) when the contributions of emissions originating with the snowpack overtake those of the underlying surface, and the ΔTb-SWE relationship inverts (Saberi et al., 2020). This effect is called saturation and can occur at snow depths of 50-100 cm for the 37 GHz channel depending on other properties of the snowpack (Kelly et al., 2003).

Effective Grain Size

The composition of snow microstructure is typically represented with the Stick Hard Spheres model in DMRT-ML (Lowe & Picard, 2015), where the size of snow grains (i.e. grain radius) is expressed as the effective grain size. Microwave interactions increase with effective grain size, allowing for more efficiency volume scattering by a snowpack with larger grains. Effective snow grain size has a large influence on the observed Tb of a snowpack, and inversion models can provide very different outputs based on the parameterization of grain size (Chang et al., 1982; Pulliainen, 2006). This presents an issue when parameterizing an RTM because snow

grain size estimates are limited to *in situ* sampling methods (Molotch et al., 2016) and are not available at the scales required for use in these models. Furthermore, effective grain size is a modelling term and is not the same as the optical grain size one might estimate in the field through visual examination (if such measurements are available). For this reason, grain size can be treated as a tuning parameter in the RTM – as in the GlobSnow method (Takala et al., 2011), where grain size is optimized given the snow depth measured at a weather station (and a set of static parameters). Alternatively, empirical relationships can be used to estimate relative changes in grain size and simulate how its contribution to scattering might change over the winter season (e.g. Kelly et al., 2003, 2019).

In DMRT-ML there is a second parameter associated with grain size called stickiness. Stickiness is an arbitrary, unitless, unmeasurable parameter which describes the likelihood of individual grains accumulating into clusters. Those grain clusters appear larger to microwaves than individual grains and increase the scattering efficiency of the snowpack (Picard et al., 2013). Similar to effective grain size, the stickiness parameter can be treated as a calibration parameter (e.g. Larue et al., 2018b). Alternatively, a non-sticky microstructure model can be implemented in DMRT-ML to negate the effect of stickiness because of the arbitrary nature of its parameterization.

Snow Density

Snow density has a somewhat counterintuitive effect in DMRT-ML. One might expect greater volume scattering within a snowpack with increasing density (as with increasing SWE) but instead it is the opposite for fractional volume of <50% (i.e. bulk density ~450 kg/m^3). In Dense Media theory, sparsely distributed grains act more like individual scatters than those packed closer together in a dense media (Picard et al., 2013), so a sparser snowpack will exhibit

more volume scattering than a denser one. An analogy for the volume scattering effect of snow density can be made with fiberglass insulation – which provides the best insulative properties when packed loosely in a cavity and loses insulative efficiency when more insulation is densely packed in the same space. The behaviour is reversed for snow densities above 50% volume fraction and volume scattering increases with density beyond that threshold, because of how the snow microstructure is represented in DMRT (representation switches from spheres-of-ice-in-air to spheres-of-air-in-ice above 50% volume fraction; Picard et al., 2013) . However, the density of seasonal snow is typically below 50% volume fraction (Sturm et al., 2010), so the effects of snow density above that threshold can largely be ignored in the model.

The given description of density parameter is for a non-sticky DMRT-ML implementation (as used in Chapter 3), and the stickiness parameter change the effect of snow density on volume scattering (Picard et al., 2013).

Snow Temperature

Snow temperature is different from the aforementioned parameters in DMRT-ML because it does not significantly impact microwave volume scattering. Instead, snow temperature is considered in the RTM to estimate emissions originating in the snowpack. This source of emission from the snowpack is important when considering higher microwave frequencies (i.e. 37 GHz) in deeper snow due to the effects of saturation. Under those conditions, the contributions from snowpack emission surpass that of volume scattering and Tb begins to increase with SWE accumulation (Saberi et al., 2017; Sturm et al., 1993). Therefore, the snow temperature parameter is important for accurately estimating the different sources of microwave emissions to isolate the volume scattering signal from the snowpack.

2.2 State of Knowledge of Tundra Snow

In the arctic region, the tundra biome is characterized by its very cold temperatures, short vegetation/lack of trees, and low precipitation (Terasmae & Reeves, 2017). Strong arctic winds redistribute most snow that is deposited in tundra, removing much of the snow from open areas (Pomeroy et al., 1997). Vegetation plays a large role in the spatial distribution of SWE over the tundra by trapping snow and preventing the redistribution of snow by wind (Sturm et al., 2001). As a result, the open tundra environment (with limited vegetation) is typical covered with a veneer of relatively shallow snow, with most of the deposited snow blown away to be caught in large drifts caused by features on the landscape (Benson & Sturm, 1993). That veneer snow (hereafter referred to simply as tundra snow) has a characteristic profile of, at least, one dense wind-packed slab layer overlying a lower-density depth hoar layer. On the other hand, drifts can be many times deeper than in the open tundra, and considerably denser; with the potential to hold the majority of SWE deposited on the landscape despite their relatively small areal extent (Marsh & Pomeroy, 1996). Ultimately, in some ways the composition of tundra snow is quite predictable owing to the *in situ* environmental controls.

2.2.1 Temporal Evolution of Tundra Snow Density

The typical evolution of tundra bulk snow density over time is described by Sturm & Holmgren (1998), and is as follows. New snow at the beginning of the season is very light when first deposited and is quickly mechanically compacted by tundra winds. The rate of initial compaction is dictated by local wind patterns and the characteristic strong winds over the open tundra result in relative high snow densities early in the season (compared to other biomes). Then, the snow gradually densifies over the course of the season as it undergoes metamorphosis. In summary, tundra snow is quite dense at the beginning of the season (initial compacted by the

wind) and should exhibit conservative fluctuations in bulk density over the remainder of the winter season.

Snow deposited in the tundra takes time to develop into the characteristic two-layer snowpack of depth hoar underlying wind slab. The top layer of snow is quickly compacted by the mechanical force of wind moving across the snow surface (Colbeck, 1982), forming a densely packed wind slab layer on the top of the snowpack soon after its initial deposition (Benson & Sturm, 1993). Whereas the lower depth hoar layer is the result of strong temperature gradients within the snowpack due to low arctic air temperatures and the insulative properties of snow (Colbeck, 1982). That temperature gradient produces a convective current which transports mass (in the form of water vapour) from lower to upper snow layers (Sturm & Benson, 1997). The stratigraphy of the snowpack can be complicated by melt-refreeze events where liquid water (originating in the top layers of the snowpack) will percolate down and accumulate at the boundary between distinct snow layers and form ice crust at those interfaces (Marsh & Woo, 1984). Additionally, the stratigraphy is further complicated as new snow is deposited on top of the existing snowpack resulting more complex layering with snow at various stages of metamorphism. Thus, the general composition of tundra snow is quite predictable, but can vary through local weather patterns.

2.2.2 Spatial Variability of Tundra Snow Density

As mentioned before, the current understanding of snow density is limited to *in situ* sampling. This is especially restrictive with regard to studying the spatial variability of snow density because of the heterogenous distribution of sampling locations; even more so in the tundra environment where SWE sampling location are especially sparse in northern Canada (Brown et al., 2019). Woo & Marsh (1978) studied snow storage of small basins in the high

arctic and found snow depth to much more variable than density across different terrain types (e.g. hilltops, slopes, valleys, etc.), concluding those local variations in density are only relevant in very small basins. In the absence of more information, the spatial variability of snow density is believed to be quite conservative on larger scales (at least much more so than snow depth; Fassnacht et al., 2010; López-Moreno et al., 2013; Sexstone & Fassnacht, 2014). This assumption of the relative conservativeness of snow density causes it to be generalized over large geographic areas (typically represented as the mean value of available density samples).

More recent efforts have attempted to characterize the spatial variability of tundra snow properties (one of which being density) through extensive field campaigns. Derksen et al. (2009) were the first to coordinate large-scale sampling efforts across the Canadian subarctic tundra (over a transect spanning Northwest Territories to Nunavut) to analyse the spatial distribution of tundra snow at a scale not previously explored. The study was subsequently expanded to include samples from Ellesmere Island, Nunavut (near Eureka) in the Canadian High Arctic (Derksen et al., 2014). Derksen *et al.* (2014) found snow density to be evenly distributed in the high arctic and normally distributed in the subarctic; additionally, snow density in the high arctic displayed more spatial variability than in the subarctic tundra/taiga. Hannula et al. (2016) found comparable spatial distributions of snow density in a similar field campaign in the Finnish Arctic.

The results of recent field campaigns (i.e. Derksen et al., 2014; Hannula et al., 2016) seem to challenge the suggestion that spatially uniform snow density estimates are sufficient for estimating SWE (Derksen et al., 2005a). Specifically, an average value may not be well suited for a high arctic setting where snow density is not normally distributed, so the mean would not be representative. This argument can be taken a step further in passive microwave context

because those same studies found the majority of tundra SWE to be stored in the wind slab layer, but the slab layer provides a negligible contribution to microwave volume scattering (Sturm et al., 1993). Thus, differences in wind slab density (and in turn SWE) would not be captured by passive microwave models concerning volume scattering. So, it can be posited that better snow density estimates in the tundra environment could improve SWE estimates from passive microwave observations.

2.2.3 Tundra Snow in a Passive Microwave Context

In a remote setting, such as the Canadian tundra, passive microwave remote sensing provides the opportunity to monitor snow in areas unsuitable for *in situ* sampling. Such an application of remote sensing benefits from detailed records of the composition of tundra snow that allow for increased confidence in geophysical modelling. Understanding the effects of snow conditions is essential for interpreting passive microwave observations in the tundra. Hence, a significant effort has gone into incorporating the understanding of the tundra landscape into passive microwave SWE models (e.g. Derksen et al., 2006, 2010; Derksen et al. 2005b; Meloche et al., 2022; Rees et al., 2010; Vargel et al., 2020).

Passive microwave observations of tundra snow typically display large ΔTb that cannot be explained by snow depth alone. Tundra snow is typical shallow, in comparison with other biomes, yet exhibits a strong volume scattering response typical of a much deeper snowpack. This unexcepted microwave signature can be attributed to the presence of depth hoar which greatly increases volume scattering within the snow (Hall, 1987; Sturm et al., 1993). The characteristics of depth hoar snow (i.e. large, cup-shaped, sparsely packed grains) contribute to a very strong microwave interaction and stronger attenuation of the scattering channel due to increased volume scattering. More recent work has reinforced the need for a two-layer snowpack

for microwave modelling and has focused on how to characterize tundra snow for use in such models (Saberi et al., 2017, 2021). It is now considered that a two-layer snowpack model is desirable to adequately represent the composition of tundra snow in a passive microwave SWE retrieval context.

2.3 Incorporating Snow Density into Microwave Models

2.3.1 Snow Density in Current Microwave Modelling

There is a lot that is still unknown about snow density in microwave modelling. Typically, snow density estimates are only available from sparsely distributed *in situ* samples, so there is little known about the spatial variability of snow density. This is especially true in the tundra environment where access for *in situ* sampling is restricted, so fewer snow density samples are available there compared to other biomes. Hence, the general composition of tundra snow is well documented, but less is known about variability across the domain. This knowledge gap about snow density introduces considerable uncertainties in generating spatial distributed density estimates, at the scales necessary for microwave modelling, causing most density estimates to be spatially uniform.

Snow density has generally been overlooked by passive microwave remote sensing models because of a lack of spatial information to parameterize retrievals. Early models assumed a spatially and temporally constant snow density as a compromise for computational efficiency (e.g. 300 kg/m^3 in the original Chang algorithm; Chang et al., 1987), and many current algorithms still hold that assumption (e.g. 240 kg/m^3 in GlobSnow V3.0; Luojus et al., 2021). The use of a constant snow density in microwave retrievals contributes to overestimation of SWE in the early season and underestimation later on (Luojus et al., 2021). Post-processing of passive microwave SWE estimates (retrieved with a constant density) with dynamic density

estimates can help to reduce that systematic bias (Venäläinen et al., 2021). However, post-processing of SWE estimates only impacts the final SD to SWE conversion and errors from assuming a constant snow density in the microwave retrieval process would still exist (Mortimer et al., 2022).

Kelly et al.'s. (2003) prototype AMSR-E algorithm was one of the first to introduce dynamic snow density estimates into the passive microwave retrieval process, albeit with a very simple densification model. The successive AMSR-2 SWE algorithm (Kelly et al., 2019) went a step further by incorporating snow density estimates from Sturm *et al.* (2010) to introduce distributed estimates of snow density into the SWE retrieval process. Those two algorithms were the only such identified to explicitly consider the effect of snow density in the passive microwave retrieval process (excluding regional analyses like Larue et al., 2018a). Recently, there has been a focus on introducing dynamic snow density estimates into passive microwave SWE algorithms, specifically the European Space Agency's Snow Climate Change Initiative (Mortimer et al., 2022; Venäläinen et al., 2021). However, as identified in section 2.2.4, the available sources for snow density estimates are limited and the most reliable estimates rely on *in situ* sampling which is not feasible for very large scales of analysis.

2.3.2 Available Snow Density Models

There are limited options to consider for distributed snow density estimates at the scales necessary for passive microwave remote sensing. One of the most prevalent ways to estimate snow density is through physical snowpack models. Those snowpack models are forced with the reanalysis of meteorological data to estimate changes in the physical properties of the snowpack once it is deposited. These snowpack models can be included in large land surface models as sub-models (e.g. ERA5 (Hersbach et al., 2020) and MERRA-2 (Gelaro et al., 2017)) or as

standalone comprehensive models (e.g. Crocus (Brun et al., 1992) and SNOWPACK (Bartelt & Lehning, 2002). Such reanalysis models have been shown to provide decent estimates of SWE on large scales (Mortimer et al., 2020). However, these models rely on very detailed meteorological data, and uncertainties in model outputs can be introduced through the errors associated with forcing data (Raleigh et al., 2015) or from different forcing datasets (Raleigh et al., 2016). Additionally, the physical processes considered in the available snowpack models (i.e. Crocus and SNOWPACK) are limited to heat and mass fluxes which do not properly simulate the expected tundra snow density profile by inverting positions of the wind slab and depth hoar layers (Domine et al., 2019).

Alternatively, Sturm *et al.* (2010) offers an empirical model which is one of the most straightforward methods for estimating spatially distributed snow density. That model estimates snow density without the need for detailed meteorology or ancillary data. Instead, temporally and spatially distributed estimates of snow density are regressed from day of the year (DOY), snow depth, and location (i.e. snow class; Sturm et al., 1995),. Snow density estimates from Sturm *et al.* (2010) are quite conservative in the Arctic and do not vary much over the area. The model does provide suitable temporal trends for the tundra snow class, by considering the day of year, but the spatial detail in the tundra is limited. The snow depth parameter not as significant to the snow density estimates, due to the relatively shallow snow found in the tundra, resulting in largely uniform estimates over most of the tundra environment. Even more conservative are the snow density estimates for the taiga snow class in Sturm *et al.* (2010), with a spatially and temporally static value of 217 kg/m^3 (regardless of DOY or snow depth).

Quite recently, Wang et al. (2023) developed a machine learning model to estimate snow density in China. That model is based on a geographically and temporally weighted neural

network (GTWNN) that is forced by meteorological data, topographic characteristics, and satellite snow products (snow albedo, snow depth, and snow cover extent). The GTWNN model was found to perform well compared to other regression models (e.g. multiple linear regression, geographically weighted regression, etc.) and ERA-5 snow density estimates. However, the GTWNN model included 20 input variables, all with weak influence on snow density, so it is difficult to explain why the model performs as it does. This is compounded because neural networks are typically regarded as black boxes, especially with so many dimensions (20), so it is very difficult to conceptualize how decisions are made. The described influence of the snow cover duration variable specifically stood out as odd because it was found to be negatively correlated with snow density in the GTWNN model. That seems counterintuitive because snow is expected to densify over time (Sturm et al., 2010). Additionally, the snow conditions in China (at least those described by Wang *et al.* (2023)) appear to be quite unique, consisting of both seasonal and ephemeral snow, characterized by very low observed snow densities (seemingly much lower than those found elsewhere). For these reasons, it appears the GTWNN model could prove useful for estimating snow density in those very specific conditions but needs further refinement before it can be applied more broadly.

So, it appears none of the described snow density models would be entirely well suited to the Canadian Arctic, each with their own shortcomings. Instead, this research proposes a novel method for estimating bulk snow density from satellite based passive microwave remotes sensing and operational AWS networks. The new algorithm would benefit from independent data (i.e. passive microwave observation) which are known to be related to snow density. Ideally, those independent observations should provide more specific information on the state of the snow density in the scene, compared to more general approaches in some of the existing methods. The

next chapter describes an experiment to test the potential for the proposed methodology to estimate bulk snow density in a tundra environment and compares its skill to the applicable methods used to estimate snow density for passive microwave SWE retrievals.

Chapter 3

Tundra Snow Density Estimates from Satellite Passive Microwave Remote Sensing and Automatic Weather Station Measurements

Abstract

Current methods for estimating bulk snow density are largely limited to *in situ* sampling and are difficult to scale spatially or temporally. Therefore, density is difficult to incorporate into snow water equivalent models and is typically generalized for model forcings. This research proposes a novel method to estimate bulk snow density using satellite based passive microwave observations and automatic weather stations (AWS) measurements. AWS measurements are used to parameterize a two-layer snowpack model (composed of a lower depth hoar layer and an upper a slab layer) to force a 3D gradient descent machine learning model that prescribes independent density estimates for each snow layer. Prior knowledge of the tundra environment was used to inform the machine learning model and constrain the algorithm's progression through the season. Climatology data are used to introduce a dynamic inertia force into the machine learning model to constrain fluctuations in density estimates. Similarly, a logical rule was established to prevent physically unlikely combinations of layer densities despite the simulated microwave response. The Canadian tundra was chosen to prototype this algorithm – validation and testing results are encouraging and demonstrate a sensitivity to bulk snow density in passive microwave observations. This method could provide synoptic estimates of bulk snow density across very large domains, to provide information at scales that were not previously observable. Future work will attempt to improve the algorithm and expand the scope of the analysis to other sites and biomes.

3.1 Introduction

Monitoring snow water equivalent (SWE) is important for sustainable water management. The SWE of a snowpack is a function of its depth and bulk density (i.e. average vertical density profile). Therefore, accurate quantification of both these phenomena is critical to make informed decisions regarding the management of SWE resources. Snow depth can be represented as a distance and there exist a number of methods to directly measure it (Kinar & Pomeroy, 2015). On the other hand, snow density must be estimated because there is a lack of instrumentation to measure it in an operational context. Traditionally, a standard volume of snow is collected and weighed to estimate snow density. That manual process can be very laboursome and difficult to increase the number of samples either spatially or temporally. As a result, current snow density estimates are limited to sparsely distributed *in situ* measurements which are used to predict spatially distributed density. Such efforts lead to uncertainties regarding the spatial representativeness of snow density in remote places where measurements were not collected.

Uncertainties in snow density estimates can lead to issues in quantifying spatiotemporal distribution in SWE (Luojus et al., 2021; Venäläinen et al., 2021). Therefore, improved snow density estimates could improve overall SWE quantification, and reduce uncertainties in estimating seasonal water storage within the snowpack. Additionally, there are applications where snow density has direct implications and estimates from this approach could be of use; for example, snow density is an important factor for caribou herd monitoring because snow that is too densely packed prevents foraging for food (Martineau et al., 2022). So, it would appear there is no shortage of uses for snow density data, but rather a lack of sources for those data.

3.2 Background

Earth observing passive microwave remote sensing is concerned with measuring naturally emitted radiation from the Earth in the microwave spectrum, to retrieve characteristics about the target. The microwave radiation measured by the radiometer is a function of a medium's physical temperature and its emissivity, expressed as brightness temperature (Tb) in kelvin. Understanding of the processes that govern the observed microwave signature enables inferences to be made about the conditions present in the scene. Early on, relationships were identified between microwave Tb and the presence of snow cover and its physical properties (Hofer & Mätzler, 1980; Rango et al., 1979). Thus, snow properties can be monitored remotely by observing Tb of a snow-covered scene, given adequate prior knowledge of processes effecting microwave interactions.

Understanding the interactions between microwave radiation and various media allows those processes to be generalized into mathematical representations within physically based models. Radiative Transfer Models (RTMs) attempt to represent the microwave interactions within a scene and produce synthetic microwave emissions that the radiometer might observe (e.g. Dense Media Radiative Transfer Multi-Layered model (DMRT-ML); Picard et al., 2013). The synthetic microwave emissions produced by the RTM are compared with the satellite observations to infer conditions on the ground. Historically, passive microwave models have used a frequency difference approach (ΔTb) by subtracting the Tb of a channel sensitive to snow volume scattering from a lower-frequency, background channel sensitive to underlying conditions to estimate SWE/snow depth (Chang et al., 1987). The relevant snowpack properties need to be parameterized for use in the RTM and the variable of interest can be optimized to minimize the difference between synthetic and observed ΔTb. Those snowpack parameters are

not necessarily straightforward to assign to the model and can lead to uncertainties in model outputs.

In general, the influence of snow density on passive microwave observations has been overlooked in large-scale SWE estimation models (Venäläinen et al., 2021). There is a lack of snow density observations at the necessary scales to constrain density parameterization, primarily because of the difficult in acquiring spatially distributed *in situ* observations. As a result, snow depth has been the focus of most analyses regarding SWE. In some cases, snow density is kept as a spatially and temporally constant across the domain (GlobSnow v3.0; Luojus et al., 2021) or conservative estimates are taken from empirical models of snow density evolution over time (Kelly et al., 2003). However, these simplified representations of snow density may not adequately represent variability across the domain at the scales in question – i.e. across the entire Canadian tundra.

This study proposes a novel method to observe snow density on large scales by leveraging existing observation networks. The optimization procedure at points of measured snow depth (at weather stations) from the GlobSnow approach (Pulliainen, 2006; Takala et al., 2011) was adapted to produce estimates of snow density, rather than effective grain size. Retrieving snow density at those points could produce data useful for a number of hydrological, thermodynamic, or ecological applications, whereas effective grain size is only relevant with respect to RTMs. Beyond informing snow density parameterization for independent SWE retrievals, these estimates could provide information about variation in snow density at scales not previously observed.

3.3 Study Area

Eight study sites were identified in the Canadian tundra, and are mapped in in Figure 3.1. These stations are further categorized into high arctic and subarctic according to latitude following the study by Derksen *et al.* (2014). Basic characteristics of air temperature, snow depth, and density data are included for each site in Table 3.1. The tundra biome was chosen to develop a prototype snow density retrieval algorithm for the following reasons that tend to simplify the retrieval process.

First, the relatively simple landscape of the tundra environment mitigates many of the sources of uncertainty inherent to passive microwave retrievals of snow properties in other biomes or landscapes. For example, the effects of complex terrain such as that found in alpine landscapes (Tong et al., 2010), and attenuation effects from forest cover (Li et al., 2020) is minimized in this region because of the flat, sparsely vegetated environment (Woo, 1998). However, the Canadian Arctic is vast, and the landscape can vary, and there are some exceptions to the aforementioned characteristics. Three sites (Alert, Eureka, and Resolute) were situated within the Arctic Cordillera ecozone but it was determined that differences in topography were minimal within the satellite observation footprints. Additionally, the Inuvik site is situated on the border of the tundra and taiga ecozones and is characterized by a more varied landscape (including forested areas) which increased the uncertainties in the passive microwave retrievals at that site.

Table 3.1 – Basic characteristics for each automatic weather station (AWS) site in the study area. Temperature (C) and snow depth (cm) from AWS measurements and density statistics from in situ reference dataset (CanSWE). Number of density samples in parathesis indicates number of samples included in performance analysis. High arctic sites are those above 70 N, and subarctic are those below, to match the classification by Derksen et al. (2014).

Class	Site	Avg. Temp	Avg. SD	Density Samples	Avg. Density	Std. Density	Min/Max Density
High Arctic	**Alert**	-15.5	31.5	139 (66)	362	51.1	188/451
	Eureka	-17.4	15.2	117 (89)	331	54.0	143/436
	Resolute	-14.4	15.3	116 (105)	383	81.1	217/690
Subarctic	**Cambridge Bay**	-13.2	24.6	288 (235)	330	54.6	180/560
	Coral Harbour	-9.6	21.9	120 (109)	400	52.8	283/598
	Iqaluit	-7.9	14.8	81 (43)	323	54.8	184/469
	Baker Lake	-9.8	11.1	108 (60)	260	63.5	127/371
	Inuvik	-7.4	57.1	93 (84)	188	43.2	69/285

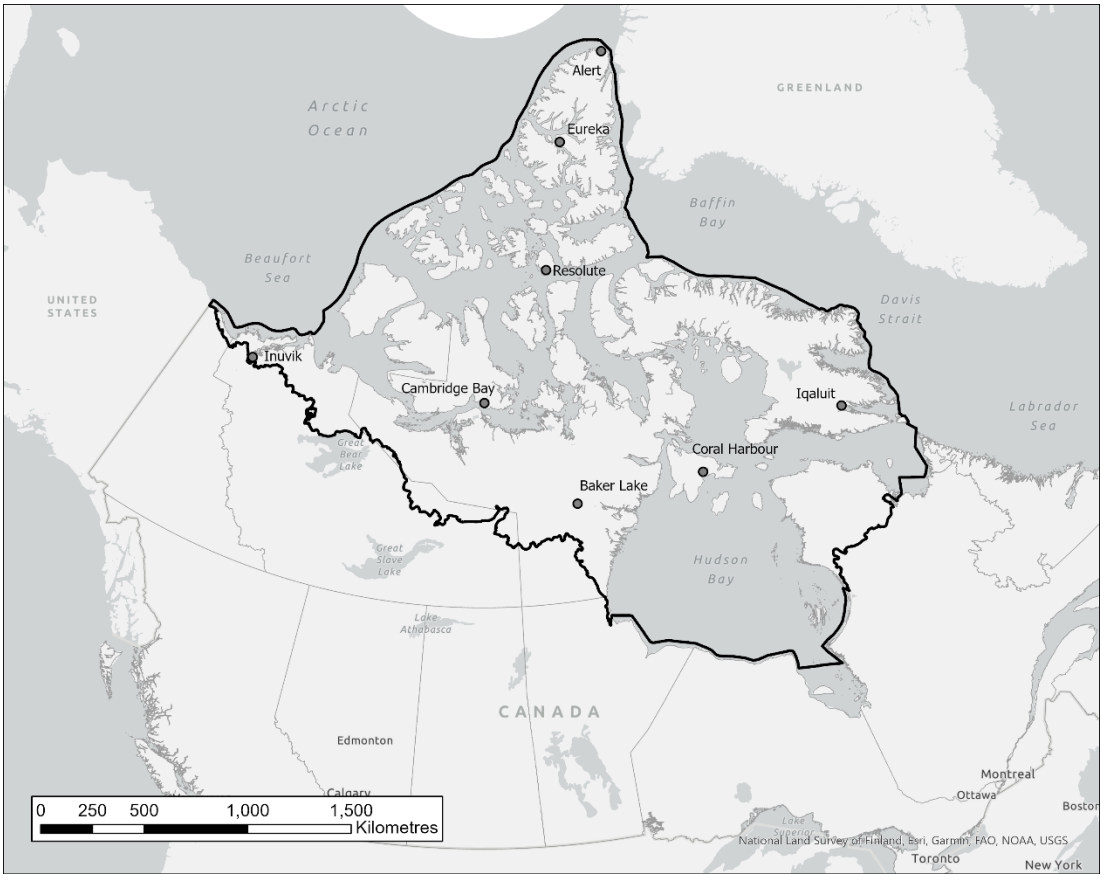


Figure 3.1 – Locations of AWS sites, distributed across the Canadian tundra (black line), used to force algorithm to estimate bulk snow density.

[ecozones shapefile] Government of Canada (n.d.). National Soil Database Ecological Framework for Canada, Retrieved from https://sis.agr.gc.ca/cansis/nsdb/ecostrat/gis_data.html

[base map] "Light Grey Canvas Background", sources: Esri, DeLorme, HERE, MapmyIndia

Second, the composition of Tundra snow is well documented (e.g. Benson & Sturm, 1993) allowing for informed decisions to be made regarding its parameterization for modelling purposes. This prior information was critical to isolate microwave volume scattering contribution of snow density because its effect is relatively low compared to the other characteristics (i.e. snow depth, grain radius, and layering).

Third, the tundra environment is aptly suited for remote sensing observations due to the practical limitations of *in situ* sampling there. This presented a compelling rationale to test this approach due to large gaps in existing SWE sampling networks in Northern Canada (Brown et al., 2019). In the future, density estimates from this approach could be used to monitor snow conditions in remote areas where it is unfeasible to manually sample snow.

3.4 Data

3.4.1 Model Forcing Data

The main forcing data for the algorithm were satellite based passive microwave observations to produce estimates of bulk snow density. Passive microwave radiometry data were acquired from the Advanced Microwave Scanning Radiometer Earth Observing System (AMSR-E) Level-2A (L2A) product, onboard the Aqua satellite (Ashcroft & Wentz, 2013). AMSR-E L2A provides twice-daily observations on a global scale, with increasing observation frequency at high latitudes, on a 25 km Equal-Area Scalable Earth (EASE) Grid. AMSR-E observations for each station were acquired from the EASE grid cell that contained the coordinates of the AWS site. Nighttime observations (~1:30 am local time at the equator) from the descending orbit track were used for the analysis, so the snowpack was most likely to be in a cold and dry state to facilitate microwave retrievals (wet conditions are unsuitable for the ΔTb approach; Chang et al., 1987). The 18.7 (19) and 36.5 (37) GHz vertical channels are used in the

ΔTb calculation to retrieve snow density. Additionally, a number of other bands were used for the detection of snow as described by Grody & Basist (1996).

The algorithm also used weather station measurements for forcing, acquired from the Environment and Climate Change Canada automatic weather station (AWS) network (Environment and Climate Change Canada & ClimateData.ca, n.d.). The AWS data were used to parameterize the electromagnetic snowpack model, where daily measurements of snow depth and air temperature provided prior knowledge of the snow conditions in scene. The AWS data were the limiting input factor in this approach, for a couple reasons. The AWS network is very sparsely distributed in Canada and is biased towards southern, more populated, areas limiting the number of sites this algorithm can be applied. Additionally, data availability is not always consistent at AWS sites because sensor problems and outages are an ongoing issue when maintenance schedules are restricted (as is this case for some northern sites).

3.4.2 Reference *In situ* Dataset

Reference measurements are necessary to evaluate the algorithm estimates of snow density. The ECCC Canadian Historical Snow Water Equivalent (CanSWE; Vionnet et al., 2021) dataset was used to provide the *in situ* snow density data. The CanSWE dataset includes a considerable number of sampling locations collocated with AWS sites which enabled a straightforward comparison of estimated and measured bulk snow density. The *in situ* CanSWE density data were collected with ESC-30 SWE Tube samples taken over standard 10-point snow courses and aggregated into a single estimate of snow depth and density. The locations specified in the CanSWE dataset for *in situ* measurements were the coordinates of the closest AWS. Thus, it is unclear where the in situ measurements were taken relative to the AWS. No additional

metadata about sampling procedure or site descriptions were available about the individual sites in the CanSWE dataset.

3.5 Methods

3.5.1 Snowpack Electromagnetic Model

The physically based modelling approach required the snowpack to be parameterized, and the relevant snowpack characteristics needed to be quantified. The DMRT-ML model, implemented with the Snow Microwave Radiative Transfer Model (SMRT) Python library (Picard et al., 2018), was adopted for use in this study. A two-layer snowpack model was configured to account for the presence of depth hoar underneath a slab layer to best represent the microwave signature of Tundra snow (Hall, 1987; Saberi et al., 2021). The composition of depth hoar (i.e. large, loosely packed grains) permits a relatively stronger interaction with microwave radiation, than the slab snow layer, and increased scattering efficiency of the snowpack. Thus, it is important to consider the presence of depth hoar because tundra snow will exhibit the microwave signature of a much deeper snowpack if this is not accounted for.

The two-layer snowpack model representation was parameterized by AWS station observations. The snow depth forcing variable was prescribed directly from the AWS observations, divided into the relative depths of the two layers. A fixed 1:2 ratio of depth hoar to slab layer depth (Saberi et al., 2017) is applied to the AWS measurement, with a maximum 10 cm of depth hoar allowed. The minimum daily air temperature from the AWS was used as a surrogate for snow temperature and was prescribed directly to each layer. Additionally, the temperature of the underlying substrate was set 5C higher than the minimum air temperature to account for the insulative properties of snow.

The microstructure model in DMRT-ML required estimates of the effective radius of the grains in the snowpack which are not acquired by operational AWS measurements. Therefore, a modified version of the grain size model from Kelly et al. (2003) was used to prescribe effective grain radius but with two notable distinctions. First, due to the high latitudes and cold temperatures of the study area, kinetic snow grain growth was assumed to dominate and the effects of equilibrium grain growth were ignored. The Sturm & Benson (1997) kinetic grain growth model was applied to estimate increases in snow grain size based on the number of days since deposition. Second, different grain growth trajectories were considered because of the two-layer nature of the snowpack model. The parameters for the "upper" and "lower" snow layers from Sturm & Benson (1997) were applied to the wind slab and depth hoar snow layers, respectively. The radii range for each of the layers in the grain growth model were scaled to reflect the effective grain radius in DMRT.

3.5.2 Machine Learning Application

The two-layer snowpack model presented an issue when prescribing independent density estimates for each layer, so machine learning was employed to simplify this issue. A three-dimensional (3D) gradient descent model was designed to isolate the influence of the distinct layers in the synthetic microwave emissions (an overview of gradient descent machine learning is provided in Aggarwal, 2020)). The gradient descent model (hereafter the model) estimates the *cost* associate with a given parameter combination (i.e. snow density profiles) as define by the *cost function*. This cost function is used to evaluate possible parameter combinations and is ultimately the only metric the model has to determine which of those combinations is most likely. At any given point, the model estimates the cost associated with the current combination of snow layer densities and, also, estimates the instantaneous rate of change in cost with respect

to the density of each snow layer. Then, the model identifies the next density estimate for each snow layer according to their respective change in cost and a prescribed learning rate (i.e. step size). The model calculates the cost associate with the new parameter combination and repeats the process as cost continues to decrease, until an exit condition is met and the process stops. The parameter combination with the lowest associated cost is the most likely representation of the scene when the observations were acquired.

Fundamentally, a machine learning model has no information available outside that upon which it is explicitly programme – to minimize outputs from the cost function. As such, any aspect of the scenario the model should consider must be included in the cost function to allow for evaluation against other possible solutions. The most basic cost metric (J) for this passive microwave remote sensing model is:

$$J(\rho_{slab}, \rho_{hoar}) = \Delta Tb_{sim}(\rho_{slab}, \rho_{hoar}) - \Delta Tb_{obs} \tag{3.1}$$

representing the difference between the observed microwave signature (ΔTb_{obs}) and the simulated signature (ΔTb_{sim}) given the prescribed wind slab (ρ_{slab}) and depth hoar layer densities (ρ_{hoar}) where:

$$\Delta Tb = Tb_{19,V} - Tb_{37,V}\ [k] \tag{3.2}$$

As a result, the cost associated for a given snow density estimate would be limited to the physical processes considered in the RTMs (i.e. volume scattering). This presented an issue in the two-layer snowpack model because different layer density combinations can produce the same synthetic emissions, resulting in a system with no global minima.

This equifinality issue in practice meant that the machine learning model was confronted by seemingly equally valid layer density combinations producing the same microwave signature.

Consequently, the algorithm's progression can be thrown off by the solution jumping between local minima. Further, the model's progression in day-to-day density estimates can vary in response to perturbations of the system from variations in forcing data that cause the solution to shift between local minima. Therefore, without any constraints the model can find solutions across an unfeasibly vast range of density combinations, corresponding to unrealistic day-to-day fluctuations in snow density. Intuitively, day-to-day changes in snow density should be mostly conservative (especially in the tundra environment; Sturm & Holmgren, 1998) and those fluctuations (on the order of hundreds of kilograms per day) experienced in the model would not represent reality. Thus, the potential snow density estimates at a given iteration of the algorithm run should be related to the estimate from the previous iteration (i.e. previous day).

3.5.3 Climatological Inertia Constraint

The climatological snow density record at the training site (Eureka) was examined to better constrain the model's evolution of snow density estimates over the course of a season. That is, the degree of change that should be allowed between subsequent algorithm estimates of bulk snow density and how it may change over time. Estimates of daily change in bulk density were calculated given the difference between subsequent *in situ* CanSWE measurements divided by the number of days between them. The CanSWE record for Eureka extends back into the 1960's, however, only observations from 2003-21 were considered due to a discontinuity in the data between 2000-03. Coincidently, the period from 2003 onwards aligns with the AMSR-E/AMSR-2 data record, beginning with the launch of the Aqua satellite in 2002.

The estimated values of daily change in snow density were then aggregated into two-week groups according to day-of-the-year (see Figure 3.2) and relationships were identified between successive *in situ* density measurements so the parameter space for the model to search

could be greatly reduced. The climatological mean change demonstrated increasing snow density at the beginning of the season, trending towards no change by the end of the season (black points in Figure 3.2). Similarly, the spread in the distributions of aggregated daily density change were found to narrow over the course of the season (green points in Figure 3.2). With these relationships identified between density measurements the range of possible solutions in the model can be limited to those which are climatological valid.

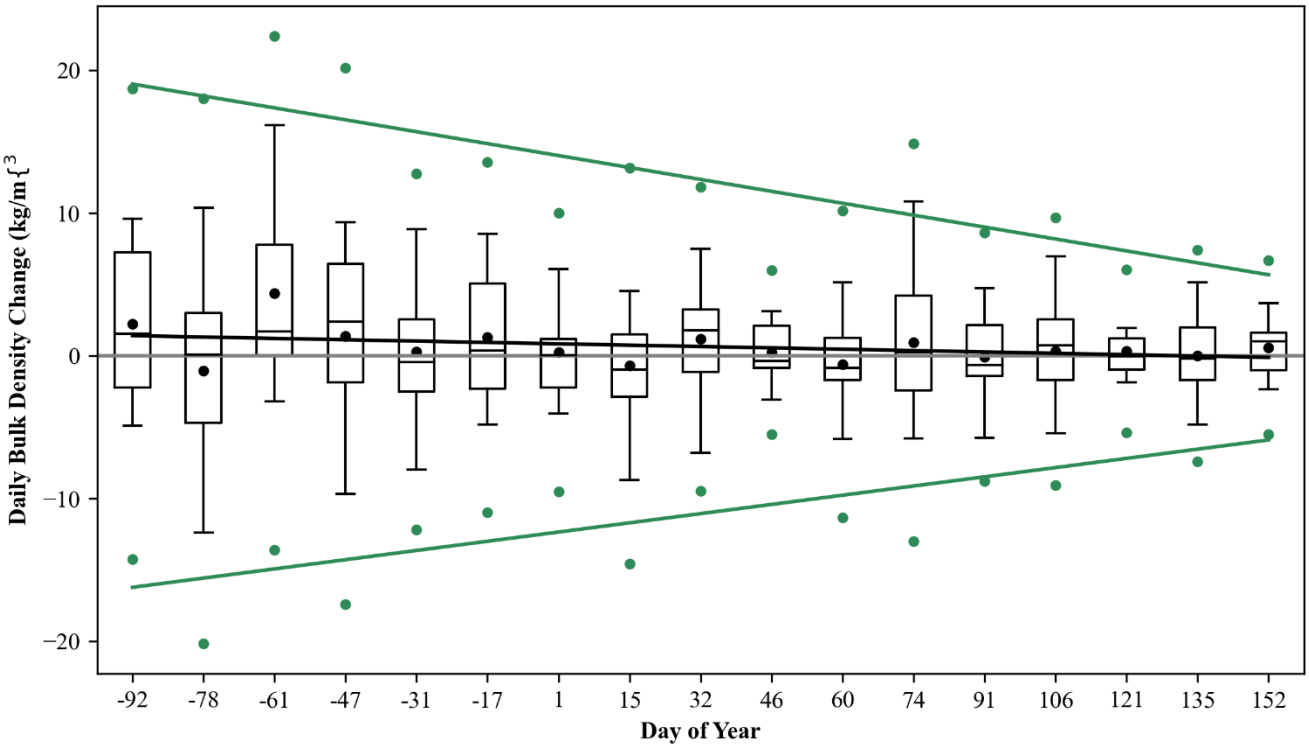


Figure 3.2 – Aggregated estimates of daily bulk density from in situ reference dataset (CanSWE) at the Eureka training site, plotted based on the day of year ("1" is January 1st). The black line is fitted through the distributions' mean values (shown as black points), representing the expected climatological change. The green lines are fitted through plus or minus three standard deviations from the expected value (shown as green points), representing the maximum allowed climatological change.

A dynamic, control test was developed to constrain density estimates based on the relationships observed between successive bulk density measurements in the reference dataset. The control test was taken from manufacturing quality control, where an output is deemed acceptable if it is within a prescribed tolerance; typically, within three standard deviations of the expected value (Manuele & Goffman, 1945). In this case, changes in bulk snow density should relatively conservative, so there should be minimal change between day-to-day estimates from the algorithm. If the scope of change in day-to-day snow density was known then the potential layer density combinations could be constrained by tolerances defined by the previous estimate of snow density. Therefore, linear relationships were fitted to the means and standard deviations of the binned distributions (see Figure 3.2), to define the tolerances for the control test. The relationships are related to the day of year (DOY), allowing them to be calculated at any iteration of the model with:

$$\mu(DOY) = -0.0061 * DOY + 0.816 \left[\frac{kg}{m^3 \cdot day}\right] \quad (3.3)$$

$$\sigma(DOY) = -0.017 * DOY + 4.57 \left[\frac{kg}{m^3 \cdot day}\right] \quad (3.4)$$

The mean change in daily snow density, from equation 3.3, introduced gradually increasing snow density estimates at the beginning of the season, trending toward zero to limit change later in the season (black line in Figure 3.2). Similarly, the range in estimated daily change in bulk density, from equation 3.4, also decreases over the course of the season (green lines in Figure 3.2); larger values allowed estimates to deviate more from climatology so to find the appropriate parameter space to converge with observed Tb, decreasing over time to restrict change in snow density later in the season.

3.5.4 Gradient Descent Model

Prior knowledge of the scene was incorporated into the model's cost function to constrain potential snow layer density combinations and improve algorithm performance. Such information provides context relevant to the scene so the model can more efficiently evaluated possible layer density combinations. There are strong environmental controls present in the tundra that contribute to a characteristic two layer snowpack composition. With an understanding of the processes in the tundra environment that affect snow properties, potential snowpack parameterizations can be discounted as they are unlikely to be representative of conditions on the ground. To that end, two components, I and B, were added to the cost function (f) of the model, such that:

$$f(\rho_{slab}, \rho_{hoar}) = \omega_1 \cdot J^2 + \omega_2 \cdot I^2 + \omega_3 \cdot B^2 \tag{3.5}$$

where the microwave modelling component (J), first introduced in equation 3.1, was modified and supplemented by inertia (I) and balancing (B) forces. The J, I, and B components are:

$$J(\rho_{slab}, \rho_{hoar}) = \frac{\Delta Tb_{sim}(\rho_{slab}, \rho_{hoar}) - \Delta Tb_{obs}}{\sigma_{obs}} \tag{3.6}$$

$$I(\rho_{slab}, \rho_{hoar}) = \frac{\Delta\rho(\rho_{slab}, \rho_{hoar}) - \mu_{\Delta\rho}}{\sigma_{\Delta\rho}} \tag{3.7}$$

$$B(\rho_{slab}, \rho_{hoar}) = \begin{cases} 0, & if\ \rho_{hoar} < \rho_{slab} \\ \frac{SD}{M} \cdot (\rho_{hoar} - \rho_{slab}), & otherwise \end{cases} \tag{3.8}$$

where each component was assigned a weight hyperparameter (ω_n) to adjust their relative importance (such that $\omega_1 + \omega_2 + \omega_3 = 1$). The modelling error ($\Delta Tb_{sim}(\rho_{slab}, \rho_{hoar}) - \Delta Tb_{obs}$), in equation 3.6, was standardized by dividing by the potential observation error (σ_{obs}) to match the function of $\sigma_{\Delta\rho}$, in equation 3.7. In this case, σ_{obs} was assigned a constant value

related to the uncertainty range of the AMSR-E instrument (± 0.6 K) and $\sigma_{\Delta BD}$ is the standard deviation in the control test, from equation 3.4. Now, J and I have the same units (i.e. standard deviations units) and are able to act relative to one another. In this way, J and I can work harmoniously to optimize synthetic microwave emissions to match observations, while remaining within the climatological scope of allowed change.

The climatological inertia force, from equation 3.7, was designed to limit fluctuations in algorithm estimates. This was not a true inertia force that limits any change but is relative to the previous algorithm's estimate of bulk snow density (ρ; weighted average of ρ_{hoar} and ρ_{slab}). Cost increases with the difference between the change in bulk density since the previous daily estimate (Δρ) and the expected climatological change ($\mu_{\Delta\rho}$, from equation 3.3). Algorithm estimates were allowed to diverge up until a boundary defined by the standard deviation from the control test ($\sigma_{\Delta\rho}$, from equation 3.4) which the estimate cannot cross on that iteration. Daily fluctuations in snow density estimates were ultimately confined to the range which is climatologically probable to ensure the model did not navigate an unrealistic distance through the parameter space on a single day.

Alternatively, the balancing component, from equation 3.8, is piecewise and only contributes when the system is deemed to be out of balance – when the depth hoar layer density (ρ_{hoar}) is higher than the slab layer density (ρ_{slab}). The microwave scattering signal of the snowpack is dominate by the depth hoar layer and, as such, fluctuations in observed ΔTb are not reflected as strongly in the wind slab layer. Thus, changes in Tb observations can cause the simulated depth hoar layer to become denser than the wind slab layer, a situation which is unlikely in the tundra. So, a logical rule was included in the cost function to avoid unlike snow density profiles by incurring a cost in those situations. This simple rule proved enough to keep

the algorithm in balance and provide realistic density estimates over the course of a season. Additionally, the balancing component was designed without hard boundaries to allow the snowpack to diverge to pseudo one-layer model when appropriate.

Now with the introduction of I and B to the cost function, the parameter space, and the number of potential layer density combinations, for the model to consider was greatly reduced. This was achieved by introducing penalties and boundaries within the parameter space, defined by the I and B components. Figure 3.3 provides a simplified representation of the restricted parameter space (without the J component) at the beginning of the season for a starting point (grey point in Figure 3.3) with equal density values for two snow layers.

The cost associated with I (shades of blue in Figure 3.3) increases as the current estimate diverges from the set of layer density combinations (dashed line in Figure 3.3) that equal the expected change in bulk density. The new estimate is allowed to diverge from the expected value up until it reaches the boundary defined by the climatological control test, which it cannot cross. If only the bulk density was considered, the boundary formed a long band that spanned across large sections of parameter space, because many layer density combinations result in the same bulk density. So, the standard deviation from the control test was applied directly to each layer to restrict the model to a specific portion of the parameter space (in the absence of detailed information on stratigraphy). This reduced the band of potential density combinations in the parameter space to a zone with boundaries in the shape of a parallelogram (black lines in Figure 3.3). That boundary was originally set at three standard deviations, with the intention of considering nearly all possible climatic scenarios, but through algorithm development it was found that that boundary was too large and afford the model too much latitude. With the boundaries set at three standard deviations the algorithm estimates were more sporadic and less

consistent. Instead, the boundary was reduced to one standard deviation which produced more conservative changes in bulk density and, in turn, more consistent algorithm outputs.

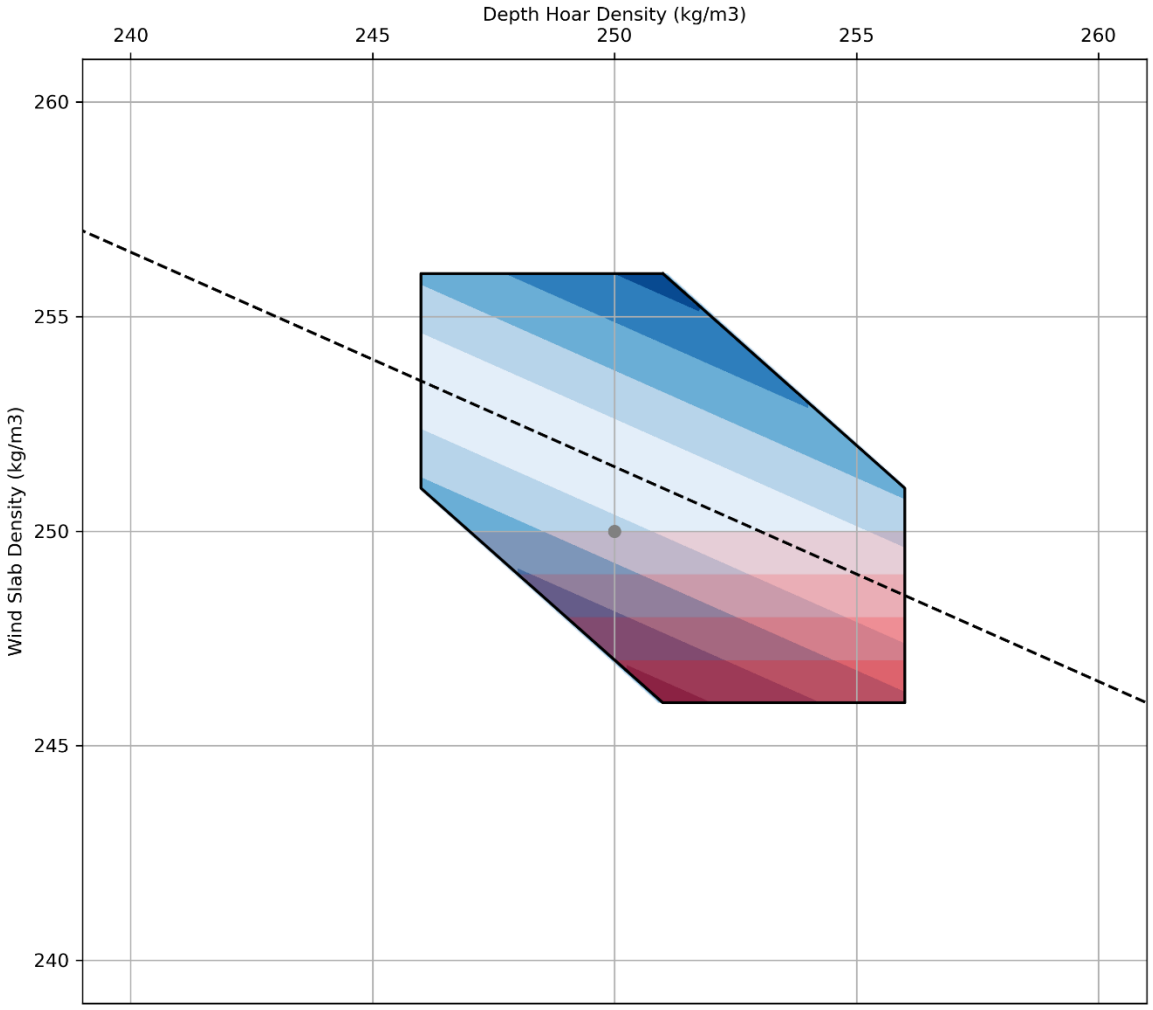


Figure 3.3 - Visual representation of the restricted parameter space by the climatological inertia (I) and balancing (B) components. The grey point indicates the starting point of layer densities for the given iteration, the dashed line is the set of layer density combinations that equal the expected climatological change, and the solid lines are the boundaries of the maximum allowed change in bulk density. The relative costs associated with the I and B components are shown in shades of blue and red, respectively (darker shades represent higher costs).

Note there is overlap of red and blue which has an additive effect on cost. Additionally, the J component is not shown, to simplify the representation, but would be superimposed on the shown parameter space.

The parameter space to search is further restricted by the B component which introduced an additional penalty when wind slab density is lower than depth hoar density (shown in shades of red in Figure 3.3), essentially disregarding those potential parameter combinations. The point where the balancing force took effect was set as the maximum density estimate for the depth hoar layer during the model's progression on that date so far, to act like a constant when estimating the rate of change in cost with respect to each layer density. In this way, decreasing the estimated depth hoar density would not be reduce the cost associated with the system being out of balance and instead the wind slab must become denser to alleviate the increased cost. The balancing force was implemented in this way because the volume scattering signal is largely dictated by the depth hoar layer, while the wind slab layer has a much lower effect and should not restrict changes in the depth hoar layer.

Beyond independent learning rates for each snow layer and the component weights in the cost function, only one other hyperparameter was included in the gradient descent model. That hyperparameter was designed to represent the likelihood of two distinct layers in the snowpack to dictate the relative strength of the balancing force. The formation of a depth hoar layer requires sufficient snow accumulation to insulate the underlying substrate and create the necessary temperature gradient for kinetic grain growth to occur (Colbeck, 1982). Thus, the strength of the balancing force was the ratio of measured snow depth at the AWS (SD) and that of a mature tundra snowpack (M) deep enough for depth hoar to exist. M was designated as a hyperparameter, to be determined by the model, in the absence of detailed information regarding stratigraphy at the training site (CanSWE is limited to bulk snow density). In the future, more detailed training data could be used to refine the balancing force.

3.5.5 Snow Density Retrieval at AWS Sites

Forcing data were required to parameterize the snowpack to be ingestible into the RTM and perform a passive microwave retrieval of snow density. The snow density retrieval process (shown in Figure 3.4) was inspired by the effective grain size optimization procedure in GlobSnow (Pulliainen, 2006; Takala et al., 2011). In this approach, points of measured snow depth from the AWS are used to provide prior knowledge of the scene to constrain passive microwave retrievals. Retrieving snow density at those points could produce data useful for a number of applications, compared to effective grain size which is only relevant with respect to RTMs. Therefore, the existing AWS network in northern Canada can be exploited to estimate snow density across the tundra.

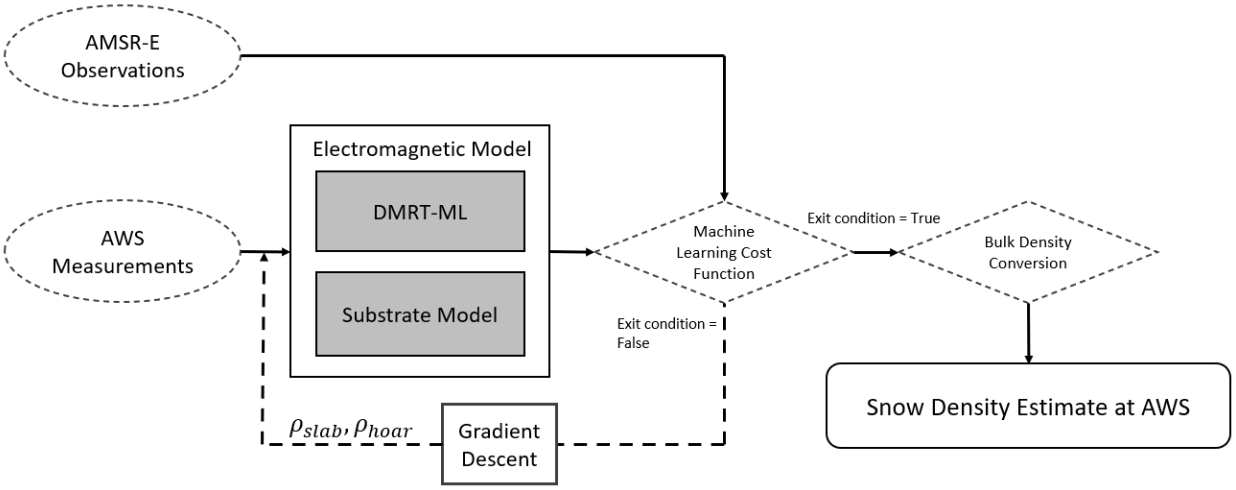


Figure 3.4 – Data flow diagram of snow density retrieval at AWS, from the AMSR-E observations (ΔTb) and AWS measurements (snow depth and air temperature) forcing data .

With the relevant AWS measurements prescribed to the model, the layer density parameters were adjusted to minimize the difference between the synthetic and observed ΔTb. Two conditions must be met before the passive microwave data were considered when estimating snow density. First, snow had to have been detected in the satellite observation

footprint using Grody & Basist (1996). Second, the AWS must have reported at least five centimetres of snow on the ground before satellite observations were considered because of the transparent nature of shallow snow to microwave radiation (Hall et al., 2002). If those conditions were not met the J term (equation 3.6) in the cost function was set to zero and the other components kept the system in check.

The algorithm was designed to include a very strong temporal element and is not intended to produce instantaneous estimates. Instead, the algorithm should be applied over the entire winter season to provide the most consistent estimates of bulk density. Unlike snow depth, snow density does not have a fixed starting value (i.e. snow density of zero kg/m^3 does not exist) and an estimate for the initial value was required to start the algorithm. Initial values were established for each site based on geographic location, with higher initial densities assigned to more northernly sites (Sturm & Holmgren, 1998). The two snow layers were initiated with equal densities, then the evolution of the snow density profile was dictated by the forcing data. The final estimate of bulk density for each day was calculated based on the densities of each snow layer and their respective depths (i.e. weighted average). This procedure required consistent forcing data throughout the season to produce robust results because of the strong link between successive snow density estimates in the algorithm. Further, in some situations the algorithm may require some time to equilibrate and evolve to the appropriate parameter space to replicate observed Tbs. Therefore, relatively high temporal sampling of the forcing data is key to successful algorithm execution while temporal gaps in input data lead to plateaus in density estimates.

3.5.6 Outlier Screening

Selected outliers in the reference density dataset were removed before analysing the algorithm outputs because they did not fit with the expected densification trajectory. In some cases, very low snow densities (<100 kg/m^3) or very high snow densities (>500 kg/m^3) were observed early in the season. It is unclear why these values were observed but they are not reflective of the expected seasonal trend in snow density. Only observations after November 1st were considered thereby removing the influence of early season variability from the analysis. In other cases, there were sporadic observations that did not fit temporally with the seasonal profile.

The Generalized Extreme Studentized Deviate (GESD) test (Rosner, 1983) was used to identify temporal outliers in model performance and remove those anomalous observations from the dataset during the analysis. GESD is an iterative test that identifies observations that deviate more from the modelled output than the other observations. At each iteration, the density estimate with the largest deviation from its reference measurement is tested and deemed a significant outlier from the remaining observations, or not. Then, the observation in question is removed and the process repeated until the maximum number of allowed outliers is reached. The significance level of the GESD test was set to p=0.05 and a maximum of two outliers were allowed during each algorithm run because of the relatively low number of observations available for each season. 18 of the 828 observations (2.18%) were identified as outliers and exclude from the analysis. Importantly, GESD is not limited to outliers that correspond with poor model performance and will remove any observations (good or bad) that significantly differ from the rest.

3.5.7 Training, Validation, and Testing

Before the model could provide suitable estimates of snow density it had to be trained. In this context, training refers to identifying the optimal hyperparameter combination for the gradient descent model, where a ***loss*** metric was used to compare the outputs from various combinations. Loss is analogous to the quality of the machine learning output (or rather the lack of quality) and different hyperparameter combinations are tried until the loss is minimized. In this case, the algorithm's loss metric is related to the difference between snow density estimates and the corresponding value in the reference dataset. Nash-Sutcliffe Efficiency (NSE) was chosen as the loss function because of its use in hydrological modelling (Nash & Sutcliffe, 1970) – values between zero and one indicate model estimates are closer to observations than they are to the observed mean; negative values indicate a constant mean estimate would perform better than the model.

The Eureka AWS was chosen to train the model because of the high-quality data available for the site. The Eureka site was chosen primarily because of Saberi et al. (2017), allowing for increased confidence in the parameterization of the snowpack (specifically the ratio of layer depths). One season of data from Eureka (2006-07) was used to training the model because it exhibited the expected gradual densification described by Sturm & Holmgren (1998) (ignoring one early season outlier). The training run was initiated with the first density observation of the 2006-07 season to ensure the algorithm started in a suitable parameter space (as opposed to all other algorithm runs where initial densities were prescribed based on site location). Training of the machine learning model was done with an interactive grid search of hyperparameter combinations to maximize NSE.

The optimal hyperparameter combination identified in training was then applied to the remain data. The remaining seven seasons were used to validate the algorithm configuration on similar, but different, data. Other AWS sites in the tundra, collocated with CanSWE sampling locations, were chosen to test the algorithm. Seven such AWS sites were identified over eight winter seasons of the AMSR-E data record. Although sparse, these sites are spread across a wide range of the Canadian tundra and provide a good representation of the conditions expected across this domain. It should be noted, the algorithm was trained on the data from a single site over one season (Eureka 2006-07) and, as such, represents a very low training sample size compared to most machine learning implementations (in this case: 2% training, 12% validation, and 86% testing).

3.6 Results

3.6.1 Performance Metrics and Evaluation

In the interest of clarity, standards are defined in this section to evaluate the algorithm's skill in estimating bulk snow density. They follow along from the previous subsection that defined the training – validation – testing sequence.

Before the analysis started, the input data were evaluated and deemed adequate, or not, for model forcing. Several of the AWS sites experienced data availability outages through the study period that disrupted the algorithm. Some of these outages were temporary and did not affect model performance, much, but others were prolonged enough to impede the model's ability to predict the temporal evolution of the snowpack. Input data were deemed ***adequate*** if AWS data were available for the majority of the winter season at a site without prolonged outages. After screening of the forcing data, 53 of the 64 datasets (eight sites over eight winter seasons) were deemed adequate and used to produce snow density estimates.

Once the input data were screened, algorithm outputs were assessed against the *in situ* reference dataset of snow density samples. A ten percent density range was applied to the reference dataset because of uncertainties inherent to manual snow density sampling (Conger & McClung, 2009; López-Moreno et al., 2020). An estimate of bulk snow density was deemed ***accurate*** if it was within the uncertainty range of the corresponding *in situ* measurement. Further, a series of density estimates was deemed ***robust*** if they were consistently within the ±10% uncertainty range over the span of multiple measurements.

Beyond the given qualitative definitions, statistical measures were used to quantify the algorithm's performance in estimating snow density, including mean absolute percentage error (MAPE), root mean square error (RMSE), and bias. NSE was also reported to quantify the algorithm's temporal performance in estimating snow density. However, there are criticisms as to the representativeness of NSE when performing inter-site or inter-season comparisons (Gupta et al., 2009), and NSE values are not necessarily comparable when considering multiple season or sites. So, an alternative metric to describe temporal algorithm performance is also reported – Kling Gupta Efficiency (KGE) – representing a standardized metric (with the similar connotations as NSE) to enable straightforward comparisons between algorithm runs for various sites or seasons.

In addition to the *in situ* reference dataset, Sturm et al.'s (2010) statistical snow density model is used as a baseline for the performance evaluation because of ease of application and applicability to passive microwave modelling (Kelly et al., 2019). Snow density is estimated by Sturm *et al.* (2010) with:

$$\rho = (\rho_{max} - \rho_0)[1 - exp(-k_1 \cdot SD - k_2 \cdot DOY)] + \rho_0 \left[\frac{kg}{m^3}\right] \qquad (3.9)$$

where bulk snow density (ρ) is related to snow depth (SD) and day-of-year (DOY), and the other parameters vary with snow class. For the tundra snow class, ρ_0 and ρ_{max} represent the minimum and maximum density values (242.5 and 363.0 kg/m^3, respectively) and $k_1$and k_2 are densification parameters for SD and DOY (0.0029 and 0.0049, respectively). When the statistical model (equation 3.9) is applied to the dataset it achieves MAPE, RMSE, bias, NSE, and KGE of 25.1%, 98.8 kg/m^3, -59.2 kg/m^3, -0.520, and -0.041, respectively.

Thresholds were established to determine the quality of the algorithm's snow density estimates, relative to existing methods. MAPE was the primary measure for absolute accuracy of algorithm runs and two thresholds were set accordingly – 10% and 25%, with respect to the ±10% uncertainty range of the *in situ* measurements and performance of the statistical model. Similarly, KGE was the primary measure for temporal accuracy and two thresholds were set accordingly – 0.00 and -0.414, with respect to the performance of the statistical model and that of a static mean estimate of snow density (-0.414 is the equivalent threshold for KGE as 0 is for NSE).

3.6.2 Training Results

Through training, the algorithm was able to very accurately replicate the observed changes in snow density over the 2006-07 Eureka winter season. The minimum loss achieved through training was an NSE of 0.388 (KGE of 0.728), corresponding to MAPE, RMSE, and bias of 5.7%, 20.5 kg/m^3, and -5.7 kg/m^3, respectively (see Table 3.2). A visual summary of that best training algorithm run is provided in Figure 3.5. Figure 3.5a shows the algorithm density estimates relative it the *in situ* measurements, where estimates were very robust with only one outside the ±10% uncertainty range of the samples on March 1st (with one outlier measurement on November 15 removed by the GESD test). The AWS forcing data, shown in Figure 3.5b&c,

experienced gradual increases in snow depth to a maximum of 16 cm and consistent sub-zero temperatures during the study period. Figure 3.5d shows the microwave modelling error between observed and synthetic emissions, which was close to zero for the majority of the training period. Thus, it appears that under ideal conditions this algorithm has the potential to produce very similar estimates of snow density to *in situ* sampling.

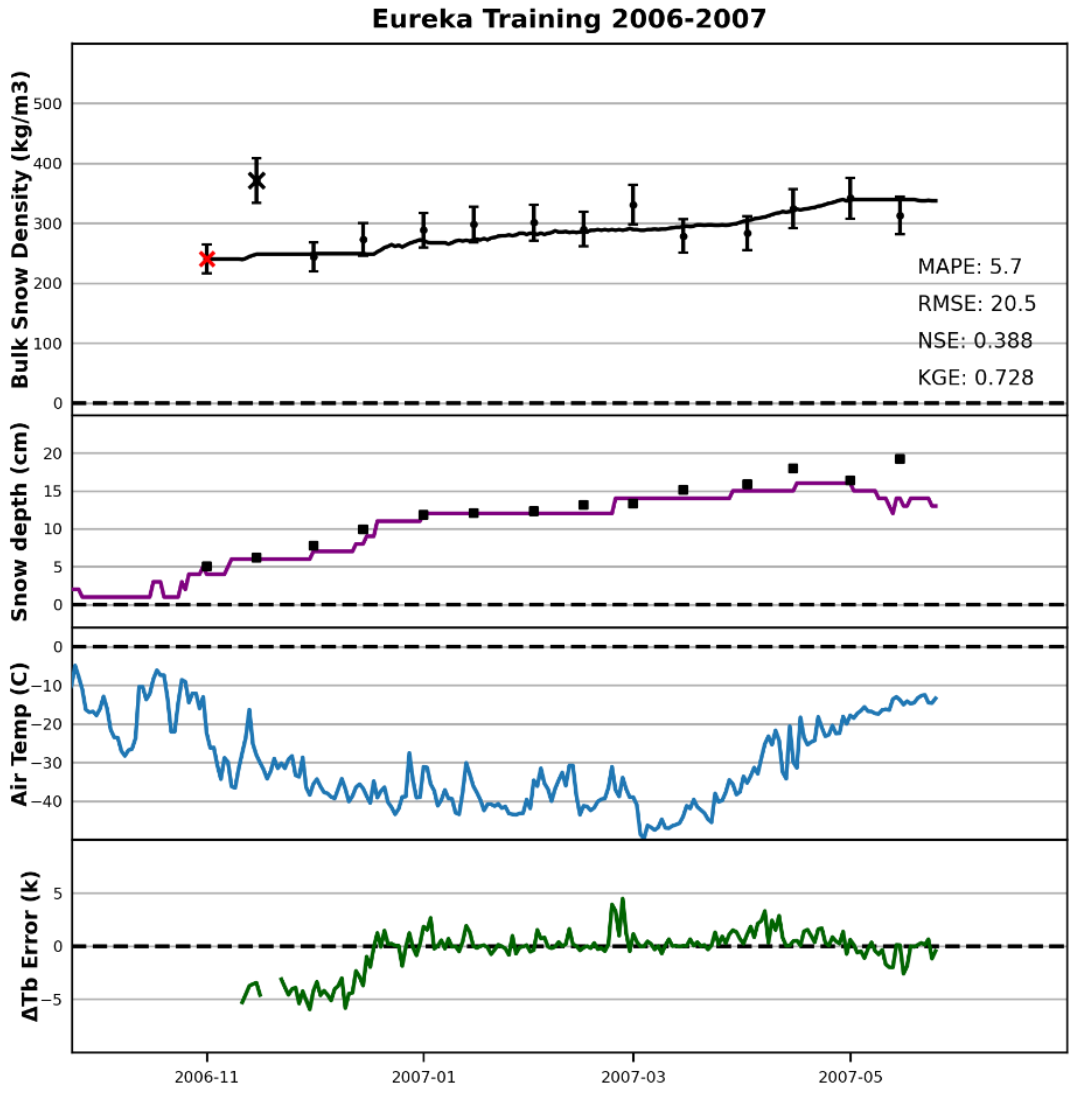


Figure 3.5 – Algorithm training results for Eureka 2006-07. a) shows bulk snow density estimates from the algorithm (black line) and CanSWE density samples (black points, with ±10% error bars). CanSWE samples marked with 'X's were removed from the analysis – red indicates the first observation used to initiate the algorithm run and black identified as an outlier by the GESD test. b) shows snow depth measured at the AWS (purple line) and CanSWE snow depth measurements (black squares). c) shows minimum daily air temperature at the AWS. d) shows the microwave modelling error ($\Delta Tb_{sim} - \Delta Tb_{obs}$).

Table 3.2 – Statistical summaries of algorithm performance (relative to the CanSWE refernce dataset) at the Eureka site, training data from the 2006-07 winter season and validation data are from 2003-06 and 2007-11.

Stage	MAPE (%)	RMSE (kg/m^3)	Bias (kg/m^3)	NSE	KGE
Training 2006-07	5.7	20.5	-5.7	0.388	0.728
Validation (remaining years)	15.5	70.1	-46.6	-1.411	0.161

3.6.3 Validation Results

Of the seven remaining winter seasons for Eureka all provided adequate input data to produce meaningful results, but for one season (2008-09) where the algorithm was hindered by intermittent snow cover. In general, the validation algorithm runs at Eureka achieved much lower absolute and temporal accuracies than in training (see Table 3.2). Algorithm performance varied over the validation seasons at Eureka, with MAPE ranging from 9 - 22% and KGE from -0.156 - 0.614 (more detail on individual runs is included in Table 3.4). The algorithm typically underestimated snow density early in the season at Eureka, but performance consistently improved over the course of each season. The algorithm was able to accurately estimate bulk snow density at peak SWE accumulations across all seasons at Eureka, despite early season performance. In all validation algorithm runs the algorithm estimates converged close to the *in situ* density samples later in the season. MAPE at peak SWE accumulation was 10.9% – just outside the ±10% uncertainty range assigned to *in situ* measurements – demonstrating a good potential for estimating end of season SWE conditions. All validation algorithm runs, but one (2003-04), surpassed the two benchmarks set by the Sturm *et al.* model (MAPE < 25% and KGE > 0; see table 3.4).

3.6.4 Testing Results

Table 3.3 – Statistical summaries of algorithm performance (relative to the CanSWE refernce dataset) for sites in the testing dataset.
** Eureka is inluded to provide context for the testing results – these data are composed of the training and validation stages.*
*** Agglomerated data for all testing sites (i.e. all sites except Eureka).*

Site	MAPE (%)	RMSE (kg/m^3)	Bias (kg/m^3)	NSE	KGE
Eureka*	14.2	65.6	-46.6	-1.077	0.186
Alert	15.3	66.9	-43.7	-0.728	0.168
Resolute	15.7	78.1	1.4	-0.13	0.042
Cambridge Bay	16.8	71.9	-32.3	-0.911	0.362
Coral Harbour	15.1	76.8	-54.8	-1.438	0.189
Iqaluit	19.7	82.8	-52.6	-1.196	0.296
Baker Lake	16.7	52.7	-28.3	0.029	0.557
Inuvik	27.3	60.8	39.2	-1.163	0.419
All Testing Sites**	17.7	71.3	-24.2	0.272	0.603

In testing, the algorithm demonstrated slightly lower absolute accuracy at sites distributed across the tundra, compared to Eureka, but significantly higher temporal accuracy (see Table 3.3). Absolute accuracies for individual sites were similar to validation with MAPE and RMSE marginally higher than those for Eureka, except for the Iqaluit and Inuvik sites which displayed considerably worse performance (see Table 3.3). Alternatively, most testing sites scored higher KGE values than Eureka. Like in the validation stage, algorithm performance consistently improved over the course of the season and MAPE at peak SWE accumulation was reduced to 14.5% (from 17.7% for the whole season). In some cases, the algorithm was able to produce very robust estimates of bulk snow density where estimates were very similar to *in situ* measurements (e.g. Figure 3.6a-c). In other cases, the algorithm estimates displayed a negative bias early in the season but converged measurements later in the season (e.g. Figure 3.6d-f). And in a few cases,

the algorithm estimates were completely different from the *in situ* densities samples (e.g. Figure 3.6g-i). there was insufficient space to describe all algorithm runs in detail, so summaries for each algorithm run were included in Appendix A.

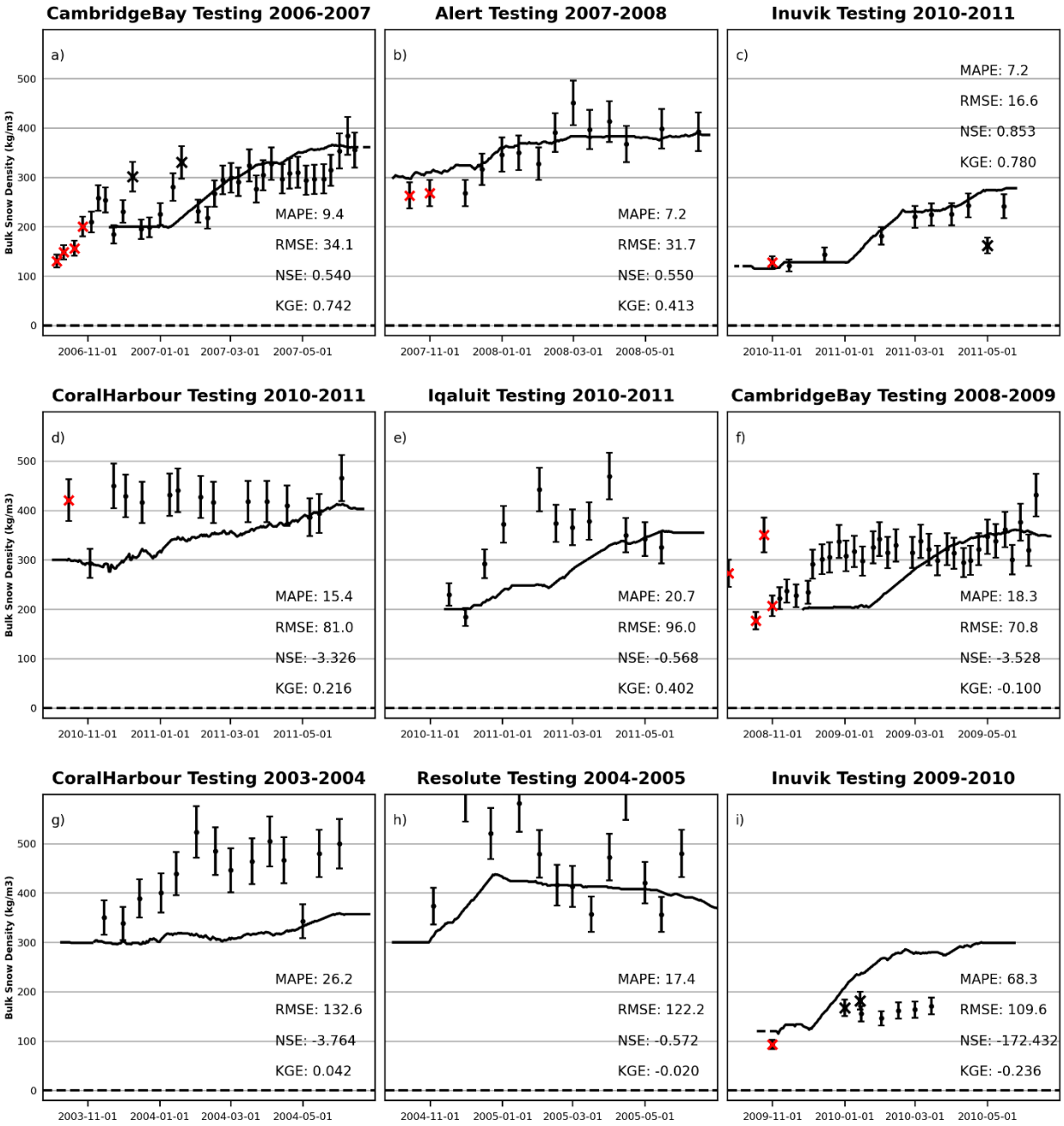


Figure 3.6 - Examples of algorithm outputs: top row (a-c) – good results, middle row (d-f) – intermediate results, and bottom row (g-i) – poor results. For each algorithm output, the plot shows bulk snow density estimates from the algorithm (black line) and the CanSWE density samples (black points, with ±10% error bars). CanSWE samples marked with 'X's were removed from the analysis – red indicates observations removed prior to Nov 1 and black identified as an outlier by the GESD test.

3.7 Discussion

A matrix of algorithm runs was created to evaluate estimation skill relative to the absolute and temporal accuracy thresholds established in section 3.5.1. The absolute accuracy of algorithm runs was more consistent at the northern sites and degraded for more southern sites (see Table 3.4). That was to be expected because the algorithm was trained at one of the most northernly sites (i.e. Eureka). Alternatively, temporal accuracy was much more consistent over the entire study area, with most algorithm runs scoring positive KGE values (see Table 3.4). Evidently, the algorithm performed better under certain conditions than others. The following subsection discusses emergent patterns in algorithm performance.

Table 3.4 – Absolute and temporal accuracies for every algorithm run, arranged in a matrix with rows ordered by descending site latitudes (horizontal black line separates high and sub arctic site designations) and columns ordered in chronological order. The training run (Eureka 2006-2007) is marked with dark borders to differentiate it from other algorithm runs.

The cells are colour coded with respect to the performance metrics outlined in subsection 3.5.1:

Absolute accuracy thresholds (MAPE) – Green = <10%, Yellow = >10% and <25%, Red = >25%

Temporal accuracy thresholds (KGE) – Green = >0, Yellow = <0 and > -0.414, %, Red = < -0.414

	Absolute Accuracy (MAPE [%])								Temporal Accuracy (KGE)							
Site	**2003-04**	**2004-05**	**2005-06**	**2006-07**	**2007-08**	**2008-09**	**2009-10**	**2010-11**	**2003-04**	**2004-05**	**2005-06**	**2006-07**	**2007-08**	**2008-09**	**2009-10**	**2010-11**
Alert			17.0	20.1	7.2		9.9	19.7			-0.030	-0.223	0.413		0.250	0.201
Eureka	20.9	21.5	13.7	5.7	14.7		12.6	9.2	-0.156	0.163	0.083	0.728	0.445		0.315	0.614
Resolute	10.8	17.4	16.9	19.0	10.9	13.9	12.0	23.8	0.313	-0.020	0.136	-0.036	-0.068	0.646	0.497	-0.051
Cambridge Bay	27.5	17.3	16.1	9.4	13.5	18.3	15.1	16.6	0.404	0.243	0.431	0.742	0.315	-0.100	0.168	-0.071
Coral Harbour	26.2	11.6	13.2	13.4	12.8	14.5	13.2	15.4	0.042	0.575	0.153	0.082	-0.396	0.138	0.368	0.216
Iqaluit		26.3			11.9		16.5	20.7		0.160			-0.009		0.061	0.402
Baker Lake				9.2	17.1	15.1	17.9	25.0				0.373	0.112	-0.006	-0.968	0.241
Inuvik	28.5	18.1	16.2	48.5	27.7	21.4	68.3	7.2	0.067	0.626	0.323	0.427	0.084	0.399	-0.236	0.780

3.7.1 High Arctic vs Subarctic Sites

The tundra is vast domain, with a varied landscape, and can be better segregated into "high arctic" and "subarctic" sub classes to better characterize different snow conditions (Derksen et al., 2014). Sites above 70N were classified as high arctic and others as subarctic (designations are provided in Table 3.1). Algorithm performance for the three high arctic sites (Alert, Eureka, and Resolute) was quite similar, with marginally lower estimation skill than Eureka at the other two sites. On the other hand, results in the subarctic where were mixed. Absolute accuracy was consistently lower at the subarctic sites, but temporal accuracy was higher. By and large, sites in the high arctic displayed better absolute estimation accuracy than those in the subarctic, and the opposite in terms of temporal accuracy (see Table. 3.5).

Table 3.5 Algorithm performance for "high arctic" and "subarctic" classes (relative to the CanSWE refernce density dataset).

Subclass	MAPE (%)	RMSE (kg/m^3)	Bias (kg/m^3)	NSE	KGE
High Arctic	15.6	72.8	-25.5	-0.343	0.312
Subarctic	18.3	70.4	-26.8	0.287	0.605

These discrepancies in algorithm performance between the high and subarctic sites could imply that a more comprehensive classification scheme should be used in the future, rather than a single algorithm configuration for the entire tundra biome. On the one hand, absolute accuracy was rather consistent in the high arctic (see Table 3.4), suggesting the algorithm was sensitive those snow conditions while not overly fit to the Eureka site. On the other hand, the subarctic sites scored lower absolute accuracies, indicating the algorithm configuration may have been less representative of those areas. The snow conditions experienced in the subarctic were more varied, and some sites had conditions similar to those in the high arctic class while other were

distinctly different. The *in situ* snow density samples for Baker Lake and Inuvik were much lower than for other sites in the subarctic (and average snow depth for Inuvik was much higher than any other, see Table 3.1), potentially warranting a sperate classification. Results for most subarctic sites were comparable to those in the high arctic but Inuvik stood out with very poor absolute accuracy. The algorithm displayed starkly different behaviour at the Inuvik site where it over estimated snow density, compared to underestimating density at all other sites (except for Resolute which had a near neutral bias, see table 3.3). As a result, MAPE values for the Inuvik site were disproportionately high because of the relatively low reference values (Makridakis, 1993). Thus, it seems the study area could be more carefully divided for future analyses, to account for differing snow conditions, and representative sites should be selected to training the algorithm for each of subclasses.

3.7.2 *In situ* vs. Estimated Densification Trajectories

In most cases, the algorithm underestimated snow density early in the season and estimates converged with the *in situ* density samples later in the season. This behaviour is apparent at most sites (see Figure 3.7) and is demonstrate distinctly at the Cambridge Bay site (see Figure 3.7d). On reason for the tendency of early season underestimation can be explained by how the algorithm runs were initiated. Anecdotally, through development the algorithm displayed somewhat monotonic behaviour, where density estimates were more likely to increase than decrease. By initiating the algorithm at a lower density it would be more likely to find the appropriate parameter space by increasing to match *in situ* samples (whereas estimates would plateau if initiated too densely, i.e. above *in situ* samples). So, initial density values for each site were set conservatively low so to ensure the modelled snowpack would evolve. This approach for initiating the algorithm came at the cost of lower absolute accuracy through underestimation

earlier in the season. However, estimation skill improved quite consistently over the course of the algorithm runs, with 25 of 53 estimates within the ±10% uncertainty range of the *in situ* measurements by the time of peak SWE accumulation.

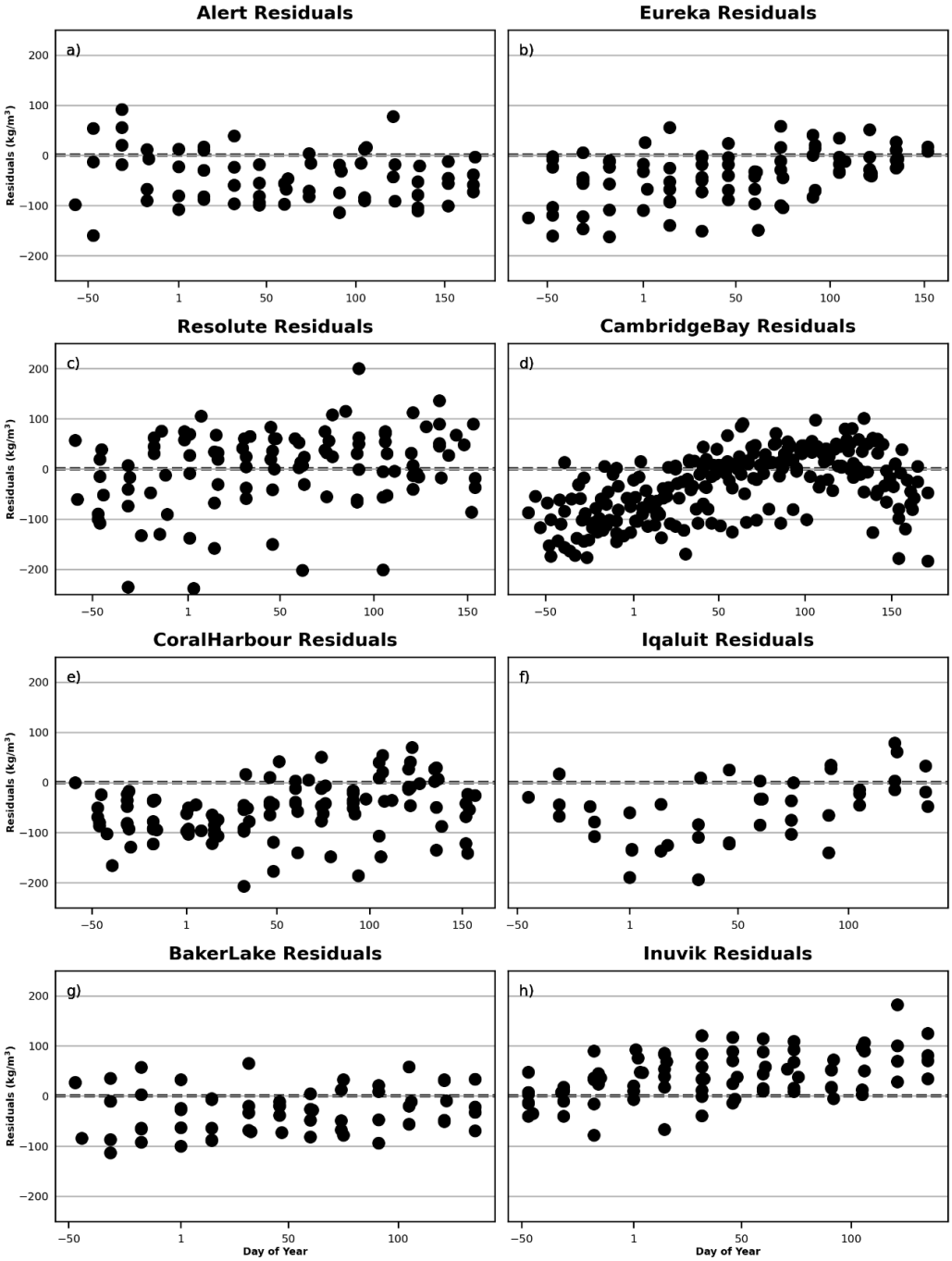


Figure 3.7 - Residuals between algorithm estimates and in situ CanSWE samples, for all algorithm runs, based on day of year.

The algorithm's ability to capture the expected densification trend varied with relation to the spatial extent of the analysis. Temporal accuracy for the testing dataset was comparable to training, whereas, testing KGE (0.603) was many times higher that of validation (0.161). This result is counterintuitive because the algorithm was trained on data from Eureka, so one would expect better performance for the validation data from the same site. However, the temporal performance metrics are related to the degree of variability in the observed data – a constant estimate will fit better to data with lower intrinsic variability than those with a degree higher variability in the observed values. For example, the *in situ* snow density samples from the 2009-10 season at Inuvik had lowest variability in the reference data (σ = 8.32 kg/m^3) and, by far, the lowest NSE of any algorithm runs (-172.4). This suggests that algorithm performance, relative to existing methods, would be significantly better when applied on a tundra-wide scale where more variability is expected (i.e. compared to the mean density across the domain), but not necessarily for a single site with less intrinsic variability (i.e. compared to the mean density from one site). Such was not only true for Eureka and every site scored lower individual KGE than the dataset as a whole.

3.7.3 Model Forcing

There were some algorithm runs which did not accurately estimate snow density, and it was not clear why. Algorithm performance did not appear to be strongly dictated by the AWS forcing data (snow depth and air temperature) and those like the training data did not guarantee accurate estimates of snow density. Specifically, it appears that the snow load did not control the densification trajectories at sites, nor was an indicator of algorithm performance (provided there was sufficient snow accumulation to permit microwave retrievals). However, the need for quality AWS forcing data should not be dismissed and they should be representative of the scene to

isolate the volume scattering contribution of snow density. Instead, other meteorological factors (e.g. wind speed; Sturm & Holmgren, 1998) would likely be more important for estimating tundra snow conditions.

From algorithm estimates, it was apparent that a minimum snow depth was required to estimate snow density using satellite passive microwave remote sensing. The algorithm was configured to consider satellite observations after five centimetres of snow was reported at the AWS, but the microwave modelling component of the cost function (J) was negligible at that level of accumulation and did not affect algorithm estimates of snow density. The inertia force (I) kept density estimates constant at the initial value until sufficient snow had accumulated to exhibit volume scattering and permit microwave retrievals. The algorithm consistently responded to a volume scattering signal in the passive microwave observations once ten centimetres of snow was reported at the AWS and the modelled snowpack would begin to evolve (i.e. change density) from that point. This minimum threshold for the retrieval of snow characteristics from passive microwave remote sensing will inform future development of this algorithm. Wind conditions may prove especially useful for estimating densification while the snowpack is still transparent to microwaves before that minimum snow depth threshold (i.e. identify the density value to initiate passive microwave retrievals). A more detailed analysis of meteorological conditions, and site characteristics, in the tundra is required to identify what drives variability in snow density at the passive microwave scale and how to properly force the algorithm in the future.

3.8 Conclusion

This research has demonstrated a sensitivity to changes in bulk snow density observed by satellite based passive microwave observations and operational AWS measurements, to provide a

novel approach for monitoring snow density in remote areas. The algorithm was applied across a very large domain with minimal training, owing to the incorporation of prior knowledge into the machine learning model to constrain the evolution of snow density estimates over time. Algorithm performance was quite good across the study area, although not entirely consistent with decreasing absolute accuracies at southern sites.

The algorithm showed especially good promise for estimating snow density at peak SWE accumulation, with nearly 50% of density estimates [at peak accumulation] within the uncertainty of the *in situ* samples. Therefore, snow density estimates from this approach could be useful for applications where the total amount of water storage within the snowpack is the primary concern. However, it was clear that more comprehensive algorithm forcing and training routine should be considered in the future, for algorithm performance to be more consistent across time and space.

Future work will build upon this prototype algorithm to improve snow density estimates from passive microwave remote sensing. The next logical step would attempt to produce spatially continuous snow density estimates from the distributed site level snow density estimates. That raises the question on how to best transition from spatially distributed density estimates to spatial continuous estimates and how to validate such a product. In a very straightforward way, the conversion could be accomplished through interpolation (like in Venäläinen et al., 2021). Alternatively, the approach could be modified to take in continuous meteorological forcing data, which could be interpolated from the AWS measurements or from climate models, to facilitate passive microwave retrievals of snow density across the entire study area.

Chapter 4

Conclusions and Future Work

This research demonstrated satellite based passive microwave remote sensing can be used to estimate bulk snow density and monitor changes over the course a season. The algorithm demonstrated good skill in estimating snow density across the study area with very minimal training, but estimation accuracy was not consistent across the domain. The spatial patterns identified in algorithm performance across the tundra suggest characteristics (either site or snowpack related) differed significantly between sites, potentially warranting different model configurations for distinct areas. Despite those issues, the algorithm consistently produced better estimates of snow density compared to the typical methods. Ultimately, this research provided a novel method for monitoring snow density in remote areas, which after further refinement might provide high quality snow density estimates in areas that were not previously observable.

With respect to the first goal, from section 1.2, the algorithm showed some sensitivity to local fluctuations in bulk snow density and produced accurate and robust estimates of snow density during some periods. In general, algorithm estimates did not reflect early season variability in the *in situ* dataset but performance improved later in the season. Two factors can explain why the algorithm could not replicate those early observations (assuming they were representative of the conditions on the ground). First, the early-season snow density fluctuations, not captured by algorithm estimates, typically occurred while the snowpack was relatively shallow (i.e. <10 cm). Therefore, the snowpack was likely transparent to microwaves and the

satellite observations could not provide any reference on the snow conditions. Second, once the snowpack had sufficiently matured to for volume scattering to dominate, the strong temporal design of the algorithm required time for the modelled snowpack to evolve to the correct parameter space. That is, the climatological inertia force slowed down the algorithm's progression, as intended, but was possibly too strong in this implementation and prevented the modelled snowpack from evolving in a timely manner. When these factors are considered together, it seems this approach could be suitable for estimating the bulk snow density of a mature snowpack, specifically at the time of peak SWE accumulation. In this way, this approach could be well suited to hydrological applications where the total water storage in the snowpack is of primary concern.

The second goal of this thesis, from section 1.2, was not explicitly tested but the algorithm demonstrated promising behaviour for estimating snow density at sites across the tundra. While this research did not attempt to convert the distributed density estimates into continuous fields (e.g. interpolate between sites; as described in Venäläinen et al., 2021), it did provide a new method for distributed snow density estimates that future research can build upon. Moving forward there may be an alternative to interpolating the snow density estimates from each AWS. Instead, the snow depth forcing data could be interpolated, similar to the GlobSnow snow depth interpolation approach (Takala et al., 2011), to provide spatially continuous forcing data for the passive microwave retrieval process. In this way, snow density could be estimated for each passive microwave grid cell, individually, rather than directly at the AWS points. This modified approach could potentially provide greater spatial detail in algorithm estimates, than an interpolated product, by making use of the full extent of passive microwave observations (at greater computational expense).

There is potentially a more significant outcome from this research, beyond estimating bulk snow density from passive microwave observations. That is, the very complex system in this algorithm (i.e. the 3D gradient descent model) was sufficiently constrained with the introduction of simple, logical rules. Those rules introduced physical knowledge into the retrieval process and made the algorithm behaviour consistent enough to be applied across a large domain with a very limited training sample. In this case, the rules were implemented in a fashion specific to gradient descent machine learning, but the principles could be applied to other models. Specifically, the concepts behind the climatological inertia force could be applied to other SD/SWE algorithms so to increase algorithm efficiency by limiting the model parameter space to search, while simultaneously preventing fluctuations in estimates.

References

Aggarwal, C. C. (2020). Optimization Basics: A Machine Learning View. In *Optimization Basics: A Machine Learning View* (pp. 141–204). Springer International Publishing. https://doi.org/10.1007/978-3-030-40344-7

Ashcroft, P., & Wentz, F. J. (2013). *AMSR-E/Aqua L2A Global Swath Spatially-Resampled Brightness Temperatures, Version 3 [Data Set]*. https://doi.org/10.5067/AMSR-E/AE_L2A.003

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(17), 303–309. https://doi.org/10.1038/nature04141

Bartelt, P., & Lehning, M. (2002). A physical SNOWPACK model for the Swiss avalanche warning Part I: numerical model. *Cold Regions Science and Technology*, *35*, 123–145. www.elsevier.com/locate/coldregions

Benson, C. S., & Sturm, M. (1993). Structure and wind transport of seasonal snow on the Arctic slope of Alaska. *Annals of Glaciology*, *18*, 261–267. https://doi.org/10.3189/s0260305500011629

Boelman, N. T., Liston, G. E., Gurarie, E., Meddens, A. J. H., Mahoney, P. J., Kirchner, P. B., Bohrer, G., Brinkman, T. J., Cosgrove, C. L., Eitel, J. U. H., Hebblewhite, M., Kimball, J. S., Lapoint, S., Nolin, A. W., Pedersen, S. H., Prugh, L. R., Reinking, A. K., & Vierling, L. A. (2019). Integrating snow science and wildlife ecology in Arctic-boreal North America. *Environmental Research Letters*, *14*(1). https://doi.org/10.1088/1748-9326/aaeec1

Bohr, G. S., & Aguado, E. (2001). Use of April 1 SWE measurements as estimates of peak seasonal snowpack and total cold-season precipitation. *Water Resources Research*, *37*(1), 51–60. https://doi.org/10.1029/2000WR900256

Brown, R. D., Fang, B., & Mudryk, L. (2019). Update of Canadian Historical Snow Survey Data and Analysis of Snow Water Equivalent Trends, 1967–2016. In *Atmosphere - Ocean* (Vol. 57, Issue 2, pp. 149–156). Taylor and Francis Ltd. https://doi.org/10.1080/07055900.2019.1598843

Brun, E., David, P., Sudul, M., & Brunot, G. (1992). A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting. In *Journal of Glaciology* (Vol. 38, Issue 128). https://doi.org/https://doi.org/10.3189/S0022143000009552

Champollion, N., Picard, G., Arnaud, L., Lefebvre, É., MacElloni, G., Rémy, F., & Fily, M. (2019). Marked decrease in the near-surface snow density retrieved by AMSR-E satellite at Dome C, Antarctica, between 2002 and 2011. *Cryosphere*, *13*(4), 1215–1232. https://doi.org/10.5194/tc-13-1215-2019

Chang, A. T. C., Foster, J. L., & Hall, D. K. (1987). Nimbus-7 SMMR Derived Global Snow Cover Parameters. *Annals of Glaciology*, *9*, 39–44. https://doi.org/https://doi.org/10.3189/S0260305500200736

Chang, A. T. C., Foster, J. L., Hall, D. K., Rango, A., & Hartline, B. K. (1982). Snow Water Equivalent Estimation by Microwave Radiometry. *Cold Regions Science and Technology*, *5*, 259–267. https://doi.org/https://doi.org/10.1016/0165-232X(82)90019-2

Chang, A. T. C., Shiue, J. C., Boyne, H., Ellerbruch, D., Counas, G., Wittmann, R., & Jones, R. (1979). Preliminary Results of Passive Microwave Snow Experiment During February and March 1978. *NASA Technical Paper 1408*.

Cohen, J., & Rind, D. (1991). The Effect of Snow Cover on the Climate. Journal of Climate, 4(7), 689–706. https://doi.org/https://doi-org.proxy.lib.uwaterloo.ca/10.1175/1520-0442(1991)004%3C0689:TEOSCO%3E2.0.CO;2

Colbeck, S. C. (1982). An Overview of Seasonal Snow Metamorphism. *Reviews of Geophysics and Space Physics*, *20*(1), 45–61. https://doi.org/10.1029/RG020i001p00045

Conger, S. M., & McClung, D. M. (2009). Comparison of density cutters for snow profile observations. *Journal of Glaciology*, *55*, 16–196. https://doi.org/https://doi.org/10.3189/002214309788609038

Deems, J. S., Painter, T. H., & Finnegan, D. C. (2013). Lidar measurement of snow depth: A review. In *Journal of Glaciology* (Vol. 59, Issue 215, pp. 467–479). https://doi.org/10.3189/2013JoG12J154

Derksen, C., Lemmetyinen, J., Toose, P., Silis, A., Pulliainen, J., & Sturm, M. (2014). Physical properties of arctic versus subarctic snow: Implications for high latitude passive microwave snow water equivalent retrievals. *Journal of Geophysical Research*, *119*(12), 7254–7270. https://doi.org/10.1002/2013JD021264

Derksen, C., Strapp, J. W., & Walker, A. (2006). Passive Microwave Brightness Temperature Scaling Over Snow Covered Boreal Forest and Tundra. *IEEE International Symposium on Geoscience and Remote Sensing*, 3762–3765. https://doi.org/https://doi.org/10.1109/IGARSS.2006.964

Derksen, C., Sturm, M., Liston, G. E., Holmgren, J., Huntington, H., Silis, A., & Solie, D. (2009). Northwest Territories and Nunavut snow characteristics from a subarctic traverse: Implications for passive microwave remote sensing. *Journal of Hydrometeorology*, *10*(2), 448–463. https://doi.org/10.1175/2008JHM1074.1

Derksen, C., Toose, P., Rees, A., Wang, L., English, M., Walker, A., & Sturm, M. (2010). Development of a tundra-specific snow water equivalent retrieval algorithm for satellite passive microwave data. *Remote Sensing of Environment*, *114*(8), 1699–1709. https://doi.org/10.1016/j.rse.2010.02.019

Derksen, C., Walker, A. E., Goodison, B. E., & Strapp, J. W. (2005). Integrating in situ and multiscale passive microwave data for estimation of subgrid scale snow water equivalent distribution and variability. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(5), 960–972. https://doi.org/10.1109/TGRS.2004.839591

Derksen, C., Walker, A., & Goodison, B. (2005). Evaluation of passive microwave snow water equivalent retrievals across the boreal forest/tundra transition of western Canada. *Remote Sensing of Environment*, *96*(3–4), 315–327. https://doi.org/10.1016/j.rse.2005.02.014

Domine, F., Picard, G., Morin, S., Barrere, M., Madore, J. B., & Langlois, A. (2019). Major Issues in Simulating Some Arctic Snowpack Properties Using Current Detailed Snow Physics Models: Consequences for the Thermal Regime and Water Budget of Permafrost. *Journal of Advances in Modeling Earth Systems*, *11*(1), 34–44. https://doi.org/10.1029/2018MS001445

Environment and Climate Change Canada, & ClimateData.ca. (n.d.). *Historic Station Data*. Retrieved June 15, 2023, from https://eccc-msc.github.io/open-data/licence/readme_en/

Farmer, C. J. Q., Nelson, T. A., Wulder, M. A., & Derksen, C. (2009). Spatial-temporal patterns of snow cover in western Canada. *Canadian Geographer*, *53*(4), 473–487. https://doi.org/10.1111/j.1541-0064.2009.00283.x

Fassnacht, S. R., Heun, C. M., López-Moreno, J. I., & Latron, J. (2010). Variability of Snow Density Measurements in The Rio Esera Valley, Pyrenees Mountains, Spain. *Cuadernos de Investigación Geográfica*, *36*(1), 59–72. https://doi.org/https://doi.org/10.18172/cig.1227

Foster, J. L., Chang, A. T. C., & Hall, D. K. (1997). Comparison of Snow Mass Estimates from a Prototype Passive Microwave Snow Algorithm, a Revised Algorithm and a Snow Depth Climatology. *Remote Sensing and Environment*, *62*, 132–142. https://doi.org/https://doi.org/10.1016/S0034-4257(97)00085-0

Fox-Kemper, B., H.T. Hewitt, C. Xiao, G. Aðalgeirsdóttir, S.S. Drijfhout, T.L. Edwards, N.R. Golledge, M. Hemer, R.E. Kopp, G. Krinner, A. Mix, D. Notz, S. Nowicki, I.S. Nurhati, L. Ruiz, J.-B. Sallée, A.B.A. Slangen, and Y. Yu, 2021: Ocean, Cryosphere and Sea Level Change. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1211–1362, doi: 10.1017/9781009157896.011.

Gan, T. Y., Barry, R. G., Gizaw, M., Gobena, A., & Balaji, R. (2013). Changes in North American snowpacks for 1979-2007 detected from the snow water equivalent data of SMMR and SSM/I passive microwave and related climatic factors. *Journal of Geophysical Research Atmospheres*, *118*(14), 7682–7697. https://doi.org/10.1002/jgrd.50507

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., … Zhao, B. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Gloersen, P., Wilheit, T. T., Chang, T. C., Nordberg, W., & Campbell, W. J. (1974). Microwave Maps of the Polar Ice of the Earth. *Bulletin of the American Meteorological Society*, *55*(12), 1442–1448. https://doi.org/https://doi.org/10.1175/1520-0477(1974)055%3C1442:MMOTPI%3E2.0.CO;2

Goodison, B. E., & Walker, A. E. (1993). Use of snow cover derived from satellite passive microwave data as an indicator of climate change. *Annals of Glaciology*, *17*, 137–142. https://doi.org/https://doi.org/10.3189/S0260305500012738

Goodison, B. E., & Walker, A. E. (1995). Canadian development and use of snow cover information from passive microwave satellite data. In B. J. Choudhury, Y. H. Kerr, E. G. Njoku, & P. Pampaloni (Eds.), *Passive Microwave Remote Sensing of Land-Atmosphere Interactions* (pp. 245–262). De Gruyter. https://doi.org/10.1515/9783112319307-015

Grody, N. C., & Basist, A. N. (1996). Global identification of snowcover using ssm/i measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *34*(1), 237–249. https://doi.org/10.1109/36.481908

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological

modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hall, D. K. (1987). Influence Of Depth Hoar on Microwave Emission from Snow in Northern Alaska. *Cold Regions Science and Technology*, *13*, 225–231. https://doi.org/https://doi.org/10.1016/0165-232X(87)90003-6

Hall, D. K., Kelly, R. E. J., Riggs, G. A., Chang, A. T. C., & Foster, J. L. (2002). Assessment of the relative accuracy of hemispheric-scale snow-cover maps. *Annals of Glaciology*, *34*, 24–30. https://doi.org/10.3189/172756402781817770

Hannula, H. R., Lemmetyinen, J., Kontu, A., Derksen, C., & Pulliainen, J. (2016). Spatial and Temporal Variation of Bulk Snow Properties in North Boreal and Tundra Environments Based on 2 Extensive Field Measurements. *Geoscientific Instrumentation, Methods, and Data Systems*, *5*(2). https://doi.org/https://doi.org/10.5194/gi-5-347-2016

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Hofer, R., & Mätzler, C. (1980). Investigations on snow parameters by radiometry in the 3- to 60- mm wavelength region (Alps). *Journal of Geophysical Research*, *85*(C1), 453–460. https://doi.org/10.1029/JC085iC01p00453

Houtz, D., Naderpour, R., Schwank, M., & Steffen, K. (2019). Snow wetness and density retrieved from L-band satellite radiometer observations over a site in the West Greenland

ablation zone. *Remote Sensing of Environment*, *235*. https://doi.org/10.1016/j.rse.2019.111361

Jones, H. G. (1999). The ecology of snow-covered systems: A brief overview of nutrient cycling and life in the cold. *Hydrological Processes*, *13*(14–15), 2135–2147. https://doi.org/10.1002/(sici)1099-1085(199910)13:14/15<2135::aid-hyp862>3.0.co;2-y

Kelly, R. (2009). The AMSR-E Snow Depth Algorithm: Description and Initial Results. *Journal of The Remote Sensing Society of Japan*, *29*(1), 307–317. https://doi.org/https://doi.org/10.11440/rssj.29.307

Kelly, R., Chang, A. T., Tsang, L., & Foster, J. L. (2003). A prototype AMSR-E global snow area and snow depth algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 230–242. https://doi.org/10.1109/TGRS.2003.809118

Kelly, R., Li, Q., & Saberi, N. (2019). The AMSR2 Satellite-based Microwave Snow Algorithm (SMSA): A New Algorithm for Estimating Global Snow Accumulation. *2019 IEEE International Geoscience & Remote Sensing Symposium*, 5606–5609. https://doi.org/https://doi.org/10.1109/IGARSS.2019.8898525

Kinar, N. J., & Pomeroy, J. W. (2015). Measurement of the physical properties of the snowpack. In *Reviews of Geophysics* (Vol. 53, Issue 2, pp. 481–544). Blackwell Publishing Ltd. https://doi.org/10.1002/2015RG000481

King, F., Erler, A. R., Frey, S. K., & Fletcher, C. G. (2020). Application of machine learning techniques for regional bias correction of snow water equivalent estimates in Ontario, Canada. *Hydrology and Earth System Sciences*, *24*(10), 4887–4902. https://doi.org/10.5194/hess-24-4887-2020

Larue, F., Royer, A., De Sève, D., Roy, A., & Cosme, E. (2018). Assimilation of passive microwave AMSR-2 satellite observations in a snowpack evolution model over northeastern Canada. *Hydrology and Earth System Sciences*, *22*(11), 5711–5734. https://doi.org/10.5194/hess-22-5711-2018

Larue, F., Royer, A., De Sève, D., Roy, A., Picard, G., Vionnet, V., & Cosme, E. (2018). Simulation and Assimilation of Passive Microwave Data Using a Snowpack Model Coupled to a Calibrated Radiative Transfer Model Over Northeastern Canada. *Water Resources Research*, *54*(7), 4823–4848. https://doi.org/10.1029/2017WR022132

Lemmetyinen, J., Schwank, M., Rautiainen, K., Kontu, A., Parkkinen, T., Mätzler, C., Wiesmann, A., Wegmüller, U., Derksen, C., Toose, P., Roy, A., & Pulliainen, J. (2016). Snow density and ground permittivity retrieved from L-band radiometry: Application to experimental data. *Remote Sensing of Environment*, *180*, 377–391. https://doi.org/10.1016/j.rse.2016.02.002

Li, Q., Kelly, R., Lemmetyinen, J., & Pan, J. (2020). Simulating the Influence of Temperature on Microwave Transmissivity of Trees during Winter Observed by Spaceborne Microwave Radiometery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 4816–4824. https://doi.org/10.1109/JSTARS.2020.3017618

López-Moreno, J. I., Fassnacht, S. R., Heath, J. T., Musselman, K. N., Revuelto, J., Latron, J., Morán-Tejeda, E., & Jonas, T. (2013). Small scale spatial variability of snow density and depth over complex alpine terrain: Implications for estimating snow water equivalent. *Advances in Water Resources*, *55*, 40–52. https://doi.org/10.1016/j.advwatres.2012.08.010

López-Moreno, J. I., Leppänen, L., Luks, B., Holko, L., Picard, G., Sanmiguel-Vallelado, A., Alonso-González, E., Finger, D. C., Arslan, A. N., Gillemot, K., Sensoy, A., Sorman, A., Ertaş, M. C., Fassnacht, S. R., Fierz, C., & Marty, C. (2020). Intercomparison of measurements of bulk snow density and water equivalent of snow cover with snow core samplers: Instrumental bias and variability induced by observers. *Hydrological Processes*, *34*(14), 3120–3133. https://doi.org/10.1002/hyp.13785

Lowe, H., & Picard, G. (2015). Microwave scattering coefficient of snow in MEMLS and DMRT-ML revisited: The relevance of sticky hard spheres and tomography-based estimates of stickiness. *Cryosphere*, *9*(6), 2101–2117. https://doi.org/10.5194/tc-9-2101-2015

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Mortimer, C., Derksen, C., Mudryk, L., Moisander, M., Hiltunen, M., Smolander, T., Ikonen, J., Cohen, J., Salminen, M., Norberg, J., Veijola, K., & Venäläinen, P. (2021). GlobSnow v3.0 Northern Hemisphere snow water equivalent dataset. *Scientific Data*, *8*(1). https://doi.org/10.1038/s41597-021-00939-2

Makridakis, S. (1993). Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting*, *9*, 527–529. https://doi.org/https://doi-org.proxy.lib.uwaterloo.ca/10.1016/0169-2070(93)90079-3

Malmros, J. K., Mernild, S. H., Wilson, R., Tagesson, T., & Fensholt, R. (2018). Snow cover and snow albedo changes in the central Andes of Chile and Argentina from daily MODIS observations (2000–2016). *Remote Sensing of Environment*, *209*, 240–252. https://doi.org/10.1016/j.rse.2018.02.072

Manuele, J., & Goffman, C. (1945). Statistical Tools for Controlling Quality. *Transactions of the American Institute of Electrical Engineers*, *64*(7), 524–528. https://doi.org/https://doi-org.proxy.lib.uwaterloo.ca/10.1109/EE.1945.6441168

Markus, T., Powell, D. C., & Wang, J. R. (2006). Sensitivity of passive microwave snow depth retrievals to weather effects and snow evolution. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(1), 68–77. https://doi.org/10.1109/TGRS.2005.860208

Marsh, P., & Pomeroy, J. W. (1996). Meltwater fluxes at an arctic forest-tundra site. *Hydrological Processes*, *10*(10), 1383–1400. https://doi.org/10.1002/(sici)1099-1085(199610)10:10<1383::aid-hyp468>3.0.co;2-w

Marsh, P., & Woo, M. -K. (1984). Wetting front advance and freezing of meltwater within a snow cover: 1. Observations in the Canadian Arctic. *Water Resources Research*, *20*(12), 1853–1864. https://doi.org/10.1029/WR020i012p01853

Martineau, C., Langlois, A., Gouttevin, I., Neave, E., & Johnson, C. A. (2022). Improving Peary Caribou Presence Predictions in MaxEnt Using Spatialized Snow Simulations. *Arctic*, *75*(1), 55–71. https://doi.org/10.14430/arctic74868

Mätzler, C., Schanda, E., & Good, W. (1982). Towards the Definition of Optimum Sensor Specifications for Microwave Remote Sensing of Snow. *IEEE Transactions on Geoscience and Remote Sensing*, *20*(1), 57–66. https://doi.org/10.1109/TGRS.1982.4307521

Meloche, J., Langlois, A., Rutter, N., Royer, A., King, J., Walker, B., Marsh, P., & Wilcox, E. J. (2022). Characterizing tundra snow sub-pixel variability to improve brightness temperature estimation in satellite SWE retrievals. *Cryosphere*, *16*(1), 87–101. https://doi.org/10.5194/tc-16-87-2022

Molotch, N. P., Barnard, D. M., Burns, S. P., & Painter, T. H. (2016). Measuring spatiotemporal variation in snow optical grain size under a subalpine forest canopy using contact spectroscopy. *Water Resources Research*, *52*(9), 7513–7522. https://doi.org/10.1002/2016WR018954

Mortimer, C., Mudryk, L., Derksen, C., Brady, M., Luojus, K., Venäläinen, P., Moisander, M., Lemmetyinen, J., Takala, M., Tanis, C., & Pulliainen, J. (2022). Benchmarking algorithm changes to the Snow CCI+ snow water equivalent product. *Remote Sensing of Environment*, *274*. https://doi.org/10.1016/j.rse.2022.112988

Mortimer, C., Mudryk, L., Derksen, C., Luojus, K., Brown, R., Kelly, R., & Tedesco, M. (2020). Evaluation of long-term Northern Hemisphere snow water equivalent products. *Cryosphere*, *14*(5), 1579–1594. https://doi.org/10.5194/tc-14-1579-2020

Naderpour, R., Schwank, M., Mätzler, C., Lemmetyinen, J., & Steffen, K. (2017). Snow Density and Ground Permittivity Retrieved From L-Band Radiometry: A Retrieval Sensitivity Analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(7), 3148–3161. https://doi.org/10.1109/JSTARS.2017.2669336

Nash, J. E., & Sutcliffe, J. V. (1970). River Flow Forecasting Through Conceptual Models Part I-A Discussion of Principles. *Journal of Hydrology*, *10*, 282–290. https://doi.org/https://doi.org/10.1016/0022-1694(70)90255-6

Picard, G., Brucker, L., Roy, A., Dupont, F., Fily, M., Royer, A., & Harlow, C. (2013). Simulation of the microwave emission of multi-layered snowpacks using the Dense Media Radiative transfer theory: the DMRT-ML model. *Geoscientific Model Development*, *6*(4), 1061–1078. https://doi.org/10.5194/gmd-6-1061-2013

Picard, G., Sandells, M., & Löwe, H. (2018). SMRT: An active-passive microwave radiative transfer model for snow with multiple microstructure and scattering formulations (v1.0). *Geoscientific Model Development*, *11*(7), 2763–2788. https://doi.org/10.5194/gmd-11-2763-2018

Pomeroy, J. W., Marsh, P., & Gray, D. M. (1997). Application of a distributed blowing snow model to the arctic. *Hydrological Processes*, *11*(11), 1451–1464. https://doi.org/10.1002/(sici)1099-1085(199709)11:11<1451::aid-hyp449>3.0.co;2-q

Pulliainen, J. (2006). Mapping of snow water equivalent and snow depth in boreal and sub-arctic zones by assimilating space-borne microwave radiometer data and ground-based observations. *Remote Sensing of Environment*, *101*(2), 257–269. https://doi.org/10.1016/j.rse.2006.01.002

Pulliainen, J., Karna, J.-P., & Hallikainen, M. (1993). Development of Geophysical Retrieval Algorithms for the MIMR. *IEEE Transactions on Geoscience and Remote Sensing*, *31*(1). https://doi.org/https://doi-org.proxy.lib.uwaterloo.ca/10.1109/36.210466

Pulliainen, J., Luojus, K., Derksen, C., Mudryk, L., Lemmetyinen, J., Salminen, M., Ikonen, J., Takala, M., Cohen, J., Smolander, T., & Norberg, J. (2020). Patterns and trends of Northern Hemisphere snow mass from 1980 to 2018. *Nature*, *581*(7808), 294–298. https://doi.org/10.1038/s41586-020-2258-0

Raleigh, M. S., Livneh, B., Lapo, K., & Lundquist, J. D. (2016). How Does Availability of Meteorological Forcing Data Impact Physically Based Snowpack Simulations? *Journal of Hydrometeorology*, *17*(1), 99–120. https://doi.org/10.1175/JHM-D-14

Raleigh, M. S., Lundquist, J. D., & Clark, M. P. (2015). Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrology and Earth System Sciences*, *19*(7), 3153–3179. https://doi.org/10.5194/hess-19-3153-2015

Rango, A., Chang, A. T. C., & Foster, J. L. (1979). The Utilization of Spaceborne Microwave Radiometers for Monitoring Snowpack Properties. *Hydrology Research*, *10*(1), 25–40. http://iwaponline.com/hr/article-pdf/10/1/25/9657/25.pdf

Rees, A., Lemmetyinen, J., Derksen, C., Pulliainen, J., & English, M. (2010). Observed and modelled effects of ice lens formation on passive microwave brightness temperatures over snow covered tundra. *Remote Sensing of Environment*, *114*(1), 116–126. https://doi.org/10.1016/j.rse.2009.08.013

Rosner, B. (1983). Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics*, *25*(2), 165–172. https://doi.org/https://doi-org.proxy.lib.uwaterloo.ca/10.2307/1268549

Saberi, N., Kelly, R., Flemming, M., & Li, Q. (2020). Review of snow water equivalent retrieval methods using spaceborne passive microwave radiometry. *International Journal of Remote Sensing*, *41*(3), 996–1018. https://doi.org/10.1080/01431161.2019.1654144

Saberi, N., Kelly, R., Pan, J., Durand, M., Goh, J., & Scott, K. A. (2021). The Use of a Monte Carlo Markov Chain Method for Snow-Depth Retrievals: A Case Study Based on Airborne Microwave Observations and Emission Modeling Experiments of Tundra Snow. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(3), 1876–1889. https://doi.org/10.1109/TGRS.2020.3004594

Saberi, N., Kelly, R., Toose, P., Roy, A., & Derksen, C. (2017). Modeling the observed microwave emission from shallow multi-layer Tundra Snow using DMRT-ML. *Remote Sensing*, *9*(12). https://doi.org/10.3390/rs9121327

Schwank, M., & Naderpour, R. (2018). Snow density and ground permittivity retrieved from L-band radiometry: Melting effects. *Remote Sensing*, *10*(2). https://doi.org/10.3390/rs10020354

Serreze, M. C., & Barry, R. G. (2011). Processes and impacts of Arctic amplification: A research synthesis. *Global and Planetary Change*, *77*(1–2), 85–96. https://doi.org/10.1016/j.gloplacha.2011.03.004

Sexstone, G. A., & Fassnacht, S. R. (2014). What drives basin scale spatial variability of snowpack properties in northern Colorado? *Cryosphere*, *8*(2), 329–344. https://doi.org/10.5194/tc-8-329-2014

Smith, T., & Bookhagen, B. (2016). Assessing uncertainty and sensor biases in passive microwave data across High Mountain Asia. *Remote Sensing of Environment*, *181*, 174–185. https://doi.org/10.1016/j.rse.2016.03.037

Sturm, M., & Benson, C. S. (1997). Vapor transport, grain growth and depth-hoar development in the subarctic snow. *Journal of Glaciology*, *43*(143). https://doi.org/10.3189/S0022143000002793

Sturm, M., Goldstein, M. A., & Parr, C. (2017). Water and life from snow: A trillion dollar science question. *Water Resources Research*, *53*(5), 3534–3544. https://doi.org/10.1002/2017WR020840

Sturm, M., Grenfell, T. C., & Perovich, D. K. (1993). Passive microwave measurements of tundra and taiga snow covers in Alaska, USA. *Annals of Glaciology*, *17*, 125–130. https://doi.org/10.3189/s0260305500012714

Sturm, M., & Holmgren, J. (1998). Differences in compaction behavior of three climate classes of snow. *Annals of Glaciology*, *26*, 125–130. https://doi.org/10.3189/1998aog26-1-125-130

Sturm, M., Holmgren, J., & Liston, G. E. (1995). A Season Snow Cover Classification System for Local to Global Applications. *Journal of Climate*, *8*(5), 1261–1283. https://doi.org/10.1175/1520-0442(1995)008<1261:ASSCCS>2.0.CO;2

Sturm, M., Mcfadden, J. P., Liston, G. E., Stuart Chapin Iii, # F, Racine, C. H., & Holmgren, J. (2001). Snow-Shrub Interactions in Arctic Tundra: A Hypothesis with Climatic Implications. *Journal of Climate*, *14*, 336–344. https://doi.org/https://doi-org.proxy.lib.uwaterloo.ca/10.1175/1520-0442(2001)014%3C0336:SSIIAT%3E2.0.CO;2

Sturm, M., Taras, B., Liston, G. E., Derksen, C., Jonas, T., & Lea, J. (2010). Estimating snow water equivalent using snow depth data and climate classes. *Journal of Hydrometeorology*, *11*(6), 1380–1394. https://doi.org/10.1175/2010JHM1202.1

Sullender, B. K., Cunningham, C. X., Lundquist, J. D., & Prugh, L. R. (2023). Defining the danger zone: critical snow properties for predator–prey interactions. *Oikos*. https://doi.org/10.1111/oik.09925

Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J. P., Koskinen, J., & Bojkov, B. (2011). Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based

measurements. *Remote Sensing of Environment*, *115*(12), 3517–3529. https://doi.org/10.1016/j.rse.2011.08.014

Tedesco, M., & Jeyaratnam, J. (2016). A new operational snow retrieval algorithm applied to historical AMSR-E brightness temperatures. *Remote Sensing*, *8*(12). https://doi.org/10.3390/rs8121037

Terasmae, J., & Reeves, A. (2017). *Tundra*. The Canadian Encyclopedia. Retrieved from https://www.thecanadianencyclopedia.ca/en/article/tundra

Tong, J., Déry, S. J., Jackson, P. L., & Derksen, C. (2010). Testing snow water equivalent retrieval algorithms for passive microwave remote sensing in an alpine watershed of western Canada. *Canadian Journal of Remote Sensing*, *36*(1), 74–86. https://doi.org/https://doi-org.proxy.lib.uwaterloo.ca/10.5589/m10-009

Vargel, C., Royer, A., St-Jean-Rondeau, O., Picard, G., Roy, A., Sasseville, V., & Langlois, A. (2020). Arctic and subarctic snow microstructure analysis for microwave brightness temperature simulations. *Remote Sensing of Environment*, *242*. https://doi.org/10.1016/j.rse.2020.111754

Venäläinen, P., Luojus, K., Lemmetyinen, J., Pulliainen, J., Moisander, M., & Takala, M. (2021). Impact of dynamic snow density on GlobSnow snow water equivalent retrieval accuracy. *Cryosphere*, *15*(6), 2969–2981. https://doi.org/10.5194/tc-15-2969-2021

Vionnet, V., Mortimer, C., Brady, M., Arnal, L., & Brown, R. (2021). Canadian historical Snow Water Equivalent dataset (CanSWE, 1928–2020). *Earth System Science Data*, *13*(9), 4603–4619. https://doi.org/10.5194/essd-13-4603-2021

Wang, H., Zhang, X., Xiao, P., Che, T., Zheng, Z., Dai, L., & Luan, W. (2023). Towards large-scale daily snow density mapping with spatiotemporally aware model and multi-source data. *Cryosphere*, *17*(1), 33–50. https://doi.org/10.5194/tc-17-33-2023

Woo, M.-K. (1998). Arctic Snow Cover Information for Hydrological Investigations at Various Scales. *Nordic Hydrology*, *29*(5), 245–266. https://doi.org/https://doi.org/10.2166/nh.1998.0014

Woo, M.-K., & Marsh, P. (1978). Analysis of Error in the Determination of Snow Storage for Small High Arctic Basins. *American Meteorological Society*, *78*, 1537–1541. https://doi.org/https://doi.org/10.1175/1520-0450(1978)017%3C1537:AOEITD%3E2.0.CO;2

Xu, X., Tsang, L., & Josberger, E. G. (2010). Dense Media Radiative Transfer theory for passive remote sensing and application to SWE retrieval. *11th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment*, 110–115. https://doi.org/10.1109/MICRORAD.2010.5559578

Xue, Y., Forman, B. A., & Reichle, R. H. (2018). Estimating Snow Mass in North America Through Assimilation of Advanced Microwave Scanning Radiometer Brightness Temperature Observations Using the Catchment Land Surface Model and Support Vector Machines. *Water Resources Research*, *54*(9), 6488–6509. https://doi.org/10.1029/2017WR022219

Yang, J. W., Jiang, L. M., Lemmetyinen, J., Pan, J. M., Luojus, K., & Takala, M. (2021). Improving snow depth estimation by coupling HUT-optimized effective snow grain size

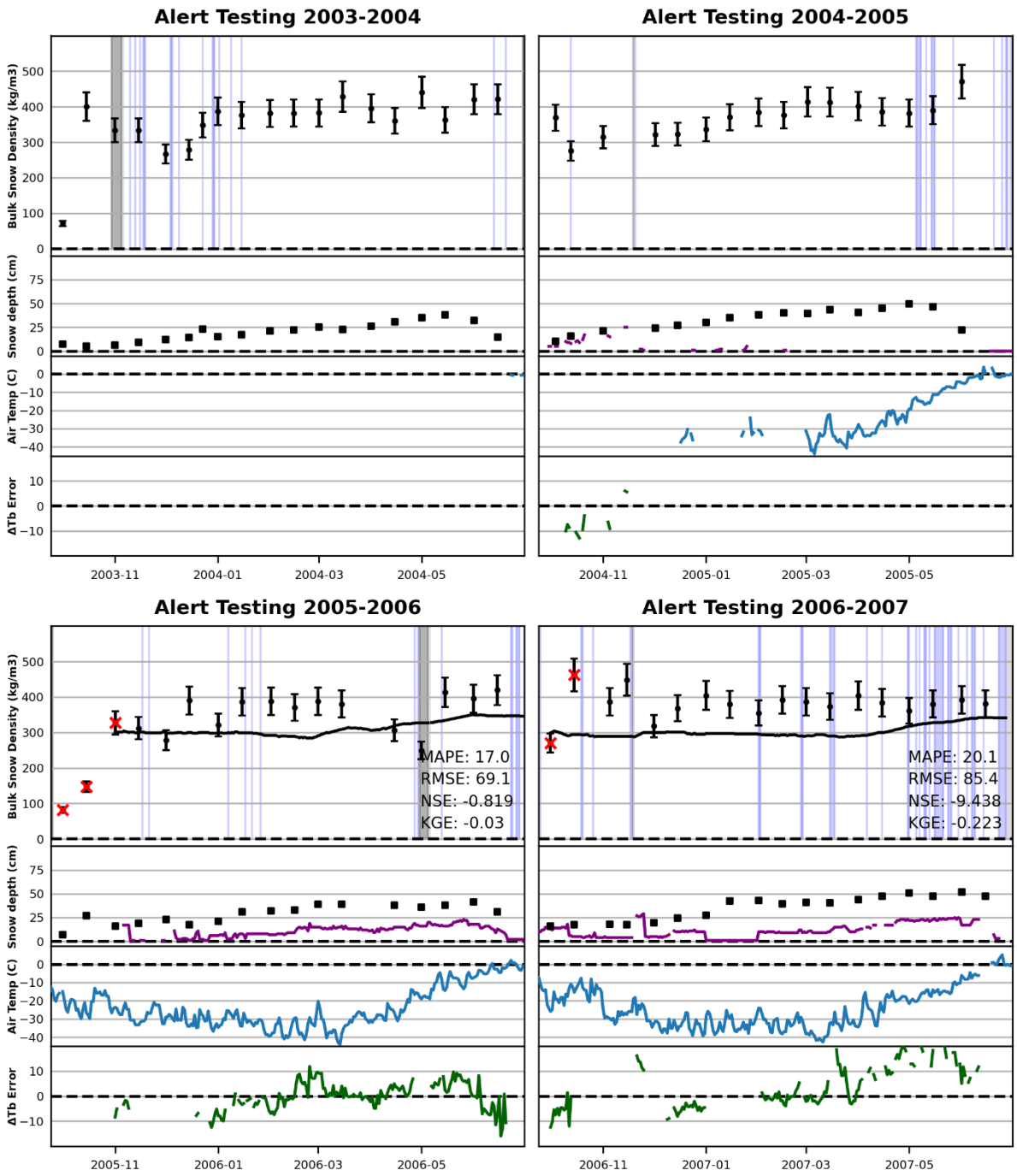
parameters with the random forest approach. *Remote Sensing of Environment*, *264*. https://doi.org/10.1016/j.rse.2021.112630

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., & Zhang, L. (2020). Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment*, *241*. https://doi.org/10.1016/j.rse.2020.111716

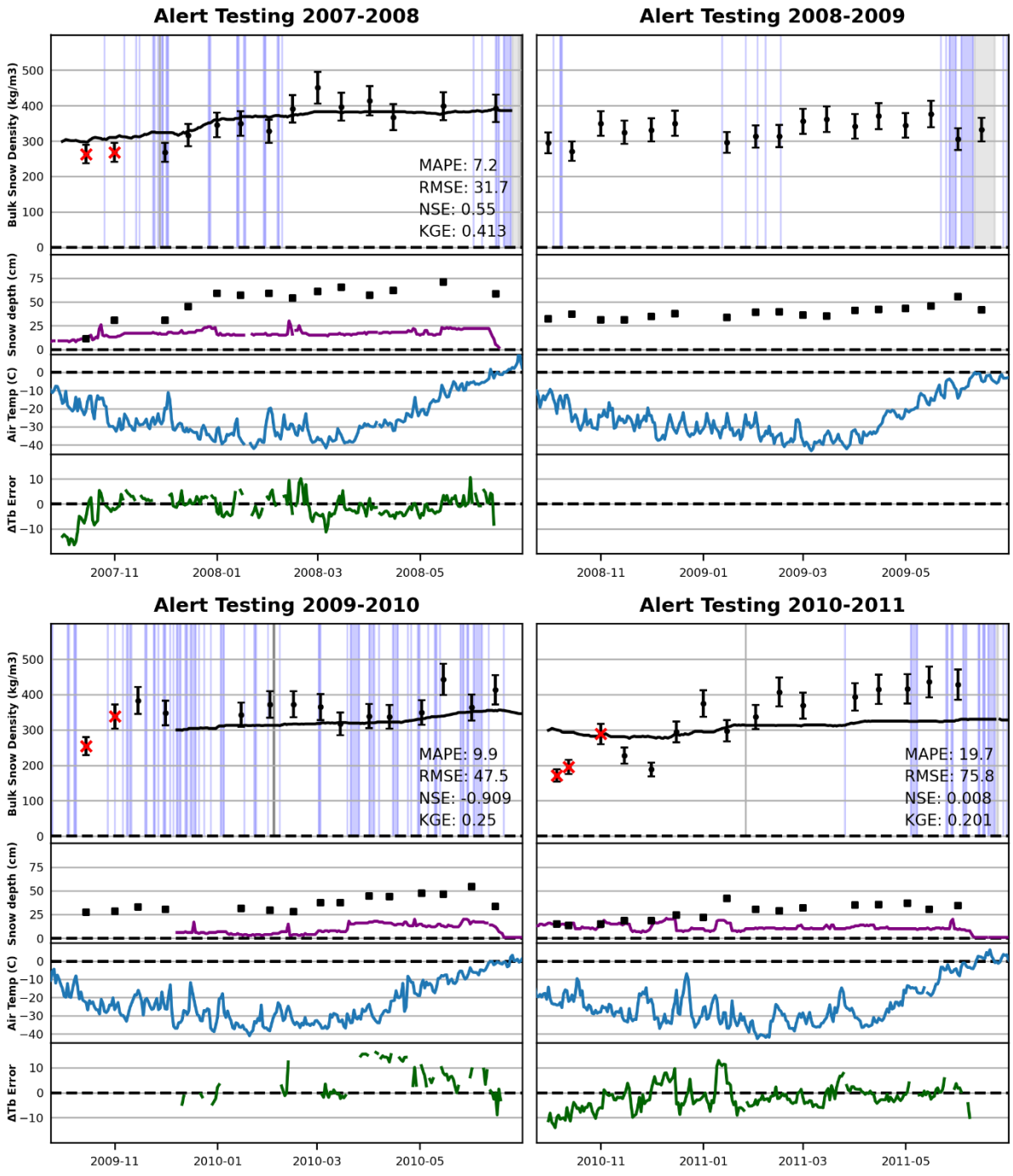
Appendix A

Appendix A contains summaries for all algorithm runs that could not be included in the text body for lack of space. The algorithm was applied at eight AWS sites over the course of eight winter seasons, for a total of 64 algorithm runs. On the following pages, summaries for algorithm runs are arranged in alphabetical order by site name, and the eight algorithm runs for each site span two pages in chronological order. Each summary includes algorithm estimates of snow density, model forcing data, and the microwave modelling error; each summary is structured as follows. The title describes the site name, stage of analysis (i.e. training, validation, or testing), and the winter season, in that order. Immediately underneath the title, the first plot shows the algorithm bulk snow density estimates (black line) and the corresponding *in situ* CanSWE density samples (black points, with ±10% uncertainty bars). CanSWE density samples marked with 'X's were removed from the analysis – red indicates samples prior to November 1st and black identified as an outlier by the GESD test. Also shown in the first plot are qualitative checks performed by the algorithm prior to the microwave retrieval process which were not passed and are represented as shaded regions (dark grey indicates missing satellite data, light grey indicates no volume scattering [i.e. cannot perform a retrieval], blue indicates snow not detected by Grody and Basit (1996) in the observation scene). The second plot shows the AWS snow depth measurements (purple line) used for model forcing and the *in situ* CanSWE depth samples (black squares) for reference. The third plot shows the air temperature measured at the AWS used for model forcing. The fourth plot shows the microwave modelling error (i.e. the difference between simulated and observed ΔTb).

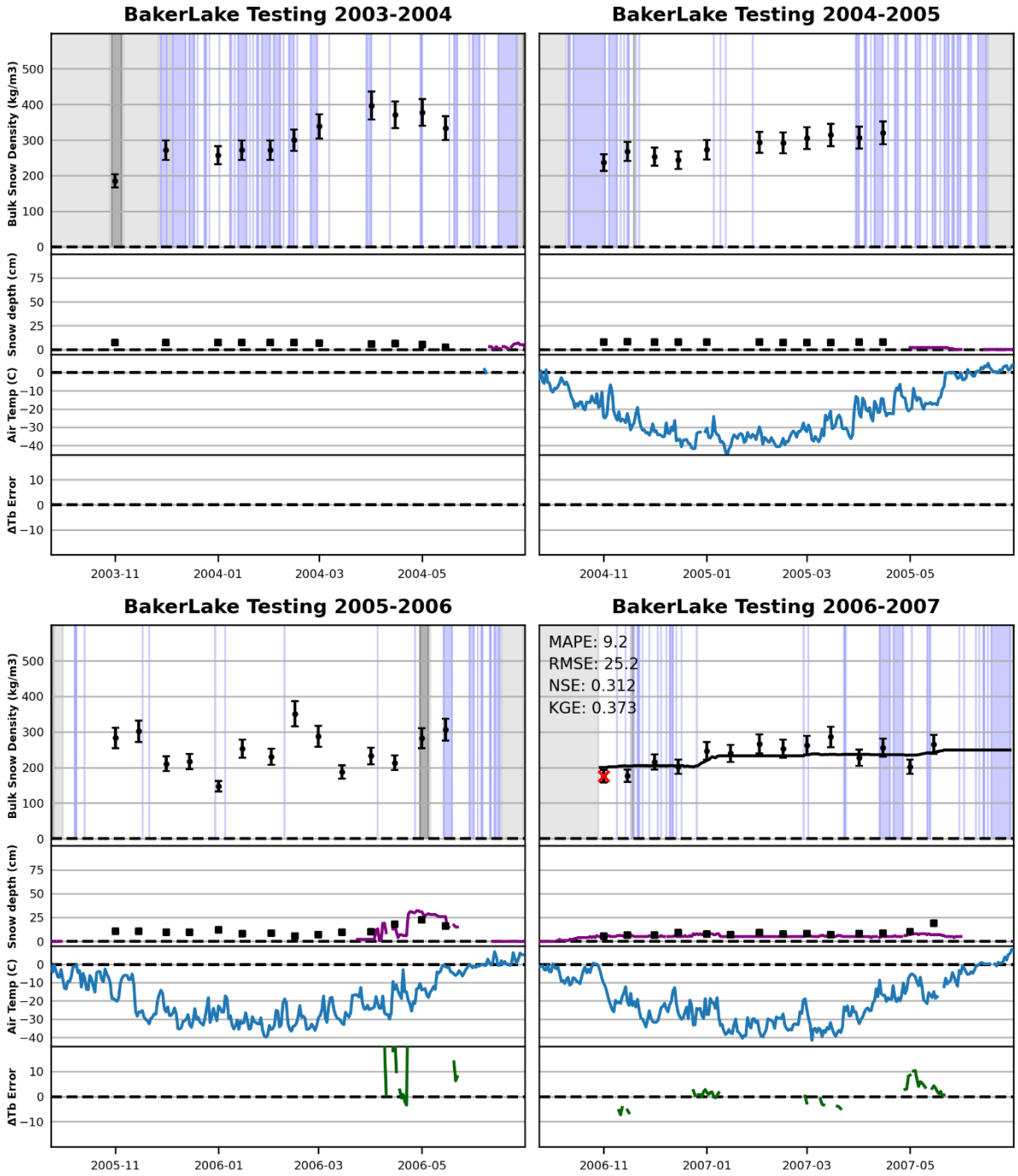
Alert Results



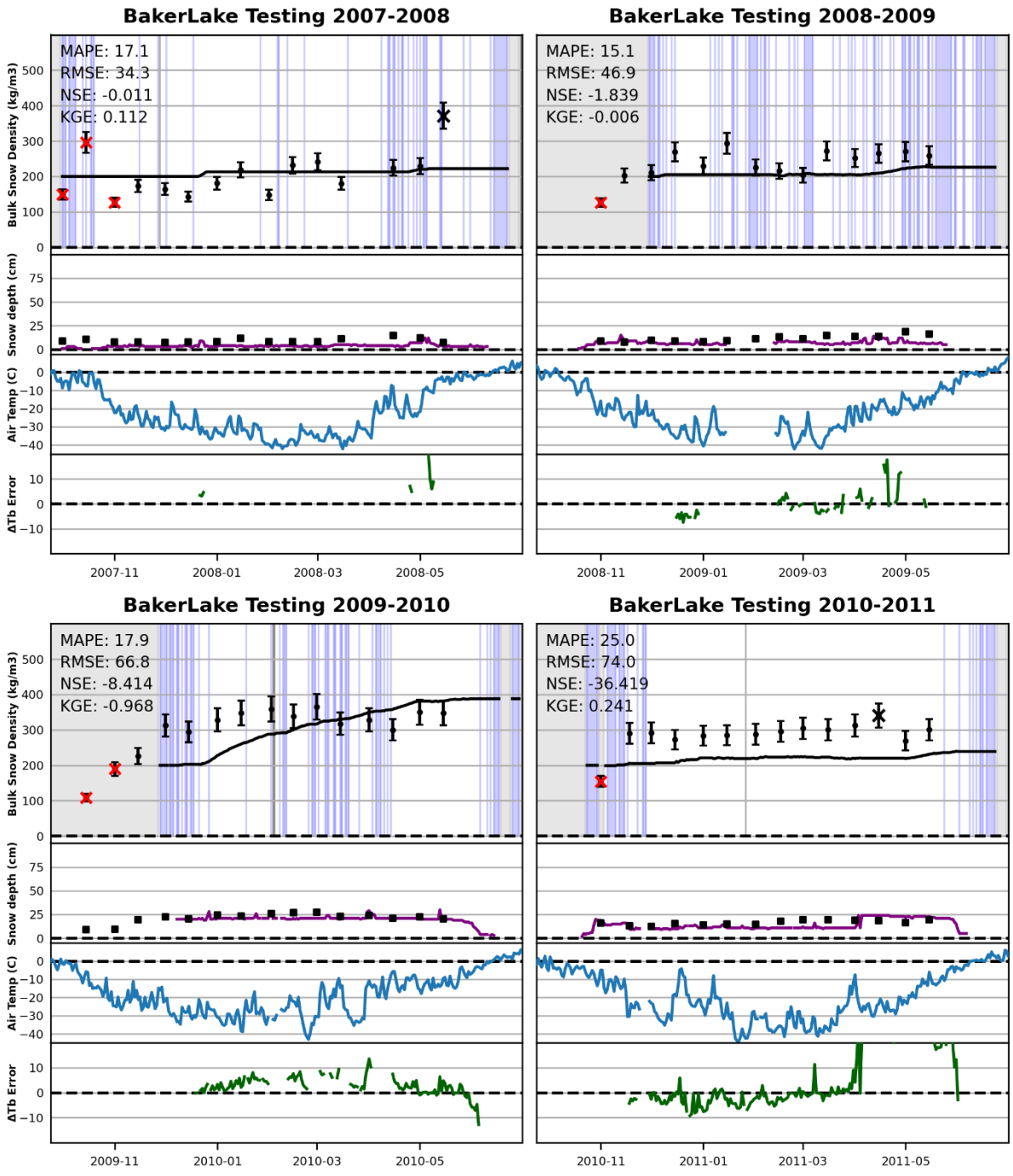
Alert Results cont'd



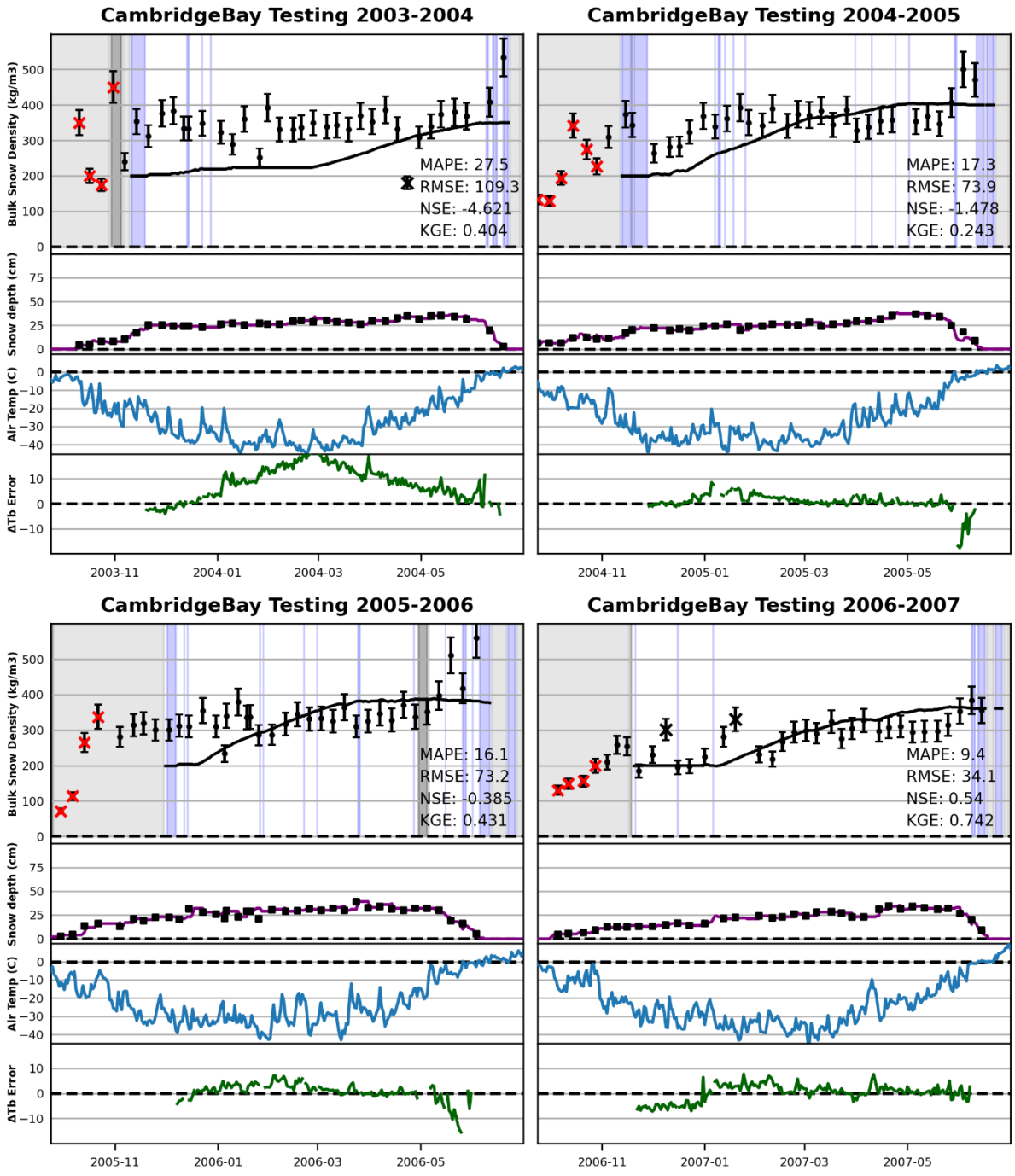
Baker Lake Results



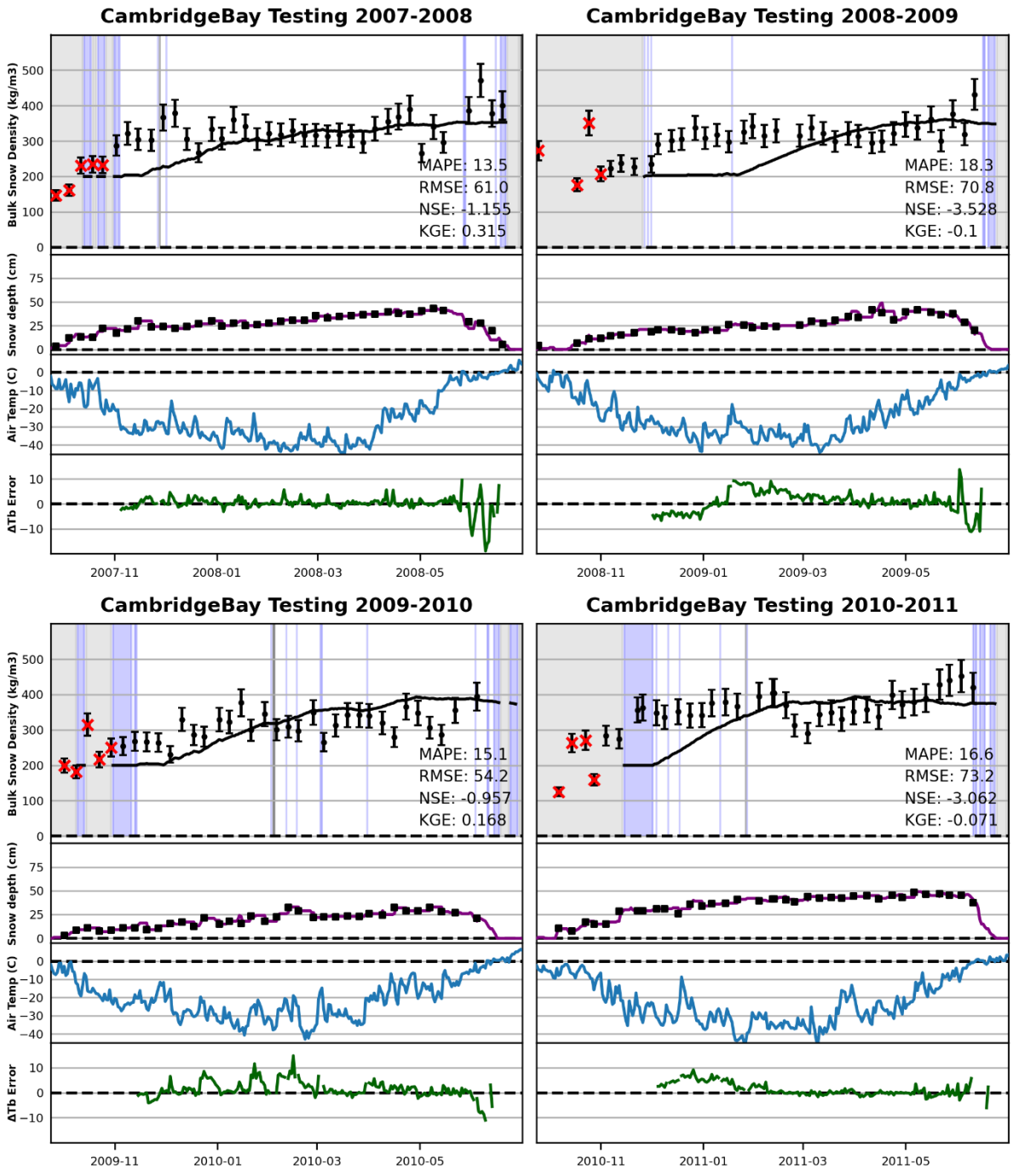
Baker Lake Results, cont'd



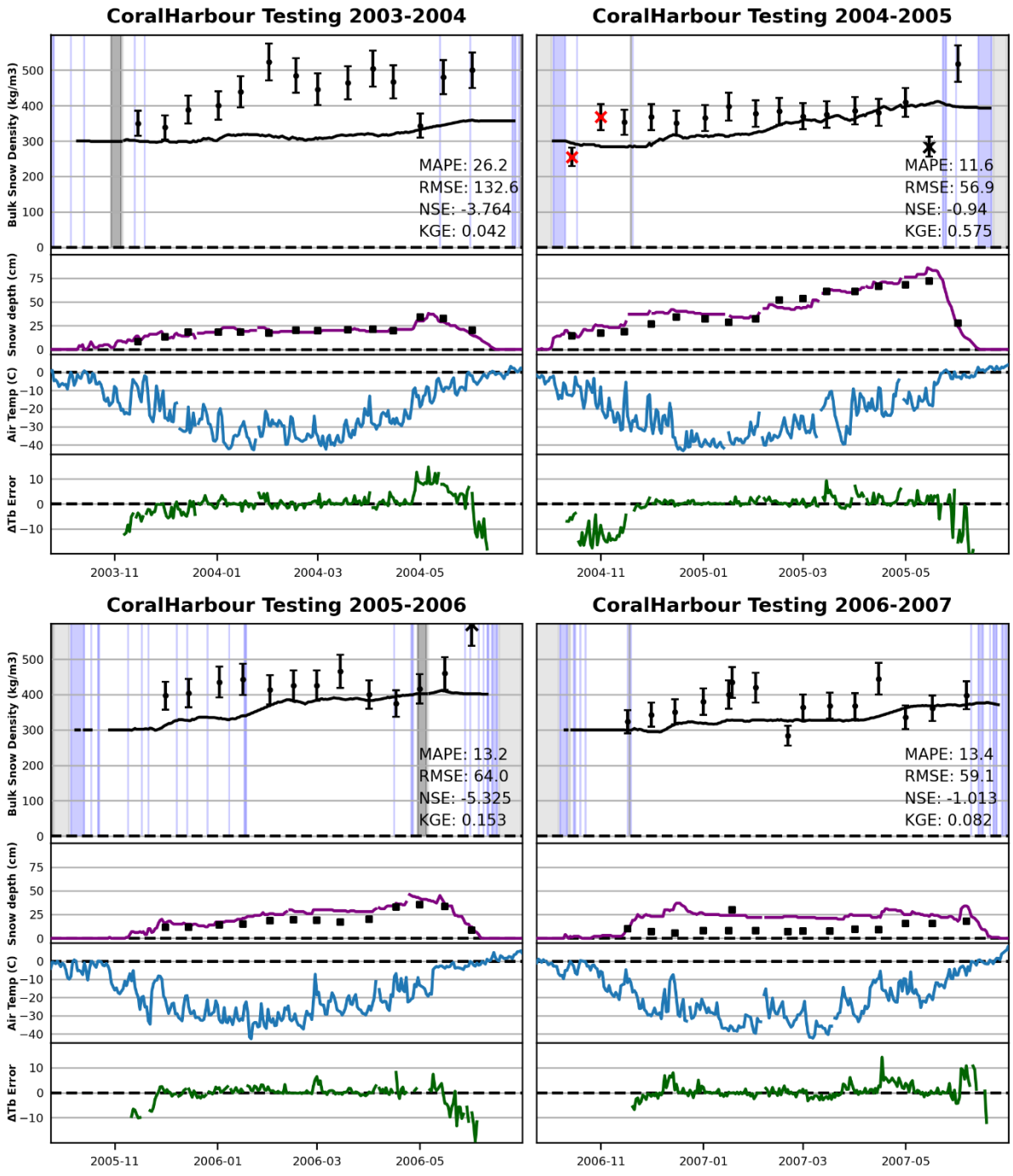
Cambridge Bay Results



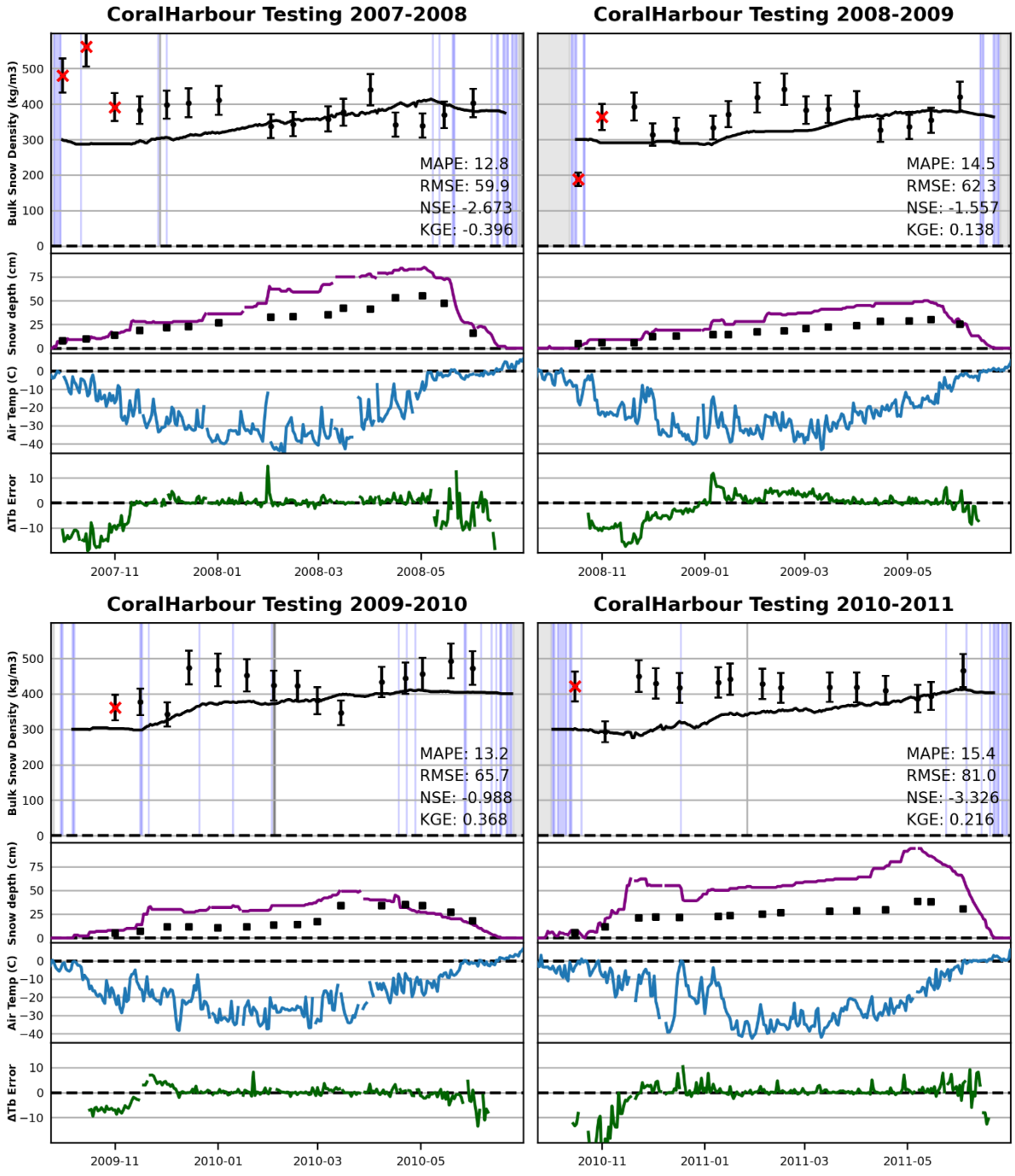
Cambridge Bay Results, cont'd



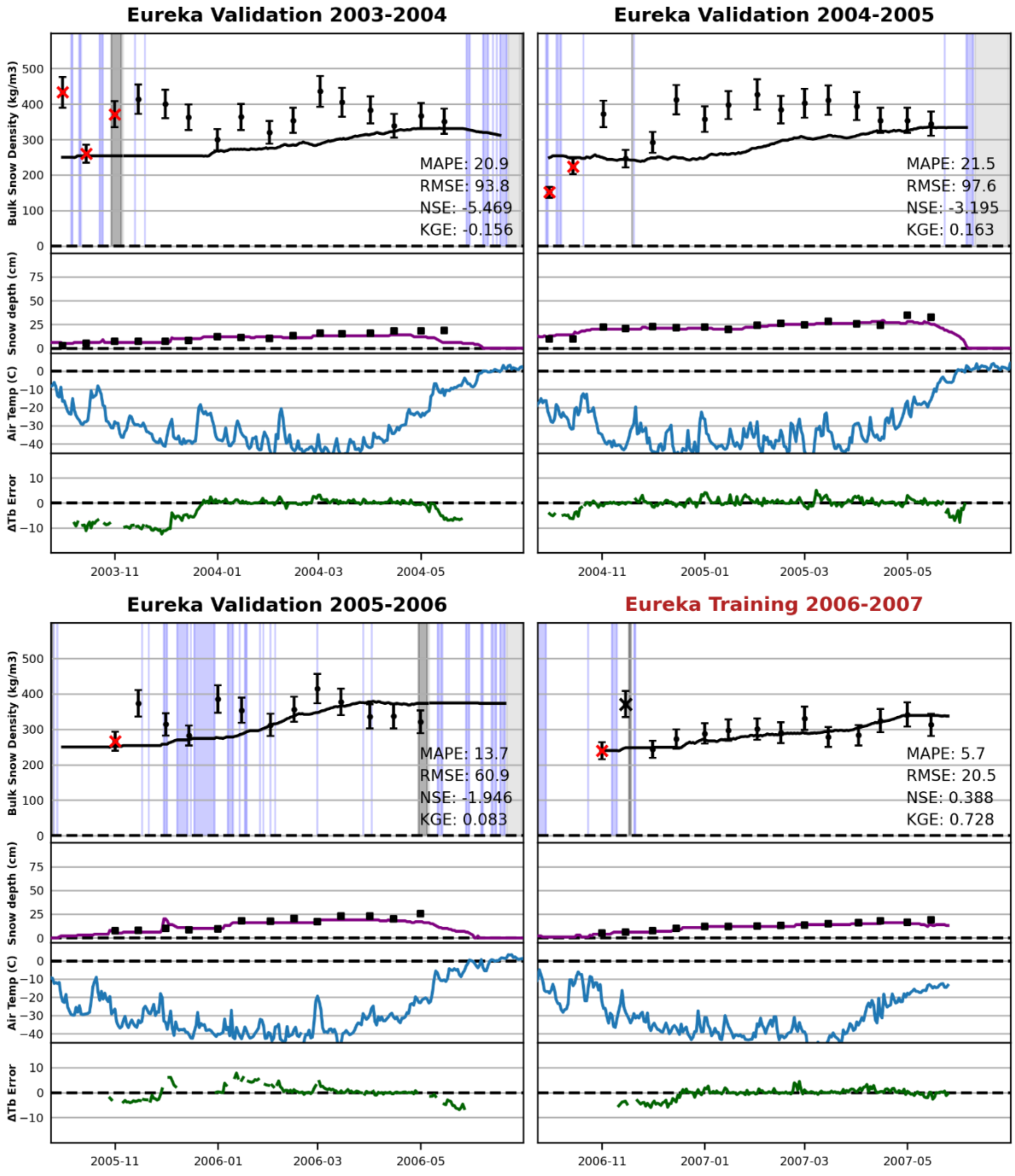
Coral Harbour Results



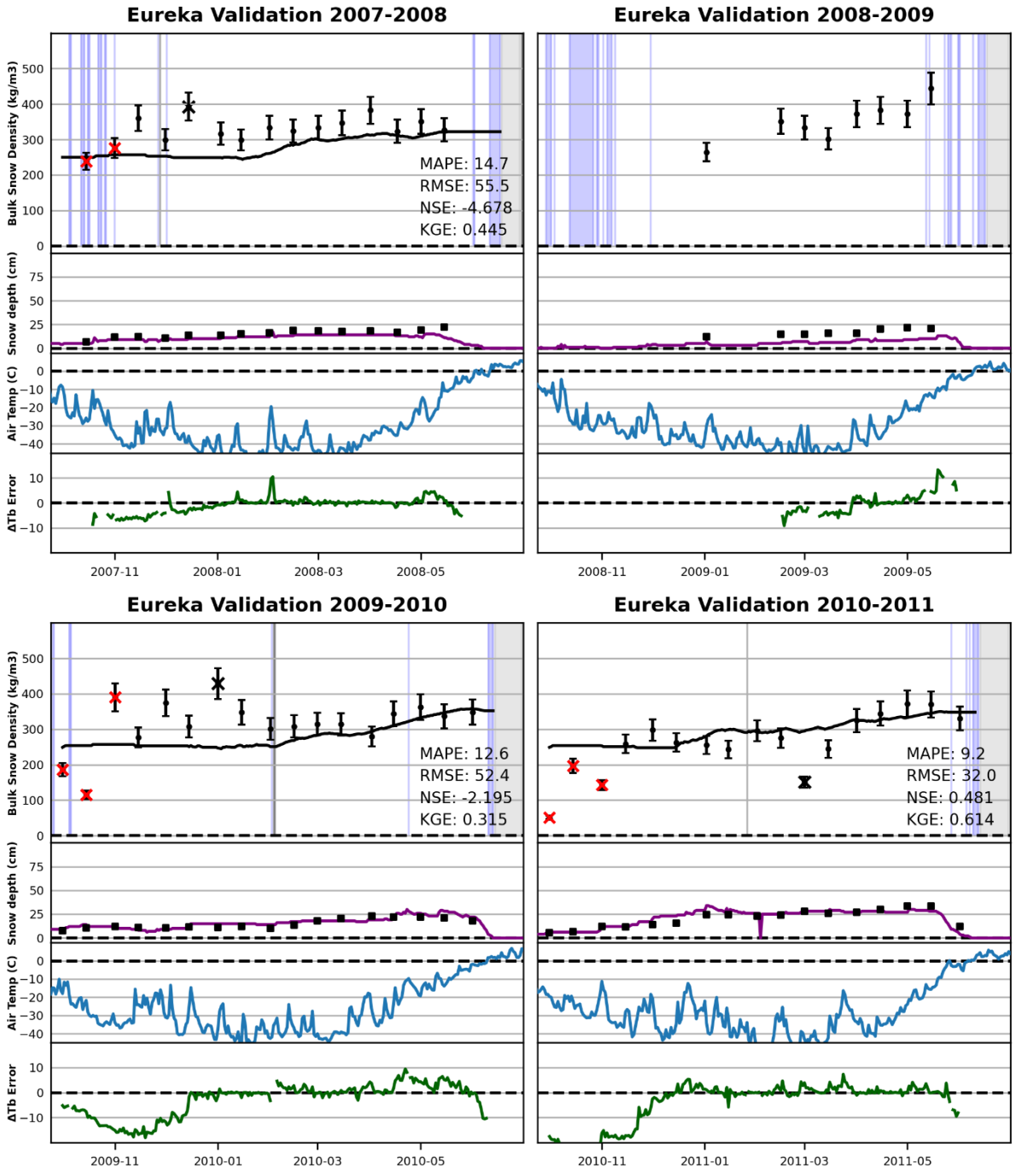
Coral Harbour Results, cont'd



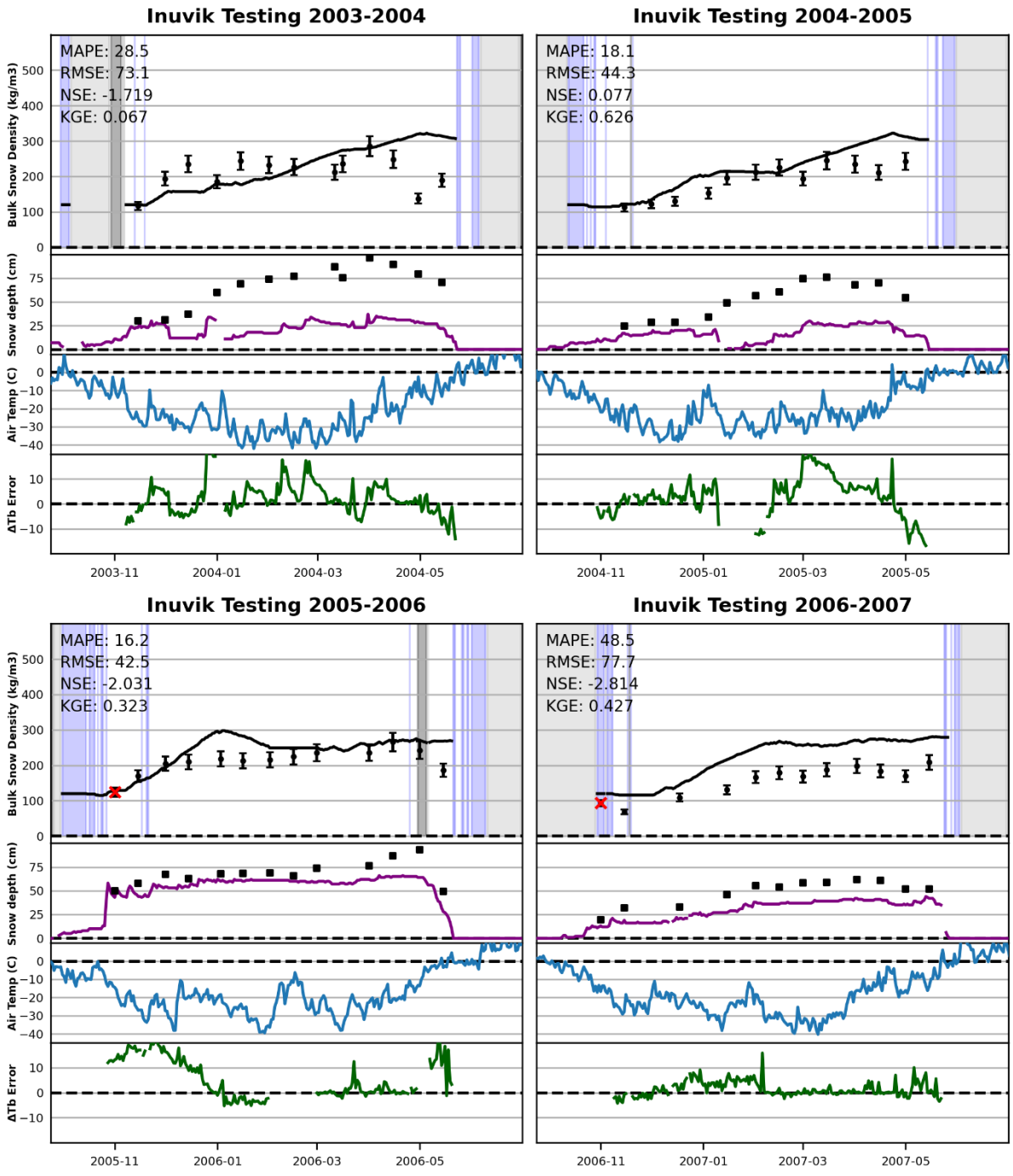
Eureka Results



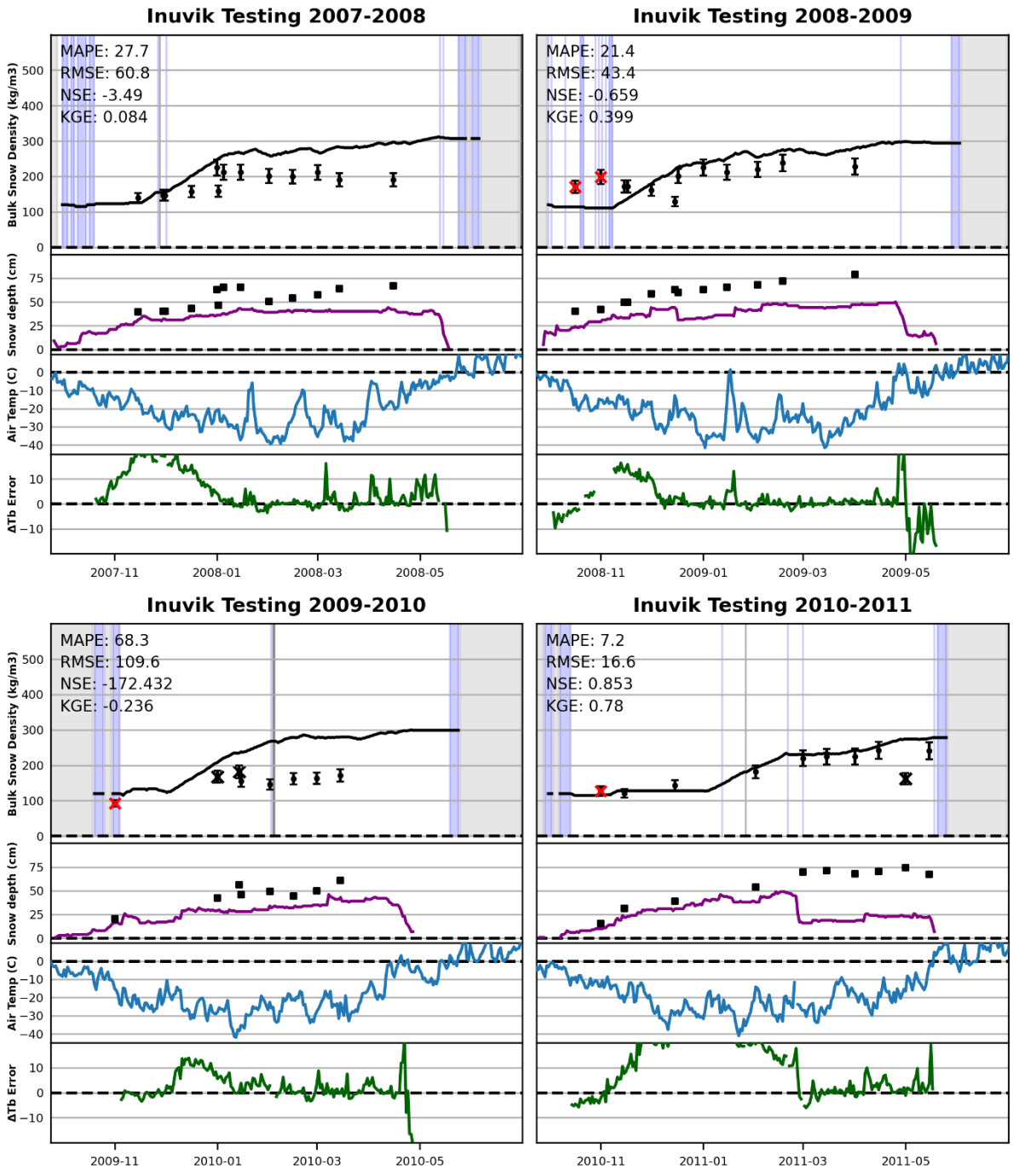
Eureka Results, cont'd



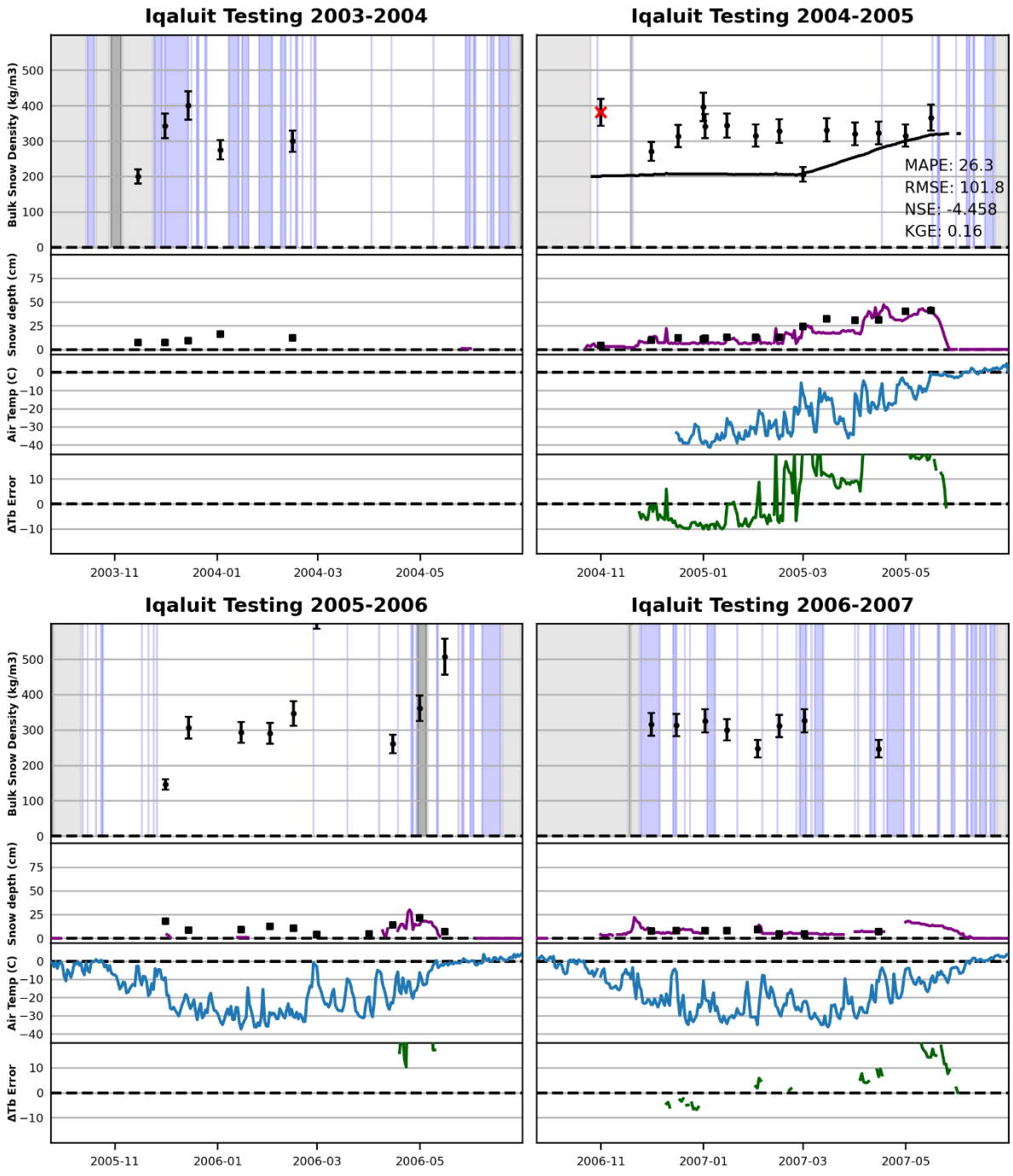
Inuvik Results



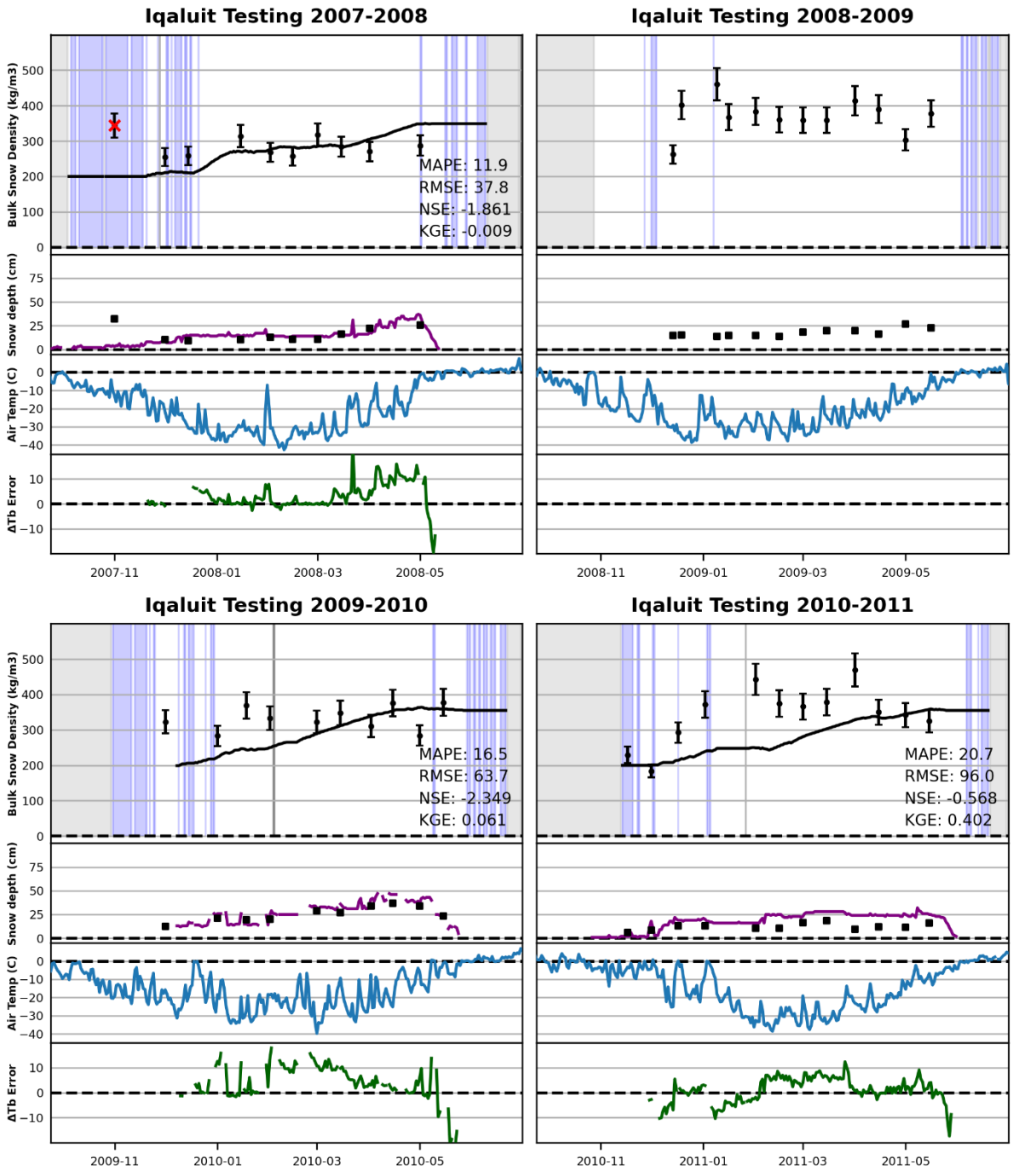
Inuvik Results, cont'd



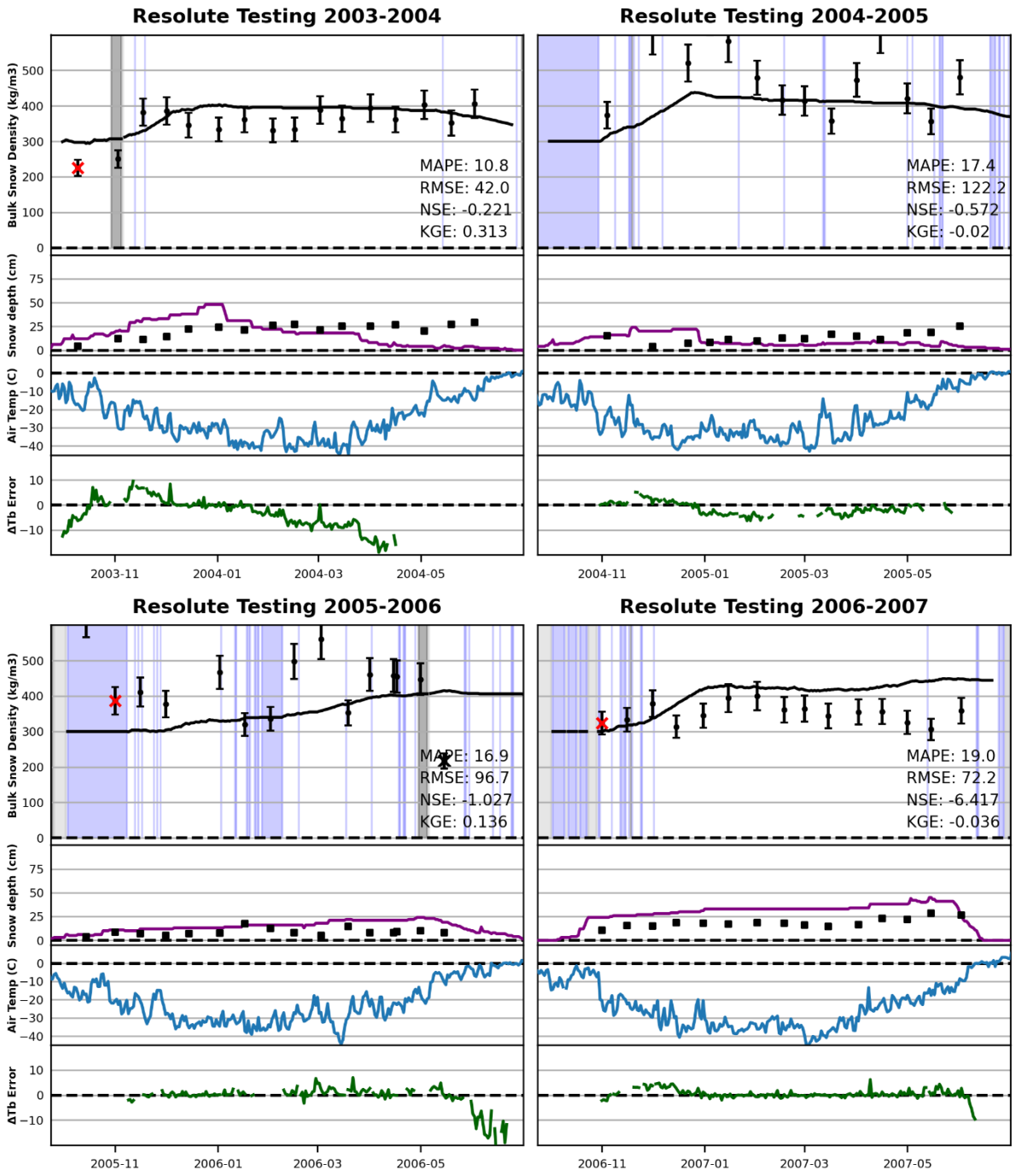
Iqaluit Results



Iqaluit Results, cont'd



Resolute Results



Resolute Results, cont'd