Improving Predictions of Precipitation Phase over Canada with the Development of a Spatially Aware Parameterization of Rain/Snow Partitioning

by

Charlie Ballantyne

A thesis
presented to the University of Waterloo
in fulfillment 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 my manuscript journal article on an improved parameterization of precipitation phase over Canada. The research presented here is my own work, with edits and comments provided by Dr. Chris Fletcher throughout the process. The final draft includes additional edits suggested by my committee members, Dr. Philip Marsh and Dr. Richard Kelly.

Abstract

Partitioning precipitation into rain or snow is an important aspect of hydrologic and climatological modelling, affecting a wide variety of downstream processes (Harpold et al., 2017). Current conventional methods typically rely on surface temperature, either as a strict threshold or to inform a probability function which predicts a 0% chance of rain at one temperature to 100% chance of rain at another (Jennings et al., 2019, Feiccabrino et al., 2015). However, recent studies have shown that variables such as wind, pressure, humidity, and atmospheric profiles of temperature can all have a significant effect on precipitation phase at the surface (Wang et al., 2019, Sims and Liu, 2015, Jennings et al., 2018). This study utilized CloudSat data of precipitation phase and associated environmental variables ground-truthed at ECCC stations across Canada to build an improved statistical parameterization of precipitation phase. Our results showed that using a random forest model with atmospheric profiles of wetbulb temperature in addition to surface wetbulb temperature, elevation, and wind resulted in a probability of detection of 97.8% across the -1, 4°C temperature interval, compared to a probability of detection of below 80% across the same interval for conventional methods. The random forest parameterization was also spatially robust, performing well on stations it had not been trained on. Additionally, adding Sturm's snow classes as an indicator variable to the model did not result in any significant improvement, indicating that a model trained on all available data is adequately able to capture spatial variability in rain-snow partitioning across the study area.

Acknowledgements

I would like to thank my parents, Andy and Laura, and my partner, Martin, for always being there for me and providing constant encouragement throughout this process. I'd also like to thank my supervisor, Dr. Chris Fletcher, for all his incredibly helpful feedback and coaching. His guidance has helped me become a stronger, more confident researcher. Finally, I'd like to thank all my other family, friends, labmates, and members of my graduate cohort for their support and general awesomeness. This research would not have been possible without you all.

Table of Contents

List of Figures

Chapter 1

Introduction

Precipitation phase partitioning, the separation of precipitation into rain or snow, is an important aspect of hydrologic and climatological modelling and inaccurate methods can lead to biases in estimates of SWE and snow depth and errors in streamflow, soil moisture, and land-atmosphere energy and water exchanges [Jennings et al., 2018, Wang et al., 2019]. Outside of the tropics, precipitation typically starts out as solid and its phase at the surface is determined by the properties of the atmosphere through which it passes as well as those of the hydrometeor itself [Harder and Pomeroy, 2013]. Conventional methods, which typically only incorporate ambient air temperatures at the surface, have been found to be inadequate at expressing the complex environmental factors which contribute to precipitation phase at the surface [Harpold et al., 2017, Harder and Pomeroy, 2013, Wang et al., 2019].

Previous research has indicated that incorporating a comprehensive set of environmental and spatial variables is an essential step for the creation of improved precipitation phase models [Harpold et al., 2017]. This first requires an understanding of the physical and spatial factors which influence the temperature at which the 50/50 rain-snow threshold occurs and the ability to incorporate this information into a parameterization which can accurately predict precipitation phase. In doing so, this will provide more realistic conditions for climate and hydrologic models and reduce bias in estimations of precipitation related factors such as snow depth and SWE.

This section will summarize previous research as it pertains to the classification of precipitation phase and the influence of various factors, both environmental and spatial, which can impact precipitation phase. It will focus on empirical estimation of precipitation phase, as that is where research is most needed in order to improve climate and hydrologic model outputs. The goal of this is to support our work on applying this knowledge to create a phase partitioning model which can be integrated into more general climate and hydrologic models.

1.1 Importance of Accurate Precipitation Phase Partitioning

While the 0°C mark is generally thought of as the point at which frozen precipitation melts, this threshold can vary depending on a variety of other factors, meaning that using a simple 0°C threshold is rarely accurate for precipitation phase partitioning [Harpold et al., 2017]. In reality, the phase of precipitation reaching the surface can be influenced by atmospheric conditions such as wind speed, humidity, the presence of inversions, lapse rate, and velocity of the hydrometeors [Feiccabrino et al., 2015]. This creates a high degree of uncertainty around the temperature at which the 50/50 rain-rain snow threshold occurs, with a reduction in skill across all methods in the range of -3C – 5C [Froidurot et al., 2014, Jennings et al., 2018].

Comparative studies of precipitation phase partitioning methods have shown that the selection of methods can significantly influence SWE and snow depth estimates. These uncertainties hold true across hydrological models. Comparisons between different phase partitioning methods using the Cold Regions Hydrological Model have shown that using an ambient temperature-based method can produce uncertainties in peak SWE of up to 160mm at the same site. In contrast, using a method which incorporates humidity can reduce these uncertainties, particularly in high elevation regions [Harder and Pomeroy, 2014]. Similarly, comparisons of the performance of wet bulb and ambient temperature thresholds ran using the Noah-MP Land Surface Model (LSM) found that the use of wet bulb temperature reduced bias in SWE estimates by 11.8% and 10.4% over the elevation ranges of 1,500-2,000m and 2,000-2,500m, respectively [Wang et al., 2019]. Finally, comparisons of a variety of methods ranging from simple thresholds to logistic regression results in the SNOWPACK model showed that the 0°C static temperature threshold and -0.5°C to 0.5°C ranged temperature threshold tended to underpredict snow accumulations while the 2°C and 3°C thresholds overpredicted it. The binary logistic regression, which incorporated both temperature and relative humidity, performed the best [Jennings and Molotch, 2019].

Accurate snowfall monitoring is especially important in cold countries such as Canada where a significant portion of precipitation falls as snow. In particular, predicting the fraction and conditions under which precipitation falls as snow is important for environmental modelling, snowpack monitoring, and flood forecasting [McAfee et al., 2014, Dai, 2008]. Changing rain/snow fraction can result in changes to snowpack and streamflow dynamics, affecting water availability for ecosystems and society [Harder and Pomeroy, 2014, Harpold et al., 2017]. Figure 1.1 illustrates how overpredicting the rain and snow portion in models can affect model predictions of the timing and magnitude of peak flow. As cold regions often rely on water stored as snowpack higher amounts of rain will result in a lowered water budget, increasing the likelihood of drought. Obtaining accurate estimates of snowpack, and by proxy snowfall, is important for water resource managers, agriculture, and wildfire prediction [Harpold et al., 2017].

1.2 Conventional Phase Partitioning Methods

Currently, there are a few different conventional methods for assessing precipitation phase, both by observation and empirically. While hydrologic and climate models rely on empirical phase partitioning methods to estimate the amount of precipitation falling as either rain or snow, it is important that these methods are tested against in-situ observations in order to ensure their accuracy. In turn, it is necessary to consider the strengths and weaknesses of in-situ observations to assess their suitability for the ground truthing of empirical methods.

One of the most common and well-established methods is the collection of in-situ ob-

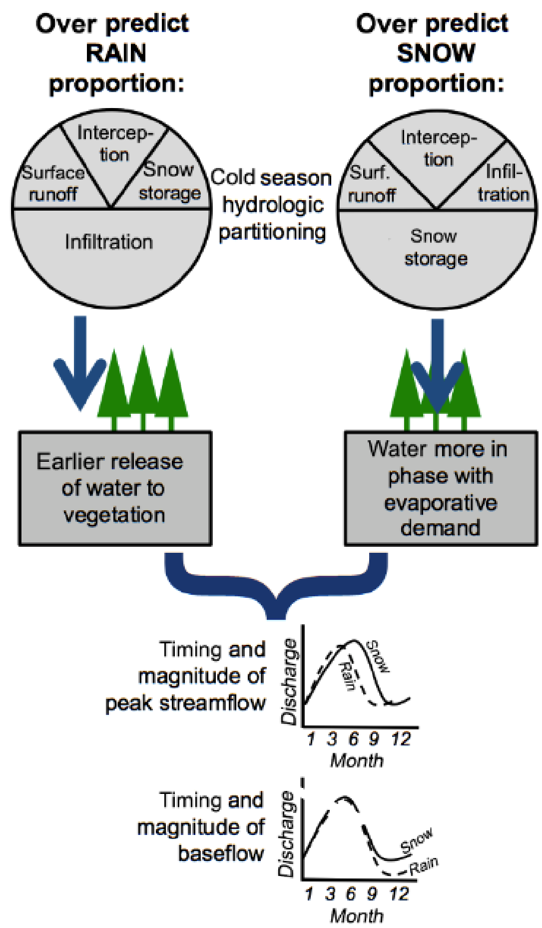

Figure 1.1: Precipitation phase has numerous implications for modeling the magnitude, storage, partitioning, and timing of water inputs and outputs. Potentially affecting important ecohydrological and streamflow quantities important for prediction. (Harpold et al., 2017)

servations from human observers, such as those stationed at manual stations or citizen scientists [Kodamana and Fletcher, 2021]. This data is generally well trusted and seen as valuable for snow-related studies but suffers issues such as the ability of observers to discern precipitation types in low-light conditions, a lack of observations in remote areas, and the difficulty in identifying mixed phase events [Harpold et al., 2017]. However, the difficulty of collecting in-situ observations is especially apparent in sparsely populated areas such as Northern Canada, where despite the importance of precipitation monitoring in the region the remoteness provides a major challenge for researchers. To get around this and other issues associated with human observations, remote sensing is also often used to detect precipitation phase [Harpold et al., 2017]. While this method still requires some degree of validation, it has been shown to be able to accurately assess precipitation phase with a high degree of accuracy [Kodamana and Fletcher, 2021]. In high latitude regions CloudSat has also been shown to be particularly suited to capturing snowfall related data due do its spatial coverage in higher latitudes and high sensitivity of the W-band CPR [Tang et al., 2017]

While satellite and ground-based radar is a valuable tool in retrieving precipitation estimates, instrumental and environmental factors can reduce its accuracy. For instance, attenuation from water vapour and hydrometeors can cause errors in CloudSat retrievals. Additionally, the accuracy of retrievals is significantly reduced in the bottom two kilometers of the atmosphere due to surface clutter caused by the interaction of the radar pulses with the surface [Kodamana and Fletcher, 2021, Harpold et al., 2017] To alleviate this, data from the bottom two kilometers is often cut out of analysis, which can create some inaccuracy in phase estimation of hydrometeors in this area. Mountainous regions in particular are vulnerable to this issue as inversions are common in these areas, meaning that hydrometeors can melt or refreeze relatively close to the surface [Feiccabrino et al., 2015].

Algorithmic precipitation partitioning using a temperature threshold is often used in modeling applications. There are a few common methods for estimating the probability of precipitation falling as either rain or snow, including a static threshold, linear transition, minimum and maximum temperature, or as a sigmoidal curve [Dai, 2008, Harpold et al., 2017, Wang et al., 2019]. A static threshold presumes that all precipitation occurring above a certain temperature is rain while all the precipitation occurring below said temperature is snow. This method ignores the occurrence of events such as mixed-phase precipitation [Kienzle, 2008]. A more in-depth approach involves the use of some probability function of precipitation phase that represents the probability of precipitation falling as snow from 0% at some temperature to 100% at another temperature. This probability function can either be linear or sigmoidal and is typically centered on the 0°C threshold [Dai, 2008, Wang et al., 2019]. In general, static threshold produce the highest error while sigmoidal relationships produced the lowest error [Harpold et al., 2017].

Modelling climate in complex terrain has always presented a major challenge to researchers and precipitation phase partitioning is no exception. Rain-snow transitions are widely variable across mountain terrain and temperature and humidity gradients along mountain slopes are poorly understood [Marks et al., 2013]. For instance, cold air damming caused by topographic barriers can create inversions and cause previously melted hydrometeors to partially or fully refreeze before reaching the ground [Feiccabrino et al., 2015]. Uncertainties are exacerbated due to the fact that networks of meteorological stations are often sparse in mountain areas and the ability of vertically pointing ground-based radar to retrieve data is significantly reduced due to the presence of physical barriers This creates a particular issue for regions that rely on water stored as snowpack in mountains, as it can lead to higher levels of error in SWE estimation [Harpold et al., 2017].

In order to address the shortcomings of current precipitation phase partitioning methods, a variety of new techniques have been introduced. These typically focus on the use of multivariate models in the form of a regression or machine learning model instead of a single or dual value ambient air temperature threshold [Jennings and Molotch, 2019]. The ability of these models to integrate multiple variables has made them a powerful alternative to conventional methods, though they are not yet commonly used operationally [Harpold et al., 2017].

1.3 Variables Relevant to Precipitation Phase

The state of the atmosphere is a major control on precipitation phase at the surface [Harpold et al., 2017]. A melting layer in the atmosphere, where ambient temperatures are above 0°C, must be of adequate size for precipitation to fully melt as it falls through in order to produce rain. The presence of an inversion, where precipitation falls through a melting layer and then refreezing layer, can generate freezing rain or sleet at the surface, Figure 1.2. Atmospheric temperature profiles can also produce mixed-phase precipitation, where precipitation is partially melted or re-frozen when it reaches the surface. Atmospheric settings with high lapse rates, where temperatures increase rapidly as hydrometeors near the surface, mean a smaller melting layer and less opportunity for precipitation to fully melt before it reaches the surface [Dai, 2008]. If surface conditions in this case are above 0°C, surface-based parameterizations might predict rain when precipitation is falling as snow.

While ambient temperature has traditionally been relied upon for phase partitioning, recent literature on precipitation phase has revealed a wide variety of complex environmental factors which contribute to phase at the surface. Factors such as humidity, which influences the energy balance of the hydrometeor, and pressure, which influences the fall rate of the hydrometeor and therefore the amount of time its spending in an above 0°C environment, all play a role in shifting the rain-snow threshold [Harpold et al., 2017]. In addition, the spatial variability of the rain-snow threshold demonstrates the need for increased spatial discrimination in phase partitioning models. It is therefore important to investigate and quantify the influence of a wide variety of environmental and spatial variables in order to determine the most effective way to improve precipitation phase models.

A growing body of research has shown that variables which incorporate both temperature and humidity, such as dewpoint and wetbulb temperature, can provide a far more accurate estimate of precipitation phase [Wang et al., 2019, Harder and Pomeroy, 2013]. Wetbulb temperature, approximated by Equation 1.1, defined as the lowest temperature achievable by the evaporation of water, has been found to be especially useful as it the closest approximation to the actual temperature of the hydrometeor [Harder and Pomeroy, 2013].

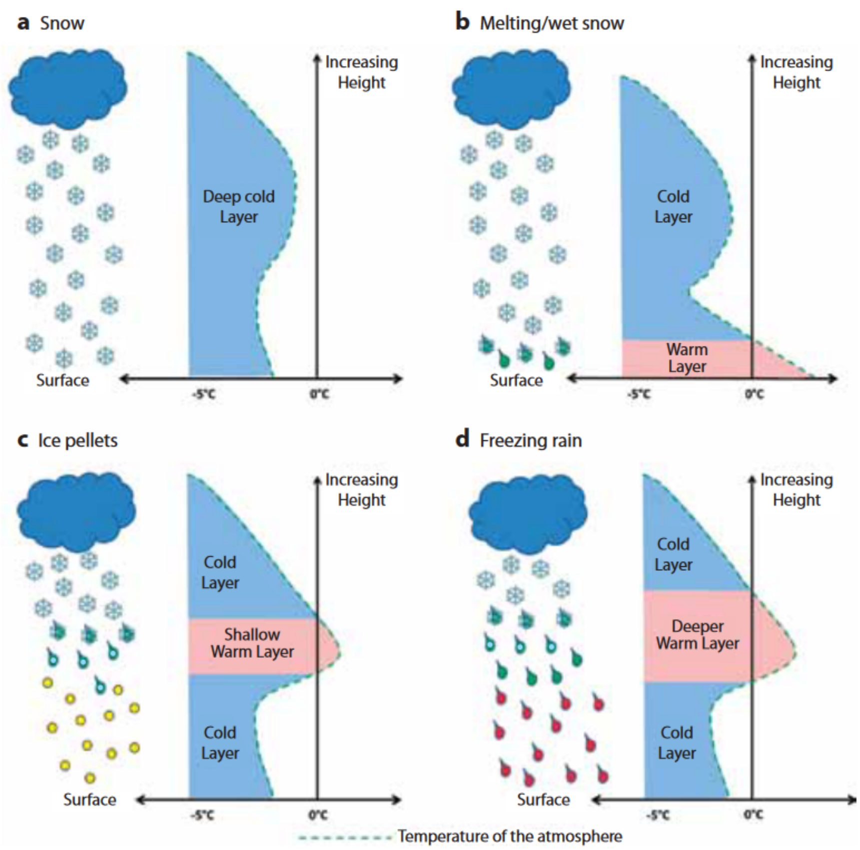

Figure 1.2: Schematic of typical temperature profiles for different precipitation types: (a) snow, (b) melting/wet snow, (c) ice pellets and (d) freezing rain (all assuming 100% relative humidity) (Forbes et al., 2014).

Evaporation and sublimation create latent heat exchanges at the surface of the hydrometeor, causing the hydrometeor surface to cool off relative to the ambient air temperature [Sims and Liu, 2015, Wang et al., 2019]. Unlike similar variables such as dewpoint temperature, approximated here by Equation 1.2 and defined as the temperature at which air becomes 100% saturated, and relative humidity, which is the percent saturation of the air, wetbulb temperature provides a direct physical relationship to the phase of the hydrometeor by taking into account the sensible and latent heat fluxes to its surface [Harder and Pomeroy, 2013, Lawrence, 2005].

$$\begin{aligned} T_w \approx T * atan(0.152 * (RH + 8.3136)^{\frac{1}{2}}) + atan(T + RH) - atan(RH - 1.6763) \\ + 0.00391838 * (RH)^{(\frac{3}{2})} * atan(0.0231 * RH) - 4.686 \end{aligned} \tag{1.1}$$

$$T_d \approx t - \frac{100 - RH}{5} \tag{1.2}$$

Low RH environments facilitate evaporative cooling through latent heat exchange, slowing down the melting of the hydrometeors [Harder and Pomeroy, 2013]. This means that the lower the RH in a given environment, the higher the rain/snow threshold. As illustrated in Figure 1.3, in observational studies across the Northern Hemisphere a 10% increase in RH has been associated with a 0.8°C decrease in the 50% rain-snow threshold. Meaning that at 40%-50% humidity the 50% rain-snow threshold corresponds to approximately 4.5°C whereas the 90%-100% humidity the 50% rain-snow threshold corresponds to approximately 0.7°C [Jennings et al., 2018]. The influence of humidity on phase is related to the thermodynamics of the hydrometeor, as hydrometeor energy balance theory dictates that low RH environments facilitates evaporative cooling through latent heat exchange, resulting in precipitation that stays frozen longer in environments above freezing [Harder and Pomeroy, 2013]. The importance of humidity in phase partitioning has been demonstrated by multiple studies which show that using a partitioning scheme which incorporates humidity related variables such as wetbulb temperature reduces biases in model outputs. For example, Wang et al. (2019) found that estimating snowfall fraction as a

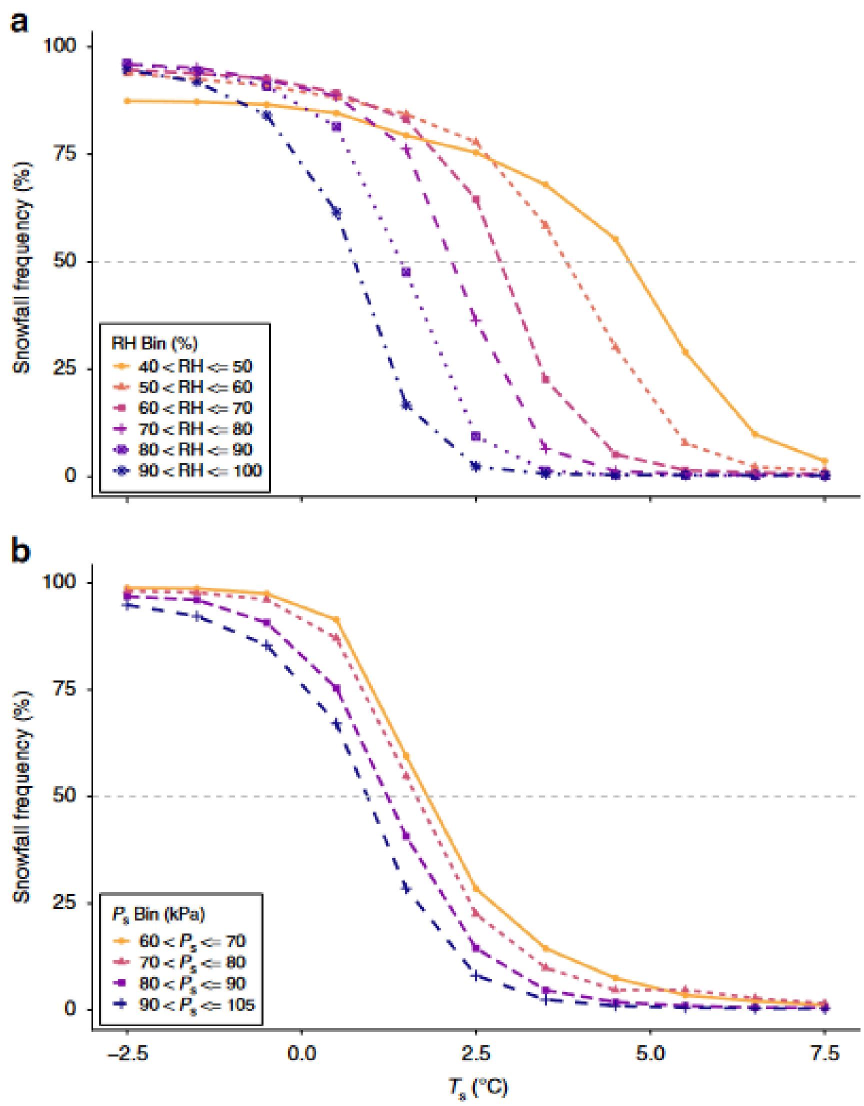

Figure 1.3: Snowfall frequency curves calculated using observations from 11,924 stations across the Northern Hemisphere (1978–2007). a) Snowfall frequency curves plotted by RH bin. b) Snowfall frequency curves plotted by P_s bin (Jennings et al., 2018).

sigmoid function of wetbulb temperature reduced bias in the Noah-MP land surface model over the Western United States by up to 11.8%.

The contribution of atmospheric humidity is particularly important in high elevation mountain regions where the average relative humidity is low [Wang et al., 2019, Jennings et al., 2018, Marks et al., 2013]. Relative to more humid regions, these areas see the greatest underestimation of snow depth when using precipitation phase methods which don't include humidity related parameters [Wang et al., 2019, Jennings and Molotch, 2019]. As the influence of evaporative cooling is stronger in these low humidity regions the rain-snow threshold tends to be higher, meaning that snow will fall at higher temperatures compared to warmer or more humid areas [Jennings et al., 2018].

As precipitation requires time to melt as it falls through warmer layers in the atmosphere, a quicker fall rate can decrease the chance of it reaching the surface in solid form. While fall rate is a lesser on influence on precipitation phase relative to humidity, it can still play a role while also being difficult to quantify [Jennings et al., 2018]. Pressure is one variable typically used to try and capture the influence of fall rate as hydrometeor speed decreases with air pressure, meaning that snow falls quicker at lower pressure. This creates the effect of lower surface pressures being more likely to see snowfall at a given surface temperature, meaning a higher temperature threshold for snow at higher elevations [Sims and Liu, 2015, Dai, 2008].

The rain/snow threshold demonstrates great variability across space, particularly in areas of complex terrain and coastal regions [Marks et al., 2013]. For instance, in high mountain areas the rain/snow threshold has been shown to behave differently relative to lower elevations due to lower air pressure and humidity [Dai, 2008, Wang et al., 2019]. While spatial variations can often be explained by the variability of moisture, temperature, and pressure along longitudinal, latitudinal, and elevation gradients, there has been evidence that there are more complex spatial relationships at play which may not easily represented by environmental variables alone. The influence of air pressure on precipitation phase, for example, has been found to be higher at higher elevations [Dai, 2008]. Additionally, satellite retrievals of precipitation phase have been shown to be less accurate at above

50°N despite experiencing no limitations on sampling with little apparent explanation as to why [Kodamana and Fletcher, 2021].

Overviews of precipitation phase models have found that certain climate regions show a higher degree of sensitivity to method selection than others. Across the Northern Hemisphere lower humidity areas, such as the Rocky Mountains in North America and the Tibetan Plateau in Asia, were found to be most sensitive to method selection when estimating snowfall frequency. In contrast, warm maritime sites have been found to be the most sensitive to method selection when modelling SWE [Jennings and Molotch, 2019]. This demonstrates the complexity in selecting optimal precipitation phase partitioning methods when dealing with climatologically complex regions and that any given option may result in different biases at different sites, indicating the need for acknowledging spatial complexity when modeling precipitation phase.

1.4 Research Gaps and Objectives

There is a major gap between the methods that climate and hydrological models use to discriminate precipitation phase and the real-world complexity of precipitation phase. A growing body of literature has shown that the incorporation of humidity-related variables in particular is a key step in improving precipitation phase partitioning models [Jennings and Molotch, 2019, Harder and Pomeroy, 2013]. The use of ambient temperature alone is inadequate as it does not consider the energy balance of the falling precipitation, but nevertheless humidity-related variables are rarely considered in precipitation partitioning methods [Harpold et al., 2017].

This work will build on previous research by using CloudSat-CPR data to train a precipitation phase partitioning model based on multiple environmental variables. The aim of this approach is to fill in gaps caused by lack of precipitation phase-related meteorological data and to address the calls for the creation and validation of phase partitioning models which utilize multiple environmental variables to improve upon conventional, mainly ambient air temperature-based methods. CloudSat-CPR data is a valuable resource in terms of spatial coverage, volume of data collected, and its accuracy at estimating precipitation phase. As such, it provides an extensive dataset from which machine learning models can be trained and tested to predict precipitation phase.

In order to determine the simplest and most effective method for predicting precipitation phase, a series of statistical prediction methods will be applied. While some previous studies have predicted precipitation phase as a melted fraction, here we will be classifying it as binary rain or snow. Our methods will advance from the simplest, logistic regression, to the more complex, machine learning. A key difference between these two methods is that logistic regression assumes a linear relationship between the independent variables and the log of the odds while machine learning does not. This may make machine learning a more effective tool for phase prediction, as there is evidence to suggest that the relationship between precipitation phase and predictor variables often varies across the landscape in a nonlinear fashion. For example, pressure has been shown to exhibit a stronger control on phase at higher elevations [Dai, 2008]. A comparison between these two methods will aid in determining which type of relationship is a more suitable representation of the relationship between precipitation phase and its predictors. The goal of this approach is to create a method that is accurate while still being simple enough to be operationally useful.

Spatial variability has hampered past efforts in phase prediction, making the application of a single model difficult across large areas. While the integration of humidity related variables has decreased biases in phase partitioning models in mountainous regions, there is still room for improvement by the creation of a spatially aware phase partitioning scheme. There are multiple options for achieving this, such as through the integration of spatial variables as predictor values or through creating a series of phase partitioning methods divided by some spatial schema. For instance, a separate precipitation phase prediction method for each snow class. Alternatively, as the creation of a gridded phase partitioning product is something which has the potential to be particularly beneficial, a gridded surface which each cell representing the rain-snow threshold based on the climatological parameters at that specific location.

In order to fill in the gaps in previous research as presented here, this work seeks to answer the following questions: (1) what is the relationship between the vertical profile of temperature and humidity and precipitation phase across Canada? (2) Does a parameterization based on these factors improve precipitation phase prediction relative to conventional methods?

Chapter 2

Data and Methods

2.1 Data Sources

Satellite remote sensing in the form of microwave sensors is commonly used for the retrieval of precipitation related information such as occurrence, rate, and phase. Sensors work by either passively detecting the radiation emitted by the earth's surface or actively emitting radiation and measuring the returns, with the degree of radiance measured by the sensor determining precipitation characteristics. Liquid hydrometeors reflect small amounts of microwave radiation, creating a bright band in the atmospheric profile. Solid hydrometeors, on the other hand, scatter microwave radiation, attenuating the radiation received by the sensor [Harpold et al., 2017]. Despite this, it is difficult to separate rain and snow based on reflectivity alone as various other factors may influence the brightness of returns. For instance, surface clutter and the presence of thick or convective clouds can all reduce the accuracy of returns [Liu, 2008] As such, additional information is required to successfully partition satellite-based retrievals of precipitation.

Of the various satellites dedicated to monitoring precipitation, CloudSat in particular is well suited at high latitudes due to its near-daily level of coverage above 80°N. For the Canadian Arctic this provides a high degree of spatial and temporal coverage in a region that is otherwise challenging to monitor by ground [King and Fletcher, 2020]. CloudSat was launched in 2006 as a joint project by the North American Space Agency, the Canadian Space Agency, and the California Institute of Technology's Jet Propulsion Laboratory. CloudSat's Cloud Profiling Radar (CPR) is a 94 GHz active radar that measures backscatter from clouds and precipitation and relates reflectivity to snowfall occurrence, rate, and phase [Liu, 2008]. This frequency allows a higher degree of sensitivity in measuring precipitation relative to other weather detecting radars [Tang et al., 2017]. The orbital track it covers is its overpass, which is made up of retrievals from its CPR consisting of radar returns recorded every 0.16 seconds from surface level to a height of 30km. Returns are split into 125 vertical bins of 240m in height [Kodamana, Rithwik, 2020].

Kodamana and Fletcher (2021) validated CloudSat-CPR retrievals of precipitation occurrence and phase by comparing estimates to ground-based human observations across Canada for the period from 2006-2016. They found that it has an approximate 65% chance of accurately detecting precipitation occurrence and an approximate 90% chance of accurately detecting the phase of precipitation occurrence. This means that while it is only moderately accurate at detecting the occurrence of precipitation when it can do so correctly it is highly accurate at detecting precipitation phase. The implications of this are that CloudSat data, when verified correctly, can provide a large amount of valuable information on precipitation phase and the surrounding environmental conditions. Which makes it a useful tool for the development of precipitation partitioning algorithms through the use of machine learning or logistic regression.

This research will focus on Canada, using meteorological data collected by CloudSat-CPR along with Environment Canada weather station data. CloudSat-CPR's 2C-PRECIP-COLUMN provides data on precipitation-related factors such as precipitation rate and phase while its ECMWF-AUX data product provides data on other atmospheric variables such as humidity, wind, and temperature. Atmospheric wetbulb temperatures were derived from CloudSat ECMWF-AUX vertical profiles of drybulb temperature, humidity, and pressure using Equation 1.1. ECMWF-AUX utilizes two datasets: an independent dataset and a reference dataset. The independent dataset is AN-ECMWF, a dataset provided by the European Center for Medium-Range Weather Forecasts which contains 3-hour forecast atmospheric state variable data on a half-degree lat/lon grid. The reference dataset is the geolocation data from the 1B-CPR product. The meteorological information from the independent dataset is interpolated to the geolocation data from the reference dataset using the procedure described in the following paragraph [Cronk and Partain, 2017].

The interpolate-to-reference algorithm operates by taking one CloudSat ray at a time and using the reference and independent dataset to find the four bounding ECMWF grid points around it. The height of the bin from the reference data is then used to find the two adjacent vertical bounding levels from AN-ECMWF at each of the grid points. Linear interpolation is then used on these four values to calculate a value for the given radar bin height at each of the bounding grid points. Using bilinear interpolation on these four values, a single value is then calculated at the location of the CPR ray at each bin height. As the AN-ECMWF data includes multiple forecasts for the same time, this procedure is repeated for both of the forecast times that bound the ray of interest and a linear temporal interpolation is performed on the two resulting values to produce the final value of the AN-ECMWF data field interpolated to the time and location of the radar bins [Cronk and Partain, 2017].

In order to classify the phase of precipitation in CPR returns, the Cloudsat algorithm utilizes a decision tree approach based on temperature and reflectivity profiles. Temperature is derived from ECMWF-AUX, which provides atmospheric state variables interpolated to CPR radar bins. Precipitation phase at the surface is classified as snow when temperatures are below 0°C, rain when temperatures are above 2°C, and mixed when temperatures are between 0°C and 2°C. Further classification of precipitation profiles is then performed using an occurrence algorithm on information from Cloudsat's 2B-GEOPROF product, which identifies radar reflectivity from hydrometeors. The occurrence algorithm utilizes the reflectivity of the fifth radar bin above the surface (1200m, Z_5), the cloud layer maximum reflectivity (Z_{max}), the cloud layer cloud base height (H_{base}), and the surface layer cross section (σ_0) [Smalley et al., 2014]. It then flags precipitation with a value from the following: 1 ("rain possible"), 2 ("rain probable"), 3 ("rain certain"), 4 ("snow possible"), 5 ("snow certain"), 6 ("mixed possible"), 7 ("mixed certain") [Wood and L'Ecuyer, 2018]. Profiles classified as rain are flagged as "rain certain" if Z_5

> 5 dBZ or if there is evidence of heavy attenuation in the variables Z_{max}, H_{base}, and σ_0. Mixed precipitation is classified as certain when Z_5 > -2.5 dBZ, with an additional check for attenuation given the likely presence of liquid precipitation in these scenarios. Snow is classified as certain when Z_5 > -5 dBZ, with no additional check for attenuation from the other three variables as attenuation by frozen precipitation at the 94 GHZ frequency tends to be small [Smalley et al., 2014].

This study will utilize the set of validated Cloudsat precipitation returns from Kodamana and Fletcher (2021). This dataset was created by first aggregating Cloudsat retrievals that fell within a 100km radius of an ECCC weather station. The proportion of solid and liquid precipitation flags from each overpass were then used to classify the phase of that overpass, where a higher proportion of solid flags classifies it as solid and a higher proportion of liquid flags classifies it as liquid. Mixed precipitation was classified as either solid or liquid based on the value of the melted mass fraction retrieved from Cloudsat's 2C-PRECIP-COLUMN, where a melted mass fraction value of ¡= 0.15 is considered snow. The phase of each overpass was then validated at the corresponding ECCC stations using a combination of human observations and observations from a Precipitation Occurrence Sensor System (POSS), a ground-based radar, upward-looking, X-band radar that provides an estimate of precipitation phase and intensity. In terms of temporal resolution, human observations are taken every hour and POSS observations are taken every minute. Cloudsat overpasses were considered temporally co-located if they fell within a half-hour of a human observation and 30 seconds of a POSS observation. Cloudsat estimates of precipitation phase were considered correct if they agreed with either the human or POSS observed phase. The final set of 26 ECCC stations utilized in this study were chosen due to having coincident datasets for Cloudsat observations, human observations, and POSS observations of precipitation phase.

2.2 Methods

While precipitation phase has commonly been determined with thresholds, a growing body of research has demonstrated that using logistic regression in order to predict phase can produce far more accurate results [Jennings and Molotch, 2019]. This approach incorporates various predictor values, such as ambient temperature, wet bulb temperature, and surface pressure, in order to estimate the probability of precipitation falling as either rain or snow. Relative to conventional threshold methods it has shown to be more accurate at predicting precipitation phase, likely due to the integration of multiple meteorological variables [Jennings and Molotch, 2019, Froidurot et al., 2014].

Froidurot et al. (2014) tested a variety of meteorological predictors in order to build a logistic regression precipitation phase partitioning model based on data from the Swiss Alps. They found that a model using a combination of wet bulb and ambient temperature performed the best at predicting precipitation phase, with a model using a combination of relative humidity and ambient temperature providing comparable performance. They also found that models which incorporated two variables performed better than those which just used one. These findings are corroborated by Jennings and Molotch (2019), who found that the use of a logistic regression model incorporating ambient temperature and relative humidity outperformed other partitioning methods.

A second recent advancement is the use of machine learning for the estimation of precipitation phase. Currently, these methods have mainly been applied to satellite data retrievals as they have been proven to be especially powerful when working with image-based data. For instance, the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) product uses machine learning techniques to estimate precipitation from infrared satellite data [Tao et al., 2016]. Despite advances in remote sensing data applications, little research has been conducted on using machine learning to predict precipitation phase from ground-based observations or in climate and hydrological models.

Our approach will utilize static variables such as latitude, longitude, and elevation as well as environmental variables such as wind speed, relative humidity, and temperature, allowing us to finetune our model according to different climatological regimes based on spatial variables. A hierarchy of classification methods will then be applied beginning with the simplest, logistic regression, and progressing to more complex machine learning methods such as Random Forest. The goal of this approach is to determine the phase partitioning method that is the simplest and most efficient while still producing a significantly positive impact on model estimates of snow accumulation.

Model performance will be assessed using the probability of detection (POD) within the -1°C to 4°C temperature range. This range was chosen by calculating model accuracy in temperature steps of 1°C and choosing the range in which minimums were most likely to occur. The advantage of this approach is that it allows us to finetune our model based on where uncertainty in precipitation phase is highest, as we found that outside this range the likelihood of misidentifying phase was so low as to be negligible.

In order to get a true assessment of how our parameterization performs, we will compare it to a set of conventional methods. This includes three different static thresholds as well as the more sophisticated method created by Sims and Liu. The Sims and Liu method utilizes look up tables to predict precipitation phase based on a combination of 2m air temperature, surface temperature, surface type, the 0-1000m lapse rate, surface pressure, and relative humidity [Sims and Liu, 2015].

In order to assess the spatial performance of the model, two different tests will be applied. The first, titled the leave-one-out test, retrains the model excluding data from the target ECCC station, then assesses model performance at that station. This test allows us to assess how well the model performs on regions on which it has not been trained. The second, titled the all-station test, tests how well the model trained on data from all stations performs at individual stations. This test allows us to determine if there are any spatial patterns to model performance.

Multiple studies have suggested that increasing the spatial discrimination of phase partitioning parameterizations might also improve their performance. In order to determine if our parameterization would benefit from this, we added Sturm snow classes as a predictor variable. Sturm's snow classification system provides a global dataset of snow cover type at a °0.5 x °0.5 resolution. It classifies snow cover into one of 7 types based on regional wintertime climate conditions including mean temperature, precipitation, and wind speed [Sturm and Liston, 2021]. Meaning that not only does it present a classification of snow type, but wintertime climates as well.

The code and data for reproducing our model is available at:

github.com/cdmballantyne/PrecipPhaseParameterization.git.

Chapter 3

An Improved Statistical Parameterization for Precipitation Phase Partitioning Across Canada

3.1 Introduction

Partitioning precipitation into rain or snow is an important aspect of hydrologic and climatological modelling, affecting a wide variety of downstream processes [Harpold et al., 2017]. However, conventional methods are often inadequate, leading to biases in model output [Jennings et al., 2018, Wang et al., 2019]. In mid-latitudes, precipitation typically starts out as solid during its formation and its phase at the surface is determined by whether it passes through the melting layer before reaching the ground. This, in turn, is affected heavily by atmospheric profiles of temperature and humidity along with fall rate and hydrometeor size [Harpold et al., 2017]. Neglecting these complexities can cause inaccurate estimates of the proportion of precipitation falling as rain or snow, leading to biases in model outputs of snow water equivalent, snow depth, and snow cover area [Imura and Michibata, 2022]. This leads to a cascading effect which can result in biases in other parts of the model such as streamflow, soil moisture, and land-atmosphere energy and water exchanges [Dai, 2008, McAfee et al., 2014]. Therefore, the ability to accurately estimate precipitation phase is a critical component of land surface modelling.

When creating precipitation partitioning parameterizations it is often most useful to consider the 50/50 rain-snow threshold, or the point at which there is a 50% chance of precipitation falling as either rain or snow. Thresholds can vary from a strict boundary at 0°C, below which all precipitation is classified as snow and above which all precipitation is classified as rain, to a linear or sigmoidal probability function ranging from 0% chance of rain at the lower threshold to a 100% chance of rain at the higher [Harpold et al., 2017]. Thresholds are typically centered on the 0°C surface temperature point, but as the 50/50 rain-snow threshold can vary around this point with factors such as humidity and pressure, this often leads to inaccuracies in precipitation phase estimation, particularly when models are evaluated at locations spread over large distances [Jennings and Molotch, 2019, Feiccabrino et al., 2015].

While temperature is a major control on the melting of hydrometeors, recent studies have shown that wet bulb, rather than ambient or dewpoint temperature, is a more effective parameter to use in precipitation phase estimation [Behrangi et al., 2018]. Latent heat exchanges at the surface of the hydrometeor through evaporation and sublimation can cause the hydrometeor surface to cool off relative to the ambient air temperature [Sims and Liu, 2015, Wang et al., 2019]. While the temperature of the hydrometeor itself would be the ideal way to determine its physical state, the environmental wet bulb temperature provides a close approximation and one that is much easier to obtain [Harder and Pomeroy, 2013]. Similar variables which incorporate humidity, such as the dewpoint temperature or relative humidity, don't consider the sensible and latent heat fluxes between the hydrometeor and the environment, therefore they do not provide the same physical relationship to phase as wet bulb temperature does [Harder and Pomeroy, 2013].

Vertical lapse rates of temperature, particularly those in the lowest 500m of the atmosphere, can influence the phase of hydrometeors in a variety of ways [Sims and Liu, 2015]. The presence of an inversion, where the temperature profile increases with height, may cause a hydrometeor to fall as rain when the temperature at the surface would classify it as snow. On average, steeper vertical temperature gradients tend to promote snowfall over rainfall, because they are associated with a shallower melting layer [Dai, 2008]. Interestingly, while the state of the upper atmosphere does influence the precipitation phase at the surface due to the aforementioned reasons, some studies have found that incorporating atmospheric predictors such as lapse rates does not necessarily present an improvement compared to surface predictors [Casellas et al., 2021, Froidurot et al., 2014] while others have found the opposite [Vionnet et al., 2022]. This suggests that additional scrutiny is required when incorporating atmospheric predictors into phase partitioning parameterizations.

Atmospheric humidity plays an important role in determining precipitation phase at temperatures close to 0°C [Harpold et al., 2017, Jennings et al., 2018]. Precipitation at lower relative humidity (RH) is more likely to fall as snow at higher temperatures relative to precipitation at higher RH, which makes the 50% rain and snow threshold in drier conditions higher than in moister conditions. In observational studies across the Northern Hemisphere, a 10% increase in RH is associated with a 0.8°C decrease in the 50% rain-snow threshold. As an example, at 40%-50% humidity the 50% rain-snow threshold is at approximately 4.5°C whereas for 90%-100% humidity the 50% rain-snow threshold is at approximately 0.7°C [Jennings et al., 2018]. The influence of humidity on phase is related to the thermodynamics of the hydrometeor, as drier environments are more readily associated with evaporate cooling through latent heat exchange, resulting in precipitation that stays frozen longer with ambient temperatures above freezing [Harder and Pomeroy, 2013]. The power of this relationship is corroborated by multiple studies showing that the integration of humidity related variables, in particular wet bulb temperature and relative humidity, has significantly improved phase estimates [Wang et al., 2019, Jennings et al., 2018, Behrangi et al., 2018].

The selection of precipitation phase partitioning method can significantly influence SWE and snow depth estimates. For instance, comparisons between different phase partitioning methods using the Cold Regions Hydrological Model have shown that using an ambient temperature-based method can produce uncertainties in peak SWE of up to 160mm at the same site [Harder and Pomeroy, 2013]. Incorporating a comprehensive set of environmental and spatial variables is essential for the creation of improved precipitation phase models. This requires an understanding of the physical and spatial factors which influence the temperature at which the 50/50 rain-snow threshold occurs and the ability to incorporate this information into a model which can accurately predict precipitation phase. This will provide more realistic boundary conditions for climate and hydrologic models and reduce bias in estimations of precipitation related factors such as snow depth and SWE. In order to achieve this goal, this study seeks to answer the following questions:

1. What is the relationship between the vertical profile of temperature and humidity and precipitation phase across Canada?

2. Does a parameterization based on diverse environmental factors improve precipitation phase prediction relative to conventional methods?

The primary goal of this work is to use CloudSat estimates of precipitation phase in order to create a statistical parameterization of precipitation phase across Canada. Section 2 describes the data and methods utilized, including how the training data was derived and the metrics for assessing model skill. Section 3 provides the results of our parameterization, including comparisons of spatial performance as well as performance relative to conventional methods. Finally, Section 5 provides a discussion and summary of our findings, limitations, and future research.

3.2 Data and Methods

3.2.1 Data

Building on the work done by Kodamana and Fletcher, 2021, this study utilized ground-truthed CloudSat-CPR returns of precipitation characteristics from 26 Environment and Climate Change Canada (ECCC) weather stations across Canada during the period from 2006-2016, the distribution of which is seen in Figure 3.1. Note that due to sampling restrictions not all stations were used in the final analysis. Observations of precipitation

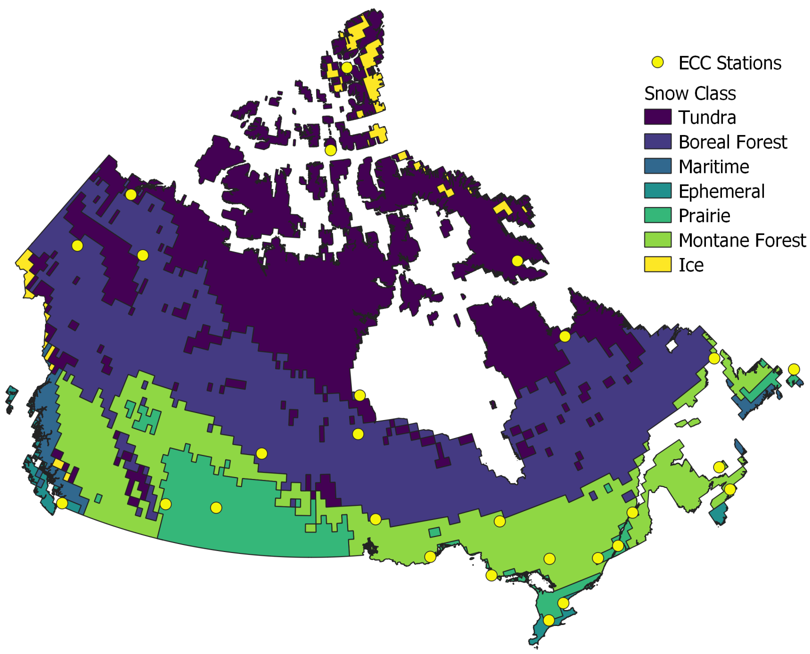

Figure 3.1: Map of Canada with Sturm snow classes and ECCC weather station locations

occurrence and type are recorded by trained human observers at these locations once every hour. CloudSat observations within a 100km radius of each station are then compared to coincident ground observations. The surface precipitation of a CloudSat profile is classified as solid when its flag indicates either snow possible or snow certain and as liquid when its flag indicates rain possible, rain probable, or rain certain. Individual profiles are then aggregated into overpasses using the weights of the solid and liquid profiles, where the weight is estimated as the inverse distance from the target ECCC weather station. If the sum of the weights of both liquid and solid precipitation profiles in each overpass exceeds 30%, the overpass is classified as precipitating. Precipitating overpasses are then classified as liquid if the weight of the liquid profiles exceeds the weight of the solid profiles and solid if the reverse is true.

CloudSat was launched in 2006 as a joint project by the National Aeronautics and Space

Administration, the Canadian Space Agency, and the California Institute of Technology's Jet Propulsion Laboratory. It carries a 94 GHz Cloud Profiling Radar (CPR) which sends radar pulses to the surface and utilizes the backscatter to form a vertical profile extending from surface level to 30 km of altitude in the atmosphere. The data utilized in this study comes primarily from its 2C-PRECIP and ECMWF-AUX products, both of which are post-processed products which contain information on environmental variables such as vertical profiles of temperature, humidity, and pressure, elevation, surface wind speed, and precipitation phase. The vertical profiles of atmospheric variables co-located with phase present one advantage of satellite data using station-based observations, as the latter often only records surface variables.

Due to the nature of Cloudsat's orbit, the number of observations at a given ECCC station increases significantly with latitude. For instance, the station with the lowest number of Cloudsat observations was Kindersley at 51.52°N, -109.18°W with a total of 8843 data points whereas the station with the highest number of Cloudsat observations was Eureka at 79.98°N, -85.93°W with 88343 data points. This can lead to issues with spatial bias, as the training data includes more observations from northern stations, potentially leading to decreased performance at more southern stations. In order to determine if spatial bias was an issue with the full dataset, the data was subset based on station, with a random selection of 8843 observations taken from each station. This resulted in a dataset with a total of 66,902 observations, down from the 341,512 observations in the full dataset. However, the difference in model accuracy for both the logistic regression and random forest models was not significantly different between the full dataset and the spatially robust dataset, indicating that spatial bias was not an issue with the full dataset. As such, the full dataset was utilized for the final analysis.

CloudSat utilizes a decision tree in order to classify precipitation phase. Using meteorological information from ECMWF_AUX which has been interpolated to CPR bins, it classifies precipitation as snow when the surface temperature is below 0°C, rain when the surface temperature is above 2°C, and mixed phase when surface temperature is between 0°C and 2°C [Smalley et al., 2014].

Atmospheric wetbulb temperatures were derived from CloudSat ECMWF-AUX vertical profiles of drybulb temperature, humidity, and pressure. CloudSat ECMWF-AUX contains AN-ECMWF modeled atmospheric state variable data interpolated to each CloudSat CPR bin at a vertical resolution of 240m. Input data is obtained from the European Center for Medium-Range Weather Forecasts' AN-ECMWF dataset. Using CloudSat's geolocation data, the four ECMWF grid points bounding the CloudSat ray are located and a linear interpolation is utilized to calculate the value of each AN-ECMWF variable of interest at the radar bin height for the four points. Bilinear interpolation is then performed to find the value of each variable at the location of the CPR ray [Cronk and Partain, 2017].

A two step process was used to derive wetbulb temperatures for the full atmospheric profile. First, relative humidity was calculated using Equation 3.1, where q is specific humidity, p is pressure, and T is air temperature.

$$RH = \frac{q * \frac{p}{0.378*q+0.622}}{6.112 * e^{((17.67*T)/(T+243.5))}} \tag{3.1}$$

Then, relative humidity was used along with drybulb temperature to calculate wetbulb temperature using the approximation shown in Equation 3.2 [Stull, 2011]:

$$\begin{aligned} T_w \approx T * atan(0.152 * (RH + 8.3136)^{\frac{1}{2}}) + atan(T + RH) - atan(RH - 1.6763) \\ +0.00391838 * (RH)^{(\frac{3}{2})} * atan(0.0231 * RH) - 4.686 \end{aligned} \tag{3.2}$$

3.2.2 Methods

A standard logistic regression based on the following equation was run in R on the full set of predictor variables as well as various iterations. This model predicted precipitation phase as a binary where 1 indicated rain and 0 indicated snow.

$$p(rain) = \frac{e^{\beta_0+\beta_1 X_1+\beta_2 X_2+...+\beta_p X_p}}{1 + \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p} \tag{3.3}$$

The Cloudsat-CPR data was then split into training and testing data based on a 80/20 split. The training data was utilized to generate the model, which was then validated by predicting the values of the reserved testing data and comparing the modeled values to the actual values. Model accuracy was calculated based on the ratio of correctly predicted values to incorrectly predicted values. As precipitation partitioning is most important around the 0°C mark a more detailed breakdown of model accuracy across the temperature gradient was required in order to get a better picture of model performance. In order achieve this, accuracy was calculated for each 1°C bin in the full range of temperatures captured in the testing data, approximately -40°C to 30°C.

The logistic regression model was evaluated using several combinations of variables in order to select the combination which produced the most accurate results. Initially, this included surface drybulb temperature, surface specific humidity, surface pressure, surface wind, and elevation Sensitivity testing was performed by running the model with one variable removed in order to determine how each variable affected model performance.

Following a similar process to the logistic regression model, a random forest model was generated using the randomForest package in R, as documented in [Liaw and Wiener, 2002]. Random forests work by growing a large number of decision trees based on the predictor variables in order to predict the response variable. Each tree is grown based on a random subset of the training data, with the number and depth of the trees being determined by the user. This presents an advantage over logistic regression as the random forest makes no assumptions about the linearity of the relationship between the predictors and the response variable and therefore is better able to represent nonlinear relationships.

Random forests were initially trained with 50 trees and 3 variables randomly sampled at each split. The best performing model was then selected to undergo additional tuning in order to determine the optimal fit. Tuning was done using the caret package in R and resulted in an optimal number of variables at each split of 2 and 75 trees. K-fold cross validation utilizing 10 folds was used in order to reduce overfitting and provide a more robust assessment of model accuracy.

Variable importance for the random forest model was assessed using the sperrorest package in R, [Brenning, 2012]. This package utilizes a permutation-based method to assess the predicative power of each variable by calculating the cross-validated mean square error (MSE_{CV}), shown in the formula below.

$$MSE_{CV} = \frac{1}{kn}\sum_{i=1}^{n}\sum_{j=1}^{k}(y_{ij} - \widehat{y_{ij}})^2 \tag{3.4}$$

Where y_{ij} represents the observed values and $\widehat{y}_{ij}$ represents the modeled values. The test data is divided into n test folds where k is the total number of values in a given test fold. Next, the MSE_{CV} is calculated for r permutations on each test fold and the mean across each permutation is taken in order to obtain a final MSE_{CV} for each predictor. Finally, the change in MSE (ΔMSE) produced by each predictor is calculated by subtracting the forecast MSE_{CV} from the mean MSE_{CV} across all permutations. A larger value of ΔMSE indicates a higher degree of predicative importance.

Spatial validation of both the random forest and logistic regression model was performed using two different tests. The first involved using all the observations related to one ECCC as the testing dataset and using the observations from the remaining observations as the training dataset. The goal of this test is to obtain an estimate of how well the model performs in climates which were not included in the training data. The second test involved evaluating how well the model trained on all the data performed at each individual station. The goal of this test is to obtain an estimate of how model performance varies over space.

Two different metrics were utilized to evaluate model performance. The first is the model accuracy over the [-1°C,4°C] interval, calculated as the ratio of hits to misses when the model is applied to the unseen testing data. The reason this interval was selected was twofold; first, calculating model accuracy over the full temperature interval does not provide the most informative measure of model performance, as there is a large portion of the temperature interval wherein uncertainty in precipitation phase is low and therefore model accuracy is high. Second, this interval is where minimums of model accuracy are most likely to occur, indicating that uncertainty is highest there and therefore the region we are most interested in when evaluating phase partitioning models.

3.3 Results

3.3.1 Surface-Only Model

We start by using two different statistical models, logistic regression and random forest, to quantify variable importance. The outcome will define the subset of surface predictors to be used in all subsequent models. Multiple versions of the logistic regression were run, one with the full suite of variables five with a single variable from the full suite left out each time. As shown in Figure 3.2, the variables which had the biggest impact on model performance were temperature and humidity. Pressure, wind, and elevation all had a very small influence on model performance.

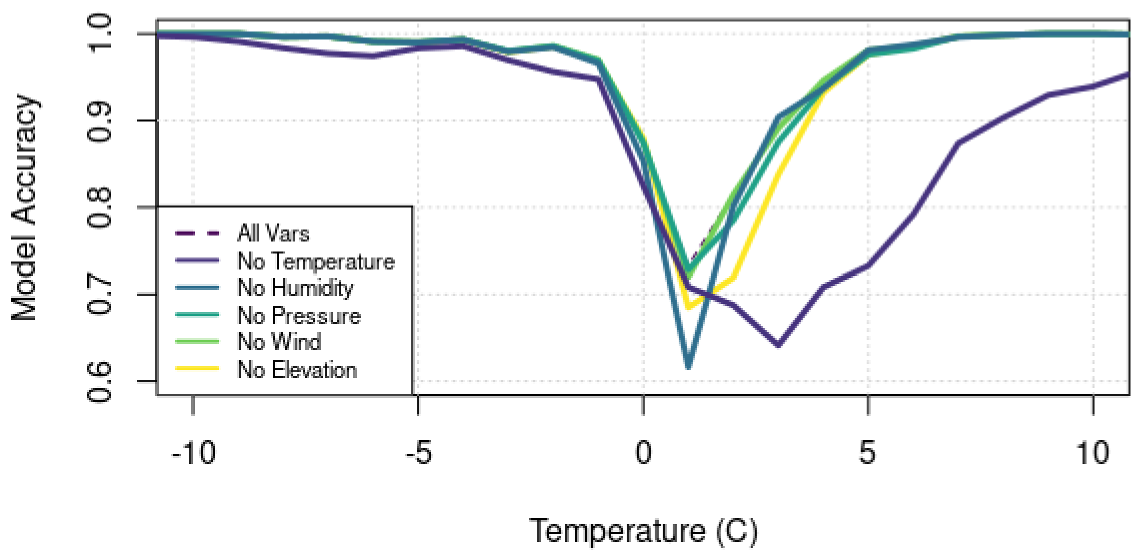

Figure 3.2: Logistic regression accuracy as a function of temperature.

Following a similar process to the logistic regression model, a random forest model was generated using the same suite of surface variables. This was done in order to determine if the random forest was better able to capture nonlinear relationships between the predictor variables. As seen in Figure 3.3, the removal of singular variables had a similar effect to the logistic regression, wherein the removal of temperature and humidity had the largest effect on model performance while the removal of wind, pressure, and elevation had a smaller effect on model performance. While all models performed well, with all maintaining an accuracy above 0.9 across the temperature range, the model with the full suite of variables performed the best. Compared to the logistic regression models, all of the random forest models showed superior performance with a significantly smaller reduction in accuracy around the 0°C mark. The improved performance of the random forest suggests that there are nonlinear relationships present between the predictors that the logistic regression fails to capture.

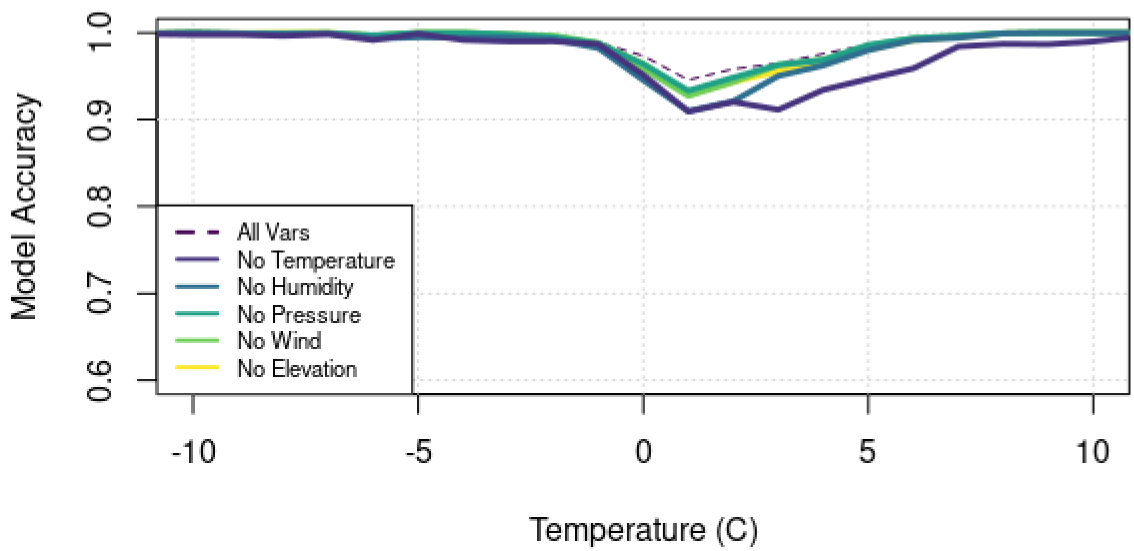

Figure 3.3: Random forest accuracy as a function of temperature.

Motivated by the knowledge that additional predictive information for rain-snow partitioning may be drawn from atmospheric variables above the surface, we next incorporated vertical profiles of T_{wb} into the model in addition to the surface variables discussed in the previous section. Additionally, we replaced surface pressure, surface temperature, and surface specific humidity with surface wetbulb temperature. This replacement was done with the goal of simplifying the final model, as wetbulb temperature is calculated using pressure, temperature, and humidity. As shown in Figure 3.4, the random forest model with atmospheric profiles of wetbulb temperature did not reveal a decrease in accuracy when surface wetbulb temperature was included over surface drybulb temperature, pressure, and humidity.

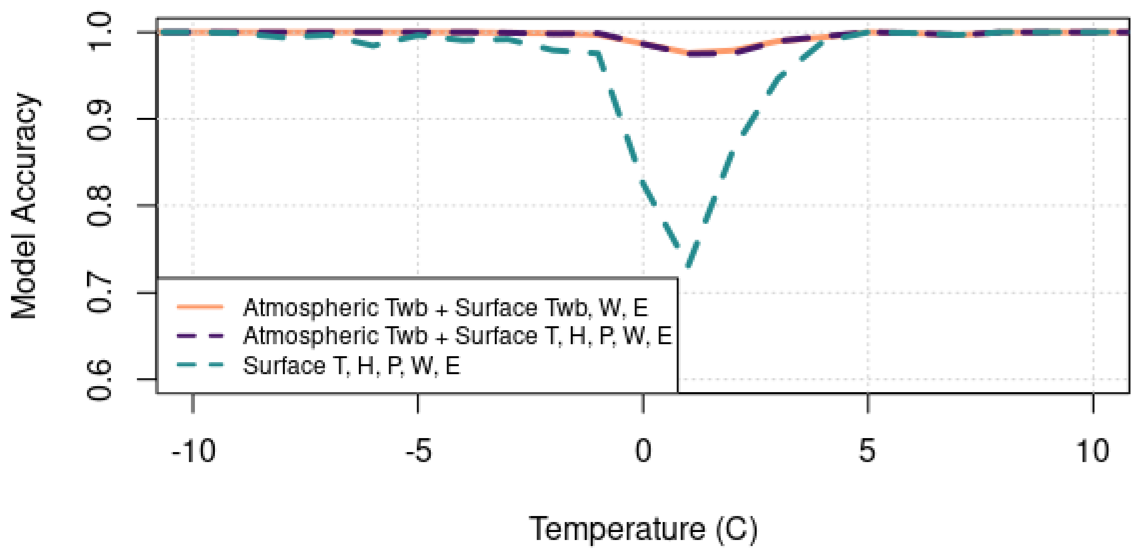

Figure 3.4: Logistic regression accuracy as a function of temperature for two models using atmospheric profiles of wetbulb temperature, one with surface wetbulb temperature, wind, and elevation, and the other with surface T, humidity, pressure, wind, and elevation, as well as one model with only surface variables.

In order to determine which levels of wetbulb temperature from ECMWF-AUX to use in the random forest and logistic regression models, a random forest model was run using all available levels, from 240m to 24,240m. As seen in Figure 3.5, the lowest seven levels from 240m-1680m had the highest mean decrease in accuracy, with a steep drop off above that. As such, these levels were selected for inclusion in the final model.

Narrowing in on the seven lowest levels of T_{wb}, the model was retrained with various combinations of these in order to determine the optimal configuration. This included running the model with only the four levels below 1000m, only the 3 levels above 1000m, and averaging the levels from 240-1440m together in sets of two to determine the impact of

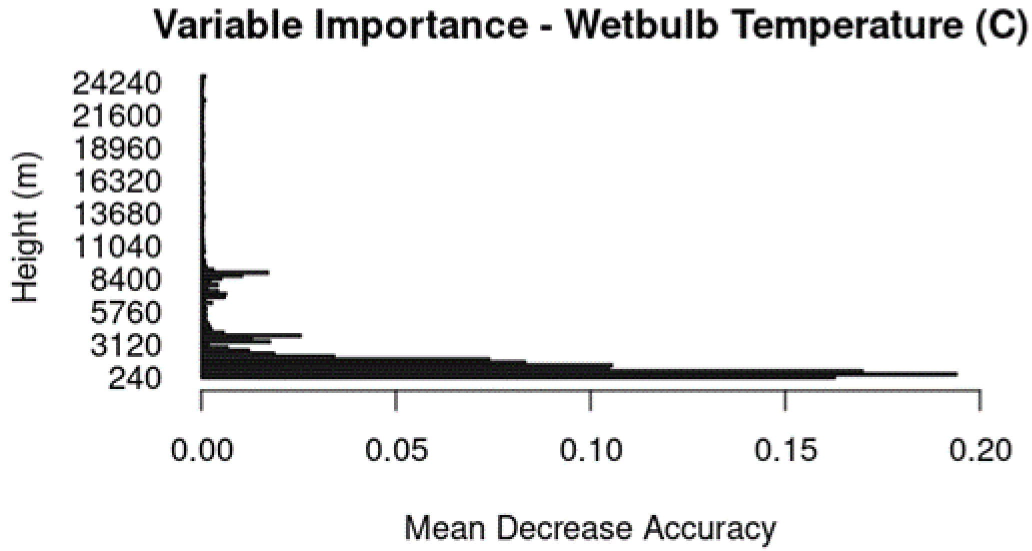

Figure 3.5: Variable importance of all atmospheric levels when used in random forest model.

sampling at a 240m vertical resolution. When performance was calculated over the entire study area, all variable combinations performed at approximately 97% accuracy. However, when broken down by station some differences began to emerge (Figure 3.6). Generally, the model with the 1680m, 1440m, and 1200m levels performed the worst out of the three. While the model with the averaged levels performed well, it still showed lower performance relative to the other two at 15 out of the 20 total stations. The model with the full suite of the seven lowest levels and the model with the four sub-1000m levels of T_{wb} performed similarly, therefore the sub-1000m level model was selected as the final parameterization.

From Figure 3.7, the wetbulb temperature levels from 240-960m were the most important for model performance. After the importance of the atmospheric T_{wb} levels was determined in the random forest, those were also added to the logistic regression model as well. This resulted in a further increase in accuracy from 96.5% to 97.8% in the random forest and 79.1% to 89.4% in the logistic regression. The final logistic regression and random forest model both included the seven predictors included in Figure 3.7.

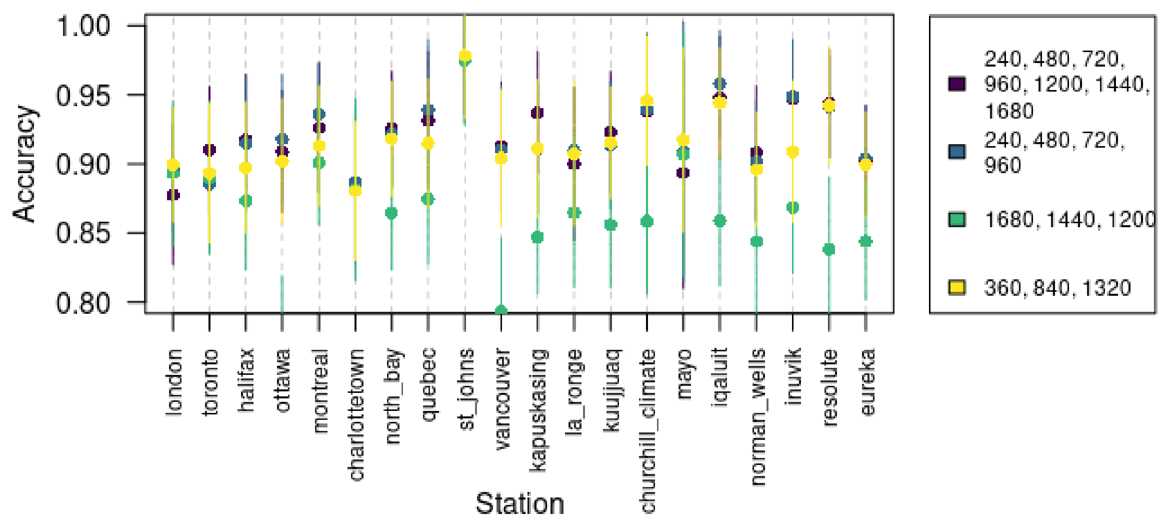

Figure 3.6: Model performance at individual stations for various combinations of atmospheric wetbulb temperature levels

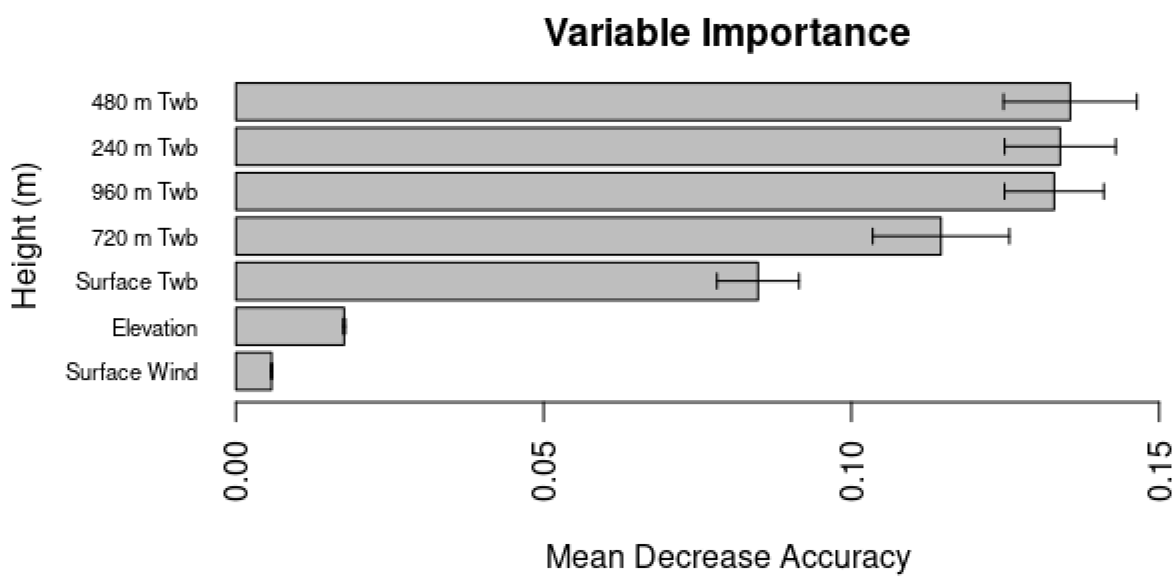

Figure 3.7: Variable importance of the full set of predictors for the random forest model.

In addition to overall improved performance, the inclusion of atmospheric T_{wb} also resulted in an improvement to the spatial robustness of the random forest model. When the leave-one-out test was applied to the random forest model without atmospheric T_{wb}, it resulted in an average accuracy across all stations of 83.4%. When it was applied to the random forest model with T_{wb}, it resulted in an average accuracy across all stations of 92.7%.

The high performance of the random forest model reinforces that it is able to capture some non-linear relationships in precipitation phase threshold which conventional models do not. The inclusion of wetbulb temperature profiles in the lowest 1000m of the atmosphere was also instrumental in improving model performance of both the logistic regression and random forest model, indicating that temperature profiles in the near-surface atmosphere are highly important for precipitation phase at the surface.

3.3.2 Model Validation

In order to gain an understanding of how well our models performed compared to conventional methods, we compared them to a set of existing methods. The first are the static temperature thresholds set at 0°C, 1°C, 2°C, and 3°C. As seen in Figure 3.8, these ranged in performance from the 3°C threshold with an accuracy over the -1,4°C interval of 62.9% to the 1°C temperature threshold, with an accuracy of 79.2%. The second method is the more sophisticated Sims-Liu phase partitioning method. The Sims-Liu method, with an accuracy of 79.4%, performs similarly to the 1°C temperature threshold. Both the logistic regression and random forest model, with respective accuracies of 89.4 and 97.8%, showed improvement over the simple threshold methods as well as the Sims-Liu method. While the logistic regression model performs better than the static threshold and Sims-Liu model but worse than the random forest model.

When a random forest was trained using the same predictors as the Sims-Liu model it produced similar overall performance, with an accuracy of 96.6%, but reduced performance for the leave-one-out test, with an average accuracy of 83.7%. This performance was similar

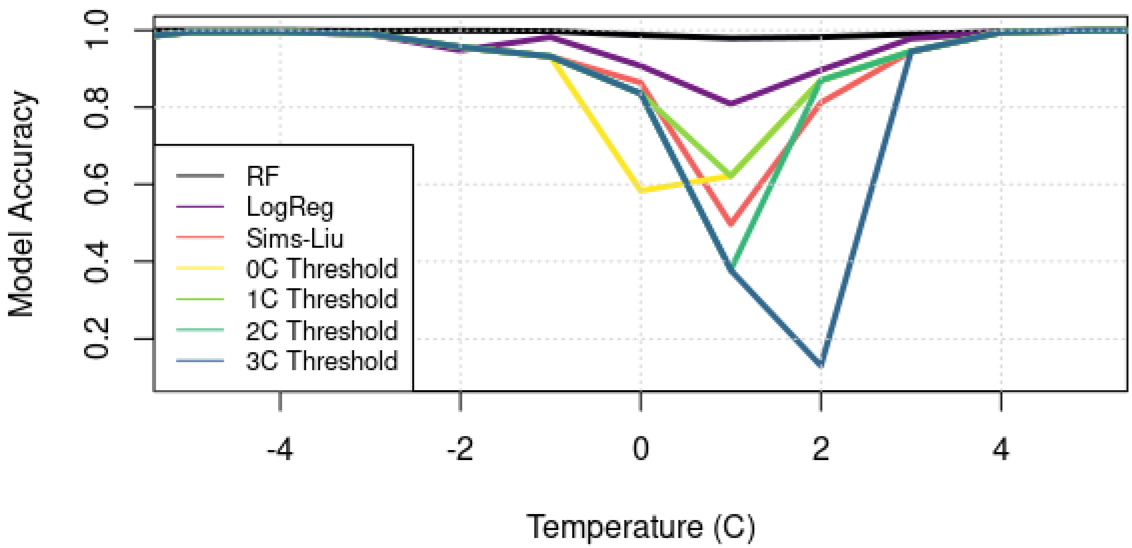

Figure 3.8: Comparison of model performance on the full set of ground-truthed Cloudsat data for the random forest, logistic regression, Sims-Liu, and simple temperature threshold methods.

to the performance of the random forest with only surface variables, suggesting that the 0-1000m lapse rate alone does not provide enough atmospheric information to improve phase partitioning.

However, it should be noted that spatial differences in training datasets may contribute to the relatively lower performance of the Sims-Liu model here. Their training dataset included global observations from weather stations and atmospheric soundings, mostly concentrated in the Northern Hemisphere. In comparison, our training dataset included solely ground-truthed satellite observations over Canada, the majority of which were concentrated in Northern Canada due to the nature of Cloudsat's orbital track.

3.3.3 Spatial Variability in Rain-Snow Partitioning

We next sought to understand whether a single parameterization worked well for the entire country or if regional variations in weather and climate make it necessary to have a regionalized parameterization. Determining the spatial variation in the skill of the rain-snow partitioning model across Canada was done in two ways. The first method was to address the skill at each individual station, including how well the model generalizes to stations it was not trained on. The second was grouping the stations by Sturm snow class in order to determine if having multiple models trained by individual snow class produced better performance to have one general model.

Figure 3.9 shows the difference in performance at individual stations between the leave-one-out and all-station tests for the random forest models. For the leave-one-out out test, mean accuracy was 92.7%, whereas for the all-station test mean accuracy was 98.5%. Generally, the difference between the two tests was smaller at the northern stations, on the right side of the plot, and slightly larger at the more southern stations, on the left side of the plot. However, this could be an artifact of the distribution of samples, since Cloudsat samples more with increasing latitude

As a potential alternative to the model fitted with all the data across Canada, a regionalized version of the random forest model was generated. This model involved fitting

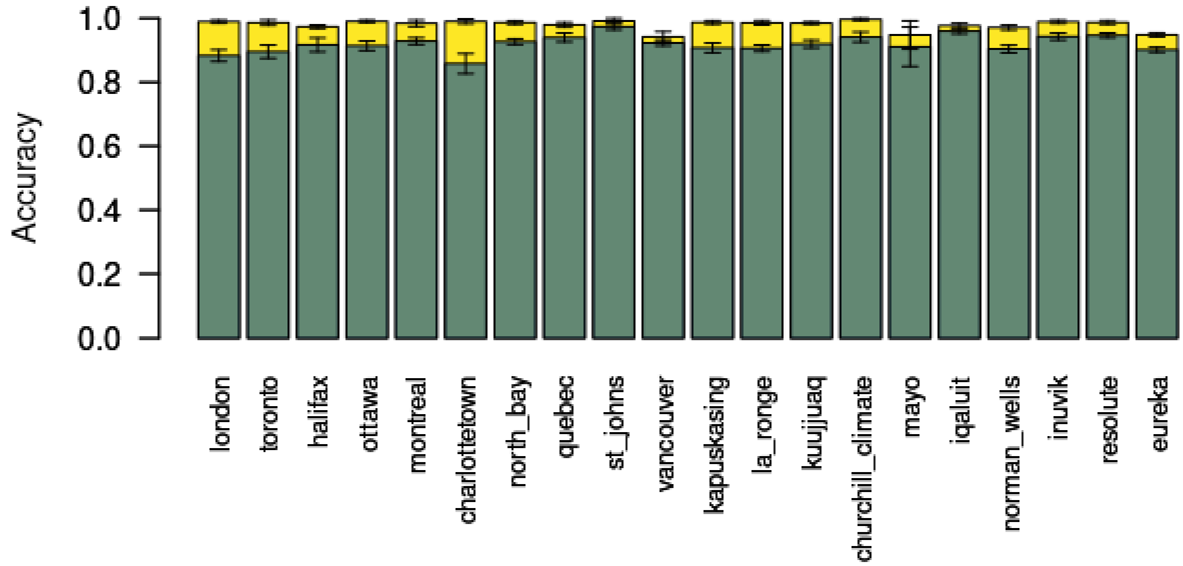

Figure 3.9: Accuracy comparison for the random forest model between the leave-one-out (green) and all-station (yellow) tests.

individual models for locations grouped by Sturm snow class, a total of five unique regions across the study area. Sturm snow classes were chosen as they are based on a combination of geographical and climatological variables. This was done in order to determine both if fitting by individual snow class improved overall model performance and if there was variation in predictor importance across snow class.

Figure 3.10 shows normalized variable importance for each snow class. Wetbulb temperatures ($T_w b$) at the 240m and 480m levels were generally the most important, followed by 720m, 960m and surface $T_w b$. Elevation and surface wind were generally the least important in terms of phase prediction, and their importance showed low variation across model runs. This lines up well with variable importance in the pan-Canada model, which showed a similar distribution.

Overall, the regionalized random forest had an accuracy averaged over all the snow classes over of 97.7%, while the pan-Canada model had an accuracy of 97.8%. Accuracy between individual snow classes ranged between 94.6% in Ephemeral to 98.7% in Montane

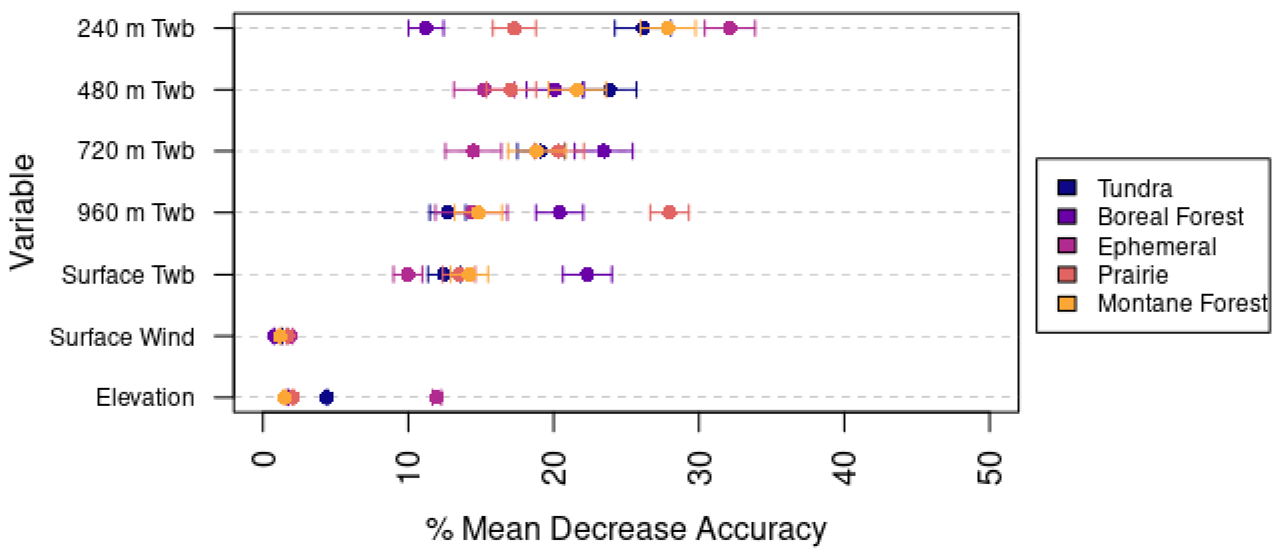

Figure 3.10: Normalized variable importance for the five snow classes included in the study region.

Forest, though significant variation in sample size between snow class meant the mean was weighted towards the higher end. For example, the class with the fewest observations, Ephemeral, had 16,454 observations while the class with the most observations, Tundra, had 186,676 observations.

Figure 3.11 shows model performance by individual snowclass compared to the non-regionalized model. With the exception of the Ephemeral snow class, model performance shows very little variation across snow classes, indicating that in terms overall model performance there is little impact of fitting models by region. It also suggests that the influence of the geophysical parameters picked for the final models does not vary significantly by region. This is in contrast to other studies which suggested some spatial component to variable importance which caused degradation in model performance in certain regions. However, these studies also mainly used surface variables and atmospheric information integrated in the form of lapse rates.

Figure 3.12 shows the performance of both the regionalized and pan-Canada models for

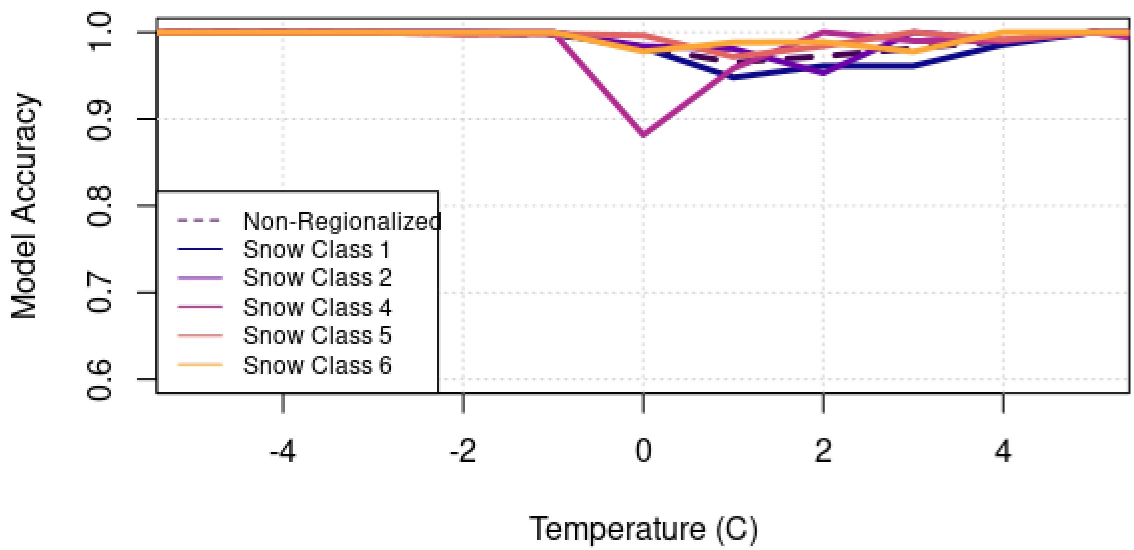

Figure 3.11: Comparison of model performance across the temperature gradients for the regionalized and non-regionalized random forest and logistic regression models.

the leave-one-out and all-station tests. Vancouver was eliminated from this section as it was the only station in its snowclass, therefore had no other stations against which it could be compared. Overall, the regionalized models performed fairly similarly to the pan-Canada models. For the leave-one-out test, the mean accuracy across stations of the regionalized model was 90.4% and the mean accuracy of the pan-Canada model was 92.7%. For the all-station test, the mean accuracy across stations of the regionalized model was 98.3% and the mean accuracy of the pan-Canada model was 98.4%. The similar performance between the regionalized and non-regionalized models for the two spatial tests agrees with the relative variable importance in each model.

3.4 Conclusion

This study presents an improved method for the parameterization of precipitation into rain and snow. This random forest parameterization has an accuracy of 97.8% over the -1°C

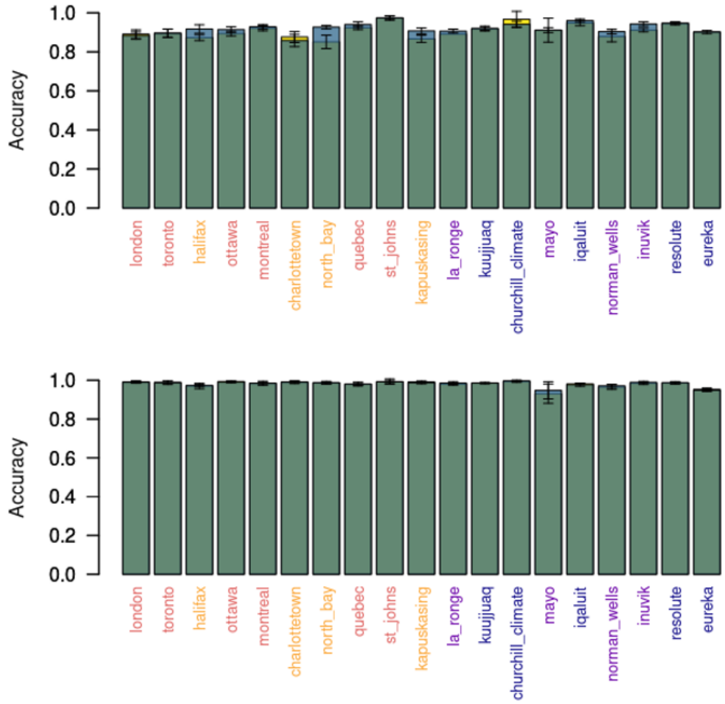

Figure 3.12: Comparisons of leave-one-out (top) and all-station test (bottom) non-regionalized (blue) and regionalized (yellow) model accuracy at each stations. Stations names have been colour-coded by snow class

to 4°C temperature range where phase uncertainty is highest. In contrast, conventional methods such as simple thresholds and the Sims-Liu parameterization range from accuracies of 62.9% to 79.4% over the same interval. This method is also spatially robust and still performs well at stations it has not been trained on, with an average drop in accuracy of 5% between when the parameterization has been trained on a station and when it has not.

This improvement in accuracy can be attributed to two main things: the first is the use of a random forest, which can model nonlinear relationships between parameters and the second is the use of atmospheric levels of wetbulb temperature. While previous studies have found mixed results on using lapse rates and vertical profiles of wetbulb temperature

[Casellas et al., 2021, Froidurot et al., 2014] this study found that not only did incorporating wetbulb temperature at four levels below 1000m improve overall accuracy, it also improved the performance of the model on stations on which it had not been trained.

The spatial robustness of our parameterization is important for the application of the parameterization over large, climatologically diverse areas such as Canada. Previous studies have found that the performance of phase partitioning methods varies over space and finding a method which performs universally well has proven difficult [Wang et al., 2019, Harpold et al., 2017]. As the spatial robustness improved when atmospheric levels of T_{wb} were added, this suggests that incorporating atmospheric information is a key component in creating phase parameterizations which generalize well across space.

One possible reason for the lower performance of the regionalized model is that the regionalization scheme selected had too much climatological variation within regions. For instance, Kindersley and Inuvik showed lower performance with the regionalized models than the pan-Canada models. This could indicate that these stations are climatological outliers compared to the rest of the stations in those snow classes, resulting in reduced model performance. This is supported by looking at the other stations included in their respective snow classes. Kindersley was included in the Prairie class, snow class 5, which also includes several eastern stations such as Quebec, Montreal, Toronto, and London. The climate in these regions tends to be more humid than Kindersley, which is located in Saskatchewan. Similarly, Inuvik is included in the Boreal Forest class, snow class 2, but is the most northern of any of the other boreal forest stations and is close to the boundary between the Boreal Forest and Tundra snow classes. Regionalization of the model has the effect of introducing more edge cases like Kindersley and Inuvik, degrading the overall performance of the model.

As precipitation phase and occurrence is ground-truthed by human observers, errors may result from misclassification by either Cloudsat or observers on the ground. An issue when it comes to light precipitation events due to the difference in sensitivity between instrumentation and the human eye. As instrumentation is far more sensitive than human observers, it is possible for human observers to classify precipitation events incorrectly [Kodamana and Fletcher, 2021]. While Cloudsat has shown to be quite accurate at classifying precipitation phase there is still a small degree of uncertainty which, particularly in larger datasets, may add up to thousands of misclassified observations [Kodamana and Fletcher, 2021].

Future work could utilize this parameterization in a climate or hydrological model to determine if it improves model outputs relative to conventional methods. Additionally, expanding this work to include mixed-phase precipitation would also be a significant improvement.

Chapter 4

Discussion and Conclusions

4.1 Summary

Building on previous work on validating Cloudsat returns of precipitation phase and investigating the environmental factors which control phase at the surface, we have created an improved parameterization scheme for partitioning precipitation into rain and snow. We started by comparing the performance of a logistic regression and random forest model using surface variables only to predict precipitation phase. Both models presented an improvement over conventional methods, with the random forest performing the best out of all. We found that, in terms of surface variables, temperature and humidity have the highest impact on model performance for both the logistic regression and random forest. As previous literature has pointed out the importance of atmospheric profiles of temperature and humidity, we then added this to the random forest model in the form of wetbulb temperature to determine whether they improved performance. We found that not only did this improve overall performance, it also improved the performance of the model at when applied to regions on which it had not been trained. Finally, we found that dividing our study area by Sturm snow class and training individual models for each class did not provide any improvement in model performance, indicating that a single model is adequate to capture the variability in the behaviour of the rain snow threshold across the study site.

The accurate partitioning of precipitation into phase is an important but often overlooked aspect of climate and hydrological modeling. Small errors in phase prediction can lead to major downstream impacts on biases in snow depth and snow cover. For example, overpredicting rain can lead to hydrological models predicting a peak discharge that's earlier in the season and lower in magnitude than in reality [Harpold et al., 2017]. This, in turn, can effect the ability of water resource managers to make decisions. In hydrologic and land surface models such as the Cold Regions Hydrological Model, Noah-MP, and SNOWPACK the proper selection of phase partitioning methods can significantly reduce biases in predicted snow depth and SWE [Wang et al., 2019, Harder and Pomeroy, 2014, Jennings and Molotch, 2019]. In climate models such as the MIROC6 GCM, phase partitioning schemes have been found to overpredict the fraction of snowfall in the arctic. As snow has a large effect on Earth's energy budget, while the effect of rain is negligible, this can impact the climate sensitivity of the model [Vionnet et al., 2022]. More accurate parameterizations which take into account the complexity of the factors which contribute to precipitation phase at the surface can, in turn, reduce these biases.

Conditions at near 0°C surface temperature can result in hazardous weather such as freezing rain and rain-on-snow events which can cause significant damage to infrastructure. As climate change warms cold regions, it is expected that many northern areas will experience near 0°C temperatures more frequently and therefore become more vulnerable to these hazards. Additionally, the timing of these events is expected to change as less near 0°C events occur in the summer and more occur in the winter [Klos et al., 2014]. Improving precipitation phase estimation can result in better prediction of these events, helping mitigate the impacts of hazardous weather [Mekis et al., 2020].

4.2 Limitations and Challenges

The training data used in this work was ground-truthed Cloudsat observations of precipitation phase which, while highly accurate, can still contain errors in the form of either misclassification by Cloudsat or observers on the ground [Kodamana and Fletcher, 2021].

In large datasets, even a small percentage of errors can add up to potentially thousands of misclassified observations. This can affect the ability of the parameterization to accurately express the physical relationships which lead to phase, as non-physical observations are included in the training dataset. Additionally, even if the parameterization does correctly classify the incorrectly ground-truthed observations, it will reduce the calculated accuracy of the parameterization.

Uncertainty in Cloudsat estimates of precipitation phase can be introduced by both instrumentation and the assumptions made by the algorithms used. As Cloudsat's CPR is designed to retrieve cloud properties rather than precipitation, the attenuation of the radar by water vapour can be significant [Wood, 2011]. High concentrations of liquid hydrometeors and high-intensity snowfall rates can both cause attenuation, with reductions in reflectivity of up to 10 DBZ for the former and 5 DBZ for the latter [Kodamana, Rithwik, 2020]. As assumptions on the distribution of hydrometeor size and shape are built into Cloudsat's scattering models, variations in these characteristics can also cause over or underestimates of reflectivity. More compact hydrometeors can result in reflectivity overestimates, while less compact hydrometeors can result in reflectivity underestimates [Wood et al., 2014]. While Cloudsat classifies all precipitation below -2°C as snow and all precipitation above 2°C as rain, with caveats based on the reflectivity of the returns, it is possible for snow to occur at above 2°C and for rain to occur at below -2°C [Wang et al., 2019]. Meaning that, for example, Cloudsat would be likely to classify heavy, wet snow occurring at surface temperatures above 2°C as "rain certain" due to the attenuation caused by the presence of liquid water, as its algorithm uses significant attenuation as evidence of rain. Alternatively, light rain occurring at temperatures below -2°C might be classified as "snow certain", as it may not cause significant enough attenuation to be flagged otherwise [Smalley et al., 2014].

The presence of surface clutter causes significant problems for Cloudsat retrievals below 1.2 km. As the surface is typically more reflective than hydrometeors, this causes a significant enhancement of the signal within the first four radar bins from the surface [Marchand et al., 2008]. To avoid this, it mainly relies on reflectivity in the fifth bin from the surface, at approximately the 1.2 km level, to classify phase uncertainty [Smalley et al., 2014]. While this may make phase uncertainties due to conditions below this level difficult to detect, the high accuracy of Cloudsat estimates of precipitation phase when compared to surface level ground-truth datasets indicates that this does not have a substantial negative impact on Cloudsat's performance [Kodamana and Fletcher, 2021].

Another limitation of this study was that the atmospheric state data was derived from the AN-ECMWF model, not taken directly from coincident observations. Meaning that it only represents an approximation of the atmospheric state at the observation time, not the actual conditions. This adds a layer of uncertainty to any physical relationships expressed in this model which could propagate into errors down the line. The use of atmospheric soundings in future work could serve to reduce these uncertainties by taking direct measurements of atmospheric conditions instead of relying on modeled data.

4.3 Future Work

In order to assess how well this parameterization improves model outputs, a next step would be to use it in a hydrologic or climate model. Previous work has found that incorporating both humidity-based measures and atmospheric profiles into phase partitioning parameterizations improves model accuracy [Wang et al., 2019, Harder and Pomeroy, 2013]. While this parameterization shows good potential compared to conventional methods, it is still uncertain whether or not it will produce the same results when applied in a hydrologic or climate model or outside of the study area.

Our work paid particular attention to the -1°C to 4°C interval, as that is where uncertainty in precipitation phase is the highest. This is particularly relevant to cold countries like Canada, where near-0°C conditions are common [Mekis et al., 2020]. Previous studies have looked at overall parameterization performance, or how well the parameterization performs when placed into a climate model, but as far as we are aware none have focused on model performance only at near-0°C conditions. Integrating this approach into future work could prove advantageous as it provides a clearer estimate of how well the model performs in the area of highest uncertainty.

While we utilized Sturm's snow classifications to test the regionalized version of the model in this study, there are other climate classification systems that could be utilized in its place. For instance, the Koppen Climate Classification system provides a more detailed division of climate regions. While our study found no benefit of breaking down the parameterization by region, it is possible that using a different regionalization scheme could prove beneficial as previous studies have suggested that regionalization could be useful in phase partitioning [Harpold et al., 2017].

Due to the nature of the ground-truthing stations, this parameterization has been trained over a limited area and not tested anywhere else. As such, there is no way of knowing if it is able to perform well on regions outside of Canada. In particular, regi0ons with very different climates from anywhere in Canada,. Assessing its performance on a global scale would be useful for determining the generalizability of the parameterization and applicability outside of Canada.

Along with an increased interest in more complex precipitation phase partitioning methods, a need for an increased level of spatial discrimination has emerged. Multiple studies have pointed out that phase partitioning accuracy varies across latitudes, potentially due to the nonlinear influence of humidity [Sims and Liu, 2015, Kodamana and Fletcher, 2021]. Despite this, very few studies have utilized spatial variables or some other spatial delineation scheme when creating phase partitioning models. Calls to address this gap through the creation of gridded phase partitioning products have been raised with the reasoning that this is one way in which spatial variability can easily be integrated into precipitation phase methods. Additionally, these gridded products could benefit climate and hydrological applications by providing spatially resolved data which can easily be integrated into models [Harpold et al., 2017]. In the past, the creation of spatially resolved precipitation phase products has been hindered by the lack of gridded meteorological data outside of ambient air temperature and precipitation, particularly at the smaller scales needed to predict precipitation phase. The emergence of new meteorological products in the form of satellite data, reanalysis products, and ground monitoring networks presents an opportunity to fill in these gaps.

References

[Auer, 1974] Auer, A. H. (1974). The Rain versus Snow Threshold Temperatures. *Weatherwise*, 27(2):67–67.

[Balogh et al., 2022] Balogh, B., Saint-Martin, D., and Ribes, A. (2022). How to Calibrate a Dynamical System With Neural Network Based Physics? *Geophysical Research Letters*, 49(8).

[Behrangi et al., 2018] Behrangi, A., Yin, X., Rajagopal, S., Stampoulis, D., and Ye, H. (2018). On distinguishing snowfall from rainfall using near-surface atmospheric information: Comparative analysis, uncertainties and hydrologic importance. *Quarterly Journal of the Royal Meteorological Society*, 144(S1):89–102.

[Brenning, 2012] Brenning, A. (2012). Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: the r package 'sperrorest'. In *IEEE International Symposium on Geoscience and Remote Sensing IGARSS*.

[Cabaj et al., 2020] Cabaj, A., Kushner, P., Fletcher, C., Howell, S., and Petty, A. (2020). Constraining Reanalysis Snowfall Over the Arctic Ocean Using CloudSat Observations. *Geophysical Research Letters*, 47(4).

[Cannon et al., 2015] Cannon, A. J., Sobie, S. R., and Murdock, T. Q. (2015). Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes? *Journal of Climate*, 28(17):6938–6959.

[Casella et al., 2017] Casella, D., Panegrossi, G., Sanò, P., Marra, A. C., Dietrich, S., Johnson, B. T., and Kulie, M. S. (2017). Evaluation of the GPM-DPR snowfall detection capability: Comparison with CloudSat-CPR. *Atmospheric Research*, 197:64–75.

[Casellas et al., 2021] Casellas, E., Bech, J., Veciana, R., Pineda, N., Miró, J. R., Moré, J., Rigo, T., and Sairouni, A. (2021). Nowcasting the precipitation phase combining weather radar data, surface observations, and NWP model forecasts. *Quarterly Journal of the Royal Meteorological Society*, 147(739):3135–3153.

[Cronk and Partain, 2017] Cronk, H. and Partain, P. (2017). Cloudsat ecmwf-aux auxiliary data process description and interface control document. Technical report, NASA.

[Dai, 2008] Dai, A. (2008). Temperature and pressure dependence of the rain-snow phase transition over land and ocean: RAIN-SNOW PHASE TRANSITION. *Geophysical Research Letters*, 35(12):n/a–n/a.

[Eldardiry and Habib, 2020] Eldardiry, H. and Habib, E. (2020). Examining the Robustness of a Spatial Bootstrap Regional Approach for Radar-Based Hourly Precipitation Frequency Analysis. *Remote Sensing*, 12(22):3767.

[Feiccabrino et al., 2015] Feiccabrino, J., Graff, W., Lundberg, A., Sandström, N., and Gustafsson, D. (2015). Meteorological Knowledge Useful for the Improvement of Snow Rain Separation in Surface Based Models. *Hydrology*, 2(4):266–288.

[Feiccabrino et al., 2013] Feiccabrino, J., Gustafsson, D., and Lundberg, A. (2013). Surface-based precipitation phase determination methods in hydrological models. *Hydrology Research*, 44(1):44–57.

[Forbes et al., 2014] Forbes, R., Tsonevsky, I., Hewson, T., and Leutbecher, M. (2014). Towards predicting high-impact freezing rain events. pages 15–21. Publisher: ECMWF.

[Foster et al., 2005] Foster, J. L., Sun, C., Walker, J. P., Kelly, R., Chang, A., Dong, J., and Powell, H. (2005). Quantifying the uncertainty in passive microwave snow water equivalent observations. *Remote Sensing of Environment*, 94(2):187–203.

[Froidurot et al., 2014] Froidurot, S., Zin, I., Hingray, B., and Gautheron, A. (2014). Sensitivity of Precipitation Phase over the Swiss Alps to Different Meteorological Variables. *Journal of Hydrometeorology*, 15(2):685–696.

[Gentine et al., 2018] Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., and Yacalis, G. (2018). Could Machine Learning Break the Convection Parameterization Deadlock? *Geophysical Research Letters*, 45(11):5742–5751.

[Harder and Pomeroy, 2013] Harder, P. and Pomeroy, J. (2013). Estimating precipitation phase using a psychrometric energy balance method: PRECIPITATION PHASE USING A PSYCHROMETRIC ENERGY BALANCE. *Hydrological Processes*, 27(13):1901–1914.

[Harder and Pomeroy, 2014] Harder, P. and Pomeroy, J. W. (2014). Hydrological model uncertainty due to precipitation-phase partitioning methods: HYDROLOGIC MODEL UNCERTAINTY OF PRECIPITATION-PHASE METHODS. *Hydrological Processes*, 28(14):4311–4327.

[Harpold et al., 2017] Harpold, A. A., Kaplan, M. L., Klos, P. Z., Link, T., McNamara, J. P., Rajagopal, S., Schumer, R., and Steele, C. M. (2017). Rain or snow: hydrologic processes, observations, prediction, and research needs. *Hydrology and Earth System Sciences*, 21(1):1–22.

[Imura and Michibata, 2022] Imura, Y. and Michibata, T. (2022). Too Frequent and Too Light Arctic Snowfall With Incorrect Precipitation Phase Partitioning in the MIROC6 GCM. *Journal of Advances in Modeling Earth Systems*, 14(12).

[Jennings et al., 2023] Jennings, K. S., Arienzo, M. M., Collins, M., Hatchett, B. J., Nolin, A. W., and Aggett, G. (2023). Crowdsourced Data Highlight Precipitation Phase Partitioning Variability in Rain-Snow Transition Zone. *Earth and Space Science*, 10(3):e2022EA002714.

[Jennings and Molotch, 2019] Jennings, K. S. and Molotch, N. P. (2019). The sensitivity of modeled snow accumulation and melt to precipitation phase methods across a climatic gradient. *Hydrology and Earth System Sciences*, 23(9):3765–3786.

[Jennings et al., 2018] Jennings, K. S., Winchell, T. S., Livneh, B., and Molotch, N. P. (2018). Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. *Nature Communications*, 9(1):1148.

[Joutsensaari et al., 2018a] Joutsensaari, J., Ozon, M., Nieminen, T., Mikkonen, S., Lähivaara, T., Decesari, S., Facchini, M. C., Laaksonen, A., and Lehtinen, K. E. J. (2018a). Identification of new particle formation events with deep learning. *Atmospheric Chemistry and Physics*, 18(13):9597–9615.

[Joutsensaari et al., 2018b] Joutsensaari, J., Ozon, M., Nieminen, T., Mikkonen, S., Lähivaara, T., Decesari, S., Facchini, M. C., Laaksonen, A., and Lehtinen, K. E. J. (2018b). Identification of new particle formation events with deep learning. *Atmospheric Chemistry and Physics*, 18(13):9597–9615.

[Kashinath et al., 2021] Kashinath, K., Mustafa, M., Albert, A., Wu, J.-L., Jiang, C., Esmaeilzadeh, S., Azizzadenesheli, K., Wang, R., Chattopadhyay, A., Singh, A., Manepalli, A., Chirila, D., Yu, R., Walters, R., White, B., Xiao, H., Tchelepi, H. A., Marcus, P., Anandkumar, A., Hassanzadeh, P., and Prabhat (2021). Physics-informed machine learning: case studies for weather and climate modelling. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194):20200093.

[Kienzle, 2008] Kienzle, S. W. (2008). A new temperature based method to separate rain and snow. *Hydrological Processes*, 22(26):5067–5085.

[King and Fletcher, 2020] King, F. and Fletcher, C. G. (2020). Using CloudSat-CPR Retrievals to Estimate Snow Accumulation in the Canadian Arctic. *Earth and Space Science*, 7(2).

[Klos et al., 2014] Klos, P. Z., Link, T. E., and Abatzoglou, J. T. (2014). Extent of the rain-snow transition zone in the western U.S. under historic and projected climate: Climatic rain-snow transition zone. *Geophysical Research Letters*, 41(13):4560–4568.

[Kodamana and Fletcher, 2021] Kodamana, R. and Fletcher, C. G. (2021). Validation of CloudSat-CPR Derived Precipitation Occurrence and Phase Estimates across Canada. *Atmosphere*, 12(3):295.

[Kodamana, Rithwik, 2020] Kodamana, Rithwik (2020). Validation of precipitation phase estimates from CloudSat-CPR across Canada. Master's thesis, UWSpace.

[Lawrence, 2005] Lawrence, M. G. (2005). The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications. *Bulletin of the American Meteorological Society*, 86(2):225–234.

[Liaw and Wiener, 2002] Liaw, A. and Wiener, M. (2002). Classification and regression by randomforest. *R News*, 2(3):18–22.

[Liu, 2008] Liu, G. (2008). Deriving snow cloud characteristics from CloudSat observations. *Journal of Geophysical Research*, 113:D00A09.

[Liu et al., 2021] Liu, W., Ikonnikova, S., Scott Hamlin, H., Sivila, L., and Pyrcz, M. J. (2021). Demonstration and Mitigation of Spatial Sampling Bias for Machine-Learning Predictions. *SPE Reservoir Evaluation & Engineering*, 24(01):262–274.

[Lynn et al., 2020] Lynn, E., Cuthbertson, A., He, M., Vasquez, J. P., Anderson, M. L., Coombe, P., Abatzoglou, J. T., and Hatchett, B. J. (2020). Technical note: Precipitation-phase partitioning at landscape scales to regional scales. *Hydrology and Earth System Sciences*, 24(11):5317–5328.

[Marchand et al., 2008] Marchand, R., Mace, G. G., Ackerman, T., and Stephens, G. (2008). Hydrometeor Detection Using Cloudsat—An Earth-Orbiting 94-GHz Cloud Radar. *Journal of Atmospheric and Oceanic Technology*, 25(4):519–533.

[Marks et al., 2013] Marks, D., Winstral, A., Reba, M., Pomeroy, J., and Kumar, M. (2013). An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin. *Advances in Water Resources*, 55:98–110.

[McAfee et al., 2014] McAfee, S. A., Walsh, J., and Rupp, T. S. (2014). Statistically downscaled projections of snow/rain partitioning for Alaska: DOWNSCALED SNOW/RAIN PROJECTIONS FOR ALASKA. *Hydrological Processes*, 28(12):3930–3946.

[Mekis et al., 2020] Mekis, E., Stewart, R. E., Theriault, J. M., Kochtubajda, B., Bonsal, B. R., and Liu, Z. (2020). Near-0 °C surface temperature and precipitation type patterns across Canada. *Hydrology and Earth System Sciences*, 24(4):1741–1761.

[Minokhin et al., 2017] Minokhin, I., Fletcher, C. G., and Brenning, A. (2017). Forecasting northern polar stratospheric variability with competing statistical learning models. *Quarterly Journal of the Royal Meteorological Society*, 143(705):1816–1827.

[Mitrescu et al., 2010] Mitrescu, C., L'Ecuyer, T., Haynes, J., Miller, S., and Turk, J. (2010). CloudSat Precipitation Profiling Algorithm—Model Description. *Journal of Applied Meteorology and Climatology*, 49(5):991–1003.

[O'Gorman and Dwyer, 2018] O'Gorman, P. A. and Dwyer, J. G. (2018). Using Machine Learning to Parameterize Moist Convection: Potential for Modeling of Climate, Climate Change, and Extreme Events. *Journal of Advances in Modeling Earth Systems*, 10(10):2548–2563.

[Panegrossi et al., 2017] Panegrossi, G., Rysman, J.-F., Casella, D., Marra, A., Sanò, P., and Kulie, M. (2017). CloudSat-Based Assessment of GPM Microwave Imager Snowfall Observation Capabilities. *Remote Sensing*, 9(12):1263.

[Rasmussen et al., 2012] Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A. P., Black, J., Thériault, J. M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., and Gutmann, E. (2012). How Well Are We Measuring Snow: The NOAA/FAA/NCAR Winter Precipitation Test Bed. *Bulletin of the American Meteorological Society*, 93(6):811–829.

[Shin and Baik, 2022] Shin, J. and Baik, J. (2022). Parameterization of Stochastically Entraining Convection Using Machine Learning Technique. *Journal of Advances in Modeling Earth Systems*, 14(5).

[Sims and Liu, 2015] Sims, E. M. and Liu, G. (2015). A Parameterization of the Probability of Snow–Rain Transition. *Journal of Hydrometeorology*, 16(4):1466–1477.

[Smalley et al., 2014] Smalley, M., L'Ecuyer, T., Lebsock, M., and Haynes, J. (2014). A Comparison of Precipitation Occurrence from the NCEP Stage IV QPE Product and the CloudSat Cloud Profiling Radar. *Journal of Hydrometeorology*, 15(1):444–458.

[Stephens et al., 2008] Stephens, G. L., Vane, D. G., Tanelli, S., Im, E., Durden, S., Rokey, M., Reinke, D., Partain, P., Mace, G. G., Austin, R., L'Ecuyer, T., Haynes, J., Lebsock, M., Suzuki, K., Waliser, D., Wu, D., Kay, J., Gettelman, A., Wang, Z., and Marchand, R. (2008). CloudSat mission: Performance and early science after the first year of operation. *Journal of Geophysical Research*, 113:D00A18.

[Stull, 2011] Stull, R. (2011). Wet-Bulb Temperature from Relative Humidity and Air Temperature. *Journal of Applied Meteorology and Climatology*, 50(11):2267–2269.

[Sturm and Liston, 2021] Sturm, M. and Liston, G. E. (2021). Revisiting the global seasonal snow classification: An updated dataset for earth system applications. *Journal of Hydrometeorology*, 22(11):2917 – 2938.

[Tang et al., 2018] Tang, G., Long, D., Behrangi, A., Wang, C., and Hong, Y. (2018). Exploring Deep Neural Networks to Retrieve Rain and Snow in High Latitudes Using Multisensor and Reanalysis Data. *Water Resources Research*, 54(10):8253–8278.

[Tang et al., 2017] Tang, G., Wen, Y., Gao, J., Long, D., Ma, Y., Wan, W., and Hong, Y. (2017). Similarities and differences between three coexisting spaceborne radars in global rainfall and snowfall estimation: COMPARING SPACE PRECIPITATION RADARS. *Water Resources Research*, 53(5):3835–3853.

[Tao et al., 2016] Tao, Y., Gao, X., Hsu, K., Sorooshian, S., and Ihler, A. (2016). A Deep Neural Network Modeling Framework to Reduce Bias in Satellite Precipitation Products. *Journal of Hydrometeorology*, 17(3):931–945.

[Tao et al., 2018] Tao, Y., Hsu, K., Ihler, A., Gao, X., and Sorooshian, S. (2018). A Two-Stage Deep Neural Network Framework for Precipitation Estimation from Bispectral Satellite Information. *Journal of Hydrometeorology*, 19(2):393–408.

[Thériault and Stewart, 2007] Thériault, J. M. and Stewart, R. E. (2007). On the effects of vertical air velocity on winter precipitation types. *Natural Hazards and Earth System Sciences*, 7(2):231–242.

[Turk et al., 2021] Turk, F. J., Ringerud, S. E., Camplani, A., Casella, D., Chase, R. J., Ebtehaj, A., Gong, J., Kulie, M., Liu, G., Milani, L., Panegrossi, G., Padullés, R., Rysman, J.-F., Sanò, P., Vahedizade, S., and Wood, N. B. (2021). Applications of a CloudSat-TRMM and CloudSat-GPM Satellite Coincidence Dataset. *Remote Sensing*, 13(12):2264.

[Vionnet et al., 2022] Vionnet, V., Verville, M., Fortin, V., Brugman, M., Abrahamowicz, M., Lemay, F., Thériault, J. M., Lafaysse, M., and Milbrandt, J. A. (2022). Snow Level From Post-Processing of Atmospheric Model Improves Snowfall Estimate and Snowpack Prediction in Mountains. *Water Resources Research*, 58(12).

[Wang et al., 2019] Wang, Y., Broxton, P., Fang, Y., Behrangi, A., Barlage, M., Zeng, X., and Niu, G. (2019). A Wet-Bulb Temperature-Based Rain-Snow Partitioning Scheme Improves Snowpack Prediction Over the Drier Western United States. *Geophysical Research Letters*, 46(23):13825–13835.

[Wen et al., 2013] Wen, L., Nagabhatla, N., Lü, S., and Wang, S.-Y. (2013). Impact of rain snow threshold temperature on snow depth simulation in land surface and regional atmospheric models. *Advances in Atmospheric Sciences*, 30(5):1449–1460.

[Wen et al., 2016] Wen, Y., Behrangi, A., Lambrigtsen, B., and Kirstetter, P.-E. (2016). Evaluation and Uncertainty Estimation of the Latest Radar and Satellite Snowfall Products Using SNOTEL Measurements over Mountainous Regions in Western United States. *Remote Sensing*, 8(11):904.

[Widmann et al., 2019] Widmann, M., Bedia, J., Gutiérrez, J. M., Bosshard, T., Hertig, E., Maraun, D., Casado, M. J., Ramos, P., Cardoso, R. M., Soares, P. M. M., Ribalaygua, J., Pagé, C., Fischer, A. M., Herrera, S., and Huth, R. (2019). Validation of spatial variability in downscaling results from the VALUE perfect predictor experiment. *International Journal of Climatology*, page joc.6024.

[Wood, 2011] Wood, N. B. (2011). *ESTIMATION OF SNOW MICROPHYSICAL PROPERTIES WITH APPLICATION TO MILLIMETER-WAVELENGTH RADAR RETRIEVALS FOR SNOWFALL RATE*. PhD thesis.

[Wood and L'Ecuyer, 2018] Wood, N. B. and L'Ecuyer, T. (2018). Level 2c snow profile process description and interface control document. Technical report, NASA.

[Wood et al., 2014] Wood, N. B., L'Ecuyer, T. S., Heymsfield, A. J., Stephens, G. L., Hudak, D. R., and Rodriguez, P. (2014). Estimating snow microphysical properties using collocated multisensor observations. *Journal of Geophysical Research: Atmospheres*, 119(14):8941–8961.

[Yuval and O'Gorman, 2020] Yuval, J. and O'Gorman, P. A. (2020). Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions. *Nature Communications*, 11(1):3295.

[Zhang et al., 2019] Zhang, J., Okin, G. S., and Zhou, B. (2019). Assimilating optical satellite remote sensing images and field data to predict surface indicators in the Western U.S.: Assessing error in satellite predictions based on large geographical datasets with the use of machine learning. *Remote Sensing of Environment*, 233:111382.

[Zhang et al., 2021] Zhang, L., Li, X., Zheng, D., Zhang, K., Ma, Q., Zhao, Y., and Ge, Y. (2021). Merging multiple satellite-based precipitation products and gauge observations using a novel double machine learning approach. *Journal of Hydrology*, 594:125969.