ERDC/CRREL TR-YY-DRAFT

US Army Corps
of Engineers®
Engineer Research and
Development Center

ERDC
ENGINEER RESEARCH & DEVELOPMENT CENTER

Program Title [[[if not applicable, delete this entire line]]]

Machine Learning Applications for FMCW Radar Layer Picking in Snow

Genetic Algorithm, Multilayer Perceptron, Support-Vector Machine Classifier

Scott S. Storms, Hans-Peter Marshall, Elias J. Deeb.

September 2021

Cold Regions Research
and Engineering Laboratory

The U.S. Army Engineer Research and Development Center (ERDC) solves the nation's toughest engineering and environmental challenges. ERDC develops innovative solutions in civil and military engineering, geospatial sciences, water resources, and environmental sciences for the Army, the Department of Defense, civilian agencies, and our nation's public good. Find out more at www.erdc.usace.army.mil.

To search for other technical reports published by ERDC, visit the ERDC online library at https://erdclibrary.on.worldcat.org/discovery.

Program Title [[[If project was not funded by a military or civil program, delete all text in this cell]]]

ERDC/CRREL TR-YY-DRAFT
September 2021

Machine Learning Applications for FMCW Radar Layer Picking in Snow

Genetic Algorithm, Multilayer Perceptron, Support-Vector Machine Classifier

Scott S. Storms, Elias J. Deeb

Cold Regions Research and Engineering Laboratory [[[always acknowledge performing lab first]]]
U.S. Army Engineer Research and Development Center
72 Lyme Road
Hanover, NH 03755

Hans-Peter Marshall

Boise State University

1910 W. University Dr
Boise, ID 83725

Final report

Prepared for Name of Sponsoring Agency, Program Office, or reimbursable customer
City, ST ZIP

Under Project Number, "Project Title"

Monitored by Cold Regions Research and Engineering Laboratory [CRs ONLY, otherwise delete block]
U.S. Army Engineer Research and Development Center
Hanover, NH 03755

Abstract

The U.S. Army Cold Regions Research and Engineering Laboratory conducts snow-science research using multiple radar systems in an effort to understand micro and macro properties of snowpack. One of these systems, FMCW radar, is particularly useful for measuring accumulated snow-depth, which in turn is used for measures such as snow-water-equivalent (SWE). The automated identification of the air/snow and snow/ground interface, generally referred to as 'Layer Picking', has proved challenging and has historically relied largely on time consuming human input, or 'Manual Picking'.

This research presents two machine learning methods to automate the picking process. These algorithms are then applied to the SnowEx 2020 Grand Mesa field campaign data in western Colorado. First, we use a Genetic Algorithm, and subsequently use these picks as the supervised data set for a single hidden layer perceptron model (the classic neural network). We then compare the calculated snow-depths with field observation snow-depths over 24 sites.

Finally, with these picks, we perform two case studies. The first is a binary classifier, which can distinguish between high and low stratification in the snowpack. We use both logistic regression and a support-vector machine and obtain 90 % accuracy on the test set. The second case study provides evidence of the ability to model snow density from the FMCW data.

Contents

Figures and Tables

Figures

Tables

Preface

This study was conducted for the Name of Sponsoring Organization under Project Number, "Title." The technical monitor was T.M. Name.

The work was performed by the Branch Name of the Division Name, U.S. Army Engineer Research and Development Center, Cold Regions Research and Engineering Laboratory (ERDC-CRREL). At the time of publication, B.C. Name was Chief; D.C. Name was Chief; and T.D. Name was the Technical Director for Technical Domain Name. The Deputy Director of ERDC-CERL was Dr. D.D. Name and the Director was Lab Director Name.

[[[Any acknowledgments should appear here; otherwise delete this text. Do not acknowledge anyone who is listed as an author or who is listed elsewhere on this page.]]]

The Commander of ERDC was COL Teresa A. Schlosser and the Director was Dr. David W. Pittman.

Acronyms and Word Usage

FMCW	Frequency modulated Continuous Wave Radar
GPS	Easting, Northing pair geolocating a trace.
Skycal	Radar response when pointed at the sky.
Trace	Power Spectral Density (PSD).
Record	Collection of traces, 43 seconds, 999 traces
Trajectory	The GPS curve traced by FMCW
Top pick	Air/snow interface
Bottom pick	Snow/ground interface
GA	Genetic Algorithm
Lower and upper bounds	Crossover points of filtered trace with filtered skycal.
Colormap	A record (or collection of traces) with the signal level represented as color.
Site	A spiral pattern of manual depth probe measurements against which the radar is compared.

1 Introduction

1.1 Background

The U.S. Army Cold Regions Research and Engineering Laboratory conducts snow-science research using multiple radar systems in an effort to understand micro and macro properties of snowpack. One of these systems, FMCW radar, is particularly useful for measuring accumulated snow-depth, which in turn is used for measures such as snow-water-equivalent (SWE). The automated identification of the air/snow and snow/ground interface, generally referred to as 'Layer Picking', has proved challenging and has historically relied largely on time consuming human input, or 'Manual Picking'.

This research presents two machine learning methods to automate the picking process. These algorithms are then applied to the SnowEx 2020 Grand Mesa field campaign data in western Colorado. First, we use a Genetic Algorithm, and subsequently use these picks as the supervised data set for a single hidden layer perceptron model (the classic neural network). We then compare the calculated snow-depths with field observation snow-depths over 24 sites.

Finally, with these picks, we perform two case studies. The first is a binary classifier, which can distinguish between high and low stratification in the snowpack. We use both logistic regression and a support-vector machine and obtain 90 % accuracy on the test set. The second case study provides evidence of the ability to model snow density from the FMCW data.

1.2 Objective(s)

We want to produce a reliable, computationally fast method to automate the layer picking process for the data obtained from an FMCW radar system. The methods developed herein are specific to this radar, in particular our calibration technique, 'skycal', but long term focus would strive to make these routines applicable across a range of radar systems, for instance Ground Penetrating Radar and Synthetic Aperture Radar.

1.3 Approach

The GA maximizes a fitness function, and when terminated returns the fittest 'individual' or the desired picks. We need to identify in the radar trace properties to maximize. The signal level itself is one such component, since the boundary reflections produces a peak in the power spectral density (PSD). This however is not sufficient. A close inspection of the radar trace, together with a calibration technique we call 'skycal', and a continuity requirement, provide the additional components to complete the fitness function.

The neural network classifier is adapted from the classic textbook example of digit recognition; given an image of a handwritten digit, the neural network will (hopefully) identify it correctly. Successive radar traces grouped together are often presented as a colormap, or image, with the color representing the signal level. This allows for a good visual interpretation of the radar response over large trajectories. The class to which a trace belongs is given by the picked value. Just as a handwritten '0' belongs to one of 10 classes, the number of classes for the picks depends on the range of picked values in the supervised set. The neural network for the top picks has fewer classes as the radar sits at a (mostly) constant distance about the air/snow interface. The neural network for the bottom picks can potentially have a large number of classes, limited only by the maximum measurable depth.

2 FMCW Radar & PSD

2.1 RF Hardware, Colormap and Trace

The FMCW radar system is a time-of-flight device, recording the arrival time of the echo as a frequency difference between the current (instantaneous) frequency and the frequency of the echo. This frequency difference, the baseband, or IF, is obtained by using an rf mixer and subsequently digitized and processed.

In our case, a frequency source, a Teledyne YIG Oscillator, ramps from 6 GHz to 18 GHz in 20 ms, and back to 6 GHz, driven by a triangle wave 0-10V input signal. The Oscillator is driven at the highest chirp rate, and our data acquisition device samples at 100 kS/s, making the maximum depth we can measure approximately 12 m.

If we suppose there is a single target, then all frequencies in the 6 -18 GHz chirp will return to the receiver at the same time, and the frequency difference will be constant throughout the chirp. Hence the Fourier transform is a natural method for analyzing radar traces. The power spectrum, or PSD, obtained from the Fourier transform will be a strongly peaked function at the IF frequency, and is proportional to the distance to the target.

We can look at individual traces of the PSD or put them together and form a colormap, as shown below in Figure 1.

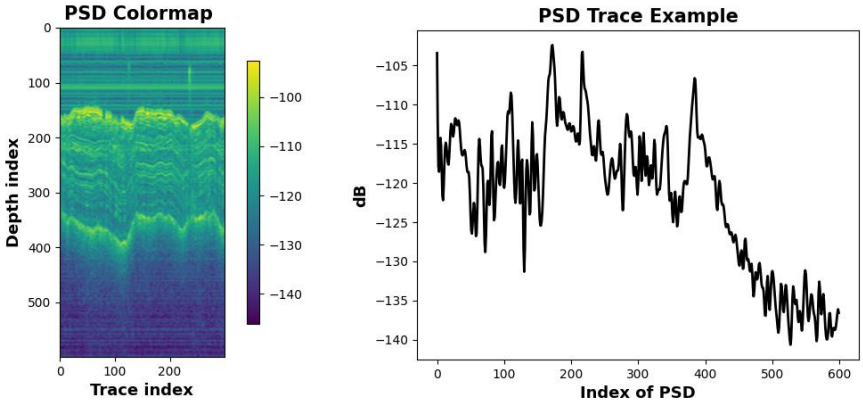

Figure 1. a.) Colormap of the PSD, with signal value identified as a color level. b.) PSD of a single trace. The index of the PSD is in 1-1 correspondence with the distance to the target and hence a Depth index.

2.2 PSD Processing

2.2.1 /skycal/PSD/Windowing

The 'skycal' is a trace obtained by running the radar while pointed at the sky. Since there are no targets, this produces a systematic response. Under these conditions, the trace-trace stability is very good and we do not perform averaging, but use a single trace. The skycal trace is subtracted from the radar trace in the time domain before the fft is applied. Figure 2 shows a comparison between a radar trace with and without the skycal. Figure 3a shows the time domain response. This energy dominates the low end of the PSD and are (mostly) removed upon time domain subtraction. Subtraction also lowers the noise floor of the PSD by about 15 dB. By including the skycal we generally insure the air/snow signal will be the maximum value of the trace. This is a desirable condition because if we can remove the overall monotonically decreasing shape of the PSD, then simple level detection can be a very useful analytical tool.

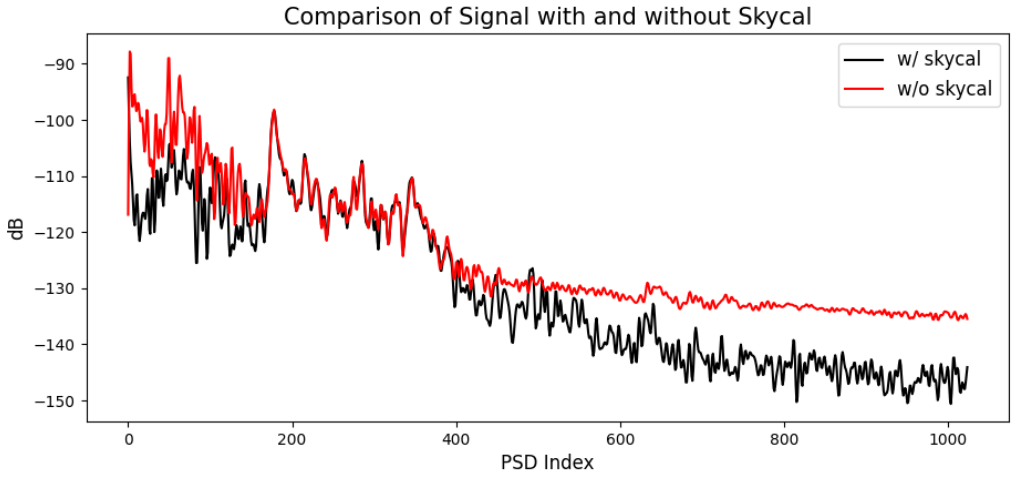

Figure 2. Effect on the trace with subtraction of the skycal in the time domain.

Figure 3 is an ensemble of all skycals used in processing, one for each day. We can use this information to help determine an ideal height above the snow to place the radar. With a little foresight (see Table 2), we know the radar is approximately 98 cm (160 index, 3.93 kHz), on average, above the snow. However, looking closely at Figure 3b, this height corresponds to

relative uncertainty in the skycal ensemble. If the height is reduced by 5 cm, we enter the small stable region (3.75 kHz).

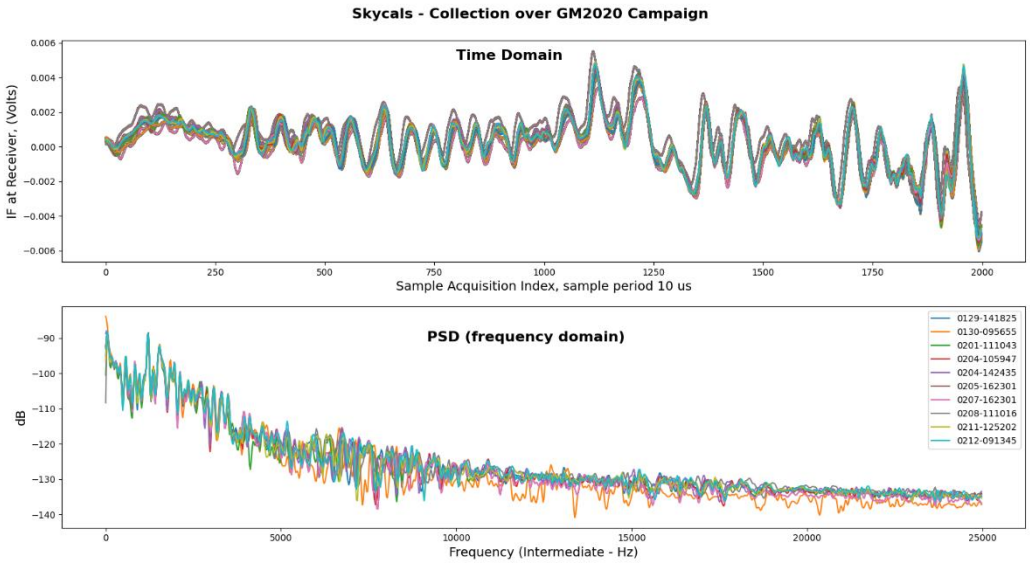

Figure 3. Time and frequency domain plots of all skycals. The signals are coherent up to ~ 4 kHz. We consider the orange trace (0130) a poor skycal and this has had consequences in our analysis.

Another consideration is to raise the radar significantly, translating the response to higher frequencies. Since the skycal has a benign response after 10 kHz, it would no longer be necessary in processing. However, the maximum depth will be impacted and the return signal weakened due to the increased distance.

Initially, it seemed desirable to eliminate the skycal measurement. It is time consuming in the field to prepare the radar and it can be processed in multiple ways, all of which need investigating. For instance, division in frequency space as opposed to subtraction in the time domain (i.e. as deconvolution to remove a systematic response resulting in a 'true' response).

For the machine learning approaches herein, the skycal became indispensable because it bounds the region where the signal occurs, allowing the top and bottom spaces to be partitioned, as shown in Figure 4a. This improves computation time for both the training of the neural network and the iteration step of the GA algorithm. Its use also aids in quantifying properties of the snow and radar by identifying features which may be used for modeling; as the snow-depth increases the radar footprint

illuminates larger areas and at different ratios between the extremes of bandwidth. This will broaden the peak of the snow/ground interface. Differences in snow density can also broaden the signal (see Case Study II).

Windowing in the time domain is a widely applied method to help suppress 'side-modes', or gain a better estimate of a continuous function. Without windowing, however, the skycal allows the signal crossover points to be well-defined, as shown below. Windowing also exacerbates the shape of the PSD; for processing it is of benefit to remove this shape. It is an artifact of semiconductor devices, particularly the mixer.

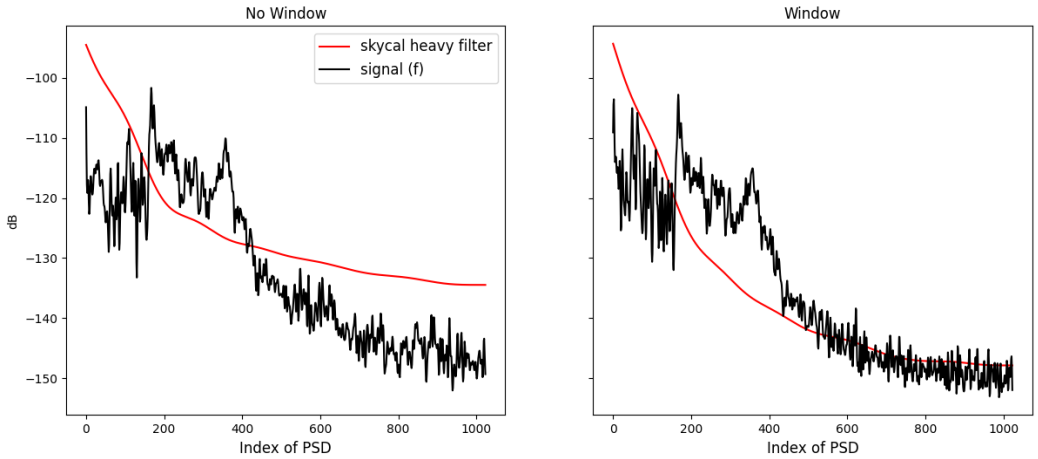

Figure 4. a.) A low pass filtered skycal bounds the signal for machine learning applications. b.) Applying a window in the time domain increases instability in the lower bound point.

2.2.2 /range resolution/frequency resolution

The range resolution of the FMCW radar, which is extreme, is inversely proportional to the bandwidth. For our system, we can typically resolve peaks in the PSD with a separation of 1.25 cm (24 Hz).

Since we are free to 'pad' the fft to achieve a desired binning, the number of points should be such that we at least match the FMCW resolution. We use n = 2^12 = 4096 which gives a bin resolution of approximately 0.5 cm. The time domain trace we are transforming has 2000 samples. If we do not pad the fft, the bin resolution will be 1 cm, which is acceptable. However, for computational efficiency we can use 2^11 = 2048. The genetic algorithm uses 2^12, whereas, the neural network was reduced to 2^11 as the number of classes was prohibitive for the bottom layer.

2.2.3 /filtering/antenna pattern

The final trace is Savitzky-Golay filtered on the record in the direction of motion, after first averaging the up and down PSD's. This filter retains the inherent sharpness of the PSD. We have previously demonstrated a correlation window over approximately 15 indices along the direction of motion and this is presumably due to the high spatial overlap of the radar footprint between adjacent traces. The window used in this filter is of length 9. The depth-axis of the record is low pass filtered at 40 % of Nyquist with a 6-pole Butterworth, based on colormap interpretations of the effect on stratigraphy. The filter is applied back-and-forth. The antenna patter, or beam divergence vs. frequency, effects the reflection signatures as much as the properties of the snow. Thus, modeling one, requires understanding the coupling with the second. Measuring in the laboratory the antenna pattern or consulting the datasheets will help in this effort.

2.2.4 /Spatial Interpolation and time synchronization

The GPS logging system runs independently of the radar. It records a location every second and is accurate to 1.000 seconds. The laptop writes data records to file every 43 seconds and the file name contains the local computer time, queried from the OS, signifying the start of the record. With this timing accuracy, there is no interpolation in the time domain. To compare the manual depth probe observations to the nearest radar locations, we interpolate between GPS locations given every 1 sec, with the number of points commensurate with motion at 5 m/s and 40 ms/ trace. This requires approximately 6000 points over the length of the parameterized spiral.

3 Grand Mesa 2020

3.1 Site Description

The GM2020 field campaign in Grand Mesa, Colorado, occurred from 29 January through 12 February, 2020. While many measurements of snow properties were obtained, our focus is to verify the functioning of the FMCW radar system against 'ground truth' measurements made with manual depth probes. In a typical scenario, a snow pit is made, and emanating from this, observers walk a spiral pattern on the snow and record the snow depth using the depth probe. After the FMCW radar data is processed, we compare the picks against this manual depth measurement. The spiral is typically 1 km in length with several hundred manually entered depths and their geospatial coordinates, measuring about 60 meters in diameter. The FMCW radar repeats this pattern as best as is practically possible. After processing, the two depth distributions can be compared, for instance with metrics such as the estimator bias, correlation coefficient, and mean standard error. Figure 5 shows a site before and after processing, so the radar and probe depth can be compared.

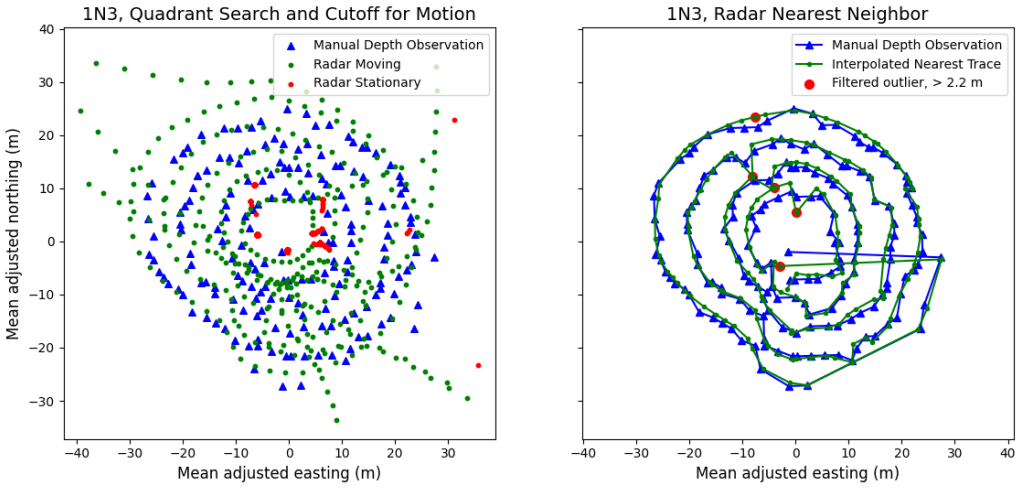

Figure 5. a.) A typical site spiral (manual depth observations) with moving and stationary radar points in the quadrant. b.) The interpolated nearest radar points, used for analysis.

4 Methods – Layer Picking

4.1 Introduction

The overall monotonically decreasing curvature of the PSD presents difficulties for layer picking. Simple thresholding, or maximum value picks are generally inadequate, even if the respective subspaces can be partitioned. For both layers, the pick is often not the signal maximum, as buried layers may reflect more strongly than the air/snow or snow/ground interface. Also, for the bottom picks, the general downward curvature of the PSD gives higher signal levels to layers within the snowpack. We propose two methods for layer picking. We first use a Genetic Algorithm to obtain the picks, and subsequently use this result as the supervised data set for a multilayer perceptron neural network with a single hidden layer.

4.2 GA

The genetic algorithm is effective because it is possible to quantify desirable features of the radar signatures. These features then become part of a fitness function. When the GA terminates, the fitness function is maximized and the picks are simply the fittest 'individual'. The algorithms for the top and bottom picks are run separately. Clearly, the signal level is an important feature and is used in both cases. The second top layer feature is a measure of the rise of the signal. From Figure 6, we can see the top layer pick generally follows a long sustained increase in signal strength and a short decrease after the pick. Quantifying this notion is the second feature. Finally, the third feature of the fitness function is a continuity requirement. Trace to trace variation in the pick is limited and we can impose a penalty if this variation is greater than a predetermined threshold. The bottom layer fitness function also uses the signal strength, but the overall curvature of the PSD can cause picks in the snowpack due to high stratigraphy. The second feature then seeks to penalize picks which are too far away from the lower bound point (the solid red line in the figure below). Finally, a continuity requirement can be applied to the bottom fitness function as well. The features are transformed to have a range from zero to one. Each feature is weighted, with the weights summing to one. The value of the fitness function is then the inner product of the weights and features. The continuity feature counts the number of times the individual (hypothesis) trace-trace pick difference exceeds the threshold.

Since we want to minimize this value, and maximize the fitness function, we subtract it from one.

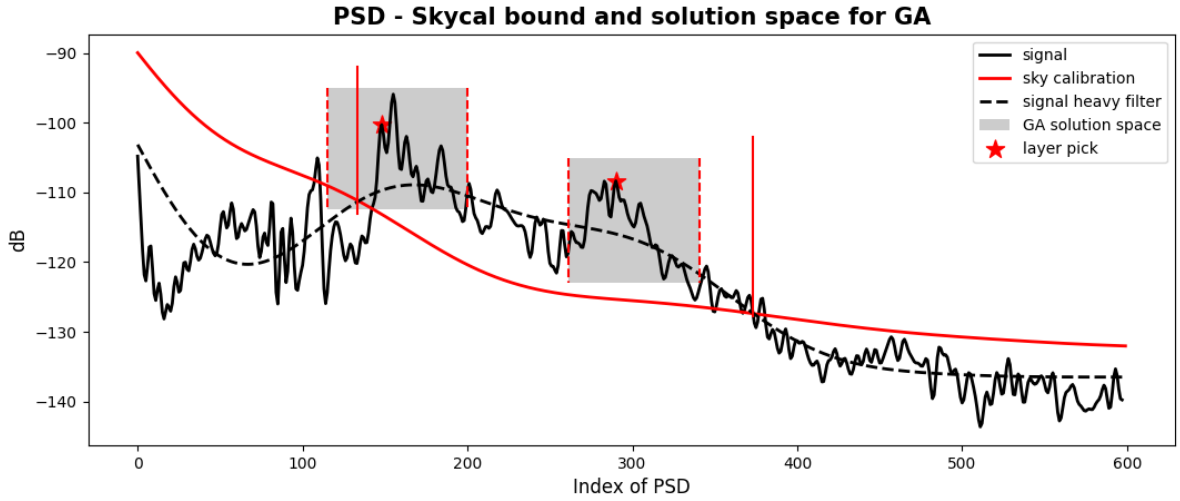

Figure 6. Using the skycal to determine the upper and lower bound points and hence the GA solution space.

4.3 NN

A neural network approach to layer picking is well suited for certain applications. The supervised set can be used across multiple seasons/field campaigns and this will have a significant impact on labor and processing time. The prediction step is a computationally fast multiply-add and hence can be used in a real-time control system, for instance, in maintaining a drone at constant height above either the air/snow or snow/ground interface. But configurations and hardware components change and the supervised data may no longer adequately represent the current conditions. By design of the network, the input layer incorporates local 2-D spatial variability seamlessly (as contrasted with the genetic algorithm and the continuity component of the fitness function). Also, the length of the input layer can be reduced as the traces are filtered with a Solvitzy-Golay filter along the direction of motion.

The approach is adapted from the classic example of digit recognition of numbers 0–9. Given a picture of a handwritten digit, we want to correctly identify the number. For numbers, 0-9, there are 10 classes, and the input layer to the neural network is a flattened vector of the 2-D black/white picture of the number. For our case, the number of classes must at least be

equal to the range of the picks (top or bottom). The input layer is a flattened vector of the PSD colormap; the trace under consideration and the two adjacent traces. Since the transmitter is mounted with a known distance above the air/snow interface, the range of the top layer picks is small. The bottom layer is limited only by the Nyquist range and hence will have a large number of classes compared to the top layer.

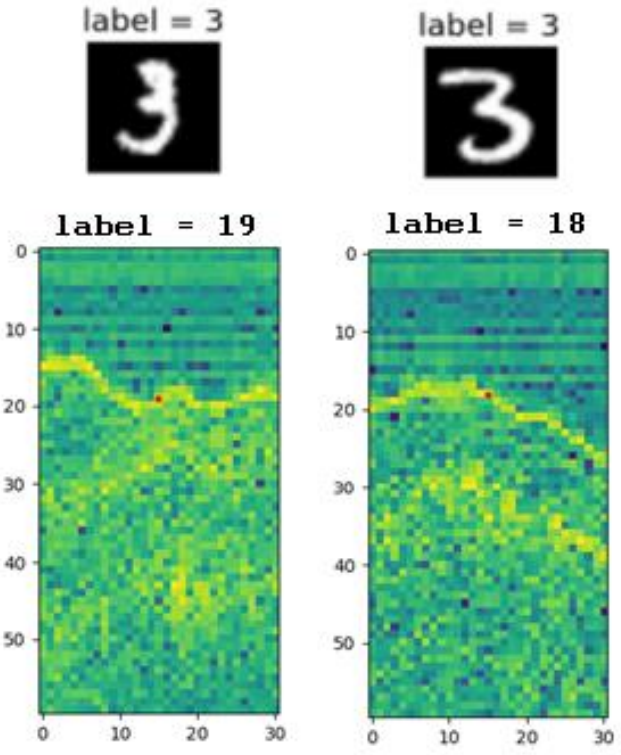

Figure 7. Supervised data structure analogous to the classic example of digit pattern recognition. The image (colormap) is flattened to a vector and become the input layer to the neural network. The class label is represented as the y axis.

We trained a neural network with a single hidden layer with approximately 8000 picks (Case study II) from the genetic algorithm. During the training step we assign 80% of the picks for training and 20% for validation. Figure 8 shows the comparison. We achieved an accuracy on the top layer of 63% and an accuracy on the bottom layer of 32%. Due to the higher number of bottom layer classes we see the accuracy is reduced compared to the top layer picks. Hence, for the same top and bottom accuracy, we will need a larger supervised data set for the bottom layer picks.

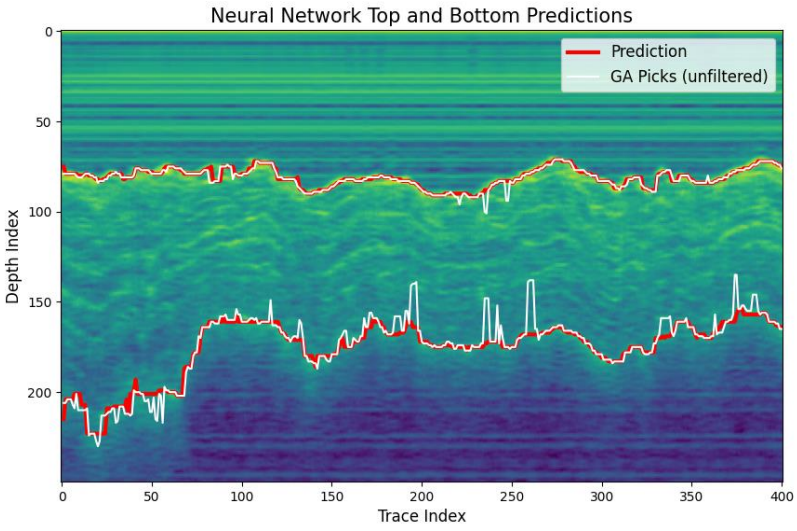

Figure 8. Results of the neural network prediction (red) vs. the class labels (white). The network filters the noisy class labels.

5 Results & Discussion

With the picks obtained from the GA we analyzed 24 sites, comparing the depth distributions of the picks and the manual depth probes. Our primary metrics for comparison are the radar bias (relative to the depth probe measurements), correlation coefficient, and mean standard error of the difference distribution. These values are shown in Table 1. Using the 3 metrics listed above, we can form their distribution by including all 24 sites and presenting it as a boxplot, as shown in Figure 9.

Site_Day	Sample size	Filter Sample size	Mean Probe Depth (cm)	Mean FMCW Depth (cm)	Bias (%)	CC	Mean Standard Error (cm)
1C5_20200212	174	165	81.971	78.691	-4.2	0.125	0.723
1N3_20200211	178	174	73.893	70.594	-4.7	0.342	0.917
1N7_20200211	287	236	81.732	74.788	-9.3	0.700	0.560
2C12_20200212	164	157	93.976	94.581	0.6	0.252	1.032
2C4_20200205	131	109	87.573	84.633	-3.5	0.239	0.980
2C9_20200205	223	200	89.211	84.403	-5.7	0.251	0.703
2N13_20200207	251	217	99.470	98.463	-1.0	0.295	1.093
2N4_20200211	209	137	101.330	93.809	-8.0	0.547	1.116
2S11_20200201	349	253	90.542	85.953	-5.3	0.502	0.845
3N26_20200208	290	279	91.879	90.248	-1.8	0.817	0.539
3S52_20200204	312	289	102.994	105.909	2.8	0.084	0.636
5S42_20200204	162	152	87.031	87.142	0.1	0.690	0.870
6N16_20200208	214	169	117.575	113.262	-3.8	0.819	0.891
6N31_20200130	207	160	104.353	103.442	-0.9	0.379	0.523
8C25_20200130	160	76	84.963	75.978	-11.8	0.447	1.757
9N28_20200208	510	398	103.337	95.808	-7.9	0.852	0.466
5C21_20200130	123	114	89.642	91.835	2.4	0.272	1.952
1C14_20200205	275	139	83.193	73.526	-13.1	0.247	0.917
1S1_20200207	317	228	87.770	81.715	-7.4	0.511	0.960
6N17_20200210	166	138	109.241	103.436	-5.6	0.218	1.435
5N11_20200210	285	239	89.000	86.880	-2.4	0.578	0.666

6S19_20200129	239	211	104.117	106.340	2.1	0.163	1.112
1C7_20200205	218	191	91.110	86.444	-5.4	0.099	0.881
6S44_20200204	297	241	102.256	103.637	1.3	0.089	1.038

Table 1

The data from Table 1 is presented as the boxplot 'Near Neigh'. The remaining 4 distributions are obtained by considering different methods or different picks altogether. 'RBF intp' generates a 2-D Surface map from the radar nearest neighbor points using radial basis functions. This functional representation is then evaluated at the ground truth coordinates. 'filter' is obtained by requiring the radar nearest neighbor points to be no further than 2.2 meters from the ground truth coordinates and discarding the outliers. This results in a reduced sample size. 'max signal' simply considers the maximum value of the PSD for both picks, in their respective subspaces. Likewise, 'bounds' considers the picks as the value of the upper and lower bounds, discussed previously.

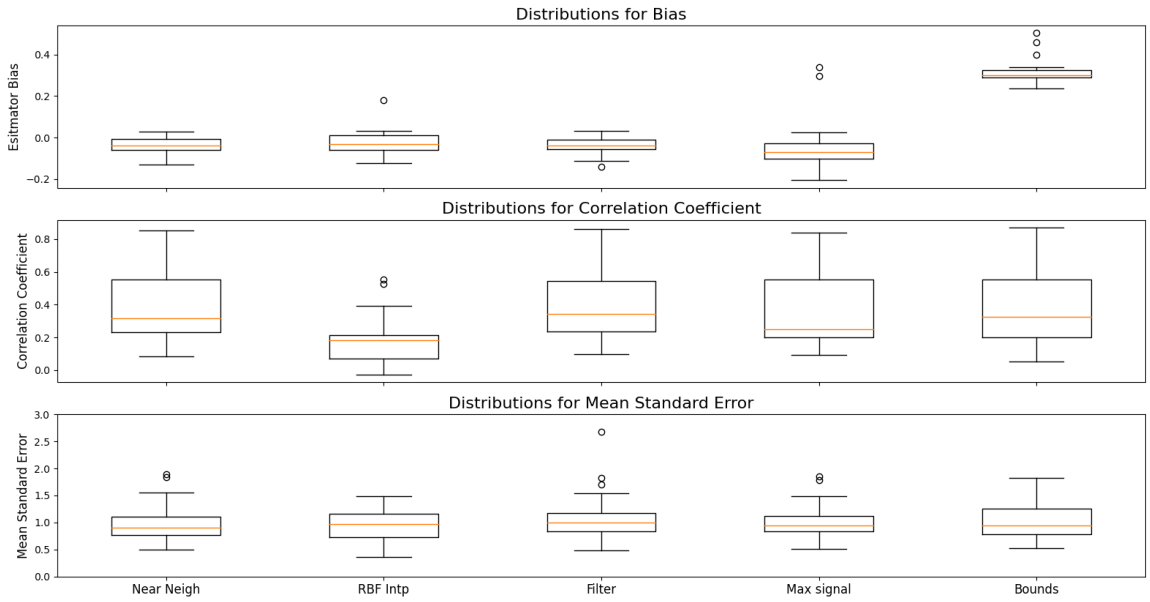

Figure 9. Three metrics: bias, correlation coefficient, and mean standard error over 24 sites for 5 processing methods.

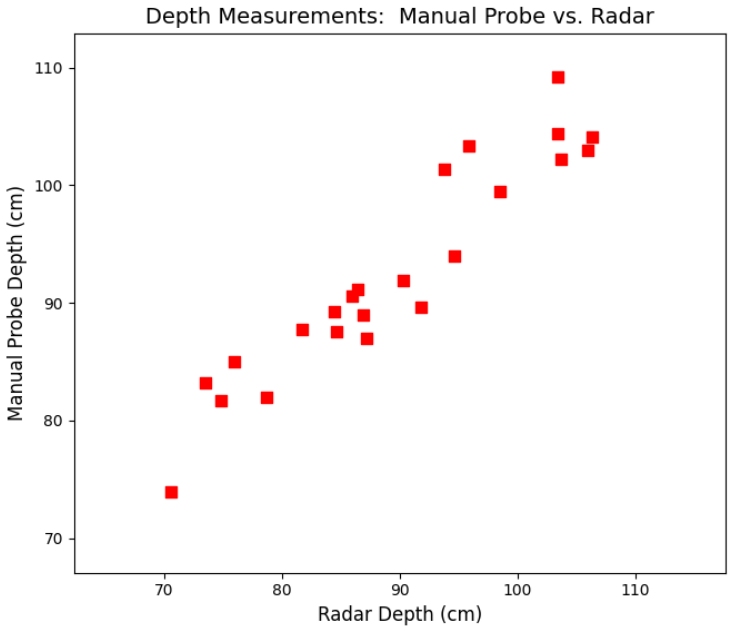

Figure 10. Scatter plot of the 24 sites analyzed, showing the mean depths.

The most consequential result from Figure 9, in relation to labor and processing time, is the benefit of using the upper and lower bounds to calculate snowpack depth. This is a simple, fast computation obviating the need for GA or NN picking. It has a high bias, as expected, but the distribution has significantly lower standard deviation than any of the other methods considered. Two of the three outliers in this distribution are the result of the bad skycal on 30 January.

The 2-D Radial Basis Function interpolation method may be useful as it has implications for use in the field. Currently, we try to position the radar as close to the manual depth probe measurement as is practical. This requires a high degree of field coordination. An RBF approach would allow the radar to be operated at points which optimally cover a 2-D area, and removes the need to obtain precise overlap. This would require some experimentation in the field, or perhaps simulation. Once the picks and their coordinates are obtained, the processing involved for this method is fast and minimal.

The 'near- neigh' vis-à-vis 'filter' analysis, in hindsight, has a couple of problems. The goal was to show improved performance when compared with the filtered case, which removes nearest neighbor outliers. This radar

system has an extremely high bandwidth and hence very good range resolution. As a result we would expect, for the radar and depth probe measurement at exactly coincident locations, very good agreement, at least in correlation. The filtered case does show marginally improved correlation, but we are comparing distributions with different sample sizes. The mean standard error, which accounts for sample size, is a little higher for the filtered case and has a higher outlier count. One possible reason is GPS accuracy. The ground truth coordinates are logged with a GPS system with less accuracy than the GPS system used with the FMCW. A better method for comparison may be to consider all the radar traces within a certain radius (commensurate with the least accurate GPS), compute these picks and average the result.

Finally, the picks themselves contain a high level of noise. We have three large picked data sets to compare: those for 24 sites and Case Studies I and II. The range of top picks for the 24 sites is large, as shown in Table 2. The site analysis was done on traces individually picked from the continuous record and ordered according to the spiral. The continuity component of the fitness function had to be dropped in this case, as adjacent traces no longer represent the next sample time. A better method for analyzing the sites would be to run the GA on the continuous record and then use the necessary picks from that record. This is currently prohibitive as the GA algorithm would take too long to process the approximately 15 records per site. A quicker method was tried, processing 15 traces on both sides of the trace under consideration, but the GA did not perform well under this condition and the results are not presented.

	Sample points	**Top pick minimum**	**Top pick maximum**	**Top pick mean**	**Top pick std**
24 Sites	5741	76	265	160.56	15.85
Case 1	7992	120	220	160.98	11.02
Case 2	4995	125	231	163.36	12.73

Table 2. The top layer picks for the 24 sites show higher noise as the continuity component for the GA fitness function is not applied.

6 Case Studies I & II

6.1 I: High vs Low Stratigraphy Classifier

6.1.1 Motivation

The presentation of the FMCW traces as a colormap offers an opportunity for binary classification based on a visual interpretation of the image. A sampling of the records indicates two distinct images, classed as high and low stratigraphy, as shown in Figure 12. The low stratigraphy case was obtained while operating the snowmachine/radar on an ice road while in transit to the testing sites.

Figure 11. The low stratigraphy examples are due to travel on the ice-road.

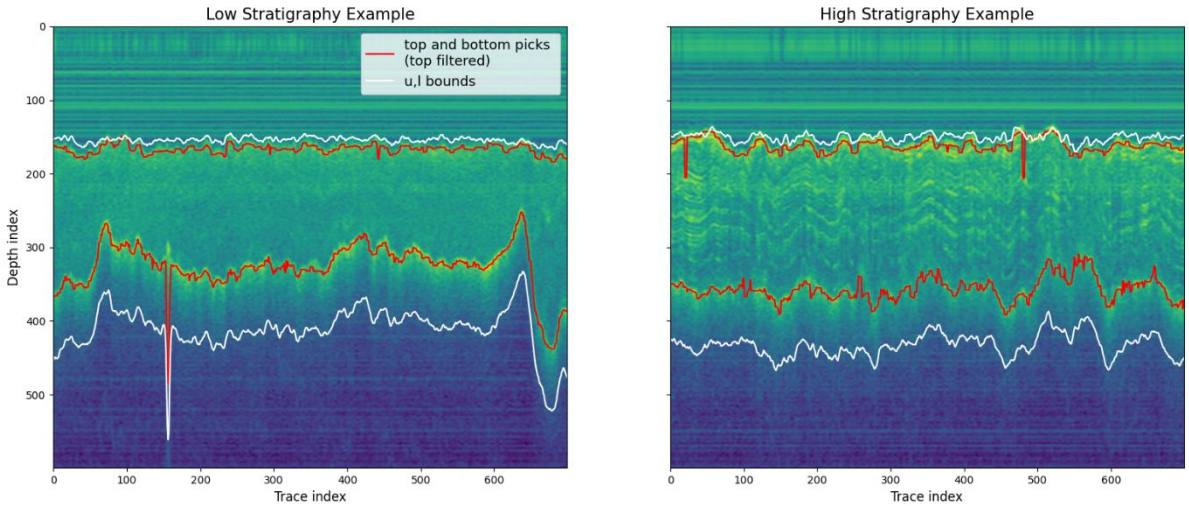

Figure 12. GA picks (red) with upper and lower bounds (white) for the low and high stratigraphy classes.

The features upon which to base the classification are properties of the signal. For instance the PSD value at each layer pick; it has been observed that in the low stratigraphy case, generally, the top layer signal value is less than for the high stratigraphy case; the bottom layer's is greater than its high stratigraphy counterpart. This suggests for the low stratigraphy case, more energy is entering the snowpack. Since we expect signal attenuation with greater snowpack depths this quantity is also included. We want the classifier to consider both absolute and average values so the integrated signal above the skycal and the integrated signal from the second pick to the lower bound add two more features. The final two features are the lengths over which the just mentioned integrations occurred. This gives a total of six features. There are some obvious ways to explore dimensional reduction, but these were not tested, nor was using spectral analysis, such as principal component analysis. This would be useful for a more rigorous approach.

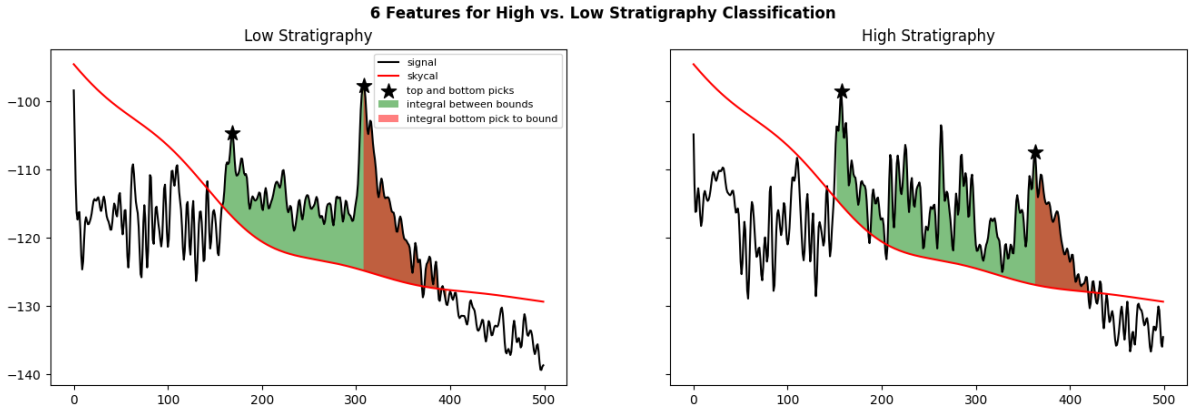

Figure 13. Identification of 6 features used to classify high and low stratigraphy classes.

6.1.2 Results and Discussion

The latter, integrated signal from the bottom pick, is included in an attempt to quantify a broadening term associated with the snow/ground interface. Establishing a way to quantify this metric, and showing its dependence would assist us in understanding the properties of the snow or the radar. If it is a radar induced effect, then the response we are seeing is not truely due to the properties of the snow. For instance, a known radar induced effect is the dependence of the transmitting antenna (gain) on frequency; the 'footprints' at 6 and 18 GHz are substantially different and are

dependent on depth. A snow induced broadening effect is from dispersion arising from density differences. From existing models of the dielectric constant (density), we can calculate what change we expect to see and verify this with the PSD. Another possible broadening mechanism is energy storage and re-radiation. The high-stratigraphy case has the potential to support standing wave excitation because the intra-distance separation between snowpack layers is commensurate with the radar's rf wavelength. The radar traces are collected at 20 ms sampling, so even with modest motion (~ 5 m/s) the transmitter is still illuminating the same footprint on successive traces. If conditions are right, a very low Q resonator can be established. As the radar passes off the original footprint, this stored energy is released and the result will be a broadening of the snow/ground interface peak, as these echoes arrive back at the receiver much later than usual due to their longer path travelled. This condition is apparent for instance when there is a pool of water on the ground under the snow, or perhaps something metallic. See Figure 12a. For our two cases here, high and low stratigraphy it is difficult to judge at present if the peak width (the 'distance' from bottom pick to lower skycal bound) is greater for the high stratigraphy case. Difficulties arise because we should compare everything at the same depth and also the low stratigraphy case has a higher PSD value at the bottom pick.

Lastly, once we have the classifier modeled we can present uncertain traces and have it classifier them. This was done on an ambiguous record, and the predictions are shown in Figure 14.

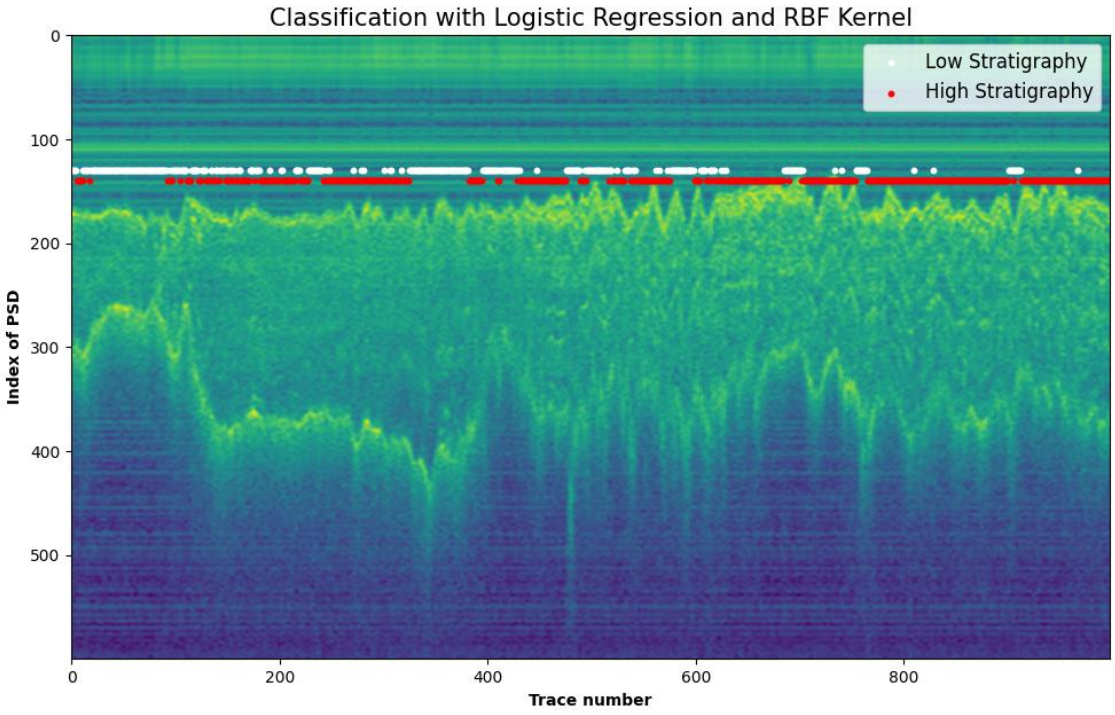

Figure 14. Trace prediction results for the classifier on an ambiguous record.

6.2 II: Towards a High vs. Low Density Classifier

6.2.1 Motivation

The FMCW radar is a time-of-flight device and time (distance) is in 1-1 correspondence with the IF frequency. If the echo is travelling through denser snow it will take longer to return to the receiver. By bounding the signal with the skycal and considering the bottom layer pick, we calculate the length of segment show in red below in Figure 15. This length we call the 'f1' feature and is considered as ½ width measure for the bottom layer response to the full bandwidth of the FMCW. This is a better controlled study of the same feature discussed in Case I.

From the GM2020 campaign, we have an average snowpack density associated with each site. The GA ran for top and bottom picks on eight records. Two records each from two low density sites and two records each from two high density sites. These are 2N13, 6N16 (~ 230 kg/m^3), and 2S3, 6N18, (~ 330 kg/m^3) respectively. We then partition the results for each class into depth bins and plot the resulting f1 histograms as shown in Figure 16. By looking at a constant depth slice, we can compare the f1

feature across the two density classes. By comparing across depth slices, we can decouple the broadening associated with the antenna pattern.

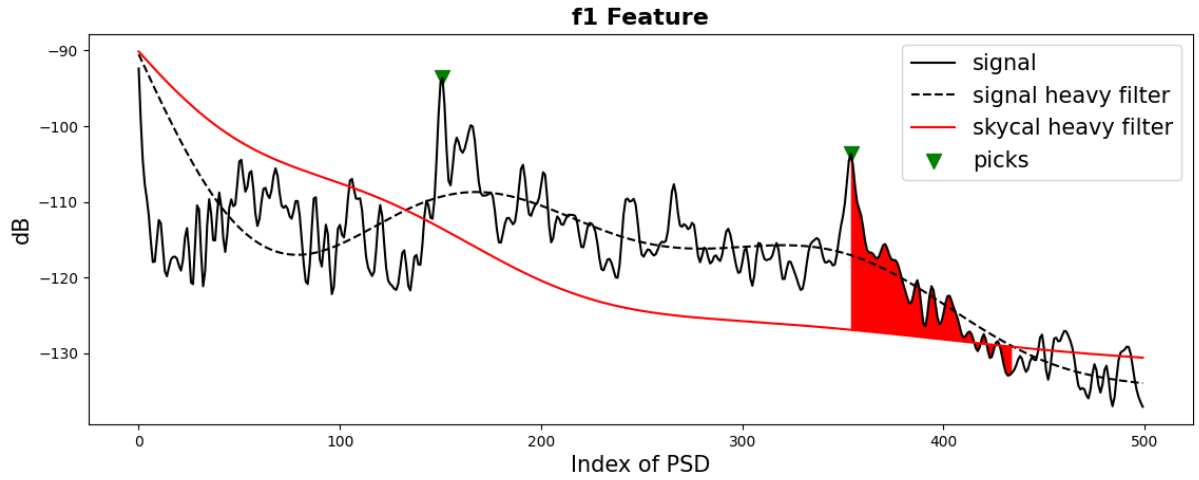

Figure 15. The ½ width (length along the x-axis) of the snow/ground signal used to analyze snowpack density.

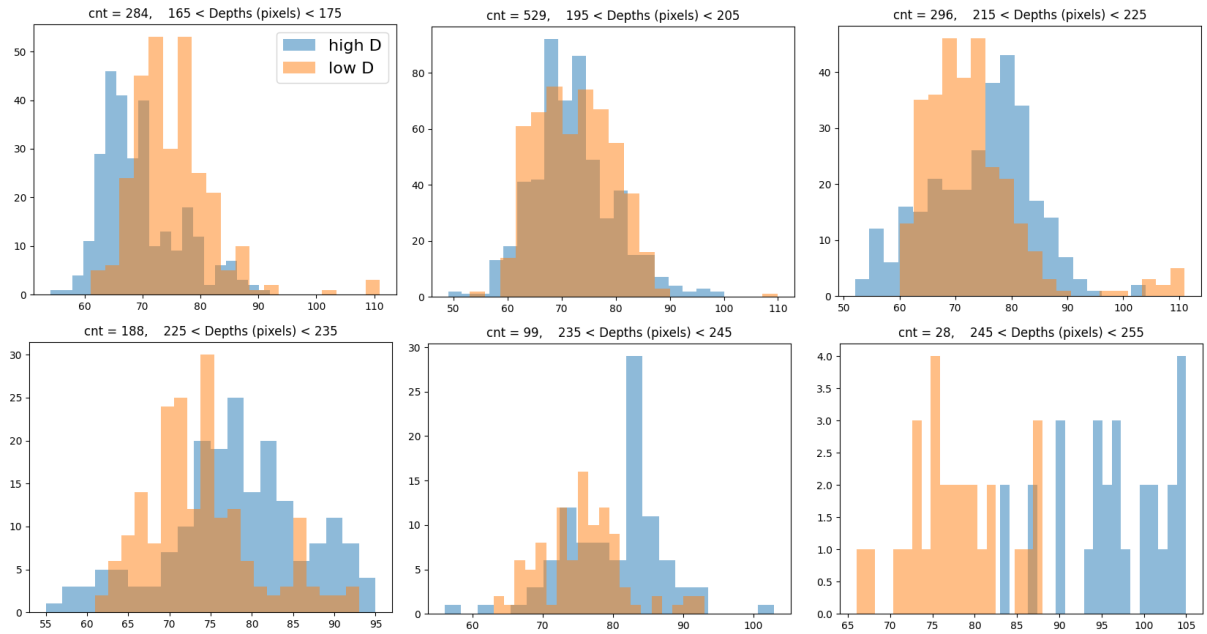

Figure 16. Histograms of the f1 feature for depth intervals (10 pixels or approximately 5 cm). The high density feature gradually moves out to higher mean values compared with the low density case.

As a control case, we can compare across the two cases in each class, and we see the distributions remain fixed relative to each other as the depth increases.

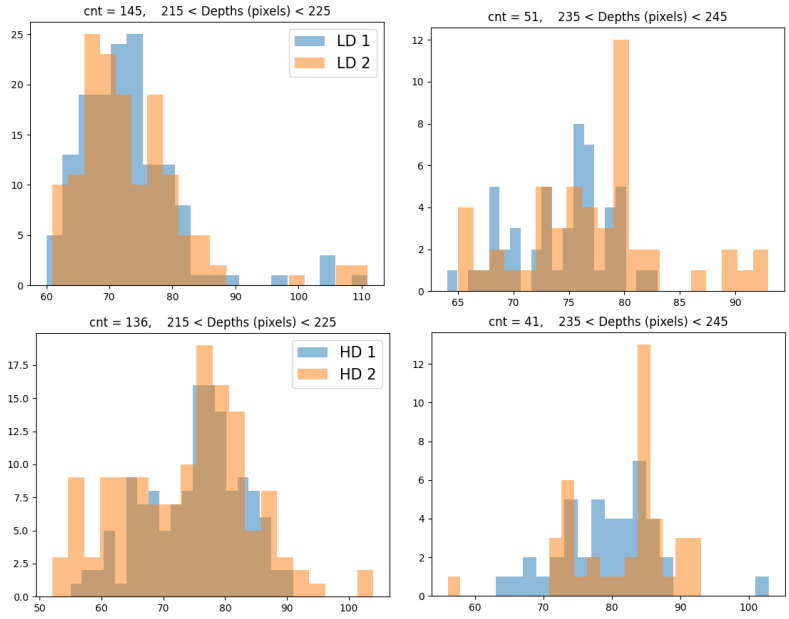

Figure 17. A control case showing the relative stationarity for low and high density cases.

The mean values are plotted as shown in Figure 18.

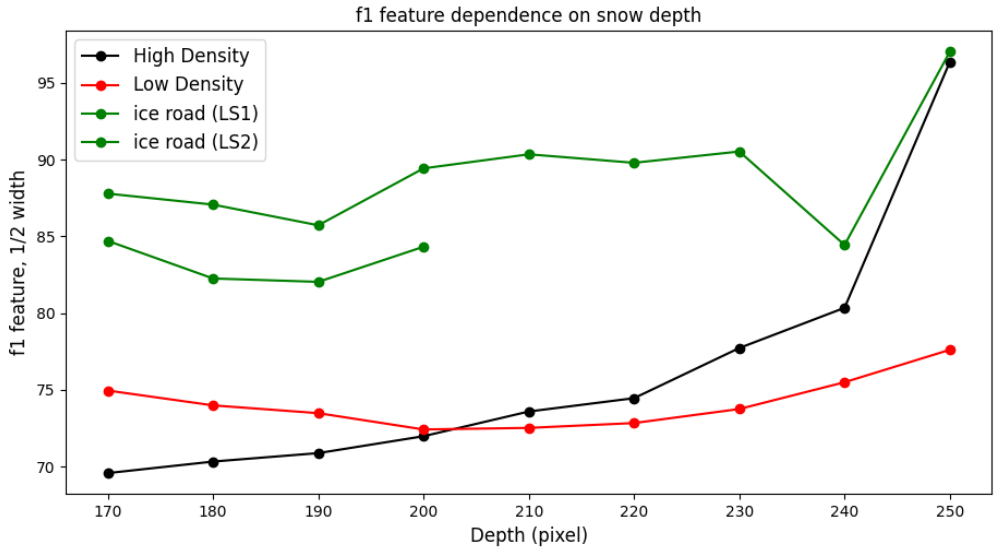

Figure 18. Plot of the mean value of the f1 feature vs Depth.

6.2.2 Results and Discussion

We can see high and low densities have different slopes and start diverging at a depth index of 200. The green plots are from the previous case study; the low stratigraphy case is an ice road, which we consider here to have a high density. The slope of the red and green curves suggest the antenna pattern dependence on depth is not great. The divergence of the high density case is curious. The intent was to demonstrate an effect, but the method is not rigorous. The sample sizes, for instance, are different across the depth slices, and the range of depths itself is not large, as the overall dataset does not contain large differences in snow-depth. Another good control may be a large sample from densities in the center of the density distribution. Also the average power feature, the integral of f1, divided by f1 may be a better feature to consider.

7 Conclusions and Recommendations

The use of the skycal to bound the solution space greatly reduces the search requirements for the genetic algorithm and the number of classes for the neural network. The GA for the bottom picks is relatively fast, requiring approximately 2 hours per record. However, the GA for the top picks require about 18 hours. This is primarily due to the calculation of the second feature. Future effort should seek to reduce this computation time. Experience has shown that the continuity feature (the number of counts above the threshold) is rarely reduced to zero. It seems the population of hypotheses may quickly lose diversity around this point and hence higher weighting on the continuity feature will not help. This is primarily why the mutation rate is so high, with the hope it will eventually settle to a better pick. The benefit of running the GA for 18 hours is marginal as it gets close well before this time. Perhaps digital filtering is a better way to reduce the run time and still have reasonable picks.

The bottom picks on records were done with the continuity weight set to zero, as some records indeed have a very high rate of change. This was a poor choice and more development work should be done on this component of the fitness function. Perhaps it should contain a window around the point instead of looking only at the subsequent trace.

The intent of the continuity feature was to 'bridge' the pick, precisely where the trace goes 'dark'. Why this occurs is not understood. An IMU unit may help clarify this anomaly, as perhaps the transmitter is poorly pointed towards the ground. This is a rare but consistently recurring problem.

The neural network is relatively simple and the code is practically cut and pasted from examples on the internet. Keras is high-level and there are many parameters which can be tested for improved performance. The code itself is only about 30 lines. The network performed surprisingly well on the bottom layer with high levels of noise in the supervised data set. The supervised data should uniformly cover the range of depths expected. If we had a representative supervised data set prior to the GM2020 campaign the the 24 sites could be picked (predicted) in a less than a second. This is contrasted with the GA which can run for a week on a handful of records.

In many applications, the skycal measurement is not viable because the system cannot be pointed towards the sky. For instance, on a drone, helicopter, etc. There is a belief the skycal is 'measuring' the in-air coupling between the transmitter and receiver, but lab measurements do not confirm this hypothesis. The transmitter and receiver pointed in opposite directions produce the same skycal as when they are colinear. Furthermore a change of mixers in the system will change the skycal. Rather, it is believed the skycal is measuring, primarily, the systematic response of the mixer while the transmitter is disabled (in mixer parlance, the 'DC Offset'). If we wish to eliminate the skycal as currently practiced, we can temporarily disable the transmitter by switching it into a 50 ohm load and measure the system's response. This way it can be used in systems where the skycal is not possible. It can be automated using a digital I/O bit to perform the switching. This systematic response can be deconvolved out of the signal, which implies division in frequency space as opposed to subtraction in the time domain [2]. Division will flatten the PSD and make processing easier.

The RF hardware would benefit from an analog low pass filter with a cutoff of 1-2 kHz. Unfortunately this is not an off-the-shelf component as it is directly connected to the IF port of the mixer (i.e. an rf device). It must have a 50 ohm input impedance, and be an absorptive (not reflective) filter with a good return loss up to several hundred MHz. Subtracting the skycal in the time domain greatly reduces the power in the low frequency band of the PSD, but it does not entirely eliminate the DC component (i.e. the first value of the PSD). Figure 6 shows how the digital low pass filter applied to the signal propagates the DC component forward to higher frequencies. This may unduly influence the calculation of the upper bound point. Furthermore, capacitive coupling has been observed between the ramp signal (0 – 10 V) and the very small IF signal (< 1 mV), so this may be adding power in the low frequency band of the PSD (assigning the ramp signal to a port as far away from the signals being digitized will help).

With our current hardware we can reduce the ADC sample rate to 80 kHz and this will allow digitization of both the I and Q channels of the mixer. The baseband (IF) signal will be complex and we compute the fft with a complex number using the I channel for the real part and the Q channel for the imaginary part. This should reduce the noise.

References

1.) Fundamentals of Radar Signal Processing, Mark A. Richards, 2nd Ed, McGraw-Hill.

2.) Numerical Recipes, Press, Teukolsky, Vetterling, Flannery, 3rd Ed, Cambridge University Press.

3.) Machine Learning, Tom M. Mitchell, McGraw-Hill

4.) Computational Statistics, Givens, Hoeting, 2nd Ed, Wiley.

Appendix A: Comments on Code

7.1.1 Files

The original data files used in the processing are matlab's .mat file and .csv files.

The processing in matlab converts the .mat file to ramp up and down traces in the time domain keeping the same record data structure of approximately 43 seconds per record. This code is in 'get_record.m'. All subsequent processing is done with python, generally from .npy files, which load fast and minimize data storage.

The raw stream GPS (lat, lon) text files are converted to utm coordinates and the time converted to local cpu time and stored as a .csv file in the ./gps folder. The two independent GPS systems on board are both processed in this way.

The skycal files are in ./skycal and there is one per day.

7.1.2 Python Code

There are three classes, GPS, FMCW, and FILE_IO which do most of the data computation and inputting of data. This way analysis and presentations can be prepared mostly by defining the file names and instantiating an object to return the desired processed data.

The GPS class has different versions depending on whether the remote data server is used or the local hard drive. Full processing has to be done on the server, as directories are walked through to obtain the necessary files. For this class, all methods are executed from '__init__'.

The FMCW class, or FMCWhr (high resolution) processes the radar data.

The genetic algorithm code runs from '__main__' and is mostly self-contained; def's are defined in the script. Top and bottom picks have separate scripts and versions.

In cases where intermediate files are generated, a suffix is added to the existing filename and saved. Generally, the .csv files were converted to .npy files for speed and the .csv files deleted.

The last existing code on the server will be placed in the software repository in the /working directory. Code used from local machine files will be placed in the /local directory.

Hardcoded values exist, unfortunately, generally appearing at the top of the file and in all caps.