Modelling Avalanche Risk Using ERA-5 Climate Data and Multiple Linear Regression

EOSC 410 – Term Project

Evan Williams & Simon Campbell

April 7th, 2020

Introduction

Research Questions:

(1) Which climate predictor(s) is/are the most important predictor(s) of avalanche risk for each elevation level in the Sea to Sky Region?

(2) Do similar important climatic predictors of avalanche risk exist between different mountain regions for the three different elevation levels (Below Treeline, Treeline, Alpine)?

Background

Avalanche

- A mass of snow / ice that loses its hold and is discharged down a slope
- Loose/wet Avalanches vs. cohesive slab Avalanches
- Slab avalanches occur when a dense layer of snow (slab) fails on an underlying weak layer in the snowpack, where the shear stress exceeds the shear strength between snow grains

Avalanche Risk

- Refers to both the likelihood and consequence of avalanches occurring in a given region, at a given time
- Depends on numerous factors such as: recent loading by precipitation and wind, warming of the snowpack by solar radiation, formation of unstable layers in the snowpack, etc…

[Avalanche Canada, 2020], [Stetham et.al, 2003]

Background

How do Climate Variables affect Avalanche Risk?

- Avalanches are caused by instabilities in the snowpack, which become activated by either human or natural triggers

Precipitation: loading of the snowpack, adding weight onto potential weak layers

Wind: loading aby re-distribution and formation of cohesive slabs due to saltation of snow crystals

Temperature: heating up of the snowpack by solar radiation increases snowpack density, which exerts more weight onto potential weak layers. Extreme temperature gradients above and below the snow surface can cause the formation of faceted crystals (surface hoar), which will be a potential weak layer once buried

Air pressure: affects wind and precipitation. Atmospheric low-pressure systems are usually associated with precipitation events.

[McClung & Schaerer, 2006]

Background

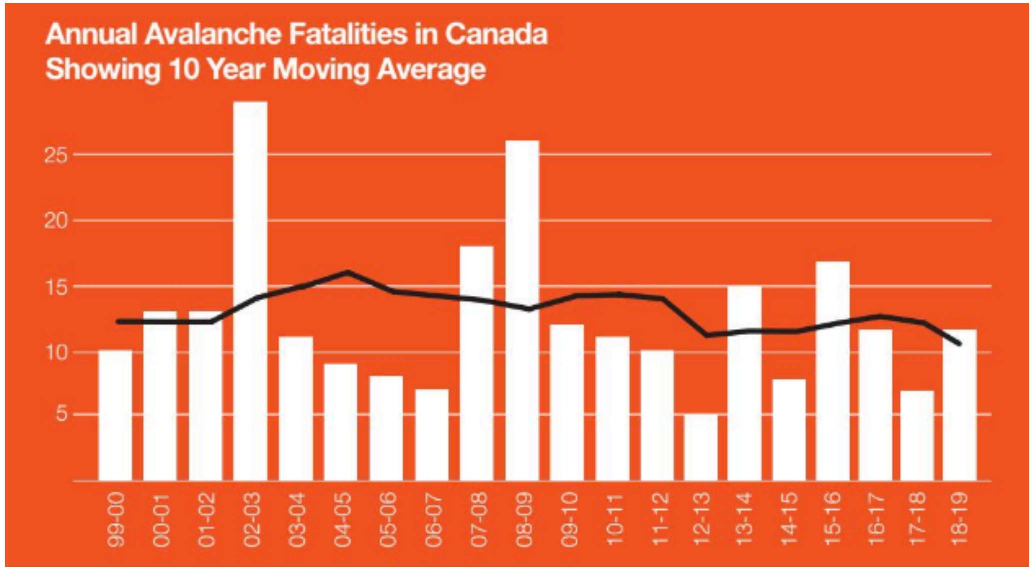
How is this Relevant?

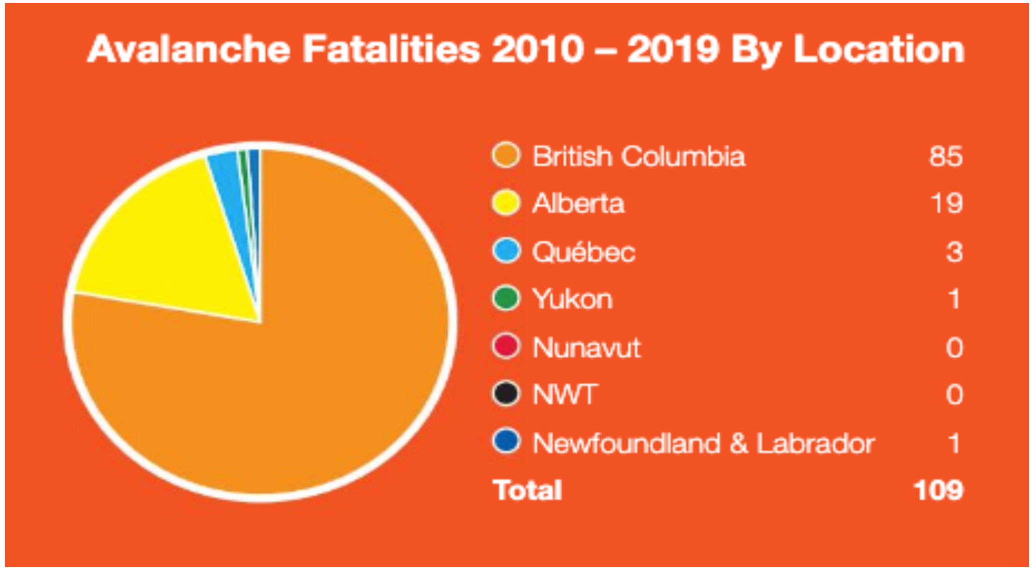
- Avalanche danger ratings are key resources for backcountry users
- Affects avalanche control measures in mountainous environments - ski resorts, guiding servicers, highways, infrastructure, resource industry, etc.

Trans-Canada Highway – Glacier National Park:

- 3000 motorists pass through daily in winter months
- Highway crosses 130 avalanche paths
- A 2-hour closure at rogers pass is estimated to result in 50 000 - 90 000$
- On average, 11 backcountry users die due to avalanche related incidents each year in Canada

[Parks Canada, 2017], [Jamieson and Geldsetzer, 1996]



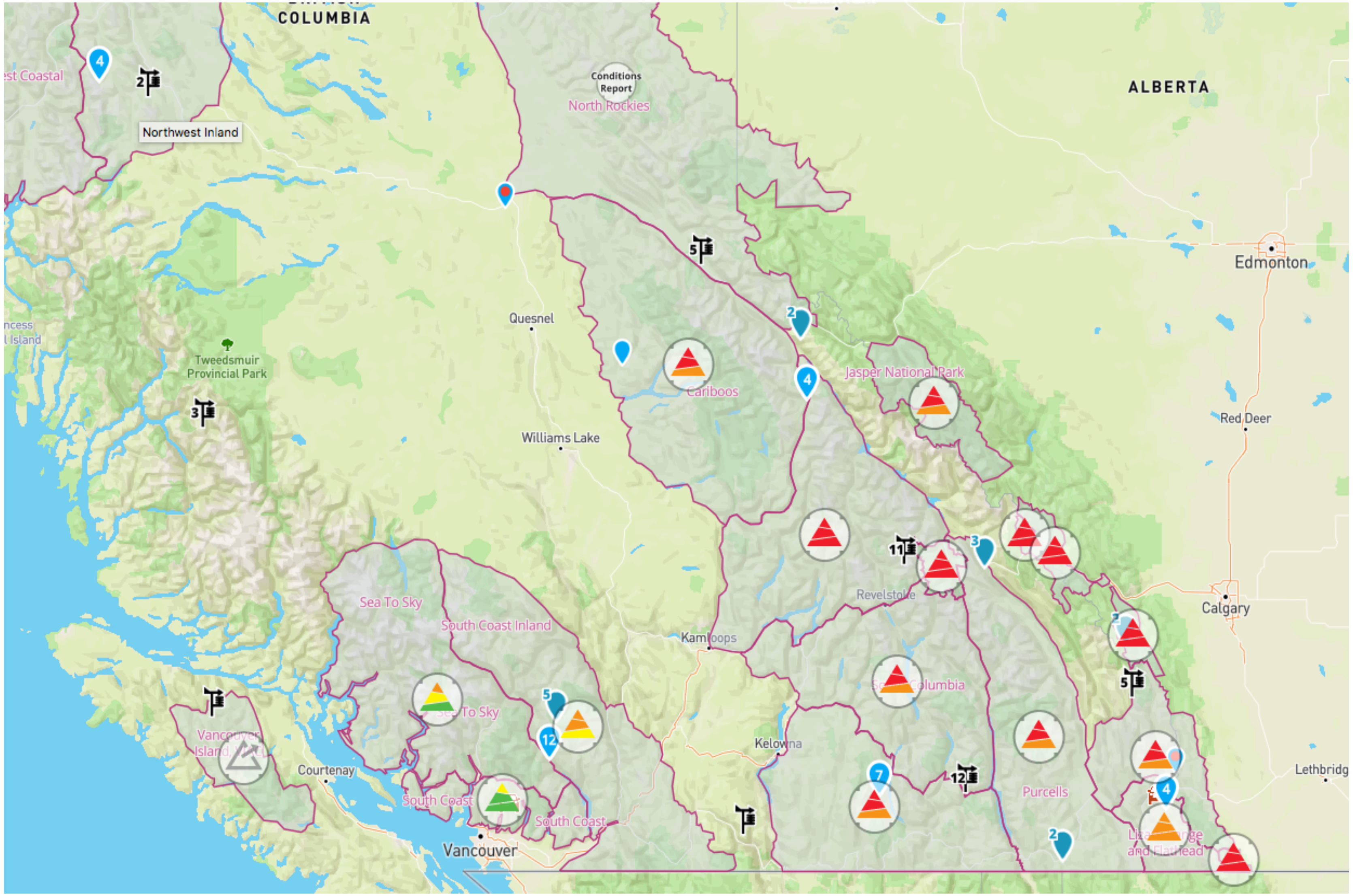


Study Area

Sea to Sky Region

- Encompasses the majority of the south coast mountains accessed from Vancouver, Squamish and Whistler
- Stretches from Lions Bay to Pemberton along Highway 99
- Has seen a massive increase in backcountry users over the last decade

[Avalanche Canada, 2020]



Study Area

Other Canadian Regions

- Span from the coast mountains on the Pacific Ocean, to the Rocky Mountains in Alberta, as well as the interior mountain ranges in BC and Yukon
- Spatial data was extracted from a specified lat/lon box that encompasses the majority of the region

[Avalanche Canada, 2020]

Data

1. Avalanche danger ratings for each mountain region:

Timeframe: January 1st, 2012 – April 30th, 2017, Daily

Source: Avalanche Canada:

Type: DF: 4 Variables, 722 Observations: Dates and associated Avalanche danger (1-5) for each elevation level (Below Treeline, Treeline, Alpine)

Mountain Regions: Sea to Sky, Cariboo, Kananaskis, Kootenay Boundary, Lizard Range, North Columbia, Northwest Coastal, Northwest Inland, Purcell, South Coast, South Columbia, South Rockies

2. ERA5 re-analysis climate data (for each mountain region):

Timeframe: January 1st, 2012 – April 30th, 2017, Every 6 hours (00:00-06:00-12:00-18:00)

Source: Copernicus Climate Change Service (C3S) Climate Date Store: https://cds.climate.copernicus.eu/#!/search?text=ERA5&type=dataset

Type: 2-metre air temperature (t2m), 10-metre U wind component (u10), 10-metre V wind component (v10), mean sea level pressure (msl), total precipitation (tp)

Spatial Boundaries: Spatial data was extracted from a manually specified lat/lon box that encompasses the majority of each regions area.

Data

1. Avalanche danger ratings for each mountain region:

Format: .csv translated from .txt

Avalanche Data

```
In [162]: avy = pd.read_csv('ACdata_northwestcoastal.csv')

In [163]: np.shape(avy)
Out[163]: (722, 4)

In [164]: print(avy)
```

	Unnamed: 0	Below Treeline	Treeline	Above Treeline
0	2012-01-01	3	4	4
1	2012-01-02	3	4	4
2	2012-01-03	4	4	4
3	2012-01-04	4	4	4
4	2012-01-05	3	3	4
5	2012-01-06	3	3	4
6	2012-01-07	2	4	4
7	2012-01-08	3	4	4
8	2012-01-09	2	3	3
9	2012-01-10	1	2	3
10	2012-01-11	1	2	2
11	2012-01-12	1	3	3
12	2012-01-13	2	3	3
13	2012-01-14	1	2	3
14	2012-01-15	2	3	3
15	2012-01-16	2	2	3
16	2012-01-17	2	3	3
17	2012-01-18	2	3	3
18	2012-01-19	2	3	3

Raw Data format, Python 3x Screenshot

Data

2. ERA5 re-analysis climate data (for each mountain region):

Format: netCDF

ERA-5 Climate Re-Analysis Data

In [166]:
```
import xarray as xr
# single file
dataDIR = 'ERA5_P_S_T2m_U_V_SLP.nc'
DS = xr.open_dataset(dataDIR)
```

In [167]:
```
#this code also works: xr.decode_cf(dataERA)
DS.values
```

Out[167]:
```
<bound method Mapping.values of <xarray.Dataset>
Dimensions:    (latitude: 61, longitude: 101, time: 7248)
Coordinates:
  * longitude  (longitude) float32 -135.0 -134.75 -134.5 ... -110.25 -110.0
  * latitude   (latitude) float32 60.0 59.75 59.5 59.25 ... 45.5 45.25 45.0
  * time       (time) datetime64[ns] 2010-01-01 ... 2019-12-31T18:00:00
Data variables:
    u10        (time, latitude, longitude) float32 ...
    v10        (time, latitude, longitude) float32 ...
    t2m        (time, latitude, longitude) float32 ...
    msl        (time, latitude, longitude) float32 ...
    tp         (time, latitude, longitude) float32 ...
Attributes:
    Conventions:  CF-1.6
    history:      2020-03-06 00:03:09 GMT by grib_to_netcdf-2.16.0: /opt/ecmw...>
```

Raw data format: Python 3x Screenshot

Methodology – Research Question #1

Research Question #1: Which climate predictor(s) is/are the most important predictor(s) of avalanche risk for each elevation level in the Sea to Sky Region?

1. Upload the Sea to Sky region Avalanche Risk data into a Pandas dataframe.

2. Split the Sea to Sky regional dataframe into elevation level dataframes (Below Treeline, Treeline, Alpine).

3. Download ERA5 re-analysis climate data for the Sea to Sky Region and upload the data into a Pandas dataframes.

4. Normalized the ERA5 re-analysis climate data.

5. Build a multiple linear regression (MLR) model for each elevation level.

6. Calculate MLR coefficients for each normalized climate predictor to see which predictors are the most important predictors of avalanche class for each elevation level.

Methodology – Research Question #2

<u>Research Question #2</u>: Do similar important climatic predictors of avalanche risk exist between different mountain regions for the three different elevation levels (Below Treeline, Treeline, Alpine)?

1. Upload the Avalanche Risk data for additional mountain regions (Cariboos, Kananaskis, Kootenay Boundary, Lizard Range, North Columbia, North Rockies, Northwest Coastal, Northwest Inland, Purcells, South Coast, South Columbia, South Rockies, Yukon) into Pandas dataframes.
2. Split the regional dataframes into elevation dataframes (Below Treeline, Treeline, Alpine) for each mountain region. (1 regional dataset = 3 elevation datasets)
3. Download ERA5 re-analysis climate data for the additional mountain regions and upload them into a Pandas dataframes.
4. Build individual MLR models for each region and elevation (14 regions x 3 elevations = 42 models).
5. Compare the results of MLR for each elevation level between mountain regions, and assess whether there might be differing significant climatic predictors of avalanche risk between mountain regions.

Results: Sea to Sky

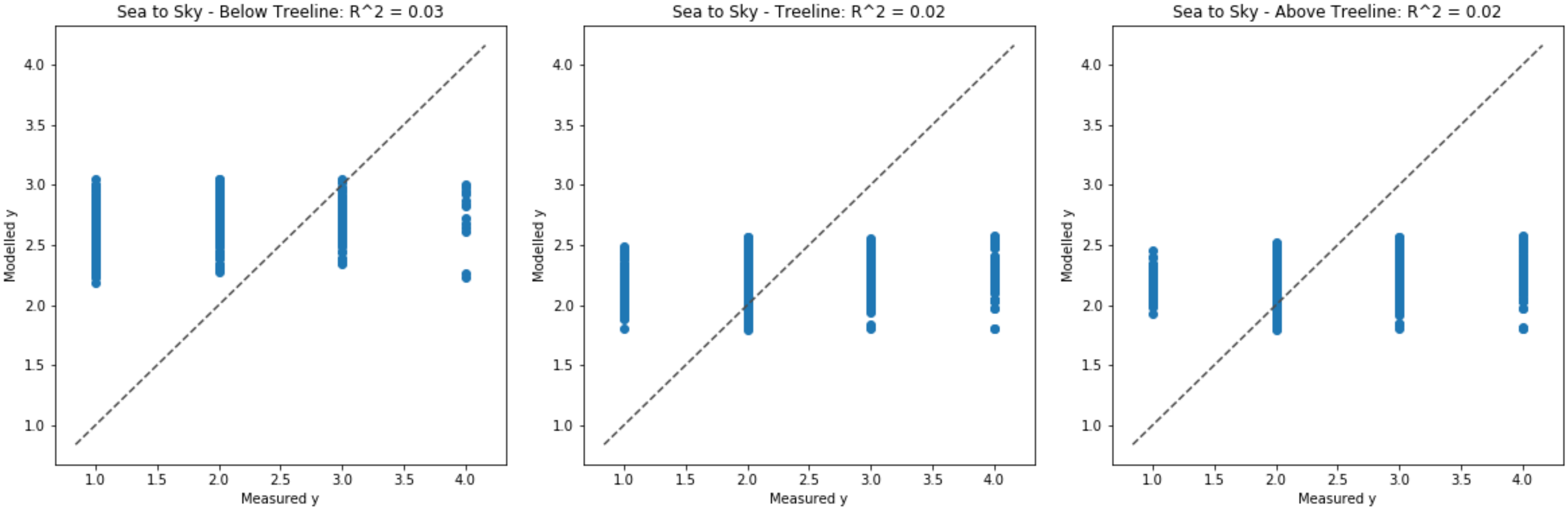


Figure 1. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the Sea to Sky region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: Lizard Range

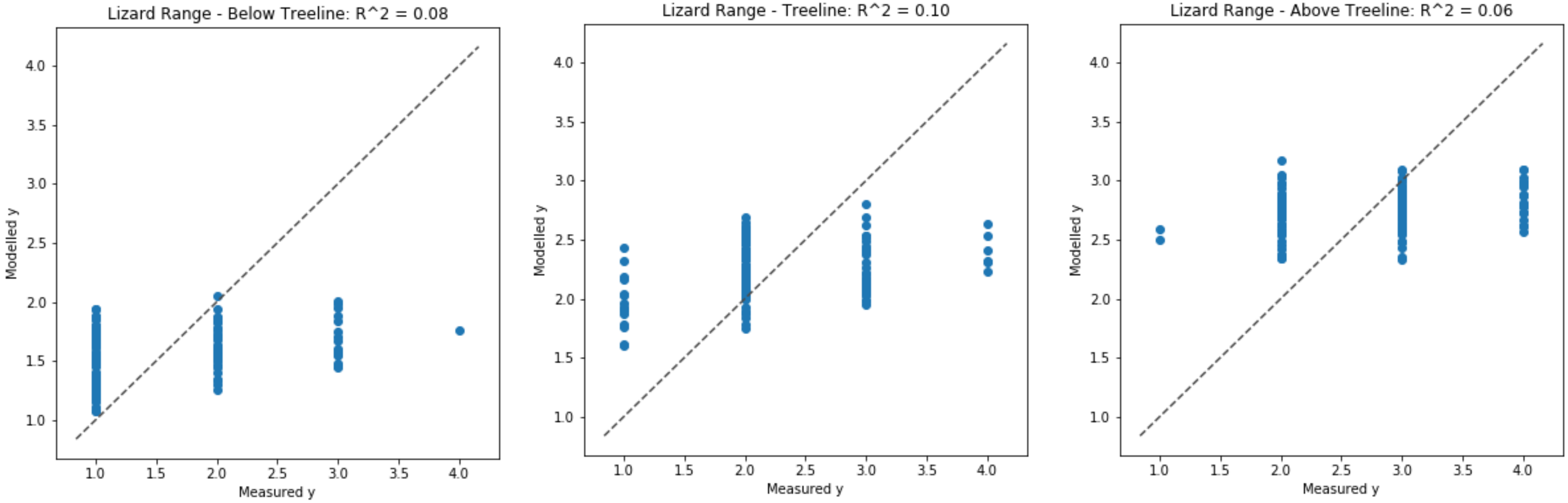


Figure 2. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the Lizard Range region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: North Columbia

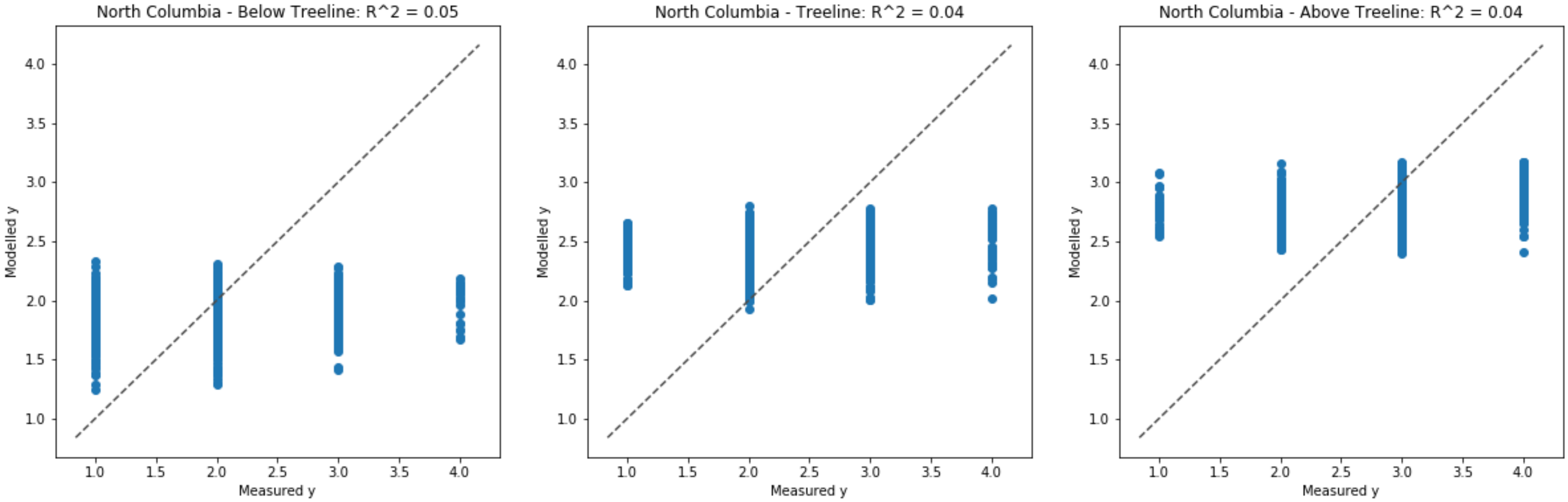


Figure 3. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the North Columbia region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: Kananaskis

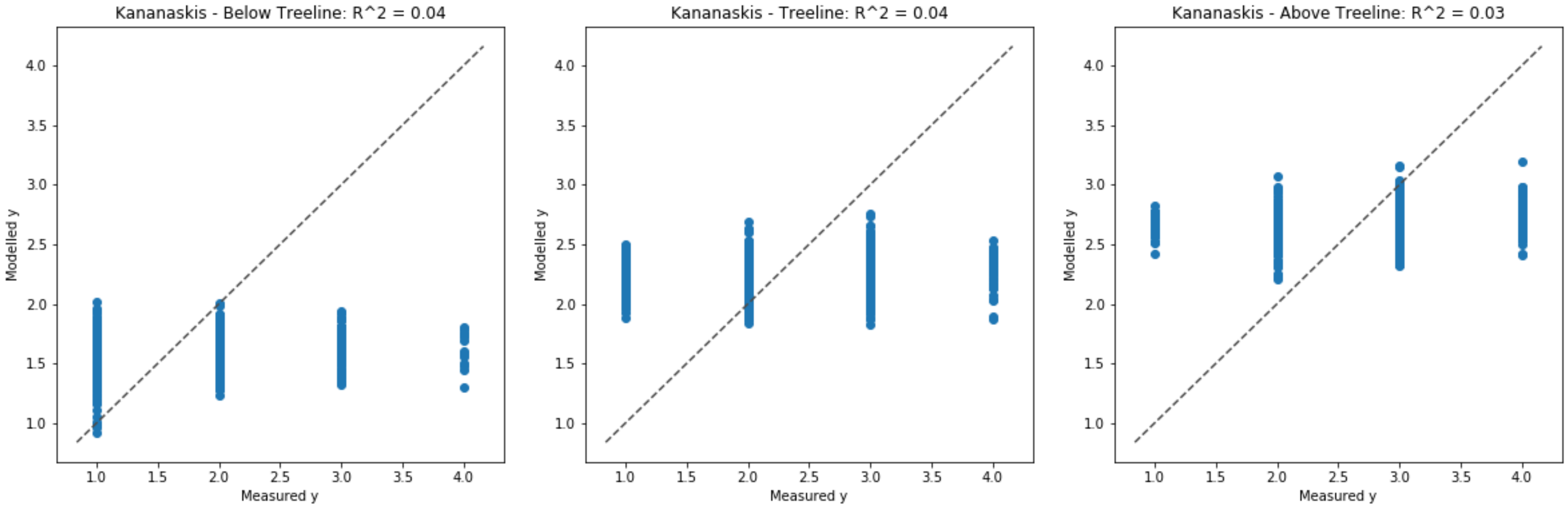


Figure 4. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the Kananaskis region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: Kootenay Boundary

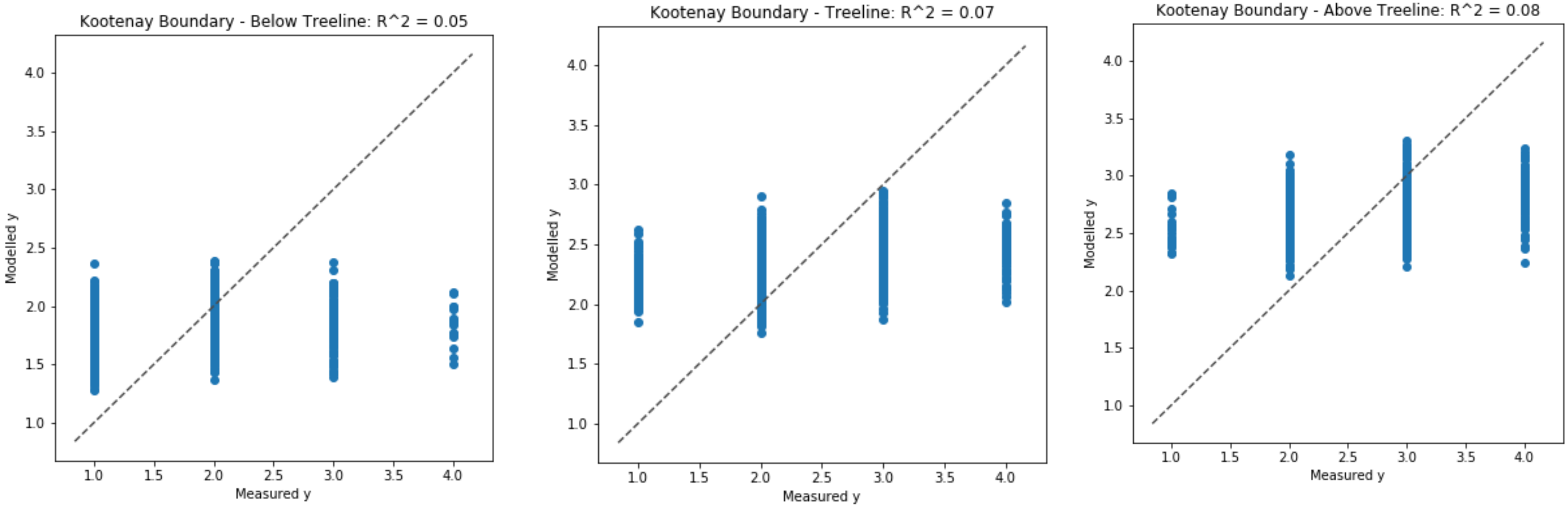


Figure 5. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the Kootenay Boundary region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: Cariboo

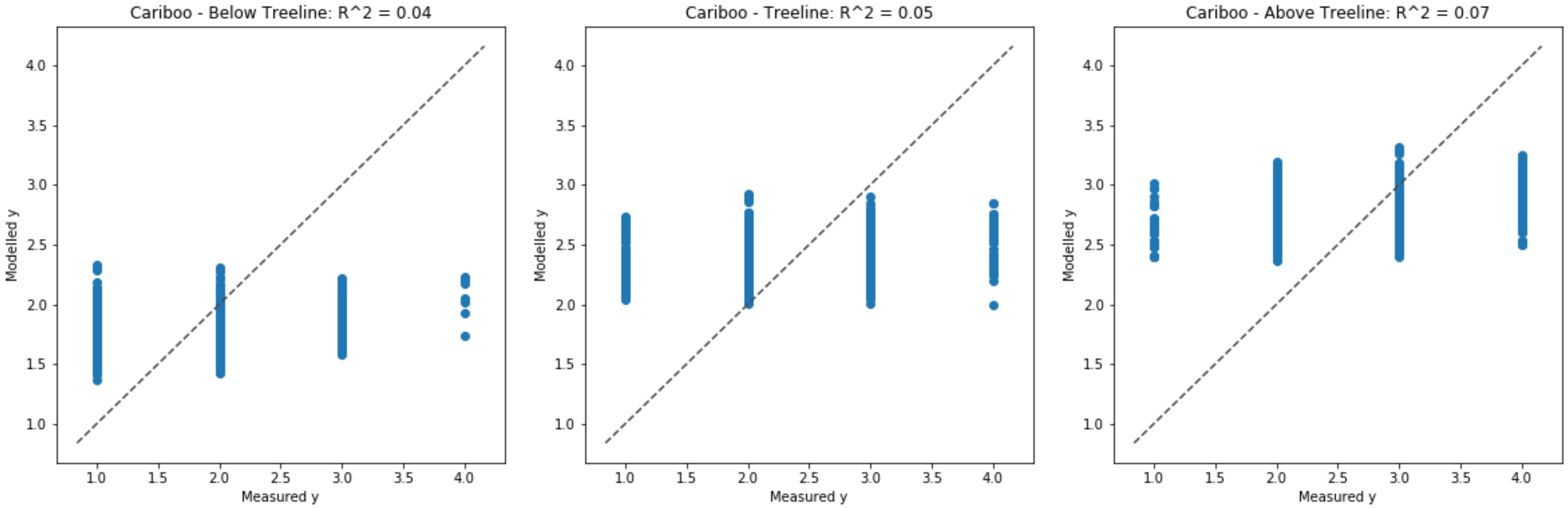


Figure 6. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the Cariboo region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: Northwest Coastal

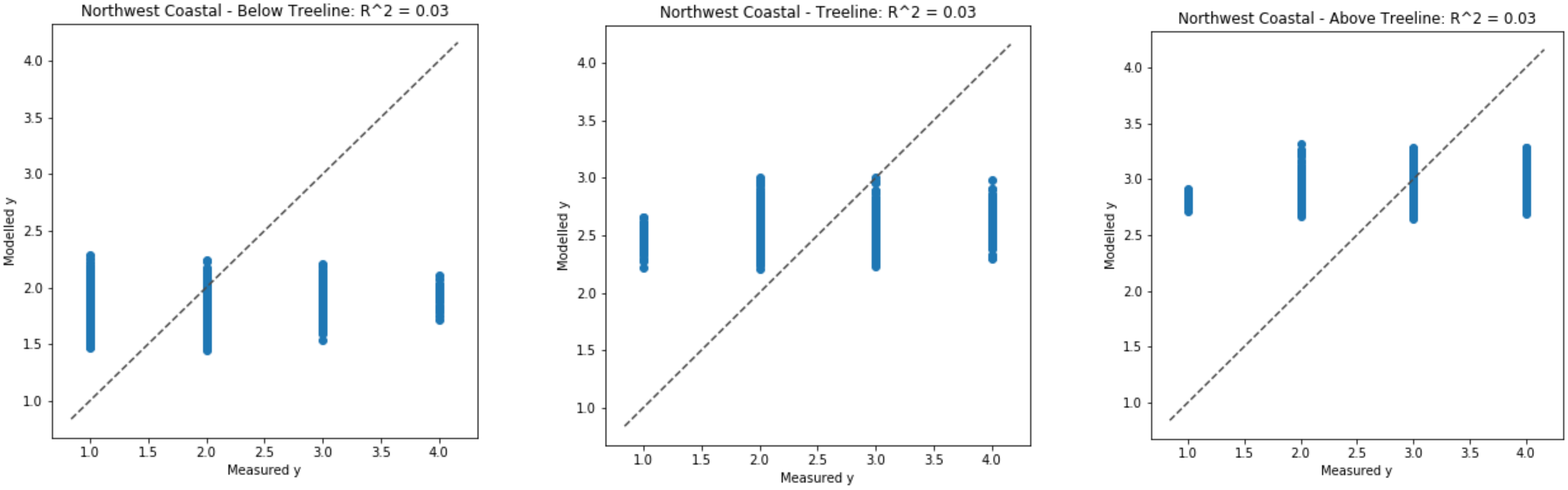


Figure 7. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the Northwest Coastal region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: Northwest Inland

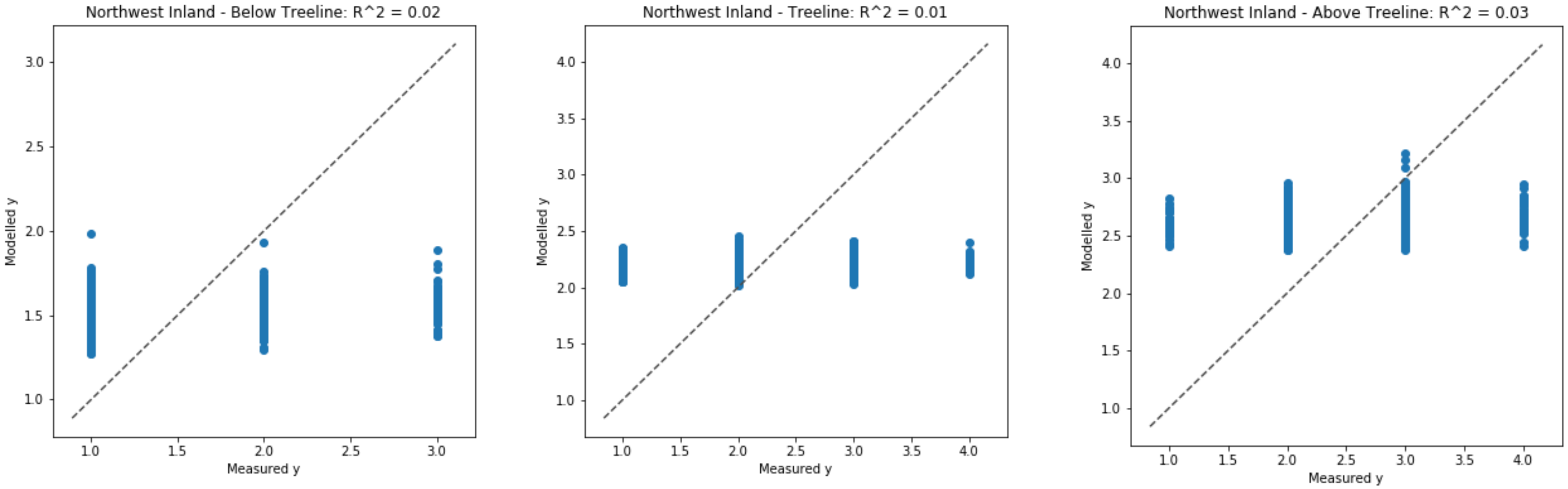


Figure 8. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the Northwest Inland region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: Purcells

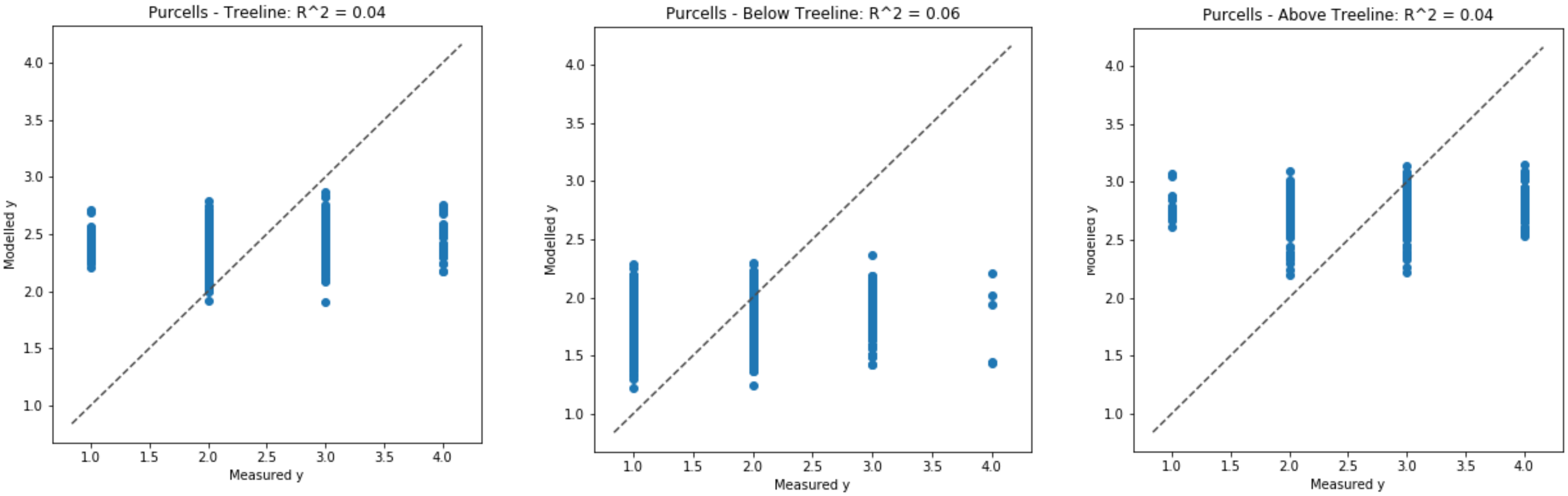


Figure 9. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the Purcells region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: South Coast

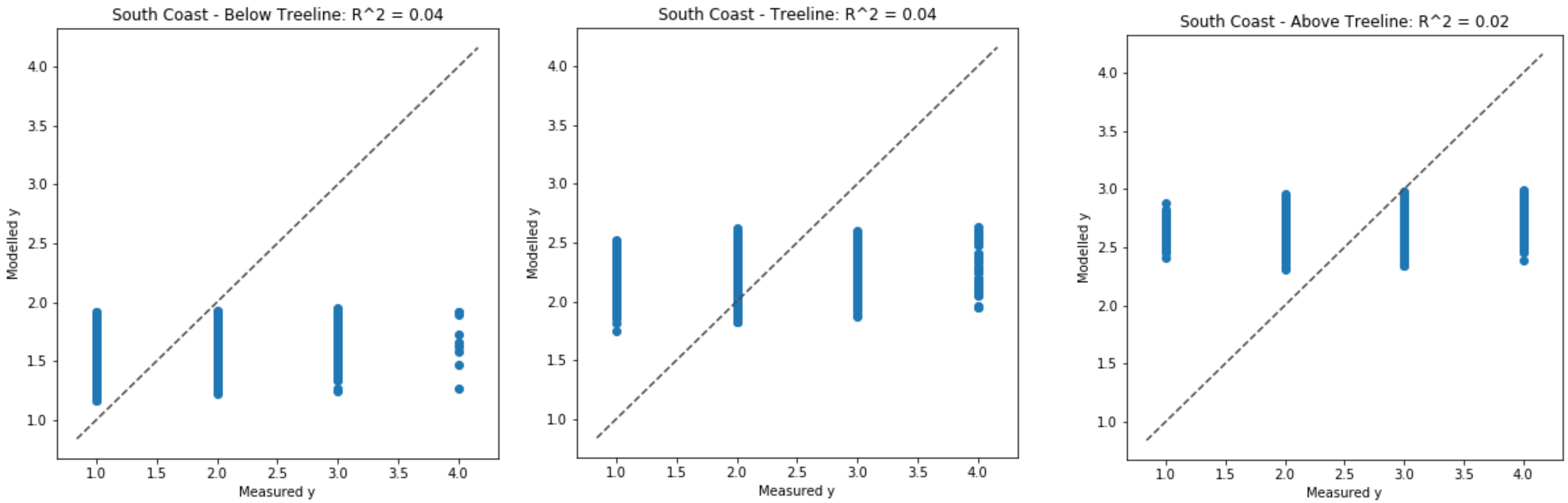


Figure 10. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the South Coast region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: South Columbia

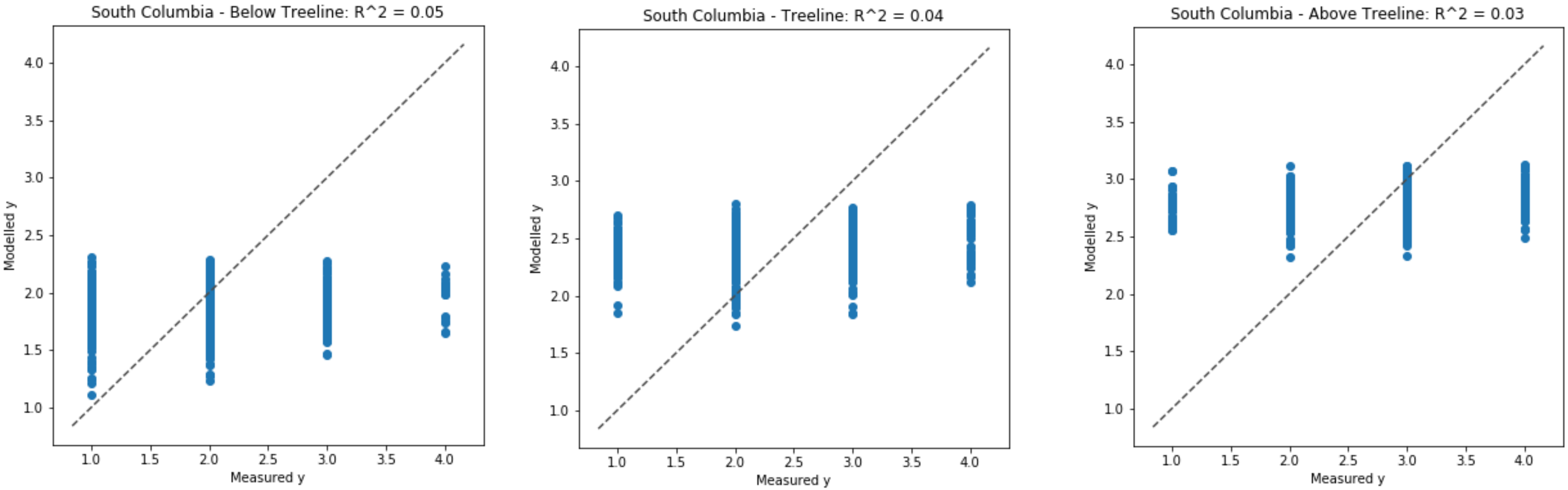


Figure 11. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the South Columbia region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Results: South Rockies

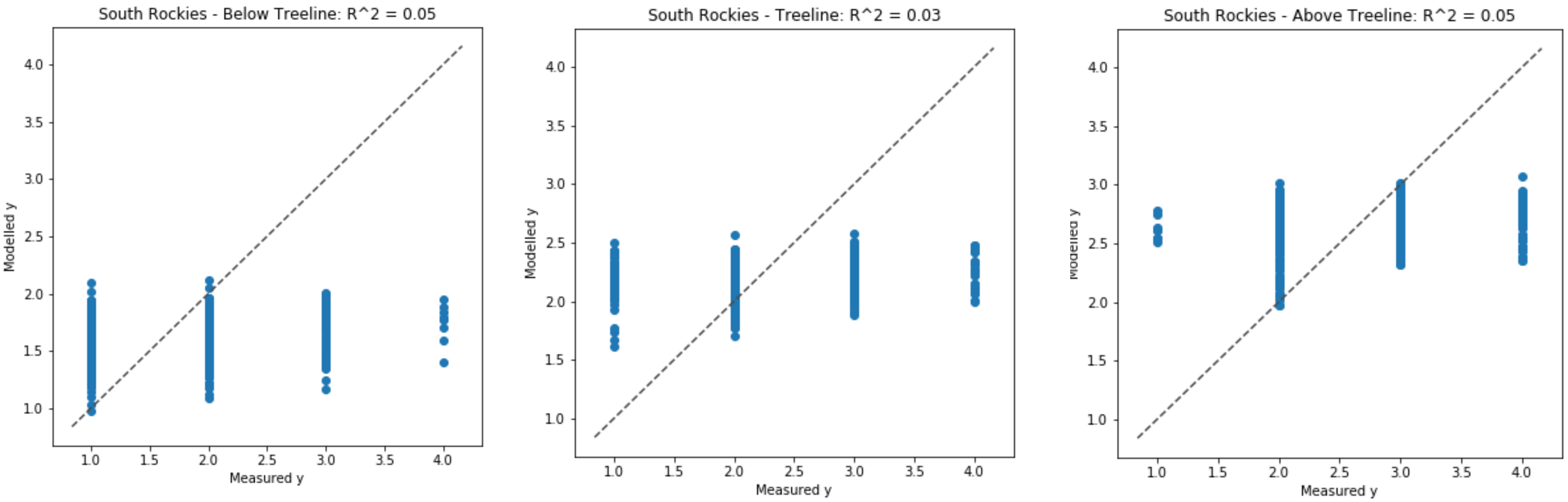


Figure 12. Scatter plot of MLR model performance between ERA5 data (t2m, u10, v10, msl, tp) and avalanche risk data for below treeline, treeline and above treeline in the South Rockies region. The low R^2 indicates that there is no significant correlation between the ERA5 data and avalanche risk data.

Discussion

(1) Which climate predictor(s) is/are the most important predictor(s) of avalanche risk for each elevation level in the Sea to Sky Region?

(2) Do similar important climatic predictors of avalanche risk exist between different mountain regions for the three different elevation levels (Below Treeline, Treeline, Alpine)?

- The low R values for below treeline, treeline and above treeline indicate that there is no significant correlation between the chosen climate predictors and avalanche danger ratings for all the mountain regions.

- In other words, 2-metre air temperature, 10-metre U wind component, 10-metre V wind component, mean sea level pressure, total precipitation are not significant predictors of avalanche danger ratings for all the mountain regions.

- Seeing that surface temperature, wind and total precipitation significantly affect avalanche danger in reality it is unlikely that the chosen climate predictors truly have minimal correlation with avalanche danger ratings. The low levels of correlation may be due the resolution of the data, the chosen statistical test (multiple linear regression), or errors in the code.

Discussion continued..

- The resolution of the ERA5 climate data has a 31 km horizontal resolution. Consequently, it is possible that the resolution of the ERA5 climate data is too low to accurately predict avalanche danger ratings within the relatively small mountain ranges.

- Multiple linear regression assumes that the independent variables are not highly correlated with each other. That being said, it likely that all the chosen climate predictors are highly correlated with each other. This might explain why the R^2 values did not show any significant correlation.

- Errors in the code, such as incorrectly defining the lat/lon dimensions of the mountain ranges may have caused the MLR test to show now significance. Conversely, it is possible that the incorrect dates in the netCDF were chosen.

- Avalanche risk is calculated based on a variety of factors affecting snowpack stability in a given region, which may not respond instantaneously. Given that a snowpack will stabilize / become unstable as a result of *cumulative* effects, daily climate variables may not be a proper indicator of stability.

Conclusion

- Multiple linear regression showed no significant correlation between 2-metre air temperature, 10-metre U wind component, 10-metre V wind component, mean sea level pressure, total precipitation and the avalanche danger ratings for all the different mountain ranges. Alternative statistical analysis should be used and/or alternative climatological data should be used as the chosen climate predictors are important for predicting avalanche risk in reality.

- For future analysis:

  1. Try additional variables, such as incoming radiation or take the vector of U wind and V wind.
  2. Adjust the temporal scale in of our model to look at weekly or even monthly trends as opposed to daily observations.
  3. Adjust the spatial scale of our model to encompass smaller areas within each region, this might help limit the spatial variability we see within the Climate Re-analysis data.

References

Avalanche Canada Website (2020). Retrieved March, 2020 from https://www.avalanche.ca/

Avalanche Canada. (2020). October 14, 2019. *Avalanche Canada 2019 Annual Report*. Retrieved March, 2020. from https://issuu.com/avalancheca/docs/ac_2019_annual_reportissuu

McClung, D., & Schaerer, P. A. (2006). *The avalanche handbook*. The Mountaineers Books.

Jamieson, B., Geldsetzer, T. (1996). *Avalanche Accidents in Canada, Volume 4: 1984-1996.* Canadians Avalanche Association – National Research Council of Canada. Retrieved March, 2020, from https://cdn.ymaws.com/www.avalancheassociation.ca/resource/resmgr/files/avalancheaccidentsv4.pdf

Parks Canada. April 04, 2017. Mountain Safety: *Highway avalanche control program.* Retrieved March, 2020, from https://www.pc.gc.ca/en/pn-np/mtn/securiteenmontagne-mountainsafety/avalanche/routes-highways

Stethem, C., Jamieson, B., Schaerer, P., Liverman, D., Germain, D., & Walker, S. (2003). Snow avalanche hazard in Canada–a review. Natural Hazards, 28(2-3), 487-515.