Biogeoclimatic Projections for British Columbia

C. Mahony*, D. Obrist*, W. MacKenzie, and K. Daust

*Equal first authors

Version 1.0

March 31, 2025

Executive Summary

Biogeoclimatic projections are a technique in which recent or projected future climate conditions are classified in terms of the climate units of the Biogeoclimatic Ecosystem Classification (BEC) system and spatially mapped across the province. They are a useful approach to anticipating how ecosystems will respond to climate change, and are becoming widely used in forestry, conservation, and land use planning.

In the Climate Change Inferred Species Selection (CCISS) tool, we use machine learning models trained on historical climate data to classify future climates based on their closest analog within the Biogeoclimatic Ecosystem Classification (BEC) system, which spans Western North America. We consider a diversity of possible future climates, according to various greenhouse gas emissions scenarios, natural variability in the climate, and uncertainty in our predictions.

In this report, we describe the methods for producing biogeoclimatic projections for British Columbia. Specifically, we detail the process of fitting a Random Forest model, including selecting points to train the model, climate variables, and model hyperparameters. We describe cases where no good analog is available, which we classify as novel climates. Finally, we discuss current known issues and provide guidance to end users of our data products.

The data products produced here are available through the spatial module of the CCISS tool, which allows users to examine changes at both local and provincial scales, and to download biogeoclimatic projections in raster format for five 20-year periods of the 21st century.

Acknowledgements

We gratefully acknowledge Dr. Tongli Wang (University of British Columbia) and Dr. Andreas Hamann (University of Alberta) for their groundbreaking research on biogeoclimatic projections and its applications in resource management.

We acknowledge the World Climate Research Program's Working Group on Coupled Modelling, which is responsible for the Coupled Model Intercomparison Project (CMIP6), and we thank the global climate climate modelling groups for producing and making available their model output.

Table of Contents

Introduction

The CCISS tool uses spatial climatic analogs to make predictions about the environmental suitability of tree species across British Columbia. A spatial climate analog is a location with a historical climate that is similar to the current or future projected climate of a different location. Biogeoclimatic (BGC) subzones and their variants are a uniquely useful set of spatial climate analogs because they are familiar to resource management practitioners and are the organizing units for site-specific ecological interpretations accumulated over many decades.

This report describes the development of biogeoclimatic projections for British Columbia. Biogeoclimatic analogs are identified by training a machine learning model to recognize BGC subzone-variants in terms of their historical (1961-1990) climatic conditions and then applying that classification model to new (current or projected) climate conditions across the province (MacKenzie and Mahony 2021). The future climates are thus matched to the closest analog of historical climates within the BEC system, a process called BGC projection.

Methods

Biogeoclimatic Ecosystem Classification (BEC)

The Biogeoclimatic Ecosystem Classification (BEC) system is a framework used in British Columbia (BC) to categorize ecosystems based on climate, vegetation, and site conditions. Within BEC, biogeoclimatic (BGC) zones define climatic regions based on mature plant associations in sites which are sites best representative of regional climate regimes. Subzones and variants of BGC zones capture finer-scale climate and vegetation differences, while site series classify local soil moisture and nutrient conditions. Our BGC projections take place at the level of BGC subzones and their variants, and we will refer to these as 'BGC units' or 'BGC subzone/variants' throughout this document.

Because BGC projections rely on climatic classification to model future ecosystem shifts, identifying climate analogs—both within and outside of BC—is essential. Some future conditions expected in BC already exist elsewhere, providing important insights into potential ecological change. We used the Ecological Classification of Alberta (Archibald et al. 1996, Beckingham and Archibald 1996, Beckingham et al. 1996) to incorporate Albertan climate analogs, and a draft BGC ecosystem classification system for the western US (Washington, Idaho, Montana, Oregon, northern California, and northwestern Wyoming) developed by Del Meidinger and Will MacKenzie (2024). Incorporating these climate analogs allows us to predict how BC's forests may change and what species may thrive there in multiple future climate scenarios.

Establishing training points

We completed the following steps to establish training points by using the function bgc_trainingSample() from the '*ccissr*' package in R (MacKenzie et al. 2024). For more information about this function, see **Appendix A1**.

Sampling schemes

We tested several different methods for establishing the points upon which to train our Random Forest model to produce the BGC projections (**Appendix A2**). Specifically, we assessed trade-offs between oversampling large BGC units and undersampling small BGC units and considered different balancing schemes. In the end, we opted to use a sampling scheme that assigned training points to BGC units proportionally to the square root of the area of each BGC unit. This method ensures that smaller BGC units receive relatively more points proportional to their total size, which prevents overly small units from being underrepresented in predictions. We scaled the number of points by the interquartile range (**IQR)** of mean annual temperature (**MAT**), which resulted in BGC units with greater spatial climatic variation receiving more training points (**Error! Reference source not found.**).

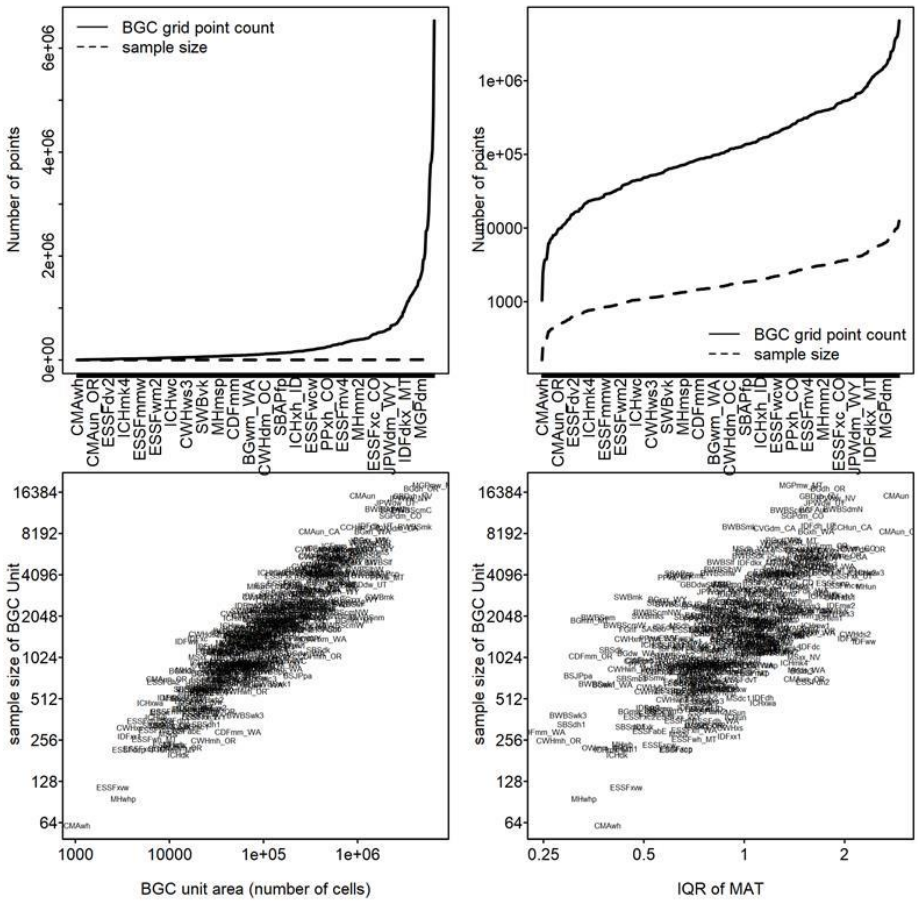

Figure 1. Diagnostic plots for sampling scheme **V2**. Top panels show the number of possible points (i.e., pixels) by BGC subzone/variant, organized by smallest to largest BGC unit area on linear (**top left**) and logarithmic scale (**top right**), represented by the solid black line. The dashed line represents the number of training points (i.e., sample size). Labels on the x-axis are a subset of the 300+ BGC units represented in the plot. Bottom panels show sample size selected by each BGC unit area (**bottom left**) and based on the inter-quartile range (IQR) of mean annual temperature (MAT) (**bottom right**).

Outlier removal and other filtering

From the established set of training points for each BGC subzone/variant, we removed points (i.e., sampled locations) where 1961-1990 climate normals of seasonal precipitation (**PPT**), annual mean daily maximum temperature (**Tmax**), and annual mean daily minimum temperature (**Tmin**) fell outside the 3-sigma range ($\alpha = 0.0027$) for points in BC and the 2-sigma range ($\alpha = 0.05$) for points located outside of BC. The rationale for this was to replicate BGC unit boundaries within BC, but only the BGC core concept outside of the province, as we retained more climatic variation within the training set for BC.

Climate variable selection

We considered and compared three different climate variable sets for our Random Forest model, all of which are detailed in **Appendix A3**. For this version of CCISS, we used 23 climate

variables referred to in this report as the ‘expert variable set’ (Table A1). All climate data were accessed through the downscale() and addVars() functions in the ‘*climr*’ package in R (Daust et al. 2024).

At the provincial scale, many variables effectively distinguish between coastal and interior climates, which is useful for broad-scale differentiation. However, these same variables can be less effective in distinguishing nearby BGC units from one another. Thus, to ensure that the retained variables were relevant for both local and regional differentiation, we combined expert knowledge with a local feature selection approach. Specifically, we pre-selected a set of ecologically relevant seasonal and annual climate variables, in addition to several derived variables and removed any highly correlated variables. We then created a list that identifies the adjacent BGC units for each individual unit. To determine the most important local climate variables, we ran a separate Random Forest model with 101 trees for each BGC unit, considering only the climate space of that BGC unit and its neighbouring BGC units. These Random Forest models used the “Gini” split rule, and variable importance was calculated based on Gini impurity. Once a list of variable importance was created for each BGC unit, we retained the most important variables from all to retain the ability to differentiate between BGC units at a local scale.

In summary, the final expert variable set was chosen based on ecological relevance, low correlation, and high predictive importance within each BGC unit. For further details, also see **Appendix A3**.

Random Forest model

To classify BGC units, we trained a Random Forest machine learning model using the ‘*ranger*’ package in R (Wright and Ziegler 2017). This model was built using 500 decision trees, the extraTrees split rule, and a minimum node size of two, with training points selected according to the sampling scheme outlined above. We used the ‘expert variable set’ of climate variables (Table A1). This model underwent validation and comparison against alternative approaches, including multiple sensitivity analyses to assess robustness, accuracy, and sensitivity to training selection schemes and various model hyperparameters (**Appendix A4**).

Global climate model ensemble

The biogeoclimatic projections incorporate three types of climate change uncertainty: modeling uncertainty, natural variability, and socioeconomic uncertainty. These uncertainties are represented by modeling biogeoclimatic projections for an ensemble of 60 potential future climate states (8 climate models x 1-3 simulation runs x 3 socioeconomic scenarios) for each future time-period. We provide biogeoclimatic projections for five global climate model simulations that represent the diversity of patterns and trends in climate change across the full ensemble of 60 simulations.

Since the 60 simulations typically do not converge on a single BGC analog, CCISS displays an *ensemble vote winner*—that is, the BGC analog most frequently selected as the convergence result across all simulations, with each independent simulation effectively casting a "vote" for a given BGC analog. We provide spatial data of these ensemble vote winner analogs at both the subzone/variant and zone levels. For the zone level map, we aggregate the results by grouping all votes for subzone/variants under their respective zones. We then sum these grouped votes to determine the zone with the highest number of votes. For example, consider a location within the IDFdk3—the Interior Douglas-fir zone, the dry, cool subzone, and Cariboo variant (Figure 2). CCISS shows that in the 2061-2080 time period, IDFmw2 received 51% of votes, IDFxh2 received 16.3%, IDFdm2 received 11.5%, PPxw_MT received 8.7%, ICHxm1 received 4.8%, and SBSmh and IDFxx2 each received 3.8%. At the subzone/variant level, IDFmw2 is displayed, because it was the result in the greatest percentage of simulations (i.e., got the most votes). To calculate the zone-level winner, we would sum up the votes by zone: IDF would win with 82.6% of votes, while PP would get 8.7%, ICH would get 4.8%, and SBS would get 3.8%.

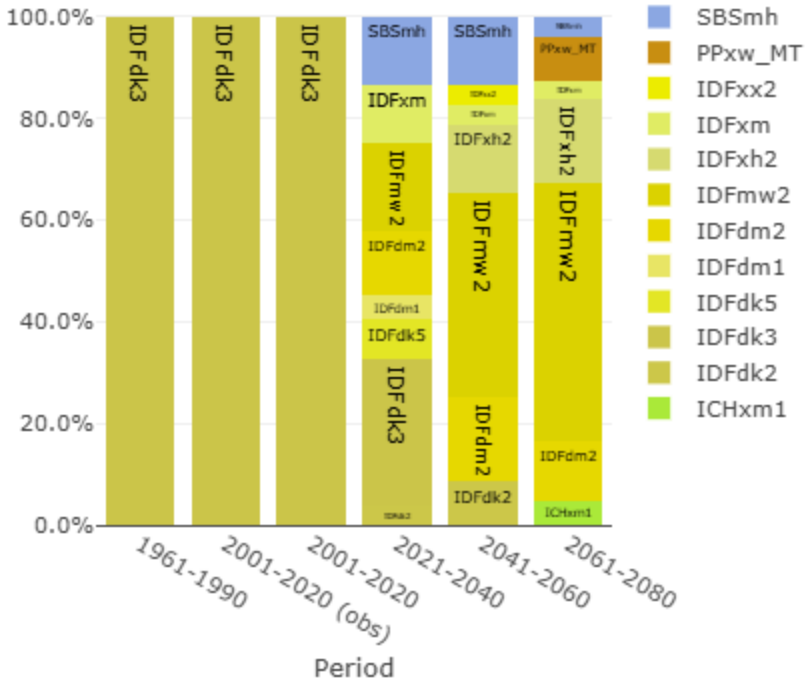

Figure 2. Screenshot from CCISS Spatial showing the distribution of simulation convergence results (i.e., votes) for a location in the IDFdk3 (Interior Douglas-fir zone, dry, cool subzone, in the Cariboo region).

Further details on the climate model ensembles are provided in Appendix D.

Novel Climate Detection

In the CCISS framework, the pool of historical climate analogs consists of BGC subzone-variants from British Columbia, Alberta and the Western United States. For each projected future climate, the BGC model chooses the most similar historical climate analog. However, the climates of western North America do not necessarily represent all potential future climate conditions. There may be substantial differences between the future climate condition and the climate conditions of the BGC analog. In such cases, the future conditions are considered a 'novel climate'. As the BGC classification model simply assigns the closest climate analog in climate space, the basic problem of novel climates (i.e. locations with poor analogs) is that they may give misleading results but are not explicitly identified by the BGC model. Detecting novel climates requires an additional step in the CCISS analysis.

CCISS measures climatic novelty relative to the climatic variation of the analog. In each BGC projection, we calculate a novelty measurement for each BGC subzone/variant that is being used as an analog. We sample the baseline (1961-1990) climates found across the current geographical range of each BGC subzone/variant, and combine this spatial variation with the 1951-1990 temporal (interannual) variation of the analog. This pooled spatiotemporal variation is then used to measure the difference between the climate of the analog and the future climates that are assigned that analog in the BGC projection. To make this measurement, we use a metric called sigma dissimilarity (Mahony et al. 2017).

Novelty detection for machine learning models is non-trivial and imprecise. The novelty metric provides an indication of where CCISS results may be unreliable or require additional expert input. However, it likely underestimates novelty in some locations and overestimates it in others. Users are encouraged to look at the broader pattern of climatic novelty around their query locations to assess whether their local novelty results are representative. Over time, the CCISS team will develop region-specific interpretations of the drivers of climatic novelty and their ecological implications.

Appendix B describes the CCISS novelty detection methods and results in further detail, outlining a coastal case study and describing results at the provincial scale. The goal of this appendix is to provide users with the ability to interpret novelty results with an understanding of the strengths and limitations of the method.

Biogeoclimatic projections data

The data produced through the BGC projections can be accessed interactively via the spatial module of the CCISS tool. This module allows users to view projections at both local and provincial scales and to download raster data for predefined subregions. BGC projections will be available for the following five 20-year periods of the 21st century:

- Ensemble vote winner BGC subzone/variant from an ensemble of 60 global climate model projections (5 rasters)
- Ensemble vote winner BGC zone (5 rasters)
- BGC projections for 5 global climate model simulations that represent the variation in the 60-member ensemble (25 rasters)
- BGC projections for observed climates of the 1961-1990 and 2001-2020 periods (2 rasters)
- Estimated climatic novelty for all BGC projections (37 rasters)

Known Issues

This version of the BGC projections is a work in progress, and future refinements will be made. As more end users engage with the CCISS tool, additional issues may come to light. Users can report any issues or discrepancies in the CCISS_Review GitHub repository.

We have already identified several areas for improvement. One key issue is that the baseline BEC maps are not being perfectly reconstructed, resulting in fuzzy boundaries, and the over- and underrepresentation of certain BGC units. For example, the Coastal Mountain-heather Alpine (CMA) zone in the Pemberton region appears to be consistently underrepresented.

Additionally, future projections appear highly sensitive to changes in sampling scheme (**Appendix A4**), climate variable sets, and Random Forest model hyperparameters. These factors will be closely examined in future iterations to improve the model's predictive ability.

Guidance and Caveats

Evaluating BGC future projections

An important consideration when evaluating the quality of any BGC projection is that, ideally, it should be compared to the baseline prediction of the same model, rather than to the original BEC linework. In some cases, the BEC linework and climate data informing the baseline model prediction might not align. A mismatch between the model and linework does not necessarily indicate an error in the BGC modeling. Discrepancies between BGC mapping and baseline BGC projections can, in some cases, indicate errors in the BGC mapping or errors in the baseline climate mapping, and should therefore be interpreted carefully on case-by-case basis rather than assumed to be BGC modeling errors.

Interpreting spatial shifts in BGC projections

Although the visual effect of BGC projections is of BGC zones and subzone/variants shifting across the map, these spatial shifts should not be taken literally. No analog is perfect, and projected analogs may be highly imperfect for several reasons. The actual future climate at any location will be a hybrid of (1) the characteristics of the analog climate, (2) novel climatic characteristics (e.g., extremes) that are not represented by the analog, and (3) enduring features of the local climate such as frost pooling, lake effects, and wind patterns. The estimated BGC classifications from any location and time period require careful consideration by the end users of CCISS products.

The role of BGC mapping in a changing climate

The misinterpretation of BGC projections as literal spatial shifts in BGC units has led to a common perception that climate change is rendering BGC mapping obsolete. This is not the case. The linework of BGC subzone/variants in many cases will remain useful as units of relative climatic variation across landscapes. The terminology we use for BGC projections can help to emphasize that BGC analogs are only approximations and that the BGC units themselves are not undergoing spatial shifts. Rather than saying “this location is becoming IDFxh1”, it is more correct to say “the future climate at this location is predicted to be similar to the historical climate of the IDFxh1.” Rather than “the IDF is moving north into the SBS”, it is better to say “the SBS is transitioning into more IDF-like climates”.

Next Steps

Most of our next steps involve further validation of the model we used here for BGC projection. In summary, these include (1) trying different statistical approaches beyond RF, (2) using different split rules within the Random Forest framework, (3) validating the current model by testing its ability to predict to 'gaps' (i.e., areas left out of the training data), (4) adding further climate variables, and (5) applying or developing novel metrics for model assessment. Further explanations on these steps are in **Appendix A5**.

Literature Cited

Archibald, J. H., G. D. Klappstein, and I. G. W. Corns. 1996. Field guide to ecosites of southwestern Alberta. Special Report 8. Natural Resources Canada, Canadian Forest Service, Northwest Region, Northern Forestry Centre, Edmonton, Alberta.

Beckingham, J. D., and J. H. Archibald. 1996. Field guide to ecosites of northern Alberta. Northern Forestry Centre, Edmonton.

Beckingham, J. D., I. G. W. Corns, and J. H. Archibald. 1996. Field guide to ecosites of west-central Alberta. Northern Forestry Centre, Edmonton.

Daust, K., C. R. Mahony, B. Tremblay, and C. Barros. 2024. climr: Downscaling climate data in R.

MacKenzie, W. H., K. Daust, C. Barros, and C. R. Mahony. 2024. ccissr: B.C. Climate Change Informed Tree Species Selection.

MacKenzie, W. H., and C. R. Mahony. 2021. An ecological approach to climate change-informed tree species selection for reforestation. Forest Ecology and Management 481:118705.

MacKenzie, W. H., and D. Meidinger. 2024. Biogeoclimatic Zones and Subzones of the Western United States. BC Data Catalogue. https://catalogue.data.gov.bc.ca/dataset/cciss-western-north-america-bec-tables

Mahony, C.R., A.J. Cannon, T. Wang, and S.N. Aitken. 2017. A closer look at novel climates: new methods and insights at continental to landscape scales. Global Change Biology 23: 3934–3955.

Wright, M. N., and A. Ziegler. 2017. ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software 77.

Appendix A. BGC Projection Methods

A1. bgc_trainingSample() function details

The bgc_trainingSample() function from the '*ccissr*' package (MacKenzie et al. 2024) provides a quick, reproducible, and easily customizable way to extract training points. It follows these steps:

1.) Extracts point data from the DEM and BGC rasters
2.) Optionally removes specified BGC units.
3.) Balances sample size across BGC units using one of two subsampling scheme, asymptotic or square-root (described below)
4.) Optionally adjusts sample size based on climatic variance
5.) Optionally removes outliers from the sample based on climate variables
6.) Generates diagnostic plots (e.g., Figures 1, A1, A2, and A3).

A2. Sampling schemes

We considered several different methods for establishing the points upon which to train the Random Forest model to produce the BGC projections. Specifically, we assessed trade-offs between spatial clustering but numerically even distributions of sampling points across BGCs, versus an even spatial distribution but a numerical imbalance due to the inherent variability in BGC area. Thus, we trained the Random Forest model with training sample data derived from asymptotic and square root balancing schemes for training point sample size. In all cases, we define the population size (**N**) as the total number of grid cells in the BEC raster belonging to each BGC unit.

Asymptotic:

V1 and V3 of our sampling scheme take an asymptotic approach. In short, V1 assigns training points to BGC units proportionally to the dispersion (interquartile range, IQR) of mean annual temperature (MAT), while V3 represents a roughly even distribution of 2000 samples per BGC unit.

To achieve this, we used an exponential subsampling function that limits the number of points per BGC unit while ensuring that larger BGC units receive proportionally more training points. Specifically, we defined an asymptote of 2000, and a threshold at which subsampling begins, which we calculated as:

$$threshold = asymptote / 20$$

For BGCs with population sizes below this threshold, values remained unchanged. For larger populations, the exponential function was applied to control the number of assigned points and

prevent excessive allocation. The transition rate, which controls how quickly subsampled values reach the asymptote, was controlled by a shape parameter, which we defined as:

shape = 1/asymptote

This resulted in the following formula, used to determine the number of sampling points per BGC:

$$n = threshold + (asymptote - threshold) \times (1 - e^{-shape \times (N - threshold)}),$$

where **n** is the number of subsampled points, and **N** is the original population size (i.e., the total number of grid cells in the BEC raster belonging to a BGC).

This approach resulted in most BGCs receiving between 1600 and 2000 points, aside from CMAwh (~800), and ESSFxvw (~1400, Figure A2, bottom right). This method yielded our sampling scheme V3, which represents a scenario where most BGC units received approximately 2000 sampling points.

For V1, we scaled the number of assigned points by the IQR of MAT, effectively weighting the sample distribution by climatic variability. The relationship between the IQR of MAT and the number of training points selected per BGC unit was linear (Figure A1, bottom right).

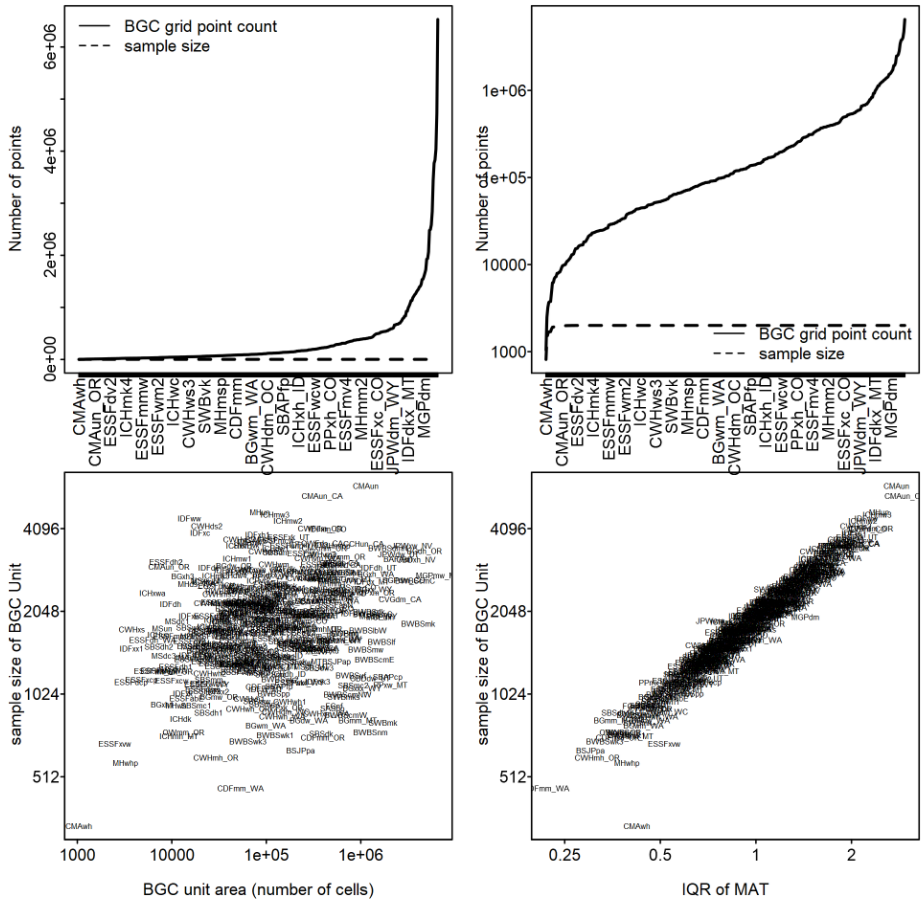

Figure A1. Diagnostic plots for sampling scheme **V1**. Top panels show the number of possible points (i.e., pixels) by BGC subzone/variant, organized by smallest to largest BGC unit area on linear (**top left**) and logarithmic scale (**top right**), represented by the solid black line. The dashed line represents the number of training points (i.e., sample size). Labels on the x-axis are a subset of the 300+ BGC units represented in the plot. Bottom panels show sample size selected by each BGC unit area (**bottom left**) and based on the IQR of MAT (**bottom right**).

Square root:

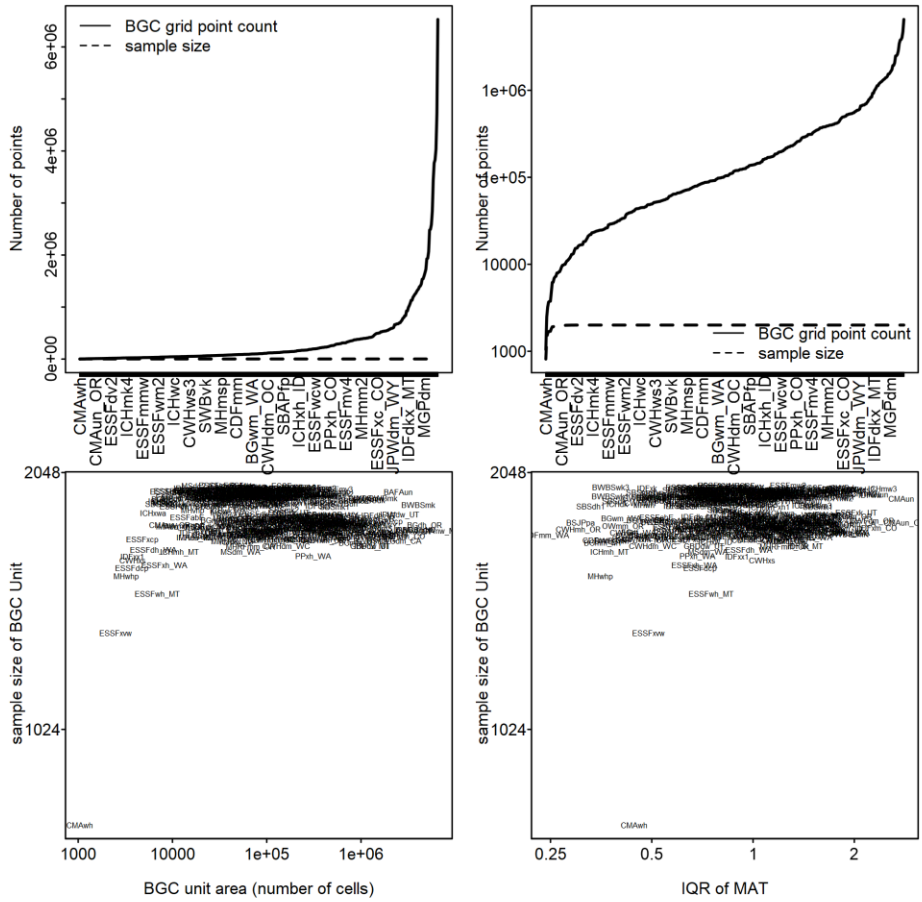

Figure A2. Diagnostic plots for sampling scheme **V3**. Top panels show the number of possible points (i.e., pixels) by BGC subzone/variant, organized by smallest to largest BGC unit area on linear (**top left**) and logarithmic scale (**top right**), represented by the solid black line. The dashed line represents the number of training points (i.e., sample size). Labels on the x-axis are a subset of the 300+ BGC units represented in the plot. Bottom panels show sample size selected by each BGC unit area (**top left**) and based on the IQR of MAT (**top right**). **Note:** In **V3**, we did not scale the sample size by IQR but here just demonstrate what the IQR looks like.

Our V2 and V4 sampling schemes take a square root approach to subsampling. In short, V2 assigns training points to BGC units proportionally to the square root of the area of each BGC, and V4 does the same, but selects twice the number of points.

Specifically, our sampling algorithm selected the number of points per BGC proportionally to the square root of the total number of grid cells belonging to each BGC:

$$n = \sqrt{N} * squareRoot.multiplier$$

This value was then multiplied by a scaling factor (squareRoot.multiplier)—in our case, we chose a multiplier of 5 for V2 and 10 for V4, the latter of which results in a 1:1 subsample where the original population size (N) is less than or equal to 100. This method ensures that smaller BGC units receive a relatively larger proportion of points compared to their total size, preventing overly small units from being underrepresented. As in V1, we scaled the number of points by the IQR of MAT, meaning that BGCs with greater climatic variability received more training points (Figure 1, Figure A3).

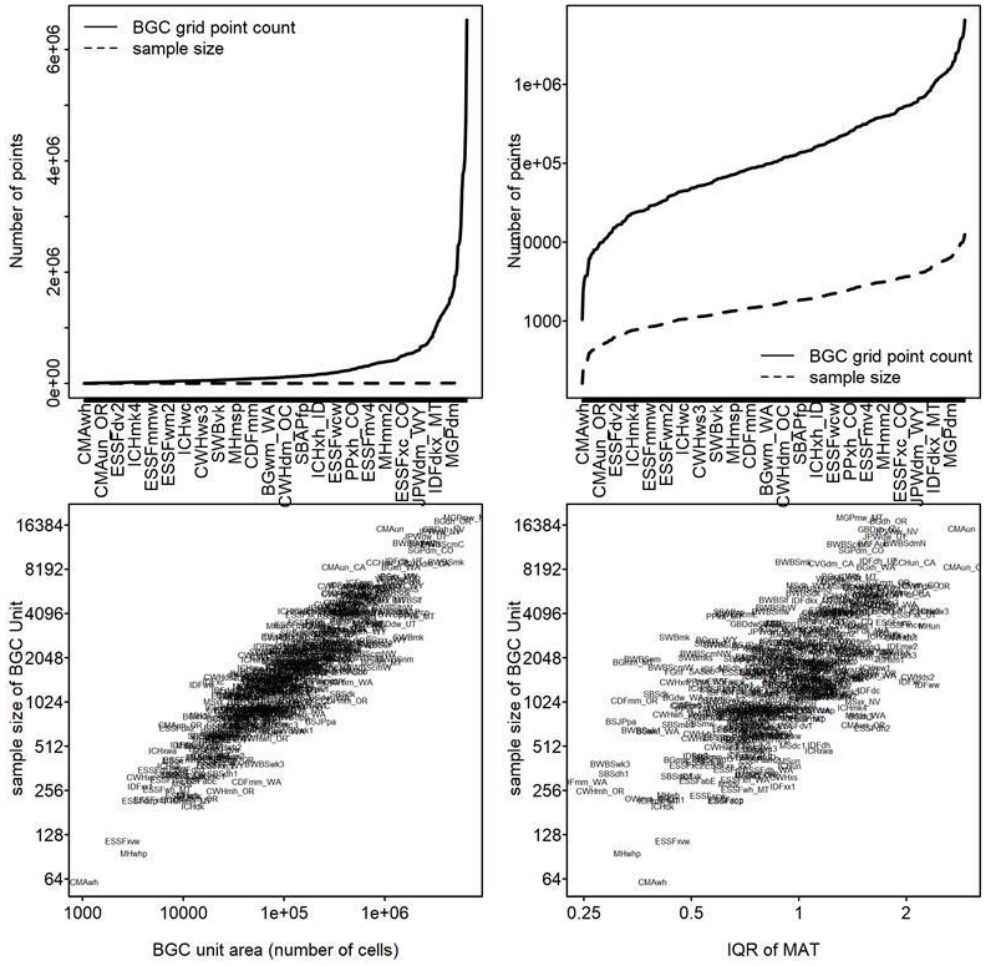

Figure A3. Diagnostic plots for sampling scheme **V4**. Top panels show the number of possible points (i.e., pixels) by BGC subzone/variant, organized by smallest to largest BGC unit area on linear (**top left**) and logarithmic scale (**top right**), represented by the solid black line. The dashed line represents the number of training points (i.e., sample size). Labels on the x-axis are a subset of the 300+ BGC units represented in the plot. Bottom panels show sample size selected by each BGC unit area (**bottom left**) and based on the IQR of MAT (**bottom right**).

A3. Climate variable sets

We considered and compared three different climate variable sets in building our Random Forest model. All climate data were accessed through the downscale() function in the '*climr*' package in R (Daust et al. 2024), except for a subset of derived variables, which were added using the

addVars() function. These variable sets include a ‘simple’ variable set, an ‘expert selection’ variable set, and an ‘all seasonal variables’ set.

Simple variable set:

This set included only precipitation (PPT), maximum temperature (Tmax), minimum temperature (Tmin) for each season: winter, spring, summer, and autumn.

Expert selection variable set:

At the provincial scale, many variables effectively distinguish between coastal and interior climates, which is useful for broad-scale differentiation. However, these same variables can be less effective in distinguishing nearby BGC units from one another. Thus, to ensure that the retained variables were relevant for both local and regional differentiation, we selected a suite of variables based on ecological relevance, low correlation, and high predictive importance within each BGC unit (Table A1). Specifically, to obtain variables for the expert variable set, we employed a local feature selection approach, where the Gini importance of a pre-selected set of seasonal and annual climate variables, in addition to a number of derived variables, was calculated separately for each BGC unit. This preliminary starting set was comprised of variables deemed ecologically relevant. Highly correlated variables were filtered at this stage.

We first created a list where each BGC unit is paired with all neighbouring BGC units. We then ran a Random Forest model with 101 trees to determine the most important climate variables for each BGC unit separately, considering only that BGC unit and its neighbouring BGC units. These Random Forest models used the “Gini” split rule, and variable importance was calculated based on Gini impurity. For classification, this means that the importance of a variable was determined by how much it reduces Gini impurity when it is used for splitting. Once a list of variable importance was created for each BGC unit, we retained the most important variables, making sure to include even those that may have only been important for a few select BGC units.

Table A1. Variables included in our ‘expert selection’ variable set.

Code	Climate variable	Time scale
CMD_sm	Hargreaves climatic moisture deficit	Summer
CMD.total	Climatic moisture deficit (total)	Annual
DD5_sp	Degree days above 5°C	Spring
DDsub0_sp	Degree days below 0°C	Spring
Tave_sp, Tave_sm	Mean temperature	Spring, summer
Tmax_sp, Tmax_sm	Mean daily maximum temperature	Spring, summer
Tmin, Tmin_at, Tmin_sm, Tmin_sp, Tmin_wt	Mean daily minimum temperature	Autumn, summer, spring, winter, annual
MWMT	Mean warmest month temperature	Warmest month
EXT	Extreme maximum temperature over 30 years	30-year maximum
PAS, PAS_sp	Precipitation as snow	Spring, annual

PPT_MJ, PPT_JAS	Total precipitation	May–June, July–September
Eref_sp, Eref_sm	Hargreaves reference evapotranspiration	Spring, Summer
CMI	Climate moisture index	Annual
SHM	Summer heat:moisture index	Summer

All seasonal variables set:

Finally, we also considered a suite where we included a large number of seasonal and annual variables for comparison (Table A2).

Table A2. Variables included in our 'all seasonal' variables set. Note: Codes listed without explicit seasonal delineation include seasonal or annual variants when the corresponding time scale column includes those designations.

Code	Climate variable	Time scale
CMD	Climatic moisture deficit	Spring, summer, autumn, winter, annual
DD18	Degree days above 18°C	Spring, summer, autumn, winter
DD5	Degree days above 5°C	Spring, summer, autumn, winter
DDsub0	Degree days below 0°C	Spring, summer, autumn, winter
DDsub18	Degree days below 18°C	Spring, summer, autumn, winter
Eref	Hargreaves reference evapotranspiration	Spring, summer, autumn, winter
NFFD	Number of frost-free days	Spring, summer, autumn, winter
PAS	Precipitation as snow	Spring, summer, autumn, winter
PPT, PPT_MJ, PPT_JAS	Precipitation	Spring, summer, autumn, winter, May–June, July–September
RH	Relative humidity	Spring, summer, autumn, winter
Tave	Mean temperature	Spring, summer, autumn, winter
Tmax	Mean daily maximum temperature	Spring, summer, autumn, winter
Tmin	Mean daily minimum temperature	Spring, summer, autumn, winter
CMD.def	CMD deficit	Annual
CMDMax	July climatic moisture deficit	Annual
CMD.total	CMD total (CMD + CMD.def)	Annual
DD_delayed	Growing season delay	Annual

A4. BGC model validation

To classify BGC units, we first trained four Random Forest machine learning models using the '*ranger*' package in R (Wright and Ziegler 2017). These models were trained with the points selected via various schemes, detailed as V1-V4 (**Appendix A2**). These models were initially trained on 250 decision trees with the extraTrees split rule, which selects splits at random, increasing speed of computation (Geurts et al. 2006). All models had a minimum node size of two. To evaluate the effect that number of trees has on the results of the Random Forest models,

we also reran the V2 model with 500 decision trees (V2.1). V2.1 was ultimately selected as the model used in the current version of CCISS. We also evaluated the effects of sample size and climate variable sets. For model validation, we decided on four study areas to more closely examine the impacts of different sampling schemes, the numbers of sampling points, different methods of handling outliers, number of trees in the Random Forest models, and which climate variables to include. These four study areas represent a wide range in BGC units and therefore climates included in the western North America broader analysis. These areas are Bamfield, Kamloops, Pemberton, and Smithers, BC. We validated models both quantitatively, by looking at raster prediction errors and qualitatively, through sensitivity analyses.

Upon comparison of the spring 2025 sampling schemes and subsequent Random Forest models with the November 2024 approach, which used a model trained on 200 randomly placed points in each BGC unit, it appears that the recent models have greatly improved in terms of raster prediction error. When comparing different sampling schemes, we found that the models trained on points derived from the V4 sampling scheme had the lowest raster prediction errors, across all smaller study areas and across the entire province (Figure A4). Although this would imply that V4 was the best sampling scheme and should be the one used in the CCISS tool, running qualitative sensitivity analyses (outlined below) demonstrated that baseline prediction error is just a small portion of model assessment. Relying solely on a single, all-encompassing metric of model performance, such as prediction error, can be misleading. While quantitative error measures provide useful insights, they do not fully capture the model's ability to make ecologically meaningful predictions. In this case, a qualitative assessment of spatial patterns and predicted distributions was essential for evaluating the model’s effectiveness.

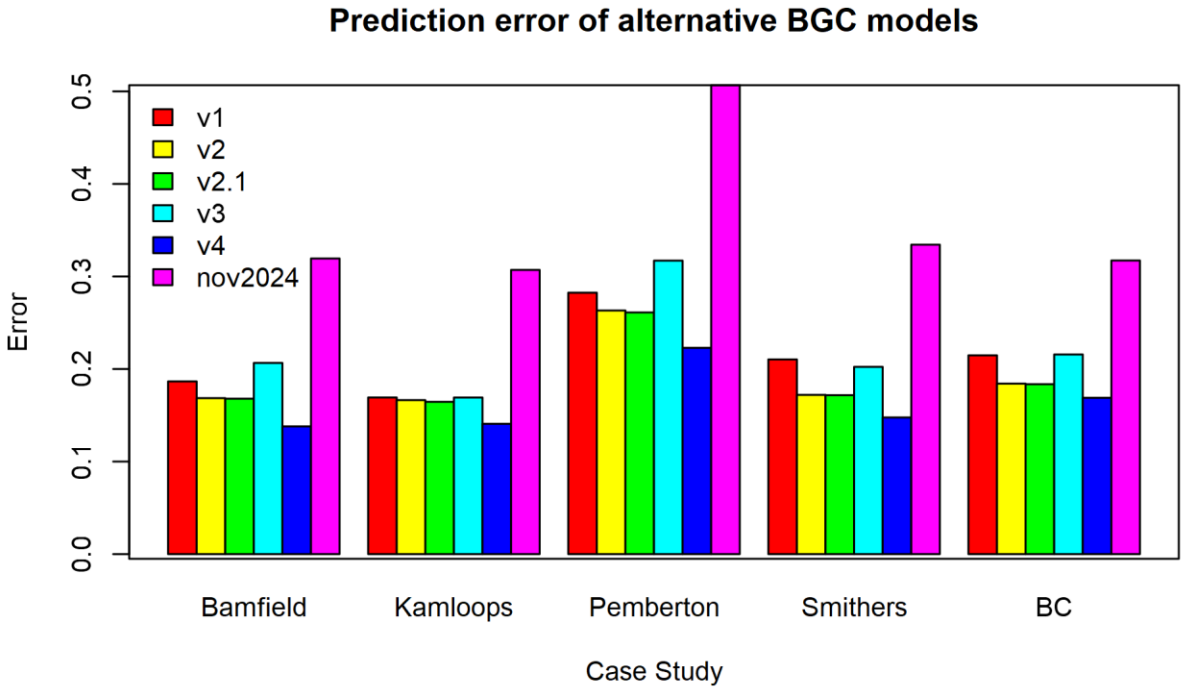

Figure A4. Raster prediction errors across four case study regions, and the entire province of British Columbia. V1 and V3 represent Random Forest models trained on points selected with an asymptotic sampling scheme, where V1 was also scaled according to variability in mean annual temperature (MAT), and **V3** essentially represents 2000 randomly drawn points across BGC units regardless of size. **V2** and **V4** represent Random Forest models trained on points selected with a square root approach to account for differences in BGC area and are also scaled according to the variability in MAT. **V2.1** is identical to **V2** with the exception that the Random Forest model was trained on 500 trees (all other models were trained on 250). The **nov2024** version represents a model trained on 200 randomly placed training points in each BGC.

Sensitivity analyses

For validation, we generated projections using climatic predictions from the 2041-2060 time period based on the EC-Earth3 general circulation model (GCM) under the ssp245 Shared Socioeconomic Pathway (SSP) scenario. Unless otherwise stated, we used the simple climate variable set (**Appendix A3**) for sensitivity analyses.

Sampling scheme

Using the training points obtained via methods described in **Appendix A2**, we trained four main Random Forest models—one for each training point sampling scheme. Initially, we fit each model with 250 trees, but after inspection and sensitivity analyses, we refit V2 with 500 trees (V2.1) and include V2.1 in further assessments.

Across the Bamfield, Pemberton, and Smithers case study areas, baseline BGC unit predictions were less sensitive to changes in sampling scheme than future projections. Key differences in baseline predictions we observed between V1-4 primarily involved the over and

underrepresentation of certain BGC units. For example, in the Bamfield case study area, we noticed an extension of the CWHvh1 into the CHMvm1 up the Nitinat Valley in the V1 and V3 schemes, a pattern that does not appear in the baseline reference map. In Pemberton, we noted a significant reduction in the alpine BGC unit, Coastal Mountain-heather Alpine (CMA) in the V1 and V3 schemes compared to V2, V2.1, and V4.

Differences between sampling schemes were more pronounced in the future projections, showing novel climate introductions, further over and underrepresentation of certain BGC units, and larger variations in the spatial organization of elevation gradients. Notably, the asymptotic sampling schemes (V1 and V3) yielded better spatial organization in elevational gradients, on average, than models fit with square root derived training points (V2 and V4). This which raises an important question—why? Interestingly, when we increased the number of trees in the V2 version of the Random Forest model to 500 (V2.1), the spatial organization recovered.

Overall, the differences we detected in the future projections across sampling schemes demonstrates the importance of incorporating several validation measures in model assessment and not just choosing the model with the lowest baseline prediction error. It is also noteworthy that the sampling design of the training points alone can have influences that are difficult to explain on BGC projections.

Number of sampling points

We evaluated the quality of both the reference baseline and the future predictions made with models trained on 50, 200, 500, 2000, 8000, and all training points. We looked at the influence of these in both the Pemberton and Kamloops study areas. Consistently, we found better predictions with higher numbers of training points, with almost no difference between the mapped BGCs and our baseline predictions when all training points were used. In the future projections, increasing the number of sampling points tidied up the elevational gradients and minimized the amount of spatial variability (i.e., where lots of pixels are classified differently from neighbouring pixels).

Outlier handling

We assessed the effects of different outlier removal thresholds on our Random Forest models, testing cutoffs ranging from no removals (0%) to increasingly strict thresholds: 0.27% (±3σ), 5% (±1.96σ), and 32% (approximately ±1σ) of the data removed. We also used the Pemberton and Kamloops study areas for this. In the future projections, the model version with no outliers removed struggled with spatial organization, especially in valley bottoms and along steep elevational gradients. The 0.27% (±3σ) model showed improvements with spatial organization, and differences between the 0.27% (±3σ), 5% (±1.96σ), and 32% (±1σ) model were minimal.

Number of trees

We evaluated the quality of both reference baseline predictions and future predictions made with models trained on 50, 100, 200, and 500 decision trees. While the elevational organization seemed generally better with higher numbers of trees, it is also important to consider the tradeoffs between computational feasibility and model quality. There is a linear relationship between computation time and number of trees in an Random Forest model (Probst et al. 2019). However, after a certain point, adding more trees only leads to marginal performance improvements. Meanwhile, adding trees means a higher computational cost (e.g., higher memory usage and slower predictions), particularly when working with large spatial datasets such as those we used here.

Climate variable set

We tested the impacts of our choice in climate variable set (i.e., simple, expert selection, or full seasonal, as described in **Appendix A3**) in both the Pemberton and Kamloops study areas. Although our choice of variables did not make a big difference in the baseline projections, we found that it had a large impact on future projections.

A5. Data sources

We used a digital elevation model (DEM) with a spatial resolution of 0.002 degrees (6.94 arcseconds), and rasterized BEC V13 polygons (*unpublished data*).

A6. Literature cited

Daust, K., C. R. Mahony, B. Tremblay, and C. Barros. 2024. climr: Downscaling climate data in R.

Geurts, P., D. Ernst, and L. Wehenkel. 2006. Extremely randomized trees. Machine Learning 63:3–42

MacKenzie, W. H., K. Daust, C. Barros, and C. R. Mahony. 2024. ccissr: B.C. Climate Change Informed Tree Species Selection

Wright, M. N., and A. Ziegler. 2017. **ranger** : A Fast Implementation of Random Forests for High Dimensional Data in *C++* and *R*. Journal of Statistical Software 77

Appendix B. Novel Climate Detection

This appendix describes the CCISS novelty detection methods and their results in both a coastal case study and a provincial scale analysis. The goal of this article is to provide users with the ability to interpret novelty results with an understanding of the strengths and limitations of the method.

B1. Definitions

The following are definitions of key terms used in this section:

- **Baseline** --- The historical climate used to compare observed climates of more recent periods or simulated future climates. The CCISS baseline period is 1961-1990.
- **Target** --- The current or future climate condition for which a historical analog is identified, and for which novelty is being measured. In CCISS, these are climatic averages over 20-year periods (2001-2020 through 2081-2100) for locations throughout BC.
- **Analog** --- The baseline climate condition assessed as being most similar to the target condition, among a set of candidate analogs. The analog pool for CCISS is the 1961-1990 climates of the BGC subzone-variants of Western North America.

B2. The novelty metric – sigma dissimilarity

The basic question we are asking in novel climate detection is: "how similar is the future climate condition to the climatic conditions of the historical climate analog." A statistical rephrasing of this question is: "what is the probability that the future climate condition was drawn from the same distribution as the climatic conditions of the historical analog?" Mahalanobis distances are commonly used to answer this type of question. Following Mahony et al. (2017), we translate Mahalanobis distances into a novelty metric called sigma dissimilarity.

Mahalanobis distance

The typical understanding of distance is Euclidean distance, in which a line of equal distance from any point has the shape of a circle or sphere. The problem with Euclidean distance for novelty detection is that it doesn't account for correlations among variables. If variables are correlated (as climate variables often are), the standardized Euclidean distance will overestimate the novelty of some points and underestimate the novelty of others. Mahalanobis overcomes this problem by accounting for correlations among variables (Figure B1). Instead of following concentric circles like Euclidian distances do, Mahalanobis distances follow ellipses with eccentricity that matches the correlation of the data. Assuming multivariate normality, Mahalanobis distances represent lines of equal probability that a new observation is an outlier. Mahalanobis distances are equivalent to standardized Euclidean distances in the principal components of the variables.

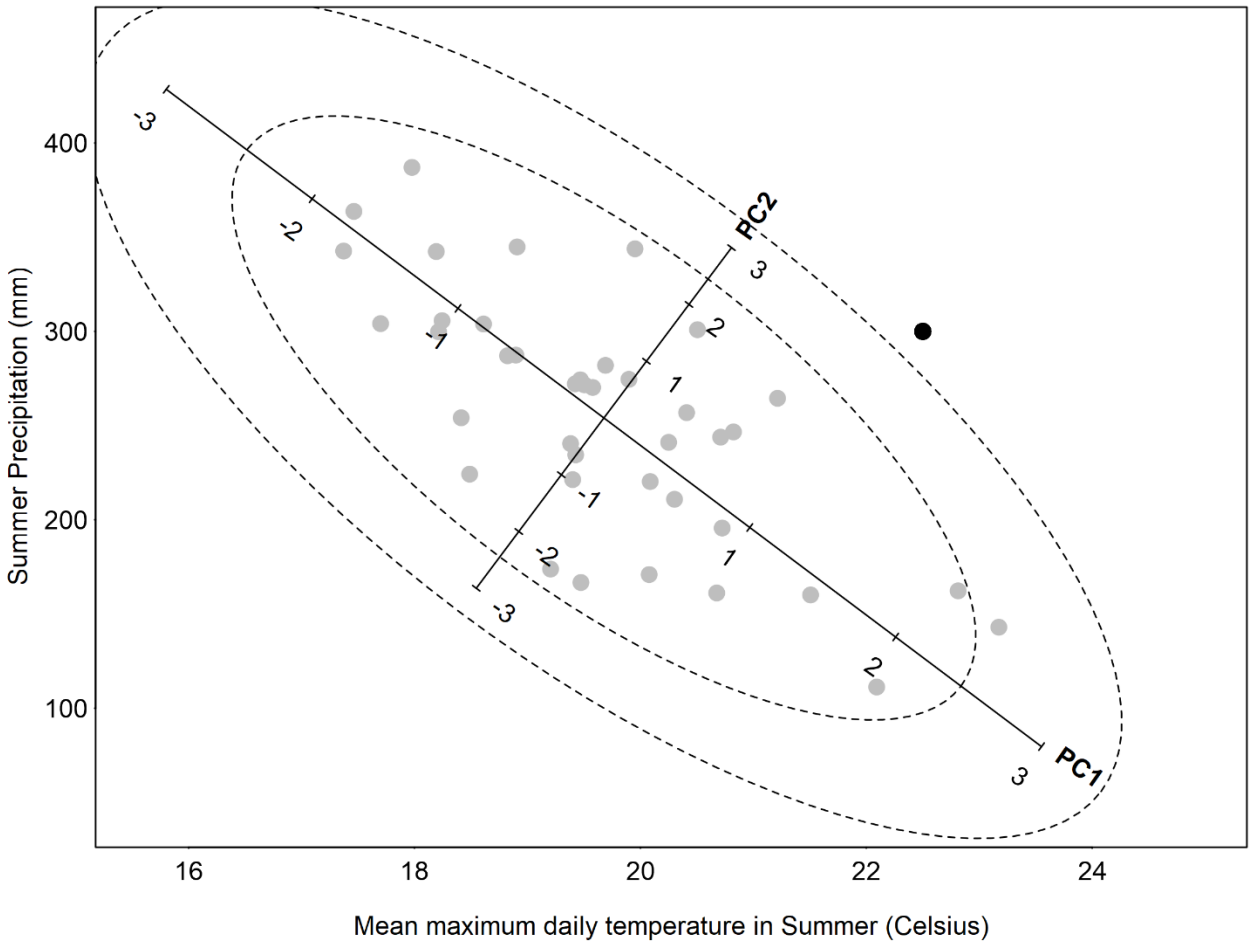

Figure B1. Illustration of novelty/outlier detection using Mahalanobis distance. Grey dots are individual years of 1951-1990 summer temperature and precipitation at Quesnel, BC. The climate variables are correlated such that there is a dominant mode of variability—cool/wet years vs. warm/dry years—that is detected by a principal components analysis (PCA) as PC1. The more unusual mode of variability (cool/dry vs. warm/wet years) is the second principal component, PC2. Mahalanobis distance is equivalent to Euclidean distance measured in the standardized principal components. The dashed ellipses are the Mahalanobis distances within which 95% and 99.7% of the years would be expected to occur, assuming multivariate normality of the data. A hypothetical new climate observation (black dot) is within the individual (univariate) distributions of historical temperature and precipitation, but it can be identified as an outlier (i.e., a novel climatic condition) because it falls outside their bivariate distribution as described by Mahalanobis distance.

Sigma dissimilarity

Mahalanobis distance is not in itself an adequate metric of novelty because the statistical meaning of distances depends on the number of dimensions in which the distances are measured. Since CCISS uses different numbers of principal components for different BGC analogs, these dimensionality effects are important. Thus to obtain an intuitive and statistically consistent novelty metric that can be used to compare across analogs, we use sigma dissimilarity (Mahony et al. 2017).

Theory of sigma dissimilarity

The effect of dimensionality on the probabilities of distances can be visualized using a random sample from a multivariate normal distribution (Figure B2a). In any one dimension (either the x or y axis), there is a 68% probability that an observation will be within one standard deviation of the mean. In other words, 68% of the points are expected to have a distance from the mean of less than one. In two dimensions, the probability that an observation will fall within a distance of one from the centroid is reduced to 39%. As the number of dimensions increases, even more observations fall outside a distance of 1, such that all observations become increasingly distant from their own mean.

The probability of multivariate normal distances (and therefore Mahalanobis distances) are described by the chi distribution with degrees of freedom equaling the dimensionality of the data (Figure B2b-c). Chi percentiles can be expressed using the terminology of univariate z-scores; i.e., 1, 2, and 3 sigma (σ) for the 68th, 95th, and 99.7th normal percentiles, respectively. The chi distribution in 1 dimension is a half-normal distribution, and the sigma levels correspond to distance. This result is expected because Mahalanobis distances in one dimension are the absolute values of z-scores. At increasing dimensionality, the sigma levels shift away from the origin. For example, 1σ (the 68th percentile) occurs at Mahalanobis distances of 1.0 in one dimension versus 1.5 in two dimensions (Figure B2b, c). By extending sigma levels into multiple dimensions, sigma dissimilarity serves as a multivariate z-score.

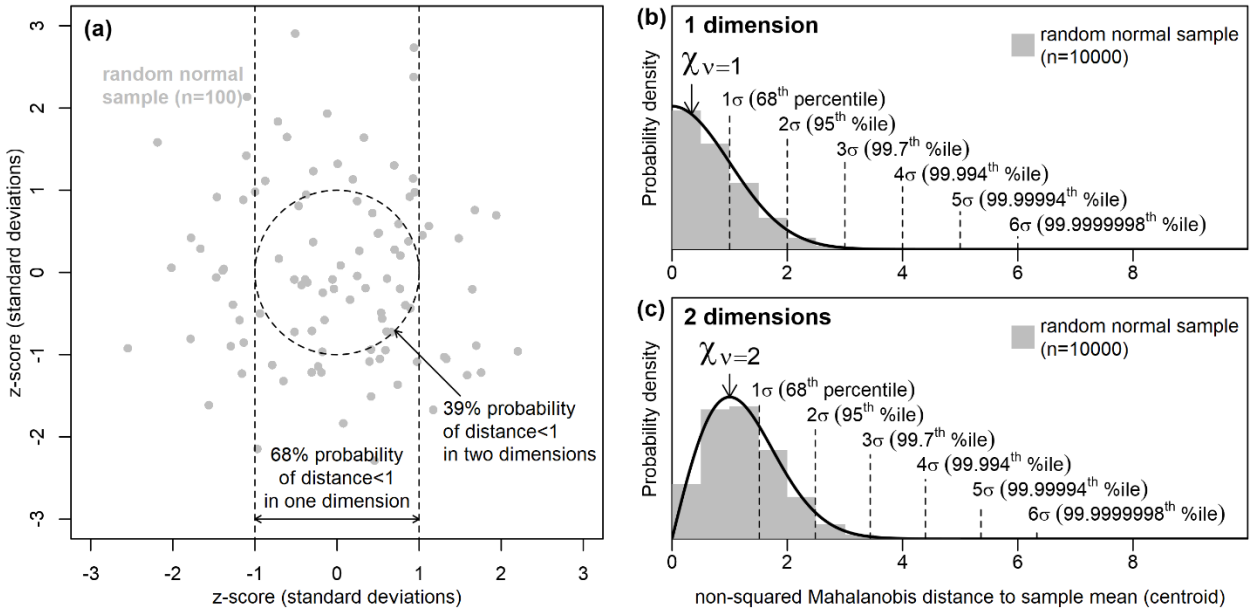

Figure B2. An illustration of the basis for sigma dissimilarity. **a**, Geometric illustration of the effect of dimensionality on the probability density of distances. **b-c**, Use of the chi distribution to describe Mahalanobis distances as sigma (σ) percentiles at increasing dimensionality.

B3. Novel climate detection in CCISS

To measure climatic novelty with sigma dissimilarity, we need to define the variation that the Mahalanobis distance is proportioned to. The question we are asking is "novel relative to what?"

Since CCISS uses BGC subzone-variants as climate analogs, it is primarily interested in whether the target (future) climate matches the baseline (historical) climate of any location within the analog. In other words, CCISS is most interested in measuring novelty relative to the *spatial climatic variation* of the analog. However, we have also found it necessary to include the *interannual climatic variability* (ICV) of the analog into the measurement (see rationale below). Therefore, we measure novelty relative to the spatiotemporal climatic variation of the analog. This is different from the approach of Mahony et al. (2017), who defined novelty relative to the interannual climatic variability of the target climate.

Novel climates detection in CCISS is done separately for each BGC analog, in three steps:

1. Principal components analysis (PCA) – conduct a PCA on an equal sample of spatial variation in baseline and target climates of the analog, and choose a subset of the PCs that represent 95% of the variance in this pooled sample;
2. z-standardization – standardize the retained PCs so that the pooled analog spatial climatic variation and ICV have a mean of 0 and a standard deviation of 1; and
3. Measure sigma dissimilarity of target (future) climatic conditions. This approach measures dissimilarity from the target climate to the average climate of the BGC analog rather than to the most similar (potentially peripheral) location within the analog.

Rationale

In step 1, why is the PCA done on a combination of the baseline and target climates?

Mahalanobis distance can be unstable if all PCs are included. For this reason, we retain only 3-6 PCs that describe 95% of the variance in the data. however, there is no guarantee that the major principal components of the analog's spatial climatic variation include the ways in which the future climate is different than the analog climate. In other words, distance measurement in the primary modes of spatial variation could be blind to the primary modes of climate change. This would cause underestimation of novelty. To ensure that the differences between the future climates and the analog climate are detected, we conduct the PCA on a combination of spatial variation in both the analog and target climates.

In step 2, why are the PCs standardized to both spatial and temporal variation of the analog?

Scaling the novelty metric to spatial variation alone has two important shortcomings. First, very small BGC subzone/variants often have much less climatic variation relative to larger ones simply due to their size. This can lead to overestimation of climatic novelty for small subzone/variants. Second, spatial variation can be very low in the modes of climate change, especially for analogs that are located on flat terrain. Measuring climatic distances in PCs standardized to very low spatial variation can result in artificially high sigma dissimilarity. To make the novelty metric more robust, we include interannual climatic variability (ICV) in the z-standardization. Specifically, we pool the analog's spatial climatic variation with the climate values of each year in the 1951-1990 period. A 40-year sample is used because it is more robust than 30 years. The ecological rationale for including ICV is that a future condition can be considered less novel if it is similar to individual years in the baseline climate instead of being entirely unprecedented. The effect of this modification is to reduce novelty in small and/or topographically simple BGC analogs.

Variable selection

Novelty should be measured in climate variables that are ecologically meaningful and relevant to the variables that were used to identify the climate analog. Superficially, it seems logical to measure novelty using the variables that were used to train the BGC model. However, there are two reasons why the input variables to the BGC model are not necessarily the most appropriate for novelty measurement.

1. Machine learning doesn't use input variables equally for differentiating each potential analog from its neighbours. For example, the variable selection and thresholds used to differentiate grassland from forest climates are likely to be different from those used to differentiate subalpine from alpine climates. Mahalanobis distance, in contrast, is blind to the variables that differentiate the focal analog from its neighbours and instead emphasizes the modes of spatial and temporal variation of the analog itself. As a result, even if both the BGC model and the novelty detection were given the same input variables, their results would not necessarily be any more consistent than a novelty detection with different input variables.
2. Mahalanobis distance relies on the assumption of multivariate normality, i.e., that the analog variation can be approximated with a hyperellipse. While this assumption generally holds for temperature and log-transformed precipitation variables, it can be severely violated by other biologically-relevant climate indices (e.g., degree-days and climatic moisture deficit) that are often used in BGC modeling. Violations of the multivariate normality assumption can cause large novelty artefacts.

For those reasons, CCISS novelty detection uses a basic 12-variable set of Tmin, Tmax, and precipitation of the four seasons.

Threshold selection

Novelty is a continuum. Further, the novelty measurement is prone to error because the relevant climates variables and ecological responses are unknown and specific to each ecosystem and tree species. Hence there is no objective basis for a threshold of sigma dissimilarity above which a target climate can be labelled as a novel climate. Nevertheless, it is necessary to establish a threshold of novelty above which BGC analogs will not be used to make inferences of future species suitability. If this threshold is set too high, CCISS will make invalid and misleading inferences. If the threshold is too low, useful information may be discarded.

The CCISS tool uses a default threshold of 5σ dissimilarity to infer a novel climate and discard the BGC analog. 5σ is a low bar to meet for analog goodness of fit. Assuming multivariate normality, it corresponds to a 1-in-3,500,000 chance that the target climate would occur within the analog's geographical range and/or interannual climatic variability. However, given the many potential sources of error in the CCISS novelty measurement, 5σ novelty could be more likely in some cases. This threshold provides a reasonable balance between the risks of making an invalid inference vs. discarding useful information.

The CCISS tool also indicates the proportion of each species suitability projection that is in the 3-5σ range. 3σ corresponds to a 1-in-370 event within the analogs spatiotemporal climatic variability. Novelty in the 3-5σ range can be interpreted as a poor analog. In this range, species suitability inferences may be useful but are also likely to be somewhat misleading.

B4. Case study – CWHxm_WA

To illustrate the interpretation of climatic novelty in BGC projections, we use the example of the CWHxm_WA subzone (Coastal Western Hemlock very dry maritime, Washington variant), which spans the Puget Lowlands in Washington State from the hilltops of the San Juan Islands to Portland, Oregon (Figure B3). The baseline 1961-1990 climate of the CWHxm_WA is widely used as an analog for current and future climates of the near-shore lands surrounding the Strait of Georgia including Greater Vancouver.

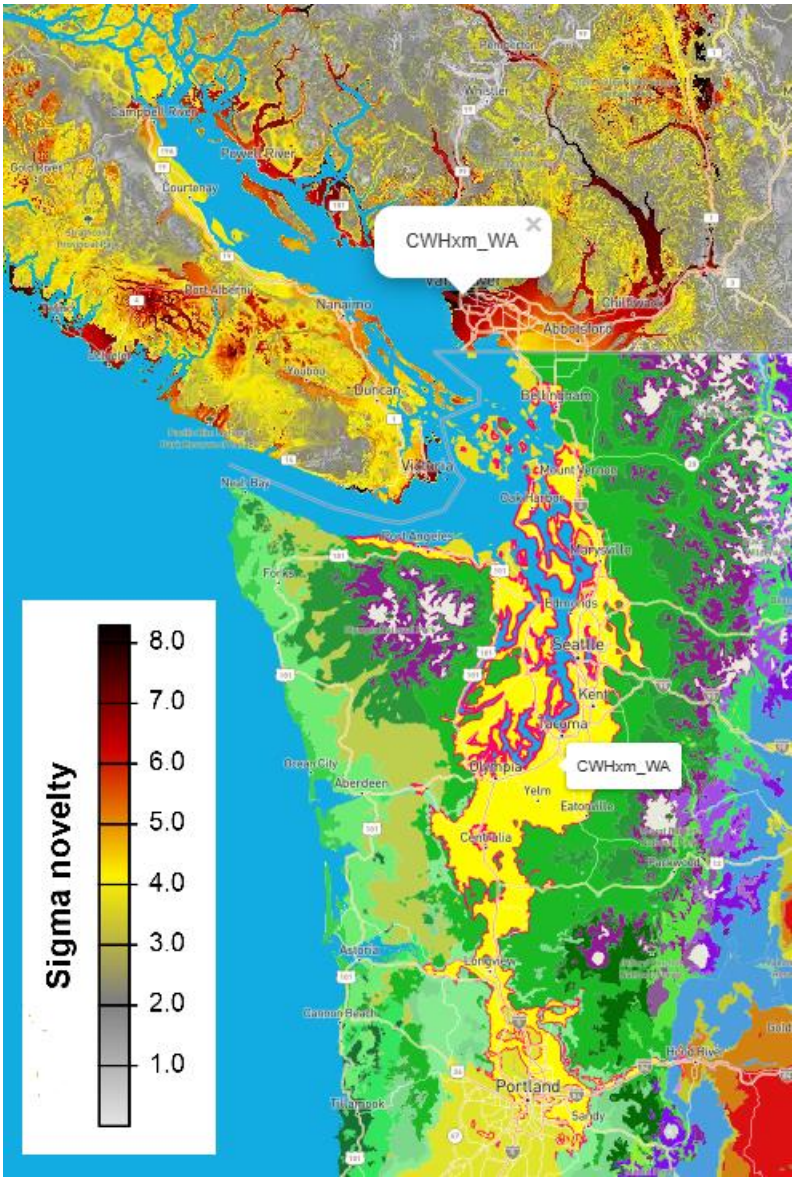

Figure B3. Location of the CWHxm_WA subzone in the Puget lowlands of Washington State. The CWHxm_WA is a common analog for novel near-future climates of the Georgia Basin, including Vancouver, BC. Areas of BC are mapped by degree of novelty for an EC-Earth3 SSP2-4.5 simulation for the 2041-2060 period. This image is a screenshot from the CCISS spatial module.

The variation of the baseline analog climates and the target climates are quantified in a PCA scree plot (Figure B4). The basic interpretation of the scree plot is that there are 12 PCs because there are 12 input variables each standardized to unit variance (st. dev. = 1). Since the PC was performed on the combined target and analog climates , PC1 is the primary mode of separation between the target and analog climates. PC2 contains the most spatial variation in the analog and target climates but very little separation of their means, as visualized in Figure B5b. Subsequent PCs contain decreasing spatial variation of analog and target climates, such that 95% of the variance in the pooled data is contained within the first four PCs. This was the criterion for

exclusion of PCs 5-12. ICV variation does not decline as quickly because it was excluded from the PCA.

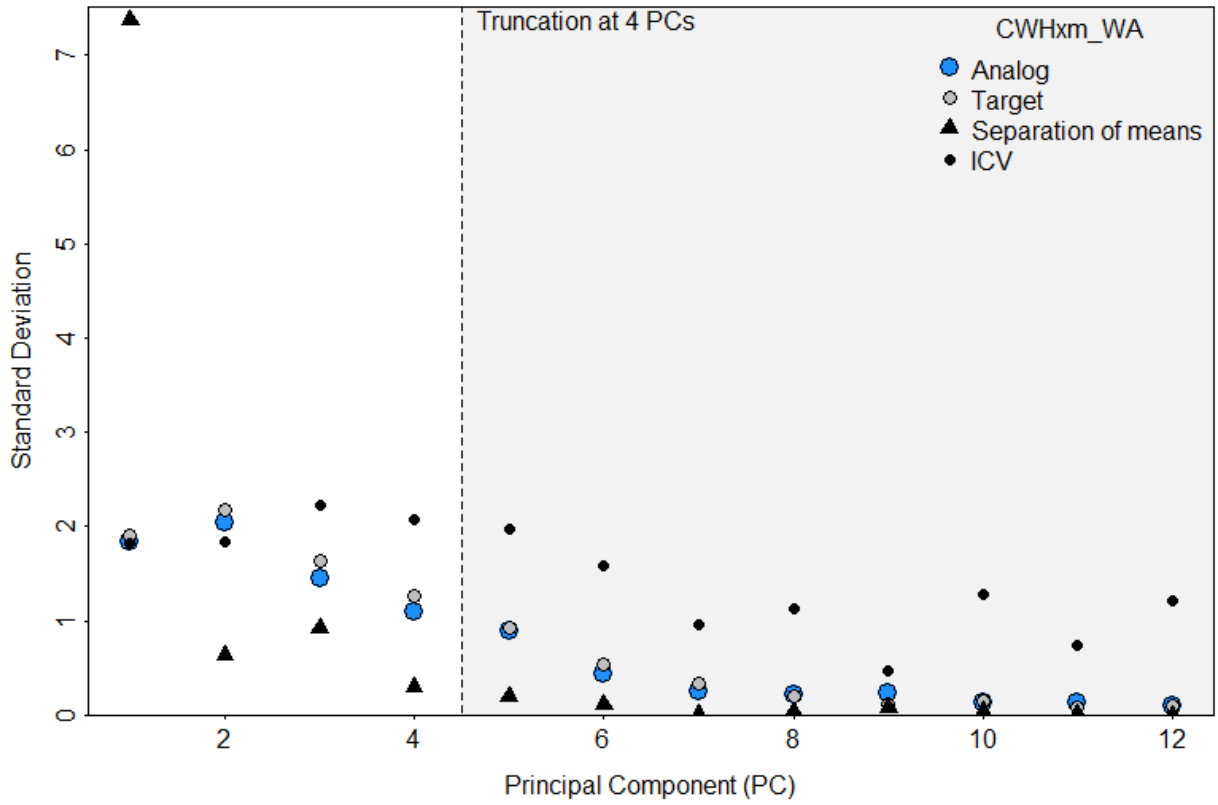

Figure B4. Scree plot of the principal components analysis used in novelty detection for future climates (EC-Earth3, 2041-2060, SSP2-4.5) assigned to the CWHxm_WA analog. Circles represent the climatic variation in each PC of the analog climates (baseline climates of the CWHxm_WA), target climates (future climates assigned to the CWHxm_WA analog), and ICV (1951-1990 interannual climatic variability of the CWHxm_WA). Black triangles are the distance between the means of the target and analog climates. 4 PCs are retained for novelty detection based on the criterion of retaining 95% percent of the pooled target and analog variance.

When viewed in the climate space of the first two PCs (Figure B5a), there is a clear separation in the baseline (1961-1990) climates of the CWHxm_WA (analog climates, blue dots) and the target (2041-2060) climatic conditions that are classified by the model as CWHxm_WA (red, orange, yellow dots). Degrees of novelty in the target climates range from 4σ (yellow) suggesting that the analog is a poor match for the target climate, to 7σ (dark red) suggesting that the analog is non-valid for these highly novel climates. The overlap between the analog and target climates is an illusion of the 2-dimensional space of PC1 and PC2: When viewed in PC3 and PC4 (Figure B5c), the yellow target climates are non-overlapping with the analog climates, demonstrating that the 4th PC was required to detect the novelty in these climates. Since PC5 was excluded from the novelty measurement, the additional modes of novelty are not detected (Figure B5d).

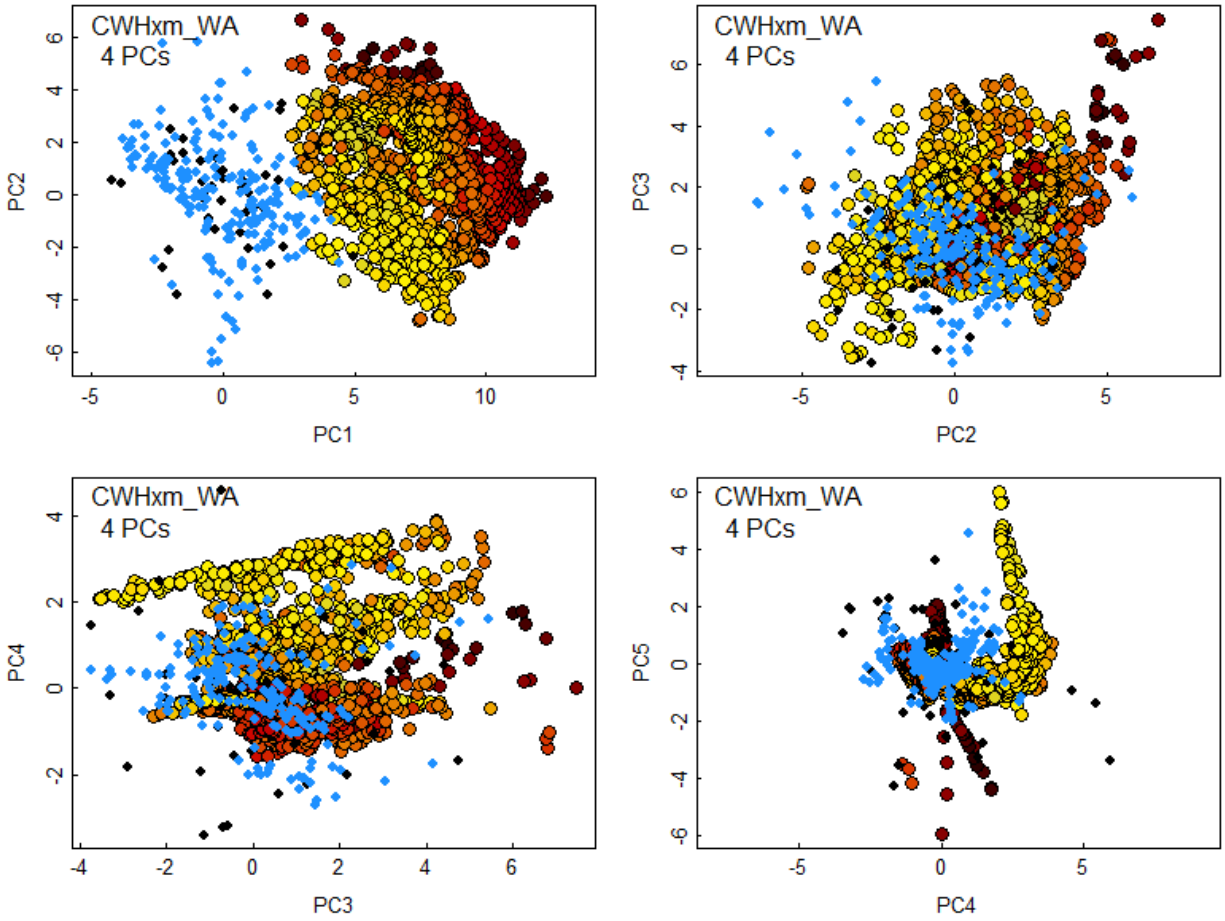

Figure B5. Demonstration of novelty of future climates classified as the CWHxm_WA analog. Blue dots are the spatial variation in baseline (1961-1990) climates of the CWHxm_WA subzone. Black dots are the 1951-1990 interannual variability at a representative location in the CWHxm_WA. Filled circles are projected 2041-2060 climates of locations in British Columbia that are assigned the CWHxm_WA analog by the BGC model, with colors indicating sigma novelty consistent with the Figure B4 legend. Each panel provides a different orthogonal (perpendicular) view on the multidimensional data cloud, with PC1 being the axis with the most variation in the pooled points, and PC5 having the least.

The climatic novelty of the CWHxm_WA target climates can be further visualized in the context of the baseline climates of Western North America (Figure B6). These projected 2041-2060 climates (yellow to red points) have no overlap with the spatial climatic variation of Western North America. The CWHxm_WA (blue points) is the best (i.e., most similar) available analog for these future climatic conditions, but it is a poor analog at best (yellow points) and a non-analog at worst (black points). The primary mode of novelty is positively correlated with mean summer nighttime temperature (Tmin_sm), but is also somewhat correlated with spring and autumn Tmin. Since the mode of novelty is relatively uncorrelated with precipitation and daily maximum temperatures, the CWHxm_WA may be an informative analog for ecological drivers,

such as drought, that are primarily dependent on these variables. This illustrates how a climate analog can be useful for some interpretations, but misleading for others.

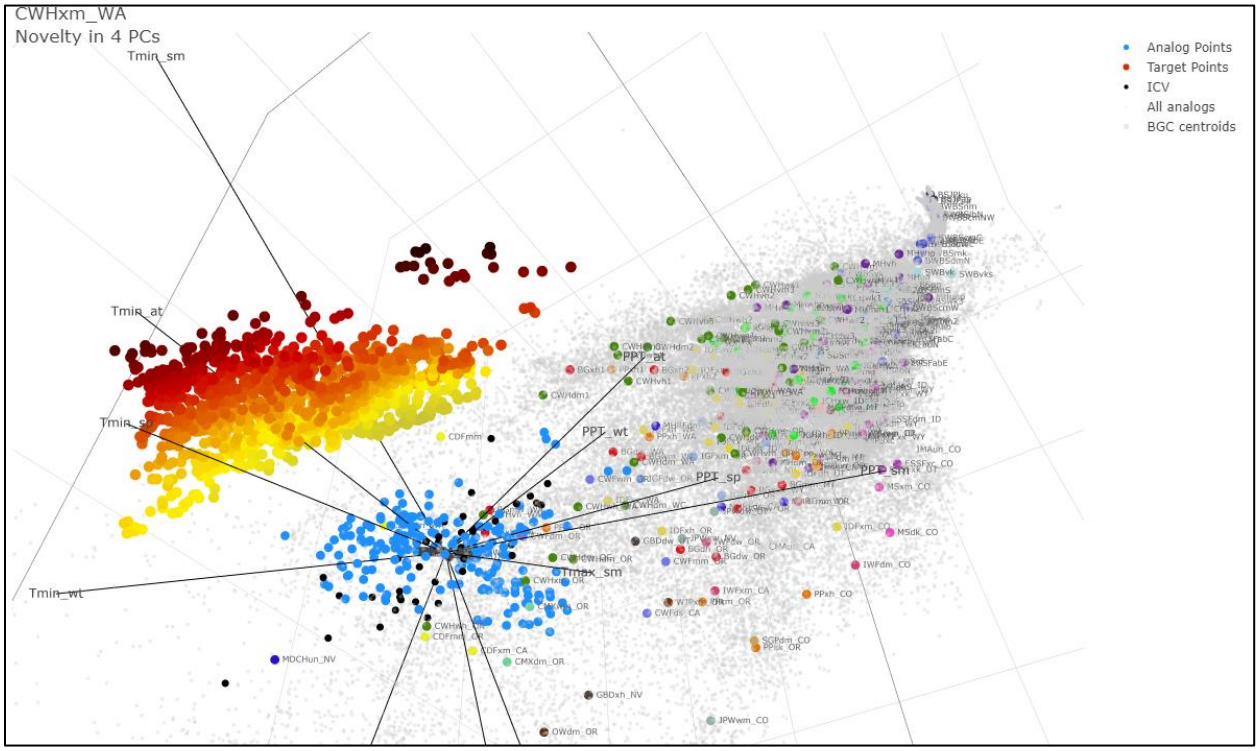

Figure B6. 3-dimensional view of the CWHxm_WA analog and target climate variation in context of other climates of Western North America. Symbology is consistent with Figure B5, with the addition of spatial variation (grey dots) and spatial means (multicolored dots) of the climates of other candidate BGC analogs. Labelled lines indicate correlations of the input variables with the 3 PCs that comprise the viewspace.

B5. Provincial-scale results

Novelty over time

Figure B7 shows the progression of climatic novelty over time in a global climate model simulation of 21st century climate change. Even in the baseline period (1961-1990), scattered locations have some degree of novelty from their model-predicted BGC unit (Figure B7a). This baseline novelty can be due to (1) spatial climatic variation within a BGC that violates the assumption of multivariate normality (isn't elliptical in shape) and (2) baseline model predictions that differ from the BEC mapping. Overall, however, there is a very low degree of novelty in the baseline period, as would be expected.

Novelty in the simulated 2001-2020 climate is also generally low (Figure B7b), with notable novelty inferred only in small areas of the Peace region foothills of Northeastern BC and the Tatshenshini-Alsek provincial park in northwestern BC. By midcentury (Figure B7c), there is widespread inference of moderate novelty (3-5σ; yellowish colours), indicating poor analogs in these areas. However, the majority of the province still has good analogs in this period (0-2σ; grey colours). By the end of the century (Figure B7d), moderately to highly novel (>5σ) climates occupy most regions of the province. Highly novel climates are projected for Northeastern BC, the coast, the Chilcotin Plateau and southern interior.

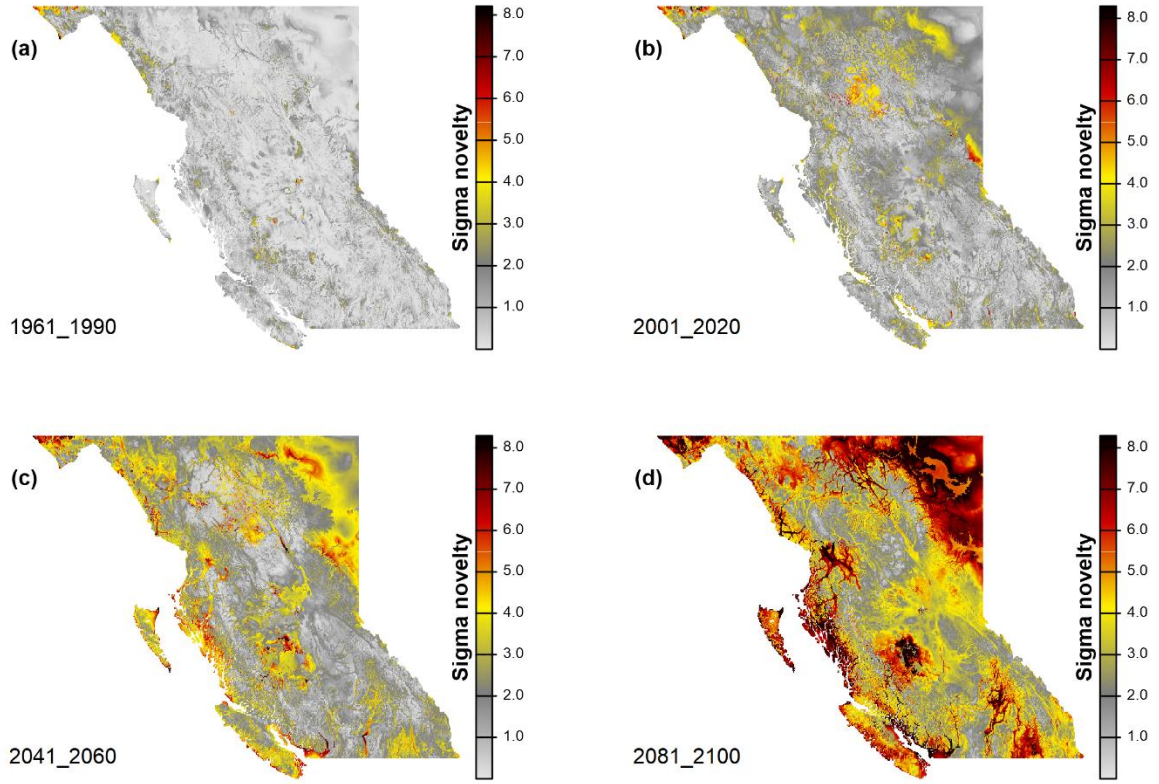

Figure B7. Progression of climatic novelty over time. (**a**) Novelty of baseline observed climate. (**b-d**) Novelty of the EC-Earth3 simulation in three 20-year periods.

Climate model variation

There is considerable variation among global climate models in the magnitude of climatic novelty (Figure B8). Some of this variation is due to different warming rates in the models, as indicated by the similarity in spatial patterns of novelty among models. However, there is also variation among models in the level of novelty at each warming level (Figure B9). For example, the MIROC6 model induces approximately half of the novelty as the GISS model at warming levels above 2\u00B0C.

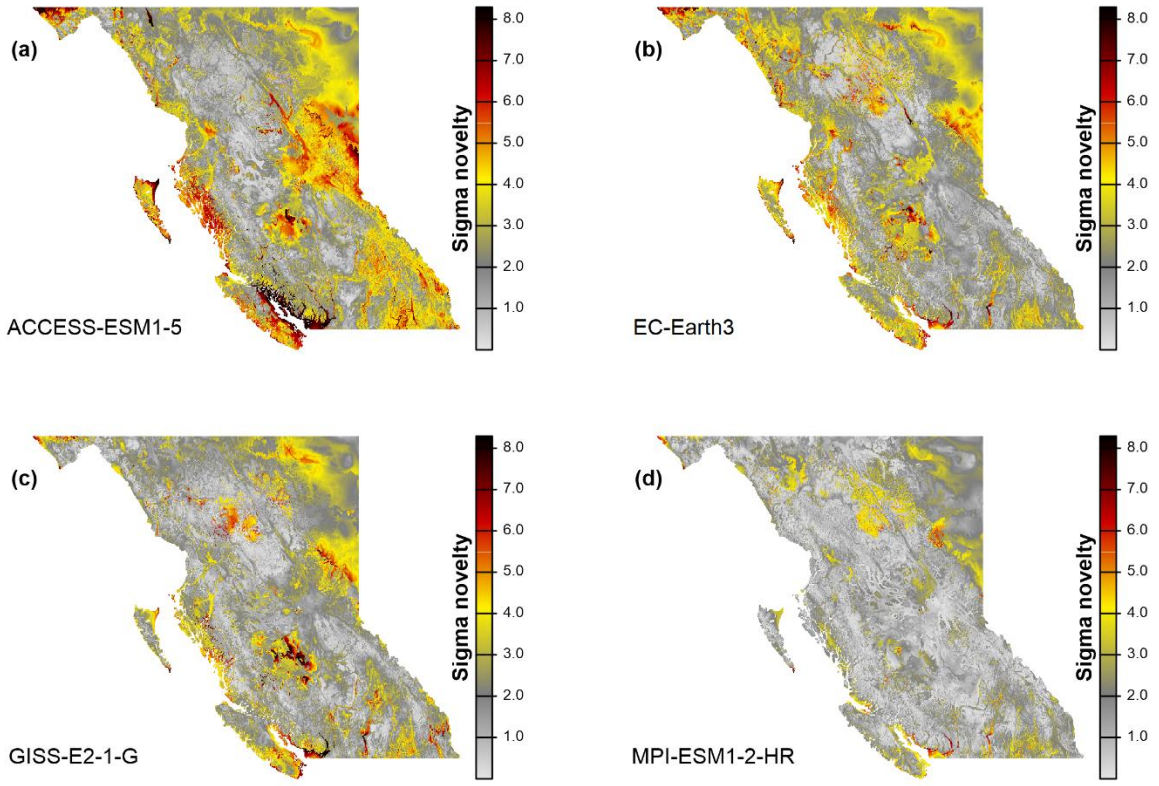

Figure B8. Variation in climatic novelty among individual simulations of global climate models, for the 2041-2060 period (SSP2-4.5 scenario).

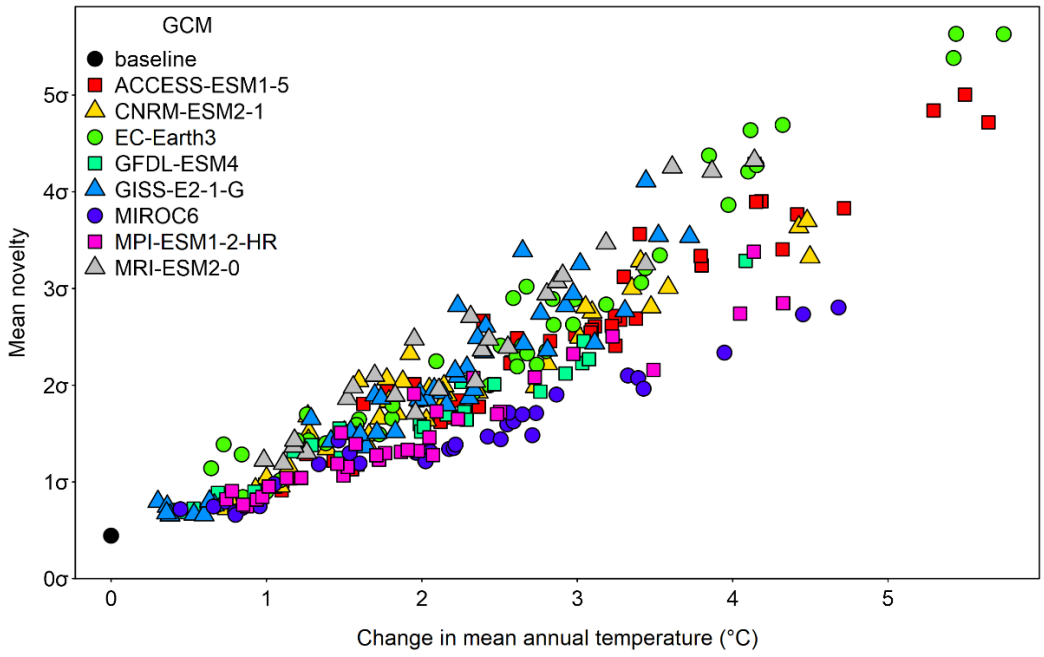

Figure B9. Relationship of BC-mean novelty to warming level. Each point is one of five 20-year periods, 3 simulations, and 3 SSPs for each model, for a total of ~45 points per model.

B6. Sensitivity Analysis

This section provides sensitivity analyses demonstrating the effect of key methodological decisions, namely variable selection, PC truncation rules, and how to balance spatial and temporal variation (ICV) in the standardization. Unless otherwise stated, sensitivity analyses use the EC-Earth3 r4i1p1f1 SSP2-4.5 simulation at the 2041-2061 time period.

Variable selection

The sensitivity of novelty to variable selection is illustrated in Figure B10. Both variable sets produce common regions of novelty, such as on the coast, southern Kootenays, and northeast BC, though the basic variables detect a lesser degree of novelty on the coast. The high-novelty region in the northern Columbia mountains in Figure B10b is likely an artefact of multivariate normality violation in degree-day and reference evaporation variables, which can have many zeros and highly skewed distributions. The disagreement between the two variable sets in the western Chilcotin is more enigmatic and deserves further investigation.

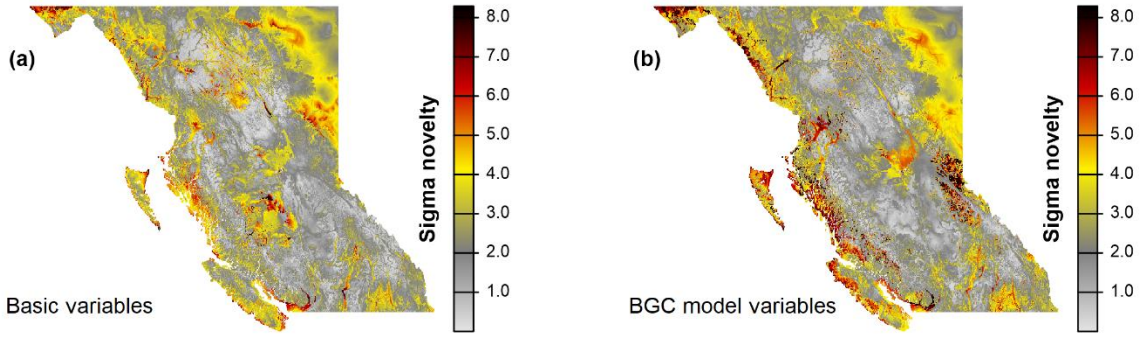

Figure B10. Sensitivity of climatic novelty to variable selection. Basic variables are Tmin, Tmax, and precipitation for the four seasons. BGC model variables are the 23 variables used to train the Random Forest model used to make the biogeoclimatic projections.

Balance of spatial and temporal variation

Novelty is highly sensitive to the relative weight of spatial and temporal variation. When interannual climatic variability (ICV) is excluded from the novelty measurement (Figure B11a), novelty is extreme in areas of interior BC with low topographic relief. This is primarily due to the modes of climate change being different from the few modes of spatial climatic variation on flat terrain. Conversely, when spatial climatic variation is excluded (Figure B11d), novelty is greater on the coast. This is due to lower interannual climatic variability in maritime climates and the prevalence of spatially large BGC subzone/variants in the coastal classification. The contrasting drivers of coastal and interior novelty balance each other at intermediate ICV weights (Figure B11b-c). We chose an ICV weight of 0.5 for novelty detection in CCISS).

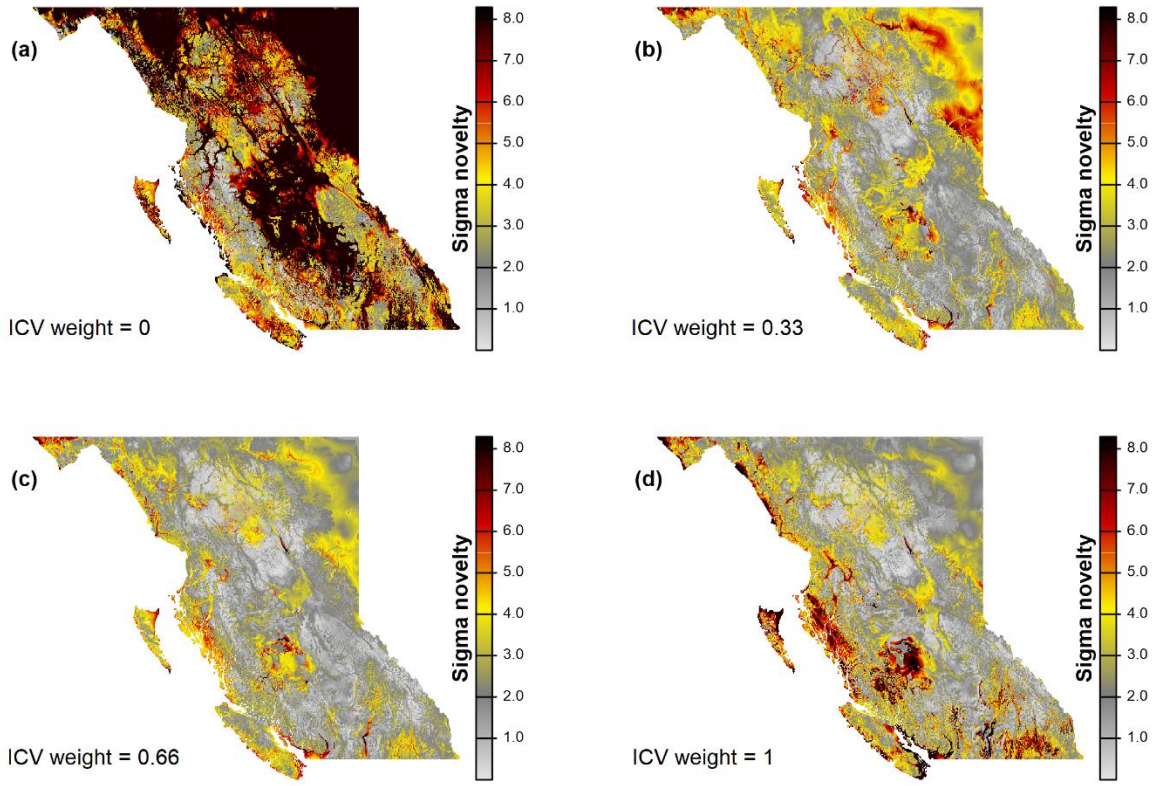

Figure B11. Sensitivity of climatic novelty to the balance of spatial vs. temporal (ICV) variation. ICV weight is the proportional influence of interannual climatic variability on the calculation of Mahalanobis distance, relative to spatial climatic variation.

Dimensionality (PC truncation rule)

The CCISS novelty method retains PCs with a cumulative variance of 95% of the total data variance. Consequently, the dimensionality of the sigma dissimilarity differs among analogs, within the range of 4-6 PCs. Figure B12 illustrates the effect of dimensionality on novelty by measuring sigma dissimilarity at uniform number of PCs throughout BC. In general, novelty increases when more PCs are retained because there are more modes of variation for future conditions to be different from their analog. However, the CCISS novelty method is not highly sensitive to the dimensionality of the measurement (Figure B12). This indicates that the results of novelty detection do not hinge on the decision to the 95% variance criterion vs. other potential truncation rules.

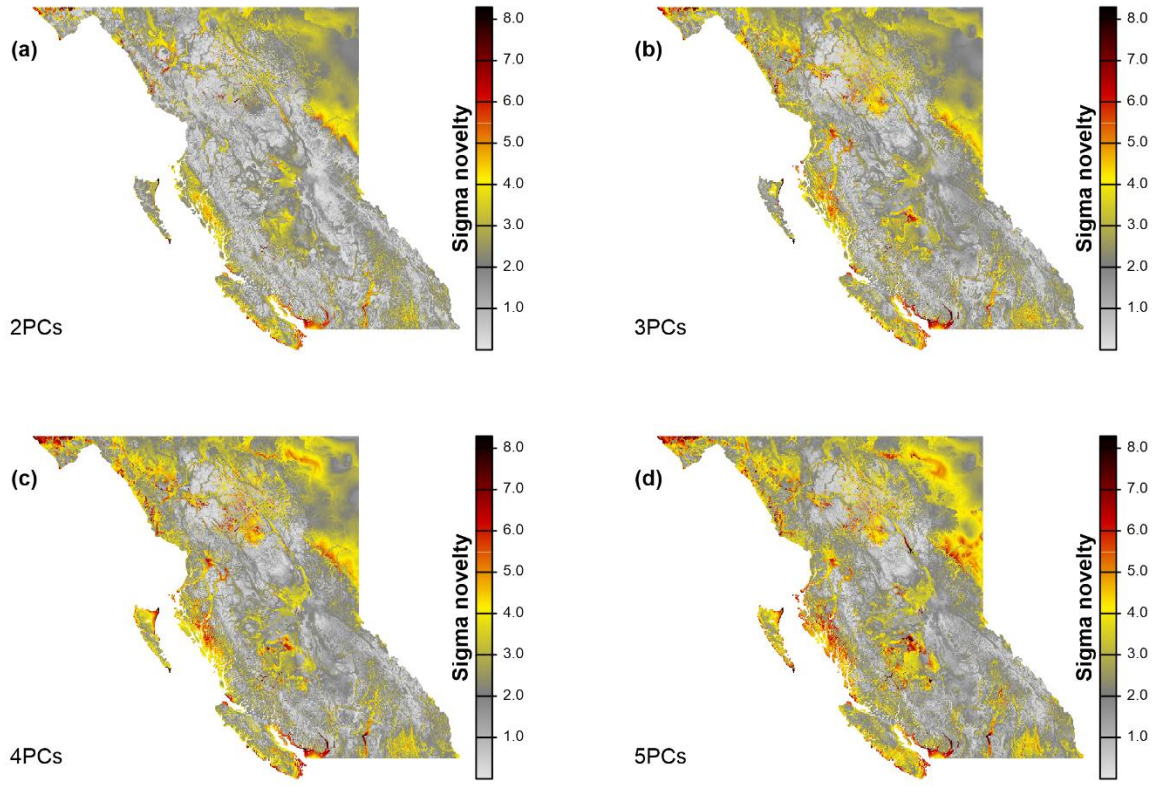

Figure B12. Sensitivity of climatic novelty to the number of principal components used to calculate Mahalanobis distance.

B7. Literature Cited

Mahony, C.R., A.J. Cannon, T. Wang, and S.N. Aitken. 2017. A closer look at novel climates: new methods and insights at continental to landscape scales. Global Change Biology 23: 3934–3955.

Appendix C. Next Steps

C1. Beyond Random Forest models

In the past, Random Forest models have commonly been employed for the purposes of BGC projection (Wang et al. 2012, MacKenzie and Mahony 2021). Indeed, Random Forest models have been found to be equal to or better than other tree-based statistical models in the context of climate envelope modeling (summarized in Wang et al. 2012). However, it will be important to investigate different classes of models. In particular, generative adversarial networks (GANs, Goodfellow et al. 2014) are a conceptually different type of modeling that make use of Convolutional Neural Networks. By treating the problem as an image generation task rather than a single point classification problem, they could improve spatial consistency of modelled BEC units, as well as removing potential challenges with spatial autocorrelation. We plan to investigate this strategy in FY 25/26.

C2. Updated hyperparameters

Gini impurity and extraTrees are two split rule algorithms commonly used in Random Forest modeling. The difference between the two is that Gini impurity selects the best possible split (i.e., the one that reduces impurity the most), while extraTrees (Geurts et al. 2006) randomly selects a split point for each feature, making it computationally much faster. An important next step will be to determine whether this trade-off for computational feasibility compromises the quality of the projections.

C3. Predicting into gaps

We aim to conduct further validation on the current (and future) BGC projections model(s) by testing its ability to predict into ‘gaps’, or areas that have not been used in developing the training data set. This is a way to cross-validate our results that minimizes the influence of spatial autocorrelation. So far, we have done this on the scale of our case study areas, and one of our next steps will be to properly document our findings from these case study areas. In the future, we would also like to do this on the WNA-scale

C4. Extreme climates

In coming months, we plan to integrate extreme climate variables into the climate variable sets. Extreme climate events are known to have large impacts on ecosystems, often even more so than shifts in climatic means (Ummenhofer and Meehl 2017, Walsh et al. 2020). Incorporating climate extremes into both the ‘*climr*’ package and then into our BGC projections is a priority.

C5. Further validation metrics

Finally, we aim to adopt or develop further metrics to better understand the quality of our BGC model projections. For example, we want to find a metric to represent how well a model can capture elevational banding. Are the predicted raster pixels discontinuous, with units crossing

over? What is the spatial contiguity of BGC units? Here, we aim to further examine available tools and adapt them to fit our needs.

C6. Literature cited

MacKenzie, W. H., and C. R. Mahony. 2021. An ecological approach to climate change-informed tree species selection for reforestation. Forest Ecology and Management 481:118705.

Wang, T., E. M. Campbell, G. A. O'Neill, and S. N. Aitken. 2012. Projecting future distributions of ecosystem climate niches: Uncertainties and management applications. Forest Ecology and Management 279:128–140.

Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. 2014. Generative Adversarial Nets. Advances in neural information processing systems 27.

Appendix D. Climate Change Projections

This section describes the rationale for the global climate model simulations used to create the ensemble biogeoclimatic projections and describes the small ensemble of five representative simulations for which biogeoclimatic projections are available for download.

D1. The climate model ensemble for representing climate change uncertainty

The biogeoclimatic projections ensemble incorporates three types of climate change uncertainty: modeling uncertainty, natural variability, and socioeconomic uncertainty. These uncertainties are represented by modeling biogeoclimatic projections for an ensemble of 60 potential future climate states (8 climate models x 1-3 simulation runs x 3 socioeconomic scenarios) for each future time-period.

Climate modeling uncertainty

Climate models are simplifications of the earth system; they involve many compromises in modeling complex processes. Consequently, an ensemble of independent climate models is required to represent modeling uncertainties about climate change outcomes over large regions. CCISS uses an ensemble of 8 global climate models (GCMs), selected by Mahony et al. (2022), for independent modeling methods that are consistent with historical climate changes and the IPCC assessed range of very likely climate sensitivity: ACCESS-ESM1.5, CNRM-ESM2-1, EC-Earth3, GFDL-ESM4, GISS-E2-1-G, MIROC6, MPI-ESM1.2-HR, and MRI-ESM2.0. This ensemble is described and visualized in the cmip6-BC app.

Natural variability

Global climate models, and the Earth system itself, have internal variability—weather at time scales of hours to decades. At any point in time, the climatic conditions in different GCMs can differ not only because of differences in how they model climate, but also due to internal variability (weather). Even 20-year averages can differ significantly in different runs of the same model (Figure D1). For this reason, we include three independent simulation runs of each climate model in the biogeoclimatic projections ensemble.

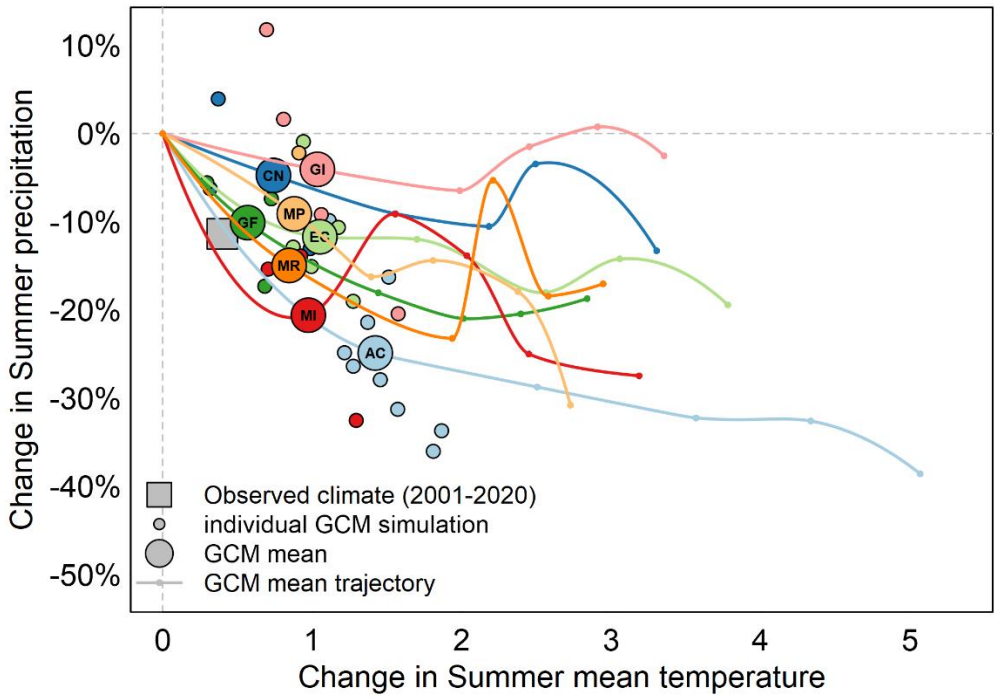

Figure D1. Trajectories of simulated and observed climate change for southern Vancouver Island, illustrating uncertainty due to natural variability (weather) and structural differences among models. Small points are the changes in climate from 1961-1990 to 2001-2020 in up to ten independent simulations for each of eight global climate models (SSP2-4.5 scenario). Larger labelled points indicate the single-model ensemble mean change in 2001-2020. Lines indicate the trajectory of each single-model ensemble mean through 2100, with dots on each line indicating the ensemble mean during the five 20-year periods of the 21st century. The large grey square is the change in observed climate from 1961-1990 to 2001-2020 averaged across weather stations for the region. Trajectories further from the dotted grey lines (no change) indicate larger projected changes in summer precipitation (y-axis) and mean temperature (x-axis) or both. Model uncertainty is driven by differences in the global climate models (i.e. different colors) as well as differences in the individual runs of the same model (i.e. small circles of same color).

Socioeconomic uncertainty

The third major category of climate change uncertainty relates to future concentrations of greenhouse gas concentrations in the atmosphere that result from global emissions policies and socioeconomic development. The climate model projections used by CCISS follow scenarios of future greenhouse gas concentrations commonly referred to as Shared Socioeconomic Pathways (SSPs). The biogeoclimatic projections ensemble includes three major SSP scenarios: SSP1-2.6, SSP2-4.5, and SSP3-7.0 (Figure D2). SSP1-2.6 assumes strong emissions reductions (mitigation) roughly consistent with the goal of the Paris Climate Accords to limit global warming to 2ºC above pre-industrial temperatures. SSP2-4.5 assumes moderate mitigation and is roughly consistent with current emissions policies and economic trends. SSP3-7.0 is representative of a broader range of "baseline" scenarios that assume the absence of mitigation policies and is characterized by a linear increase in the rate of greenhouse gas emissions. Collectively, SSP1-2.6, SSP2-4.5, and SSP3-7.0 provide a reasonable representation of optimistic, neutral, and

pessimistic outlooks (respectively) on global GHG emissions reduction efforts (Hausfather and Peters 2020).

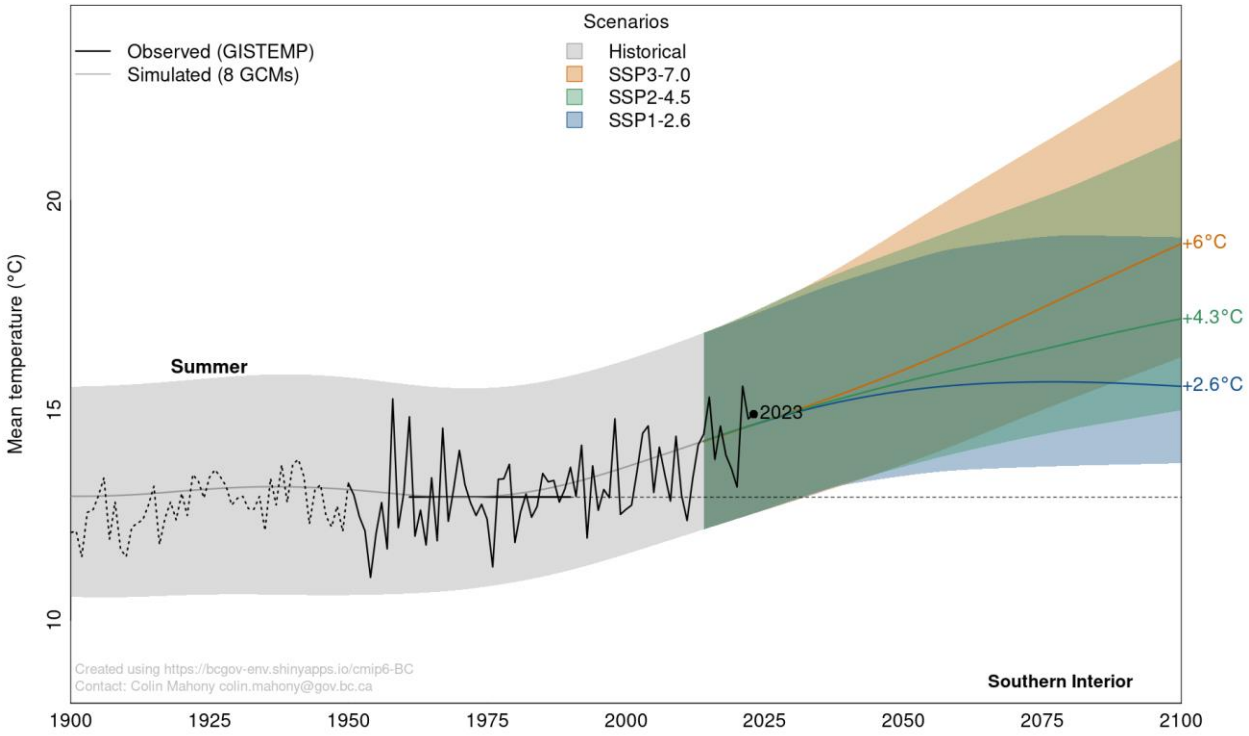

Figure D2. Projected change in summer mean temperature for the Southern Interior Ecoprovince of BC, showing the ensemble mean and range of the 8-model climate ensemble for the three greenhouse gas concentration scenarios used as a default setting in the CCISS tool.

D2. The Small Ensemble of GCM runs

We provide biogeoclimatic projections for five GCM simulations that represent the diversity of patterns and trends in climate change across the full ensemble of 60 GCM simulations. This small ensemble of representative simulations was selected using the KKZ algorithm (Cannon 2015), which selects the model simulations that span as large a range of climate changes as possible. We implemented this method using the mean of the changes in the biogeoclimatic model predictor variables from the 1961-1990 reference period to the 2081-2100 time period under the SSP2-4.5 scenario.

The five simulations provide a reasonable representation of the ensemble variation in temperature and precipitation changes (Figure D3). These simulations are not consistent in their position with respect to the rest of the ensemble; for example, the EC-Earth3 simulation has more warming than the ensemble mean in summer (Figure D3b) but less warming in winter (Figure D3c). Further, there is substantial spatial variation in precipitation change that can be obscured by the BC average (Figure D4 and Figure D5). For example, the MPI-ESM1 simulation has very strong reduction in summer precipitation in southern BC that is balanced by increases in

summer precipitation in Northern BC (Figure 4). For these reasons, it is not possible to characterize the simulations in simple contrasting extremes such as cooler/wetter vs. hotter/drier. Nevertheless, the five members of the small ensemble can be generically characterized as follows:

- MIROC6 r2—representative of the centroid (average) of the ensemble.
- MPI-ESM1-2-HR r1—least warming with regionally variable precipitation change (e.g., Figure 4).
- GISS-E2-1-G r2—lower-than-average warming with large precipitation increase.
- EC-Earth3 r4—higher-than-average warming with regionally variable precipitation change.
- ACCESS-ESM1-5 r1—highest warming with strong summer drying in southern interior BC.

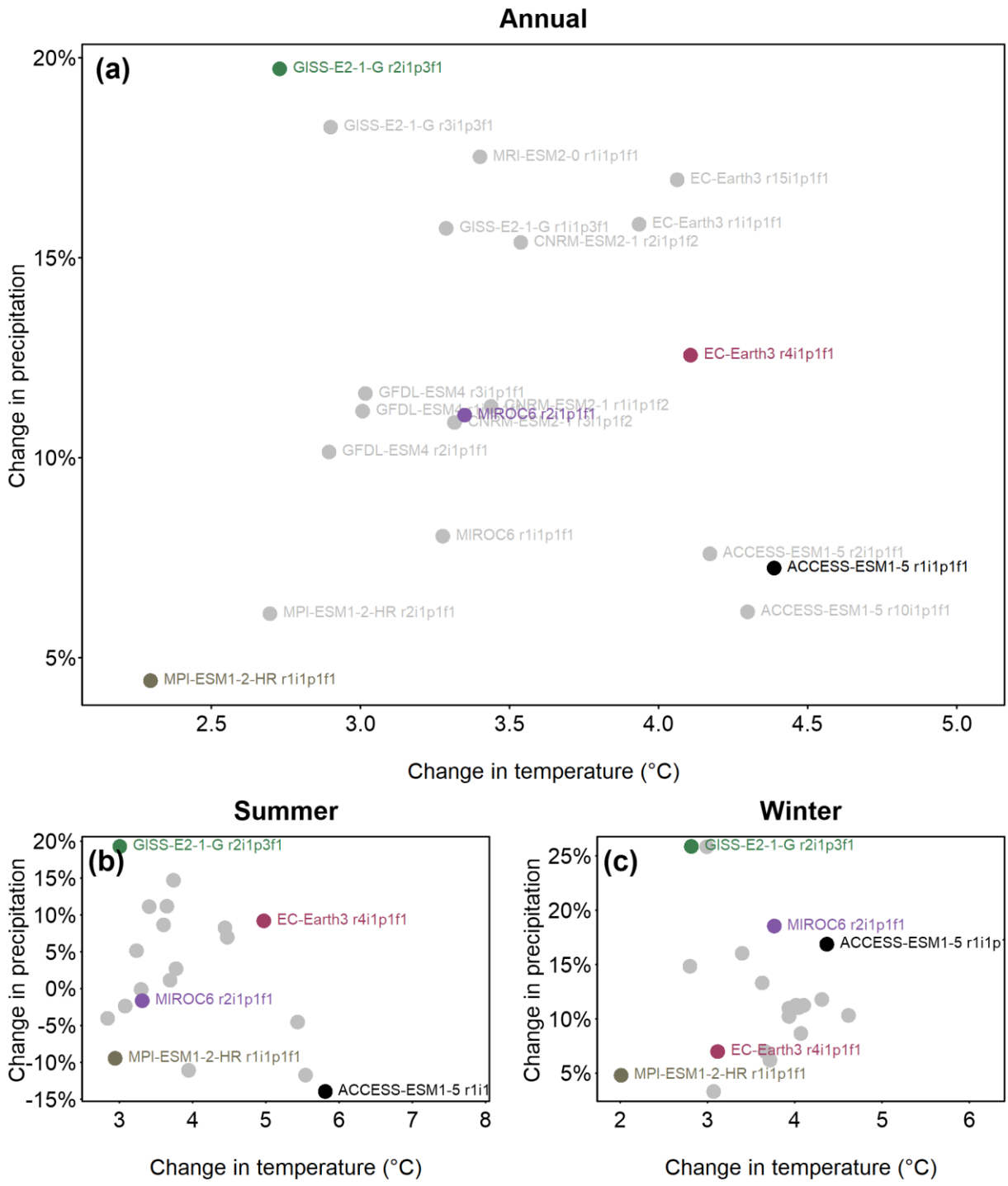

Figure D3. Change in annual and seasonal temperature and precipitation of the small ensemble of GCM runs, relative to the full CCISS ensemble. Changes are from the 1961-1990 reference period to 2081-2100 under the SSP2-4.5 scenario, and are averaged across British Columbia. (a) Change in mean annual temperature vs. annual precipitation. (b) Change in Summer mean daily maximum temperature vs. summer precipitation. (c) Change in winter mean daily minimum temperature vs. winter precipitation. Simulation names (e.g., r2i1p3f1) indicate the realization (run), initialization, physics scheme, and forcing scheme.

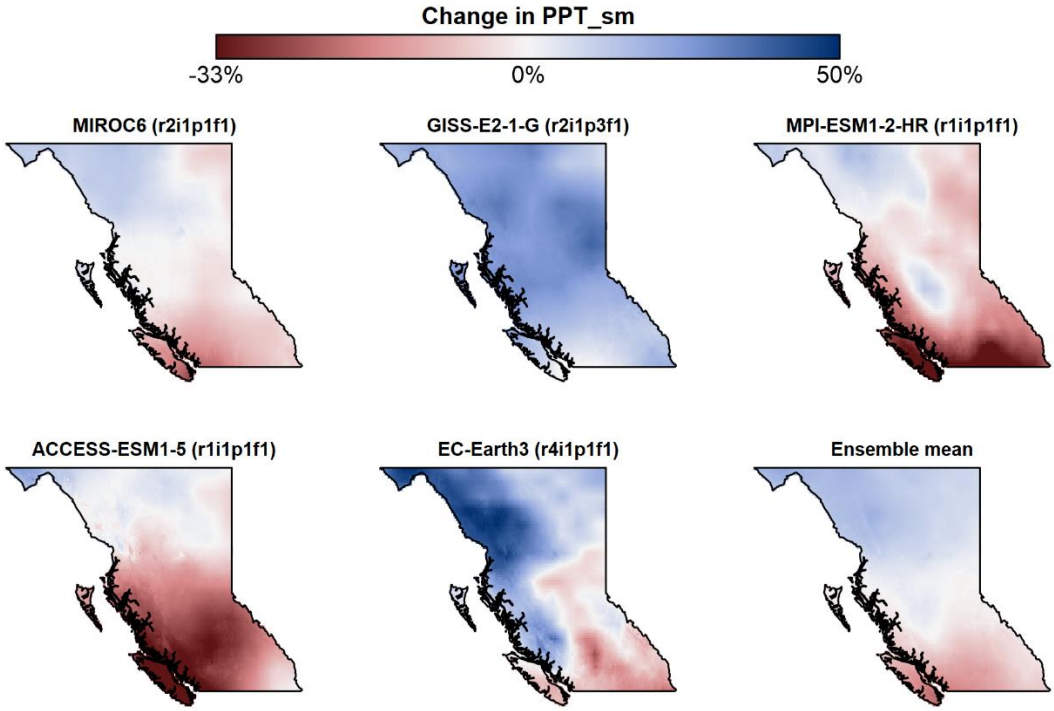

Figure D4. Spatial pattern of changes in summer precipitation in the small ensemble of GCM runs, and of the full ensemble mean. Changes are from the 1961-1990 reference period to 2081-2100 under the SSP2-4.5 scenario.

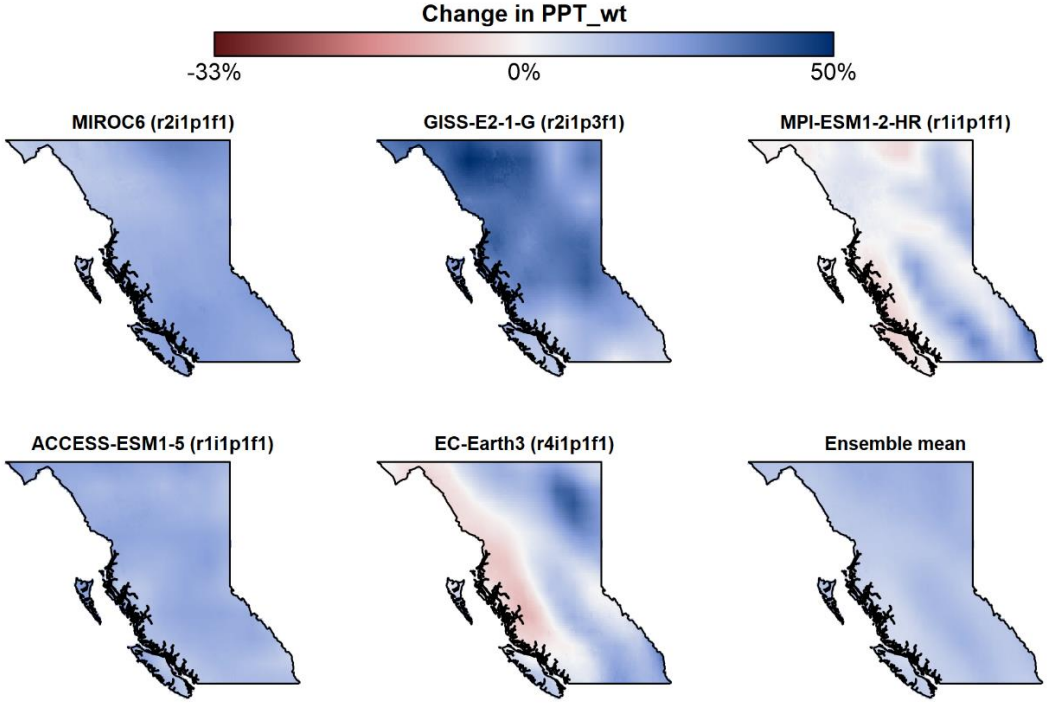

Figure D5. Spatial pattern of changes in Winter precipitation in the small ensemble of GCM runs, and of the full ensemble mean. Changes are from the 1961-1990 reference period to 2081-2100 under the SSP2-4.5 scenario.

D3. Literature Cited

Cannon, A. J. 2015. Selecting GCM scenarios that span the range of changes in a multimodel ensemble: Application to CMIP5 climate extremes indices. Journal of Climate 28:1260–1267.

Hausfather, Z., and G. P. Peters. 2020. Emissions - the “business as usual” story is misleading. Nature 577:618–620.

Mahony, C.R., T. Wang, A. Hamann, and A.J. Cannon. 2022. A global climate model ensemble for downscaled monthly climate normals over North America. International Journal of Climatology. 42:5871-5891. doi.org/10.1002/joc.7566