

Climate Data for Field Trials: Onsite Micro Stations Versus ClimateNA

Zhengyang Ye

Government of British Columbia

Gregory O'Neill

BC Ministry of Forests: British Columbia Ministry of Forests Lands Natural Resource Operations and Rural Development

Tongli Wang (✉ tongli.wang@ubc.ca)

The University of British Columbia Faculty of Forestry <https://orcid.org/0000-0002-9967-6769>

Research

Keywords: onsite weather station, response function, genealogy, ClimateNA, model prediction accuracy

Posted Date: September 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-898518/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Studies in diverse environmental fields require accurate climate data for point locations that are often distant from reliable public weather stations. 'Onsite' micro weather stations can be established directly at research locations, but purchase, establishment, and maintenance costs and data gaps can limit their feasibility. Alternatively, climate data for point locations can be predicted from ClimateNA, a publicly available software package, but the prediction accuracy in remote and mountainous locations is uncertain.

Results

We compared ClimateNA predictions with observations from onsite weather stations located at 11 interior spruce provenance trials in British Columbia, Canada. We found that ClimateNA predictions were highly accurate for temperature variables (average prediction error 0.77°C; most R^2 values > 0.99) but moderate for precipitation variables (average prediction error 27mm; $0.21 < R^2$ values < 0.58) when compared with onsite weather data (with random errors identified). Growth response functions developed with the two data sources showed similar patterns for temperature variables.

Conclusions

Our results suggest that 1) temperature variables can be accurately predicted at remote and mountainous locations using ClimateNA; 2) precipitation variables are more accurately predicted with ClimateNA than with onsite weather stations, which were considerably affected by random factors; and 3) response functions provide an effective, independent tool to assess alternative sources of climate data. Our results recommend the use of ClimateNA over onsite weather stations, except where highly accurate precipitation data are required, in which case, high-quality onsite weather stations must be established and carefully maintained.

1 Background

With the growing importance of climate-change-related studies in ecological and biological sciences, the demand for accurate climate data increases. This is particularly the case in genecology, which studies relationships between genetic variation among populations and their climatic origins to better manage genetic resources. Genecological analyses often require accurate climate data for remote test sites and provenance locations, distant from high quality, 'standard' government weather stations established and maintained by specialist crews. The need for climate data at remote locations gave rise to several publicly available climate models that interpolate climate data from standard weather stations to provide gridded climate estimates (*e.g.*, PRISM (Daly et al. 2008), WorldClim (Fick and Hijmans 2017), and PCIC

(Pacific Climate Impacts Consortium 2021)), with grid cells ranging from 800 m to 10 km. With grid cells of these sizes, differences in climate values between a grid cell average and specific point locations within a grid cell can be substantial in mountainous regions. Thus, 'onsite' micro weather stations (or onsite weather stations) are sometimes used when high-quality climate data is required at remote locations. However, the costs of the establishment, maintenance, and data downloading can be high. Also, maintenance can be irregular at remote sites, and data gaps are frequent due to vandalism, wildlife, and equipment failure (Krebs et al. 2018).

The recently developed software package ClimateNA (Wang et al. 2016) provides an alternative to onsite weather stations as it generates scale-free climate data for specific point locations for historical and future periods by interpolating from standard weather stations. Through dynamic local downscaling, the scale-free climate data has significantly improved accuracy over the gridded climate data and provides over 200 climate variables for any location in North America (Wang et al. 2016). The accuracy of ClimateNA predictions has been evaluated high against observations from standard weather stations over the entire continent of North America (Wang et al. 2016) and found to be compromised in remote areas or mountainous locations where standard weather stations used to develop and test ClimateNA are sparse, such as the Yukon (Krebs et al. 2018). Understanding the accuracy of climate data generated by ClimateNA versus climate data collected from the onsite weather stations will assist researchers in assessing the most appropriate option for obtaining climate data in remote locations.

Comparisons between the climate data from onsite weather stations and site-specific model predictions from ClimateNA are important to evaluate the accuracy of model-predicted climate data against observations and to determine the necessity of establishing onsite weather stations for local studies. There is only one study that compared ClimateWNA (a Predecessor of ClimateNA) predictions with observations from seven sites in forests in the southwestern Yukon (Krebs et al. 2018), where the density of weather stations used in developing the ClimateWNA baseline data was extremely low (Wang et al. 2012). Other studies compared climate models but did not evaluate them against observations from onsite weather stations (Jiang et al. 2018).

This study compared some major climate variables from the two data sources based on a series of provenance trials for interior spruce in western Canada. Provenance trials are common-garden experiment plantations that comprise populations originating from various provenances (or "origins") planted in one or more test environments. The original objective of provenance trials was to determine the optimal seed source for reforestation sites (Morgenstern 1996). More recently, information from provenance trials has been used to develop (1) transfer functions, which relate population performance and population transfer distance (Mátyás 1994; O'Neill et al. 2008), and (2) climate response functions, which reflect the performance of a population over a range of climate (*i.e.*, test site) conditions (Rehfeldt et al. 1999; Wang et al. 2006). British Columbia's interior spruce provenance trials provide the ideal material to evaluate the accuracy of ClimateNA predictions in mountainous regions in terms of climate variables themselves and their effect on the quality of response functions developed using the two climate data sources. It is

expected that response functions fitted with more accurate climate data will generate stronger relationships in what we call a “bio-assessment”.

We hypothesized that monthly temperature and precipitation variables predicted by ClimateNA and observations from the onsite weather stations are strongly correlated and show no bias. We further hypothesized that population response functions developed using predicted climate variables from ClimateNA and onsite weather stations would be similar. To test our hypotheses, we first evaluated the accuracy of ClimateNA predictions against observations from onsite weather stations in terms of their directly predicted values. Second, as prediction errors, more or less, are expected for all model predictions, including ClimateNA, we investigated if such errors considerably affect the performance of response functions developed using the climate data from the two sources.

2 Material And Methods

2.1 Provenance trials

Materials used in this study were from the Interior spruce climate change/genecology provenance trial (EP 670.71.12) of the British Columbia (BC) Ministry of Forests, Lands, Natural Resource Operations, and Rural Development (FLNRORD). The trial consists of 128 natural stand and genetically selected populations of Interior spruce (*Picea glauca* (Moench) Voss, *P. engelmannii* Parry ex Engelm., and their hybrids) from western North America, that were planted as 1-year old seedlings at 18 test sites (17 extant) in British Columbia, Alberta, and the Yukon Territory (Fig. 1, Table 1) in the spring of 2005. Each of the 17 test sites employed an incomplete block design containing 16 incomplete blocks of 9 four-tree row plots at 1 X 2m spacing in each of 8 replicates. Details regarding the populations, test sites, and experimental design can be found in O'Neill et al. (2014). Height and survival were assessed after 3, 6, and 10 years of growth in the field (*i.e.*, in fall 2007, 2010, and 2014). We use the term “population” to refer to a group of inter-mating individuals and “provenance” to indicate the geographic origin of a population. We selected three populations (N Rossland, Lower Canjilon, and McGregor, Fig. 1, Table 2) to sample the temperature range of the tested populations. Tree height increment between ages 6 and 10 (*i.e.*, fall 2010 to fall 2014) of each tree was used to calculate the mean increment of the 3 populations at each test site.

Locations of 15 test sites (with onsite weather stations, red and green dots) and 3 selected provenances (blue dots). The test sites used in this study are shown in green. The other 4 sites in red were excluded due to missing values and lack of adjacent sites for imputation

Table 1

Test site (with onsite weather station) geographical information (latitude, longitude, and elevation) and 6-year and 10-year tree height (average of the 3 selected populations) at each test site

Site name	Latitude (°)	Longitude (°)	Elevation (m)	Average 6-year tree height (cm)	Average 10-year tree height (cm)
Cranbrook	49.410	-115.958	1370	143	200
Duncan Lake	50.382	-116.922	640	199	269
Harrison*	49.342	-121.996	166	250	363
High Level	59.142	-117.568	334	121	153
Jordan River*	48.426	-124.023	120	217	295
Mayo*	63.543	-137.342	456	108	118
Nakusp	49.887	-117.890	1107	170	219
Parsnip	54.537	-122.028	805	168	266
Pine Pass	55.147	-122.776	811	169	275
Revelstoke	50.766	-117.958	910	195	266
Skimikin	50.784	-119.422	506	213	326
Terrace	54.485	-128.569	167	189	286
Tete Jaune	52.964	-119.417	820	187	312
Wells	53.155	-121.556	1230	131	183
Whitecourt*	54.056	-115.791	816	159	269
*Site excluded due to missing data periods and lack of adjacent test sites for the imputation of the missing periods					

Table 2

Geographical information (latitude, longitude, and elevation), mean annual temperature (MAT), and mean annual precipitation (MAP) from ClimateNA (Wang et al. 2016) for the three selected provenances

Provenance name	Latitude (°)	Longitude (°)	Elevation (m)	MAT (°C)	MAP (mm)
N Rossland	49.3	-118.00	1650	3.2	1036
Lower Canjilon	36.57	-106.33	2941	3.6	647
McGregor	54.67	-120.58	1219	1.0	962

2.2 Climate data from onsite weather stations

Onsite micro weather stations consisting of a precipitation gauge (S-RGB-M002), temperature/relative humidity sensor (S-THB-M002), and a HOBO Micro Station data logger (H21-002) from Onset Computer Corporation were established at 15 of the 18 provenance test sites in fall 2010 (one site was abandoned due to fire, one site was too remote to allow for annual data downloading, and one site did not establish). Temperature and precipitation were recorded hourly. Recording started in fall 2010 and data was available through 2018 for most test sites. We extracted weather records for January 1, 2011 to December 31, 2014 to match the age-6 to age-10 growth increment period (Table 1).

However, data gaps consisting of 9.6% of the total data recording period, which were generally 3 to 11 months in duration, were detected. The gaps arose due to vandalism, animal damage, dead batteries, or equipment malfunction, and did not appear to change in frequency over time. There were no data gaps in the first year of data records, and no gaps in any years of four test sites: Duncan Lake, Nakusp, Terrace, and Tete Jaune.

Three monthly climate variables were calculated from the observed hourly data: monthly minimum temperature, monthly maximum temperature, and monthly precipitation. For the comparisons and response function analysis, data gaps at the onsite stations were filled by imputation from other onsite stations that were geographically adjacent (< 150km) and at a similar altitude (difference < 300m). Missing temperature and precipitation values were imputed for months with missing data using values of the same month in the previous year and the following year. Data gaps could not be imputed at the Mayo, Jordan River, and Harrison test sites because they lacked geographically adjacent test sites. Thus, those 3 test sites were removed from the analysis. Also, the onsite weather station at test site Whitecourt did not record any rainfall from 2013 to 2016 and was therefore removed. Thus, 11 of the 15 sites were used for further analyses (see green marks in Fig. 1), and the onsite missing data rate decreased to 7.6% after the deletion of the 4 test sites. Because temperature in degree Celsius is an interval variable with no true zero, and precipitation records contained numerous 0 values which were not amenable to multiplication, imputation was achieved using the temperature and precipitation differences between adjacent years using the following formula:

$$P_{xij} (\text{°C or mm}) = \{[O_{yij} - O_{yi(j-1)}] + O_{xi(j-1)}\} + \{O_{xi(j+1)} - [O_{yi(j+1)} - O_{yij}]\}/2$$

Where P_{xij} is the value of the missing monthly climate variable to be imputed for month i of year j at site x , O_{xi} is the observed value of month i at site x . x identifies the test site for which a value is to be imputed, and y represents the neighboring test site.

Additional climate variables calculated from the primary monthly climate variables are sometimes found to be more strongly related to population differentiation than the primary monthly climate variables. These included mean annual temperature (MAT), mean coldest month temperature (MCMT), mean

warmest month temperature (MWMT), mean annual precipitation (MAP), annual heat-moisture index (AHM), average temperature in each season, and total precipitation in each season following Wang et al. (2016). Therefore, those additional climate variables were calculated for each month at each site using the onsite climate observations for use in the response function analysis.

2.3 Climate data from ClimateNA

Monthly and annual climate variables were predicted for 2011–2014 for the 11 test sites with ClimateNA v6.20 (Wang et al. 2016; ClimateNA 2020) using test site coordinates and elevation.

2.4 Climate data from Environment Canada standard weather stations

Since we noticed some random factors affecting the accuracy of onsite weather stations, especially for precipitation, we sought Environment Canada standard weather stations located < 20 km and within 100 m elevation of our onsite stations to assess the accuracy of the onsite and ClimateNA data as a reference. We assume climate to be very similar across these short distances. We found two standard stations that met these criteria. The Duncan Lake Dam standard weather station (50.239° N, 116.972° W, 549 m) is located 2 km from, and 22 m higher than, the Duncan Lake test site, while the Terrace A standard weather station (54.469° N, 128.578° W, 217m) is located 16 km from, and 91 meters below, the Terrace test site. Monthly climate data for the period 2011–2014 were downloaded for these two standard weather stations (Environment and Climate Change Canada 2021).

2.5 Statistical analysis

To examine the first hypothesis, we compared the monthly temperature and monthly precipitation of remote onsite weather stations and ClimateNA predictions. Comparisons were conducted for each test site individually, using linear regression analysis. Root mean square error (RMSE), the coefficient of determination (R^2 value), and P values were used to evaluate the prediction accuracy. Also, the same comparisons were conducted between monthly temperature and monthly precipitation from Environment Canada standard weather stations and the two climate data sources for the two locations with reference standard weather stations (Duncan Lake and Terrace), which, coincidentally, were 2 of the 4 test sites with complete weather data. Thus, we chose those two test sites for graphical demonstrations in figures.

To assess the second hypothesis, we developed and compared the patterns of the response functions for each of 3 climatically disparate provenances using a univariate quadratic function to relate tree height increment between ages 6 and 10 with observed and predicted climate variables. Quadratic functions are widely used in genecology because of their simplicity and effectiveness (Rehfeldt et al. 1999). The proportions of variation explained by the regression models (R^2 values) and prediction errors (RMSE) were used to determine the association between climate variables and growth. We also developed bivariate quadratic regressions to combine a temperature variable and a precipitation variable that performed well in the previous univariate regressions as the explanatory variables.

3 Results

3.1 Comparisons between onsite micro stations and ClimateNA

ClimateNA predictions of temperature were strongly related to observations from the onsite weather stations (most R^2 values > 0.99 and all P values < 0.0001). Average prediction errors (RMSE values) for monthly minimum and monthly maximum were 0.77°C ($0.48\text{--}1.12^\circ\text{C}$), as shown in Fig. 2 for Terrace and Duncan Lake test sites. Statistics for relationships at the remaining test sites are provided in Fig. S1 in additional files. Regression slopes were close to 1, and intercepts were near zero. All statistics suggested that ClimateNA predictions of temperature variables were highly accurate.

ClimateNA predictions of monthly precipitation were moderately well related to observations from the onsite weather stations (R^2 values ranged from 0.21 to 0.58). Prediction errors across all sites varied between 14.1–36.7 mm, with an average of 27.0 mm. Although monthly precipitation showed significant relationships between the observed and predicted values across all test sites, the regression lines deviated from the 1:1 line as shown in the last column of Fig. 2 (the slopes were 0.68 and 0.56 for Duncan Lake (c) and Terrace (f) test sites, respectively) and the range was much broader for ClimateNA predictions at Terrace site (20–240 mm vs. 0–180 mm), suggesting that precipitation was considerably overestimated across most of the test sites by ClimateNA. Plots showing the relationships for the rest of the sites are in Fig. S1 in additional files.

3.2 Comparisons with standard weather stations

For monthly temperature, both regression lines of onsite observations and ClimateNA predictions were close to the 1:1 line, with most of the R^2 values greater than 0.99 suggesting that both onsite and ClimateNA temperature values are unbiased (Fig. 3, a, b, d, e). Prediction errors from onsite observations were slightly smaller than ClimateNA predictions (average 0.53°C and 0.63°C for onsite observation and ClimateNA prediction respectively), suggesting that onsite weather stations performed slightly better than ClimateNA on temperature variables.

For monthly precipitation, both regression lines of onsite observations and ClimateNA predictions deviated from the 1:1 line. ClimateNA predictions were more strongly related to standard weather data than were onsite observations (Fig. 3, c, f) at both sites (average R^2 values of 0.537 and 0.283, respectively) and smaller prediction errors were calculated for ClimateNA than for onsite observations (average RMSE = 33.4 mm and 41.5 mm, respectively). Several exceptionally low monthly precipitation values (close to zero) were observed at the onsite station at the two sites, particularly at Duncan Lake, suggesting substantial recording errors at these onsite stations.

3.3 Comparisons using response functions

Results of the response function analysis are shown only for mean annual temperature (MAT) and precipitation in spring (PPT_sp) as these were the temperature and precipitation variables (respectively) most strongly related to height increment, and the values were relatively well distributed across the variable gradients (Fig. 4, Table 3).

Response functions developed with MAT were slightly stronger for onsite observations than for ClimateNA predictions (mean $R^2 = 0.507$ versus 0.380 when averaged across the 3 populations). For PPT_sp, response functions built using ClimateNA predictions explained a slightly larger amount of total variation in tree height increment than those build using onsite observations (mean $R^2 = 0.330$ versus 0.304 when averaged across the 3 population, Table 3). Despite the differences in statistics, the relationships showed similar patterns between using the two data sources for all the three provenances when using MAT as the predictor (Fig. 4, a, b, c). However, the range of the precipitation variable, PPT_sp, was much wider for the ClimateNA predictions (60–310 mm) compared with the onsite observations (30–230 mm), resulting in substantially different response curves (Fig. 4, d, e, f), likely due to the random error in the observed precipitation as mentioned previously.

Table 3

Results of univariate response function analyses. Relationships of population height increment with onsite observed and ClimateNA predicted site mean annual temperature (MAT) and precipitation in spring (PPT_sp) for 3 populations tested at 11 test sites

Provenances	Statistics	MAT (°C)		PPT_sp (mm)	
		Observed	Predicted	Observed	Predicted
N Rossland	R^2	0.506	0.359	0.302	0.348
	p	0.024	0.069	0.097	0.074
	RMSE	0.205	0.234	0.244	0.236
Willow Creek	R^2	0.425	0.387	0.308	0.284
	p	0.044	0.058	0.094	0.073
	RMSE	0.189	0.195	0.207	0.211
Lower Canjilon	R^2	0.345	0.316	0.284	0.300
	p	0.076	0.089	0.107	0.099
	RMSE	0.191	0.195	0.200	0.198

The bivariate response functions developed using combinations of MAT and PPT_sp as predictors were considerably stronger than univariate response functions, and stronger when built with ClimateNA climate data rather than with onsite observations (R^2 averaged across 3 populations = 0.765 versus 0.745 , respectively (Fig. 5). The patterns of the response surfaces along each climate variable remained

alike for both climate data sources, particularly in the dimension of the temperature variables, although the difference in the range of PPT_sp appeared as expected.

4 Discussion

Accurate climate data for specific locations are essential for ecological modeling and genecology. Cost-effectiveness is another major concern for climate data collection. Observations from standard weather stations provide the most accurate climate data, but they are mostly distant from the forest or ecological sites of interest. Onsite weather stations are widely applied to obtain local climate data, but their purchase and maintenance cost is high. ClimateNA can also predict climate variables for specific locations for free and has been widely used as well. Although ClimateNA has been evaluated against weather stations over entire North America (Wang et al. 2016), its accuracy remains a concern, particularly at remote and mountainous areas where standard weather stations, upon which ClimateNA is built, are sparse. In this study, we found that ClimateNA can generate climate variables at reasonably high accuracy, particularly for temperature variables. The response functions developed using the climate variables from the two sources showed similar patterns. Therefore, our results suggested that ClimateNA can serve as a competitive alternative to generate climate data.

4.1 The accuracy of ClimateNA predictions for temperature variables

We found that the predictions of temperature variables by ClimateNA to be highly accurate. Temperature-related climate variables are often found to be the primary drivers of local adaptation (Hu et al. 2019) and the major environmental factor affecting the performance of plant populations (Rehfeldt et al. 1999; Wang et al. 2006). Further, temperature is critical to moisture-related variables as it affects evapotranspiration, and increasing temperatures are expected to contribute significantly to future drought stress (Tam et al. 2019). Thus, the high accuracy of the ClimateNA predictions for temperature variables is important and encouraging.

For univariate response functions using temperature variables as predictors, we found MAT was the most important climate variable. MAT was widely used as an explanatory variable in previous studies (Rehfeldt et al. 1999; Wang et al. 2010). We found that the response curves derived from the two climate data sources were moderately strong and very similar in shape using MAT as the predictor. These results suggest that the use of such climate variables from ClimateNA is not likely to lead to bias in developing climate response functions of tree populations.

4.2 The accuracy of ClimateNA predictions for precipitation variables

Compared with temperature variables, precipitation variables predicted by ClimateNA were less accurate, which is in agreement with the results of a study testing local accuracy of ClimateWNA in the southern

Yukon (Krebs et al. 2018). The same situation also occurs in a study using gridded data to predict agricultural yields in the US (Mourtzinis et al. 2017). The team found that gridded weather models produced good predictions only for temperature-related variables, while the results for precipitation-related variables are not satisfying. These results indicated that precipitation is more difficult to predict than temperature, especially in mountainous areas.

In the present study, several problems were found regarding observed onsite precipitation data. Aside from missing data for both precipitation and temperature records due to uncontrollable reasons, we found that the recorder took precipitation record as 0 for more than a year at two test sites (Jordan River, 48.426°N, -124.023°W, on Vancouver Island and Whitecourt, 54.056° N, -115.791°, in Alberta), which is not realistic. Tipping buckets (also called “rain gauges”) were sometimes plugged by wasp and bird nests, causing the precipitation loggers to miss rainfall events, and no precipitation was recorded during periods of freezing weather because snow is not released from the tipping arm. These factors may have contributed to the underestimate of onsite observed precipitation detected when compared against precipitation records from adjacent standard weather stations (Fig. 3). Consequently, climate response functions developed using precipitation variables from the two climate data sources differed significantly in univariate response functions, suggesting that using precipitation data from onsite weather stations might lead to considerable bias in ecological modeling.

Bivariate response functions, using the two climate variables for building univariate response functions separately, significantly improved the goodness of model fit. However, the problem of the range difference in precipitation remained in the bivariate response functions. Importantly, we found no clear evidence that the observed data is better than the predicted data for establishing either univariate or bivariate response functions.

Overall, the highly accurate monthly temperatures and moderately accurate monthly precipitation predicted by ClimateNA are probably attributable to the scale-free feature of the model as it allows climate variables to be predicted for specific locations, whereas other climate data sources rely on grid cell averages. Results of this study suggest that ClimateNA can yield the same level of accuracy for temperature variables and a higher level of accuracy for precipitation variables compared to the observed data from the onsite weather stations in mountainous locations. Underestimates of onsite precipitation could lead to erroneous predictions of forest productivity when response functions are constructed using onsite precipitation. Therefore, particular caution needs to be practiced in using the precipitation data from onsite weather stations.

4.3 Cost-effectiveness of climate models comparing to onsite weather stations

Most temporary field research projects that require climate data use inexpensive ‘onsite’ weather stations maintained by non-specialist crews. Unlike standard weather stations operated by weather agencies, onsite stations generally lack wildlife and vandal exclusion fences, predisposing them to damage. And the high cost of accessing remote weather stations can limit the frequency with which they are

maintained. Together, these factors can contribute to lower quality data and greater data gaps compared with standard weather station data. The frequency of data gaps did not increase with station age, suggesting that equipment deterioration was not at fault, nonetheless, 11 of 15 stations had at least one data gap. All these problems combined make it harder and costlier to get precipitation data of high accuracy.

Climate data can be acquired free online from climate models such as ClimateNA, or for a fee considerably less than the cost of establishing and maintaining remote onsite stations. Also, climate models can predict climate variables for past and future periods and can derive biologically relevant climate variables. Nowadays, using climate models as a tool for tree growth modeling has become a major trend, and we believe those models like ClimateNA are valuable alternatives to onsite weather stations.

4.4 Limitations and reflection

We used ClimateNA to predict climate data at 11 geographically widespread test sites, most of which are located in remote or mountainous terrain where standard weather stations are sparse, and obtaining accurate predictions can be challenging. Nonetheless, we were able to develop moderately strong response functions, despite a relatively narrow climate range among the test sites (MAT – 1.1 to 7.2°C; PPT_sp 65 to 235 mm). For precipitation data, both ClimateNA and onsite weather stations have weaknesses. Thus, if high-quality local precipitation data are required, it may be necessary to establish high-quality onsite weather stations and maintain them carefully.

5 Conclusion

Results of this study indicate that temperature variables can be accurately predicted at remote and mountainous locations using ClimateNA and onsite weather stations. Precipitation variables are more accurately predicted with ClimateNA than with onsite weather stations. Given the greater accuracy of ClimateNA precipitation predictions, and its free availability, and lack of data gaps, we recommend its use over onsite weather stations, except where highly accurate precipitation data are required, in which case, high-quality onsite weather stations must be established and carefully maintained.

The use of response functions to evaluate the accuracy of alternative weather data sources appears to be effective and presents a novel, independent approach to assessing climate models, particularly where independent climate data are lacking. The response functions developed using ClimateNA data generated similar accuracy statistics and shapes as those developed with onsite climate data for temperature variables and even better for precipitation variables, suggesting that ClimateNA climate predictions are a reliable option to onsite climate data in genealogical and other environmental research requiring climate estimates in remote locations.

Abbreviations

MAT: mean annual temperature

MCMT: mean coldest month temperature

MWMT: mean warmest month temperature

MAP: mean annual precipitation

AHM: annual heat-moisture index

PPT_sp: precipitation in spring

Declarations

Availability of data and material:

Growth and climate data for the interior spruce provenance trials were provided by the Forest Improvement and Research Management Branch, Ministry of Forests, Lands, Natural Resource Operations and Rural Development, Province of British Columbia.

Conflicts of interest:

Not applicable.

Funding:

Funding for this study was provided by the Natural Sciences and Engineering Research Council of Canada – Discovery Grants program to Tongli Wang (RGPIN-2018-04643).

Authors' contributions:

TW conceived of the presented idea. GAO provided the provenance trials data. ZY imputed and analyzed the data. ZY wrote the manuscript with the support from TW. GAO provide suggestions and helped modify the language. All authors provided critical feedback and helped shape the analysis and manuscript.

Acknowledgement:

Not applicable.

References

1. Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP (2008) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Int J Climatol* 28:2031-2064. doi:10.1002/joc.1688.
2. Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302-4315. doi:10.1002/joc.5086.
3. Hu X, Mao J, El-Kassaby YA, Jia K, Jiao S, Zhou S, Li Y, Coops NC, Wang T (2019) Local adaptation and response of *Platycladus orientalis* (L.) Franco populations to climate change. *Forests*. doi:10.3390/f10080622.
4. Jiang Y, Kim JB, Still CK, Kerns BK, Kline JD, Cunningham PG (2018) Inter-comparison of multiple statistically downscaled climate datasets for the Pacific Northwest, USA. *Sci Data*. doi:10.1038/sdata.2018.16.
5. Krebs CJ, Henry JD, Kenney AJ, Hofer EJ (2018) Empirical test of the ClimateWNA model for local accuracy in the Kluane Lake area of the southern Yukon, Canada. *Clim Res* 75:111-115. doi:10.3354/cr01511.
6. Mátyás C. Modeling climate change effects with provenance test data (1994) *Tree Physiol* 14:797-804. doi:10.1093/treephys/14.7-8-9.797.
7. Morgenstern EK (1996) *Geographic variation in forest trees: genetic basis and application of knowledge in silviculture*. University of British Columbia Press, Vancouver.
8. Mourtzinis S, Rattalino Edreira JI, Conley SP, Grassini P (2017) From grid to field: Assessing quality of gridded weather data for agricultural applications. *Eur J Agron* 82:163-172. doi:10.1016/j.eja.2016.10.013.
9. O'Neill GA, Hamann A, Wang T (2008) Accounting for population variation improves estimates of the impact of climate change on species' growth and distribution. *J Appl Ecol* 45:1040-1049. <https://doi.org/10.1111/j.1365-2664.2008.01472.x>.
10. O'Neill GA, Stoehr M, Jaquish B (2014) Quantifying safe seed transfer distance and impacts of tree breeding on adaptation. *For Ecol Manage* 328:122-130. doi:10.1016/j.foreco.2014.05.039.
11. Rehfeldt GE, Ying CC, Spittlehouse DL, Hamilton DA Jr (1999) Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecol Monogr* 69:375-407. doi:10.2307/2657162.
12. Tam BY, Szeto K, Bonsal B, Flato G, Cannon AJ, Rong R (2019) CMIP5 drought projections in Canada based on the Standardized Precipitation Ecapotranspiration Index. *Can Water Resour J* 44:90-107. doi:10.1080/07011784.2018.1537812..
13. Wang T, Hamann A, Spittlehouse D, Carroll C (2016) Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS ONE* doi:10.1371/journal.pone.0156720.
14. Wang T, Hamann A, Spittlehouse DL, Murdock TQ (2012) ClimateWNA—high-resolution spatial climate data for western North America. *J Appl Meteorol Climatol* 51:16-29. doi: 10.1175/JAMC-D-11-043.1.

15. Wang T, Hamann A, Yanchuk A, O'Neill GA, Aitken SN (2006) Use of response functions in selecting lodgepole pine populations for future climates. *Glob Chang Biol.* 12:2404-2416. doi:10.1111/j.1365-2486.2006.01271.x.
16. Wang T, O'Neill GA, Aitken, SN (2010) Integrating environmental and genetic effects to predict response of tree populations to climate. *Ecol Appl* 20:153-163. doi:10.1890/08-2257.1.
17. Pacific Climate Impacts Consortium (2021). <https://www.pacificclimate.org>. Accessed 10 Sept 2021.
18. ClimateNA (2020). <http://www.climate-na.ca>. Accessed 10 Sept 2021.
19. Environment and Climate Change Canada (2021) Historical Data. https://climate.weather.gc.ca/historical_data/search_historic_data_e.html. Accessed 10 Sept 2021.

Figures



Figure 1

Locations of 15 test sites (with onsite weather stations, red and green dots) and 3 selected provenances (blue dots). The test sites used in this study are shown in green. The other 4 sites in red were excluded due to missing values and lack of adjacent sites for imputation

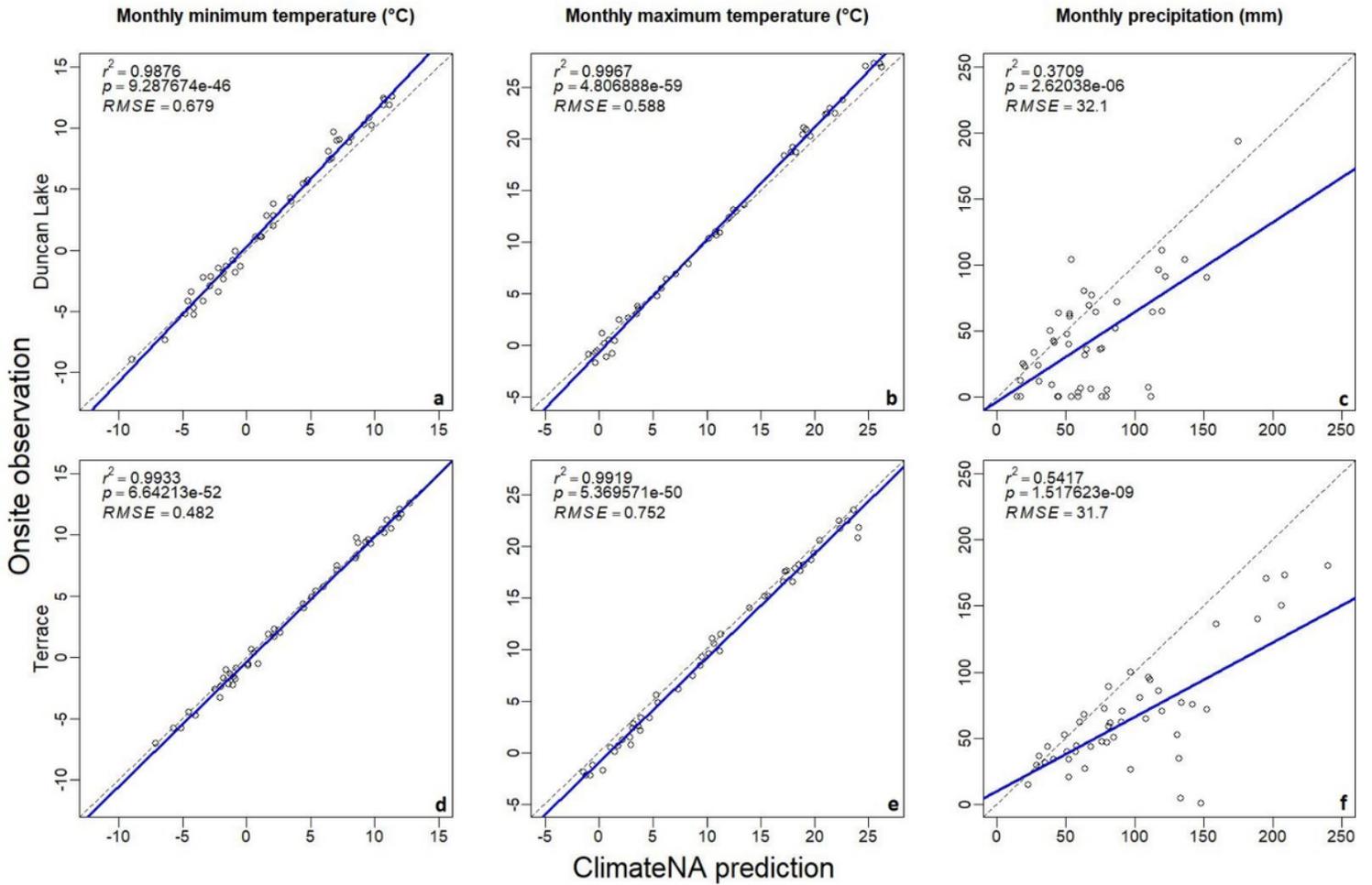


Figure 2

Relationships between observed and predicted (by ClimateNA) monthly minimum, monthly maximum, and monthly precipitation at the Duncan Lake (the upper row, a, b, c) and the Terrace (the lower row, d, e, f) test sites from 2011 to 2014. The blue lines are regression lines, the dashed lines are the 1:1 line. There are 48 (= 12 months X 4 years) points per scatterplot

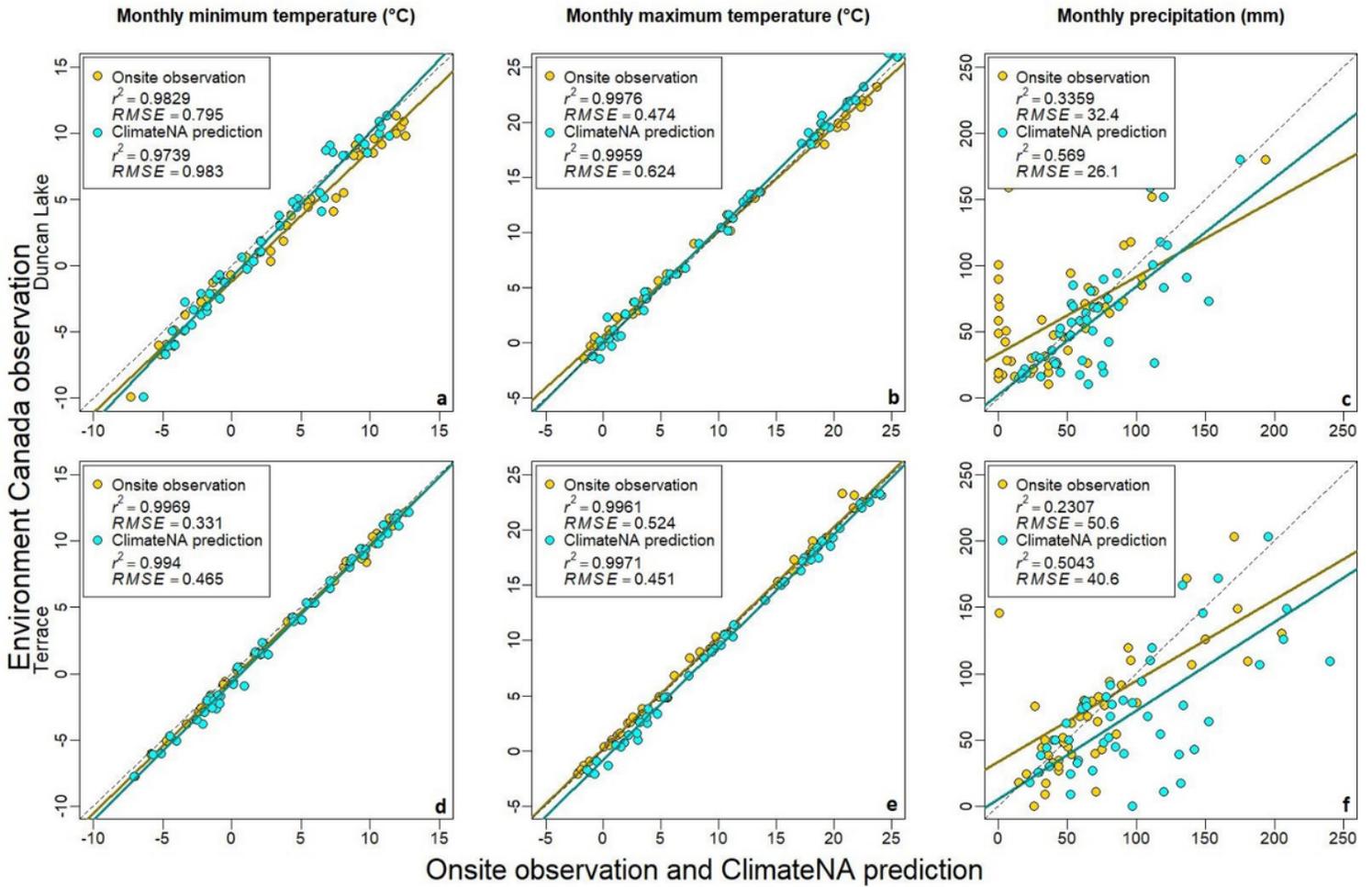


Figure 3

Comparisons of onsite observations and ClimateNA predictions with the standard weather station data for monthly minimum and monthly maximum temperature, and monthly precipitation at the Duncan Lake (the upper row, a, b, c) and the Terrace (the lower row, d, e, f) test sites from 2011 to 2014. The gold and cyan lines are regression lines representing onsite observations and ClimateNA predictions, respectively, the dashed lines are the 1:1 line. There are 96 (= 2 data sources X 12 months X 4 years) points per scatterplot

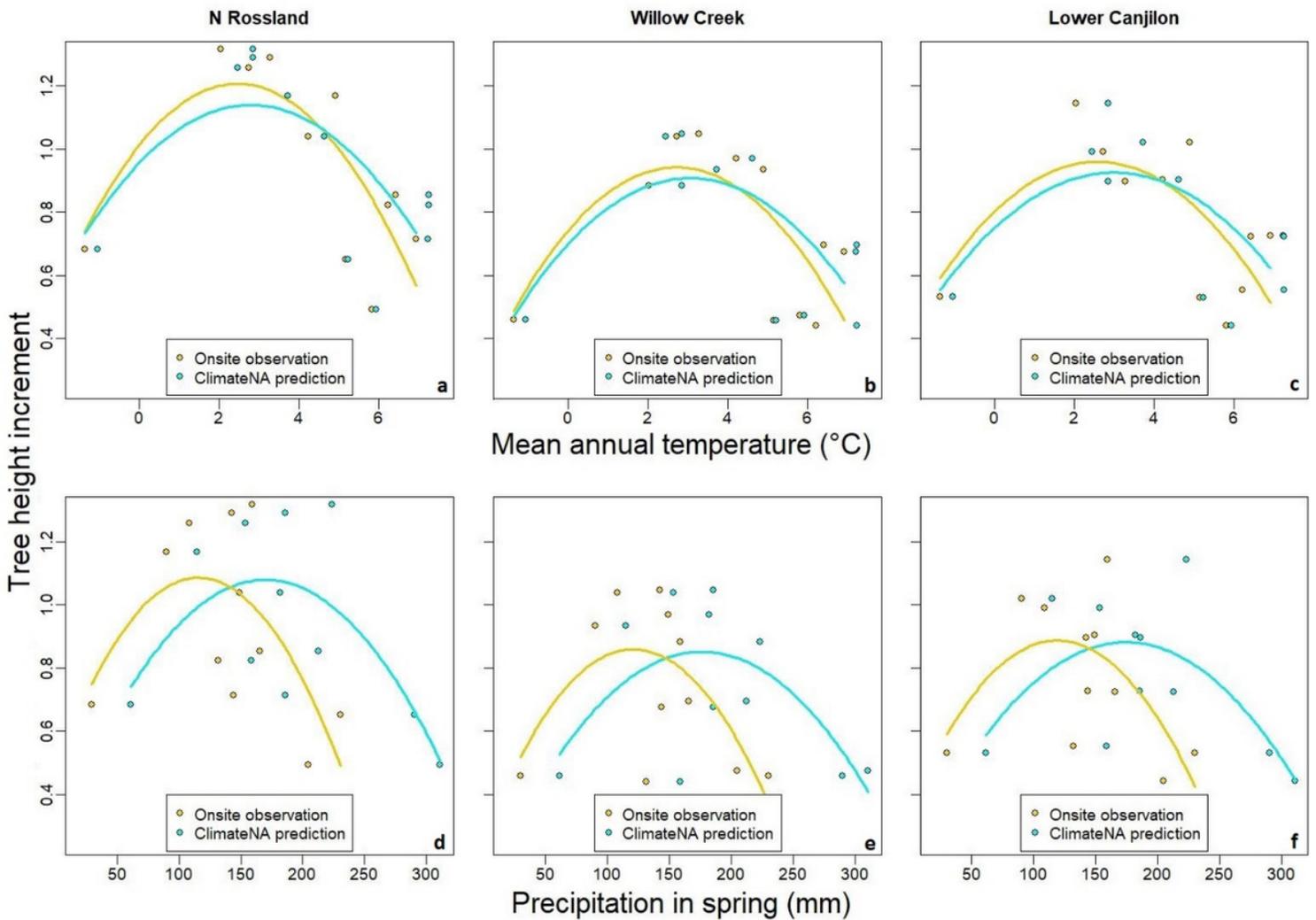


Figure 4

Univariate response functions of height increment in response to onsite observed and ClimateNA predicted mean annual temperature (MAT, a, b, c) and precipitation in spring (PPT_sp, d, e, f) for 3 populations (N Rossland, Lower Canjilon, and McGregor) from 2011 to 2014. There are 22 (= 11 test sites X 2 provenances/test site) observations for each plot, except for population McGregor which was not planted at two test sites (Revelstoke and Tete Jaune)

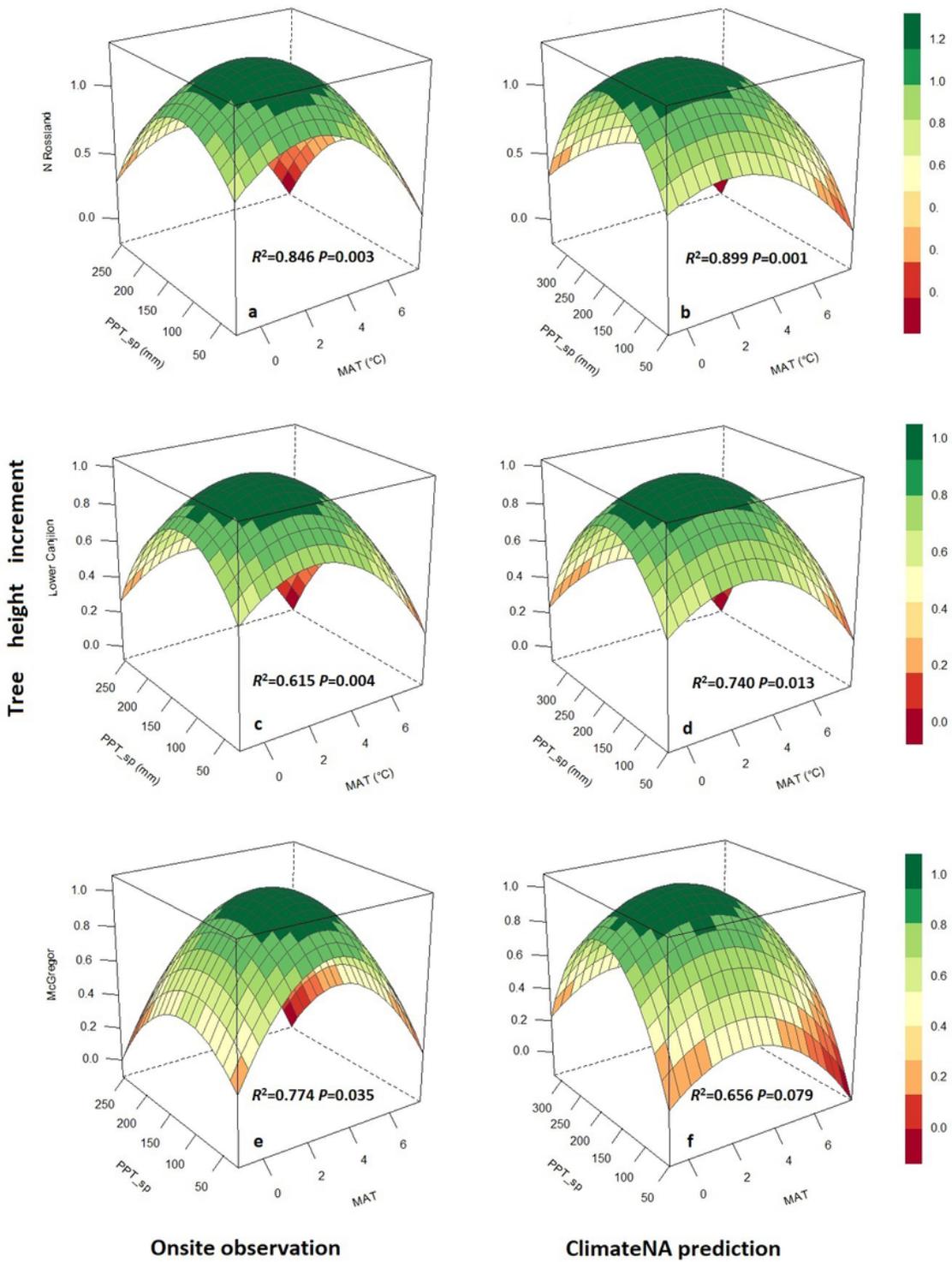


Figure 5

Bivariate response functions of height increment in response to onsite observed (left, a, c, e) and ClimateNA predicted (right, b, d, f) mean annual temperature (MAT) and precipitation in spring (PPT_sp) for 3 populations (N Rossland, a, b, Lower Canjilon, c, d, McGregor, e, f) from 2011 to 2014 based on 11 test sites

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additional1.docx](#)