# Regional Climate Model Simulation of Surface Moisture Flux Variations in Northern Terrestrial Regions

Ross Fischer[1], John E. Walsh[2], Eugénie S. Euskirchen[3], and Peter A. Bieniek[2]

[1]University of Colorado, Boulder, CO
[2]International Arctic Research Center, University of Alaska, Fairbanks, AK
[3]Institute of Arctic Biology, University of Alaska, Fairbanks, AK

Corresponding author:

John Walsh

International Arctic Research Center

University of Alaska, Fairbanks, AK 99709

jwalsh@iarc.uaf.edu

Submitted to the ~~Journal of Hydrometeorology~~Atmospheric and Climate Sciences

~~September~~ December 2017

## ABSTRACT

The wetness of high-latitude land surfaces is strongly dependent on the difference between precipitation (P) and evapotranspiration (ET). If climate models are to capture the trajectory of surface wetness in high latitudes, they must capture the seasonality, the variations and the drivers of variations in the surface moisture fluxes. In this study, a~~Variations of high-latitude ET over daily, seasonal, and interannual timescales are poorly documented, as are their relationships to key drivers such as precipitation and temperature, T. A~~ combination of regional climate model output and eddy covariance measurements from flux tower locations in Alaska is used to evaluate model simulations of the surface moisture fluxes and their variations. In particular, we use the model output and the field measurements to test the ~~test the~~ hypothesis that temperature is the key driver of variations of ET in tundra regions underlain by permafrost, while precipitation plays a greater role in boreal forest areas. Although the model depicts a stronger hydrologic cycle (larger P, larger ET) relative to the *in situ* measurements at all the sites, the strong seasonal cycles of P, T, and ET are captured by the model. The tower measurements from all sites show a short period (one or two months) of negative P-ET during summer, indicative of surface drying, although the model does not show this period of drying in the tundra region in the foothills of the Brooks Range. At all the tundra sites, both the flux tower data and the model simulations show that daily and warm-season totals of ET are largely temperature-driven. Daily ET shows a weak negative correlation with precipitation ~~at~~ in the measurements and in the model simulations of all the sites~~all the sites~~. Precipitation is the main driver of year-to-year variations of the seasonally integrated net moisture flux at all the sites, implying that precipitation will be at least as important as temperature in the future trajectory of surface wetness.

*Keywords*: Evapotranspiration, precipitation, Arctic, tundra, boreal forest, moisture budget

## 1. Introduction

Some of the largest and most consequential uncertainties in the trajectory of the Arctic climate system are associated with the hydrology of the Arctic terrestrial surface. The challenge facing the research community is to provide a scientifically sound response to the fundamental question: Will Arctic landscapes become wetter or drier as climate changes? Because this question pertains to the future, model simulations must be relied upon for guidance. While there are important scaling issues ~~that are discussed below~~associated with the validation and use of model output, in situ measurements of surface moisture flux measurements in recent decades now makes it possible to perform direct comparisons of measured and model-derived fluxes at a variety of high-latitude terrestrial sites.

~~a~~ At the heart of the model validation issue is the ability of the model to capture the net surface moisture flux, which is the difference between ~~major uncertainty surrounds the surface moisture budget and the relationship between~~ precipitation (P) and evapotranspiration (ET). If P (including both rain and snow) exceeds ET over a period of time, the excess goes into runoff or storage. If ET exceeds P, the surface moisture deficit leads to drying unless there is sufficient recharge from below. A drying surface leads to decreased water supplies, increased wildfire risk, and moisture stress on vegetation, all of which have consequences for terrestrial ecosystems and human activities.

~~In a recent synthesis of Arctic terrestrial hydrology, Bring et al. (2016, p. 642) conclude that "We are still lacking a comprehensive picture of evapotranspiration across Arctic regions and how it will interact with climate and landscape changes". Arctic hydrology is constrained by permafrost in its northern subregions, and changes in permafrost as well as vegetation will undoubtedly affect the trajectory of Arctic terrestrial hydrology, particularly the wetness of the surface. This paper focuses on evapotranspiration, especially variations relative to precipitation and temperature (T), two of the key drivers of terrestrial hydrology.~~

The most comprehensive assessment of recent P and ET trends based on historical data and model simulations appears to be that of Rawlins et al. (2010), who used a variety of precipitation datasets, atmospheric reanalyses, land surface model output, and global climate models. Trends in P, ET, P-ET and river discharge were generally positive in the observational data, for which record lengths ranged from 20 to 50 years. However, trends of P-ET, computed as differences between historical P datasets and satellite-derived (AVHRR GIMMS) ET, showed no significant trend. The nine global climate models examined by Rawlins et al. (2010) showed statistically significant trends of *terrestrial* pan-Arctic P-ET over the period 1950-1999 in eight of the nine cases, and in all nine cases for the period 1950-2049. Trends for the historical period were smaller than for the future period in the climate model output. All results were for annual means. The positive trends in annual mean P-ET contrast with the expectation that longer and warmer summers will increase ET sufficiently to favor summer drying, which is indicated by projected decreases of high-latitude soil moisture in major climate change assessments such as the IPCC (2013, Figs. 11.14 and 12.23). Anticipated increases of high-latitude wildfire activity (Partain et al., 2016; Flannigan et al., 2015) are consistent with this expectation, highlighting the mixed picture of future surface wetness trends in the Arctic.

Laine et al.'s (2014) more recent evaluation of global climate model projections, although global rather than Arctic in scope, highlights the challenge of assessing changes in high-latitude surface wetness. While Laine et al. found that the projected changes of ET also show the expected pattern of increases over Arctic land areas, primarily during summer, the projected changes of P-ET over the Arctic were much less spatially coherent and less robust than the changes of P and ET separately. In particular, summer P-ET is projected to decrease over northern Canada, increase over Alaska, decrease over the western and northern Eurasian subarctic, and increase over parts of northeastern Russia, including Chukotka. Over most of these areas, the sign of the projected change was not robust across the models at the 95% confidence level. The spatial pattern over high-latitude land areas is very consistent with the projected changes in soil moisture obtained by Dirmeyer et al. (2014) using 15 of the same models. Dirmeyer et al.'s Figure 1 shows reductions of summer soil moisture over northern Canada and north-central Russia, but not over Alaska and eastern Siberia, from the preindustrial to the 20$^{th}$ century. The results of these studies highlight the uncertainty in the trajectory of surface wetness in Arctic land areas in the present generation of global climate models. Given this uncertainty and the reliance on models for anticipation of future changes in surface wetness, it is imperative to (1) understand the drivers of variations and changes in high-latitude ET, and (2) assess the ability of models to capture the relationship between the atmospheric drivers and variations of ET. The present paper addresses both of these needs.

The modeling studies highlighted above were based on global climate model simulations. There has been little evaluation of corresponding simulations by regional climate models, which offer several advantages relative to global climate models. First, regional climate models enable finer resolution, by up to an order of magnitude, relative to global climate models. Second, when

driven at the lateral boundaries by historical reanalyses, regional climate models are constrained to observations, at least at the lateral boundaries, while global models are freer to drift to their own model climatologies. For these reasons, the present study makes use of a regional climate model driven at the lateral boundaries by an atmospheric reanalysis.

A key measure of the validity of the ET simulations by climate models is their ability to capture the sensitivity of ET to variations of precipitation and temperature. During the growing season, ET can be expected to increase with temperature in biomes that are not moisture-limited (e.g., tundra underlain by permafrost). One might also expect ET to respond positively to precipitation events in areas where ET is moisture limited (e.g., permafrost-free boreal forest sites). The expectation of influences by temperature (T) and precipitation (P) is supported by variations in yearly ET, T, and P at tundra sites in Alaska (Euskirchen et al., 2012) and Canada (Lafleur and Humphreys, 2007). Other variables undoubtedly play a role in temporal variability of ET. For example, solar and longwave radiation are included in ET formulations (e.g., Penman-Monteith) and are important drivers of Arctic ecosystem production (e.g., Lafleur et al., 2012). Changes in subsurface moisture, including moisture made available by permafrost thaw, can also be an important driver (Ohta et al., 2008). However, because direct measurements of these quantities are greatly limited compared to T, P, and ET, the diagnostic evaluation in the present study focuses on P and T as key drivers. Against this backdrop of prior studies, hypotheses and data constraints, we address two key questions:

1) How well are the surface moisture fluxes~~, variations, and relationships to key drivers~~ over timescales of days to seasons captured by a state-of-the-art regional climate model?

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

2) ~~How strongly do temperature and precipitation influence variations of~~Are the relationships between temperature, precipitation and surface moisture ~~fluxes over daily to seasonal timescales?~~ flux variations reproduced by the regional climate model?

~~2.~~ Both these questions point to the paper's emphasis on model validation. Because inter-variable relationships are part of the validation in (2), an assessment of processes and drivers is implicitly part of our model evaluation strategy. In that respect, the following sections include discussion of the drivers and processes relevant to variations of the surface moisture budget. In all cases, however, the consistency (or lack thereof) between the simulated and observed relationships is the underlying thread of the discussion.

**Formatted:** Normal, Indent: First line: 0.31", No bullets or numbering

**Formatted:** Font: (Default) Times New Roman, 12 pt

~~3. How well are the surface moisture fluxes, variations, and relationships to key drivers captured by a state-of-the-art climate model?~~

The study focuses on the tundra and boreal forest biomes of Alaska, where *in situ* measurements of the moisture fluxes and drivers are available for several sites in each biome. A regional climate model provides the simulated fluxes and driving variables for the same years and locations. In this study we refer to temperature and precipitation as "drivers" of ET and P-ET only in a proxy sense. In reality, ET is driven by radiative fluxes (for which temperature may be regarded as a proxy), boundary layer stability (which is also affected by surface air temperatures), relative humidity and vertical humidity gradients (which are affected by precipitation), and wind speed. Our emphasis on temperature and precipitation is dictated in part by the availability of these variables, especially from the nearby weather observing sites used for the infilling of moissing values at the tower sites.

Both of these sources of hydrologic information are described in Section 2Section 2 describes the two sources of hydrologic information: the regional climate model and the in situ measurements. The Section 3 describes the processing of the model output and the *in situ* measurements, including their quality-control is, described in Section 3. Section Section 4 presents the results in terms of seasonal climatologies as well as interannual variations derived from both sources. The relationships of ET and P-ET to the atmospheric drivers are evaluated in Section 5 for seasonal and daily timescales. Discrepancies between the observational data and the model results are highlighted in both Sections 4 and 5. Section 6 summarizes the primary conclusions.

## 2. Sources of data

### *2.21. Regional climate model output*

Global model output is available from several dozen climate modeling centers, and the hydrologic output from the current generation of these models (CMIP5, Coupled Model Intercomparison Project, version 5) has been evaluated by Rawlins et al. (2010), Laine et al. (2014) and Yao et al. (2016). For comparison with site-specific measurements such as those in Section 2.1, , finer resolution of model output is highly desirable. Dynamical downscaling by a regional climate model can achieve finer resolution by approximately an order of magnitude. In this study, we analyze P, T, and ET from a widely used regional climate model, the Advanced Research (ARW) core of the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008). The model makes use of a thermodynamic sea ice model (Zhang and Zhang 2001) and the Noah land-surface model (Yang et al., 2011) used within WRF in order to model the thermal conditions at the surface. It is driven at the lateral boundaries by observationally-based

**Formatted:** Indent: Left: 0", Numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 2 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

**Formatted:** Font: (Default) Times New Roman, 12 pt, Italic

**Formatted:** Font: (Default) Times New Roman, 12 pt, Italic

**Formatted:** Normal, No bullets or numbering

**Formatted:** Font: (Default) Times New Roman, 12 pt

reanalysis fields (ERA-Interim) in order to avoid introduction of additional biases arising from the driving model. The ERA-Interim reanalysis was selected as it has been successfully downscaled using WRF in other regions of the world (e.g. Gao et al. 2015; Srivastava et al. 2014, 2013; Soares et al. 2012), has been used in previous Arctic WRF simulations and analyses (e.g. Liu et al. 2014) and is among the best performing reanalysis data sets for Alaska (Lader et al. 2015) and the wider Arctic (Lindsay et al. 2014).

As described by Bieniek et al. (2016), the downscaling covered a domain with 262x262 grid points that encompassed all of Alaska and portions of far eastern Russia and northern Canada at 20 km horizontal resolution (Figure ~~4~~1) with 49 vertical model levels. The model was reinitialized every two days and was integrated for a total of 54 hours after each reinitialization. Each initialization occurred at 18 UTC (09 AKST). After each reinitialization, the first 6 hours of output (a "spin-up" overlapping with the final 6 hours of the previous 54-hour integration) were discarded. The frequent reinitialization to the ERA-Interim reanalysis, together with a continuous nudging to the ERA-Interim upper air fields during each 54-hour simulation, ensured that the model was tightly constrained by the observationally-based reanalysis.

The output from the 2-day simulations was combined together to form the final downscaled product, which spanned the 1979-2015 period. Hourly WRF output was saved and used to produce daily mean values of downscaled variables. The daily ET values were calculated from the daily values of latent heat flux in the archive of model output. (The same conversion was performed with the ~~*eddy covariance*~~*in situ* measurements described below). The daily mean values of T, P, and ET from the model's 20-km grid cells corresponding to the tower sites of Section 2.1 were extracted for the years of the available ~~tower data (Section 3).~~validation data.

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Font: Italic

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Font: (Default) Times New Roman, 12 pt

~~4.~~

Formatted: Font: (Default) Times New Roman, 12 pt, Bold

## ~~2.1~~2.2 *In situ measurements*

Eddy covariance estimates of ET based on flux tower measurements are available for a network of sites in the Arctic and subarctic land areas. Many of these are accessible from the archives of the AmeriFlux database, hosted by the U.S. Department of Energy's Lawrence Berkeley Laboratory (http://ameriflux.lbl.gov/). The eddy covariance estimates of the fluxes of moisture (as well as energy and trace gases such as $CO_2$) are based on measurements made several meters to tens of meters above the surface (Figure ~~1~~2). In addition to instrumentation for measuring vertical fluxes, the towers include measurements of temperature and precipitation, enabling evaluations of the relationships between ET, P, and T presented in Section 5. The horizontal footprint of the measurements is typically several tens to 100-200 meters, so the measured fluxes are representative of vegetation in the immediate vicinity of the measurement site. However, one of the main limitations of a comparison of tower measurements and climate model output is that the footprint of the tower measurements is orders of magnitude smaller in scale than the grid cells of climate models. For this reason, we stress the relationships between ET flux variations and the associated drivers (T and P) computed separately from each of the two information sources. ~~Recognizing the~~The scale discrepancy~~, we also choose for our comparison~~ is also the main motivation for our use of a regional climate model ~~with higher resolution~~rather than a global climate model~~s (Section 2.2).~~ .

Despite the scale discrepancy, several recent studies have compared tower measurements of ET with corresponding model output. Yao et al. (2016) evaluated ET fluxes simulated by global climate models (with resolutions of 100-200 km) against tower measurements at 240 globally

distributed sites. The aggregate statistics computed by Yao et al. were heavily weighted towards the stations outside the Arctic. Tower measurements at Canadian sites were used by Wang et al. (2015) in a Canada-wide comparison with water budgets, remote sensing products and a land surface models. These studies did not address relationships between ET and the driving variables. These relationships are the focus of the present paper.

In order to assess the moisture fluxes and their drivers across Arctic and subarctic biomes, we present results for a set of four locations in Alaska: two on the tundra (Barrow and Imnavait Creek) and two in the boreal forest (Bonanza Creek and Poker Flat). Figure ~~2~~ 3 shows the locations of the four sites. The Barrow site is located at the Barrow Environmental Observatory approximately midway between Barrow on the Chukchi Sea coast and the Beaufort Sea (Eilson Lagoon). The coastline is about 3 km to the northeast and 5 km to the northwest of the tower site. Winds from the west, north, and east are onshore, advecting marine air over the site; northeast winds predominate during the summer. The tundra vegetation at Barrow is a mixture of vascular plants such as sedge and nonvascular constituents such as moss and lichens.

The Imnavait site is approximately 200 km south of the Arctic Ocean at the base of the foothills of the Brooks Range near the Arctic Long-Term Ecological Research (LTER) field station at Toolik Lake. While maritime air occasionally impacts the site during summer, the climate is considerably warmer than at Barrow: annual (summer) temperatures are -8ºC (+9ºC) at Toolik/Imnavait and -11ºC (+4°C) at Barrow. Cloudiness is also less pervasive at Imnavait than at Barrow. We use data from two flux tower sites at Imnavait. The 200 m footprint of the "sedge" site, located in the valley bottom, is classified as 52% wet sedge and 47% tussock tundra, with the remainder bare soil or open water. The footprint of the "heath" site is classified

as 20% heath, but also includes 72% tussock tundra, with the balance made up of sedge meadow and bare soil. Figure 3 shows the vegetation in the vicinity of the two Imnavait tower sites.

Both the two Imnavait sites and the Barrow site are underlain by continuous permafrost that is several hundred meters deep. During summer the maximum active layer depths are approximately 40-60 cm at the Imnavait sites and 20-30 cm at Barrow. Snow typically covers the ground from October through May, with typical snow depths of 30-40 cm at the start of the spring thaw, although considerable blowing and drifting results in a heterogeneous distribution of snow depths.

The boreal forest tower sites are at Bonanza Creek and Poker Flat, located about 30 km southwest and 50 km northeast of Fairbanks, respectively. Both locations are considerably warmer than the tundra sites in summer, with June-August mean temperatures of approximately 15ºC. Winter temperatures are comparable to those at the tundra sites, although with considerably less wind, and the spring snowmelt occurs 4-6 weeks earlier than on the ~~tundra~~ tundra. The Bonanza Creek tower, part of the Bonanza Creek LTER, is located in a lowland area underlain by permafrost. The vegetation is mature black spruce forest on a permafrost plateau (Figure 3). Poker Flat is an upland area of discontinuous permafrost, although the tower itself is underlain by permafrost. Black spruce is the dominant vegetation type, ~~although there are~~with some white spruce and birch in the area.

For purposes of this study, a key limitation of the tower measurements is the lack of useful data during the winter months when snow and icing preclude measurements of P and ET. For this reason, our study focuses on the warm season, May through September, when the full suite of measurements (T, P, ET) can be documented, analyzed, and compared with the model output.

## ~~2.2 Regional climate model output~~

~~Global model output is available from several dozen climate modeling centers, and the hydrologic output from the current generation of these models (CMIP5, Coupled Model Intercomparison Project, version 5) has been evaluated by Rawlins et al. (2010), Laine et al. (2014) and Yao et al. (2016). For comparison with site-specific measurements such as those in Section 2.1, finer resolution of model output is highly desirable. Dynamical downscaling by a regional climate model can achieve finer resolution by approximately an order of magnitude. In this study, we analyze P, T, and ET from a widely used regional climate model, the Advanced Research (ARW) core of the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008). The model makes use of a thermodynamic sea ice model (Zhang and Zhang 2001) and the Noah land-surface model (Yang et al., 2011) used within WRF in order to model the thermal conditions at the surface. It is driven at the lateral boundaries by observationally-based reanalysis fields (ERA-Interim) in order to avoid introduction of additional biases arising from the driving model. The ERA-Interim reanalysis was selected as it has been successfully downscaled using WRF in other regions of the world (e.g. Gao et al. 2015; Srivastava et al. 2014, 2013; Soares et al. 2012), has been used in previous Arctic WRF simulations and analyses (e.g. Liu et al. 2014) and is among the best performing reanalysis data sets for Alaska (Lader et al. 2015) and the wider Arctic (Lindsay et al. 2014).~~

~~As described by Bieniek et al. (2016), the downscaling covered a domain with 262x262 grid points that encompassed all of Alaska and portions of far eastern Russia and northern Canada at 20 km horizontal resolution (Figure 4) with 49 vertical model levels. The model was reinitialized every two days and was integrated for a total of 54 hours after each reinitialization. Each initialization occurred at 18 UTC (09 AKST). After each reinitialization, the first 6 hours~~

~~of output (a "spin-up" overlapping with the final 6 hours of the previous 54-hour integration) were discarded. The frequent reinitialization to the ERA-Interim reanalysis, together with a continuous nudging to the ERA-Interim upper air fields during each 54-hour simulation, ensured that the model was tightly constrained by the observationally-based reanalysis.~~

~~The output from the 2-day simulations was combined together to form the final downscaled product, which spanned the 1979-2015 period. Hourly WRF output was saved and used to produce daily mean values of downscaled variables. The daily ET values were calculated from the daily values of latent heat flux in the archive of model output. (The same conversion was performed with the eddy covariance measurements). The daily mean values of T, P, and ET from the model's 20-km grid cells corresponding to the tower sites of Section 2.1 were extracted for the years of the available tower data (Section 3).~~

## ~~9.~~3. ~~~~Methods

Formatted: Indent: Left: 0", Numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 2 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

The flux tower data were downloaded as 30-minute means that are processed from the high frequency (e.g., 10 Hz) ~~-~~eddy covariance measurements. The data for Barrow and Poker Flat were obtained from the AmeriFlux archive (http://ameriflux.lbl.gov), which is maintained by Lawrence Berkeley Laboratory. The Imnavait and Bonanza Creek data were downloaded from the Arctic Observing Network archive maintained by the Institute of Arctic Biology of the University of Alaska, Fairbanks, http://aon.iab.uaf.edu/ The measurements and data processing are described in more detail for Imnavait by Euskirchen et al. (2017) and for Bonanza Creek by Euskirchen et al. (2014). For all the sites, the 30-minute values were aggregated into 24-hour values (averages for T; totals for P, and ET) using MATLAB. ~~In the summation of the 30-minute P and ET into daily values, a missing 30-minute value was counted as zero.~~

Because the measurements from the tower sites are subject to instrumental outages and occasional malfunctions, it was necessary to perform several layers of quality-control. The quality-control procedure was applied to the warm-season months (May through September), which are the months with above-freezing mean air temperatures and with most of the yearly ET, as shown in Section 4. First, if occasional 30-minute segments of a day were missing (e.g., ET data during periods of rain), the daily values were computed as means of all 30-minute values that were available for the day. Second, we omitted a site-year if one or more of the sensors were inoperative, were known to be miscalibrated or repeatedly reported values out of range during a particular year. The years that survived this stage of the quality control are listed in Table 1. It is apparent that the available data from the various sites span different time periods. Such differences in temporal coverage, which are characteristic of eddy covariance sites operated worldwide, result from the diversity of the operators and funding sources of flux tower sites.

A special case of the quality-control was the precipitation at the Imnavait site, where three gauges (tipping buckets) were located within a 1 $km^2$ area. Calibration and malfunction issues led to sufficiently long and frequent gaps in the three sets of reliable precipitation measurements that it was necessary to combine the precipitation records from the three gauges. This consolidation utilized the most reliable single-gauge value that was available for each day. ~~Analysis indicated that when~~When data were available from all three precipitation gauges, analysis indicated that the measurements were equivalent as they were all within the same 1 $km^2$ area.

For the site-years that were retained, data were missing for some days or for some longer periods. If the gap was only a day or several days, the missing values were interpolated from available values for surrounding days. For longer intervals (e.g., 5-15 days) of missing

temperature and/or precipitation data, values were substituted from nearby station data: the Barrow airport data were used for the Barrow tower site, and the Fairbanks airport data for the Bonanza Creek and Poker Flat sites. No such substitution was possible for the Imnavait sites because there is no nearby reporting station. For the yearly accumulations of P, T, and ET described in Section 4, the values for a site-year are labeled as "estimated" if more than 45% of the daily values for May-September were filled in by interpolation or station-substitution.

In order to evaluate relationships between ET, T, and P, cross-correlations were evaluated at various lags (including zero). These cross-correlations, which are presented in Section 4, were based on departures from the daily averages for each site. The daily averages were computed using all available years of data for a site, but were replaced by 15-day running means of the single-date averages. The 15-day running means were used because the relatively small sample (< 10 years) for each site results in "climatological" seasonal cycles characterized day-to-day jumps. In the case of temperature, for example, the period-of-record means of the unsmoothed daily temperatures sometimes ~~varied~~ can vary by 2° to 3°C from one day to the next.

The model output from the simulation described in Section 2.~~2~~ 1 was available through 2015. The model output was obtained as daily totals of P and ET (and daily averages of T), so there was no need for infilling of missing values, averaging of 30-minute values, nor correcting for instrumental malfunctions. For comparison with the in situ measurements, the values of T, P, and ET for the model grid cells containing the tower sites were extracted for the years corresponding to the available tower data.

## ~~10.~~4. Seasonal and interannual variations

**Formatted:** Font: (Default) Times New Roman, 12 pt, Bold

**Formatted:** Indent: Left: 0", Hanging: 0.19", Numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 2 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

In the following sections, we present comparisons of the simulated values and the corresponding measurements. Because the results include four variables (ET, P, P-ET and T), four measurement sites (including one with flux towers in two vegetation sites), two sources of each variable foe each location, and seasonal climatologies in addition to variations over timescales from daily to seasonal, it is not feasible to present graphical displays of all results for all sites. Therefore, several of the following figures are based on samples of results chosen because they convey the most information relevant to the regional climate model performance. A more comprehensive diagnostic assessment focused on hydrologic processes rather than model evaluation would require additional figures as additional sites and will pursued in future work.

Formatted: Indent: First line: 0.31"

Sample time series of monthly values illustrating the key features of the ~~measured and model-derived~~model-derived and measured T, P, ET, and P-ET are shown for the tundra sites in Figure 5 and the forest sites in Figure 6. In each figure, the corresponding time series from the model (blue) can be compared with the tower measurements (red) ~~can be compared with the model's values (blue)~~ for the grid cell containing the tower site. All the measured variables (T, P, ET) undergo strong seasonal cycles, which have larger amplitudes than the interannual variations. It is apparent from Figures 5a and 6a that the seasonal cycles of temperature are well simulated by the model, although the model's temperatures tend to be too high by several degrees during summer at Barrow.

The interannual variations in Figures 5a and 6a show generally good correspondence between the model and the measurements ~~and the model~~, although there are discrepancies, especially in the summer temperatures at Barrow. Because the model was forced by observed lateral and ocean surface/sea ice boundary conditions and was also reinitialized to observational data over Alaska at 48-hour intervals, the year-to-year (and day-to-day variations in the model

383 output should, in principle, agree with the corresponding variations in the measurements. Model
384 errors (resolution-related as well as formulational) and measurement limitations cause the values
385 from the two sources to differ. In the case of the Barrow temperatures, the discrepancies are also
386 attributable to the proximity of the Barrow tower to the coastline (and onshore advection of cool
387 maritime air during summer), while the model's temperatures for the Barrow grid cell are
388 averages for a land area 20 km on a side.

389 The comparison of measured and model-derived precipitation is limited to the warm season
390 because, as noted earlier, the instrumentation at the tower sites does not enable meaningful
391 estimates winter precipitation amounts. Winter values of P are therefore shown as zero in Figures
392 5 and 6. However, Figures 5b and 6b show that the models over-simulate warm-season
393 precipitation by a significant amount. The interannual variations of P are captured to some extent
394 at the Poker Flat forest site, especially the extremely wet year of 2014, but there is little
395 correspondence between the interannual variations of measured and modeled P at Barrow.
396 However, the ET variations at Imnavait (Figure 5c) show generally good correspondence
397 between the model and the measurements, including the interannual variations. Figure 5c also
398 shows that ET is similar at the sedge and heath sites in terms of both the mean seasonal cycle and
399 the interannual variations, although the interannual variations at the sedge site are somewhat
400 closer to the model's values. The most notable discrepancy between the model and the
401 measurements at Imnavait is in the net surface flux, P-ET (Figure 5d). The model's P-ET
402 generally remains positive throughout the summer, except for single-month excursions that
403 barely reach into negative values in a few years. The measurements, on the other hand, show
404 stronger excursions into negative P-ET (net moisture loss), often for more than one month, at
405 both tower sites, especially at the sedge site. In only one year (2014) of the eight years of

406 measurements did the monthly tower-derived ET remain positive through the summer. The
407 model's excessive summer wetness (positive P-ET) is attributable to the model's much larger P
408 relative to the in situ measurements, as the differences in ET are not as large.

409 At the forest sites, the model shows greater P as well as greater ET relative to the
410 observations (Figure 6). The over-simulations of P and ET act to offset each other somewhat,
411 resulting in warm-season P-ET that is comparable in the model and the data, with more than one
412 month of drying (negative P-ET) indicated by each source in most years Figure 6d even shows
413 some model-data correspondence in the interannual variations of P-ET. 2013 and 2015 were
414 relatively dry years according to both sources of information.

415 The biases of the model are clearly apparent in the monthly means of the hydrologic
416 variables, which are shown in Figure 7 for Barrow and Poker Flat. The results for Imnavait and
417 Poker Flat (not shown) are similar. P and ET are over-simulated by the model in all months for
418 which the tower data are available, and the over-simulation is greater in the boreal forest than on
419 the tundra. The excess ET in the model is almost certainly driven in part by the excess P,
420 although the model-data discrepancies in ET at Barrow are larger than the model-data
421 discrepancies in P at Barrow. Figure 7 also makes the point that, *in the model*, the cold-season
422 fluxes are much smaller than the warm-season fluxes, especially in the boreal forest. The tower
423 data do not permit summer-winter comparisons of P or ET.

424 The climatological (mean over all years) seasonal cycles of P-ET, the net surface moisture
425 flux, show interesting differences among the various sites and in the model's ability to capture
426 the seasonal cycle of P-ET. As shown in Figure 8a, there is a net moisture loss during June and
427 July at the Barrow tower site and during June at the Imnavait tower sites. The differences

between the two Imnavait sites are small in all calendar months. However, there are large discrepancies between the P-ET of the model and the tundra tower sites. Figure 8a shows that the model's summer P-ET is much more negative than the tower-derived values at Barrow, and much more positive than the tower-derived values at Imnavait. In other words, the model shows excessive summer drying at Barrow and excessive summer wetting at Imnavait. The excess drying at Barrow is attributable to the model's excess warmth at Barrow (Figure 5a), while the model's excess wetting at Imnavait is attributable to the model's excessive precipitation (Figures 5c and 5d). These different reasons for the model's biases in the critical quantity, P-ET, point to the challenges in obtaining credible model simulations of the surface moisture budget in Arctic tundra regions.

Figure 8b shows the corresponding P-ET climatologies for the forest sites, for which the model and the measurements are in surprisingly good agreement – given the discrepancies in the simulated and measured P. The outstanding feature of Figure 8b is the difference in sign between summer P-ET at the two sites: a net moisture loss at Bonanza Creek and a net moisture gain at Poker Flat in both the model output and the tower data. This difference is consistent with greater subsurface moisture storage at Poker Flat, where the active layer is deeper and the permafrost is discontinuous. However, an additional factor contributing to the difference is the inclusion of a very high precipitation year, 2014 (Fig. 6b) in the climatology for Poker Flat but not for Bonanza Creek. While Bonanza Creek's P was available for 2014 (and was well above its mean, as shown in Section 5.1), ET was not available because of an instrument outage. Hence P-ET could not be included in the Bonanza Creel multiyear averages in Figure 8b.

## ~~11.~~5. Relationships to atmospheric drivers

Formatted: Indent: Left: 0", Numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 2 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

In order to assess the linkages between ET and two of its key drivers (T and P), we focus on variations over two timeframes: (1) yearly totals, computed as accumulations over the June-September "warm season" when the ground is generally snow-free and ET is largest, and (2) variations of ET over short periods of one to several days.

### *5.1 Warm-season integrals of ET and drivers*

~~*11.1*~~ Seasonally accumulated totals of ET provide a means to evaluate the relationships between interannual variations of ET and its drivers. In this section we present accumulation curves for the different variables based on observational measurements in order to provide estimates unaffected by model biases. We then correlate the interannual variations of the seasonal totals of E and ET with corresponding totals of precipitation and temperature anomalies in order to assess the relative contributions of T and P to the ET and P-ET. Finally, we compare correlations based on the model-output and the on the observational measurements in order to assess the model's ability to capture the drivers of interannual variations of ET and P-ET.

Warm-season accumulations of (a) daily temperature anomalies, (b) precipitation, and (c) ET are shown in Figures 9 for a tundra site (Imnavait heath) and in Figure 10 for a forest site (Bonanza Creek). In all cases, the summations are performed for the four-month period June 1 through September 30. The accumulated temperature anomalies in Figure 9 show that some years at Imnavait were relatively warm (2010, 2012), while others were relatively cool (2014, 2015). For scaling purposes, we note that a 4-month accumulated temperature departure of 122ºC corresponds to a seasonally averaged departure of 1ºC. 2015's value of -210ºC in Fig. 9a therefore represents a summer temperature departure of approximately -1.7ºC (-3.1ºF). Figure 9b shows that the wettest year (2012) had more than twice the precipitation of the driest year

**Formatted:** Indent: Left: 0.06", Hanging: 0.19", Outline numbered + Level: 2 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5", Tab stops: 0.38", Left + 0.44", Left

**Formatted:** No bullets or numbering

**Formatted:** Font: Not Italic

(2013). Examination of Figure 10c shows that 2009 and 2010 (warm years) were the years with the greatest ET, while 2015 (a cool year) was the year with the smallest ET.

Table 2 lists the correlations (across the available years) between the warm-season totals of P, E, P-E and the T anomaly. The table includes values computed from the model simulations for the same years, enabling a comparison of the relative importance of T and P for the hydrologic fluxes, ET and P-ET. Figure 11 is a graphical comparison of the measurement- and model-based correlations for the Imnavait and Bonanza Creek sites, highlighting the model's ability to capture the contributions of T and P to the interannual variations at the sites with the most coherent signals between the atmospheric and hydrologic variables. For the observational results, the ~~When the full set of years is considered, the~~ correlations between seasonally accumulated T and total ET at the Imnavait ~~heath site is is 0.63,~~sites are 0.63 (heath) and 0.71 (sedge), indicating that seasonally integrated evapotranspiration tends to be greater in warmer years and smaller in cooler years. ~~This correlation is based on an admittedly small sample of nine years, but it is even larger for the Imnavait wet sedge site and in the model output for Imnavait (Figure 11).~~ The corresponding value is 0.77 for the model's 20 km x 20 km Imnavait grid cell, which contains both heath and sedge tundra. For the sample size of N = 9 (years), the 95% significance level is approximately 0.69, so the correlations obtained from the sedge measurements and the model simulation are statistically significant despite the small sample size. By contrast, there is little correspondence between the seasonally accumulated P and seasonal ET: the correlations between the yearly P and ET at Imnavait are -0.~~13~~ 12 (heath measurements) and -0.~~40~~ 48 (sedge measurements), and -0.24 (model), none of which are statistically significant. We conclude that the primary driver of interannual variations of total warm-season ET at Imnavait is temperature. The implication is that a warming climate will lead to greater ET,

even if P also increases in a warming climate as projected by global models. However, the measurements from Barrow show no meaningful correlation between T and ET, but a strongly negative correlation between P and ET. Given the saturated surface conditions at Barrow, this negative correlation with P may point to the tendency for ET to be greater when cloudiness is reduced, allowing for greater insolation. This hypothesis is speculative and requires further examination with additional data (radiative fluxes and winds).

At the forest sites, there is also a positive correlation between seasonally integrated measurements of T and ET (r = +0.67 at Bonanza Creek, r = +0.76 at Poker Flat). The corresponding model-derived values are +0.58 and +0.81. At Bonanza Creek, ET correlates even more strongly with P (r = +0.73) than with T, pointing to a role of recycling of moisture through ET. ~~This~~ However, this signal is not apparent at Poker Flat, ~~where there are only four years of reliable ET measurements~~nor in the model results for either forest site.

A universal feature of the measurement- and model-derived results for all four sites is the strong positive correlation ~~However, there are strong correlations~~ between seasonal totals of P and P-ET ~~at the forest sites~~ (Table 2 and Figure 11). While the inclusion of P in P-ET virtually guarantees some correlation, the magnitudes of the P vs. P-ET correlations (>0.9 in most cases) indicates that the effects of temperature ~~(negatively correlated with P, r = -0.55)~~ do little to offset the effects of P on surface moisture exchange. Figure 11 shows that the model also reproduces the high correlation between seasonally integrated P and P-ET at Bonanza Creek despite the model's strong correlation between T and P-ET, r = -0.88, which is much more negative than in the tower data. Even at the tundra sites, the seasonally integrated P-ET correlates more strongly with P than with T, despite the strong dependence of ET on T. We conclude that precipitation is the main driver of P-ET at all sites on a year-to-year basis, implying that future changes of the

surface moisture budget will be determined by future changes in P more than by changes in T. The robustness of this conclusion is supported by the consistency between the model results and *in situ* measurements.

Formatted: Font: Italic

### *5.2 Short-term variations of ET and its drivers*

The seasonal relationships summarized in the previous subsection represent integrations of shorter-term linkages between ET and its drivers, T and P. In order to quantify the shorter-term linkages, we computed cross-correlations between ET and T as well as ET and P with the variables averaged over periods of 1, 3 and 30 days. We show the results as cross-correlation functions of lead/lag, whereby the ET leads or lags the T and P variations by 0, 1, 2,…30 days. Figure 12 shows these cross-correlation functions computed from the daily values of T, P, and ET for Imnavait Creek, Bonanza Creek and Poker Flat. Results are shown for both the tower data (solid lines) and the model (dashed lines). In all cases, the daily values have been converted to departures from the daily means, thereby removing the effects of the seasonal cycle from the correlations.

Two peaks appear consistently in the correlation functions: a positive correlation of ET with temperature, centered at or close to zero lag, and a negative correlation of ET with precipitation, also centered at zero lag. The positive temperature correlation is consistent with the seasonal results in Section 5.1: ET is greater when the temperature is higher than its daily average. The correlation functions for temperature decay rapidly on either side of lag zero, although the values still stand out above the noise-level at temperature leads and lags of a day or two, consistent with the autocorrelations of daily temperature anomalies. The zero-lag correlations between T and ET are slightly stronger in the model (r = +0.4 to +0.5) than in the tower data (r = +0.3 to +0.4 at

Imnavait and Bonanza Creek), in agreement with the relative strength of the seasonal correlations between T and ET in Figure 11 and Table 2. The lower zero-lag correlations at Barrow suggest that the high humidity and cloud coverage limit the importance of air temperature in daily variations of ET.

The zero-lag peaks in Figure 12's correlations between P and ET are weaker than those of the temperature correlations. The negative sign of the zero-lag spikes is not attributable to a direct effect of falling precipitation, as the eddy covariance fluxes during precipitation events have removed from the database because the measurements are problematic when the instruments are wet. The most plausible explanation for the negative peaks at zero-lag is that temperatures tend to be depressed on days with precipitation and clouds, so the negative spikes in the P curves may actually be indirect manifestations of the temperature correlations. The fact that the Imnavait correlations with ET are larger for both T and P at the sedge site compared to the heath site supports the connection between the two drivers. Aside from the spike at zero lag, precipitation shows little association with ET, consistent with the seasonal results presented earlier.

While the short-term associations with T and P are largest at zero lag, one may hypothesize cumulative effects of T or P anomalies may cause a stronger signal if the T vs. ET and P vs. ET relationships are evaluated over timescales longer than a day (Figure 12) but shorter than a season (Figure 12). To test this hypothesis, we experimented with the use of running means of T, P, and ET in the correlative analysis. The averaging period for the running means varied from several days to 30 days. Figure 13 shows the correlation functions based on 3-day means for the two Imnavait sites and the corresponding model output. While the correlation functions decay from their zero-lag values at a slower rate with the smoothed data, the maximum values are essentially the same as in Figure 12a. Even the relative magnitudes of the sedge-vs.-heath tundra

and measurement-vs.-model peaks are little changed from the results based on single-day values. Experiments with longer averaging periods produced no enhancement of the associations and even led to a degradation of the correlations as the averaging period approached 30 days. We conclude that any multiday cumulative effects are not strong enough to enhance the concurrent (zero-lag) associations inherent in the daily data.

## ~~12.~~6. Conclusion

The results in the preceding sections lead to the following conclusions:

- Both P and ET are considerably larger in the model for all sites, indicating that the model's hydrological cycle is stronger than the observed. This over-simulation of the high-latitude hydrologic cycle is consistent with the known tendency of global models to simulate more precipitation than is observed in northern high latitudes (Walsh et al., 2002; de Boer et al., 2012), subject to uncertainties in the observational data.
- 
- The model output and the tower measurements ~~from all sites~~ show a short period (one or two months) of negative P-ET during summer, indicative of surface drying, although the model does not show this period of drying ~~in~~ at the Imnavait tundra sites in the foothills of the Brooks Range.
- ~~-~~At all sites, interannual variations in the warm-season surface water balance (P-ET) are determined primarily by variations in precipitation. The dominance of P as a driver is especially apparent at the forest sites, and it is apparent in the model output as well as in the observational data.

**Formatted:** Indent: Left: 0", Numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 2 + Alignment: Left + Aligned at: 0.25" + Indent at: 0.5"

**Formatted:** List Paragraph, Indent: Left: 0.04", Bulleted + Level: 1 + Aligned at: 0.06" + Indent at: 0.33"

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** List Paragraph, Indent: Left: 0.31"

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Normal, Indent: Left: 0", Hanging: 0.31", No bullets or numbering

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Font: (Default) Times New Roman, 12 pt

**Formatted:** Normal, Indent: Left: 0", Hanging: 0.25", Space After: 0 pt

- At both the Imnavait tundra and Poker Flat forest sites, the model and the measurements are consistent in showing that variations of evapotranspiration are controlled primarily by temperature. The dominance of temperature as a driver is supported especially by the interannual variations of the seasonal totals, but also by correlations on the daily timescale, in both the model results and the observational data. Only the model shows this dominance of temperature at the Barrow tundra and Bonanza Creek forest sites.

- ~~At all sites, interannual variations in the warm-season surface water balance (P-ET) are determined primarily by variations in precipitation. The dominance of P as a driver is especially apparent at the forest sites, and it is apparent in the model output as well as in the observational data.~~

- ~~Both P and ET are considerably larger in the model for all sites, indicating that the model's hydrological cycle is stronger than the observed. This over-simulation of the high-latitude hydrologic cycle is consistent with the known tendency of global models to simulate more precipitation than is observed in northern high latitudes (Walsh et al., 2002; de Boer et al., 2012), subject to uncertainties in the observational data.~~

~~Taken at face value, the results imply that a warming climate will lead to greater warm-season ET at tundra sites because summers will be warmer and longer. At all sites, however, the seasonal net surface water flux (P-ET) is more sensitive to precipitation than to temperature,~~

Formatted: Indent: First line: 0"

~~implying that future changes in atmospheric exchanges with the land surface will be largely controlled by changes in precipitation. Such findings are based on a small sample of only two tundra and two forest locations and fewer than ten years of data for each location, so they are largely exploratory. Furthermore, changes in actual wetness of the ground surface will also depend on future changes in active layer and surface drainage, which may be affected by thermokarst in areas of permafrost thaw.~~

~~Given the small sample of locations and years, a high priority is the extension of this evaluation to other high-latitude sites. Existing flux data networks (AmeriFlux, FluxNet) contain several dozen Arctic or subarctic stations with record lengths spanning several years to a decade or more. A related need is to address the uncertainties in the precipitation data. Precipitation is notorious for small-scale variations as well as instrumental challenges (e.g., gauge undercatch), so the robustness of conclusions based on precipitation data is open to question.~~

~~Finally, t~~The discrepancies between the tower and model values of P and ET found here are sufficiently large that there is a need to determine whether other models show similar biases. A related need is to address the uncertainties in the precipitation data. Precipitation is notorious for small-scale variations as well as instrumental challenges (e.g., gauge undercatch), so the robustness of conclusions based on precipitation measurements is open to question.~~The WRF model may or not be an outlier in this regard. Whether or not it is, the reasons for the discrepancies need to be explored in terms of the processes included in the land model and the parameterizations used to represent key processes related to evapotranspiration.~~

Taken at face value, the results imply that a warming climate will generally lead to greater warm-season ET ~~at tundra sites~~ because summers will be warmer and longer. At all sites, however, the seasonal net surface water flux (P-ET) is more sensitive to precipitation than to

**Formatted:** List Paragraph

temperature, implying that future changes in atmospheric exchanges with the land surface will be largely controlled by changes in precipitation. Such findings are based on a small sample of only two tundra and two forest locations and fewer than ten years of data for each location, so they are largely exploratory. Furthermore, changes in actual wetness of the ground surface will also depend on future changes in active layer and surface drainage, which may be affected by thermokarst in areas of permafrost thaw. Such processes are not yet included in most climate models. Because models will continue to be the primary tools for anticipating future changes, it is important that Arctic terrestrial simulations be extended to additional models, especially as models evolve to include additional processes relevant to Arctic terrestrial hydrology.

**ACKNOWLEDGMENTS**

Ross Fischer was supported by the National Science Foundation (NSF) through Grant ARC-1560372, an REU (Research Experience for Undergraduates) grant to the University of Alaska, Fairbanks. Funding for the eddy covariance data collection at the Imnavait flux tower sites was through the NSF Division of Polar Programs, Arctic Observatory Network grant numbers 1503912, 1107892 and 0632264. Funding for the Bonanza Creek flux tower measurements was provided by the NSF through the Bonanza Creek Long Term Ecological Research Program and the U.S. Geological Survey.

650 **REFERENCES**


651 Bieniek, P. A., U. S. Bhatt, J. E. Walsh, T. S. Rupp, J. Zhang, J. R. Krieger, and R. Lader,
652 2015: Dynamical downscaling of ERA-Interim temperature and precipitation for Alaska. *J.*
653 *Appl. Meteor. Climatol.*, 55(3), 1276-1289, DOI: 10.1175/JAMC-D-15-0153.1.

654 Bring, A., I. Federova, Y. Dibike, L. Hinzman, J. Mard, S. H. Mernild, T. Prowse, O. Semenova,
655 S. L. Stuefer, and M.-K. Woo, 2016: Arctic terrestrial hydrology: A synthesis of processes,
656 regional effects, and research challenges. *J. Geophys. Res.: Biogeosciences*, doi:
657 10.1002/2015JG003131.

658 de Boer, G., W. Chapman, J. E. Kay, B. Medeiros, M. D. Shupe, S. Vavrus, and J. Walsh, 2012:
659 A characterization of the present-day Arctic atmosphere in CCSM4. *J. Climate, 25*, 2676-

660 Dirmeyer, P. A., Y. Jin, B. Singh, and X. Yam, 2013: Trends in land-atmosphere interactions
661 from CMIP5 simulations. *J. Hydrometeorology*, *14*(3), 829-849.

662 Euskirchen, E. S., M. S. Bret-Harte, G. R. Shaver, C. W. Edgar, and V. E. Romanovsky, 2017:
663 Long-term release of carbon dioxide from arctic tundra ecosystems in northern Alaska.
664 *Ecosystems.20*(5), 960-974, doi:10.1007/s10021-016-0085-9.

665 Euskirchen. E. S., M. S. Bret-Harte, G. J. Scott, C. Edgar, and G. R. Shaver, 2012: Seasonal
666 patterns of carbon dioxide and water fluxes in three representative tundra ecosystems in northern
667 Alaska. *Ecosphere,* (1):4. http://dx.doi.org/10.1890/ES11-00202.1.

668 Euskirchen, E. S., C. Edgar, M. R. Turetsky, M. P. Waldrop, and J. W. Harden, 2014:
669 Differential response of carbon fluxes to climate in three peatland ecosystems that vary in the
670 presence and stability of permafrost. *J. Geophys. Res.: Biogeosciences*, *119*, *1576–1595*,
671 doi:10.1002/2014JG002683.

672

673 Flannigan, M. D., B. M. Wotton, G. A. Marshall, W. J. de Groot, J. Johnston, N. Jurko, and A. S.
674 Cantin, 2015: Fuel moisture sensitivity to temperature and precipitation: Climate change
675 implications. *Climatic Change*, doi: 10.1007/s10584-015-1521-0.

676 Gao, Y., J. Xu, and D. Chen, 2015: Evaluation of WRF mesoscale climate simulations over the
677 Tibetan Plateau during 1979-2011. *J. Climate, 28,* 2823-2841.

678 IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Report of Working Group 1,
679 Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, U.K.,
680 1535 pp., doi:10.1017/CBO9781107415324.

681 Lader, R., U. S. Bhatt, J. E. Walsh, T. S. Rupp, and P. A. Bieniek, 2015: Evaluation of
682 Reanalysis Products for Alaska to Facilitate Climate Impact Studies. *J. Appl. Meteor. Climatol.*,
683 *55*(4). 901-922. doi: 10.1175/JAMC-D-15-0162.1.

684 Lafleur, P. M., and E. R. Humphreys, 2007: Spring warming and carbon dioxide exchange over
685 low Arctic tundra in central Canada. *Global Change Biology*, *14*, 740-756.

686 Lafleur, P. M., E. R. Humphreys, V. L. St. Louis, M. C. Myklebust, T. Papakryiakou, L.
687 Poissant, J. D. Barker, M. Pilote, and K. A. Swystun, 2012: Variation in peak growing season net
688 ecosystem production across the Canadian Arctic. *Env. Science and Technology*, *46*(15), 7971-
689 7977.

690 Lainé, A., H. Nakamura, K. Nishii, and T. Miyahasaka, 2014: A diagnostic study of future
691 evaporation changes projected in CMIP5 climate models. *Climate Dynamics*, 42, 2745-2761.

692 Lindsay, R., M. Wensnahan, A. Schweiger, and J. Zhang, 2014: Evaluation of Seven Different
693 Atmospheric Reanalysis Products in the Arctic. *J. Climate*, **27**, 2588–2606.

694 Liu F., J. R. Krieger, and J. Zhang, 2013: Toward producing the Chukchi/Beaufort High-
695 Resolution Atmospheric Reanalysis (CBHAR) via the WRFDA data assimilation system. *Mon.*
696 *Wea. Rev.*, *142*(2), 788-805.

697 Niu, G.-Y., Z.-L. Yang, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, A. Kumar, D. Niyogi, E.
698 Rosero, M. Tewar,i and Y. Xia, 2011: The community Noah land surface model with
699 multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale
700 measurements. *J. Geophys. Res., 116*, D12109, doi:10.1029/2010JD015139.

701 Ohta, T., T. C. Maximov, A. J. Dolman, T. Nakai, M. K. van der Molen, A. K. Kononov, A. P.
702 Maximov, T. Hiyama, Y. Iijima, E. J. Moors, H. Tanaka, T. Toba, and H. Yabuki, 2008:
703 Interannual variation of water balance and summer evapotranspiration in an eastern Siberian
704 larch forest over a 7-year period (1998-2006). *Agricul. and Forest Meteorol.*, *148*, 1941-1953.

705 Partain, J. L. Jr., S. Alden, U. S. Bhatt, P. A. Bieniek, B. R. Brettschneider, R. T. Lader,, P. Q.
706 Olsson, T. S. Rupp, H. Strader, R. L. Thoman, Jr., J. E. Walsh, A. D. York, and R. H. Ziel, 2016:
707 An assessment of the role of anthropogenic climate change in the Alaska fire season of 2015.
708 *Bull. Amer. Meteor. Soc.*, *97*(12), S14-S18.

709 Rawlins, M. A., M. Steele, and 28 coauthors, 2010: Analysis of the Arctic system for freshwater
710 cycle intensification: Observations and expectations. *J. Climate*, 23, 5715-5737.

711 Skamarock, W. C. and Coauthors, 2008: A Description of the Advanced Research WRF Version
712 3. NCAR Tech Note, NCAR/TN–475+STR, 113 pp.

713 Soares, P. M. M., R. M. Cardoso, P. M. A. Miranda, J. de Medeiros, M. Belo-Pereira, and F.
714 Espirito-Santo, 2012: WRF high resolution dynamical downscaling of ERA-Interim for Portugal.
715 *Clim. Dyn., 39***,** 2497-2522.

716 Srivastava, P. K. D. Han, M. A. Rico-Ramirez, and T. Islam, 2014: Sensitivity and uncertainty
717 analysis of mesoscale model downscaled hydro-meteorological variables for discharge
718 prediction. *Hydrol. Processes, 28***,** 4419-4432.

719 Srivastava, P. K., D. Han, M. A. Rico, and T. Islam, 2013: Comparative assessment of
720 evapotranspiration derived from NCEP and ECMWF global datasets through Weather Research
721 and Forecasting model. *Atmos. Sci. Let., 14,* 118-125.

722 Walsh, J. E., V. M. Kattsov, W. L. Chapman, V. Govorkova, and T. Pavlova, 2002: Comparison
723 of Arctic climate simulations by uncoupled and coupled global models. *J. Climate, 15*, 1429-
724 1446.

725 Wang, S., M. Pan, Q. Mu, X. Shi, J. Mao, C. Brummer, R. S. Jassal, P. Krishnan, J. Li, and T. A.
726 Black, 2015: Comparing evapotranspiration from eddy covariance measurements, water budgets,
727 remote sensing, and land surface models over Canada. *J. Hydrometeorology, 16*, 1540-1560.

728 Yao, Y., S. Liang, and 25 coauthors, 2016: Assessment and simulation of global terrestrial latent
729 heat flux by synthesis of CMIP5 climate models and surface eddy covariance observations.
730 *Agricultural and Forest Meteorology, 223*, 151-167.

731 Zhang, X. and J. Zhang, 2001: Heat and freshwater budgets and pathways in the Arctic
732 Mediterranean in a coupled ocean/sea-ice model. *J. Oceanogr., 57*, 207–237.

733

TABLE 1: Years of available data from each tower site

| | |
|---|---|
| Barrow | 1998-2007 |
| Imnavait Cr. | 2007-2015 |
| Bonanza Cr. | 2010-2013, 2015-2016 |
| Poker Flat | 2011-2014 |

Table 2. Correlations between total seasonal accumulations of P, ET, P-ET and temperature (T) anomaly. Accumulations span June 1 through September 30 as in Figures 9 and 10.

| | | ET vs. T | ET vs. P | P-ET vs. T | P-ET vs. P |
|---|---|---|---|---|---|
| Barrow | measurements | -0.22 | -0.89 | 0.38 | 0.98 |
| | model | 0.67 | 0.13 | -0.29 | 0.86 |
| Imnavait | measurements (heath) | 0.63 | 0.12 | 0.43 | 0.81 |
| | measurements (sedge) | 0.71 | -0.48 | 0.26 | 0.99 |
| | model | 0.77 | -0.24 | -0.12 | 0.97 |
| Bonanza Cr. | measurements | 0.67 | 0.73 | -0.03 | 0.98 |
| | model | 0.58 | -0.07 | -0.88 | 0.89 |
| Poker Flat | measurements | 0.76 | -0.39 | -0.90 | 0.99 |
| | model | 0.81 | -0.36 | -0.91 | 0.98 |

Formatted: Indent: Left: 0", Hanging: 0.63", Line spacing: Double

## FIGURE CAPTIONS

Figure 41. The domain of the WRF regional climate model simulation. Blue dots are spaced at intervals of 20 km, the horizontal resolution of the model.

Formatted: Left, Indent: Left: 0", Hanging: 0.69"

Figure 12. Flux towers at Poker Flat (left) and Bonanza Creek (right).

Figure 23. Locations of the flux tower sites: Barrow and Imnavait Creek on the tundra, and Bonanza Creek and Poker Flat in the boreal forest.

Figure 34. Vegetation in vicinity of flux towers at Bonanza Creek (left) and Imnavait Creek (right). Imnavait photos show heath tundra (upper right) and wet sedge tundra (lower right).

~~Figure 4. The domain of the WRF regional climate model simulation. Blue dots are spaced at intervals of 20 km, the horizontal resolution of the model.~~

Figure 5. Sample time series of tower measurements (red/yellow) and corresponding model output (blue) for tundra sites averaged over monthly periods: (a) air temperatures at Barrow, (b) precipitation, P, at Barrow, (c) evapotranspiration, ET, at Imnavait, with separate curves for heath tundra (red) and sedge tundra (yellow), and (d) P-ET at Imnavait, with separate curves for heath and sedge tundra. Note that the towers do not provide meaningful values of P and ET during the cold season.

Figure 6. As in Figure 5, but for forest sites: (a) air temperature at Poker Flat, (b) precipitation, P, at Poker Flat, (c) evapotranspiration, ET, at Bonanza Creek, (d) P-ET at Bonanza Creek.

795 Figure 7. Monthly climatologies (averages over all available years) of tower measurements
796 (blue) and corresponding model simulations (yellow) of (a) precipitation at Barrow,
797 (b) evapotranspiration at Barrow, (c) precipitation at Poker Flat, and (d)
798 evapotranspiration at Poker Flat.

799 Figure 8. Climatological (average over all available years) monthly values of P-ET at tundra
800 sites (upper panel) and forest sites (lower panel). Values derived from tower
801 measurements are shown by solid lines, model-simulated values by dashed lines.

802 Figure 9. Seasonal (June through September) accumulations of (a) air temperature anomalies,
803 (b) precipitation and (c) evapotranspiration at the Imnavair heath site. All values are
804 based on tower measurements. Different years are color-coded; dashed lines are
805 estimated values (see text) for years with large amounts of missing data.

806 Figure 10. As in Figure 9, but for seasonal accumulations at Bonanza Creek.

807 Figure 11. Sample correlations between yearly warm-season totals of ET, P-ET, and driving
808 variables (T, P). Blue and green bars are correlations based on tower data; yellow
809 bars are correlations based on model simulation.

810 Figure 12. Cross-correlation functions of daily ET with daily temperature (red) and daily
811 precipitation (blue) for (a) Imnavait Creek, (b) Bonanza Creek and (c) Poker Flat.
812 Correlations are plotted as a function of the lag of P and T relative to ET (i.e., T and P
813 lead ET to the left of zero lag; T and P lag ET to the right of zero lag). Solid lines are
814 based on tower measurements, dashed lines on model output.

Figure 13. As in Figure 12, but for cross-correlations computed from 3-day running means of T, P, and ET at Imnavait Creek. Solid lines are based on tower measurements, dashed lines on model output. Values for wet sedge and heath are plotted in different colors.

Figure 41. The domain of the WRF regional climate model simulation. Blue dots are spaced at intervals of 20 km, the horizontal resolution of the model.

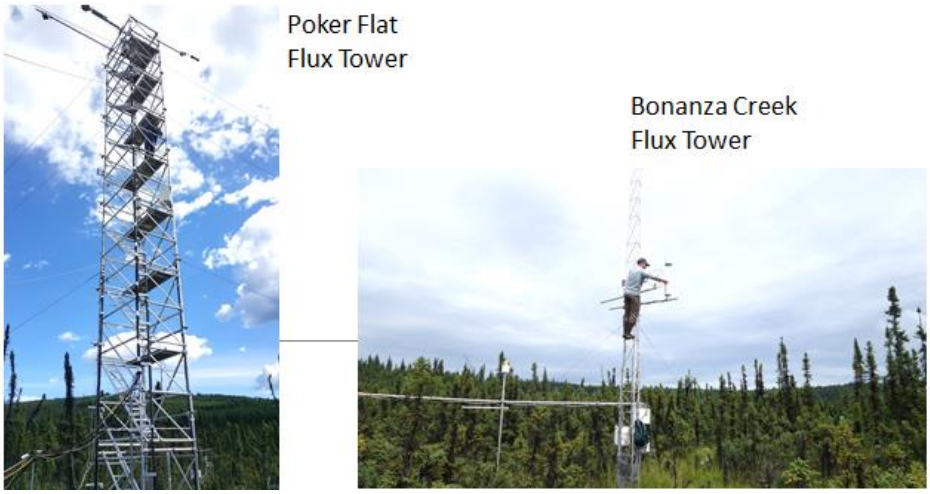



Figure ~~1~~2. Flux towers at Poker Flat (left) and Bonanza Creek (right).

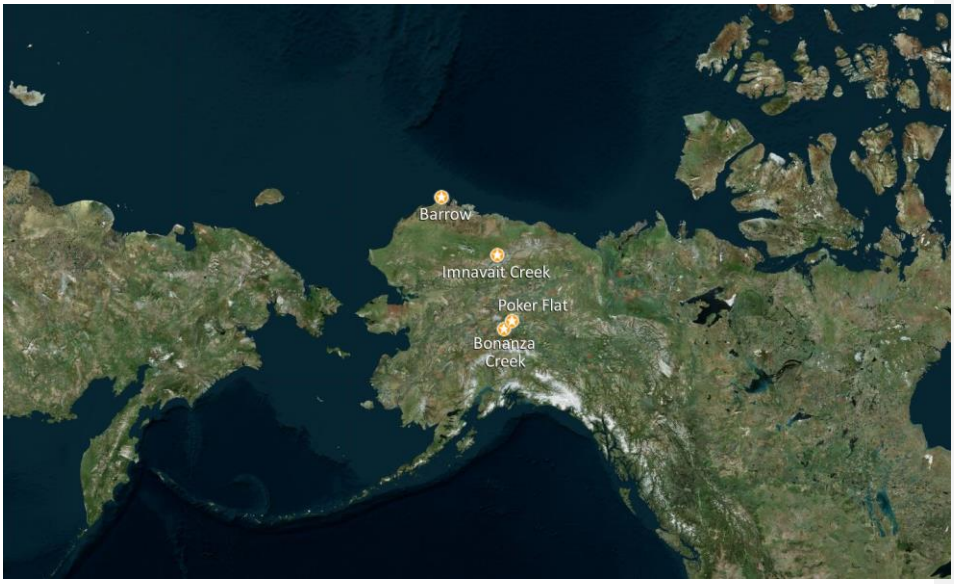



Figure ~~2~~ 3. Locations of the flux tower sites: Barrow and Imnavait Creek on the tundra, and Bonanza Creek and Poker Flat in the boreal forest.

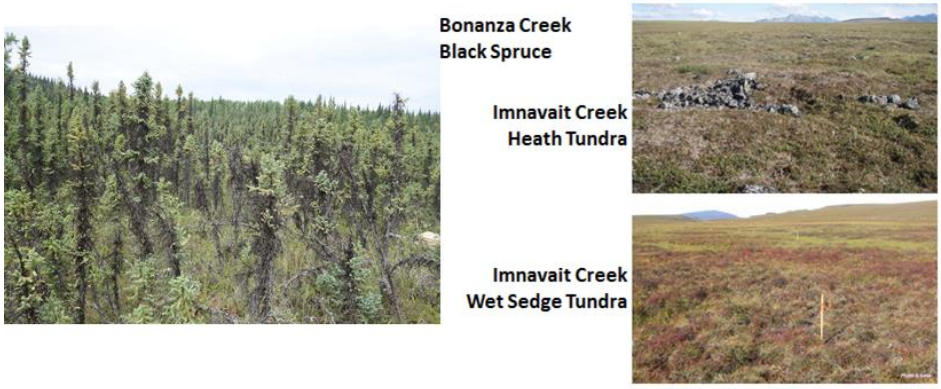
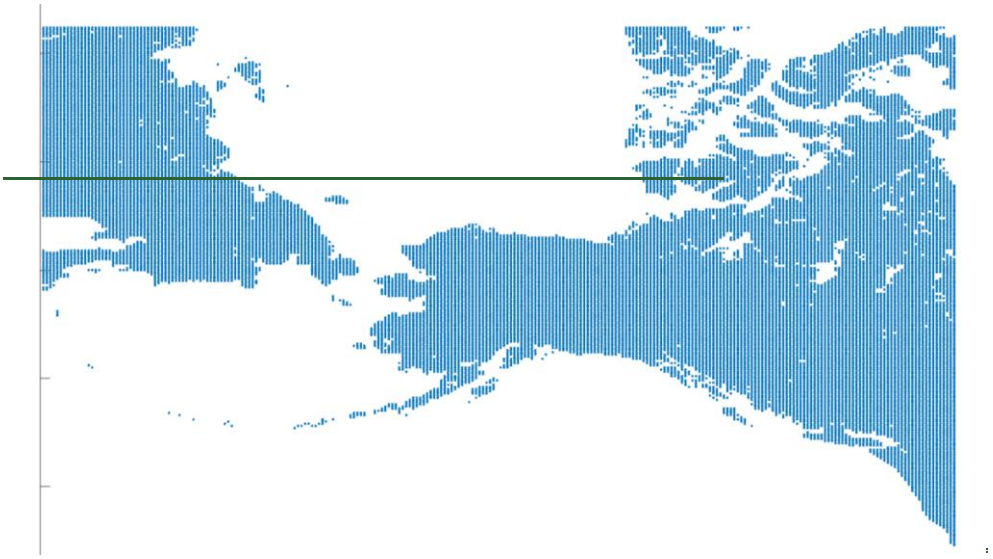



Figure ~~3~~4. Vegetation in vicinity of flux towers at Bonanza Creek (left) and Imnavait Creek (right). Imnavait photos show heath tundra (upper right) and wet sedge tundra (lower right).



~~Figure 4. The domain of the WRF regional climate model simulation. Blue dots are spaced at intervals of 20 km, the horizontal resolution of the model.~~

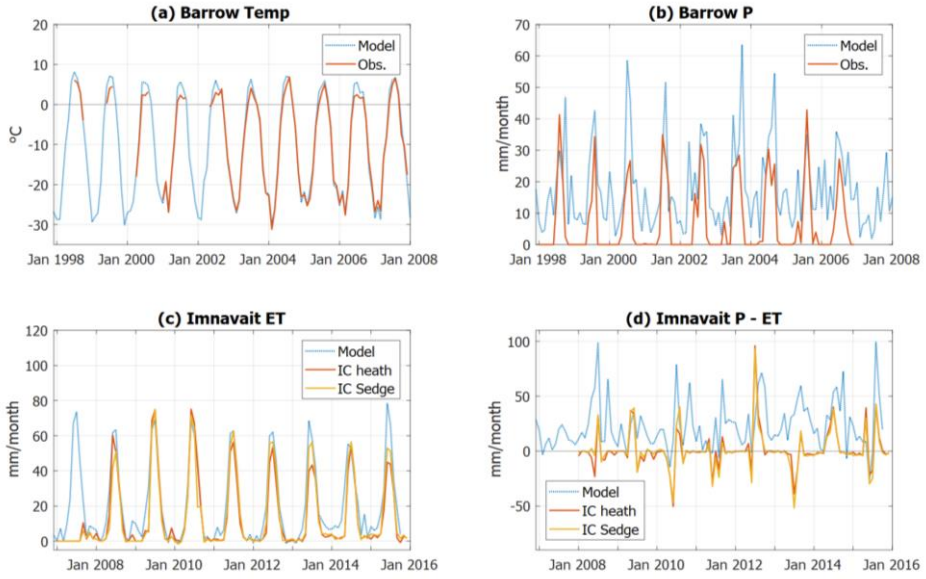



Figure 5. Sample time series of tower measurements (red/yellow) and corresponding model output (blue) for tundra sites averaged over monthly periods: (a) air temperatures at Barrow, (b) precipitation, P, at Barrow, (c) evapotranspiration, ET, at Imnavait, with separate curves for heath tundra (red) and sedge tundra (yellow), and (d) P-ET at Imnavait, with separate curves for heath and sedge tundra. Note that the towers do not provide meaningful values of P and ET during the cold season.

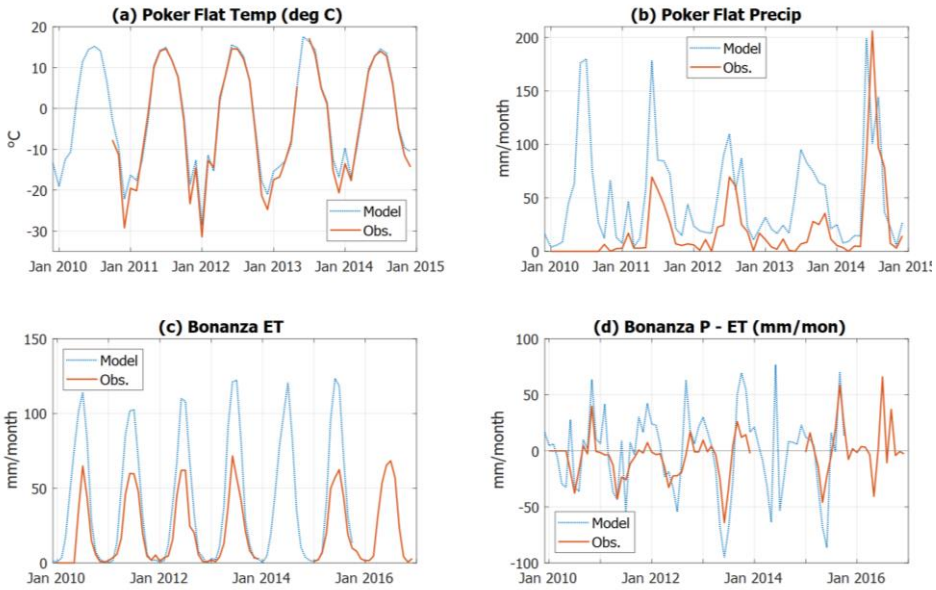



Figure 6. As in Figure 5, but for forest sites: (a) air temperature at Poker Flat, (b) precipitation, P, at Poker Flat, (c) evapotranspiration, ET, at Bonanza Creek, (d) P-ET at Bonanza Creek.

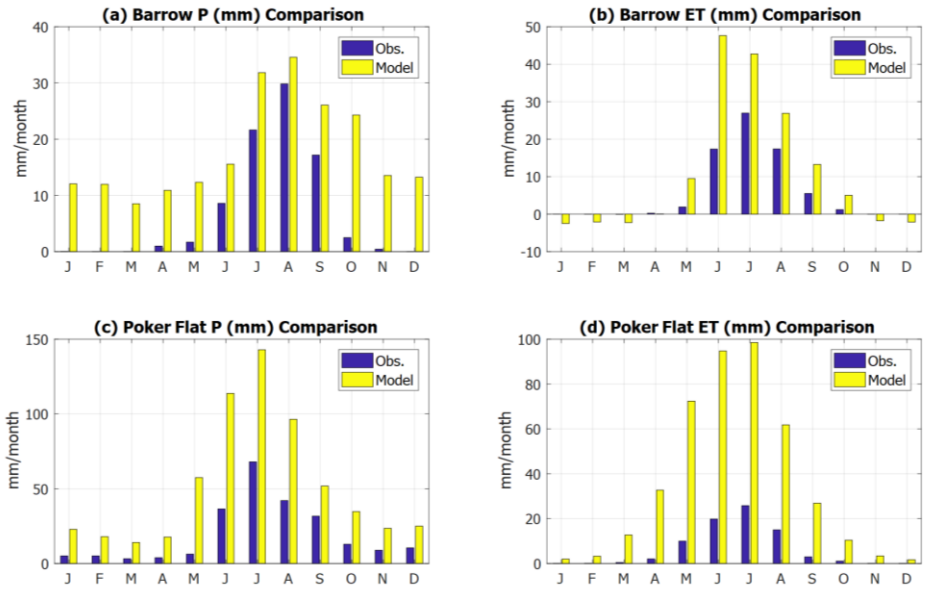



Figure 7. Monthly climatologies (averages over all available years) of tower measurements (blue) and corresponding model simulations (yellow) of (a) precipitation at Barrow, (b) evapotranspiration at Barrow, (c) precipitation at Poker Flat, and (d) evapotranspiration at Poker Flat.

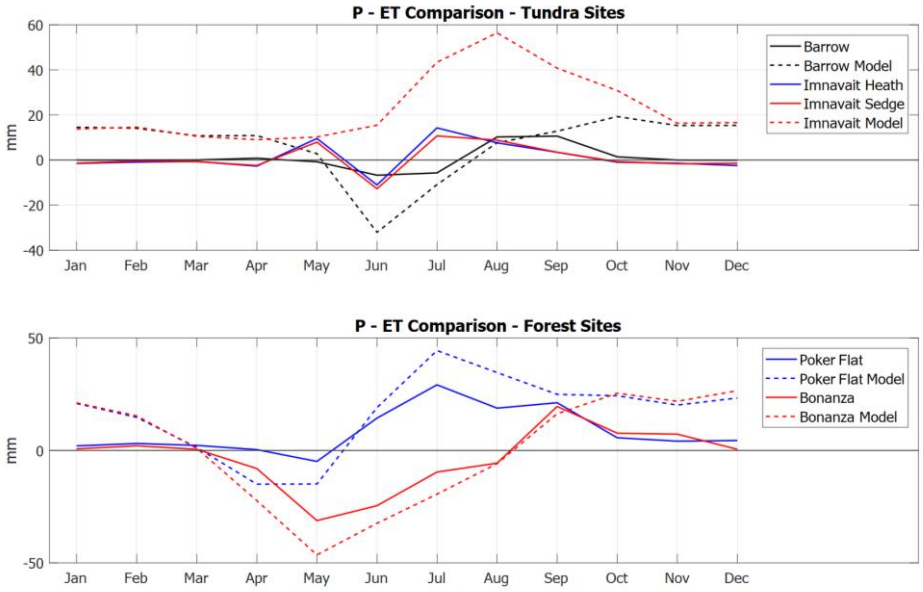



Figure 8. Climatological (average over all available years) monthly values of P-ET at tundra sites (upper panel) and forest sites (lower panel). Values derived from tower measurements are shown by solid lines, model-simulated values by dashed lines.

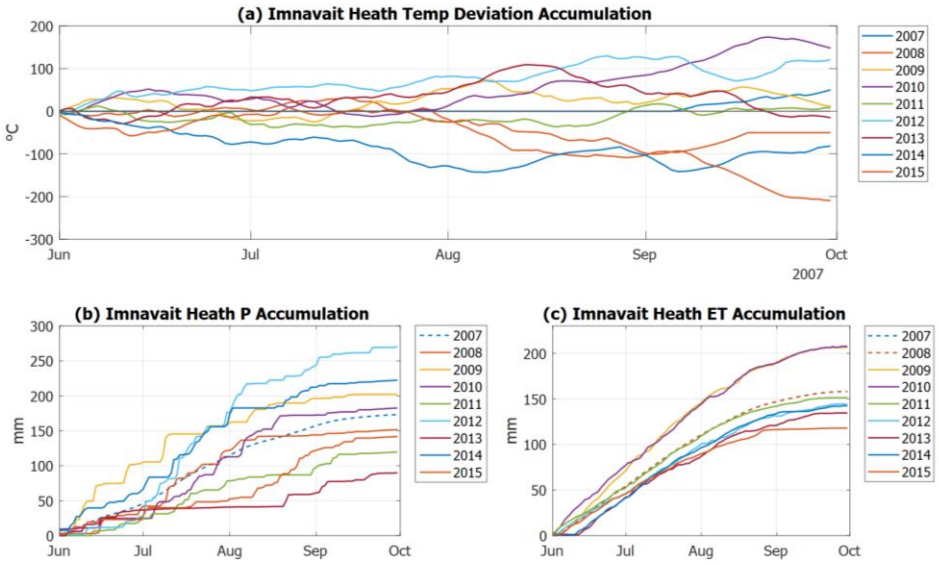



Figure 9. Seasonal (June through September) accumulations of (a) air temperature anomalies, (b) precipitation and (c) evapotranspiration at the Imnavair heath site. All values are based on tower measurements. Different years are color-coded; dashed lines are estimated values (see text) for years with large amounts of missing data.

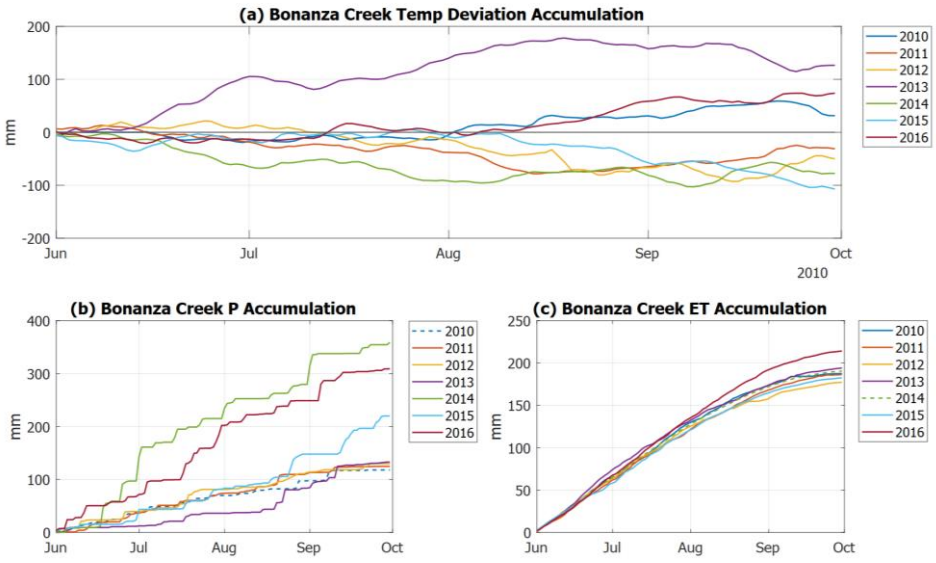



Figure 10. As in Figure 9, but for seasonal accumulations at Bonanza Creek.

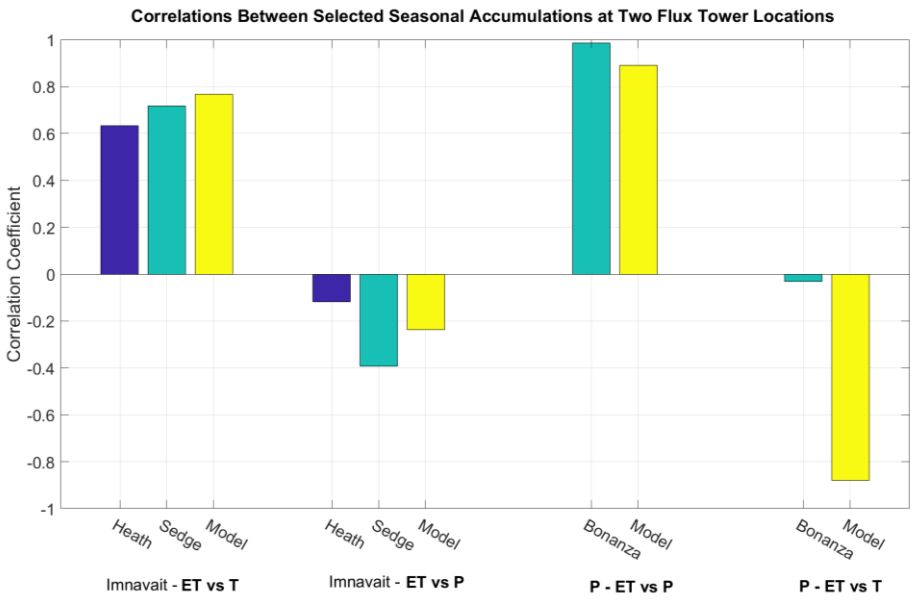



Figure 11. Sample correlations between yearly warm-season totals of ET, P-ET and driving variables (T, P). Blue and green bars are correlations based on tower data; yellow bars are correlations based on model simulation.

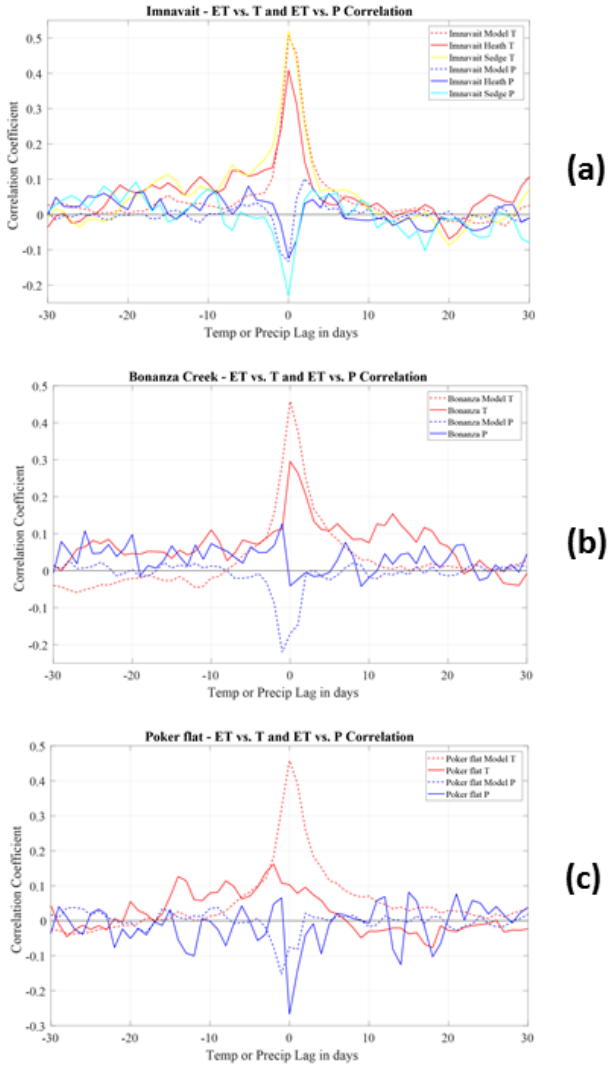



Figure 12. Cross-correlation functions of daily ET with daily temperature (red) and daily precipitation (blue) for (a) Imnavait Creek, (b) Bonanza Creek and (c) Poker Flat. Correlations are plotted as a function of the lag of P and T relative to ET (i.e., T and P lead ET to the left of zero lag; T and P lag ET to the right of zero lag). Solid lines are based on tower measurements, dashed lines on model output.

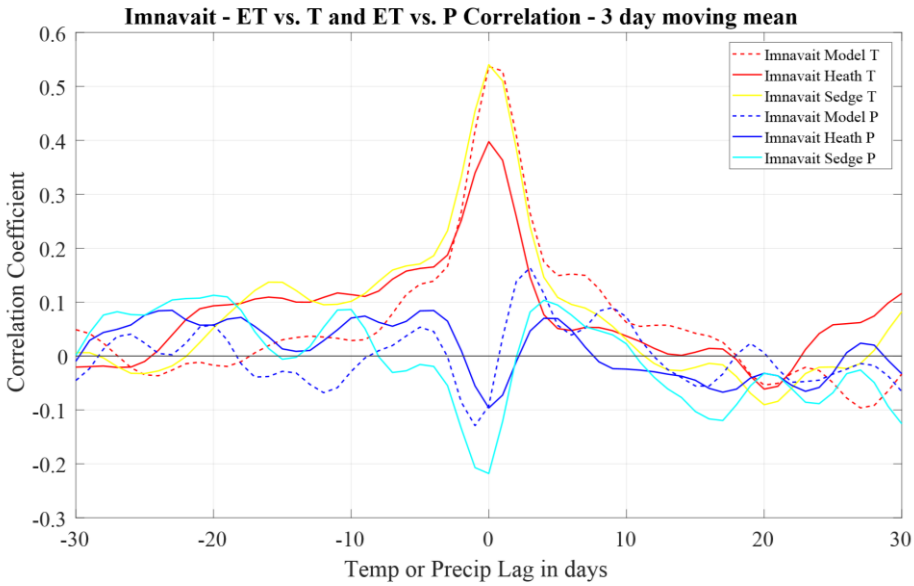



Figure 13. As in Figure 12, but for cross-correlations computed from 3-day running means of T, P, and ET at Imnavait Creek. Solid lines are based on tower measurements, dashed lines on model output. Values for wet sedge and heath are plotted in different colors.