

Attributing northern high-latitude precipitation change over the period 1966–2005 to human influence

Hui Wan, Xuebin Zhang, Francis W. Zwiers, & Seung-Ki Min
2015

Pacific Climate Impacts Consortium (PCIC)

PCIC Publications

© 2015 Wan, Zhang, Zwiers, & Min. This is an open access article distributed under the terms of the Creative Commons CC BY 4.0 License:

<https://creativecommons.org/licenses/by/4.0/>.

Original citation:

Wan, H., Zhang, X., Zwiers, F. W., & Min, S.-K. (2015). Attributing northern high-latitude precipitation change over the period 1966–2005 to human influence.

Climate Dynamics, 45(7–8), 1713–1726. <https://doi.org/10.1007/s00382-014-2423-y>

Downloaded from UVicSpace Research & Learning Repository

dspace.library.uvic.ca



**University
of Victoria**

Libraries

Attributing northern high-latitude precipitation change over the period 1966–2005 to human influence

Hui Wan · Xuebin Zhang · Francis Zwiers ·
Seung-Ki Min

Received: 4 July 2014 / Accepted: 17 November 2014 / Published online: 28 November 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract Using an optimal fingerprinting method and improved observations, we compare observed and CMIP5 model simulated annual, cold season and warm season (semi-annual) precipitation over northern high-latitude (north of 50°N) land over 1966–2005. We find that the multi-model simulated responses to the effect of anthropogenic forcing or the effect of anthropogenic and natural forcing combined are consistent with observed changes. We also find that the influence of anthropogenic forcing may be separately detected from that of natural forcings, though the effect of natural forcing cannot be robustly detected. This study confirms our early finding that anthropogenic influence in high-latitude precipitation is detectable. However, in contrast with the previous study, the evidence now indicates that the models do not underestimate observed changes. The difference in the latter aspect is most likely due to improvement in the spatial–temporal coverage of the data used in this study, as well as the details of data processing procedures.

Keywords Climate change · Detection and attribution · High-latitude precipitation

H. Wan · X. Zhang (✉)
Climate Research Division, Environment Canada, Toronto,
ON M3H 5T4, Canada
e-mail: xuebin.zhang@ec.gc.ca

F. Zwiers
Pacific Climate Impacts Consortium, University of Victoria,
Victoria, BC V8W2Y2, Canada

S.-K. Min
School of Environmental Science and Engineering, Pohang
University of Science and Technology, Pohang,
Gyungbuk 790-784, Korea

1 Introduction

Climate in the Arctic and the northern high latitudes is and will be strongly affected by human-induced forcing from factors such as the increasing concentration of greenhouse gases in the atmosphere. For example, due to meridional heat transport and positive snow/ice-albedo feedback, warming in the high latitudes is more rapid than in lower latitudes (Alexeev et al. 2005; Cai and Lu 2007). Warming has resulted in reductions in snow cover and sea-ice extent, and in thawing permafrost (Vaughan et al. 2013), which in turn is causing wide spread changes in the environment, affecting the very vulnerable ecosystems in the region. There is strong evidence indicating an anthropogenic contribution to the very substantial warming in the Arctic land surface temperature over the past 50 years (Gillett et al. 2008; Bindoff et al. 2013; Najafi et al. submitted), in intensification of Arctic hydrologic cycle (Peterson et al. 2002; Min et al. 2008a), and in the reduction of Arctic sea ice extent (Min et al. 2008b). Changes in high-latitude climate affect the global climate. For example, changes in the high-latitude water cycle affect the fresh-water influx to the Arctic Ocean and northern North Atlantic, potentially altering ocean convection in the subarctic seas and the strength of the Atlantic Ocean meridional overturning circulation. It is therefore important to understand the possible causes of observed changes in high-latitude water cycle and precipitation in particular.

Human influence has been detected in multiple facets of the global water cycle. Anthropogenic influence has been detected in the pattern of annual and seasonal precipitation over global land and sea areas (Zhang et al. 2007; Noake et al. 2012; Marvel and Bonfils 2013; Terray et al. 2012). Observed increases in surface and atmospheric humidity have also been robustly attributed

to human influence on the climate system (Willett et al. 2007; Santer et al. 2007, 2009). The increase in lower tropospheric moisture content enhances poleward atmospheric moisture transport, resulting in an increase in high-latitude precipitation. Indeed, high latitude precipitation has showed an increase in various datasets (e.g., Zhang et al. 2007; Min et al. 2008a; Noake et al. 2012; Polson et al. 2013) though trends in different datasets differ (Polson et al. 2013). An increase in high latitude precipitation is also a robust feature of climate model projected precipitation response to anthropogenic forcing in the future (Scheff and Frierson 2012).

We have previously linked the observed intensification of high-latitude land area precipitation to human influence using simulations from a limited set of climate models participating in the Coupled Model Intercomparison Project Phase 3 (CMIP3) and observations with limited spatial coverage ending in year 1999 (Min et al. 2008a). These models underestimated the observed changes in high-latitude land area precipitation, reducing confidence in the understanding of the causes of the observed changes.

It is now appropriate to revisit the detection and attribution analysis of Min et al. (2008a) because essential data for the detection analysis have been improved. These include observations for the characterization of past changes and model simulations for the estimation of climate response to external forcing and internal variability. Better availability of Russian climate data, especially for more recent years, substantially improves spatial and temporal coverage of observational data over the high latitudes. The simulations conducted with the climate models participating in the Coupled Model Intercomparison Project Phase 5 (CMIP5, Taylor et al. 2012) are now also available. CMIP5 simulations have several advantages relative to their CMIP3 counterparts: CMIP5 models generally have higher resolution than their predecessors, and they generally include a more complete representation of physical and biogeochemical climate processes; CMIP5 simulations cover the more recent period; and there are many more models, providing more robust estimates of climate responses to external forcings as well as better quantification of internal variability. The main objective of this paper is therefore to revisit the detection and attribution analysis of our earlier study (Min et al. 2008a) using more recent and spatially complete observational data and CMIP5 simulations. We pay particular attention to the robustness of the results by using various sensitivity tests. The remainder of the paper is structured as follows: we describe observational and climate model simulated precipitation data in Sect. 2. Detection methods and data processing procedures are detailed in Sect. 3. The main results are presented in Sect. 4 together with results from sensitivity tests. Conclusions and discussion are given in Sect. 6.

2 Data

2.1 Observational data

The limited availability of observational data both in space and time is a significant challenge for studying changes in high-latitude and Arctic precipitation. Coverage from the existing global precipitation data sets such as the Global Historical Climate Network monthly precipitation data version 2 (Peterson and Vose 1997) is very limited over the high latitudes especially for the more recent years. For example, the GHCN dataset includes only 16 Canadian stations and 50 Russian stations in 2005. Precipitation trends computed from different datasets differ due to differences in station coverage and in data processing methods (Noake et al. 2012). To reduce these uncertainties, we compile station precipitation data from various sources to produce a gridded product for the region with improved spatial and temporal coverage that is based on as many stations as possible.

We used monthly precipitation data for more than 450 Canadian stations from the Second Generation Adjusted Precipitation dataset (Mekis and Vincent 2011). We also extracted long-term Alaskan station data from NOAA's National Climate Data Center (NCDC) US monthly precipitation dataset. This data set contains 737 Alaskan stations, but most have only very short records; we therefore use 82 Alaskan stations with long-term records that satisfy the criteria described below.

The GHCN monthly precipitation data set has Russian data for about 200 stations until 1995, but this number drops to about 50 stations after 1996. For this reason, we use monthly precipitation data from the Russian Met Service, available online at <http://meteo.ru/english/climate/precip.php>, in place of GHCN Russian data. This Russian dataset consists of 518 stations with data starting from 1966. The dataset providers do not include monthly data prior to 1966 due to data homogeneity issues. Groisman and Rankova (2001) indicated that in 1966–1967, the Hydrometeorological Services of the former Soviet Union introduced a wetting correction to each nonzero precipitation measurement (0.2 mm for liquid and 0.1 mm for frozen precipitation) when at least one drop was extracted from the gauge to account for light precipitation “to the last drop”. This correction resulted in a substantial increase in the number of precipitation days and affected annual total precipitation. This inhomogeneity in Russian precipitation data is difficult to remove because the number of trace events in different years is hard to determine. Fortunately, no documented changes in data measurement and processing techniques have occurred subsequently, and thus time series of regionally averaged precipitation totals after that year can likely be considered to be homogeneous even if

individual station data may not be homogeneous. Regional averaging over large domains, as conducted in this study, would likely include offsetting effects from inhomogeneities causing individual stations to either over- or under-report precipitation due to changes in station location. This dataset offers a substantial improvement in coverage over the GHCN dataset, especially after the late 1990s. We also used data from 4 high-latitude Chinese stations from the GHCN.

High latitude European data are obtained from the European Climate Assessment and Dataset (ECA&D, <http://www.ecad.eu>). The ECA&D team collects, quality controls, and tests data homogeneity for daily observations at meteorological stations throughout Europe and the Mediterranean. The team receives data from 62 countries in the region. The number of stations in the ECA&D dataset is about 7000 for monthly precipitation amounts, with best data coverage between 1960 and 2000 (Klok and Klein Tank, 2008). The ECA&D dataset also contains some Russian stations, which are excluded because we use station data directly available from the Russian Meteorological Service.

Among all station data used in this study, only Canadian data have been adjusted to correct wind under-catch. Since much of the high-latitude winter precipitation falls as snow, and as temperature has been increasing, precipitation data that are not corrected for wind under-catch may contain small spurious upward trends related to changes in catch efficiency. This would occur because the relative frequency of liquid rather than frozen precipitation would rise with increasing temperature, thereby improving the catch efficiency of the gauge. However, the adjustments that have been made to Canadian data do not change annual precipitation trends much because the transition season, i.e., a season when precipitation can fall as rain or snow, is short and contributes only a small fraction of the total precipitation. Given this, we do not expect the changes in wind under-catch would significantly alter precipitation trends reported here.

In this analysis, we use stations located in north of 50°N since much of precipitation received in this region discharges to the Arctic basin (Peterson et al. 2002). We also conducted the analyses using the spatial domain of Min et al. (2008a), i.e., stations that are located north of 55°N, and found that results were not sensitive to the choice of the southern boundary of the spatial domain. As Russia covers more than a third of the spatial domain and Russian precipitation data are considered to be homogeneous only from 1966, we restrict our analysis to the time period starting from 1966. Since most historical climate simulations end in 2005, the time period used in this study is therefore 1966–2005. Overall, a total of 3832 stations from various data sources are compiled and used in this study (station

selection criteria are discussed in Sect. 3.1). Compared with the data used in (Min et al. 2008a), which was based on GHCN monthly precipitation combined with adjusted Canadian precipitation data, this compilation most triples the number of available stations, providing substantial improvements in spatial coverage (Figs. 1, 2).

2.2 Climate model simulations

We use monthly precipitation totals from the CMIP5 archive to estimate precipitation responses to different external forcings and to estimate natural internal variability. At the time of analysis, the following simulations were available and are used in this study: 158 simulations conducted with 32 models under historical forcing (ALL) that includes both natural and anthropogenic forcing combined, 59 simulations conducted with 14 models under historical natural forcing only (NAT), 41 simulations conducted with 8 models under historical anthropogenic forcing only (ANT), 45 simulations under historical greenhouse gas forcing only (GHG) from 11 models, 33 simulations under historical anthropogenic aerosol forcings only (AA) from 7 models and 15 simulations under land use forcing only (LU) from 3 models. All of the forced simulations start in the late nineteenth century and end in the early twenty-first century. Over 24000 model-years of pre-industrial control (CTL) simulations conducted with 48 models are also available from the Program for Climate Model Diagnosis and Intercomparison (PCMDI) website. The number of simulations used from each model for different forcing experiments is listed in the Table 1. Different horizontal resolutions have been used by different models; we interpolated monthly precipitation from individual models to the 5° × 5° grid of the observations using a simple distance weighting average remapping method implemented in the Climate Data Operators (<https://code.zmaw.de/projects/cdo>) prior to the analyses. While other interpolation methods could be used, the choice of interpolation method should not have much impact on the long-term variation of large-region precipitation averages and thus should not affect final results.

3 Methods and data processing

3.1 Data processing

We conduct detection and attribution analyses based on the space–time evolution of observed precipitation amounts, considering both annual and two half-year amounts for winter (from October to March) and summer (from April to September). Anomalies of annual and half-year amounts are computed for individual stations as the average of the

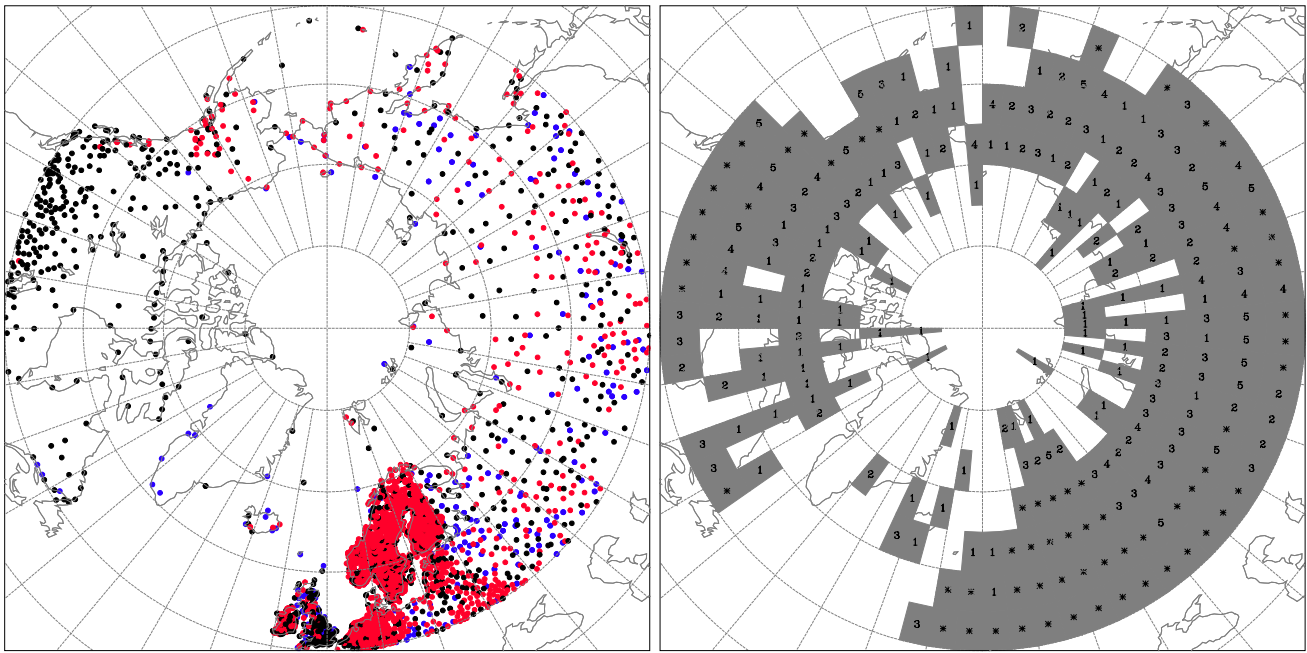


Fig. 1 Left: spatial distribution of long-term precipitation stations. Long-term stations (with at least 20 year non-missing values during 1966–1995) used in this study only are marked as *red dots*. While those used in Min et al. (2008a) only are marked as *blue dots*, and stations used in both studies are marked as *black dots*. Right spatial

distribution of grid boxes with sufficient data for the analyses. The number of stations within a $5^\circ \times 5^\circ$ longitude by latitude grid box is marked by a numerical value; grid boxes with more than 5 stations are marked with *asterisks*

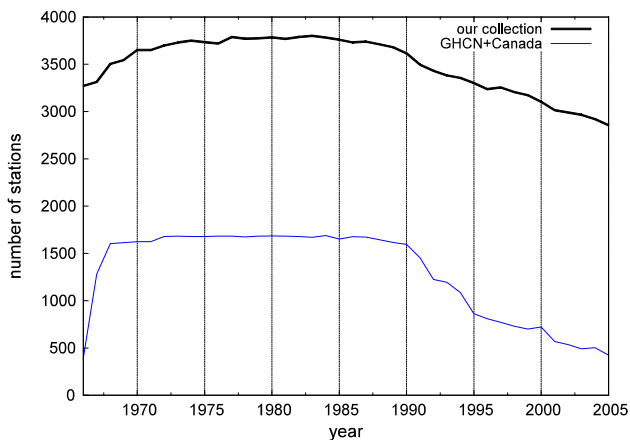


Fig. 2 Time series of the number of stations with monthly reports used in this study (*black*) and the GHCN combined with Canadian stations (*blue*) for the period 1966–2005. Only long-term stations that have at least 20 year data in the 1966–1995 base period are considered in both datasets

available monthly anomaly values if there are at least 7 or 4 months of data within the year or the half-year respectively. Estimates of annual or half-year precipitation anomalies were obtained by multiplying the annual or half-yearly averages of monthly values by 12 or 6 respectively. Station anomalies were computed relative to the 1966–1995 mean

for stations that have at least 20 years of non-missing data during this period, and were gridded at the $5^\circ \times 5^\circ$ latitude-longitude resolution by averaging available station anomaly values within each grid box. Grid box values were averaged into non-overlapping 5-year mean annual and half-year precipitation anomalies for 1966–1970, 1971–1975, to 2001–2005. A 5-year mean grid box value is only computed if annual or half-year values are available for the grid box for at least three of the five years. Grid boxes that have at least 6 non-missing five-year means during 1966–2005 are retained for the subsequent analysis. There are 256 such grid boxes north of 50°N (Fig. 1). Among the 256 grid boxes, 207 are located north of 55°N , giving 20 % increase in number of grid boxes compared to Min et al. (2008a) who had only 171 grid boxes in the region.

As will be described further in Sect. 3.2, we assume that observed precipitation change consists of an externally forced component plus additive “noise” that reflects the effects of natural internal variability in the climate system. Detection and attribution analysis is usually based on a low dimension description of observed changes so that there is enough model control run output to estimate the covariance structure of this noise. Thus the dimensionality of the space in which the analysis is conducted is typically much smaller than the space–time dimensionality of the original raw data. Dimension reduction involves filtering data

Table 1 List of models, experiments, and ensemble sizes available for this analysis

Model	ALL	ANT	NAT	GHG	AA	LU
ACCESS1-3	3					
Bcc-csm1-1	3 (3)					
Bcc-csm1-1-m	3 (3)					
CNRM-CM5	10 (10)	10	6 (6)	6		
CSIRO-Mk3-6-0	10	5	5 (5)	5	5	
CanESM2	5 (5)		5 (5)	5	5	5
CCSM4	6		4	3		
CESM1-CAM5-1-FV2	4					
CESM1-CAM5	3					
CESM1-FASTCHEM	3					
EC-EARTH	7 (5)					
FGOALS-g2	5		3			
FIO-ESM	3					
GFDL-CM2p1	10 (10)					
GFDL-CM3	5	3	3	3	3	
GFDL-ESM2G	3					
HadCM3	10					
HadGEM2-ES	4 (3)		4 (4)	4		
IPSL-CM5A-LR	6	3	3 (3)	3		
IPSL-CM5A-MR	3		3 (3)	3		
MIROC5	5 (5)					
MIROC-ESM	3		3	3		
MPI-ESM-LR	3					
MPI-ESM-MR	3					
MRI-CGCM3-p1	3 (3)					
NorESM1-M	3 (3)					
GISS-E2-R-p1	6 (5)	5 (5)	5 (5)	5	5	5
GISS-E2-R-p2	5					
GISS-E2-R-p3	6 (6)	5 (5)	5 (5)		5	
GISS-E2-H-p1	5 (5)	5 (5)	5 (5)	5	5	5
GISS-E2-H-p2	5					
GISS-E2-H-p3	5	5 (5)	5 (5)		5	
SUM (models)	158 (32)	41 (8)	59 (14)	45 (11)	33 (7)	15 (3)

The experiments include simulations under historical forcing (ALL), historical anthropogenic forcing only (ANT), historical natural forcing only (NAT), historical greenhouse gas forcing only (GHG), historical anthropogenic aerosol forcing only (AA) and land use forcing only (LU). Numbers in the parentheses are the ensemble sizes of extended simulations that end in year 2012

in space and/or time. There is currently not a commonly accepted objective method to perform dimension reduction.

Here, we reduce temporal dimension by using 5-year mean values. We reduce the spatial dimension by computing area-weighted averages over different sizes of land regions. These includes (1) one region, i.e., considering the entire spatial domain as one region, (2) two regions, i.e., dividing the spatial domain into southern (50°N–60°N) and northern (north of 60°N) subdomains, (3) three regions, dividing the spatial domain into northern North America (40°W–180°W, NA), western Eurasia (30°W–60°E, EU) and eastern Eurasia (60°E–180°E, AS), (4) six regions, further sub-dividing each of these three regions into southern (50°N–60°N) and northern (north of 60°N) sub-regions. That is, our detection and attribution analysis is conducted

on space–time series of non-overlapping 5-year means for 1966–2005 over the land regions in four different configurations that are designated by the number of sub-regions (1, 2, 3 or 6 respectively). This produces descriptions of high-latitude precipitation change of 8, 16, 24 and 48 space–time dimensions respectively.

The forced response, or signal, corresponding to an external forcing is estimated by calculating the multi-model ensemble mean under that forcing, computed by first calculating ensemble means from individual models and then averaging over available models. Model-simulated monthly precipitation data under various external forcings, after interpolation onto the common 5° × 5° grid, are masked using the availability of the gridded observed precipitation data to mimic the availability of observations. Annual and

half-year precipitation anomalies relative to 1966–1995 are then computed in a way similar to that for the observations. Figure 3 shows the 5-year mean time series of

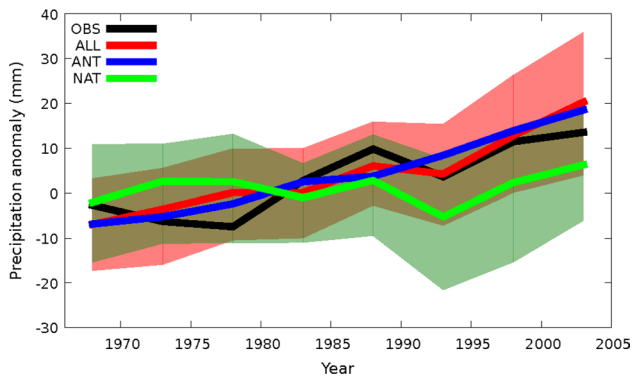


Fig. 3 Five-year mean precipitation anomalies (in mm, relative to 1966–1995) in the northern high-latitudes from observations (*black line*) and multi-model ensembles under “ALL” forcing (*red line*), “ANT” forcing (*blue line*) and “NAT” forcing (*green line*). Shading shows the 5–95 % ranges of the individual model simulations

annual precipitation anomalies (relative to the 1966–1995 climatology) averaged across the northern high-latitude land area from observations and multi-model simulations. There is a good match in the long-term changes of precipitation anomalies between observations and the multi-model ensemble-mean ALL forcing response, but greater discrepancy between observations and the multi-model ensemble mean NAT forcing response.

Figure 4 shows the spatial distribution of annual precipitation trends computed from observations and from the ensemble mean of model-simulated responses to ALL, ANT, and NAT forcings. Observations show a marked increase over the high-latitude land areas except in eastern Siberia and the Canadian Prairies. Trends in model-simulated responses to ALL and ANT forcings are largely homogeneous over space, with very little spatial variation in magnitude. Trends in the NAT forcing response are of very small magnitude and of mixed sign across the domain. The fact that model simulated trends in ALL and ANT are spatially homogeneous means that spatially averaging data over large domains would enhance the signal to noise

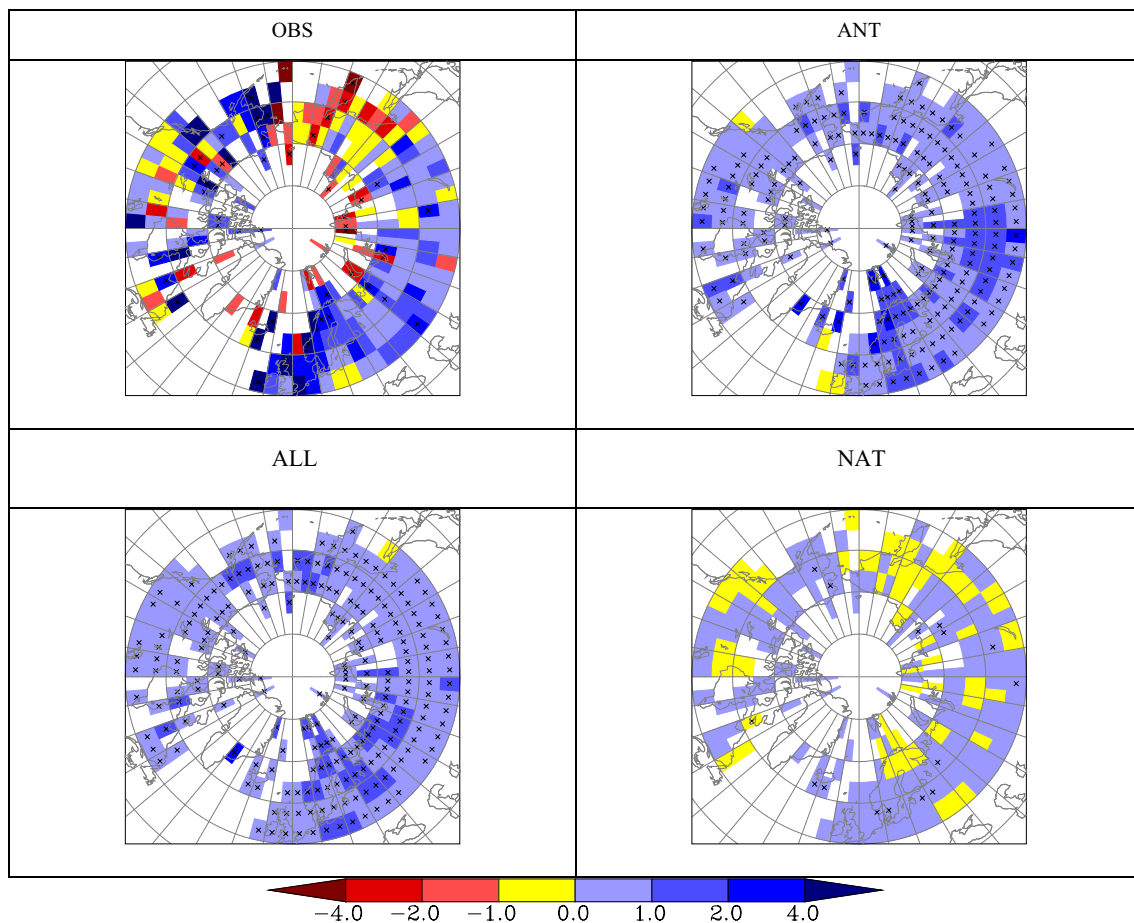


Fig. 4 Annual precipitation trends (mm/year) during 1966–2005 computed from the observations (OBS) and model simulated responses to ALL, ANT, and NAT forcings. Trends were computed only for grid boxes with at least 30 years of data

ratio. On the other hand, conducting detection and attribution analyses with some spatial structure provides greater opportunities to evaluate model-simulated internal variability via the residual consistency test (discussed in Sect. 3.2), thereby potentially increasing confidence of detection results.

The estimates of internal variability required in the detection analysis are obtained from within-ensemble differences as well as pre-industrial control simulations. Forced simulations are typically available for the entire twentieth century, so we also make use of within-ensemble differences for the earlier part of the twentieth century simulations for estimating internal variability. The monthly precipitation data from forced simulations for 1926–1965 are processed in the same way as the simulations for 1966–2005, with 1926 treated as 1966, 1927 and 1967, and so on. In this way, considering both the 1926–1965 and 1966–2005 periods, the forced ALL, NAT, ANT, GHG, AA, and LU simulations yield a total of 316, 118, 82, 90, 66, and 30 40 year chunks of within-ensemble differences respectively. When using these differences to estimate the covariance matrices of internal variability, we account for the fact that a given ensemble of size n provides only $n-1$ independent realizations of within ensemble differences. Pre-industrial control simulations are also divided into multiples of 40-year chunks, with chunks shorter than 40 years discarded. There are 604 40 year chunks in total from control simulations. When using the control run chunks, we also account for the reductions, by one, in the numbers of independent realizations that come about because control run chunks from different control simulations are treated as ensembles. Thus, when combined, 1306 40 year chunks of noise data are available, each of which is processed as the gridded observations to obtain space–time averages of anomalies appropriate for the different domain configurations that are considered. This very large sample of model-simulated internal variability provides an unprecedented opportunity to confidently estimate the covariance structure of internal variability, even in the highest dimensional (48-dimension) analysis space considered in this studied.

3.2 Detection methods

Space–time series of observations and model simulated precipitation responses to external forcings (the “signals”) are first compared qualitatively by computing correlation coefficients between observations and simulations. The statistical significance of the correlations is determined by comparing them with correlations between observations and each of the 1306 chunks of noise data. This simple method does not optimize the signal-to-noise ratio nor provide a quantitative measure of the magnitude of model-simulated response relative to that in the observations.

Nevertheless, it provides an easy-to-understand view of the similarity between observed and model-simulated changes. Santer et al. (1995) used a correlation-based method to conduct detection and attribution analysis for temperature changes.

The generalized linear regression based optimal fingerprint method (Allen and Stott 2003) is used to quantify the magnitude of model-simulated responses to observed changes. This method establishes a linear relationship between the observations (\mathbf{y}) and the climate responses to the external forcings (\mathbf{X}) such that $y = (\mathbf{X} - \mathbf{v})\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ where $\boldsymbol{\beta}$ is a vector of regression coefficients or scaling factors, $\boldsymbol{\varepsilon}$ is the regression residual representing internal variability, and \mathbf{v} represents the effect of internal variability that remains in \mathbf{X} since multi-model averaging of forced runs is not able to remove all vestiges of internal variability. In our analysis, vector \mathbf{y} and the columns of \mathbf{X} have dimension 8, 16, 24 or 48 depending upon whether we account for the time evolution of regionally averaged 5-year mean precipitation in 1, 2, 3 or 6 Northern Hemisphere subdomains north of 50°N. We estimate $\boldsymbol{\beta}$ using the total least square method of Allen and Stott (2003). As the number of runs from different models may differ, the variance in the multi-model ensemble mean \mathbf{X} due to noise, that is, the variance of \mathbf{v} , is not equal to the internal variability divided by the total number of runs used. It is the average of variance from individual model ensemble mean.

We conduct single-signal and two-signal analyses in which one and two response patterns are considered respectively. The response patterns are estimated from multi-model ensemble means computed from all available models under the ALL, ANT, and NAT forcings to detect the presence of response to a particular forcing in the observations. Two-signal analyses are conducted to examine whether the responses to different forcings can be separately detected. In this case, the response patterns are estimated from multi-model ensemble means computed from 8 GCMs that produced both ANT and NAT simulations. Alternatively, ANT could have been estimated as the difference between ALL and NAT as determined from the full ensemble of models that provided ALL and NAT simulations. However, this would have introduced the possibility that some features of the estimated ANT could have been due to differences in the collections of models participating in the ALL and NAT ensembles. Using signals from 8 common models avoids that possibility. Two independent estimates of the internal variability covariance matrix \hat{C}_{N1} and \hat{C}_{N2} required for optimization and uncertainty analysis (Hegerl et al. 1996) are estimated from the noise data consisting of within ensemble differences and pre-industrial control simulations constructed as described in Sect. 3.1. \hat{C}_{N1} and \hat{C}_{N2} are each estimated from 653 40 year chunks of noise data, with half of the available control run chunks

and within-ensemble differences allocated to each sample of noise chunks.

Optimal detection and attribution analysis very often requires a reduction of dimensionality. This is typically done by projecting both observations and simulations onto leading empirical orthogonal functions (EOFs) of internal variability and using the residual consistency check to determine the number of EOFs to be retained in the analysis (Allen and Tett 1999, Allen and Stott 2003). Because we have a relatively small number of dimensions (i.e., $8 - 1 = 7$ for the one-region series and $6 \times (8 - 1) = 42$ for six-region series) and large samples of noise data, our analysis is performed in the full space that is spanned by the observations. Note that one temporal degree of freedom is lost in each region because all data are expressed as anomalies. Consequently, we do not require a subjective determination of EOF truncation point. Results of the residual consistency tests (Allen and Stott 2003) do not suggest inconsistency between the model simulated variability and the regression residual that is obtained.

We also perform some sensitivity tests to understand differences between our results and those reported in Min et al. (2008a). There are many differences between the two studies, including the differences in the analysis period, space-time coverage of observational data, and in the climate models used to estimate signals and internal variability.

4 Detection results

Table 2 summarizes the correlation coefficients between the observations and the model-simulated responses, and their corresponding statistical significance levels. Results show that for annual, winter, summer, or two half-year (winter-summer) combined precipitation, the model-simulated responses to ALL or ANT forcing are generally detected in the observations at below the 5 % significance level regardless of spatial configuration. One exception to this general finding is that ALL was detected in winter precipitation at the 6.3 % significance level when the computation includes a 2-region domain. NAT is not detected in either annual or winter precipitation, even at the 10 % level. Response to ALL and ANT is not generally detected in summer precipitation. Response to NAT seems to be sometimes detectable but the detection is not robust across different regional configurations. Overall, the use of the correlation-based, non-optimal method is able to produce robust detection of precipitation responses to ALL and ANT in annual and winter precipitation with different spatial configurations.

Figure 5 shows scaling factors and their 90 % confidence intervals for single-signal detection of ALL and ANT under different spatial configurations. Results shown here are based on full rank estimates of the covariance matrix. The

Table 2 The correlations between observed precipitation amounts and model simulated responses

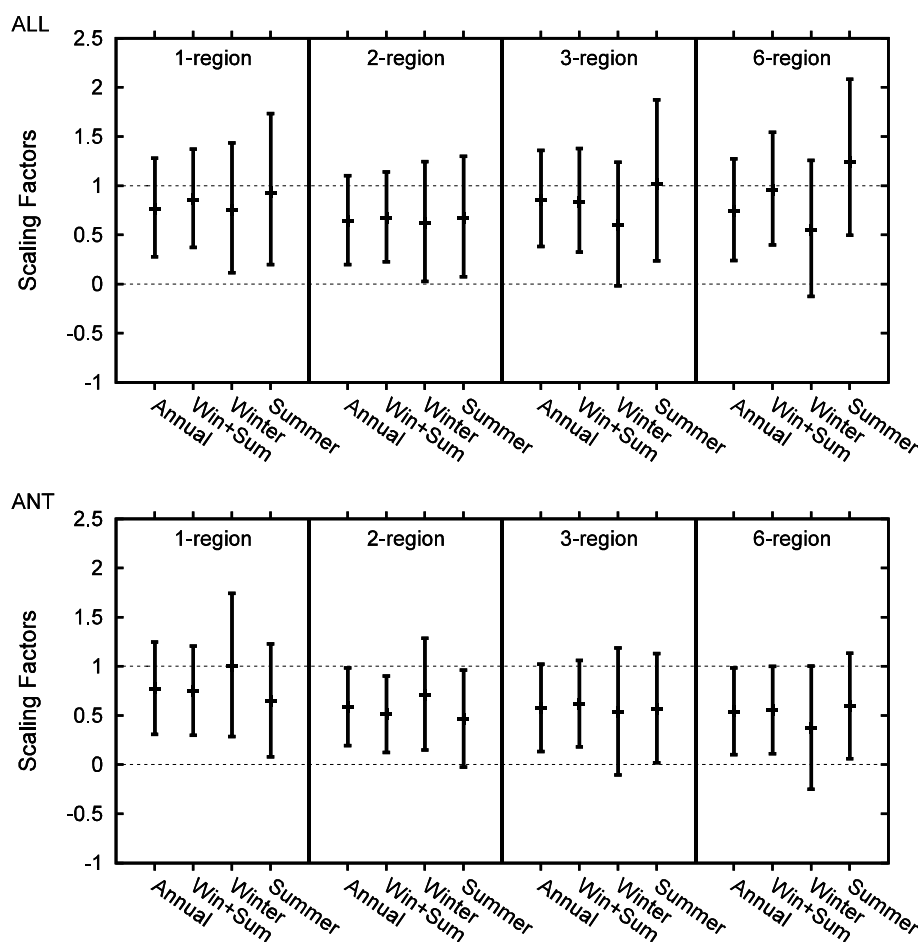
	ALL	ANT	NAT
1-region			
Annual	0.86 (0.4 %)	0.88 (0.2 %)	0.30 (27.1 %)
Winter	0.62 (6.3 %)	0.70 (3.1 %)	−0.06 (54.3 %)
Summer	0.71 (3.0 %)	0.57 (7.4 %)	0.66 (4.1 %)
Winter + Summer	0.68 (0.2 %)	0.69 (0.1 %)	0.15 (32.4 %)
2-region			
Annual	0.73 (0.5 %)	0.72 (0.6 %)	0.29 (24.0 %)
Winter	0.64 (2.8 %)	0.71 (1.5 %)	0.02 (48.2 %)
Summer	0.64 (2.4 %)	0.57 (4.4 %)	0.31 (19.5 %)
Winter + Summer	0.65 (0.2 %)	0.65 (0.2 %)	0.14 (31.2 %)
3-region			
Annual	0.50 (1.6 %)	0.41 (4.8 %)	0.28 (12.9 %)
Winter	0.47 (3.5 %)	0.46 (3.8 %)	0.08 (38.5 %)
Summer	0.31 (7.7 %)	0.24 (12.9 %)	0.33 (6.7 %)
Winter + Summer	0.41 (1.0 %)	0.36 (2.1 %)	0.20 (14.3 %)
6-region			
Annual	0.32 (2.9 %)	0.32 (3.1 %)	0.21 (11.5 %)
Winter	0.41 (1.2 %)	0.37 (1.8 %)	0.06 (37.9 %)
Summer	0.21 (8.0 %)	0.16 (14.9 %)	0.31 (2.4 %)
Winter + Summer	0.31 (0.3 %)	0.26 (1.8 %)	0.17 (9.7 %)

Shown in parentheses is the highest significance level (in percent) at which a 1-sided test of the null hypothesis of zero correlation can be rejected as determined by a Monte-Carlo method using the 1,306 chunks of noise data as pseudo observations. Note that smaller values indicate greater significance. The correlations significant at the 5 % level based on one-sided test are shown in bold

residual consistency check does not show evidence of inconsistency between the regression residual and model-simulated variability in any configuration. The scaling factors for ALL for annual or two half-year combined and summer precipitation are significantly greater than zero (5 % one-sided test) and the 90 % confidence intervals on the scaling factors include 1. Thus the combined effect of anthropogenic and natural forcings is robustly detected in the observed annual and summer precipitation, and is consistent with observed changes. The effect of ALL on winter precipitation is also detectable, but not robustly. The 90 % confidence intervals of the scaling factors for ANT and annual precipitation all lie above zero, but in some cases have upper bounds that are less than one. This indicates that the model-simulated responses to ANT forcing may be larger than the observed annual precipitation change. ANT seems to be detectable in summer and winter precipitation as well, but not robustly.

Figure 6 displays the 90 % confidence regions and the corresponding marginal confidence intervals from the two-signal detection analyses based on model simulated responses to ANT and NAT forcings by 8 GCMs for the 5-year non-overlapping means of annual or two half-year

Fig. 5 Scaling factors and their 5–95 % uncertainty ranges for ALL (*upper panel*), ANT (*lower panel*) in single-signal analyses for annual, two half-year combined, winter, and summer precipitation and for different spatial configurations



combined, winter and summer half-year precipitation with different spatial configurations. Results based on the model-simulated responses to ALL and NAT are very similar (not shown). For annual and winter precipitation, the influence of ANT can be detected in the 1- and 2-region analysis configurations, but the response to NAT is not detected. For summer precipitation, NAT is detected in the 6-region configuration, but ANT is not detected. Two-signal analyses conducted using multi-model ensemble means computed from all available model runs produce essentially the same results. Overall, the tenuous NAT detection results should not be considered to be reliable given the inability to detect NAT in annual means and unphysical scaling factor estimates for winter NAT signals. As noted previously, the estimated NAT signals contain little spatial or temporal structure, and therefore provide little in the way of detection power. Higher dimensional analyses (e.g., using 3-year or 1-year annual and half-year means rather than 5-year means, and therefore requiring higher dimension detection spaces) also fail to detect NAT (not shown) and tend to deteriorate detection results.

Overall, the analyses based on 5-year mean precipitation with different spatial dimension reductions indicate that the simulated response to combined external forcing

(ALL) on high-latitude precipitation can be detected in annual and summer half-year accumulations. The scaling factor is in general consistent with one, indicating a consistency between observations and the model simulated response. There is also some evidence of the influence of ALL in winter precipitation. The detection result on annual precipitation accumulations is robust to the use of different detection methods (e.g., optimized versus unoptimized) and to different ways of considering spatial information. The influence of anthropogenic (ANT) forcing can also be detected in annual precipitation. Two-signal detection analysis seems to be able to detect ANT forcing in annual and winter precipitation, and possibly NAT in summer precipitation, but the detection results are not as robust.

We repeated the detection analyses using 8- and 10-year mean precipitation anomalies. Results are very close to those of 5-year mean series. One exception is that NAT is not detectable in either single-signal or two-signal analyses (not shown). As some models extended their simulations to 2012 and human influence is expected to increase with time, we also conducted an analysis for the period 1966–2012. Fewer model runs available for the signal estimation for this period (as indicated in the parentheses in Table 1);

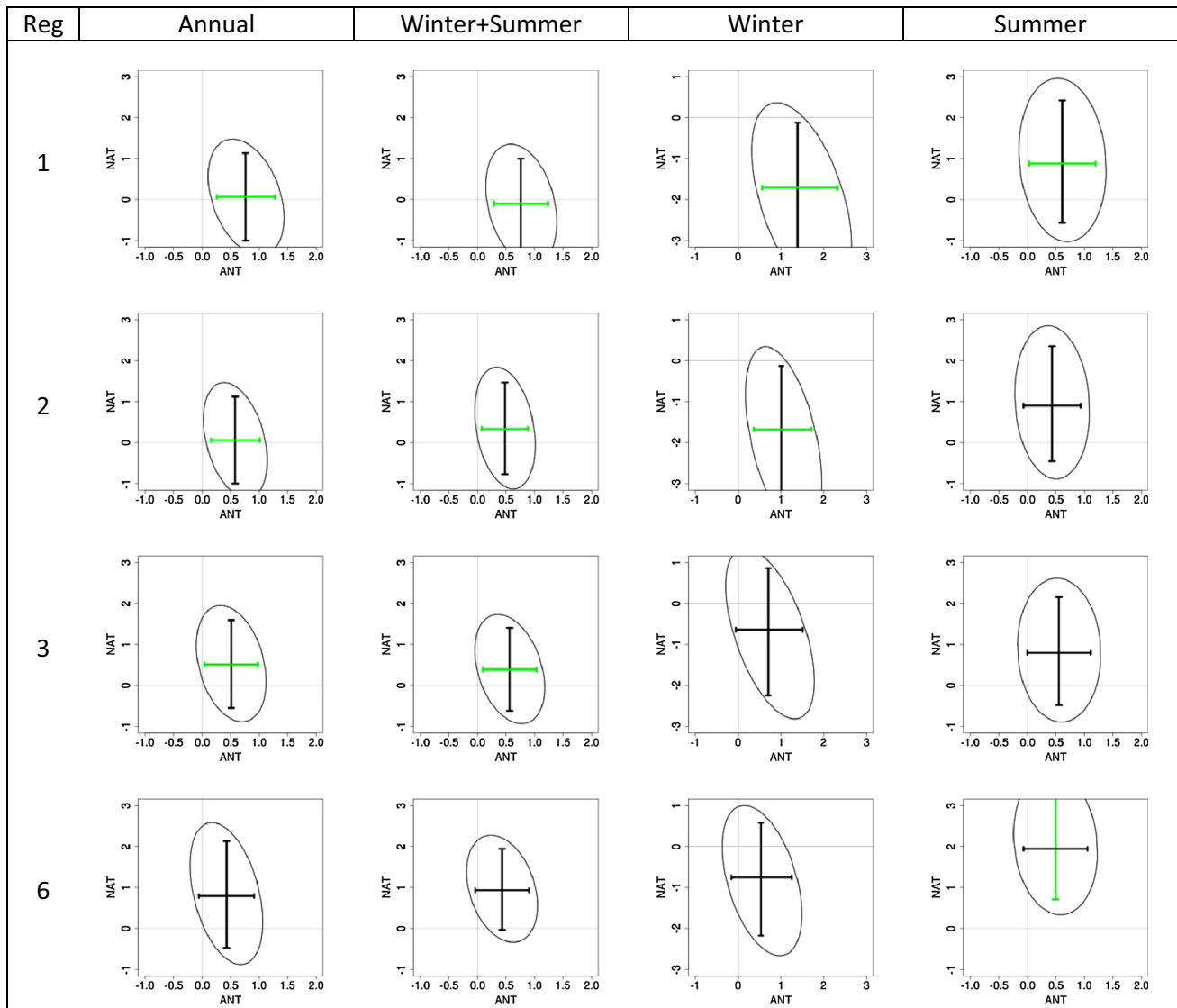


Fig. 6 Scaling factors and their joint 90 % uncertainty regions from two-signal analyses with model simulated responses to ANT and NAT. The marginal 5–95 % confidence intervals for the scaling factors are also shown. Scaling factors are shown in *green* when the forcing is detected

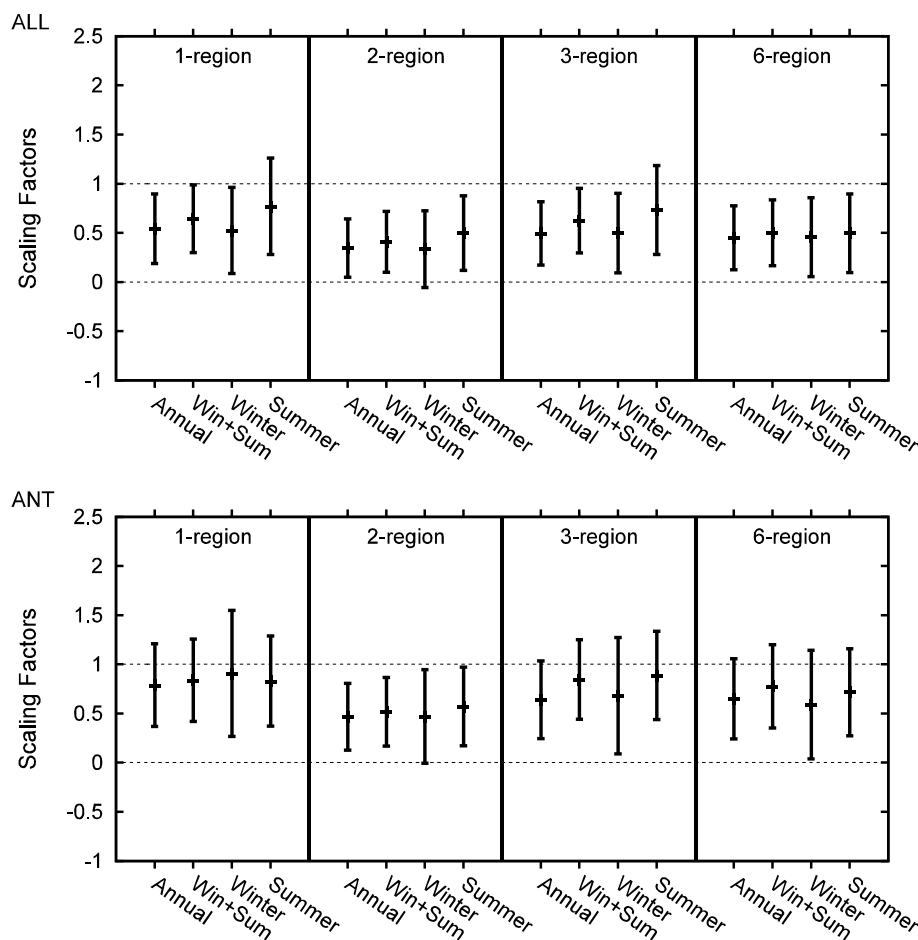
in total, 66 historical ALL runs are available from 13 models, together with 35 ANT runs from 6 models and 46 NAT runs from 10 models. In this case, detection analysis is conducted in 6-year non-overlapping mean series with the last data values being average from 5-year. Results from the extended analysis are essentially the same. As an example, Fig. 7 shows scaling factors and their 90 % confidence intervals for single-signal detection of ALL and ANT under different spatial configurations.

5 Comparison with earlier results

Min et al. (2008a) concluded that anthropogenic forcing from greenhouse gases and sulfate aerosols combined

had contributed to the observed high-latitude precipitation increase during the latter half of the twentieth century and that climate model simulated precipitation responses to anthropogenic forcing are weaker than in the observations. The updated analyses discussed here confirm Min et al. (2008a)'s finding with regard to the detection of anthropogenic influence in northern high-latitude precipitation. However, there is also a notable difference with regard to the magnitude of model-simulated changes: the Min et al. (2008a) results suggested models underestimated observed changes by a factor of 2 or more, whereas this study shows that model simulated responses are consistent with observations, if not too large. Many aspects are different between the two studies, including different model simulations, sampling uncertainties in the estimation of model

Fig. 7 Same as Fig. 5 but using data over 1966–2012



responses and internal variability, different space–time coverage of observational data, and different data processing procedures. All these factors could have contributed to the difference in the conclusions between the two studies, as we will briefly discuss in the following.

5.1 CMIP3 and CMIP5 simulations

We compare the average precipitation responses to ALL, ANT, and NAT forcings over all land grid boxes north of 50°N as simulated by the CMIP3 models used in Min et al. (2008a) and by the CMIP5 models used here. In each case, we compute the time series of non-overlapping 5-year mean annual precipitation anomalies relative to the 1961–1990 climatology for individual runs, then average these individual runs to obtain individual model ensembles, and finally compute multi-model ensemble averages. Figure 8 displays the time series for 1951–2000, including the multi-model ensemble means and model simulated spread (5–95 % ranges). Overall, long-term trends and model spread from CMIP5 models are very similar to those from CMIP3 models. The large-scale patterns of trend for ALL and ANT between CMIP3 and CMIP5 simulations are generally

similar. Judging by these, similarities in the precipitation responses to external forcings, differences between CMIP3 and CMIP5 models does not appear to be an important contributor to the difference in the scaling factors.

5.2 Sampling uncertainty

Sampling uncertainty in covariance estimation can have profound impacts in the estimation of scaling factors and their confidence intervals as larger numbers of independent noise data samples are associated with smaller sampling errors (Ribes and Terray 2013). Zhang et al. (2007, Table S2) showed that the use of different combinations of noise data for optimization and testing may result in different scaling factors and their associated confidence intervals. Noise data that were used to estimate covariance matrices in Min et al. (2008a) were relatively limited, with about 200 samples in contrast to 1,306 samples available for this study.

To provide some understanding of the magnitude of sampling uncertainty in the scaling factors due to sampling errors in covariance matrix estimation, we conducted detection analyses using the ALL signal computed from the

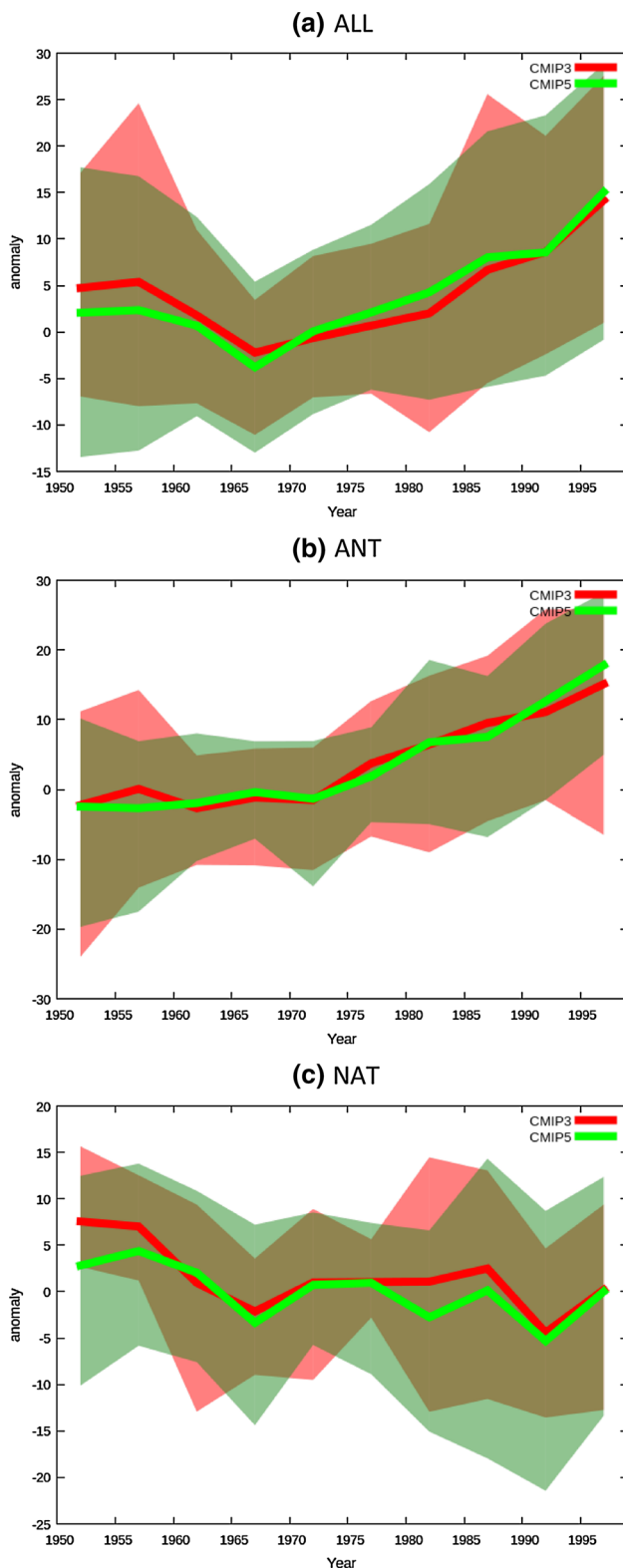


Fig. 8 Five-year mean northern high-latitude (north of 50°N) land precipitation anomalies (in mm, relative to 1961–1990) simulated by all CMIP5 models (*green line*) that are used in this study and simulated by CMIP3 models that were used in Min et al. (2008a) (*red line*). *Thick lines* represent multi-model ensemble means while *shading* shows the 5–95 % ranges from relevant ensembles

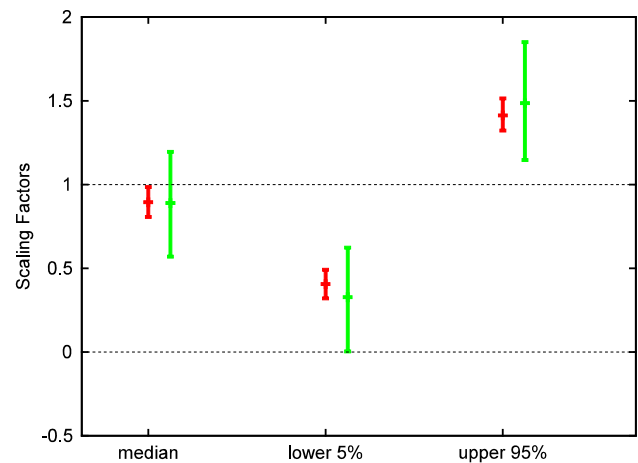


Fig. 9 Influence of sampling uncertainty in the covariance estimation on the scaling factors and their lower 5 % and upper 95 % confidence bounds for ALL using 3-region 5-year mean annual precipitation. The median scaling factor, median 5 % lower bound and median 95 % upper bound for 1,000 re-samplings of the noise data are shown, together with the 5–95 % range of the 1,000 realizations of each of these three statistics. Red bars are for the results based on 653 piece samples of noise data while *green bars* describe results based on 100 piece samples of noise data. See text for additional details

32 CMIP5 models with different noise samples. In one case, we use randomly chosen samples of 653 noise data segments for optimization and the remaining noise data for testing. In another case, we use randomly chosen samples of 100 noise data segments for optimization and additional randomly chosen samples of 100 noise data segments for testing. Both cases were repeated for 1,000 times. Results are summarized in Fig. 9. As expected, sampling uncertainty in both the scaling factor and confidence interval width are much larger when covariance matrices are estimated from small 100-chunk noise samples. However, the 90th percentile of scaling factors estimated with 100 samples of noise data is still much smaller than those obtained in Min et al. (2008a), which suggests that sampling errors in the covariance matrix estimation is also not a plausible explanation for the difference in scaling factors between the two studies. Figure 9 also shows that our annual detection result (based on the 3-region configuration) is very robust against sampling uncertainty in covariance estimation; the 5th percentile of the lower bound of the 90 % confidence intervals of the scaling factor is still above zero.

5.3 Data coverage and processing procedures

The time periods (1950–1999) used in Min et al. (2008a) and spatial coverage of data used in the two studies are both different. This would have an effect on the scaling factor estimate though the size of this effect is difficult to quantify. The main results shown in Min et al. (2008a) are based

on the projection of both observations and model simulated signals onto the first several leading space–time EOFs of the model simulated internal variability. This data processing procedure is different from what has been conducted here. Min et al. (2008a) also showed, in the supplementary information, a detection result based on precipitation averaged over the entire domain. They found in that analysis that the model simulated response was consistent with the observations. This finding is consistent with the results that have been reported here. It therefore appears that differences in data pre-processing procedures play an important role. Combined with the different time periods, they can produce a large difference in scaling factor, possibly explaining differences between Min et al. (2008a) and this study.

6 Conclusion and discussion

Using an optimal fingerprinting method, we have compared annual and semi-annual precipitation from observations and CMIP5 simulations for 1966–2005 over northern high-latitude land north of 50°N. We conclude that the multi-model simulated responses to the effects of anthropogenic forcing or anthropogenic and natural forcing combined are consistent with observed changes. These detection results are robust to the use of a correlation-based method and the optimal fingerprint method. We also find that the influence of anthropogenic forcing may be separately detected from that of natural forcings, albeit less robustly, and that observed changes may be attributable to anthropogenic forcing.

To avoid potential data inhomogeneity due to changes in observing practices in the former USSR and to have as much spatial coverage as possible, we have compiled high quality station data from different sources and restricted our analysis primarily to 1966–2005. This provides much better data coverage over Northern Hemisphere high latitude land areas when compared with our earlier study (Min et al. 2008a). The model simulations for the estimation of precipitation responses to external forcings and internal variability are completely different from those used in Min et al. (2008a). Details in data analysis procedures are also different from those used earlier. Yet, the main finding of Min et al. (2008a) that influence of combined effect of natural and anthropogenic forcings, or of anthropogenic forcing alone, can be detected in northern high-latitude land precipitation, is confirmed. In contrast with Min et al. (2008a), who suggested that models may have underestimated the observed precipitation response to external forcing, this study shows that the model simulated response is consistent with observations, if not too large. The difference is unlikely due to the use of different model simulations but is more plausibly due to differences in the

space–time coverage of observations and in the details of data processing procedures.

Acknowledgments We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the World Climate Research Programme's (WCRP's) Working Group on Coupled Modeling for their roles in making available WCRP CMIP5 multi-model data. We appreciate helpful comments from Nathan Gillett, Brigitte Mueller and two anonymous reviewers. FZ and XZ acknowledge funding from the Natural Sciences and Engineering Research Council of Canada's Climate Change and Atmospheric Research initiative via the Canadian Sea Ice and Snow Evolution (CanSISE) Network. S-KM was funded by the Korea Meteorological Administration Research and Development Program under Grant CATER 2012-3081.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Alexeev VA, Langen PL, Bates JR (2005) Polar amplification of surface warming on an aquaplanet in “ghost forcing” experiments without sea ice feedbacks. *Clim Dyn* 24(7–8):655–666
- Allen MR, Stott PA (2003) Estimating signal amplitudes in optimal fingerprinting. Part I: theory. *Clim Dyn* 21:477–491
- Allen MR, Tett SFB (1999) Checking for model consistency in optimal fingerprinting. *Clim Dyn* 15(6):419–434
- Bindoff, et al. (2013) Detection and attribution of climate change: from global to regional. In: Stocker TF, et al. (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge
- Cai M, Lu JH (2007) Dynamical greenhouse-plus feedback and polar warming amplification. Part II: meridional and vertical asymmetries of the global warming. *Clim Dyn* 29:375–391
- Gillett PN, Stone DA, Stott PA, Nozawa T, Karpechko AY, Hegerl GC, Wehner MF, Jones PD (2008) Attribution of polar warming to human influence. *Nat Geosci* 1:750–754
- Groisman PY, Rankova EY (2001) Precipitation trends over the Russian permafrost-free zone: removing the artifacts of pre-processing. *Int J Clim* 21(6):657–678
- Hegerl GC, von Storch H, Hasselmann K, Santer BD, Cubasch U, Jones PD (1996) Detecting greenhouse gas-induced climate change with an optimal fingerprint method. *J Clim* 9:2281–2306
- Klok EJ, Klein Tank AMG (2008) Updated and extended European dataset of daily climate observations. *Int J Clim*. doi:10.1002/joc.1779
- Marvel K, Bonfils C (2013) Identifying external influences on global precipitation. *PNAS*. doi:10.1073/pnas.1314382110
- Mekis É, Vincent LA (2011) An overview of the second generation adjusted daily precipitation dataset for trend analysis in Canada. *Atmos Ocean* 49(2):163–177
- Min S-K, Zhang X, Zwiers F (2008a) Human-induced Arctic moistening. *Science* 320(5875):518–520
- Min S-K, Zhang X, Zwiers F, Agnew T (2008b) Human influence on Arctic sea ice detectable from early 1990s onwards. *Geophys Res Lett* 35:L21701. doi:10.1029/2008GL035725
- Noake K, Polson D, Hegerl G, Zhang X (2012) Changes in seasonal land precipitation during the latter twentieth-century. *Geophys Res Lett*. doi:10.1029/2011GL050405

- Peterson TC, Vose RS (1997) An overview of the Global Historical Climatology Network temperature database. *Bull Amer Meteor Soc* 78B:2837–2849
- Peterson BJ, Holmes RM, McClelland JW, Vörösmarty CJ, Lambers RB, Shiklomanov AI, Shiklomanov IA, Rahmstorf S (2002) Increasing river discharge to the Arctic Ocean. *Science* 298:2171–2173
- Polson D, Hegerl GC, Zhang X, Osborn TJ (2013) Causes of robust seasonal land precipitation changes. *J Clim*. doi:[10.1175/JCLI-D-12-00474](https://doi.org/10.1175/JCLI-D-12-00474)
- Ribes A, Terray L (2013) Application of regularised optimal fingerprint to attribution. Part II: application to global near-surface temperature. *Clim Dyn*. doi:[10.1007/s00382-013-1736-6](https://doi.org/10.1007/s00382-013-1736-6)
- Santer BD, Mikolajewicz U, Brüggemann W, Cubasch U, Hasselmann K, Höck H, Maier-Reimer E, Wigley TML (1995) Ocean variability and its influence on the detectability of greenhouse warming signals. *J Geo Res* 100(C6):10693–10725. doi:[10.1029/95JC00683](https://doi.org/10.1029/95JC00683)
- Santer BD et al (2007) Identification of human-induced changes in atmospheric moisture content. *Proc Natl Acad Sci* 104:15248–15253
- Santer BD et al (2009) Incorporating model quality information in climate change detection and attribution studies. *Proc Natl Acad Sci* 106(35):14778–14783
- Scheff J, Frierson D (2012) Twenty-first-century multimodel subtropical precipitation declines are mostly midlatitude shifts. *J Clim* 25(12):4330–4347
- Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. *Bull Amer Meteor Soc* 93:485–498. doi:[10.1175/BAMS-D-11-00094.1](https://doi.org/10.1175/BAMS-D-11-00094.1)
- Terray L, Corre L, Cravatte S, Delcroix T, Reverdin G, Ribes A (2012) Near-surface salinity as nature's rain gauge to detect human influence on the tropical water cycle. *J Clim* 25:958–977
- Vaughan et al (2013) Observations: cryosphere. In: Stocker et al. (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge
- Willett KM, Gillett NP, Jones PD, Thorne PW (2007) Attribution of observed surface humidity changes to human influence. *Nature* 449(7163):710–712
- Zhang X, Zwiers FW, Hegerl GC, Lambert FH, Gillett NP, Solomon S, Stott PA, Nozawa T (2007) Detection of human influence on twentieth-century precipitation trends. *Nature* 448(7152):461–465