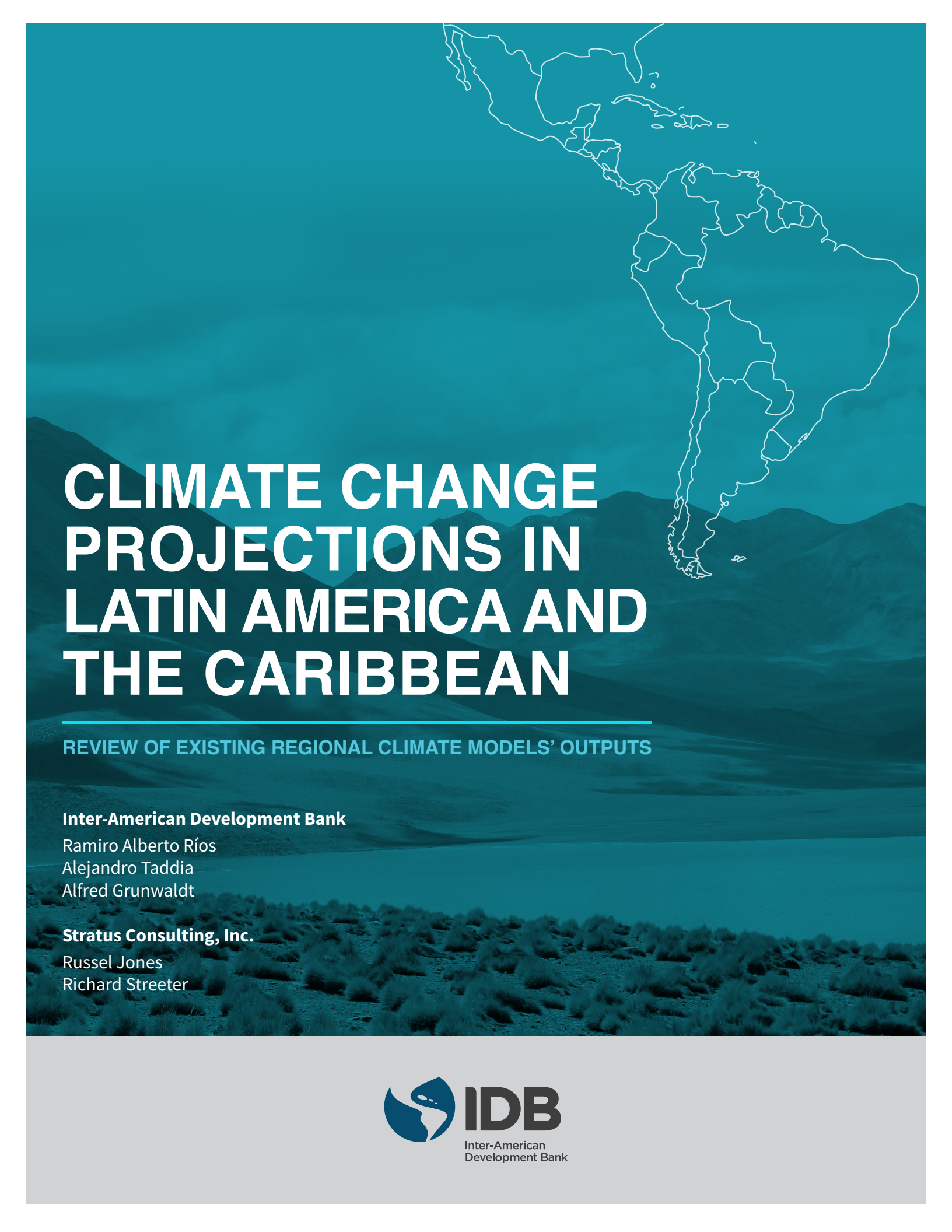




CLIMATE CHANGE PROJECTIONS IN LATIN AMERICA AND THE CARIBBEAN

REVIEW OF EXISTING REGIONAL CLIMATE MODELS' OUTPUTS

Inter-American Development Bank

Ramiro Alberto Ríos
Alejandro Taddia
Alfred Grunwaldt

Stratus Consulting, Inc.

Russel Jones
Richard Streeter

Cataloging-in-Publication data provided by the Inter-American Development Bank Felipe Herrera Library

Climate change projections in Latin America and the Caribbean: Review of Existing Regional Climate Models' Outputs / Ramiro Alberto Ríos, Alejandro Taddia, Alfred Grunwaldt, Russel Jones, Richard Streeter.

p. cm. — (IDB Monograph ; 444) Includes bibliographic references.

1. Climatic changes-Latin America-Forecasting. 2. Climatic changes-Caribbean Area-Forecasting.
3. Transportation-Environmental aspects-Latin America. 4. Transportation-Environmental aspects-Caribbean Area. I. Ríos, Ramiro Alberto. II. Taddia, Alejandro. III. Grunwaldt, Alfred. IV. Jones, Russel. V. Streeter, Richard. VI. Inter-American Development Bank. Transport Division. VII. Serie. IDB-MG-444

Publication Code: IDB-MG-444 JEL Codes: R42; Q54; O18; H54; N76 Keywords: Climate Change Adaptation; Latin America and the Caribbean; Infrastructure; Resilient Development.

Copyright © 2016 Inter-American Development Bank.

This work is licensed under a Creative Commons IGO 3.0 Attribution-NonCommercial-NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) and may be reproduced with attribution to the IDB and for any noncommercial purpose. No derivative work is allowed.

Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IDB's name for any purpose other than for attribution, and the use of IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this CC-IGO license.

Note that link provided above includes additional terms and conditions of the license.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.



Inter-American Development Bank
1300 New York Avenue, N.W.
Washington, D.C. 20577
www.iadb.org

1	Glossary of Key Terms	04
2	Executive Summary	05
3	Methods Climate Change Projections in Latin America and the Caribbean	07
	Baseline Data	07
	Emissions Scenarios	08
	Climate Sensitivity	08
	Global Climate Models (GCMs)	08
4	Average Annual Change Conditions Analysis	11
	Extreme Event Precipitation Analysis	11
	<i>Baseline Event Magnitudes</i>	11
	<i>Future Extreme Precipitation Event Projections</i>	12
	Extreme Event Maximum Temperature Analysis	13
	Extreme Event Sea Level Rise Analysis	13
5	Generation of Maps	15
6	Annex 1 Model Evaluation for Central and South America	21
	Overview	21
	Drivers of Latin American and the Caribbean Climate	22
	Summary of Key Studies	22
	Limitations of Model Evaluation Studies	23
	Summary of Rationale for Using all Available Models, Rather than a Selection	23
7	Annex 2 Comparison of Climate Station Data to Princeton Reanalysis Data in Mesoamerica	25
8	References	29

1

Glossary of Key Terms

Climate Sensitivity

Estimated amount that global average temperature is projected to increase assuming a doubling in carbon dioxide concentrations in the atmosphere.

CIMP5

Coupled Model Intercomparison Project Phase 5.

Generalized Extreme Value (GEV)

GEVs are used to model the smallest or largest value among a large set of independent, identically distributed random values representing measurements or observations, in this case extreme precipitation.

GCM

Global Climate Models - used to project changes in earth's climate.

Princeton Reanalysis

Dataset that provides near-surface meteorological data for land surface models and blends reanalysis data with observations and disaggregates in time and space.

RCP6

Representative concentration pathway 6.0 assumes a greenhouse gas concentration stabilization scenario in which total radiative forcing is stabilized shortly after 2100 by the application of a range of technologies and strategies for reducing greenhouse gas emissions.

RCP8.5

Representative Concentration Pathway 8.5 assumes a high energy-intensive scenario as a result of high population growth and a lower rate of technology development.

SimClim 2013

Software designed to facilitate the assessment of risks from climate change

SLR

Sea Level Rise

2

Executive Summary

Transport infrastructure throughout Latin America and the Caribbean (LAC) Region exhibits significant vulnerabilities to extreme weather events including coastal flooding, intense rains and landslides, and extreme temperatures. These events can disrupt and damage vital communication infrastructure that provide people's access to socio economic opportunities and that are basic for development. These impacts often affect most heavily on vulnerable populations, particularly in areas where the availability of alternative routes or other transport options is insufficient. The reality of a changing climate means that many extreme weather impacts are likely to grow in magnitude and severity, creating increasing burdens on people, governments and the private sector.

The Inter-American Development Bank (IDB) acknowledges the challenges the region faces and is developing a set of publications, guidelines and tools to increase knowledge of the vulnerability of transport infrastructure. The objective of this document is to disseminate the methods used by IDB to develop climate change projections based on key vulnerability assessment maps for 26 countries in LAC. Realistic Greenhouse gas emissions scenarios were used to develop different sets of climate projections (precipitation, temperature and sea level rise) and their potential impact on the transport sector.

This analysis used baseline data from Princeton Reanalysis database, which includes more than 40 Global Climate Models (GCMs) from modeling centers in Asia, Europe and North America. These GCMs simulate how the atmosphere, oceans and land surface interact globally to shift earth's climate over the years. GCMs were also used to project changes in earth's climate depending on greenhouse gas concentrations in the atmosphere and how these changes may present vulnerable conditions for transport infrastructure in the medium and long term.

An Average Annual Change Condition Analysis was developed using these climate scenarios specific to the LAC region. Climate projections for average annual maximum temperature (C°), precipitation (%) and sea level rise (cm) were developed from the baseline climate (1986–2005) for the years 2040 and 2070. Representative Concentration Pathways (RCP) 6.0 and 8.5 were used for each 0.5° grid box over land (not ocean) over the whole LAC region, resizing the grid boxes to 0.25° (approximately 25 x 25 km) for consistency with the baseline data.

The climate scenarios and modeling approaches described were used to generate a set of maps showing changes in maximum temperature (C°), rainfall (%) and Sea Level Rise (cm) to illustrate vulnerability in the transport sector. The selection of these maps was based on the following criteria: (i) type of use to be given to the information provided by each map (e.g. support to the decision making process in initial infrastructure design), (ii) lifetime of specific types of infrastructure, (iii) level of uncertainty of the information presented under the different scenarios that limited the real use of each map, (iv) type of infrastructure to be designed. Priority has been given in this study to culverts, bridges, pavement and coastal infrastructure.

Finally, the study includes a reasonableness check to understand how well the climate modeling data fitted actual climate station data in the region. Such exercise concluded that the climate modeling approach used provides adequate justification for using the Princeton Reanalysis dataset as the baseline climate proxy, particularly given the lack of other data currently available.

The average annual change analyses developed in this study illustrate the potential risks that climate change pose on infrastructure. These projections are not meant to be used as forecasts during the infrastructure planning phase given the complexity and uncertainty in climate projections.

3

Methods

Climate Change Projections in Latin America and the Caribbean

This report describes the methods Stratus Consulting used to provide future projections of climate change for 26 countries in Latin America and the Caribbean (LAC).¹ The projections will assist IADB with assessing, in particular, the vulnerability of the transportation sector.

For this work, projections of climate change were developed for the climate stressors shown in Table 1, including temperature, precipitation and sea level rise. The methods and the sources of data were used in this analysis are described below.

Table 1. Climate Stressors Addressed

Stressor	Average	Threshold/Value 1	Threshold/Value 2	Threshold/Value 3
Temperature (Princeton reanalysis baseline data)	Annual (for RCP6/8.5 in 2040/2070 using wet, dry, and median models)	Change in days > 29.5°C	Days > 35°C	Days > 41°C
Precipitation (Princeton reanalysis baseline data)	Annual (for RCP6/8.5 in 2040/2070 using wet, dry, and median models)	Change in return interval of today's 20-, 30-, 50-, 100-, and 300-year events	—	—
Sea Level Rise	2040 and 2070 by RCP6/8.5	Low (5th percentile regional scalar from 24 GCMs applied to global average)	Mean (50th percentile regional scalar from 24 GCMs applied to global average)	High (95th percentile regional scalar from 24 GCMs applied to global average)

Baseline Data

In Latin America and the Caribbean, few climate stations have historical climate data records that would be sufficiently long and continuous for use in this type of analysis. Thus, these data cannot serve as the sole baseline data source. Instead reanalysis data was used to provide a more complete set of baseline data. Reanalysis data are datasets typically built by calibrating a regional climate model to observed historical conditions. These products generate gridded “reanalyzed” historical data to attempt to

create a geographically and temporally complete dataset of baseline conditions. Reanalysis datasets are synthetic, but they have similar statistics, such as means and variance, as observational datasets. For this analysis, a daily historical reanalysis dataset from Princeton University was used for the purpose of evaluating climate risks in Latin America and the Caribbean. This dataset contained continuous gridded daily climate data from 1948 to 2008 at a 0.25° horizontal resolution.

¹ The region included Argentina, the Bahamas, Barbados, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, the Dominican Republic, Ecuador, El Salvador, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Suriname, Trinidad and Tobago, Uruguay, and Venezuela.

Emissions Scenarios

To determine climate change projections, two greenhouse gas concentration trajectories, or representative concentration pathways (RCPs) were selected from the Intergovernmental Panel on Climate Change's (IPCC's) Fifth Assessment Report (AR5). The RCPs measure radiative forcing by the year 2100 in watts per square meter (Moss et al., 2010; IPCC, 2014). There is no consensus on which of the four RCPs (RCP2.6, RCP4.5, RCP6, or RCP8.5) is most likely; the IPCC considers all RCPs to be within the likely range of actual radiative forcing. Although RCP4.5 is widely used, we believe that extraordinary global policy measures would be needed to control emissions to the degree necessary to achieve this lower level of greenhouse gas concentrations (measures which hopefully are implemented). However, for the purpose of adaptation planning, it would be prudent to plan for more realistic scenarios. For this analysis, RCP6 and RCP8.5 were selected simply because they are most consistent with recent global trends in greenhouse gas emissions.

Climate Sensitivity

Climate sensitivity is the estimated amount that global average temperature is projected to increase, assuming a doubling of carbon dioxide concentrations in the atmosphere. IPCC's Fourth Assessment Report (AR4) revised the likely range to 2–4.5°C, with 3°C considered as the most likely value (IPCC, 2007). IPCC's most recent report, AR5 (IPCC, 2014), did not give a most likely climate sensitivity, but stated that the value is likely 1.5–4.5°C and very likely between 1°C and 6°C. It is deferred to the AR4 to use 3°C as the most likely climate sensitivity value.

Global Climate Models (GCMs)

Global Climate Models (GCMs) are used to project changes in the Earth's climate. GCMs simulate how the atmosphere, oceans, land surface, and ice over the entire Earth interact globally to shift climate over long periods of time (decades and centuries). GCMs are typically run with various changes to forcing conditions, such as increased greenhouse gas concentrations in the atmosphere.

The prevailing generation of models associated with AR5 from the Coupled Model Intercomparison Project Phase 5 (CMIP5), includes more than 40 models from more than 20 modeling centers in North America, Europe, and Asia. For this project, it was relied on expert judgment from Columbia University researchers to conduct a literature review to evaluate whether all the models available should be used- or whether the list should be culled to remove any that perform poorly over the LAC region. The GCMs were evaluated based on their ability to simulate large-scale climate fields, such as global average surface temperature; or key climate processes, such as dynamics associated with rising air and cloud formation in the tropics and sinking air in the subtropics (Flato et al., 2013). The general approach used by many researchers is to compare observed data with climate model outputs from "hindcast" simulations that reproduce the approximate climate forcing present in the atmosphere during the period of comparison. The review concluded that there is a case to exclude four of the available models. The full report containing this review appears in Appendix 1.

To reduce the number of projections to a more manageable number, while still capturing the range of future projections, the projected annual change in precipitation for year 2070 was plotted, the most distant year projections were provided for, under IPCC's highest-emissions scenario (RCP8.5), using output from SimCLIM 2013 software at a 0.5° resolution² spatially averaged over the 26 LAC countries. Once this information was plotted, groups of GCMs were selected to capture a wide range of projections

from the models. The models were selected based on their proximity to a wide range of percentiles (end-members) of change as follows: dry (5th% Precipitation), wet (95th% Precipitation), and median (50th% Precipitation). Additionally, if a model selected on the basis of percentile did not also have projections of extreme precipitation, that model was excluded³ (Figure 1). Five GCMs per end-member grouping were averaged because averaging over several models reduces the “noise” (e.g., high month-to-

Projected Change in Precipitation from Baseline

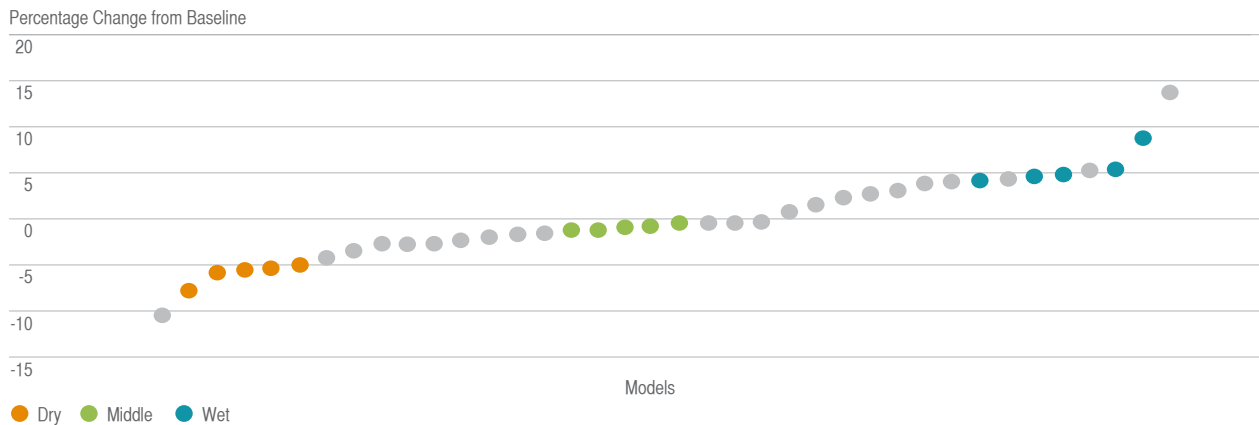


Figure 1: Distribution of average annual change in precipitation over Latin America by 2100 under the RCP 8.5 scenario for 38 GCMs. Models grouped by 5 closest models to 5th, 50th and 95th percentile representing “dry” (orange), “middle” (green), and “wet” (blue) models, respectively. Note that the wet models selected are not sequential due to the requirement that the selected models needed to have extreme estimates as well as average annual estimates.

² SimCLIM (Warrick, 2009) is a licensed software package that uses MAGICC (Wigley, 2008) and pattern scaling to estimate change in regional climate.

³ Extreme precipitation projections were available for 22 GCMs.

month variance) that can come from using a single model (Figure 2). The three groups of models (dry, wet, and median) were used for average change climate projections as well as change in extreme conditions. It is referred to the *dry, wet, and median* groups as *climate scenarios*.

It is important to stress that there is a degree of uncertainty in both the baseline climate data as well as the future climate projections from the GCMs. Due to the sparsity and lack of continuity of observed climate station data, baseline reanalysis data were used as a reasonable proxy for historical climate. It is also noted that the climate projections produced are not forecasts of future climate

as there is a high degree of uncertainty inherent in any climate projection. There is confidence about the direction of climate change for climate variables such as temperature, but have much less confidence about the magnitude of change. For other variables such as precipitation, even the direction of change can be uncertain in many areas. Therefore, scenarios have been developed about projected climate in the future and provided the range of projections that bound the outputs from the GCMs used, although the actual climate could even fall outside of this range. In summary, the projections provided should not be used as a climate forecast, but rather to provide examples (scenarios) of possible climates in the future.

Projected Change in Precipitation from Baseline

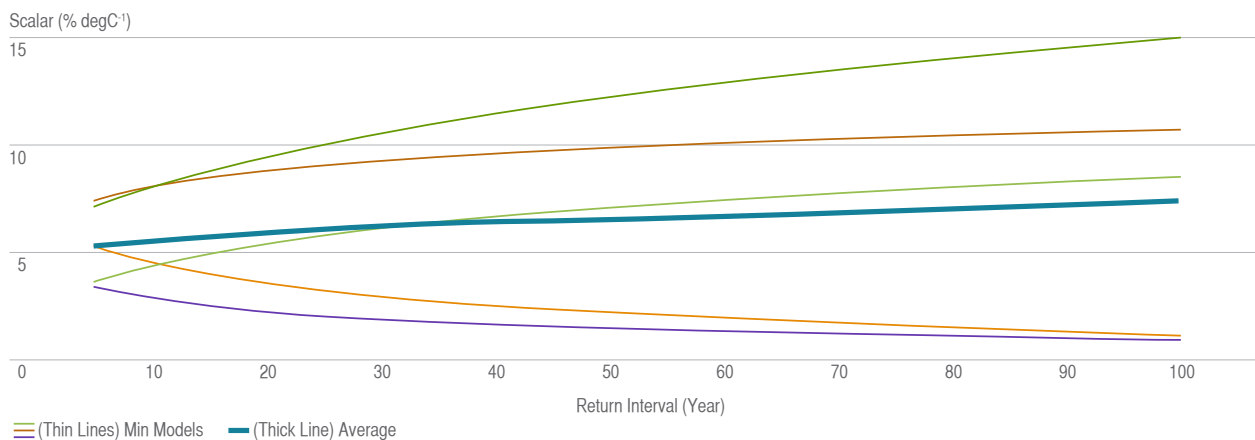


Figure 2: Cell-level output for extreme precipitation changes by return interval for five GCMs (thin colored lines) and models average (thick blue line). Note that this figure was not generated for a cell within Latin America, but exemplifies the reduction in “noise” by using a model average rather than individual GCMs.

4

Average Annual Change Conditions Analysis

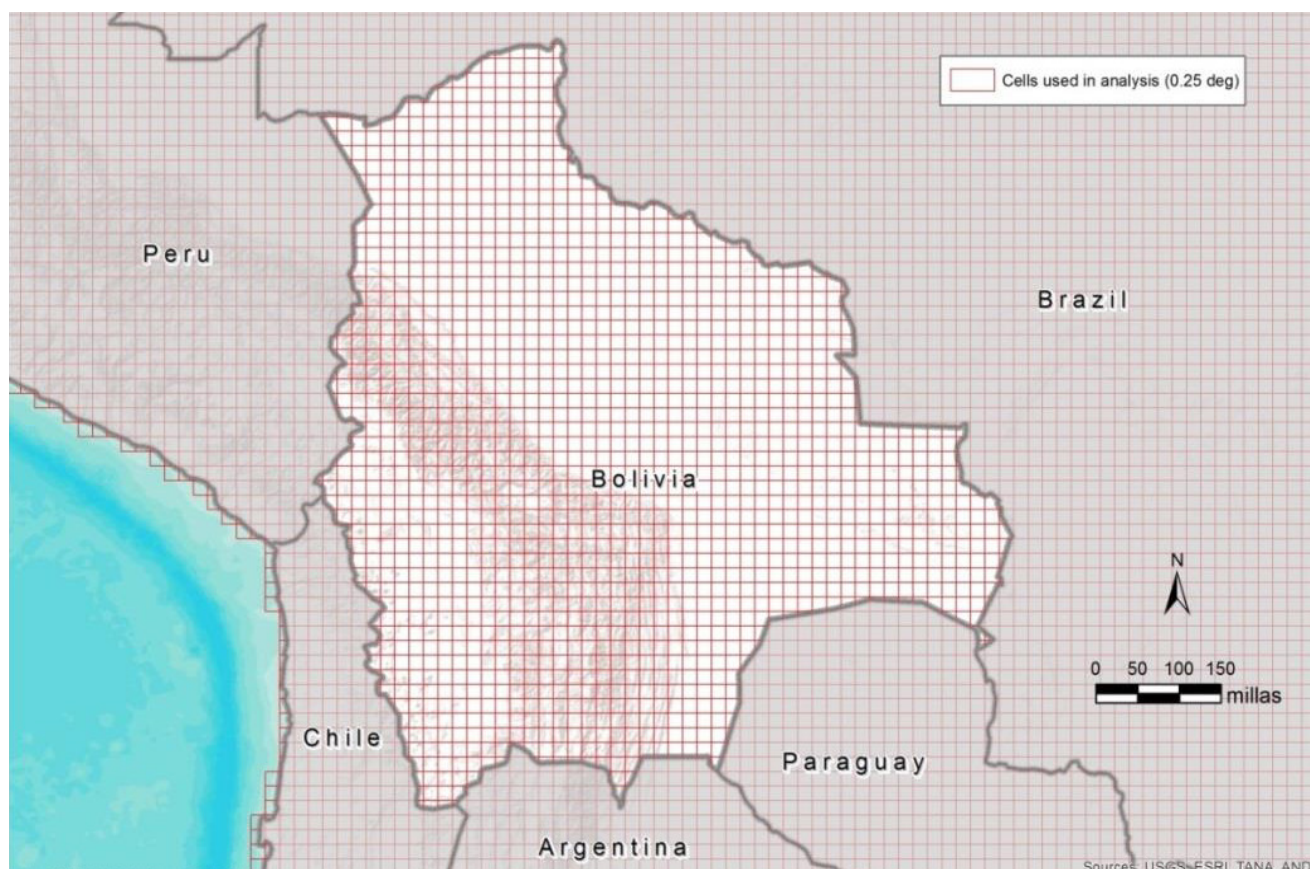


Figure 3. Study area showing the outline of 0.5° x 0.5° cells used in the analysis

Using these climate scenarios, climate projections of change to average annual maximum temperature (C°) and precipitation (%) were developed from the baseline climate (1986–2005) for the years 2040 and 2070. RCP6.0 and RCP8.5 were used for each 0.5° grid box over land (not ocean) in each country, resizing the grid boxes to 0.25° for consistency with the baseline data. Future time periods represent average conditions over a 20-year window, centered on the respective target year: 2040 represents conditions from 2031 to 2050 and 2070 represents conditions from 2061 to 2080. Figure 3 shows the 0.25° cell outlines covering the country of Bolivia.

Extreme Event Precipitation Analysis

Baseline Event Magnitudes

In addition to average conditions, projected changes to extreme precipitation for all grid boxes containing land in Latin America were generated. This was done by country for each climate scenario, using baseline (1948–2008) gridded daily climate data from Princeton University (Sheffield et al., 2006) at a 0.25° resolution. After this a generalized extreme value (GEV) curve was fitted to the observed annual maximum time series for each 0.25° cell. Using the GEV “fits,” the 24-hour maximum precipitation expected for six return intervals was estimated: 10, 20, 30,



50, 100, and 300 years. These estimates represented the magnitude of events that would be expected to occur in any given year with probabilities of 10%, 5%, 3%, 2%, 1%, and 0.3% (i.e., 1:10, 1:20, 1:30, 1:50, 1:100, 1:300), respectively.

Future Extreme Precipitation Event Projections

In the analysis, it was projected how the baseline return intervals would change under each of the climate change scenarios examined. It was examined how the baseline return interval would change (e.g., a 100-year event could become an 80-year event), as well as the change in magnitude of events at a specific return interval (e.g., the amount of precipitation that would occur in a 100-year event in the future).

Future GEV curves were derived using spatially explicit scalars indicating the changes in magnitude of extreme precipitation events over each 0.25° cell. This was done using a technique called pattern scaling, which assumes that changes in regional climate for individual models follow a pattern (e.g., a region may get drier or wetter, with the severity in linear proportion to the change in global mean temperature). CLIMSystems, Ltd. calculated these pattern scalars that are part of the SimCLIM software package (Warrick, 2009).⁴ The scalars differ from GCM to GCM, from grid cell to grid cell, and by the specific return period of interest. Resulting scalars for extreme precipitation events were then applied to the baseline peak magnitudes of annual 24-hour rainfall events (i.e., the maximum-rainfall events within a year) with the same return intervals. The choice to use two RCPs and three GCMs – wet, dry, and median – provided a range of changes in extreme precipitation.

⁴ From CLIMPACTS to SimCLIM: Development of an integrated assessment model system. In *Integrated Regional Assessment of Global Climate Change*, C.G. Knight and J. Jäger (eds.). Cambridge University Press, Cambridge, UK. pp. 280–311.

Extreme Event Maximum Temperature Analysis

The average annual number of days above each of the temperature threshold values was derived, as shown in Table 1; Princeton University's daily reanalysis data (0.25° resolution) was used, covering the 1948–2008 time period.

Climate projections of change in average monthly maximum temperature (°C) from the baseline data (1986–2005) were generated for each emissions scenario, time period, and model group as gridded output from SimCLIM. Future projections of monthly change in maximum monthly temperature to each day in the reanalysis data were then applied by the respective month. For example, the monthly change in temperature projected for January 2070 under the RCP8.5 emissions scenario for the wet model group to each January day in the historical Princeton University data was applied (January 1–31 from 1948 to 2008). After this was completed, the average annual number of days above each of the temperature thresholds was calculated in the same manner as for the baseline data.

Extreme Event Sea Level Rise Analysis

We estimated the amount of sea level rise (SLR) for all 0.25° cells along the coast of Latin America and the Caribbean. Our analysis considered the eustatic (global average) amount of SLR as well as regional variation caused by processes such as ocean currents or salinity that may cause sea level rise to be greater or less than the global average. These regional variations

are expressed as GCM-specific scalars (each scalar is regional change relative to the eustatic amount of SLR for a given time period and climate scenario) as output from the SimCLIM climate software. As indicated in Table 1, we bounded our estimates using the 5th, median, and 95th percentile scalars from the 24 GCMs with SLR estimates. To estimate the total amount of SLR at a given cell, we multiplied the appropriate scalar by the eustatic SLR projected for each time period (2040, 2070) and emission scenario (RCP 6.0 and 8.5). We did not apply the SLR projections to baseline elevation data because the vertical resolution is, in most cases, less than future SLR. Sea level rise estimates are shown in Map 4 for the median scenario for RCP 6.0 by 2040.



5

Generation of Maps

Methodologies and conceptual approaches previously described in sections one and two were instrumental in generating a series of maps with specific climate variables (see maps 1 to 7), recurrence periods for each Representative Concentration Pathways (RCPs) and different time-scales (e.g. 2040 and 2070) and recurrence times for the case of 24-hr intense precipitation. Generated maps were then revised and a group was then selected, including a regional and a national scale, taking as an example Bolivia. The selection of these maps was based on the following criteria: (i) type of use to be given to the information provided by each map (e.g. support to the decision making process in initial infrastructure design), (ii) lifetime of specific types of infrastructure, (iii) level of uncertainty of the information presented under the different scenarios that limited the real use of each map, (iv) type of infrastructure to be designed (for this specific study priority has been given to bridges, roads and drainage elements).

With respect to the last point in the past paragraph, the type of climate information prioritized for each specific infrastructure has been differentiated. For example, for the case of bridges, precipitation intensity with emphasis on a 100 return period has been selected under a 2070 time horizon under the RCP 6.0 scenario. The reason for the selection of the 2070 time horizon and the 100 return period is the lifetime of this type of infrastructure; for the case of roads drainage elements, the variable of interest has been set as the precipitation intensity as well, but with a shorter return period (20 years in this case) and a different time horizon (2040) to take into consideration its service lifetime.

Generated maps are also classified broadly in two categories, regional and country-scale maps; for the later the country of Bolivia has been selected as a pilot. For this case, climate variables under the RCP 6.0 scenario were selected as follows: (i) 24-hours intense precipitation with different recurrence times and under different time horizons (e.g. 2040 and 2070) and (ii) number of days with temperature over 29.5 C.

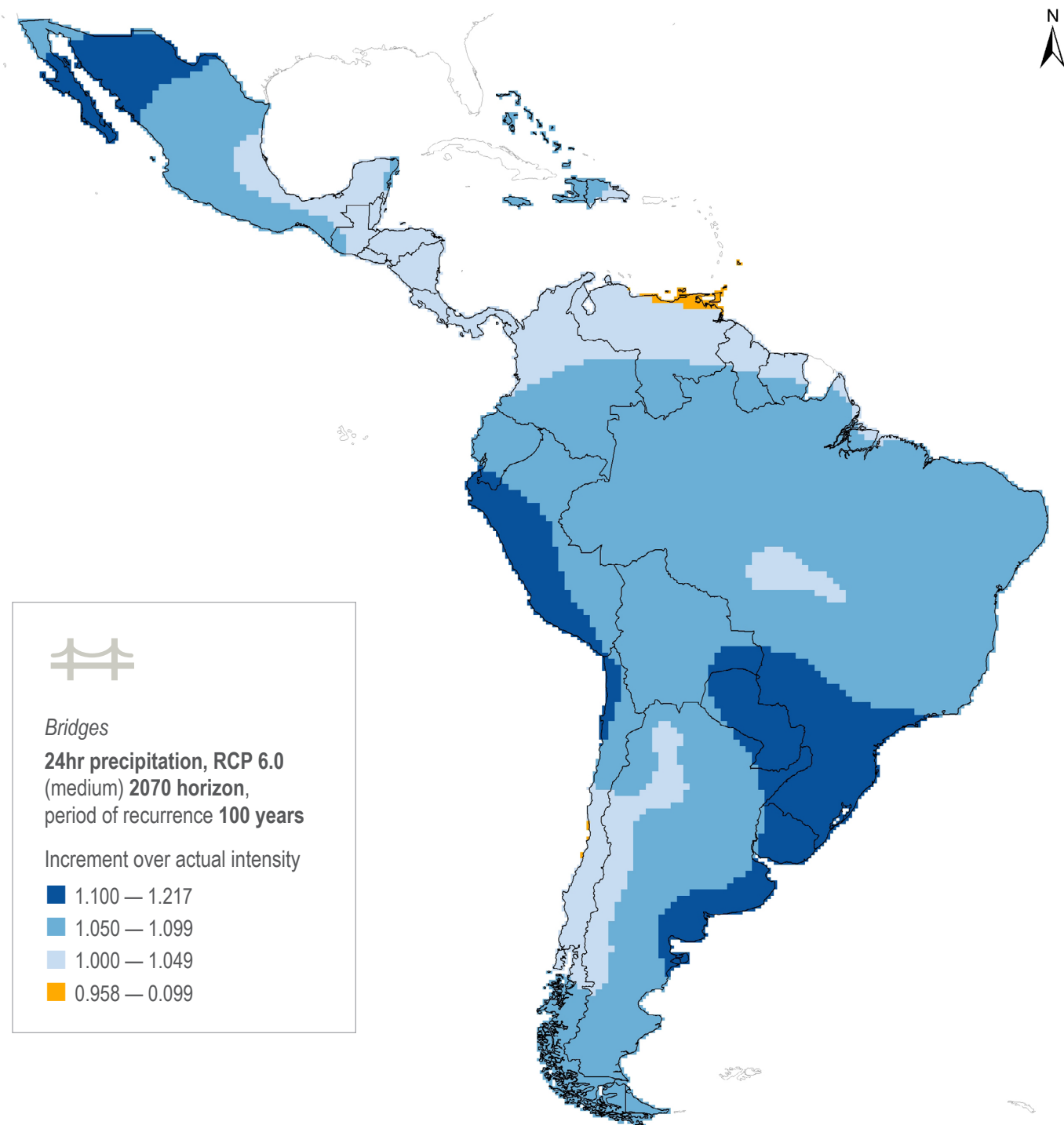
Map 1.

Climate change scenario proposed for roads drainage elements
24-hr precipitation using RCP 6.0;



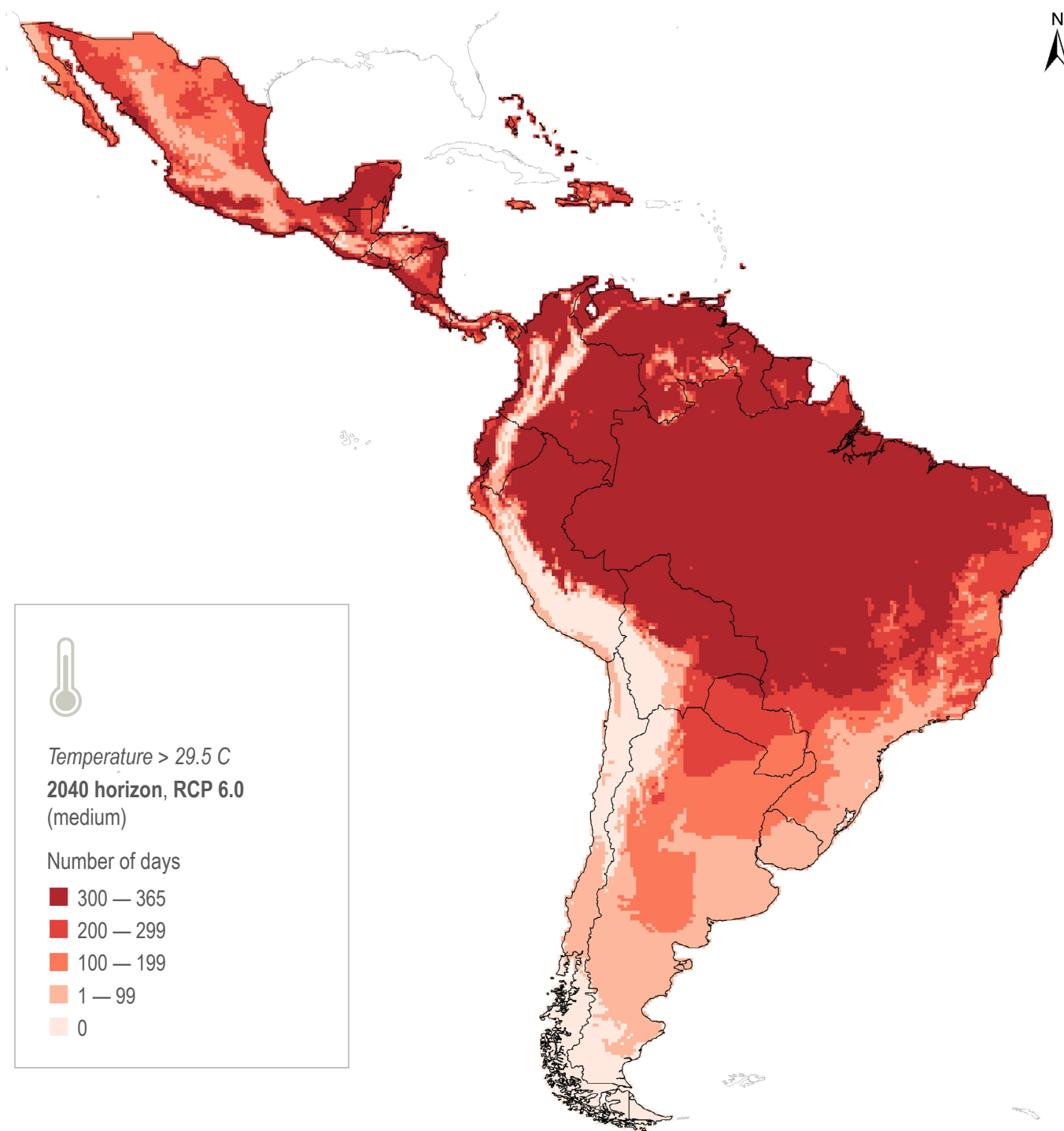
Map 2.

Climate change scenario proposed for bridges
24-hr precipitation using RCP 6.0;



Map 3.

Number of days with temperature higher than 29.5 C



Map 4.

Sea level rise increment based on the RCP 6.0



6

ANNEX 1 Model Evaluation for Central and South America

By Dr. Radley Horton

Dr. Radley Horton provided the information in this appendix. Dr. Horton used previously published studies to evaluate and select among a wide suite of CMIP5 models available, which Stratus Consulting used to simulate future climate in Latin America and the Caribbean for IADB's transportation vulnerability project.

Overview

Several studies have assessed the performance of global climate models in Central and South America, both in data from IPCC's AR4⁵ and AR5 reports. In general, the studies find some improvements in performance in the CMIP5 models, which are known for their higher spatial resolution, more complex coupling between model components, and updated parameterizations based on improved physical understanding (Flato et al., 2013).

Models are generally evaluated based on their ability to simulate large-scale climate fields, such as global average surface temperature, or key climate processes such as dynamics associated with rising air and cloud formation in the tropics and sinking air in the subtropics (Flato et al., 2013). The general approach in evaluation studies is to compare observed data with climate model outputs from "hindcast" simulations that reproduce the approximate

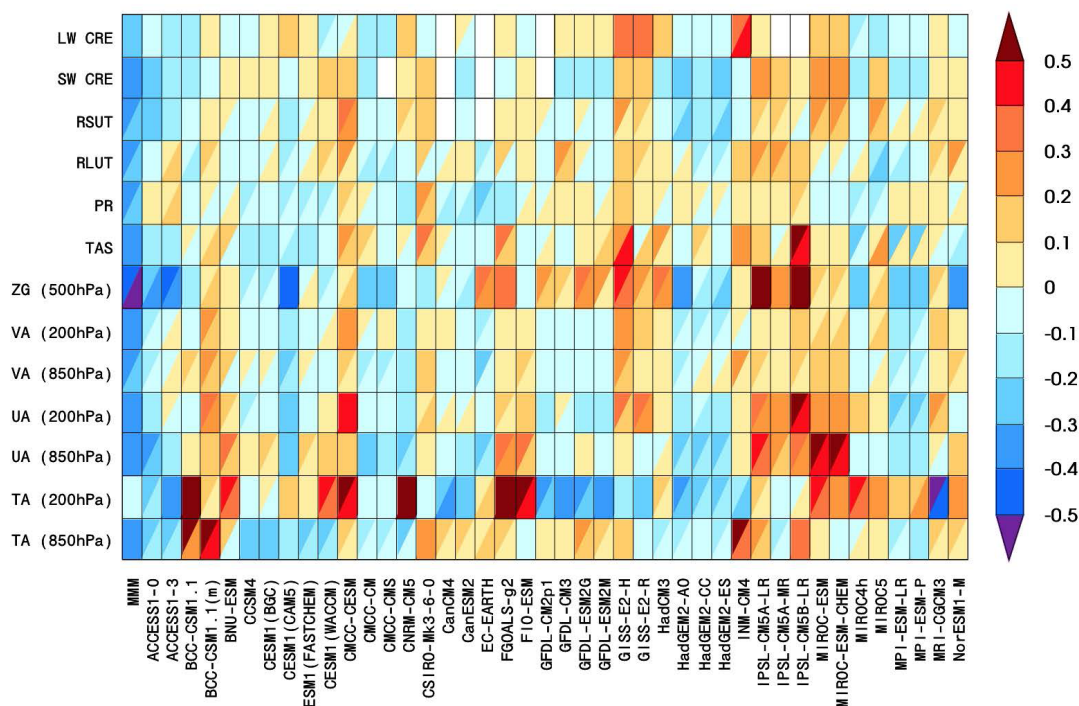


Figure A1.1

Model evaluation metrics, taken from Flato et al., 2013. Red = below average skill; Blue = above average skill.

Relative error measures of CMIP5 model performance, based on the global seasonal-cycle climatology (1980–2005) computed from the historical experiments. Rows and columns represent individual variables and models, respectively. The error measure is a space-time root-mean-square error (RMSE), which, treating each variable separately, is portrayed as a relative error by normalizing the result by the median error of a model results (Gleckler et al., 2008). For example, a value of 0.20 indicates that a model's RMSE is 20% larger than the median CMIP5 error for that variable, whereas a value of -0.20 means the error is 20% smaller than the median error. No color (white) indicates that model results are currently unavailable. A diagonal split of a grid square shows the relative error with respects to both the default reference data set (upper left triangle) and the alternate (lower right triangle). The relative errors are calculated independently for the default and alternate data sets.

⁵ Data associated with AR4 and AR5 are from the Coupled Model Intercomparison Project Phase 3 (CMIP3) and Phase 5 (CMIP5), respectively.

climate forcing present in the atmosphere during the period of comparison. Figure A1.1 shows the IPCC AR5 results from this type of global approach; the data reveal that although some models tend to perform better than others, it is highly variable-dependent, even for global-scale quantities.

Drivers of Latin American and the Caribbean Climate

Key drivers of Central and South American climate include (1) the annual monsoon cycle, which generally follows (with a delay) the latitude of maximum incoming solar radiation in the tropics; and (2) the effects of land surface features and topography, including the Amazon forest, the Andes Mountains, and the influence of the Atlantic and Pacific Oceans. Climate variability in the region is further influenced by major ocean-atmospheric modes, such as the El Niño Southern Oscillation (ENSO), and the Atlantic Multi-decadal Oscillation (AMO).

Summary of Key Studies

Unsurprisingly, with such a disparate set of processes, locations, and variables, few models are able to stand out consistently, in either a positive or negative way, in such a variety of contexts. However, we summarize a few key studies here.

As reported by Flato et al. (2013) in the IPCC AR5, the CMIP5 models generally show improvement over the CMIP3 models in the simulation of the monsoons and the critical west-to-east mean temperature gradient in the tropical Pacific. In contrast, many of the latest models still do not reproduce the east-west temperature gradient in the tropical Atlantic. In terms of variability, the CMIP5 models feature improved characterization of the ENSO, despite poor simulation of both ENSO teleconnections outside the tropical Pacific and Atlantic Ocean variability more generally. In summary, improvements are present in some areas but not others, and overall biases for some processes remain large in many individual models.

Several of the studies reviewed here have focused their evaluation on model performance in the Central and South American region specifically. Most of the region-specific evaluation studies focus on basic variables that are of most direct interest to society – temperature and precipitation. However, a few studies consider additional variables that especially influence precipitation, such as upper- and low-level winds and specific humidity (e.g. Carvalho and Jones, 2013; Jones and Carvalho, 2013). One challenge is that, given the large size of the region, studies have varied in the metrics they assess to evaluate model performance in Central and South America. Common process-oriented foci include the South American Monsoon System (SAMS), the Central American rainy season (e.g., Hidalgo and Alfaro, 2014), the El Niño Southern Oscillation (Flato et al., 2013; Steinhoff et al., 2015), and the Atlantic Meridional Overturning Circulation (AMOC; Cheng et al., 2013). For example, regional studies have assessed the Caribbean (e.g., Ryu and Hayhoe, 2014) and Amazon (e.g., Yin et al., 2013), as well as the Andes, northeast Brazil, and southern South America (Marengo et al., 2014).

In a sub-region-by-sub-region model evaluation for all of Central and South America, Marengo et al. (2014) report that the CMIP5 models largely eliminate the bias toward dry conditions that CMIP3 displays over northern South America and the Amazon. However the authors report no improvements over the Andes or western South America for the CMIP3 models. A different team of authors, Hidalgo and Alfaro (2014) conducted a preliminary evaluation of 48 CMIP5 models over Central America and identified and evaluated 13 models as top performers. The models showed skill for simulating mean temperature; for mean precipitation patterns, however, they only performed well for the March–May season. Only a few models reproduced the standard deviation of observed temperature and precipitation with high accuracy.

A few studies highlighted results from specific models. Gulizia and Camilloni (2015) found that the high-resolution MIROC4 model had superior representation of

South American precipitation, speculating that a relatively high-spatial resolution might work in the model's favor. In an evaluation of precipitation and associated dynamics focused on the tropical part of South America, Yin et al. (2013) identified the superior dynamical simulation of HadGEM2-ES, singling out GFDL-ESM2M as having a high precipitation bias. Jones and Carvalho (2013) identified individual CMIP5 models that performed poorly in specifically simulating South American monsoon precipitation in various sub-regions. However, across a range of metrics (e.g., simulation of the mean, variability, and timing of the end of the rainy season), they highlighted only CSIRO as consistently performing poorly.

Limitations of Model Evaluation Studies

Model evaluation studies are important: climate models are the primary tools used to develop future projections. However, there are major limitations to model evaluation, especially based on comparison to observed climate at the regional scale. First, across much of the region, there are large observational uncertainties because of limited high-quality, long-duration station data, especially in large remote areas such as the Amazon. Second, natural variability, which cannot be reproduced in the observed sequence by coupled climate models, can play a dominant role from yearly to multi-decadal scales. Third, some important drivers of climate in the region – such as land-use changes and aerosol formation associated with deforestation and fire – are not adequately and consistently included in climate model experiments (e.g., Zhang et al., 2009). These missing drivers increase the odds that any best-performing models for a given metric or region may be performing well for the wrong reasons. Finally, non-stationarity is an issue: even a model that performs well in the baseline climate might not perform as well in a future climate with increasing greenhouse gas concentrations.

More technical, study-specific limitations are also apparent. With only a few exceptions (e.g., Flato et al.,

2013; Hidalgo and Alfaro, 2014), the studies described here only assessed a small subset of the universe of available GCMs, thus limiting the scope of the studies' conclusions. Another limitation of the model evaluations for the region is that they have focused on yearly and monthly statistics, not on the daily outputs used for projections in this work.

Summary of Rationale for Using All Available Models, Rather than a Selection

Overall, a case can be made for erring toward an inclusive approach to model selection because of (1) the large size of the region, which, based on the literature seems to preclude identification of models that are consistently “good or “bad” across the region; (2) the temporal scale mismatch between the evaluations in the literature (yearly and monthly) and this study (daily); (3) inconsistencies in the evaluations, such as only evaluating a small number of models within a suite of models in some cases; (4) the inherent limitations of hindcasting, even in a perfectly designed evaluation as a predictor of future performance; and (5) the risk-management argument to include as broad a range of results as possible, especially given that models may underestimate the full range of possible outcomes.

Given the above arguments, we limited our exclusions to just four models. Specifically, we eschewed GISS-E2-H, IPSL-CM5A-LR, and IPSL-CM5B-LR due to poor overall performance across a broad range of global metrics shown in Figure A.1.1. The poor performance is noted especially for surface temperature and precipitation, although it is hardly confined to these two metrics. We also eschewed CSIRO-Mk3-6-0 due to poor performance for surface temperature and precipitation globally (see Figure A1.1), with an added argument provided by the fact that this model was singled out for poor performance across a variety of metrics (e.g., mean, variability, and timing) in the Jones and Carvalho (2013) evaluation of model simulation of South American monsoon precipitation.

7

ANNEX 2

Comparison of Climate Station Data to Princeton Reanalysis Data in Mesoamerica

This annex describes the process followed for a screening-level evaluation of the Princeton daily climate dataset as proposed by Stratus Consultants. The Climate change analyses related to the Inter-American Development Bank (IADB) Meso-America project was used. The evaluation was focused on assessing the degree of correspondence between the gridded Princeton dataset and the sparse station data that are available for this region. Figure A2.1 shows the climate stations used in the evaluation. While a thorough statistical analysis of the degree to which the gridded data and the station data agree was not conducted, it is believed that the performed evaluation provides adequate justification for using the Princeton dataset as the baseline climate proxy, particularly given the lack of other data currently available.

Because of the sparsity of observed climate station data that are available, the evaluation was focused on the maximum daily temperature time series. The station data was used as the base dataset, and the corresponding Princeton data were extracted for each date on which a station data point is also available.

As Figures A2.2–A2.4 show, the Princeton data do not exactly replicate the observed data; however, there are a number of reasons to expect an imperfect Tmax match between these two datasets. In particular, the station data represent climate conditions at a single point, which in most cases will not match the average conditions within a $0.5^\circ \times 0.5^\circ$ grid cell because of variations in elevation, proximity to water bodies, shading, etc. Therefore,

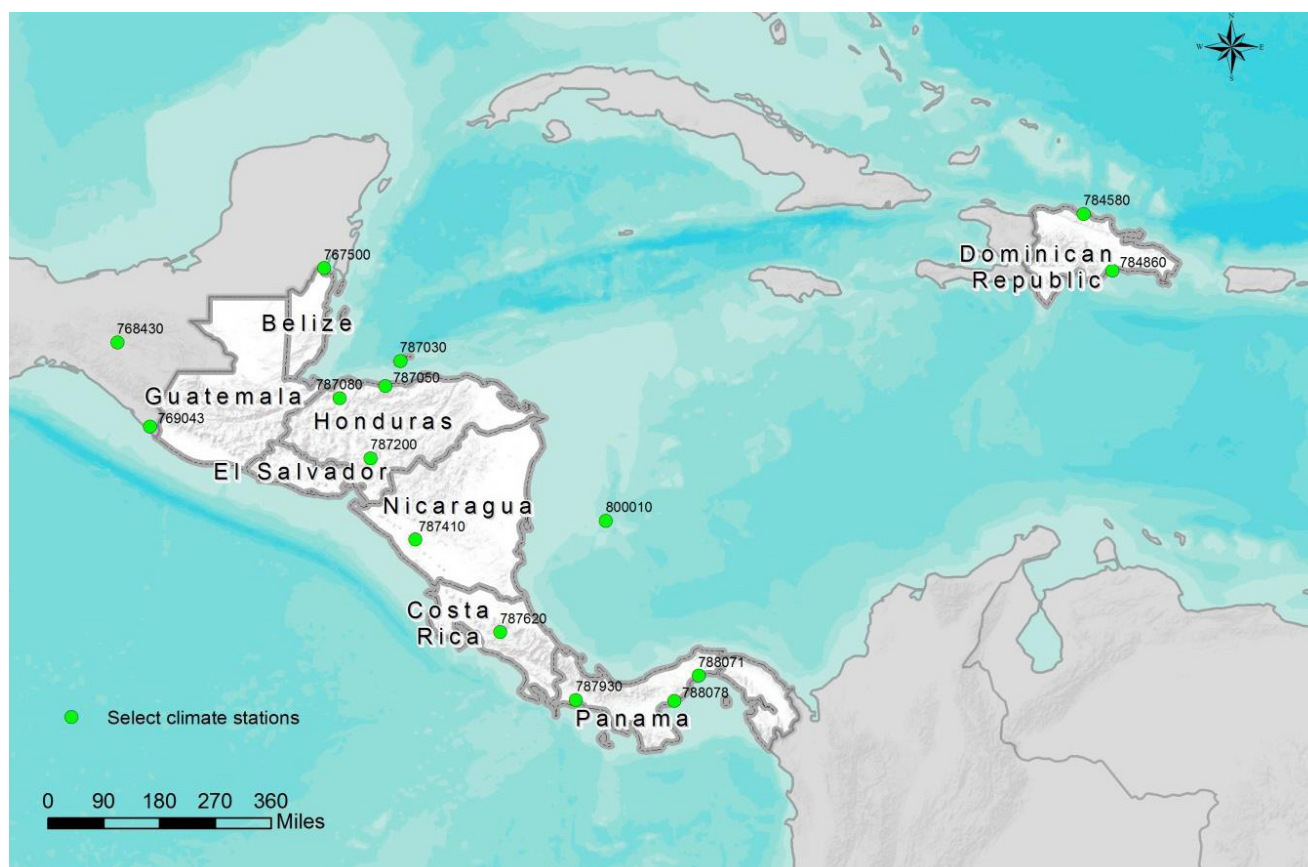


Figure A2.1. Climate stations with available information (green dots shown)

the analysis focused on evaluating whether there are systematic biases in the Princeton dataset that would preclude their use for climate change projections. A systematic bias would indicate a potential problem with using the Princeton dataset as a proxy for Tmax. If no systematic bias exists, i.e., if the differences appear to be random, then it would be reasonable to conclude that it is acceptable to use the Princeton data.

The screening-level evaluation of the Princeton data and the station data indicates there is no systematic bias introduced by estimating Tmax from one dataset versus another. In the set of data examined, errors appear to be random, with some individual stations showing a positive bias and others showing a negative bias. Therefore, it could be concluded that using the Princeton data does not introduce a systematic bias across the study area. Attached are a few examples of the results from the evaluation

carried out. On each of the plots there are three subplots:

- A scatterplot of station Tmax versus Princeton Tmax for all dates where overlap exists between the two, with a 1:1 line for reference
- A time series going from 1980 to 2008 showing Princeton Tmax data (blue) and station Tmax data (red) through time
- A zoomed-in view of the time series from 2003 to 2008 showing Princeton Tmax data (blue) and station Tmax data (red) through time.

Note that plots 2 and 3 were made with the Princeton data shown as a line and the station data as dots. This is not to imply that the Princeton data are continuous, but just to make it easier to distinguish between the two datasets. Figures A2.2 – A2.5 show a few aspects, as described below.

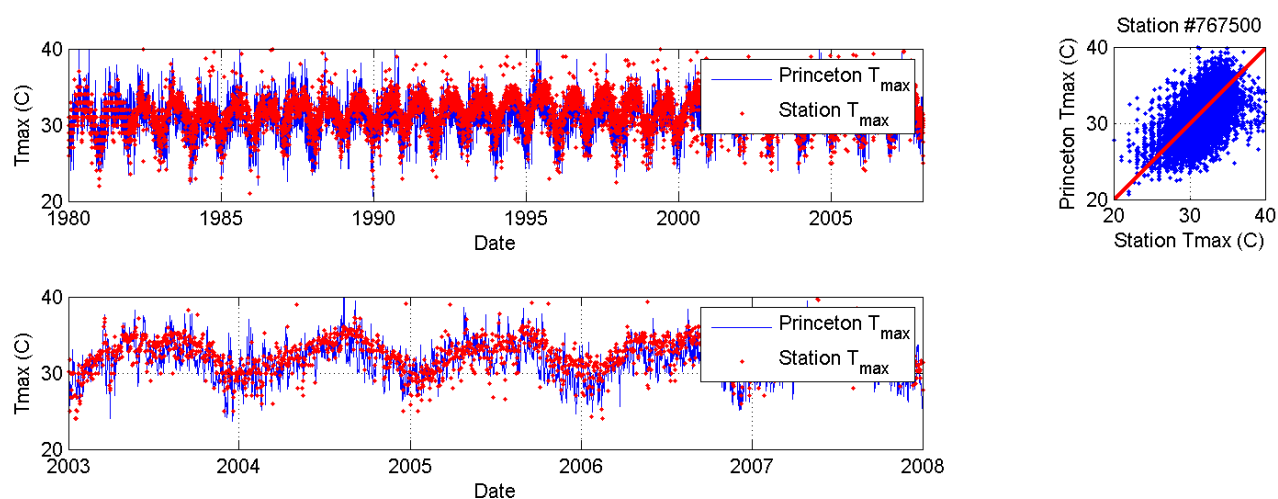


Figure A2.2. Station #767500.

Shown is a fairly typical example of how the two datasets compare: the scatterplot shows that while the two datasets do not always overlap perfectly (we would not expect them to), there is no systematic bias from the Princeton dataset. That is, there are about an equal number of points to the bottom right of the red line as there are to the top left. Looking at the time series (lower two plots), the two datasets also track each other quite well through time.

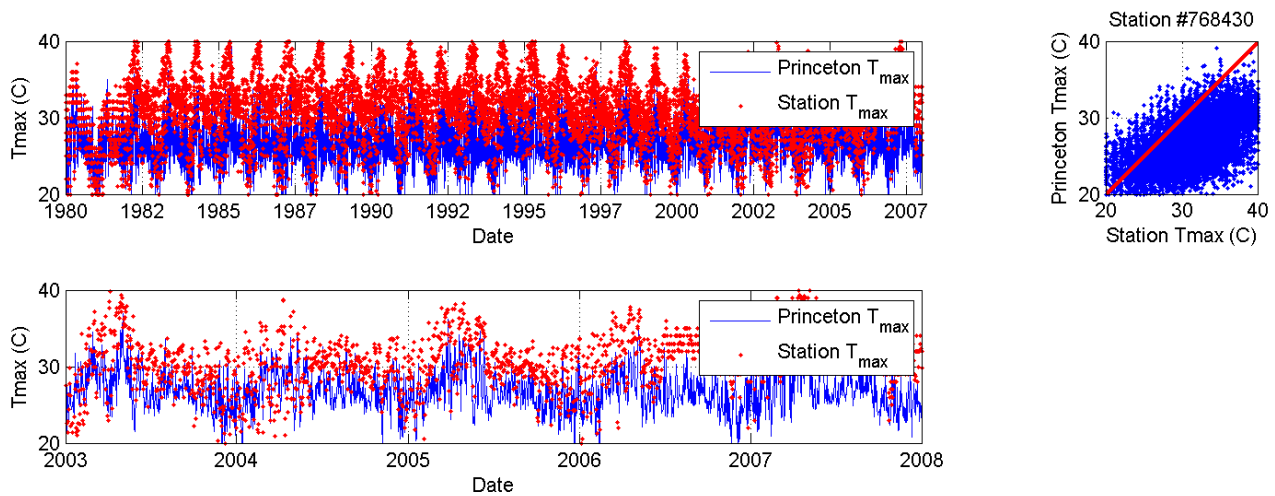


Figure A2.3. Station #768430.

Shown is an example of negative bias in the Princeton data relative to the station data; this can be seen in the higher number of points to the lower right of the 1:1 line in the top plot, and in both of the time series plots, where the red dots are typically above the blue line. On average, this bias looks to be approximately 5°C.

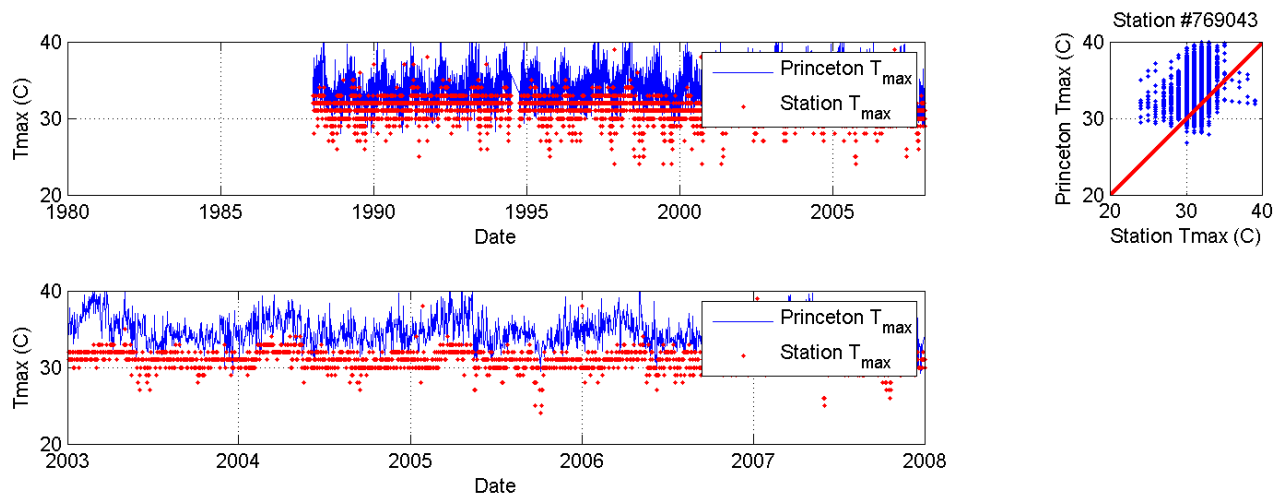


Figure A2.4. Station #769043.

Shown is an example suggesting a positive bias in the Princeton data relative to the station data. The upper plot shows more points to the upper left of the 1:1 line than the lower right; the two time series plots indicate that the station estimates are typically lower than Princeton estimates. This example also illustrates that some of the stations apparently report only integer values for temperature; note that all the station data dots fall on lines corresponding to whole-number increments.

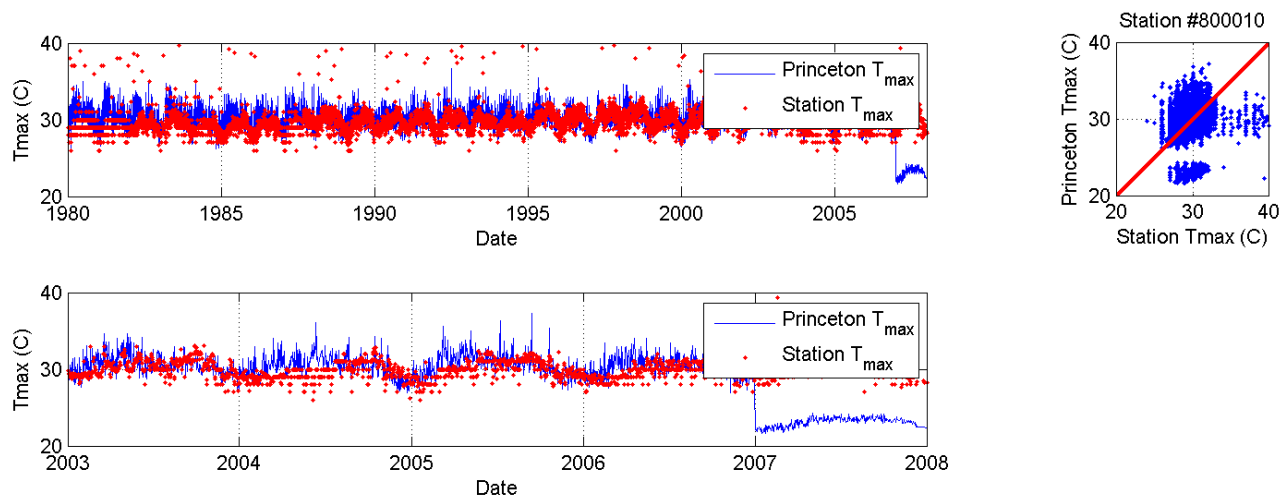


Figure A2.5. Station #800010.

Shown is the only instance we found that appears to indicate something unusual related to the Princeton dataset. First, there appears to be a time lag between the modeled and measured temperatures: Princeton modeled Tmax values peak 2–3 months before measured values peak, as can be seen in the lowermost subplot. In addition, there is a downward step in the Princeton Tmax values by approximately 5–7°C in early 2007. Based on our evaluation of available data, this 5–7°C shift appears to be real, and must reflect some artifact in the Princeton reanalysis. While the time lag should not significantly affect the estimate of Tmax return intervals, allowing this temperature shift to enter into our calculations could potentially create problems if our time series include these measurements. However, we should have the ability to flag these issues when generating our generalized extreme value (GEV) distribution curves and synthetic climate futures.

In summary, based on the evaluation carried out, it is reasonable to say that the Princeton dataset should provide a reasonably good proxy for observed maximum temperatures in the Meso-American region. A key factor in the assessment is that no other database can substitute for the Princeton dataset in terms of its spatial and temporal continuity. Going forward, it may be a good option to choose to calculate GEV curves based on both datasets

in some instances, as a sensitivity analysis to evaluate how extreme event projections might differ when using one dataset vs another. However, since the Princeton data provide the only spatially and temporally continuous data from which to evaluate return intervals, it would be appropriate to proceed with this dataset as our primary climatic data source.

8

References

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, and M. Rummukainen. 2013. Evaluation of climate models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley (eds.). Cambridge University Press, Cambridge, UK and New York.

IPCC. 2007. Summary for policymakers. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller (eds.). Cambridge University Press, Cambridge, UK and New York.

IPCC. 2013. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley (eds.). Cambridge University Press, Cambridge, UK and New York.

IPCC. 2014. *Summary for policymakers. In 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, New York.

Moss, R.H., J.A. Edmonds, K.A. Hibbard, M.R. Manning, S.K. Rose, D.P. van Vuuren, T.R. Carter, S. Emori, M. Kainuma, T. Kram, G.A. Meehl, J.F.B. Mitchell, N. Nakicenovic, K. Riahi, S.J. Smith, R.J. Stouffer, A.M. Thomson, J.P. Weyant, and T.J. Wilbanks. 2010. The next generation of scenarios for climate change research and assessment. *Nature* 463:747–756.

Sheffield, J., G. Goteti, and E.F. Wood. 2006. Development of a 50-yr high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate* 19(13):3088–3111.

Warrick, R. 2009. From CLIMACTS to SimCLIM: Development of an integrated assessment model system. In *Integrated Regional Assessment of Global Climate Change*, C.G. Knight and J. Jäger (eds.). Cambridge University Press, Cambridge, UK. pp. 280–311.

Wigley, T.M.L. 2008. MAGICC/SCENGEN 5.3. National Center for Atmospheric Research, Boulder, CO. Available: <http://www.cgd.ucar.edu/cas/wigley/magicc/>.

Annex References

Carvalho, L.M. and C. Jones. 2013. CMIP5 simulations of low-level tropospheric temperature and moisture over the tropical Americas. *Journal of Climate* 26(17):6257–6286.

Cheng, W., J.C. Chiang, and D. Zhang. 2013. Atlantic meridional overturning circulation (AMOC) in CMIP5 models: RCP and historical simulations. *Journal of Climate* 26(18):7187–7197.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, and M. Rummukainen. 2013. Evaluation of climate models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley (eds.). Cambridge University Press, Cambridge, UK and New York.

Gulizia, C. and I. Camilloni. 2015. Comparative analysis of the ability of a set of CMIP3 and CMIP5 global climate models to represent precipitation in South America. *International Journal of Climatology* 35(4):583–595.

Hidalgo, H. G. and E.J. Alfaro. 2014. Skill of CMIP5 climate models in reproducing 20th century basic climate features in Central America. *International Journal of Climatology*. doi: 10.1002/joc.4216.

Jones, C. and L.M. Carvalho. 2013. Climate change in the South American monsoon system: Present climate and CMIP5 projections. *Journal of Climate* 26(17):6660–6678.

Marengo, J.A., S.C. Chou, R.R. Torres, A. Giarolla, L.M. Alves, and A. Lyra. 2014. Climate Change in Central and South America: Recent Trends, Future Projections, and Impacts on Regional Agriculture. Working Paper No. 73. CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). Copenhagen, Denmark. Available: www.ccafs.cgiar.org.

Ryu, J. H. and K. Hayhoe. 2014. Regional and large-scale influences on seasonal to interdecadal variability in Caribbean surface air temperature in CMIP5 simulations. *Climate Dynamics* 45(1–2):455–475.

Steinhoff, D., A.J. Monaghan, and M.P. Clark. 2015. Projected impact of twenty-first century ENSO changes on rainfall over Central America and northwest South America from CMIP5 AOGCMs. *Climate Dynamics* 44:1329–1349.

Yin, L., R. Fu, E. Shevliakova, and R.E. Dickinson. 2013. How well can CMIP5 simulate precipitation and its controlling processes over tropical South America? *Climate Dynamics* 41(11–12):3127–3143.

Zhang, Y., R. Fu, H. Yu, Y. Qian, R. Dickinson, M.A.F.S. Dias, P.L.D. Dias, and K. Fernandes. 2009. Impact of biomass burning aerosol on the monsoon circulation transition over Amazonia. *Geophysical Research Letters* 36(10).



**CLIMATE CHANGE PROJECTIONS
IN LATIN AMERICA & THE CARIBBEAN**

Review of Existing Regional Climate
Models' Outputs



Inter-American Development Bank
1300 New York Avenue, N.W.
Washington, DC 20577