Centro Internacional de Agricultura Tropical
International Center for Tropical Agriculture
Consultative Group on International Agriculture Research

Agricultura Eco-Eficiente para Reducir la Pobreza

May 2010

Downscaling Global Circulation Model Outputs: The Delta Method
Decision and Policy Analysis Working Paper No. 1

Julian Ramirez-Villegas[1,2,3] and Andy Jarvis[1, 2]

[1] *International Center for Tropical Agriculture, CIAT, Cali, Colombia*
[2] *CGIAR Research Program on Climate Change, Agriculture and Food Security, CCAFS*
[3] *School of Earth and Environment, University of Leeds, UK*

Summary

There has been significant scientific discord over what the best resolution for forecasting the impacts of climate change on agriculture and biodiversity is. Several researchers (particularly climatic researchers) state that original GCM (General Circulation Model) resolution should be kept in order to manage, understand and not bias or alter uncertainties produced by GCMs themselves; however, a coarse resolution of 100 or 200km (or even more) is simply not practical for assessing agricultural landscapes, particularly in the tropics, where orographic and climatic conditions vary significantly across relatively small distances. Moreover, changes in topography and climate variables are not the only factors accounting for variability in agriculture; soils and socioeconomic drivers, also often differ over small distances, influencing agro-ecosystems, increasing uncertainties, and making forecasting and assessment models more inaccurate and complicated to calibrate. Here we present a downscaling method as well as a global database on climate change data that can be used for crop modeling, niche modeling, and more generally, for assessing impacts of climate change on agriculture at fine scales, using any approach that might require monthly maximum, minimum, mean temperatures and monthly total precipitation (from which a set of bioclimatic indices were also derived). This database (with a total of 441 different scenarios –the sum of 24, 20 and 19 GCMs, times 7 time-slices) complements other existing databases that also use downscaling but are only available either for a limited set of GCMs, time-slices, regions, or for variables or at coarser resolution. As such, we provide the most current and comprehensive set of climate change ready-to-use datasets, available online at https://ccafs-climate.org.

Introduction

There has been significant scientific discord over what the best resolution for forecasting the impacts of climate change on agriculture and biodiversity is. Several researchers (particularly climatic researchers) state that original GCM (General Circulation Model) resolution should be kept in order to manage, understand and do not bias or alter uncertainties produced by GCMs themselves; however, a coarse resolution of 100 or 200km (or even more) is simply not practical for assessing agricultural landscapes, particularly in the tropics, where orographic and climatic conditions vary significantly across relatively small distances (Wilby et al., 1998; Tabor and Williams, 2010; Hijmans et al., 2005). Moreover, changes in topography and climate variables are not the only factors accounting for variability in agriculture soils and socioeconomic drivers also often differ over small distances, influencing agro-ecosystems, increasing uncertainties, and making forecasting and assessment models more inaccurate and complicated to calibrate.

Global Circulation Models (GCMs) are large-scale representations of the atmosphere and its processes. A GCM reproduces, with certain accuracy, mass and energy fluxes and storages that occur within the atmosphere, by using an analysis unit. This unit is often called a "cell." These cells are three-dimensional objects within which a number of equations are applied by means of high performance computing units. Given the time and processing capacity required for applying these equations in a single cell (taking into account its interactions with neighbor cells), GCM cells cannot be unlimitedly small; rather, they are restricted to a size of 100-300km. Currently, more than a dozen centers around the world develop climate models to enhance our understanding of climate and climate change and to support the IPCC activities (IPCC, 2001, 2007). There are marked differences between the models, which employ different numerical methods, spatial resolutions, and subgrid-scale parameters (IPCC, 2001 2007; Govindan et al., 2002). Because of these incongruities, researchers making use of climate models output data assess uncertainties using all available GCMs instead of selecting a subset of GCMs.

However, despite the considerable effort done by climate modeling centers, GCM outputs are still too coarse to assess impacts on biodiversity, ecosystem services, agricultural systems, species distributions, conservation planning and other landscape and agriculture related matters (Tabor and Williams, 2010; Zhang, 2006; Fowler et al., 2007; Salathé et al. 2007; Kremen et al., 2008; Jones and Thornton, 2003; Jarvis et al., 2008). To meet that need, scientists have developed various downscaling methods. .

Downscaling techniques allow researchers to obtain regional predictions of climatic changes, ranging from smoothing and interpolation of GCM anomalies (e.g. Tabor and Williams, 2010, among others), to neural networks, and regional climate modeling (Giorgi, 1990). The different downscaling techniques vary in accuracy, output resolution, computational and time requirements, and climatic science robustness (i.e. theoretical background). Regional Climate Models provide 20 to 50km surfaces by re-modeling GCM outputs and are thus only applicable to a limited number of GCMs (for which boundary conditions are available), and require considerable processing capacity, time and storage for obtaining a single scenario-by-period output, thus making it barely feasible to get RCM outputs for most assessment offices and agricultural researchers.

Statistical downscaling, on the other hand, provides an easy-to-apply and much more rapid method for developing high resolution climate change surfaces for high resolution regional climate change impact assessment studies. However, climatologists have lambasted the procedure for degrading data, since downscaling tends to reduce variances (and thus alter uncertainties) and to give off a false sense of increased accuracy, when in actuality, it only provides a smoothed surface of future climates.

However, the price of not disaggregating, interpolating, smoothing, downscaling, or doing whatever possible to increase GCM resolution to a finer scale, could be greater than the inherent degradation of GCM data involved with these approaches. Without high-resolution data inputs, the performance of important assessment tools for conservation planning, niche modeling, crop modeling, and agricultural production and/or biodiversity assessment would be considerably affected.

Here we present a simple downscaling method (named delta method), based on the sum of interpolated anomalies to high resolution monthly climate surfaces from WorldClim (Hijmans et al., 2005). The method produces a smoothed (interpolated) surface of changes in climates (deltas or anomalies) and then applies this interpolated surface to the baseline climate (from WorldClim), taking into account the possible bias due to the difference in baselines. The method assumes that changes in climates are only relevant at coarse scales and that relationships between variables are maintained towards the future. While these assumptions might hold true in a number of cases, they could be wrong in highly heterogeneous landscapes where topographic conditions cause considerable variations over relatively small distances.

The method was applied over 24 different GCMs from the IPCC Fourth Assessment Report (2007), directly downloaded from the Earth System Grid (ESG) data portal, for the emission scenarios SRES-A1B (24 GCMs), SRES-A2 (19 GCMs), and SRES-B1 (20 GCMs), and for 7 different 30 year running mean periods (i.e. 2010-2039 [2020s], 2020-2049 [2030s], 2030-2059 [2040s], 2040-2069 [2050s], 2050-2079 [2060s], 2060-2089 [2070s], and 2070-2099 [2080s]). Each dataset (SRES scenario – GCM – timeslice) comprises 4 variables at a monthly time-step (mean, maximum, minimum temperature, and total precipitation), and at 4 different spatial resolutions (30 arc-seconds, 2.5 arc-minutes, 5 arc-minutes, and 10 arc-minutes). The data is freely available https://ccafs-climate.org.

The downscaling method

Here we apply a downscaling method based on thin plate spline spatial interpolation of anomalies (deltas) of original GCM outputs. Anomalies are interpolated between GCM cell centroids and are then applied to a baseline climate given by a high resolution surface (WorldClim; Hijmans et al., 2005). The method makes the following gross assumptions:

1. Changes in climates vary only over large distances (i.e. as large as GCM side cell size)
2. Relationships between variables in the baseline ("current climates") are likely to be maintained towards the future

We acknowledge that these assumptions might not hold true in highly heterogeneous landscapes, where topography could cause considerable variations in anomalies (i.e. the Andes); however, the assumption is useful for relatively homogeneous or very homogeneous areas such as the Sahara, the Amazon, and other global areas with homogeneous landscapes.

The process consists of the following steps:

1. Gathering of baseline data (current climates corresponding to WorldClim)
2. Gathering of full GCM timeseries
3. Calculation of 30 year running averages for present day simulations (1961-1990) and 7 future periods
4. Calculation of anomalies as the absolute difference between future values in each of the 3 variables to be interpolated (minimum and maximum temperature, and total precipitation)
5. Interpolation of these anomalies using centroids of GCM cells as points for interpolation
6. Addition of the interpolated surfaces to the current climates from WorldClim, using absolute sum for temperatures, and addition of relative changes for precipitation
7. Calculation of mean temperature as the average of maximum and minimum temperatures

WorldClim and full GCM timeseries are freely available in the internet, whilst all other calculations are carried out by means of Geographic Information Systems (GIS) software. Used formats are NetCDF (for GCM outputs), ESRI-GRID (for WorldClim and final downscaled data), and ESRI-ASCII grids for providing standard and easy-of-use outputs to potential users of the data.

Baseline data

In order to obtain credible, high resolution surfaces, we use WorldClim (Hijmans et al., 2005, available at http://www.worldclim.org/), a global database of climate surfaces at 30 arc-second spatial resolution (~1km at the Equator). This database was developed from compiled monthly averages of climate as measured at weather stations from a large number of global, regional, national and local sources, mostly from the 1950-2000 period. We employed the Thin Plate Smoothing Spline (TPS) algorithm (Hutchinson, 1995) to yield climate surfaces for monthly maximum, minimum, mean temperatures and total monthly precipitation.

WorldClim contains data from the Global Historical Climate Network Dataset (GHCN), the WMO Climatological Normals (CLINO), the FAOCLIM global climate database, a database assembled in the International Center for Tropical Agriculture (CIAT), and additional databases from Latin America and the Caribbean (R-Hydronet), the Altiplano in Peru and Bolivia (INTECSA), the 'Nordic Countries' in Europe (Nordklim), Australia (BOM), New Zealand, and Madagascar.

WorldClim climate surfaces were developed from 47,554 locations with precipitation records; 24,542 locations with mean temperature records; and 14,835 locations with minimum and maximum temperature records. Other global datasets have been produced using fewer locations for both temperatures and precipitations (New et al., 2002), but WorldClim has the advantage of having higher spatial resolution (Figure 1).

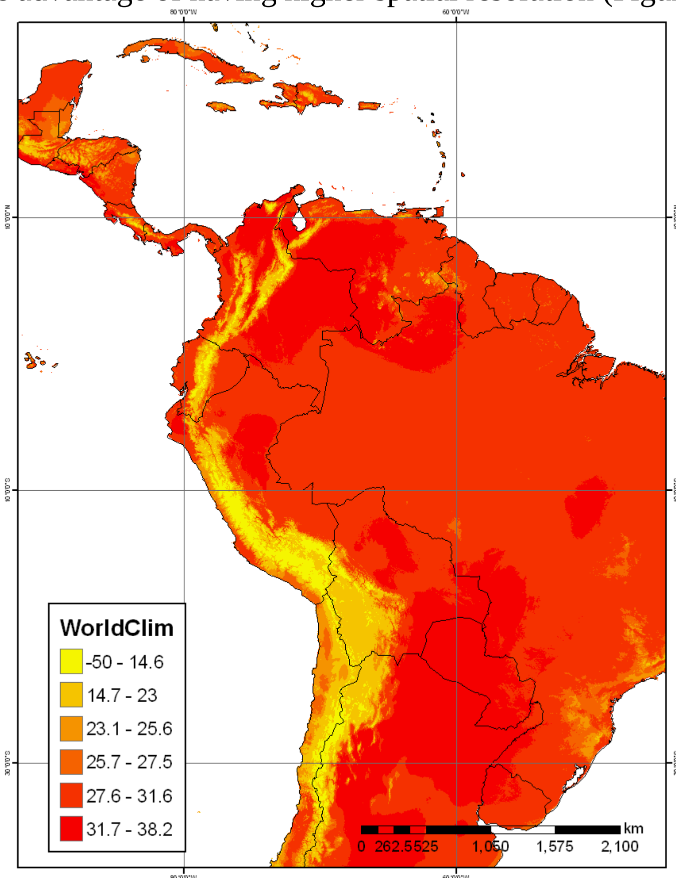

Figure 1 WorldClim surface corresponding to maximum temperature in January, at 30 arc-seconds spatial resolution

While we recognize that the dataset might not be perfect and/or accurate in all parts of the world, it does represent to a considerable extent current climates, as reported by instrumental records, at a scale that permits application of any modeling technique at a site-specific level. Critical areas where very low number of locations was used for interpolations are: the Amazon, the Sahara, Russia, Greenland, and some places in the mid-east, among others (see Hijmans et al., 2005 for further detail).

In addition, WorldClim has been used considerably by modelers, conservationists and agricultural researchers because of its high resolution. The dataset has been cited more than 500 times in peer reviewed publications. For all the above reasons, we chose to use WorldClim for our baseline data, representing the 1961-1990 period (current climates hereafter).

Future GCM predictions

As stated before, GCMs are representations of earth processes and are performed on powerful computers by climatic research centers over the world. To date, a variety of GCMs (with their respective versions) have been developed and tested, and their results have been made available to the public (IPCC, 2001, 2007). 24 Different GCMs have been used in the Fourth Assessment Report (IPCC, 2007), each with different parameterization (Table 1, see atmosphere and ocean columns indicating resolutions). These GCMs have been run under different SRES emission scenarios (IPCC, 2000), but not under all of them. Outputs have been produced for the SRES A1B, A2 and B1 emission scenarios.

Table 1 Available GCMs and principal characteristics (resolutions, references)

Model	Country	Atmosphere	Ocean	Reference
BCCR-BCM2.0	Norway	T63, L31	1.5x0.5, L35	N/A
CCCMA-CGCM3.1 (T47)	Canada	T47 (3.75x3.75), L31	1.85x1.85, L29	Scinocca et al. (2008)
CCCMA-CGCM3.1 (T63)	Canada	T63 (2.8x2.8), L31	1.4x0.94, L29	Scinocca et al. (2008)
CNRM-CM3	France	T63 (2.8x2.8), L45	1.875x(0.5-2), L31	Salas-Mélia et al. (2005)
CSIRO-Mk3.0	Australia	T63, L18	1.875x0.84, L31	Gordon et al. (2002)
CSIRO-Mk3.5	Australia	T63, L18	1.875x0.84, L31	Gordon et al. (2002)
GFDL-CM2.0	USA	2.5x2.0, L24	1.0x(1/3-1), L50	Delworth et al. (2004)
GFDL-CM2.1	USA	2.5x2.0, L24	1.0x(1/3-1), L50	Delworth et al. (2004)
GISS-AOM	USA	4x3, L12	4x3, L16	Russell et al. (1995)
GISS-MODEL-EH	USA	5x4, L20	5x4, L13	Schmidt et al. (2005)
GISS-MODEL-ER	USA	5x4, L20	5x4, L13	Schmidt et al. (2005)
IAP-FGOALS1.0-G	China	2.8x2.8, L26	1x1, L16	Yu et al. (2004)
INGV-ECHAM4	Italy	T42, L19	2x(0.5-2), L31	Gualdi et al. (2006)
INM-CM3.0	Russia	5x4, L21	2.5x2, L33	Diansky et al. (2002)
IPSL-CM4	France	2.5x3.75, L19	2x(1-2), L30	Marti et al. (2005)
MIROC3.2-HIRES	Japan	T106, L56	0.28x0.19, L47	Hasumi and Emori (2004)
MIROC3.2-MEDRES	Japan	T42, L20	1.4x(0.5-1.4), L43	Hasumi and Emori (2004)
MIUB-ECHO-G	Germany/Korea	T30, L19	T42, L20	Grötzner et al. (1996)
MPI-ECHAM5	Germany	T63, L32	1x1, L41	Jungclaus et al. (2005)
MRI-CGCM2.3.2A	Japan	T42, L30	2.5x(0.5-2.0)	Yukimoto et al. (2001)
NCAR-CCSM3.0	USA	T85L26, 1.4x1.4	1x(0.27-1), L40	Collins et al. (2005)
NCAR-PCM1	USA	T42 (2.8x2.8), L18	1x(0.27-1), L40	Washington et al. (2000)
UKMO-HADCM3	UK	3.75x2.5, L19	1.25x1.25, L20	Gordon et al. (2002)
UKMO-HADGEM1	UK	1.875x1.25, L38	1.25x1.25, L20	Johns et al. (2006)

Different Coupled Models Intercomparison Projects (CMIPs) have been created in order to support and enhance knowledge of GCM-related science. The last existing CMIP is the CMIP-3 (PCMDI, 2007; IPCC, 2007), comprising the evaluation of some 22 to 24 different GCMs on a global scale. CMIP-3 also set up a platform for providing GCM outputs to the public, under the Earth System Grid (ESG) online platform (https://esg.llnl.gov:8443/index.jsp).

The IPCC-data portal (http://www.ipcc-data.org) provides some GCM outputs as well, but the most comprehensive dataset is provided by the ESG, including complete timeseries of future simulations (2000-2100) at monthly time-steps, daily data for specific periods (e.g. 2020s, 2050s), yearly data, and 30 year running averages. The IPCC-data portal only provides the last one.

We have downloaded data from ESG corresponding to full timeseries (1850-2100) of all available GCMs (24), at monthly time-steps, for the same 4 variables of interest to us (minimum, maximum, mean temperature, and total precipitation), for the 20CM3 (20th century simulation), and the SRES-A1B, A2 and B1 emission scenarios. Not all GCMs have been run under all emission scenarios (Table 2).

Table 2 Available (o) and not available (x) GCM runs under baseline and three SRES scenarios

Model	20C3M	SRES-A1B	SRES-A2	SRES-B1
BCCR-BCM2.0	o	o	o	o
CCCMA-CGCM3.1-T63	o	o	x	o
CCCMA-CGCM3.1-T47	o	o	o	o
CNRM-CM3	o	o	o	o
CSIRO-MK3.0	o	o	o	o
CSIRO-MK3.5	o	o	o	o
GFDL-CM2.0	o	o	o	o
GFDL-CM2.1	o	o	o	o
GISS-AOM	o	o	x	o
GISS-MODEL-EH	o	o	x	x
GISS-MODEL-ER	o	o	o	o
IAP-FGOALS1.0-G	o	o	x	o
INGV-ECHAM4	o	o	o	x
INM-CM3.0	o	o	o	o
IPSL-CM4	o	o	o	o
MIROC3.2.3-HIRES	o	o	x	o
MIROC3.2.3-MEDRES	o	o	o	o
MIUB-ECHO-G	o	o	o	o
MPI-ECHAM5	o	o	o	o
MRI-CGCM2.3.2A	o	o	o	o
NCAR-CCSM3.0	o	o	o	o
NCAR-PCM1	o	o	o	x
UKMO-HADCM3	o	o	o	o
UKMO-HADGEM1	o	o	o	x
Total	24	24	19	20

An additional issue regards the availability of GCM outputs. Due to a lack of a clear agreement, not all research centers have provided outputs on all variables. Some have instead decided to selectively provide variables, causing a bottleneck for non-climatic research centers hoping to use these data. Hence, minimum and maximum temperatures were not available for all GCMs, but only for 11 (20C3M, A1B, B1) and 9 (A2). For those GCMs for which no maximum and minimum temperature data were available, we used the Multi Model Mean (MMM) of all the other GCMs. While we acknowledge this might reduce variance among the different GCMs, we prefer to provide MMM-based outputs over the alternative of not simply not providing data for those models.

Anomalies: how and why?

Using the full present day (20C3M) monthly timeseries, we calculated 30 year running means around 1985 (1961-1990) as a baseline, for each of the GCMs and the 4 variables of interest. Then we calculated 30 year running means for each of the emission scenarios and seven periods, so that the complete timeseries were reduced to 8 different 30 year averaged periods, as follows:

1. 1961-1990: The baseline climate, also referred to as 20C3M, or ‘current climates’
2. 2010-2039, referred to as 2020s
3. 2020-2049, referred to as 2030s
4. 2030-2059, referred to as 2040s
5. 2040-2069, referred to as 2050s
6. 2050-2079, referred to as 2060s
7. 2060-2089, referred to as 2070s
8. 2070-2099, referred to as 2080s

For each of the 7 future periods, the anomaly or delta with respect to the baseline climate was calculated for each of the variables and months. These anomalies were then interpolated using a thin plate spline interpolation (Franke, 1982; Mitas and Mitasova, 1988).

The basic minimum-curvature technique is also referred to as thin plate interpolation. Thin plate interpolations have been used several times in climatology (Hijmans et al., 2005; Hutchinson, 1995; Hutchinson 1984; Hutchinson and de Hoog, 1985). The procedure ensures a smooth (continuous and differentiable) surface together with continuous, first-derivative surfaces. Rapid changes in gradient or slope (the first derivative) may occur in the vicinity of the data points. The spline method performs a two-dimensional minimum curvature spline interpolation on a point data set resulting in a smooth surface that passes exactly through the input points (Eqn. 1).

$$S(x,y) = T(x,y) + \sum_{j=1}^{N} \lambda_j R(r_j) \qquad \text{[Eqn. 1]}$$

Where,

$j = 1,2,\ldots,N$
N is the number of points,
λ_j are coefficients found by the solution of a system of linear equations
r_j is the distance from the point (x,y) to the jth point

While T(x,y) and R(r) are defined differently depending upon whether the spline is to be calculated by incorporating first or third derivatives into the minimization criteria. Here we use third derivatives (Eqn. 2, 3):

$$T(x,y) = a1 + a2x + a3y \qquad \text{[Eqn. 2]}$$

$$R(r) = \frac{1}{2\pi}\left\{\frac{r^2}{4}\left[\ln\left(\frac{r}{2\tau}\right) + c - 1\right] + \tau^2\left[K_o\left(\frac{r}{\tau}\right) + c + \ln\left(\frac{r}{2\pi}\right)\right]\right\} \qquad \text{[Eqn. 3]}$$

Where,

τ^2 is a parameter defined manually
r is the distance between the point and the sample
K_o is the modified Bessel function
c is a constant equal to 0.577215
a_i are coefficients found by the solution of a system of linear equations

Original GCM cells are transformed to points with position equal to the centroid of the cell, and the thin plate spline interpolation is applied across these points, using 8 points as neighborhood and τ^2 equal to 0.5. The target resolution of this interpolation is 30 arc-seconds, in order to fit with WorldClim, and the procedure is carried out in Arc/Info Workstation 9.3 (ESRI, 2008).

This interpolation procedure yields a 30 arc-second surface of changes in climates for each of the 12 months and 3 variables (we applied the function only for minimum and maximum temperatures, and total rainfall change, since mean temperature is calculated from the average of minimum and maximum temperatures, assuming a normal distribution). A total of 36 interpolated surfaces of monthly changes in climates are produced per GCM and period (Figure 2).

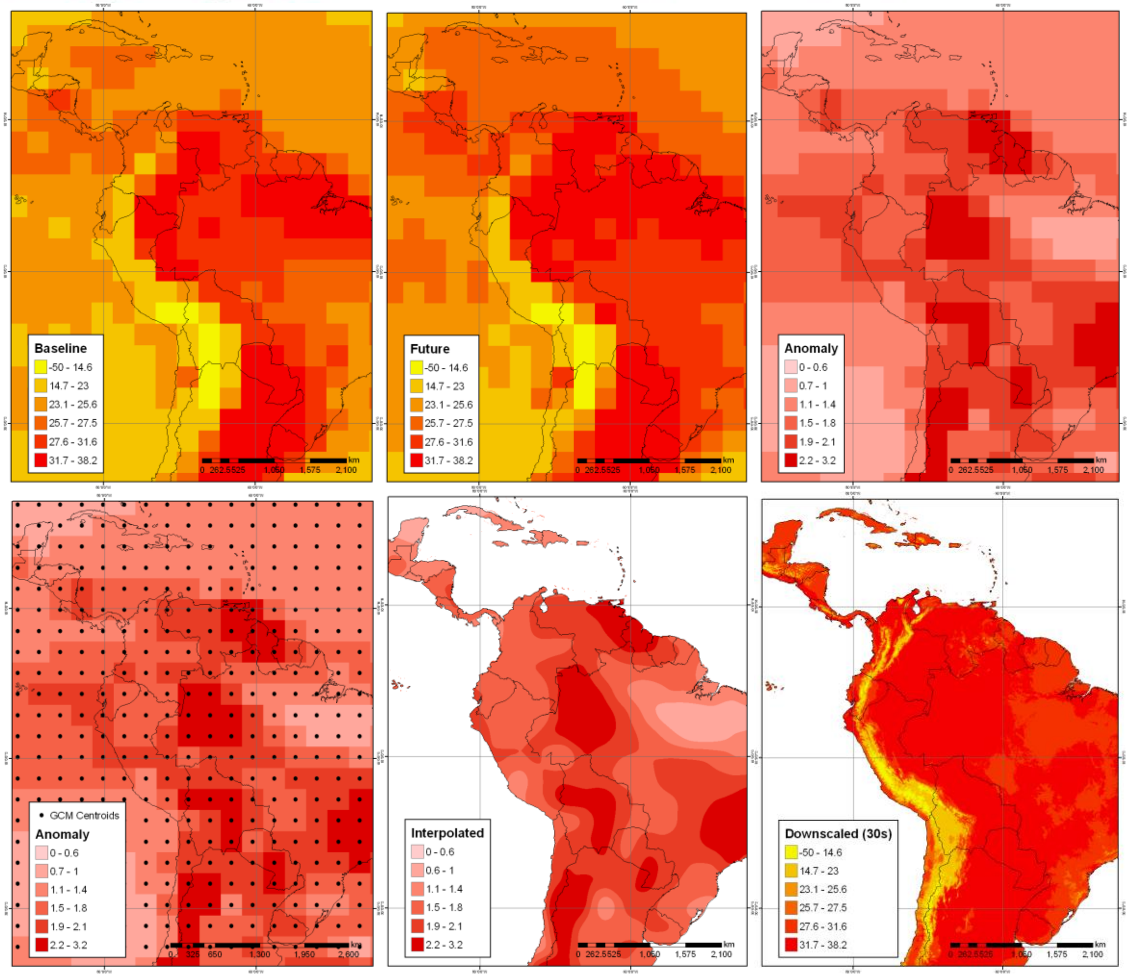

Figure 2 Illustration of the downscaling process with January maximum temperature using the BCCR-BCM2.0 GCM pattern: (a) Baseline data (20C3M), (b) future data for 2050s (2040-2069 average), (c) delta or anomaly by 2050s, (d) delta or anomaly by 2050s with GCM centroids (points) overlaid, (e) 30 arc-s interpolated anomaly, and (e) future downscaled climate surface at 30 arc-second spatial resolution

These surfaces are then applied to the baseline climates from WorldClim. In the case of temperatures (minimum and maximum temperatures) for each pixel, the anomalies in degree Celsius are simply "added" to the actual value in degree Celsius reported in WorldClim. Differences in baselines are neglected for temperatures (Eqn. 4), but taken into account for precipitation [Eqn. 5].

$$X_{F.i} = X_{C.i} + \Delta X_{I.i} \qquad \text{[Eqn. 5]}$$

$$X_{F.i} = X_{C.i} * \left| 1 + \frac{\Delta X_{I.i}}{X_{C.i} + 1} \right| \qquad \text{[Eqn. 6]}$$

Where,

$X_{F.i}$ is the future value of the pixel for the variable *X* (i.e. precipitation, temperature), in the month *i*,
$X_{C.i}$ is the current value (i.e. from WorldClim) of the pixel for the variable *X*, in the month *i*,
$\Delta X_{I.i}$ is the interpolated value of the delta or anomaly corresponding to the pixel, for the variable *X*, in the month *i*,

We add 1 millimeter to the denominator in Eqn. 6 in order to avoid indetermination in areas where current precipitation equals to 0. In Eqn. 6, we use the absolute value of the change relative to the baseline period (i.e. WorldClim) in order to avoid monthly precipitation values going below 0, and maintain homogeneities with WorldClim.

After calculating the corresponding future values for each of the 36 interpolated surfaces, we calculate mean temperatures, assuming a normal distribution in temperatures during the day (Eqn. 7).

$$T_{M.i} = \frac{T_{X.i} + T_{N.i}}{2} \qquad \text{[Eqn. 7]}$$

Where,

$T_{M.i}$ is the mean temperature of the month *i*,
$T_{X.i}$ is the maximum temperature of the month *i*,
$T_{N.i}$ is the minimum temperature of the month *i*,

All these calculations were done in Arc/Info (ESRI, 2008); however, they could have been done under any other automatable GIS software or any other package with the proper libraries (e.g. R, GRASS, Python, Java).

Future downscaled climate surfaces

Our datasets, then, comprise the most up-to-date (with climate science) and comprehensive downscaled set of climate change scenarios, with a total of 441 different scenarios (sum of 24, 20 and 19 GCMs, times 7 time-slices), at 30 arc-seconds spatial resolution. As a whole, our assumptions might lead to uncertainties, and therefore, we suggest that users of these data perform a detailed uncertainty analysis in order to determine if these data in fact fulfill their requirements.

We acknowledge the risk of providing 30 arc-seconds future climate data, but we applied the downscaling functions to the original WorldClim dataset in order to maintain its original condition. However, since 30 arc-s future climate scenarios might create a false sense of accuracy, after all these calculations are carried out, we aggregate the 30 arc-s future data to 2.5, 5, and 10 arc-minute resolutions using nearest neighbor interpolation (Figure 3).

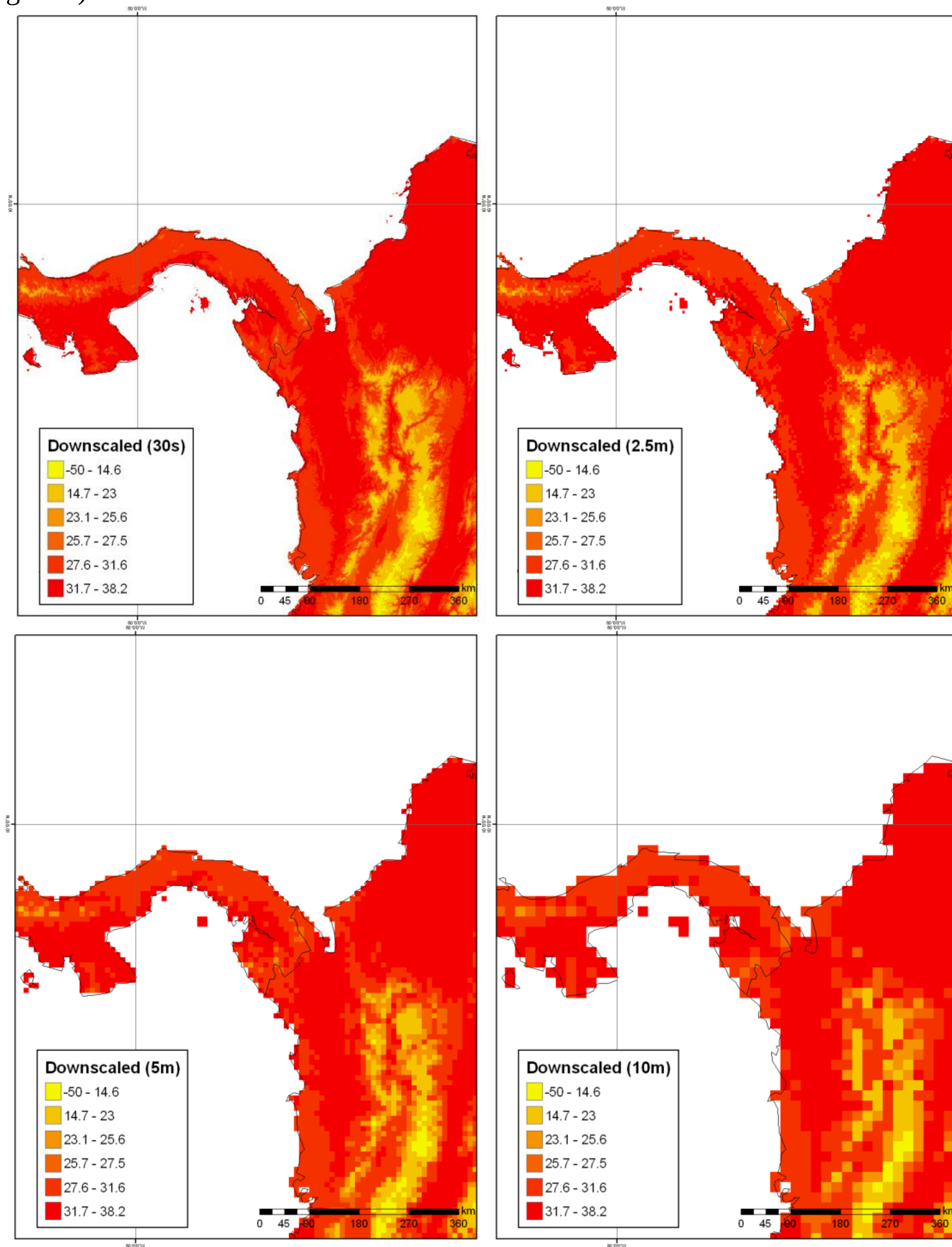

Figure 3 Comparison between downscaled surfaces at different spatial resolutions for an area in north-western Colombia including the Andes: (a) 30 arc-seconds, (b) 2.5 arc-minutes, (c) 5 arc-minutes and (d) 10 arc-minutes. All other datasets (b, c, d) are derived from the 30 arc-s dataset.

We still provide 30 arc-s data, but users of these data should be aware of the risks involved with using these data, due to the assumptions we made in producing them. We caution users regarding the uncertainties involved in our processes, and in no case should users understand these projections as the most accurate surfaces until a deep analysis is done to assess the impact of interpolating anomalies between GCM cell centroids to a higher resolution.

Significant differences are of course present between 30 arc-s and 10 arc-m spatial resolutions. The former is the original WorldClim resolution, providing considerable detail on climatic patterns according to orography, whilst the latter (which is actually the maximum resolution achieved by an RCM), retrieves a credible downscaled surface, but with less level of detail. Values within cells are averaged

Processing and storage capacity in research centers making use of these datasets might also be a limiting factor when using these data. We therefore suggest research centers to download the appropriate resolution datasets that suit to their studies.

Globally and freely available

A webpage has been created for any global user to download the datasets we have produced. This webpage is hosted at Cali, Colombia on CIAT's web server (https://ccafs-climate.org) and contains a brief description of the data. It also provides links to information of all GCM patterns that were downscaled (provided by the IPCC-CMIP3 data portal). and to the datasets in the following formats:

- ESRI Arc/Info binary grids for data at 2.5 arc-m, 5 arc-m, and 10 arc-m spatial resolution
- ESRI ASCII grids for data at 30 arc-s, 2.5 arc-m, 5 arc-m, and 10 arc-m spatial resolution

Beyond the monthly data, we also calculated 19 bioclimatic indices (see Nix, 1986; Busby, 1991), which are often used for niche and crop modeling and are related to the biology and geography of species. These indices provide descriptions of annual trends (i.e. annual mean temperature, total annual rainfall), seasonality (temperature range, temperature and precipitation standard deviations), and stressful conditions (precipitation during dry or wet periods, temperatures during hot and cold periods). These data are also provided on our webpage.

Conclusions

With the recent and rapid spread of ecological niche modeling (ENM) techniques, crop modeling, and geographic information systems (GIS), and the knowledge that climate change is a reality, the need for a detailed dataset of environmental characterizations to assess the impacts of climate change on agricultural production, biodiversity, conservation, water resources, soils, et al, has increased. GCM outputs provide credible surfaces of changes in climates during the 21st century, but these surfaces are too coarse in resolution to be used to characterize very heterogeneous landscapes or to assess the impacts of climate change over these areas.

Downscaling of GCM outputs has been therefore largely used in order to cope with this difficulty. Different downscaling techniques do exist, ranging from smoothing of GCM data, to neural networks and Regional Climate Models (RCMs). Here we used the so-called Delta Method, and created a set of 441 different future climate scenarios at four spatial resolutions (including 30 arc-second [~1km]). The datasets are up-to-date and freely available, but must be used carefully (particularly those at 30 arc-s spatial resolution), given the assumptions we made when creating them. We therefore suggest that users relying on these data complete detailed and comprehensive analyses of uncertainty that appropriately acknowledge the issues surrounding the methods we used here.

References

Busby, J.R. (1991) BIOCLIM – a bioclimatic analysis and prediction system. *In:* C.R. Margules & M.P. Austin (Eds.) *Nature conservation: cost effective biological surveys and data analysis*, pp. 64–68. Canberra, Australia, Commonwealth Scientific and Industrial Research Organisation (CSIRO).

Collins, W.D., Bitz, C.M., Blackmon, M.L., Bonan, G.B., et al. (2005) The Community Climate System Model (CCSM). *Journal of Climate*, special issue.

Delworth, T.L., et al. (2004) GFDL's CM2 global coupled climate models –part 1 Formulation and simulation characteristics. *Journal of Climate*.

Diansky, N.A. and V.B. Zalensky (2002) Simulation of present-day climate with a coupled Atmosphere-ocean general circulation model. *Izvestiya, Atmospheric and Ocean Physiscs*, V.38, No. 6, pp. 732-747.

Fowler, H.J., Blenkinsop, S. and C. Tebaldi (2007) Linking climate change modeling to impact studies: recent advances in downscaling techniques of hydrological modeling. *International Journal of Climatology*, 27:1547-1578.

Franke, R. (1982) Smooth Interpolation of Scattered Data by Local Thin Plate Splines. *Computers & Mathematics with Applications*, Vol. 8. No. 4. pp. 237 - 281. Great Britain.

Giorgi, F (1990) Simulation of regional climate using a limited area model nested in a general circulation model. *Journal of Climate* 3, 941–963.

Gordon, C., Cooper, C., Senior, C.A., Banks, H.T., Gregory, J.M., Johns, T.C., Mitchell, J.F.B. and R.A. Wood (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without adjustments. *Climate Dynamics*, 16: 147-168.

Gordon, H. B., Rotstayn, L. D., McGregor, J. L., Dix, M. R., Kowalczyk, E. A., O'Farrell, S. P., Waterman, L. J., Hirst, A. C., Wilson, S. G., Collier, M. A., Watterson, I. G., and T.I. Elliott (2002) The CSIRO Mk3 Climate System Model [Electronic publication]. Aspendale: CSIRO Atmospheric Research. (CSIRO Atmospheric Research technical paper; no. 60), 130 pp.

Govindan, R.B., Vyushin, D., Bunde, A., Brenner, S., Havlin, S. and H.J. Schellnhuber (2002) Global climate models violate scaling of the observed atmospheric variability. *Physical Review Letters* 89, 028501.

Groetzner, A., Sausen, R., and M. Clausen (1996) The impact of sub-grid scale sea-ice inhomogenities on the performance of the atmospheric general circulation model ECHAM3. *Climate dynamics*, 12: 447-496.

Gualdi, S., Scoccimarro, E. and A. Navarra (2006) Changes in tropical cyclone activity due to global warming: Results from a high-resolution coupled general circulation model. *Journal of Climate*

Hasumi, H. and Emori S. (Eds.) (2004) K-1 Coupled GCM (MIROC) Description. K-1 Technical Report No. 1, CCSR, NIES and FRCGC, September 2004.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., and A. Jarvis (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25:1965-1978.

Hutchinson, MF (1984) A summary of some surface fitting and contouring programs for noisy data.*CSIRO Division of Mathematics and Statistics*, *Consulting Report ACT* 84/6. Canberra, Australia.

Hutchinson, MF, de Hoog FR (1985) Smoothing noisy data with spline functions. *Numerische Mathematik* 47: 99-106.

Hutchinson, MF (1995) Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographic Information Systems*, 9:385-403.

Intergovernmental Panel on Climate Change (2000) IPCC Special Report Emissions Scenarios, Summary for Policymakers, *IPCC*, Geneva.

Intergovernmental Panel on Climate Change (2001) IPCC Third Assessment Report: Climate Change 2001, *IPCC*, Geneva.

Intergovernmental Panel on Climate Change (2007) IPCC Fourth Assessment Report: Climate Change 2007, *IPCC*, Geneva.

Jarvis, A., Lane, A., and R.J. Hijmans (2008) The effect of climate change on crop wild relatives. *Agriculture, Ecosystems & Environment* 126:13-23.

Johns, T.C., Durman, C.F., Banks, H.T., Roberts, M.J., McLaren, A.J., et al. (2006) The new Hadley Centre climate model HadGEM1: Evaluation of coupled simulations. *Journal of Climate*, 19(7): 1327-1353.

Jones, P.G. and P.K. Thornton (2003) The potential impacts of climate change on maize production in Africa and Latin America in 2055. *Global Environmental Change* 13:51–59

Jugnclaus, J.H., Botzet, M., Haak, H., Keenlyside, N., Luo, J-J., Latif, M., Marotzke, J., Mikolajewics, U. and E. Roeckner (2005) Ocean circulation and tropical variability in the AOGCM ECHAM5/MPI-OM. *Journal of Climate*

Kremen, C., et al. (2008) Aligning conservation priorities across taxa in Madagascar with high resolution planning tools. *Science* 320: 222-226

Marti, O., Braconnot, P., Bellier, J., Benshila, R., Bony, S., et al. (2005) The new IPSL climate system model: IPSL-CM4. Institute Pierre Simon Laplace, technical report.

Mitas, L., and H. Mitasova (1988) General Variational Approach to the Interpolation Problem. *Computers & Mathematics with Applications*, Vol. 16., No. 12., pp. 983 - 992. Great Britain.

New, M., Lister, D., Hulme, M., and I. Makin (2002) A high-resolution data set of surface climate over global land areas. *Climate Research*, 21:1-25

Nix, H.A. (1986) A biogeographic analysis of Australian elapid snakes. In: R. Longmore (Ed.). Atlas of elapid snakes of Australia. Australian Flora and Fauna Series 7. Australian Government Publishing Service, Canberra.

PCMDI (2007) IPCC Model Output. Available at: http://www-pcmdi.llnl.gov/ipcc/about_ipcc.php (accessed 26 September 2009).

Russell, G.L., Miller, J.R., and D. Rind (1995) A coupled atmosphere-ocean model for transient climate change studies. *Atmosphere-Ocean* 33 (4): 683-730.

Salas-Mélia, D., F. Chauvin, M. Déqué, H. Douville, J.F. Gueremy, P. Marquet, S. Planton, J.F. Royer and S. Tyteca (2005) Description and validation of the CNRM-CM3 global coupled model, CNRM working note 103.

Salathé, P.Jr., Mote, P.W., and M.W. Wiley (2007) Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States pacific northwest. *International Journal of Climatology*, 27:1611-1621.

Schmidt, G.A., et al. (2005) Present day atmospheric simulations using GISS ModelE: Comparison to in-situ, satellite and reanalysis data. *Journal of Climate*, 19: 153-192.

Scinocca, J.F., McFarlane, N.A., Lazare, M, Li, J. and D. Plummer (2008) The CCCma third generation AGCM and its extension into the middle atmosphere. Atmospheric Chemistry and Physics, 8, 7055-7074.

Tabor, K and J.W. Williams (2010) Globally downscaled climate projections for assessing the conservation impacts of climate change, *Ecological Applications*, 20(2): 554-565.

Washington, W.M., Weatherly, J.W., Meehl, G.A., Semtner, A.J.Jr., Bettge, T.W., Craig, A.P., Strand, W.G.Jr., Arblaster, J.M., Wayland, V.B., James, R. and Y. Zhang (2000) Parallel Climate Model (PCM) control and transient simulations. *Climate dynamics*, 16: 755-774.

Wilby, R.L., Wigley T.M.L., Conway D., Jones, P.D., Hewitson, B.C., Main, J. and D.S. Wilks (1998) Statistical downscaling of general circulation model output: a comparison of methods. *Water Resources Research*, 34: 2995-3008.

Yu Yongqiang, Zhang Xuehong, Guo Yufu (2004) Global coupled ocean-atmosphere general circulation models in LASG/IAP. *Advances in Atmospheric Sciences*, 21: 444-455.

Yukimoto, S., Noda, A., Kitoh, A., Sugi, M., Kitamura, Y., Hosaka, M., Shibata, K., Maeda, S. and T. Uchiyama (2001) The new Meteorological Research Institute coupled GCM (MRI-CGCM2), -model climate and variability. *Papers in Meteorology and Geophysics*, 51: 47-88.

Zhang, X.C. (2006) Spatial downscaling of global climate model output for site-specific assessment of crop production and soil erosion. Agricultural and Forest Meteorology, 135:215-229.