

Climate Analogues:

Finding tomorrow's agriculture today

Working Paper No. 12

CGIAR Research Program on Climate Change,
Agriculture and Food Security (CCAFS)

Julián Ramírez-Villegas

Charlotte Lau

Ann-Kristin Köhler

Johannes Signer

Andy Jarvis

Nigel Arnell

Tom Osborne

Josh Hooker



Working Paper

Climate Analogues: Finding tomorrow's agriculture today

Working Paper no. 12

CGIAR Research Program on Climate Change,
Agriculture and Food Security (CCAFS)

Julián Ramírez-Villegas
Charlotte Lau
Ann-Kristin Köhler
Johannes Signer
Andy Jarvis
Nigel Arnell
Tom Osborne
Josh Hooker

Suggested citation:

Ramírez-Villegas J, Lau C, Köhler A-K, Signer J, Jarvis A, Arnell N, Osborne T, Hooker J. 2011. Climate analogues: finding tomorrow's agriculture today. Working Paper no. 12. Cali, Colombia: CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). Available online at: www.ccafs.cgiar.org

Titles in this Working Paper series aim to disseminate interim climate change, agriculture and food security research and practices and stimulate feedback from the scientific community.

Published by the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS).

CCAFS Coordinating Unit - Department of Agriculture and Ecology, Faculty of Life Sciences, University of Copenhagen, Rolighedsvej 21, DK-1958 Frederiksberg C, Denmark. Tel: +45 35331046; Email: ccaafs@life.ku.dk

Creative Commons License



This Working Paper is licensed under a Creative Commons Attribution – NonCommercial–NoDerivs 3.0 Unported License.

Articles appearing in this publication may be freely quoted and reproduced provided the source is acknowledged. No use of this publication may be made for resale or other commercial purposes.

© 2011 CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS) 2011
CCAFS Working Paper no. 12

DISCLAIMER:

This Working Paper has been prepared as an output the Theme 1: Adaptation to Progressive Climate Change, under the CCAFS program and has not been peer reviewed. Any opinions stated herein are those of the author(s) and do not necessarily reflect the policies or opinions of CCAFS. The geographic designation employed and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of CCAFS concerning the legal status of any country, territory, city or area or its authorities, or concerning the delimitation of its frontiers or boundaries.

All images remain the sole property of their source and may not be used for any purpose without written permission of the source.

Abstract

The analogues approach, developed by CCAFS in R programming, is a novel way of supporting climate and crop models with on-the-ground empirical testing. In essence, the analogues tool connects sites with statistically similar ('analogous') climates, across space (i.e. between locations) and/or time (i.e. with past or future climates). A CCAFS dissimilarity index or Hallegatte index can be used to systematically identify climate analogues across the world, for certain regions, or among specific locations. Users may use default criteria or choose from a variety of global climate models (GCMs), scenarios, and input data. Once analogue sites are identified, information gathered from local field studies or databases can be used and compared to provide data for further studies, propose high-potential adaptation pathways, facilitate farmer-to-farmer exchange of knowledge, validate computational models, test new technologies and/or techniques, or enable us to learn from history. Users may manipulate the tool in the free, open-source R software, or access a simplified user-friendly version online.

Keywords

Climate change; Adaptation; Global climate models; Agriculture; Uncertainty; Climate analogues.

About the authors

Julián Ramírez-Villegas is a postgraduate researcher at the Institute for Climate and Atmospheric Science (ICAS), School of Earth and Environment, University of Leeds, UK. Julián also works with CCAFS' *Theme 1: Adaptation to Progressive Climate Change* and the Decision and Policy Analysis (DAPA) program at the International Center for Tropical Agriculture (CIAT), where he assesses the impact of climate change on agriculture in developing countries, primarily in sub-Saharan Africa and South Asia. Other topics on which he works include plant genetic resources conservation, biodiversity conservation, downscaling and assessment of climate model outputs, and applications of the analogues methodology. He played a key role in creating the CCAFS analogues tool and currently maintains the R package.
Email: j.r.villegas@cgiar.org

Charlotte Lau is a researcher and project coordinator for the CCAFS Indo-Gangetic Plains region. She coordinates the *Farms of the Future* project.
Email: c.lau@cgiar.org

Ann-Kristin Köhler is a postdoctoral fellow at the Institute of Climate and Atmospheric Science (ICAS), School of Earth and Environment, University of Leeds, UK. Ann's work focuses on climate-change impact assessment of agricultural systems in Africa and Asia, using the General Large-Area Model for annual crops (GLAM).
Email: A.K.Koehler@leeds.ac.uk

Johannes Signer was a visiting researcher at CIAT and formerly maintained the analogues R package. He has a strong background in modelling, programming, geoinformatics, and ecology, and played an instrumental role in building the analogues methodology.

Andy Jarvis is co-leader of CCAFS' *Theme 1: Adaptation to Progressive Climate Change*. He is also leader of the Decision and Policy Analysis (DAPA) program at CIAT. His research has focused on the application of spatial analysis and modelling to biodiversity conservation, adaptation of livelihoods to climate change, and maintenance of ecosystem services.
Email: a.jarvis@cgiar.org

Nigel Arnell is the director of the Walker Institute for Climate System Research, a professor in the Department of Meteorology at the University of Reading, UK, and member of the Lead Expert Group for the report *Foresight: Migration and Global Environmental Change*. He is also a lead author for the ‘freshwater’ chapter in the IPCC Working Group II’s *Fifth assessment report (AR5)* and for the *IPCC Special report on managing the risks of extreme events and disasters*. His expertise lies in climate change, hydrology, and impact on water resources.
Email: n.w.arnell@reading.ac.uk

Tom Osborne is a postdoctoral research fellow with NCAS–Climate, Department of Meteorology, University of Reading, UK. He has significant experience with coupled climate-crop modelling, particularly to project the impact of climate change on crop production.
Email: t.m.osborne@reading.ac.uk

Josh Hooker is a postdoctoral research fellow with the Departments of Meteorology and Agriculture, University of Reading, UK. His research involves crop models, the impact of climate change on biodiversity, and statistical analysis. He is currently developing a crop-and-land-surface model that will be incorporated into the QESM earth system model.
Email: j.hooker@reading.ac.uk

Contents

1. Introduction: the gaps in climate science	8
2. Conceptual framework	9
2.1 Climate analogues	9
2.2 Added value	9
3. The climate analogues tool	10
3.1 General description.....	10
3.2 Methodology.....	12
3.3 Technical description.....	14
3.4 Analysing environmental dissimilarity.....	15
3.5 Quantification of uncertainties	16
3.6 Known limitations and proposed workarounds.....	18
3.7 How to get the analogues tool	19
3.8 Collaborators	19
3.9 Future plans and developments	20
4. Case study: finding analogues for Lawra-Jirapa, Ghana	20
4.1 Measuring dissimilarities using different variables.....	21
4.2 Combining variable-specific dissimilarities	23
4.3 Testing sensitivities to the importance of different variables	25
4.4 Quantifying dissimilarity uncertainty.....	27
4.5 Beyond the tool: an agronomic perspective	30
5. Final analysis: results and identified analogues.....	30
5.1 Site-specific comparison of dissimilarities.....	30
5.2 Point-based analysis: mapping dissimilarities across many sites.....	33
5.3 Case study lessons: future research and adaptation pathways.....	35
6. CCAFS projects: from model to field.....	38
6.1 Farmer-to-farmer exchanges.....	38
6.2 Field trial sites	38
Conclusions and recommendations.....	39
References.....	39

Acronyms and abbreviations

CCAFS	CGIAR Research Program on Climate Change, Agriculture and Food Security
CGIAR	Consultative Group on International Agricultural Research
CIAT	Centro Internacional de Agricultura Tropical (<i>International Center for Tropical Agriculture</i>)
CV	coefficient of variation
DAPA	Decision and Policy Analysis program at CIAT
DTR	diurnal temperature range
GCM	global climate model
GDAL	Geospatial Data Abstraction Library
GIS	geographic information system
GLAM	General Large-Area Model for annual crops
GRASS GIS	Geographic Resources Analysis Support System—a free open-source GIS software
ICAS	Institute for Climate and Atmospheric Science, University of Leeds, UK
IPCC	Intergovernmental Panel on Climate Change
LJG	Lawra-Jirapa, Ghana
MD	multiplying dissimilarities
NDVI	normalized difference vegetation index
NSZ	no standardization
RE	range among predictions (<i>statistical</i>)
RS	relative similarity
SD	standard deviation
SZ	standardization
SRES A1B	one of four scenario families used by the IPCC in its <i>Special reports on emissions scenarios</i>

1. Introduction: the gaps in climate science

Scientific evidence gathered over the last couple of decades suggests that climate conditions are changing rapidly and that this trend is likely to continue and even accelerate (IPCC 2007, Moss et al. 2010). These anticipated changes in climate baseline, variability, and extremes will have far-reaching consequences on agricultural production, posing additional challenges to meeting food security for a growing world population (Lobell et al. 2008, Roudier et al. 2011). Future farming and food systems will face substantial, albeit distinct, changes in their environments. Some regions (the few winners) may benefit from more favourable climate conditions for production, while others (the larger group of losers) will face increased climate-change-related biotic and abiotic stresses (IPCC 2007).

Where conditions improve, the traditional farming systems will be challenged to exploit the additional production potential; and where conditions deteriorate, accelerated adaptation will be vital, as centuries-old coping mechanisms used by farmers become insufficient or obsolete (Jarvis et al. 2011). As climate ‘migrates’ between regions, it will disproportionately affect resource-poor and marginalized farmers who have less adaptive capacity but depend entirely on agriculture for their livelihoods (Hitz and Smith 2004, Thornton et al. 2011). Hence, men and women farmers alike will need to enhance their capacity to adapt. Research can help in this effort by improving farmers’ (and scientists’) understanding of climate projections and adaptation pathways.

Another major research gap concerns human behaviour, and the cultural and institutional vehicles or barriers to adaptive change (Thornton et al. 2011). In political and development realms, national plans and policy decisions on climate change adaptation are increasingly based on assessments that rely heavily on projections made by mechanistic computational models (e.g. general circulation models, crop response models, and agricultural trade models). Despite advances in climate science in the past decade and the emergence of more complex, integrated models, substantial uncertainty still exists (Challinor and Wheeler 2008, Challinor et al. 2009).

By definition, modelling predictions cannot be fully validated until the projected year actually arrives. As such, significant errors can occur when relying too much on models to understand the agricultural future (Challinor and Wheeler 2008). Climate and crop models can provide projections of biophysical change, but they cannot adequately consider human behaviour, particularly farmers’ capacity to innovate and respond to emerging threats. Computational models cannot tell us what kind of farming systems, supported by projected future conditions, might exist in a given location (Lobell and Burke 2008).

Substantial research funds and energies have been invested in creating more resilient crop varieties and helping farming communities adopt site-specific adaptive practices. But not enough has been done to aggregate inventories of existing local adaptive knowledge or to facilitate farmer-to-farmer exchange of that knowledge between communities facing similar challenges.

2. Conceptual framework

2.1 Climate analogues

The analogues approach is a novel way of supporting modelled policy recommendations with on-the-ground empirical testing. In essence, the analogues tool connects sites with statistically similar (‘analogous’) climates across space (i.e. in other geographic locations) and/or time (i.e. with historical or projected future climates). It helps answer the following questions:

Where can I find sites that...	<ul style="list-style-type: none"> • are at present • were in the past (x year) • are projected to be in the future (y year) 	... analogous to my selected site...	<ul style="list-style-type: none"> • at present? • in the past (z year)? • in the future (projected n year)?
--------------------------------	---	--------------------------------------	---

For instance, if the given place of interest is Lawra-Jirapa, Ghana, then we can locate another site elsewhere in the world whose climate today is analogous to Lawra-Jirapa’s predicted future climate (in *n* year), or vice versa. A ‘spatial analogue’ is therefore a location whose climate today appears as a likely analogue to the projected future climate of another location. Thus, the two sites represent promising areas for comparative research on adaptation plans.

‘Temporal analogues’ use historical data, allowing us to identify historical events that may provide insight into the possible future consequences of climate change, as well as learn how farmers adapted to climatic shifts in the past. However, the current version of the analogues tool does not yet enable historical searches, but CCAFS expects this important function to be added in the future.

Taken as a whole, the analogues approach can help link top-down global models with targeted field trials or visits.

2.2 Added value

Once analogue sites are identified, information gathered from local field studies or databases can be used and compared to develop further studies or propose high-potential adaptation pathways.

Comparisons between present-day farming systems and their spatial or temporal analogues can be useful for:

- *Facilitating farmer-to-farmer exchange of knowledge.* By identifying and connecting analogous sites, research can enable farmers to better envision how their site-specific agricultural future might look. Accordingly, they can facilitate the creation of a knowledge chain through which strategies and farming information can be passed down or shared. In particular, through this network of innovative farmers who learn by doing, other farmers can interact and learn strategies to more effectively adapt to climate change.
- *Validating computational models and testing new technologies and/or techniques.* The analogues tool, coupled with on-the-ground studies (e.g. field trials, samples, farmer-to-farmer visits, and household surveys), link computational models with existing farm realities. This allows researchers to better understand which agricultural systems can survive specific climatic and other conditions, and why. The tool connects what otherwise would be esoteric projections of the future with realities already existing today. Thus, the analogues tool also permits targeted field testing of the climate resilience of cropping systems or technologies, which, if effective, could then be implemented in analogous locations in the future.
- *Learning from history.* Development interventions often fail because they lack adequate information about rural household behaviour and decision-making patterns under climate stress. The analogues methodology can use historical data to show us when and how different agricultural communities, which shared similar climates or experienced similar climatic shifts, successfully (or unsuccessfully) adapted their production systems. These case studies can then be analysed for lessons learned, thus building understanding of the best ways to improve climate resilience and enable adaptation.

3. The climate analogues tool

3.1 General description

The climate analogues tool identifies areas where either the climate today corresponds to the future climate projected at another location, or the projected future climate corresponds to the current climate of another site. Users specify a location (the ‘reference’ location); variables (e.g. temperature and precipitation); and, for future-related analyses, one or more climate scenarios. A climate scenario is defined as a combination of a year (e.g. 2020, 2030, or 2050), an SRES

emissions scenario (IPCC 2000, Moss et al. 2010), and a global climate model (GCM) (IPCC 2007, PCMDI 2007). Data for variables can be on any time step (i.e. hourly, daily, monthly, or yearly).

When searching for global present-day analogues for the 2050 climate of a given reference site A, the analogues tool would first use the climate scenario to forecast A’s climate. It would then compare the present-day climate for all points in the world where data exist with A’s projected climate, and place the results on a dissimilarity index. The comparison would be based solely on the variables specified by the user. Measures of dissimilarity would then be exported as either tables or gridded datasets, which could be later imported into any other statistical or geographic information system (GIS) software for data analysis and plotting.

While only climatic data are considered here, other variables such as soils, crops, and socio-economic characteristics can also be taken into account. The tool permits users to manipulate a range of inputs and outputs, as summarized in table 1:

Table 1. Summary of the different inputs or variables that the user provides or selects, and the options for determining the processes the tool will follow to find analogues.

Inputs or variables	Process and outputs
<ul style="list-style-type: none"> • Reference site: For which location do you want to find an analogue? • Search range: Where do you want to search, for example, the whole world or a set of locations? • Variables: For example, temperature, precipitation, or soils (according to data availability) • Period or year: Historical, present, or future (e.g. 2030 or 2050) • Climate scenario(s) (only for future-related analyses): GCM, SRES, specific or average scenarios • Weight: See sections 3.3 and 3.4 below • Time step: Days, months, or seasons 	<ul style="list-style-type: none"> • Dissimilarity indices: for identifying analogous sites, for example, the CCAFS or Hallegatte: <i>See</i> next section 3.2 • Uncertainty measures: For example, standard deviation, coefficient of variation, range among predictions (RE), and agreement (stippling): <i>See</i> section 3.5 • Thresholds for determining the levels of (dis)similarity that qualify a site as being an analogue

The tool is coded entirely as a library for the R environment for statistical computing (R Development Core Team 2011). It is implemented through raster, rgdal, sp, maps, and maptool packages. It also makes use of the stringr, akima, grid, and rimage packages (all available, free of charge, at <http://www.r-project.org/>) to permit compatibility, easy export of outputs, and efficient use of the memory to the extent possible with R. The package is being optimized, using

the R package `sgrass6`, for large datasets with GRASS GIS, a free, open-source, GIS package available at <http://grass.fbk.eu/>.

With the climate analogues tool, calculations can be done and outputs generated for any geographic region at any resolution equal to or above 1 km. However, depending on the processing power of the computer being used, and on the spatial resolution and amount of data (i.e. number of climate scenarios or time steps) being analysed, computational times may range from a few seconds to several hours or days. Hence, we strongly recommend that those using our package properly design their experiments to avoid large computational delays.

A more user-friendly, web-based version of the climate analogues tool is also available, with an interface that can receive and process queries. This platform is intended to provide simplified but critical insights into the vulnerability of crops to climate change. It can be accessed at <http://gismap.ciat.cgiar.org/Analogues>.

3.2 Methodology

We use two different dissimilarity measures: the ‘CCAFS dissimilarity’, which is similar to that described by Williams et al. (2007), and the ‘Hallegatte dissimilarity’ (Hallegatte et al. 2007). Both measures can be used for any variable for which data are available for several time steps (often days or months, although this can be reduced to a growing season of interest).

3.2.1 The CCAFS dissimilarity measure

Future and present climates are described as vectors of m sequential mean values for v variables (V) and v weights (W). Dissimilarity is then calculated as a weighted euclidean distance between the variables’ vectors for the reference (f) and target (p) scenarios. In Equation 1, each variable difference is weighted (i.e. multiplied) with its corresponding weight to account for the different units used. The weight can also be used to alter, or better reflect, the relative importance of the variable in the dissimilarity value:

$$CCAFS = \left(\sum_{i=1}^m \sum_{j=1}^v W_{ij} \cdot (V_{ijf} - V_{ijp})^z \right)^{1/z} \quad [\text{Eq. 1}]$$

where, z is a parameter that, when equal to 2, produces euclidean distances, and can be changed to perform sensitivity analyses.

A weight can be either a single number or the rate of change in another climate variable (X) with the same time step (m). In that case, the weight is defined as the reference value for that variable (X), divided by its respective target value (Equation 2):

$$W_{ij} = \frac{X_{ijf}}{X_{ijp}} \quad [\text{Eq. 2}]$$

Seasonal variations do not occur concurrently in different locations (e.g. the rainy season in southern Africa does not occur at the same time as it does in the Mediterranean. The analogues tool can account for this by searching across all m time steps for the minimum dissimilarity, using a time lag (lag, Equation 3):

$$CCAFS = \min_{0 \leq \text{lag} \leq m-1} (CCAFS_{i=1...m}) \quad [\text{Eq. 3}]$$

To illustrate the above, let us assume that we have gathered data for the present day as the target scenario (p) and for a given future scenario (reference scenario f). The data consist of rainfall, mean temperatures, and diurnal temperature range (DTR) for 12 months (average climatology for both scenarios). We want to calculate dissimilarity, using total monthly rainfall (P) and mean temperature (T), weighting P with a factor of 1 and weighting T, using the diurnal temperature range (DTR). We use DTR as a weight because it acts as a surrogate for within-month temperature variability. As we are analysing 12 months, $m = 12$, z is again a parameter that can be varied, and we will also account for the lag. By replacing Equations 1 and 2 with Equation 3, we arrive at our CCAFS dissimilarity calculation (Equation 4):

$$CCAFS = \min_{0 \leq \text{lag} \leq 11} \left(\sum_{i=1}^{12} \frac{DTR_{if}}{DTR_{ip}} \cdot (T_{if} - T_{ip})^z + 1 \cdot (P_{if} - P_{ip})^z \right)^{1/z} \quad [\text{Eq. 4}]$$

We use the term ‘dissimilarity’ instead of ‘similarity’ only for convenience in scaling the CCAFS measure: the higher the value, the more dissimilar the two sites are for that particular pair of climate scenarios. In other words, the lowest value represents the best climate analogue.

3.2.2 The Hallegatte dissimilarity measure

Hallegatte et al. (2007) described a method to identify current analogues for future city climates. The projected climate of a given city is compared with a range of current city climates, again defined by vectors of monthly values for a given set of variables with monthly time steps. We generalize it to work with any pair of climate scenarios and with variables with any number of time steps (m). Again, we define the target (p) and reference (f) scenarios. Using this methodology, analogues are defined as candidate locations that meet all of the following three characteristics:

- *Relative difference between total values < a*

$$\frac{\sum_{i=1}^m V_{if} - \sum_{i=1}^m V_{ip}}{\sum_{i=1}^m V_{ip}} < a \quad [\text{Eq. 5}]$$

- *Mean absolute relative differences between mean values for m steps < b*

$$\frac{\sum_{i=1}^m \frac{\text{abs}(V_{if} - V_{ip})}{V_{ip}}}{m} < b \quad [\text{Eq. 6}]$$

- *Mean absolute difference between total values for m steps < c*

$$\frac{\sum_{i=1}^m \text{abs}(V_{if} - V_{ip})}{m} < c \quad [\text{Eq. 7}]$$

In the original Hallegatte et al. (2007) analysis, the first two conditions (Equations 5 and 6) were used for monthly rainfall totals, with $a = 0.15$ and $b = 0.3$, and the third condition was used for mean monthly temperatures, with $c = 1$. Our default configuration includes this original configuration but, as we found it too restrictive for many sites, we also allow users to choose their own variables, conditions, and values. That is, our tool allows users to switch off conditions, so that analogues are the sites that meet the selected conditions. This adds flexibility to the method.

Because the Hallegatte et al. (2007) method does not automatically take into account seasonality on its own, we also implement a lagged search for analogue sites. The method is slightly different to the CCAFS measure as, in this case, we would consider a site to be an analogue if at least one of the lagged sequences of m time steps for the given variable meets the selected criteria.

3.3 Technical description

R is a flexible software package that allows the integration of new features without necessarily modifying the core functions of the software itself. This functionality is provided by means of packages or libraries, which are built using a basic template. R packages can incorporate new methods and functions, and even introduce new concepts (i.e. new object types or classes).

A package often consists of two main components: scripts and documentation. Scripts are R code files that contain the instructions and commands to be executed, whereas documentation

provides detailed information on what variables, arguments, and specific details need to be known when using the functions in the script files. An R package can ‘export’ one or many functions and, for each exported function, documentation must necessarily exist. Hence, users can see and use all exported functions or methods, whereas all the other non-exported features are, by default, running in the background.

The analogues tool is based on the application of the equations described in section 3.2 on a given structure of data (i.e. an R list). Two types of data can be loaded into R for a given dissimilarity analysis: gridded data (raster objects, from the package raster), and matrices (built-in objects in R). These data can be loaded for any variable (V), any weight (W), with any time step (m), and for any combination of climate scenarios (*see* section 3.2). Therefore, users of this tool can analyse dissimilarities of any commonly available non-climatic variables, even when values represent averages of many years (e.g. soil variables), or have different measurement periods (i.e. quarterly; yearly, as for bioclimatic indices; or 16-day periods, as for the normalized difference vegetation index [NDVI]).

3.4 Analysing environmental dissimilarity

Two types of analyses can be performed with the analogues tool: grid-based and point-based analyses. With the former, a single location is compared with the whole geographic domain. In this case, data for the variables must be provided as geographic gridded data (i.e. raster datasets) and are loaded, using the available drivers in the R package ‘rgdal’. The weights may be provided as single numbers or again as raster datasets.

For point-based analyses, a number of points in a given matrix are compared with each other. In this case, data for the variables must be loaded onto the computer’s memory as matrices, and weights may be provided as single numbers or as matrices in the computer’s memory.

3.4.1 Grid-based dissimilarity analysis

Two basic inputs are required in the analogues tool for any grid-based dissimilarity analysis: variables and weights. From these two inputs, variables must be geographic gridded data, and weights may be either geographic gridded data or single values. Any geographic data inputted into this analysis must have exactly the same spatial resolution and the same geographic coverage (i.e. extent). They must also be time consistent (i.e. m must be equal for all these data).

Once a pair of climate scenarios has been selected and the proper paths to the data have been specified, the grid-based dissimilarity analysis function in the analogues tool will load the data for the specified variables, using a combination of [*climate scenario*]-[V]-[m].[*ext*], where V and

m are defined as in section 3.2, climate scenario as in section 3.1, and ext refers to the file extension in the computer's file system. This extension must be in a format supported by the Geospatial Data Abstraction Library (GDAL), or data will not load. Hence, we strongly recommend that users convert their data beforehand to a GDAL-compatible format. However, advanced users with other data formats may create raster objects out of their data instead.

The weights are then loaded, using the same combination $[climate\ scenario]-[W]-[m].[ext]$, where W is defined as in section 3.2, except when the weight is specified to be a single number. In this case, the tool will not load data from the file system; instead, it will accept the value entered as the weighting value for the corresponding variable.

The output of this function is a RasterLayer object, with the same geographic characteristics (i.e. resolution and extent) of the input data used to drive the calculation. These data can be further stored in the file system or visualized in R, using methods and functions already implemented in the R package raster.

3.4.2 Point-based dissimilarity analysis

For point-based dissimilarity analyses, data for variables must already be loaded onto the memory as matrices. Weights may be specified as single numbers, or must already be loaded onto the memory as matrices, as with the data for variables. These matrices should have m columns and s rows, where s is the total number of sites being analysed (from 1 to any number). All matrices must be of equal dimensions.

Data for point-based analyses can exist in any R-readable format within the file system and must be loaded beforehand by the user as a variable with the appropriate name. Currently, R has built-in functions to read data in comma-separated-values format (.csv), Fortran-formatted text files (.txt), and as tab-delimited data files (.dat).

Once a pair of climate scenarios has been selected, the point-based dissimilarity analysis function in the analogues tool will look for the data on the specified variables within the system's memory. The variables must be properly named when data are being loaded beforehand by the user, with a combination of $[V].[climate\ scenario]$, where V is defined as in section 3.2 and the climate scenario as in section 3.1. Otherwise, the tool will fail to find data or will search, using the wrong data.

3.5 Quantification of uncertainties

The analyses explained in section 3.4 can be applied to any combination of climate scenarios. Therefore, when searching for future-related analogues, the user may take into account a variety

of climate projections by choosing a set of different GCMs, under one or more emissions scenarios. This also permits the quantification of uncertainties.

However, before we review methods of quantifying uncertainties, two caveats should be recognized. First, because uncertainties can be estimated by using many different case-specific measures, a measure that works as an effective indicator in one instance may not do so in others (Challinor and Wheeler 2008). Second, one uncertainty indicator does not necessarily represent all other uncertainty indicators. That is, one measure's indication of 'high uncertainty' does not mean that the projection is uncertain in absolute terms, but simply uncertain, relative to that particular uncertainty measure. Even statistical dispersion measures derived from similar concepts, do not agree all the time (Jarvis et al. 2010).

Users are free to use R to explore a broader range of uncertainty quantification approaches, as we provide individual-run results. In this R package, we decided to implement two measures of uncertainty: the standard deviation (SD) and the coefficient of variation (CV). Below, we briefly review these and two other uncertainty measures.

- *Standard deviation (SD)*. High SD values are associated with high uncertainties, and are also associated with high values in the target variable. Hence, a site with higher values in the target variable would inevitably have a higher uncertainty (SD), relative to a site with lower values.
- *Coefficient of variation (CV)*. In contrast, the CV can be extremely high if the values in the target variables are very close to zero, even if the variability is low.
- *Range among predictions (RE)*. The use of RE avoids this scaling issue, but it is also highly sensitive to outliers. For example, if a certain prediction is 90% concentrated in a small part of the range and outliers exist in the distribution of values, the RE might be high, but uncertainty would necessarily be low, as the probability of predicting a single outcome is high.
- *Agreement (or stippling)*. Agreement, or stippling, partly solves this problem, as it quantifies the number of model predictions that agree on a given condition. That is, stippling shows the number of cases that agree on the direction (e.g. higher versus lower rainfall) of a given condition, but not on the differences in numerical magnitude (e.g. by how many millimetres). This method therefore only partly informs on uncertainties. It also fails when the values are evenly distributed.

3.6 Known limitations and proposed workarounds

The current approach has several known limitations:

First, several technical issues are involved when analysing very large areas at very high resolution, and calculations are slow, taking hours or even days to complete. A workaround for this is the future implementation of a GRASS GIS interfacing function that can handle heavier datasets without loading them onto the computer's memory.

Second, although we use weights to account for differences between the scales of the variables (e.g. millimetres for rainfall and Celsius degrees for temperature), we must remember that the CCAFS measure combines two different scales. Two possible solutions to this issue are to:

- Standardize (scale) the data by using the average and standard deviations. However, this may introduce biases by reducing the degree of dissimilarity between sites and/or altering the regional importance of the variables, as standardization essentially adds another layer of weighting. Note that with spatio-temporal data, it is possible to standardize temporally (all sites at one time), or spatially (all times at one site). In this paper, we deal with two single time slices (i.e. baseline and 2030s) rather than a climate time series, so we chose to perform a spatial standardization, as a temporal standardization on a pixel basis would have been meaningless. We standardized the pixel values using the overall (future, present) mean of all months, individually for temperature, diurnal temperature range and rainfall. However, the response of the analogue tool to different ways of standardizing variables, space, and time could be further tested.
- Analyse dissimilarities separately for each of the variables of interest, and then combine. The decision of how to combine them, however, lies with the user. Particular care must also be taken, as combining the independent variables this way may result in analogues at different time lags.

We recommend that users also perform sensitivity analyses to determine how scaling of the variables affects dissimilarity values for different analogous sites.

Third, other measures of dissimilarity exist. The CCAFS and Hallegatte indices are relatively easy to implement and are robust measures of environmental distance. However, other approaches can also be used and applied in our R package.

Fourth, the user must determine to what extent sites need to be similar to be considered ‘analogous’, that is, what the threshold is. This is a subjective process. We provide two options for thresholding, based on defined and relative (dis)similarity:

- *Defined*: A user defines the threshold according to a range of values, selecting those areas that have values within that range.
- *Relative*: A user selects areas that are relatively analogous to their reference site—that is, sites within the first X% of values in the probability distribution of dissimilarity values. This allows selection of priority sites and may prove especially useful when applied to a limited geographic domain.

Both methods have their own pros and cons. ‘Defined dissimilarity thresholding’ requires the user to establish a dissimilarity value under which sites are considered to be ‘analogues’. This can be difficult because it involves a scale that the user may be unfamiliar with, but has the advantage of being highly accurate when well calibrated (i.e. if the threshold is well set, rates of false negatives and false positives would be very low).

‘Relative thresholding’ avoids the difficulty of defining a dissimilarity value threshold but may fail at selecting analogues if the reference site is unique (i.e. if there are few or no analogues). Other methods of thresholding, mostly case-specific, need to be explored.

Fifth, quantification of uncertainties needs to be further and more deeply explored, as described briefly in section 3.5.

3.7 How to get the analogues tool

We created a Google code project as the best means for codeveloping the tool from various parts of the world at the same time. The source code and the latest compiled version of the package are available at <http://code.google.com/p/ccafs-analogues/>, together with tutorials and sample climate data.

3.8 Collaborators

The analogues R package is currently maintained by Julián Ramírez-Villegas, a researcher at the International Center for Tropical Agriculture (CIAT) and PhD candidate at the School of Earth and Environment at the University of Leeds, UK.

The methodology and broad application concept was jointly developed by the Walker Institute at the University of Reading (UK), CIAT, and the Climate Impacts Group at the University of Leeds, with the support and funding of CCAFS.

In particular, Professor Nigel Arnell, Dr Josh Hooker, and Dr Tom Osborne, all from the Walker Institute, developed the original idea and the coding foundation for core functions within the package. Johannes Signer, formerly a visiting researcher at CIAT in Cali, Colombia, helped develop the initial version of the analogues R package. Eike Luedeling, from the World Agroforestry Centre, an R developer, also played an important role in the tool's evolution by providing ideas and developing his own workaround for calculating analogues at a daily time step. Ernesto Giron, a GIS developer, was largely responsible for expertly implementing the package on the ArcGIS Server as a web-based interface.

3.9 Future plans and developments

The tool will eventually be able to identify temporal analogues as well as spatial ones—that is, users will be able to identify contemporary analogues for historical climates.

The climate analogues team is also very keen to collaborate with other researchers in the field who wish to add further functions to the current analogues tool; improve the support documentation, for example, through tutorials or case studies; or explore the possible applications of the tool, including agronomic validation of results.

4. Case study: finding analogues for Lawra-Jirapa, Ghana

The climate analogues tool can be used to detect analogues across the globe or just within specific locations or regions. To illustrate this process and the broad range of approaches possible with this tool, we present a case study, in which we searched for climate analogues to Lawra-Jirapa, Ghana (LJG), focusing mostly on Africa and South Asia. We used monthly data for total rainfall, mean temperatures, and diurnal temperature ranges (DTRs), together with a set of 19 bioclimatic variables (Busby 1991). We looked for analogues in both the present day and the future (for the 2030s), based on 24 different GCMs under the SRES A1B emissions scenario, as developed by the Intergovernmental Panel on Climate Change (IPCC) (IPCC 2000, Ramírez and Jarvis 2010). For the monthly analyses, we assumed a growing season from April to October. Below, the technical process is described in more detail.

4.1 Measuring dissimilarities, using different variables

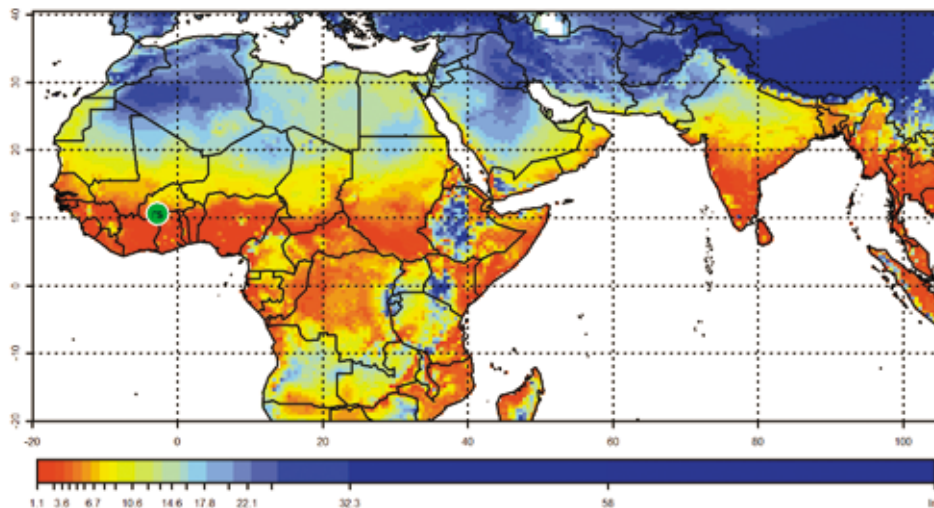
We analysed the dissimilarity between LJG and each (0.5 degree) pixel, for temperature (weighted by DTR) (figure 1A), rainfall (figure 1B), and the 19 bioclimatic variables used together (figure 1C). We then overlaid the areas where the three independent analyses indicated that a pixel was within the 25% lowest values of dissimilarity (figure 1D).

This step demonstrated the importance of including more than one variable in a given analysis. For LJG, when considering temperature alone, the closest analogues were located towards the southern part of West Africa; in some humid areas of central Africa; and in the lowlands in southern India, where the seasonal temperature patterns matched those of LJG (figure 1A).

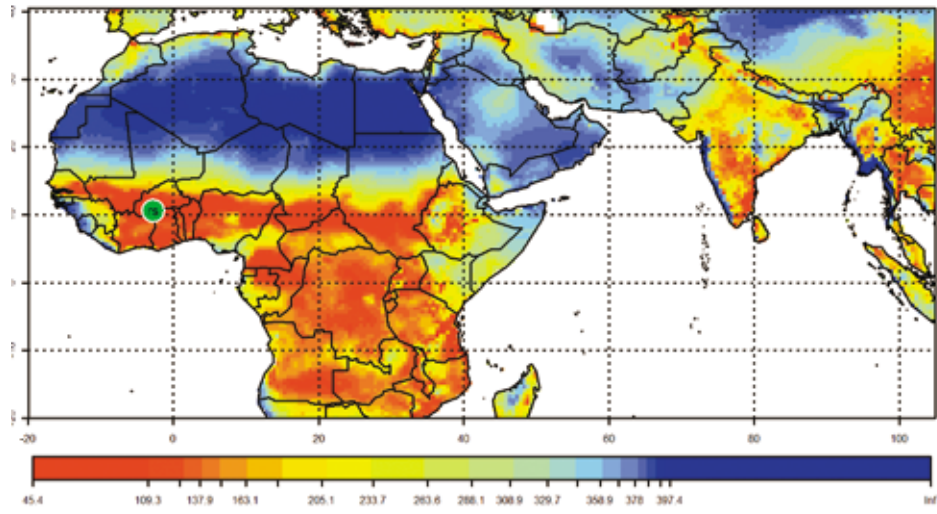
In contrast, for the precipitation variable alone, the most similar areas were located near LJG, extending towards the Sahel; and in very small pockets of southwestern India, Central Africa, and East Africa (figure 1B). In contrast to the temperature results, the precipitation variable showed much of southern India to be highly dissimilar (figures 1A and 1B). The map for only the bioclimatic variables apparently integrated the information shown in figures 1A and 1B, probably because these variables already combined rainfall and temperature. This measure showed that the most similar sites surrounded LJG (figure 1C), expanding towards southern West Africa (figure 1A) and eastern Africa (figure 1B).

Figure 1. Dissimilarity calculations as averages of 24 global climate models (GCMs) for LJG's future climate and all pixels' current climates for (A) monthly temperature, weighted by diurnal temperature range, (B) monthly rainfall, (C) the 19 bioclimatic indices (standardized), and (D) the overlay of areas in which any combination of A, B, and C occurred. LJG refers to Lawra-Jirapa, Ghana, and is the green point on the maps. Color classification is based on quantiles spaced every 5%.

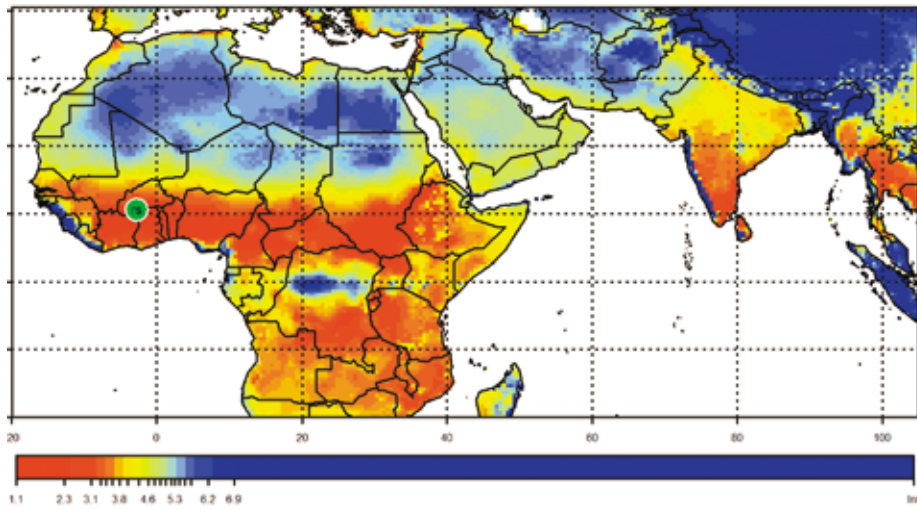
(A)



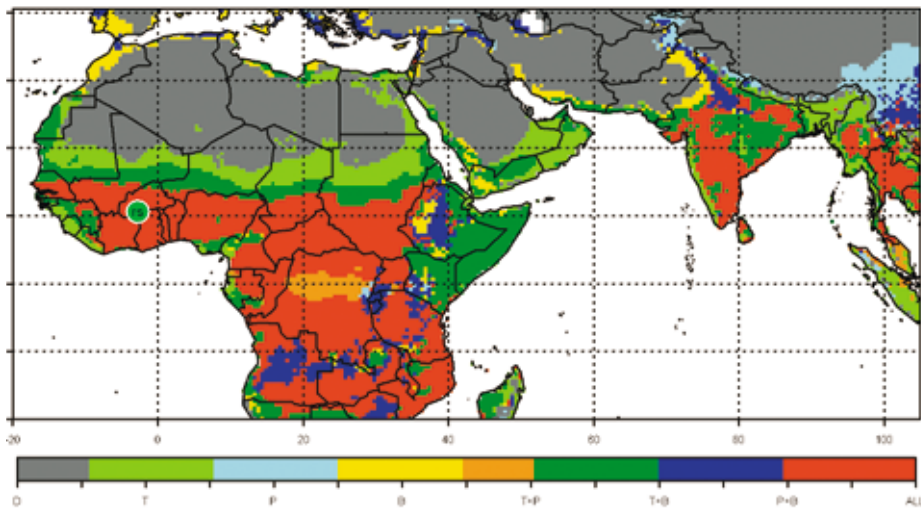
(B)



(C)



(D)



Some areas appeared more governed by differences in rainfall, whereas others were more governed by differences in temperatures. The Sahara Desert was not highly similar in any of the cases. Agreement between all the different variables (figure 1D) indicated that more than half of the area within the domain of analysis was considered to be ‘analogous’ (at 25% probability for that geographic domain) for the three variables (red areas in figure 1D). In figure 1D, areas in green are more strongly influenced by the seasonal rainfall pattern, whereas the areas in blue are more strongly influenced by temperatures.

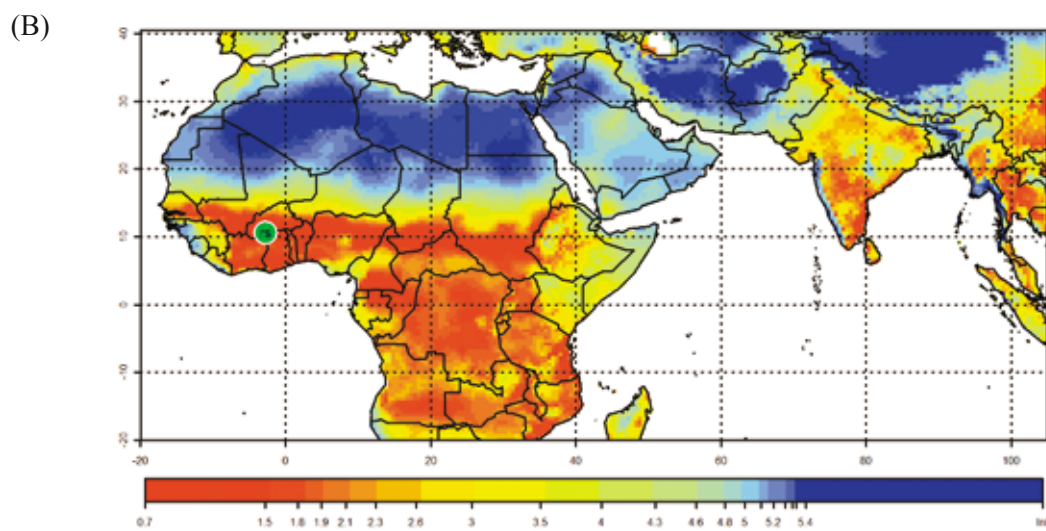
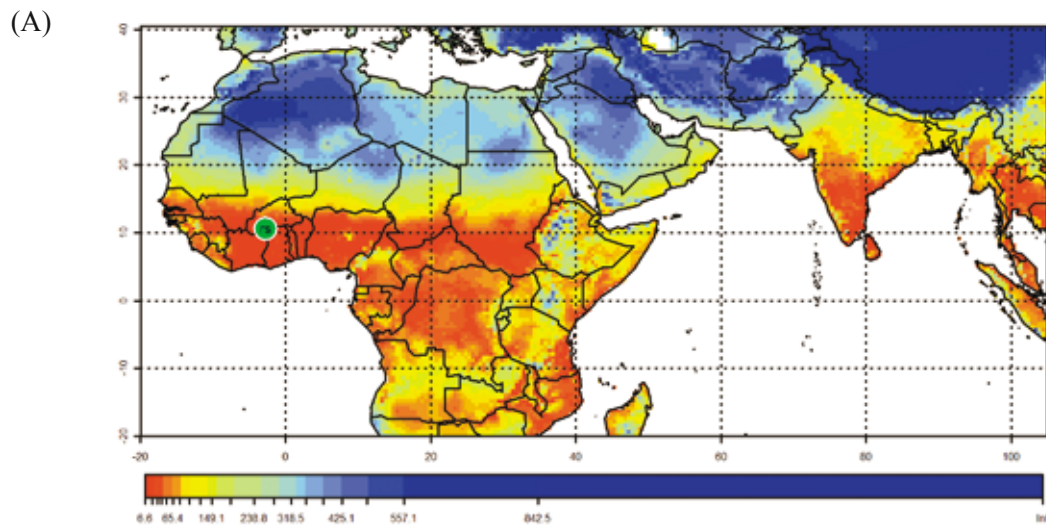
4.2 Combining variable-specific dissimilarities

We calculated temperature-rainfall dissimilarity in three ways: (1) multiplying dissimilarities (MD), for which we multiplied the individual dissimilarities for monthly temperatures (weighted by monthly DTRs) and monthly rainfall as calculated separately (with standardization); (2) standardization (SZ), for which we calculated temperature-rainfall dissimilarities with standardization; and (3) non-standardization (NSZ), for which we calculated temperature-rainfall dissimilarity without standardization.

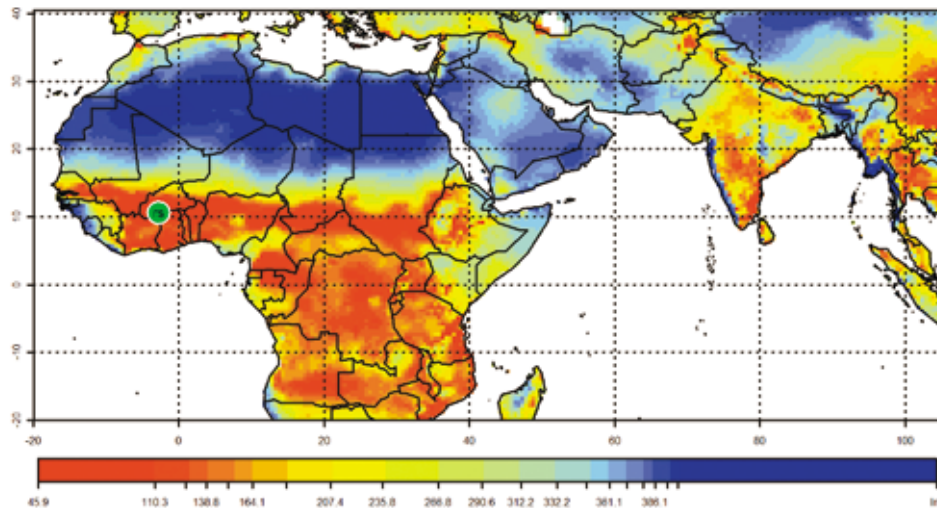
Individual rainfall and temperature dissimilarities are shown in figures 1A and 1B, respectively, regardless of whether they were standardized or not. However, combining the results of the two variables in different ways led to substantially different outcomes. All the combinations tried for LIG identified the areas surrounding it, extending towards the eastern Sahel, as highly analogous (areas in orange and red, figure 2). That said, the three approaches yielded considerable disagreement in other areas. Such variation was probably due to scaling in the two variables and to the dominance of one variable over the other, depending on season and geographic area. The NSZ was much more dominated by rainfall dissimilarities, as its map (figure 2C) resembles figure 1B much more closely than it resembles either figure 1A or 1C.

In contrast, SZ apparently underestimated the effect of rainfall by showing areas of central Africa as having relatively low dissimilarities. This probably indicated that weighting was needed to compensate for the scaling effects due to normalization in precipitation seasonality (figure 2B).

Figure 2. Dissimilarity calculations as averages of 24 global climate models (GCMs) for LJG's future climate and all pixels' current climates for (A) the product between monthly temperature dissimilarity (weighted by diurnal temperature ranges [DTRs]) and monthly precipitation dissimilarity; (B) standardized temperature (weighted by DTR) and rainfall dissimilarity combined in Equation 4; and (C) non-standardized temperature (weighted by DTR) and rainfall dissimilarity combined in Equation 4. LJG refers to Lawra-Jirapa, Ghana, and is the green point on the maps.



(C)



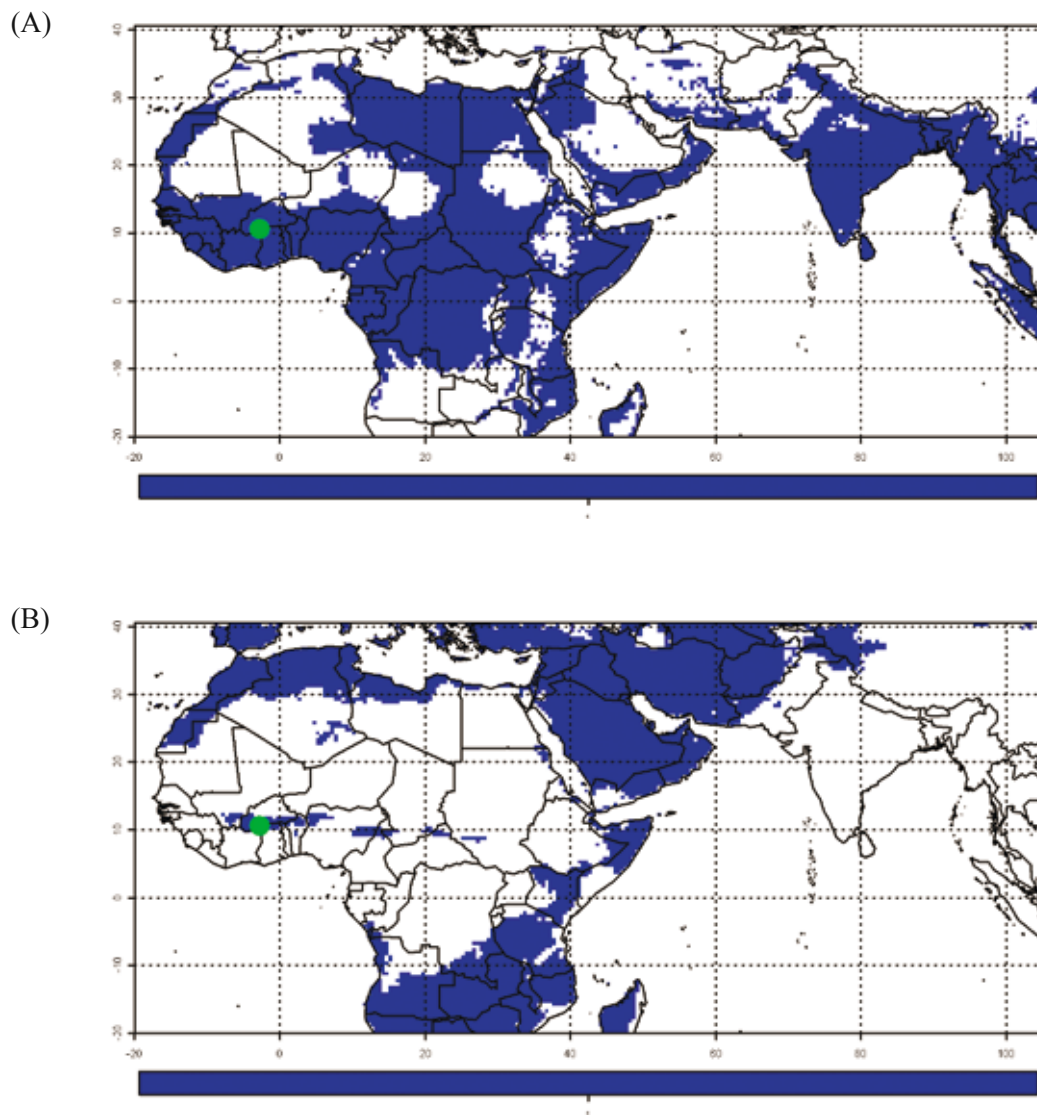
4.3 Testing sensitivities to the importance of different variables

We used the standardized rainfall-temperature CCAFS dissimilarity measure to test the sensitivity of the points within our study area to the weighting in precipitation. We used monthly totals and means, and averaged the dissimilarity results over 24 GCMs. Temperature was weighted by DTR, except for one run where temperature was weighted by zero to calculate the dissimilarity for precipitation only. We further thresholded each result to the 25% probability points and assigned 1 to each pixel that was within the 25% probability threshold and 0 to each pixel that was not. Results of all runs were then summed up into a single grid, in which the higher the value, the more insensitive the pixel was to different precipitation weightings (blue in figure 3).

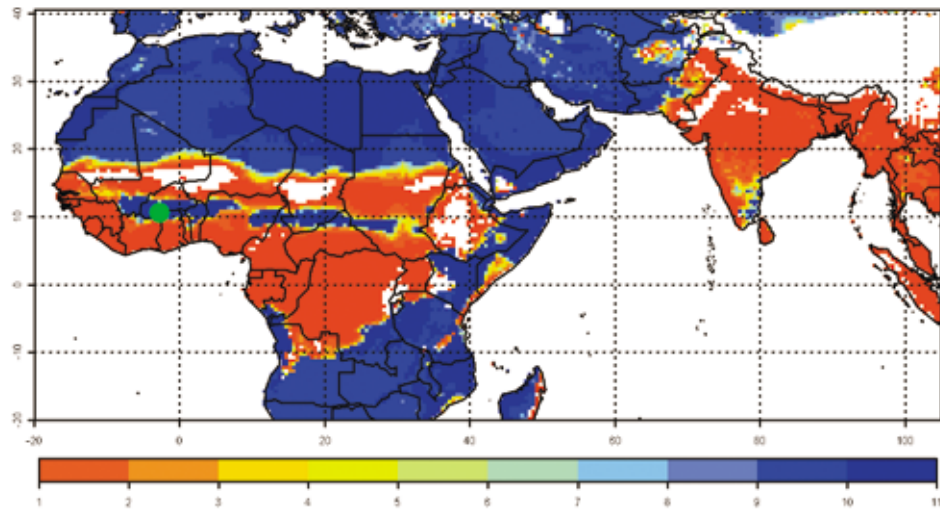
Figure 3 shows the results for the temperature-only dissimilarity (A); precipitation-only dissimilarity (B); precipitation weighted from 0 to 10, increasing by 1 (C); and precipitation weighted from 0 to 100, increasing by 5 (D). Notably, for most of the 25% best analogous regions, using a weighting of precipitation up to 10, either almost all runs agreed (figure 3C, blue) or only one run included a particular region in the best 25% analogues (figure 3C, red). The red areas in figures 3A and 3B correspond to the analogues with precipitation weighted 0. Large areas of central Africa and India were therefore only analogous to LJG when precipitation was not included in the dissimilarity calculation. For precipitation weights from 0 to 100 (figure 3D), results were still similar to figure 3C, except for parts of the Sahara where the weighting of precipitation apparently crossed a critical threshold, after which this area was no longer included in the best 25% analogues.

Hence, for LJJ, including precipitation was important but results were fairly insensitive to precipitation weightings of up to about 8 and, for most areas except the Sahara, even to much higher precipitation weightings.

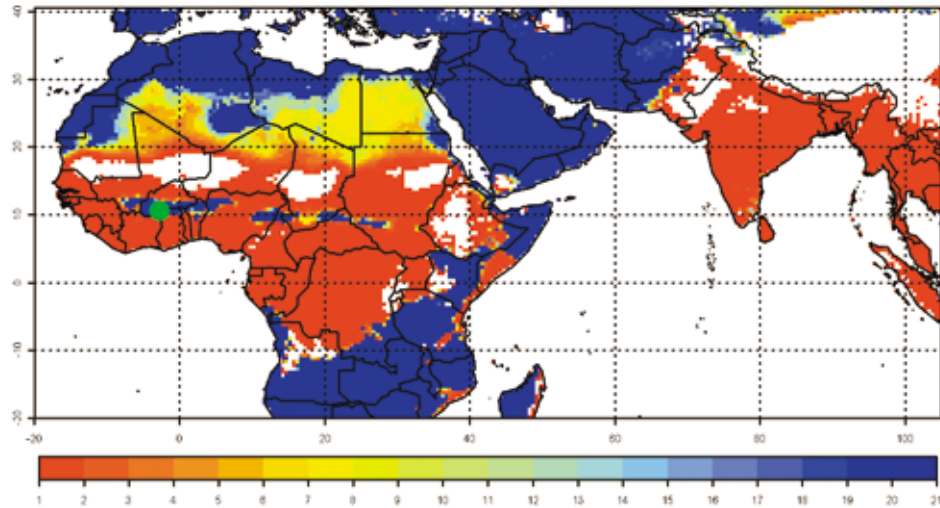
Figure 3. Dissimilarity calculations as averages of 24 global climate models (GCMs), counting the pixels that agree for the best 25% analogues for different precipitation weightings. Shown are the best 25% analogues with (A) temperature weighted by diurnal temperature range (DTR) and precipitation by 0; (B) temperature weighted by 0 and precipitation by 1; (C) temperature weighted by DTR and precipitation by 0 to 10, increasing by 1; and (D) temperature weighted by DTR and precipitation by 0 to 100, increasing by 5. The site is Lawra-Jirapa, Ghana, and is the green point on the maps.



(C)



(D)



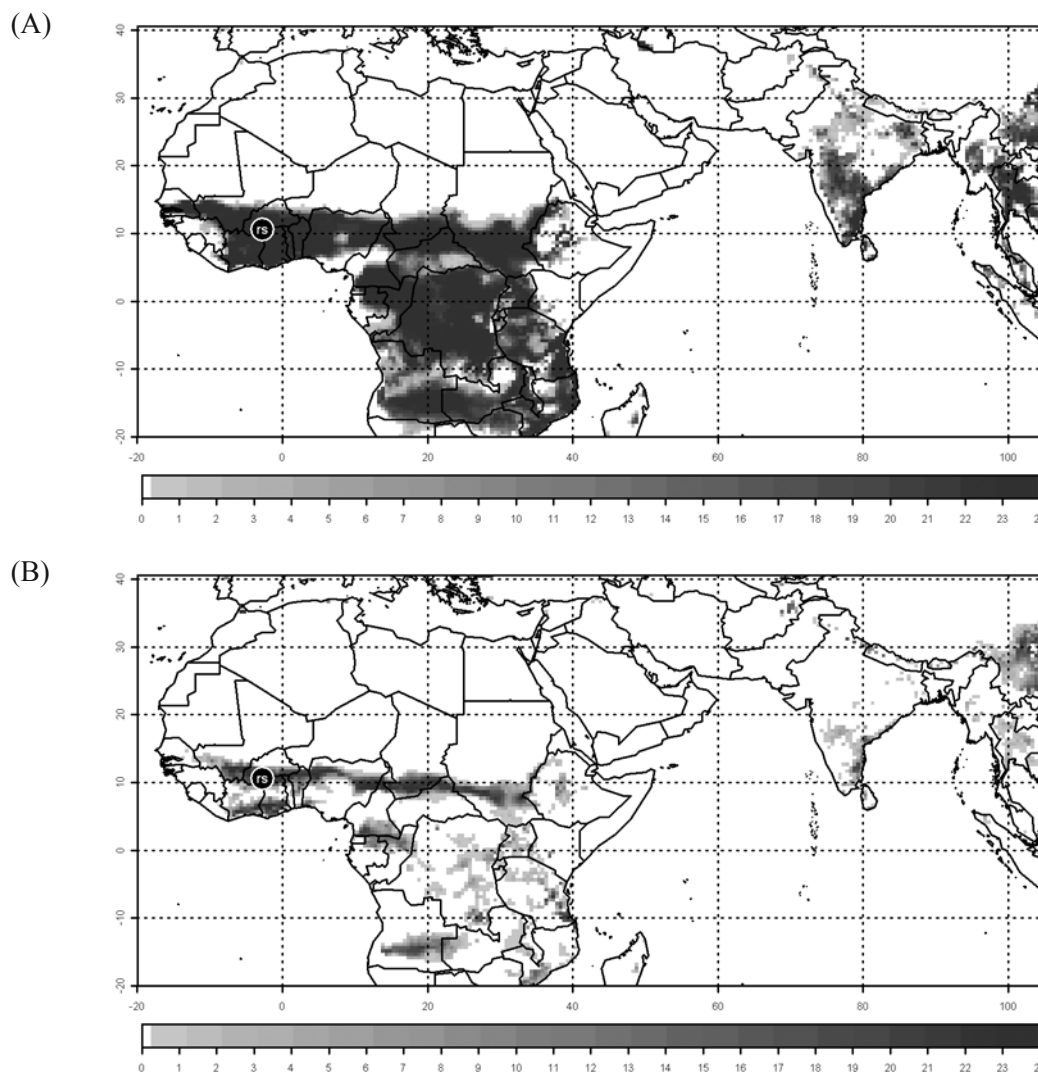
4.4 Quantifying dissimilarity uncertainty

We calculated Hallegatte's dissimilarity measure (using only conditions 1 and 2) for LJG, against all pixels in the study area for the 24 GCMs. We then quantified the uncertainties by counting the number of times each pixel was selected as an 'analogue' site of LJG, with $a = 0.15$ and $b = 0.30$ for precipitation (see Equations 5 and 6). We did not account for any dissimilarity in temperatures, as aggregating the three conditions proved too restrictive. We then compared this result with the uncertainty given by the CCAFS dissimilarity measure (temperature-rainfall standardized dissimilarity), thresholded by 10% probability of occurrence (i.e. number of times a given pixel is within that range).

For each Hallegatte GCM-specific result and for the mean of all GCMs (MEAN), we extracted 1000 random locations over the areas flagged as ‘analogous’ (TRUE), and ‘non-analogous’ (FALSE), and then extracted the corresponding CCAFS dissimilarity values at those locations for that particular GCM. We then calculated a boxplot in which boxes were permutations of the two Hallegatte conditions (i.e. TRUE or FALSE) and the GCMs.

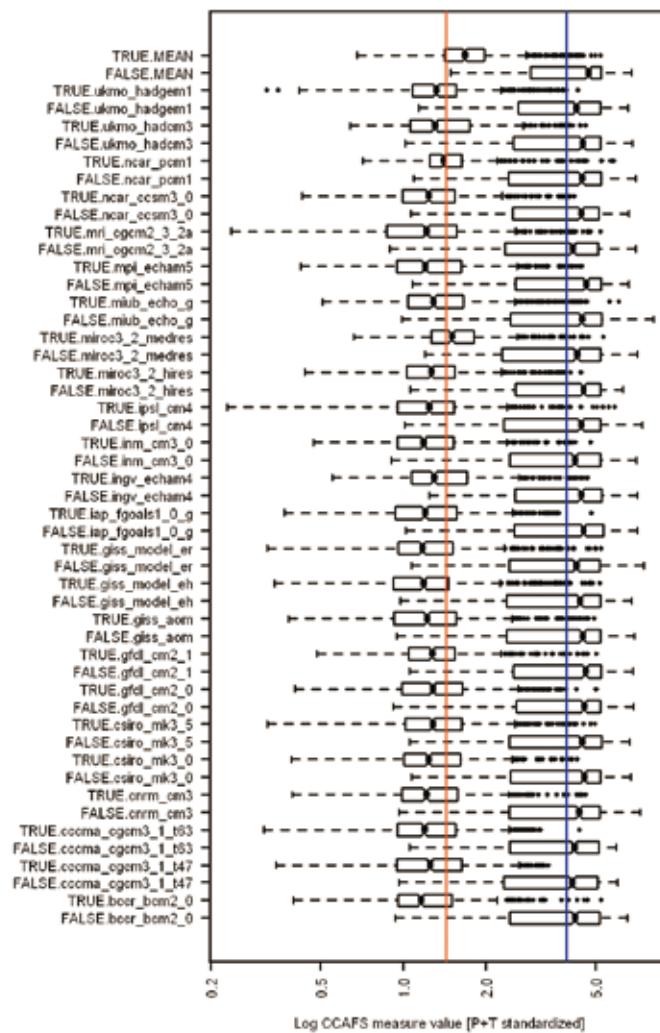
We found considerable agreement between the two measures. In both cases, areas found as climatically ‘similar’ were flagged by most GCMs (>20), signifying strong agreement about analogous areas (figure 4).

Figure 4. Uncertainty as measured by the number of global climate models (GCMs) that (A) flag each pixel as being in the closest 10% to the site’s (Lawra-Jirapa, Ghana) CCAFS rainfall-temperature dissimilarity measure, and (B) flag a site as being an ‘analogue’, according to the Hallegatte measure, with $a = 0.15$ and $b = 0.30$, only for precipitation. The site is shown by a dark ring on the maps.



In addition, the two measures showed considerable agreement, with more than 90% of the areas that the Hallegatte method deemed as analogous (figure 4B) also identified by the 10% closest thresholded CCAFS measure as analogous (figure 4A). In terms of the boxplot (figure 5), the mean value of the TRUE areas was, overall and for each individual GCM, lower than the average of the FALSE areas, confirming agreement between the methods.

Figure 5. Comparing the Hallegatte with the CCAFS method. Each boxplot shows the distribution of 1000 random CCAFS dissimilarity values over areas considered as ‘analogous’ (TRUE) or ‘non-analogous’ (FALSE) to Lawra-Jirapa, Ghana, for each global climate model (GCM) and for the average of all GCMs-MEAN (y axis). Notches in the boxes indicate the median of each distribution, boxes the interquartile range, and whiskers the 5% and 95% of the distribution. The blue line is the average of the CCAFS measure of all GCMs value over non-analogous (FALSE) areas (as flagged by all GCMs), and the red line is the average of the CCAFS measure over analogous (TRUE) areas (as flagged by at least 1 GCM).



For all sites in the globe, we found that the Hallegatte measure was much more restrictive than the CCAFS measure thresholded by the 10% closest pixels (figure 4), which could be overcome by restricting the probability to 5% or less. In the boxplot, the CCAFS dissimilarity measure also showed a larger number of higher values, as well as a higher mean value in the FALSE areas than in the TRUE areas, compared with the Hallegatte measure. These held true, regardless of the GCM used (figure 4).

Statistically significant differences were also found in the two variables for the average of all GCM-specific results, indicating that the dissimilarity measures were apparently consistent, irrespective of the uncertainties arising from the usage of different climate models.

4.5 Beyond the tool: an agronomic perspective

The analogues tool's limitations and appropriate applications should be recognized. The dissimilarity measures developed and used here must be translated into something agriculturally meaningful, relating either to the use of certain practices by farmers or to actual crop yields. In other words, we need to look *through the eyes of the crop*.

For agricultural systems, we need to quantify dissimilarities at a scale that can be interpreted in the context of agricultural production. Other agriculturally relevant measures of climate conditions such as total growing days, evapotranspiration, relative humidity, and associated soil conditions need to be included when analysing dissimilarities from an agronomic perspective. On-the-ground policies, and agronomic and infrastructural specificities (e.g. existence of irrigation systems) may also affect a production system's suitability and sustainability. Unfortunately, unless the user has access to these data, which are often unavailable or blocked by restricted access, such layering of variables is impossible. Consequently, the results of the analogues tool should be used with care.

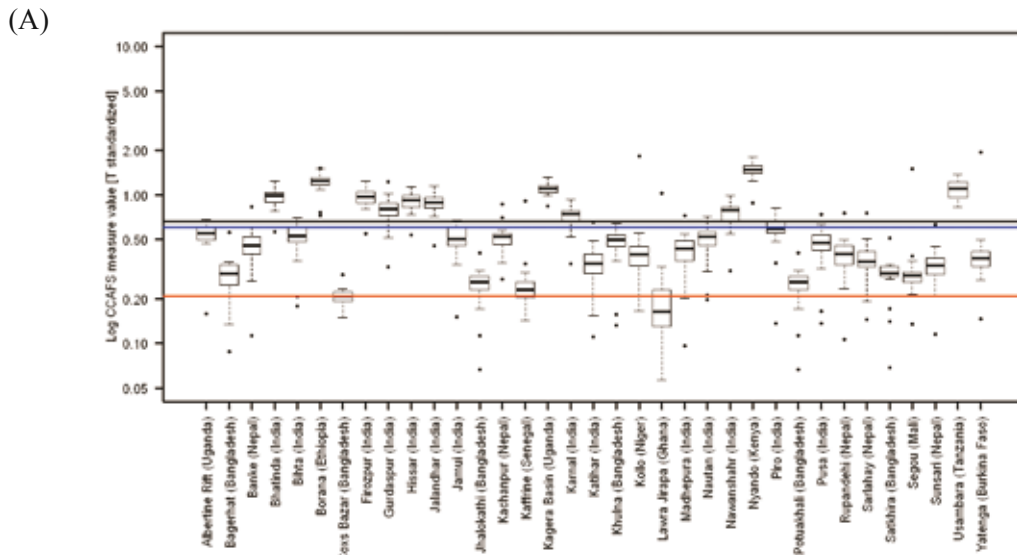
5. Final analysis: results and identified analogues

5.1 Site-specific comparison of dissimilarities

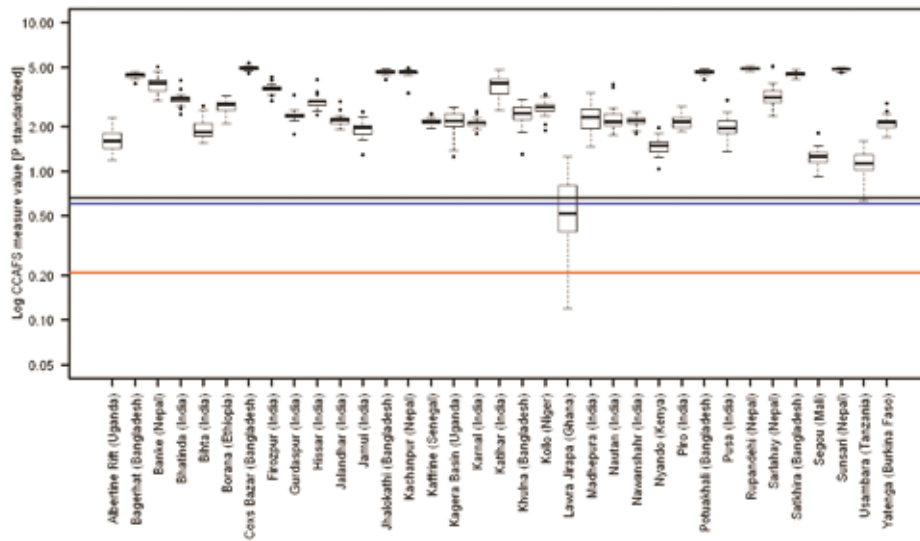
To illustrate the use of identified analogues for informing further ground research, we searched among CCAFS benchmark sites (see <http://www.amkn.org>) to find the best present-day climate analogues to LJJ in 2030. We used the CCAFS dissimilarity measure, based on standardized monthly temperatures, rainfall, and temperatures and rainfall together. We then plotted the dissimilarities between LJJ's future climate and the current climates of all other sites, using a boxplot (figure 6).

When the two variables were used, the present-day site most similar to LJG’s future was LJG itself, but it was also the most uncertain (i.e. individual GCM values were the most spread out) (figure 6C). Sites in India (Bihta in Bihar and Bhatinda in Punjab) and sub Saharan Africa (Ségou in Mali, Usambara in Tanzania, the Albertine Rift in Uganda, and Borana in Ethiopia) were also highly analogous (i.e. with low dissimilarity values). Using the temperature variable alone, dissimilarities were overall lower but displayed a much higher degree of uncertainty (all boxes were more spread out) (figure 6A). The most closely analogous sites were Cox’s Bazar (Bangladesh), LJG itself, Jhalokathi (Bangladesh), Patuakhali (Bangladesh), Kaffrine (Senegal), Bagerhat (Bangladesh), Kanchanpur (Nepal)—that is, almost all sites were in South Asia.

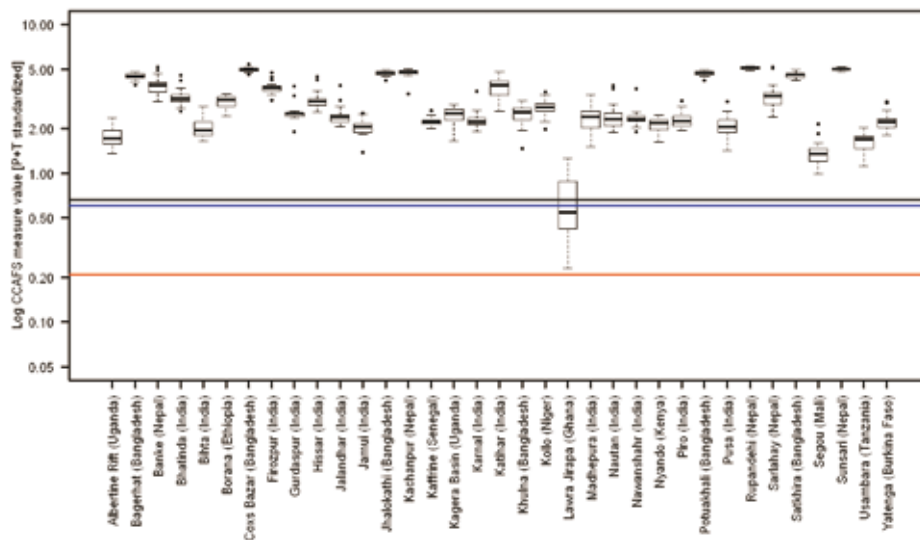
Figure 6. Comparing the future conditions of Lawra-Jirapa, Ghana with the present-day conditions in each of the other CCAFS benchmark sites, using the CCAFS measure and (A) standardized monthly mean temperature (weighted by the diurnal temperature range [DTR]); (B) standardized monthly total rainfall; and (C) standardized temperature-rainfall, as in Equation 4. Boxes are the distribution of the 24 global climate models (GCMs) used in the analyses. Black lines in the boxes indicate the median, boxes the interquartile range, and whiskers the 5% and 95% of the range of values. Red, blue, and black lines are, respectively, the averages of temperature, rainfall, and temperature-rainfall dissimilarities between LJG’s future and present in the site itself.



(B)



(C)



Rainfall dissimilarities seemed to dominate the overall dissimilarity measure, as the two produced similar results (figure 6B). Looking at that variable alone, LJG itself was again the closest analogue at present to LJG's future conditions, but with considerable uncertainty. Usambara (Tanzania), Borana (Ethiopia), Ségou (Mali), and the Albertine Rift (Uganda) were also close analogues, but interestingly, all sites in southern Asia were found to have high dissimilarity values (figure 6B).

Table 2 summarizes these results in numerical terms, with RS (%) being the relative similarity (calculated as 1 minus the site's dissimilarity divided by the maximum dissimilarity among all sites). Hence, the higher the RS value, the more similar the site is to LJG. As this relative similarity value is merely a normalization of the dissimilarity values, its values do not reflect, in absolute terms, how analogous the climates are. These values cannot be compared directly with dissimilarity values of non-benchmark sites, or with dissimilarity values calculated by using other model configurations. Moreover, as CCAFS sites were intentionally chosen to be distinct from one another to capture the range of agro-ecological zones and climates, the best analogue may often or always be the site itself; this characteristic is not representative of other point-based searches.

As discussed in section 3.5 on uncertainty, the coefficient of variation (CV, %) is likely to be biased when the SD is biased towards low values of dissimilarity (e.g. see the value of CV for LJG). However, given the agreement between the range and the SD, biases towards high SD values are unlikely in this case. Uncertainty is tricky in any case, and one could calculate the likelihood of the mean value, for example, or use the mode instead of the mean. In this case, for simplicity's sake, we chose the CV.

5.2 Point-based analysis: mapping dissimilarities across many sites

We extracted the corresponding values for all CCAFS benchmark sites (35 sites in total) and computed dissimilarities, using the CCAFS method with temperature and rainfall together, and with standardization turned on. We performed the analysis individually for each GCM and then averaged the results to reduce the number of outputs. We therefore had a matrix of dissimilarities consisting of s sites by s sites. We then used the approach, described by Paul Butler at <http://paulbutler.org/archives/visualizing-facebook-friends/>, to build a map of analogues.

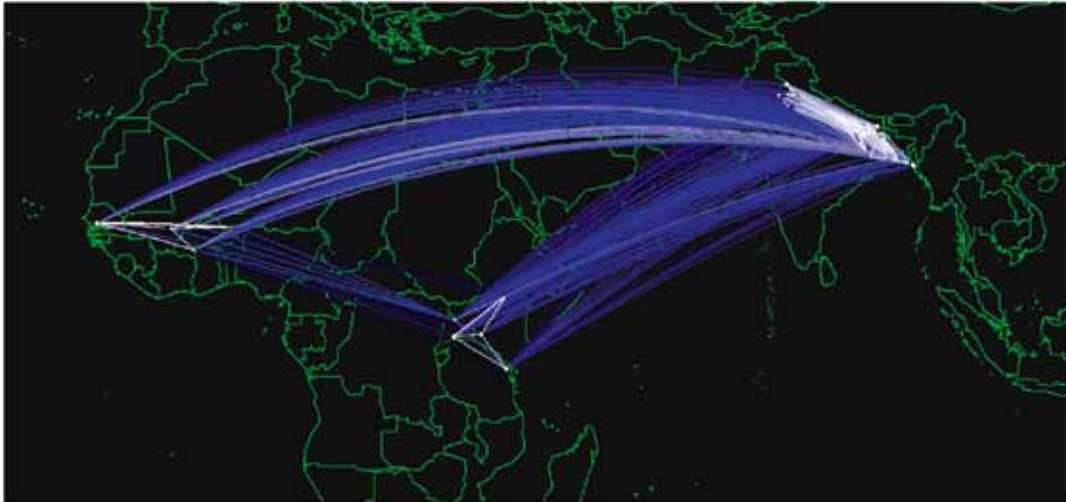
The resulting 'analogues map' showed strong similarities among sites within the same region (i.e. within South Asia, West Africa, or East Africa), thus highlighting the importance of effective within-region coordination for adapting agriculture to future climates. The map also showed strong connectivity across regions, that is, between some sites in West Africa (particularly in Mali and Ghana) and other sites in South Asia (figure 7).

Table 2. The best analogue sites to Lawra-Jirapa, Ghana (LJG), according to the CCAFS dissimilarity index, when considering mean diurnal temperature and monthly rainfall together (resulting in LJG itself, followed by five other sites). Also shown is the best analogue site when considering only mean diurnal temperature (Cox's Bazar, Bangladesh). Presented for each analogue are geo coordinates, relative similarity (RS) values for the variables together and independently, and uncertainty measures (standard deviation or SD, and coefficient of variation or CV). Values in boldface are the relative similarity values for the 'best' analogues, according to the variable(s) used.

Site	Geo coordinates	RS (%), both temp. and precip.	SD	CV	RS (%), temp. only	CV	RS (%), precip. only	CV
Lawra-Jirapa, Ghana	10.596, -2.768	87.0	0.31	46.4	85.9	88.8	87.7	49.3
Ségou, Mali	-5.762, 13.368	73.1	0.25	17.9	77.7	76.9	74.4	14.1
Usambara, Tanzania	4.820, 38.359	68.1	0.20	12.4	25.7	13.9	76.6	19.4
Albertine Rift, Uganda	31.501, 1.49	65.3	0.25	14.2	63.6	19.1	66.8	17.0
Bihta, India	25.541, 84.902	59.9	0.33	16.1	65.2	25.1	60.9	16.2
Jamui, India	24.959, 86.202	58.6	0.25	12.1	66.6	26.7	60.7	13.1
Cox's Bazar, Bangladesh	21.619, 91.927	2.7	0.15	3.0	86.0	12.7	0.0	3.1

a. Relative similarity is calculated as a percentage, and is equivalent to 1, minus the dissimilarity value divided by the maximum dissimilarity value in the 35 sites searched.

Figure 7. Analogues map. CCAFS benchmark sites dissimilarities: the whiter the lines, the stronger the future-current analogues are between sites. The original idea for this map came from Paul Butler, an intern with Facebook's data infrastructure engineering team (http://www.facebook.com/note.php?note_id=469716398919).



Similarities between East and West Africa were weaker than individual similarities between these regions and southern Asia (figure 7). The region with the most distinct climate thus appeared to be East Africa.

5.3 Case study lessons: future research and adaptation pathways

Analogues can be used to suggest various worthwhile policy interventions, development initiatives, and opportunities for participatory fieldwork or research. In this paper, we used the CCAFS analogues tool to locate present-day climate analogues to LJG's future.

Because we searched solely within CCAFS benchmark sites, these definitely do not reflect the closest existing analogues worldwide. Moreover, we do not have the on-the-ground data needed to determine the threshold level at which CCAFS dissimilarity values produce very close analogues. Hence, we simply placed our results on a relative index (as shown in table 2), and examined the best five analogues. More detailed assessments, or the inclusion of more data points, would therefore be needed to draw truly valid policy or research recommendations.

With these caveats in mind, we now use the simplified dataset of CCAFS benchmark sites to draw some example lessons or hypotheses from our case study. For each benchmark site, we have data on production systems and farmers' adaptive strategies. In table 3, we compare the production systems in Lawra-Jirapa, Ghana (LJG) with those of its best analogue sites, as determined by the CCAFS dissimilarity value.

Lawra-Jirapa, Ghana, located in the northwestern Guinea savanna zone, currently produces maize, rice, groundnuts, sorghum, cowpea, soybean, and yam. Its farming systems are threatened by high climatic variability, particularly unpredictable rainfall. Future projections show higher temperatures and greater overall precipitation, but also water stress partly because of increasing population pressures. Consequently, negative yield impacts are anticipated for rainfed maize, rice, and wheat, although the more drought-resistant millet may remain largely unaffected (Roudier et al. 2011). Below, we discuss adaptive strategies used in LJG's analogues, as potential lessons to be learned:

- *Lessons on maintaining current cultivation.* Crop model projections show declining future yields for maize, rice, and wheat in LJG, but as the site's analogues all currently grow at least one of these three crops, local knowledge may be available on how best to manage these crops in the future climate to continue their cultivation. Exchange of farmer knowledge (e.g. through field visits) and better inventories of traditional strategies may be useful here. In those cases where rice is not grown (e.g. Ségou and Cox's Bazar), research may examine whether rice was ever grown in the past and, if so, how production systems adjusted to its disappearance.
- *Crops that may not be viable in 2030.* Some crops currently grown in LJG are not cultivated in its analogues, for example, yam, indicating that climatic factors may inhibit future production of these foods, or that a wider range of analogue sites (in addition to the CCAFS benchmark sites) are needed. Of course, these conclusions must be considered jointly with research on other biophysical (e.g. soils), politico-economic (e.g. subsidies and infrastructure), or cultural (e.g. food preferences) variables not captured in this study. For instance, differences may be due to market and demand issues, rather than because of unsuitable climate. Therefore, to minimize the effects of non-climatic variables, examining analogues in geographically closer sites (where food habits and economies are similar) may be especially useful.
- *Possible substitution crops.* Crops that are grown in many or all of the analogues, but not in the reference site, may be the most suitable for diversifying. For instance, aquaculture (fisheries for both Indian sites) or wheat (for India and Bangladesh) may constitute an adaptive pathway. However, because site-specific, non-climatic obstacles to adoption may exist, complementary research should be conducted.
- *Validating crop modelling projections.* If, for instance, crop models indicate that wheat will not be a suitable crop in a certain location in 2030, yet wheat production systems exist

Table 3. Production systems in Lawra-Jirapa, Ghana (LJG), and in sites determined as analogous by the CCAFS analogues tool.

Benchmark site	Current agricultural production		Climate and water projections	
LJG	Maize, rice, groundnuts, sorghum, cowpea, soybean, yam		<ul style="list-style-type: none"> • Avg max temp increases by 2.4°C by 2050; avg min temp by 2.8°C • Reduced groundwater • Increase of 98 mm in total annual rainfall (but substantial uncertainty) 	
Analogue benchmark site	<i>Similarities:</i> Crops in both LJG and analogue	<i>Differences:</i> Crops in LJG, but not in analogue	<i>Differences:</i> Crops in analogue, but not in LJG	Adaptive strategies (lessons to learn?)
Ségou, Mali	Maize, sorghum, cowpea, groundnut	Rice, soybean, yam	Millet, pearl millet, groundnut, sesame	Extended migrating households in Côte d'Ivoire and within Mali (diversification spreads risk)
Usambara, Tanzania	(No data)	(No data)	(No data)	Mixed crop-livestock systems; intensive farming at high altitudes
Albertine Rift, Uganda	Maize, rice	Groundnuts, sorghum, cowpea, soybean, yam	None?	Adaptive crop varieties, climate forecasts, soil and water conservation
Bihta, India	Maize, pulses (e.g. cowpea and groundnut), rice	Sorghum, soybean, yam	Fisheries, wheat	Irrigation
Jamui, India	Maize, pulses (e.g. cowpea and groundnut), rice	Sorghum, soybean, yam	Fisheries, wheat	Irrigation
Cox's Bazar, Bangladesh	Maize, oilseeds (e.g. soybean), pulses (e.g. cowpea and groundnut)	Rice, sorghum, yam, sugarcane,	Wheat, mustard, onions, jute	Zero tillage

within the range of climate analogues, then either the model itself may require tweaking, or lessons are to be learned from the analogue sites on what adaptive practices permitted the crop's continued cultivation. Thus, comparisons of computational outputs and on-the-ground realities of analogue sites may better inform models and reduce uncertainties.

6. CCAFS projects: from model to field

To further validate the analogues tool and better target our research, CCAFS plans to apply the analogues methodology to several projects in 2011–2012, as reviewed below.

6.1 Farmer-to-farmer exchanges

The analogues tool may be applied to social science fieldwork, for example, through connecting farmers to their possible climate futures via farm visits. The CCAFS Farms of the Future project, planned for 2011–2012 in all three focus regions, will strive to do just this. This novel approach of farmers visiting farmers in spatially analogous sites will integrate participatory learning principles to promote knowledge sharing between farming communities. Hence, farmers will be able to envision how their site-specific agricultural future might look, and potentially put into practice the new adaptive strategies they had witnessed firsthand.

The approach improves the state of research knowledge, while building capacity on the ground. Guided exchanges between farmers permit the participatory diagnosis of capacities and needs, thus aiding in the design of community-appropriate adaptation strategies that actually fit farmers' needs, cultures, and resources. For researchers, such exchanges can improve understanding of local practices and available tools for enabling change, and provide opportunities for examining whether successful adaptation options in one place are transferrable to its future-climate analogue site. In that vein, research will seek to identify possible social, cultural, institutional, or economic obstacles to adaptive change. Moreover, our researchers will then be able to validate what is currently a computational model and thus ground mathematical projections in qualitative real world assessments.

6.2 Field trial sites

For decades, cultivar testing in trial sites has proven to be an efficient and valuable methodology for varietal improvement and targeted dissemination. However, poor data management and failure to make information publicly accessible precluded the systematic evaluation of trialled materials and adaptation options across multiple geographic locations and terrains. CCAFS is therefore supporting the establishment of a broad database and online repository of compiled

and standardized, multisite, trial data. This online platform will be used in conjunction with the climate analogues tool to better understand how crops respond to (a) different climates, or (b) analogous climates in different national settings. In the latter case, the distinct responses of production systems to similar climates may shed light on other sorts of dissimilarities, that is, in biophysical characteristics or social, economic, and institutional settings.

In South Asia in particular, work is already under way to map analogues among wheat production systems, and then potentially establish field trial sites in these locations to test different varieties and/or interventions.

Conclusions and recommendations

Considerable uncertainty remains, regarding projections of future climates and their resultant impact on farming systems, especially at the local level. The intrinsic adaptive capacity of rural communities is rarely taken into account in the global or regional models on which policy makers often rely. These uncertainties and gaps can prevent researchers, representatives from development agencies, and government officials from being able to assess what future climate change may look like, and how it may shape agriculture in 2050 or 2100. The use of climate analogues for locating future climates today can ground models in field-based realities, significantly enhancing our knowledge of adaptation capacity and supporting the identification of appropriate interventions. Although region-specific tweaks, validation processes, and additional data will be needed to produce more robust results, the CCAFS analogues tool offers an important platform for future research and decision making.

We encourage researchers to contact any of the lead authors for more information on the tool or to propose collaborative research.

References

- Busby JR. 1991. BIOCLIM—a bioclimate analysis and prediction system. *Plant Protection Quarterly* 6:8–9.
- Challinor AJ, Ewert F, Arnold S, Simelton E, Fraser E. 2009. Crops and climate change: progress, trends, and challenges in simulating impacts and informing adaptation. *Journal of Experimental Botany* 60(10):2775–2789.
- Challinor AJ, Wheeler TR. 2008. Crop yield reduction in the tropics under climate change: processes and uncertainties. *Agricultural and Forest Meteorology* 148(3):343–356.

- Hallegatte S, Hourcade J-C, Ambrosi P. 2007. Using climate analogues for assessing climate change economic impacts in urban areas. *Climatic Change* 82(1):47–60.
- Hitz S, Smith J. 2004. Estimating global impacts from climate change. *Global Environmental Change* (Part A) 14(3):201–218.
- IPCC (Intergovernmental Panel on Climate Change). 2000. Special report on emissions scenarios. Geneva, Switzerland: IPCC. (Available at www.ipcc.ch/ipccreports/sres/emission/index.htm)
- IPCC (Intergovernmental Panel on Climate Change). 2007. IPCC fourth assessment report: climate change 2007 (AR4). Geneva, Switzerland: IPCC.
- Jarvis A, Lau C, Cook S, Wollenberg E, Hansen J, Bonilla O, Challinor A. 2011. An integrated adaptation and mitigation framework for developing agricultural research: synergies and trade-offs. *Experimental Agriculture* 47(Special Issue 02):185–203.
- Jarvis A, Ramírez J, Anderson B, Leibing C, Aggarwal P. 2010. Scenarios of climate change within the context of agriculture. In: Reynolds MP, ed. *Climate change and crop production*. Wallingford, Oxfordshire, UK: CAB International. (A CAB eBook available at: <http://www.cabi.org/CABeBooks/default.aspx?site=107&page=45&LoadModule=PDFHier&BookID=521&ActiveTabIndex=0&PageNo=0>)
- Lobell DB, Burke MB. 2008. Why are agricultural impacts of climate change so uncertain? The importance of temperature relative to precipitation. *Environmental Research Letters* 3(3):034007.
- Lobell DB, Burke MB, Tebaldi C, Mastrandrea MD, Falcon WP, Naylor RL, 2008. Prioritizing climate change adaptation needs for food security in 2030. *Science* 319(5863):607–610.
- Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riahi K, Smith SJ, Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ, 2010. The next generation of scenarios for climate change research and assessment. *Nature* 463(7282):747–756.
- PCMDI (Program for Climate Model Diagnosis and Intercomparison). 2007. IPCC model output. Lawrence, CA, USA: PCMDI. (Available from <http://www.pcmdi.llnl.gov/ipcc/about/ipcc.php>) (Accessed on 30 October 2011.)
- R Development Core Team. 2011. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ramírez J, Jarvis A. 2010. Downscaling global circulation model outputs: the delta method. Decision and Policy Analysis Working Paper No. 1. Cali, Colombia: Decision and Policy Analysis (DAPA) Program of the International Center for Tropical Agriculture (CIAT).
- Roudier P, Sultan B, Quirion P, Berg A. 2011. The impact of future climate change on West African crop yields: what does the recent literature say? *Global Environmental Change* 21(3):1073–1083.
- Thornton PK, Jones PG, Ericksen PJ, Challinor AJ. 2011. Agriculture and food systems in sub-Saharan Africa in a 4°C+ world. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 369(1934):117–136.
- Williams JW, Jackson ST, Kutzbach JE. 2007. Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences* 104(14):5738–5742.



The CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS) is a strategic initiative of the Consultative Group on International Agricultural Research (CGIAR) and the Earth System Science Partnership (ESSP), led by the International Center for Tropical Agriculture (CIAT). CCAFS is the world's most comprehensive global research program to examine and address the critical interactions between climate change, agriculture and food security.

For more information, visit www.ccafs.cgiar.org

Titles in this Working Paper series aim to disseminate interim climate change, agriculture and food security research and practices and stimulate feedback from the scientific community.

