IST-Africa 2012 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2012
ISBN: 978-1-905824-34-2

Harnessing Cyber-infrastructure for Local Scale Climate Change Research in Africa

Anwar VAHED[1], Francois ENGELBRECHT[1], Ingo SIMONIS[2], Mogesh NAIDOO[1], Bolelang SIBOLLA[1], Terence VAN ZYL[1], Graeme MCFERREN[1]
[1]*Council for Scientific and Industrial Research, Meiring Naude Road, Brummeria, 0184, South Africa*
Tel: +27 (0)12 841 2954, Fax: +27 (0)12 841 4720, Email: avahed@csir.co.za
[2]*International Geospatial Services Institute GmbH, 73 Werner Heisenberg Str., Emden, 26713, Germany*
Tel: +49 6171 2088.551, Fax: +49 6171 2088.552, Email: ingo.simonis@igsi.eu

Abstract: Climate change poses a major threat to environmental sustainability. Africa in particular, is vulnerable with projected worsening food security, increased threats to public health, increased stress on surface water resources and a general increase in extreme weather events. Political commitment regarding the development of effective mitigation and adaptation strategies needs to be complemented by a more comprehensive understanding of the effects of climate change. However, comparatively little is known about the projected effects of climate change for Africa at the local scale. We present data sets of climate-change projections downscaled to a resolution of 0.5° over the African continent, and data sets downscaled further to an 8 km grid size for areas covering selected urban centres in Africa. These data sets constitute a novel and unique resource for research. We describe a set of web services that provide access to aggregations of selected key climate-change variables derived from these data sets. Furthermore, we propose to incorporate these resources in a scientific workflows platform, EO4VisTrails. EO4VisTrails augments the functionality of existing scientific workflow systems by utilizing distributed computing resources and providing interfaces, services and analytical tools for working with geospatio-temporal data. EO4Vistrails enables an accelerated pace of research by facilitating rapid experimentation, provenance management and parameter exploration.

Keywords: Climate change, environmental sustainability, scientific workflows.

1. Introduction

Climate change poses a major threat to environmental sustainability and socio-economic development. In Africa in particular, factors such as extreme poverty, lack of access to resources, especially natural resources such as water and food, exacerbate already dire socio-economic and environmental conditions. These circumstances are compounded by fragile economies, rapid growth of population, and weak governmental structures. In many urban areas the rates of economic growth and infrastructure development have lagged urbanization rates, resulting in increasing levels of poverty and worsened impacts on the environment.

There has been strong African political commitment to addressing the challenge of climate change through various inter-governmental forums, for example, by the African Partnership Forum (APF) and NEPAD, the New Partnership for Africa's Development [30]. Numerous global efforts are aimed at the provision of information pertinent to climate-change mitigation and adaptation to policy makers. Examples of these are the Global Climate Observing System (GCOS) Implementation Plan for the Global Observing System for Climate, updated in 2010 [1], and the 10-year programme entitled, 'Climate Information for Development Needs: an Action Plan for Africa' [2]. The most prominent global initiative is the Intergovernmental Panel on Climate Change (IPCC) that provides assessment reports on impacts, adaptation, vulnerability and mitigation of climate change at sub-continental and regional scales [3]. These reports are generally regarded as the authoritative resource for global climate-change projections. Some examples of climate-change data and information providers focused on the regional level are:

- the National Center for Atmospheric Research (NCAR) GIS Initiative that serves global Climate Change System Model (CCSM) projections as well as statistically downscaled projections and derived data products of key meteorological variables for the contiguous United States [4]; and
- the UK Climate Projections Report (UKCP09) providing climate change projections on a range of climate-change variables at a 25 km grid for three emission scenarios [5].

In South Africa, regional projections of future climate change are disseminated through the South African Risk and Vulnerability Atlas (SARVA) [6]. However, while most reports find Africa to be particularly vulnerable to the effects of climate change, its real impact at the local scale is still poorly understood [7].

Generally, the types of information available at these providers are observed climate data, typically 20th and 21st century historical information about key variables (temperature, precipitation, sea-surface temperatures and sea-level), climate change projections (for variables such as temperature, precipitation, air pressure, cloud cover and humidity), and marine and coastal projections (sea level rise, storm-surge frequencies, sea-surface and sub-surface temperature, salinity and currents). These data sources underpin numerous advances in knowledge about the growing threat of, and mitigation and adaptation strategies for climate change. However, merely improved and available climate change data does not automatically lead to a better understanding of future national, regional or local changes in the environment. Environmental processes are highly complex by nature; a scientific workflows environment can help to improve understanding of these processes and to derive adaptation and mitigation strategies by integrating knowledge, climate change data and cyber-infrastructure.

We describe the aims and objectives of this research in the following section, followed by the methodology for generating downscaled data sets in Section 3. The generated data sets are discussed in Section 4. A number of derived products aggregating results for key climate-change variables are generated from these data sets; these are presented in Section 4.1. In Section 5, a scientific workflows platform, EO4VisTrails, is presented. EO4Vistrails augments the functionality of VisTrails (an open-source scientific workflow and provenance management system developed at the University of Utah [8][9]) by incorporating services and interfaces for accessing, processing and analysing earth observation data and derived products, as well as components that facilitate distributed computing. Section 5 also presents some examples of scientific exploration and visualisation of climate change data in EO4VisTrails. In the concluding section, we consider possible extensions to EO4VisTrails and the inclusion of access to other climate change data resources for the purpose of research.

2. Objectives

There is a notable lack of research output on climate change issues pertinent specifically to urban centres in Africa. One reason is the lack of data at appropriate resolution – a key requirement for effectively addressing the threat of climate change to environmental sustainability. Another reason is that researchers remote from the generated data sets find it difficult to access appropriate computing infrastructure and analytical tools. This situation represents a real and sometimes insurmountable impediment to exploring and analysing climate change data. Moreover, climate change data sets are often large which limit their transfer to sites with weak connectivity.

The objectives of this research are twofold:

- to generate and provide climate change projection data sets at resolutions of 0.5° (about 60 to 80 km) for the African continent, and at 0.1° (about 8 km) for selected urban centers in Africa. These projections include all Essential Climate Variables (ECV) as described for GCOS [1]
- to implement an ICT environment, a data exploration cyber-workbench, that can be used by researchers to manage and compose scientific workflows for the spatio-temporal domain and that are optimized to handle earth observation data. These scientific workflows represent connected and parameterised processes that can access data sets, utilise high performance computing infrastructure and apply analytical tools.

The aim is to provide an environment that enables more rapid exploration of the environmental, social and economic impacts and risks of climate-change-induced hazards likely to affect urban areas. Such hazards include floods, sea-level rise, droughts, heat waves, desertification, storms, vegetation fires as well as epidemics and pandemics.

3. Methodology

The generation of the climate-change projection data presented here, also constitutes a contribution to a project entitled “Climate Change and Urban Vulnerability in Africa” (CLUVA), co-funded by the European Commission under the Seventh Framework Programme call on Environment (including climate change) [10]. This work involves the downscaling of an ensemble of six global projections of future climate change, as obtained from coupled global climate models (CGCMs) that contributed to the IPCC Assessment Report 4 (AR4), to a higher resolution of about 60 to 80 km over the entire African continent [34]. A dynamic regional climate model is used for this purpose. The downscalings are performed for the period 1961-2100. Through the process of multiple nudging, these projections are downscaled further to a resolution of about 8 km over areas of interest that include the cities selected as case studies in CLUVA. These cities are Addis Ababa (Ethiopia), Dar es Salaam (Tanzania), Ouagadougou (Burkina Faso), St Louis (Senegal) and Douala (Cameroon). That is, an ensemble of six very high resolution simulations is constructed for areas surrounding all the CLUVA areas of interest.

Through the collaboration between the Euro-Mediterranean Center for Climate Change (CMCC) in Italy and the Council for Scientific and Industrial Research (CSIR) in South Africa, a multi-model ensemble of eight high-resolution simulations of climate and climate change are derived (two at CMCC and six at the CSIR). That is, a multi-model ensemble of simulations of present-day and future climate is available for each of the cities. The data sets produced, will be most useful to describe the range of uncertainty associated with future climate for each of the cities.

The model simulations of present-day climate are verified against observations and an associated model inter-comparison study are performed. Some impact studies over urban areas require plausible scenarios of future climate at ultra-high spatial resolution of one kilometer or less. For this purpose the regional scenarios will be downscaled further to 1-2 km resolution by CMCC, using specific and accurate statistical techniques, such as a novel spectral-based stochastic downscaling technique. Ultra-high resolution projections are to be performed over one or two of the cities under consideration by the CSIR, through the dynamic downscaling of the 8 km resolution simulations.

3.1 Experimental Design

The conformal-cubic atmospheric model (CCAM) of the Commonwealth Scientific and Industrial Research Organisation (CSIRO) is used to perform the dynamic downscalings. CCAM is a variable-resolution atmospheric global circulation model. When applied in stretched-grid mode, it effectively functions as a regional climate model. The model is highly suitable for the purpose of regional climate modelling, due to its computational efficiency and variable-resolution formulation that allows great flexibility in downscaling CGCM data, through the application of a multiple nudging technique [11]. The model's ability to simulate the present-day characteristics of African climate has been investigated over Africa [12, 34] and for various other climatological regions (e.g.[13], [14]).

A four-phase multiple downscaling strategy is followed to obtain 200 km, 60-80 km and 8 km and eventually, 1 km resolution simulations. The first phase in the downscaling procedure requires that the model be forced with the sea-surface temperature (SST) and sea-ice output of a number of different CGCMs used in AR4 of the IPCC, for the period 1961-2100. All the projections performed at the CSIR are for the A2 emission scenario of the Special Report on Emission Scenarios (SRES). The multiple-nudging strategy followed first involves the integration of CCAM globally at quasi-uniform C48 resolution (about 200 km in the horizontal as shown in Figure 1), forcing the model with the SSTs and sea-ice of each host model, and with CO_2, sulphate and ozone forcing consistent with the A2 scenario.

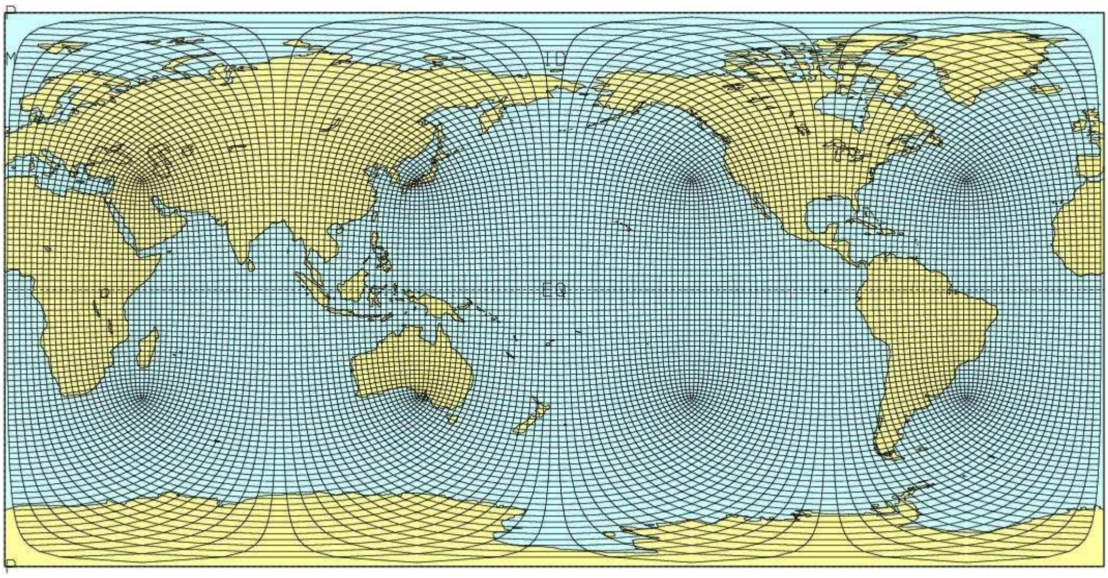

Figure 1: Quasi-uniform C48 CCAM grid with a horizontal resolution of about 200km

In a second phase of the downscaling, the model is integrated in stretched-grid mode over Africa (Figure 2), with the modestly-stretched grid providing a resolution of about 50 km to 80 km over Africa. The higher resolution simulations are nudged within the quasi-uniform C48 simulations, through the application of a digital filter using a 4000 km length scale. The filter was applied at six-hourly intervals and from 900 hPa upwards. These first two phases of the downscaling procedure have been completed. Model output for the 200 km resolution ensemble members is available at six-hourly intervals, whilst output for the 60 km to 80 km resolution simulations is available at hourly intervals for selected variables. Subsequent downscaling to 8 km and eventually 1 km resolution is to be performed.

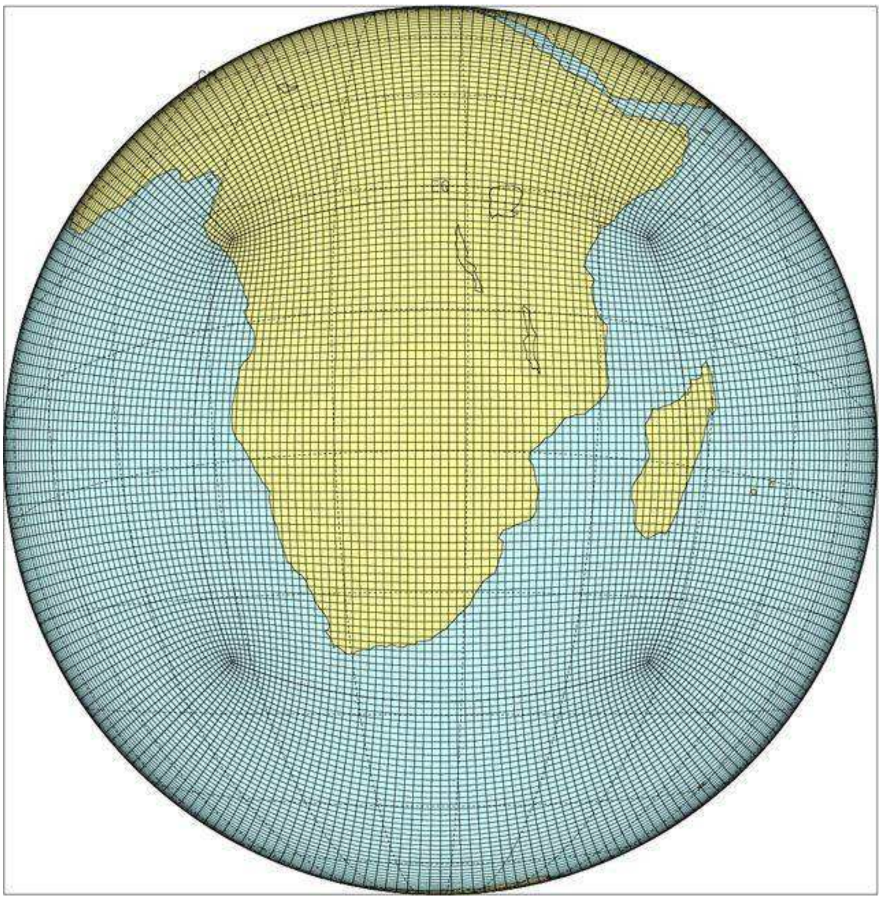

Figure 2: Stretched high-resolution conformal-cubic-grid, horizontal resolution of about 60-80km over Africa

4. Model Output Statistics and Dissemination

The available computing power at the CSIR has allowed the completion of six high-resolution (60 km to 80 km horizontally) simulations over the African continent. The simulations cover the period 1961-2100. Analysis of these simulations and their subsequent downscaling to very high spatial resolution (8 km) over the CLUVA areas of interest are currently in progress. Figure 3 shows the model projections of the annual change in rainfall totals over a part of East Africa (including the Dar es Salaam region), for the period 2071-2100 versus the base-line period 1961-1990. The six ensemble members project a robust pattern of rainfall increases over this region. Note that the projections are available continuously for any time-slab within the period 1961-2100.

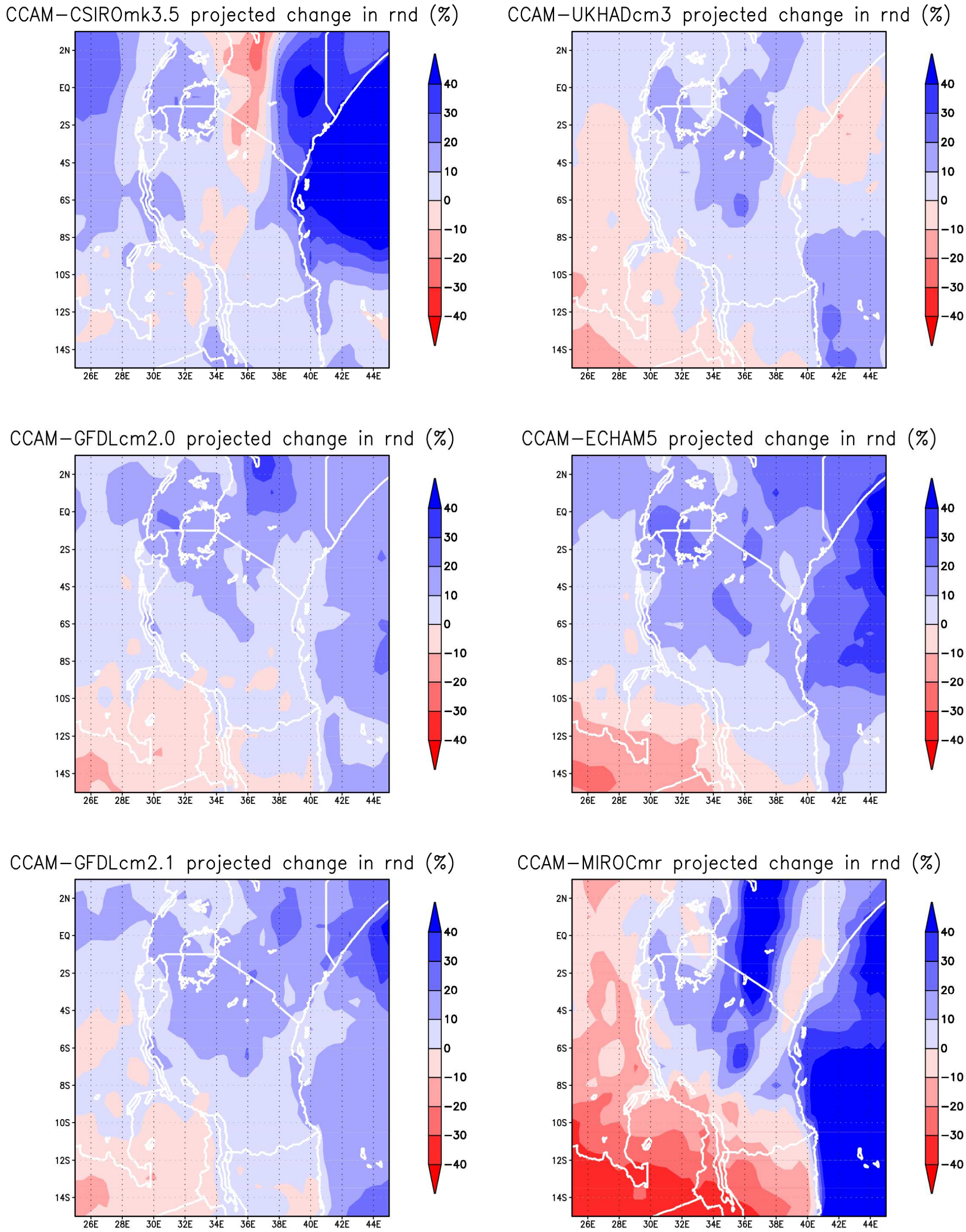

Figure 3: Six CCAM projections of changing annual rainfall patterns (expressed as a % change) over a part of East Africa (resolution about 60 km), for the period 2071 -2011 vs 1961-1990.

The coupled models for which downscalings are presented in Figure 3, are as follows.

CCAM-CSIRO mk3.5: Version 3.5 of the Coupled Global Climate Model (CGCM) by the Commonwealth Scientific and Industrial Research Organisation in Australia (CSIRO) [11]

CCAM-UKHADcm3: The third version of the United Kingdom's Met Office Hadley Centre Coupled Ocean-Atmosphere Global Climate Model [40]

CCAM-GFDLcm2.0 and CCAM-GFDLcm2.1: Versions 2.0 and 2.1 of the Coupled Global Climate Model of the Geophysical Fluid Dynamics Laboratory (GFDL) of the National Oceanic and Atmospheric Administration in USA [41]

CCAM-ECHAM5: The fifth generation climate model developed by the Max Planck Institute in Germany [39]

CCAM-MIROmr: The MIROC3.2-medres Model for Interdisciplinary Research on Climate 3.2, medium resolution version, of the Japanese Agency for Marine-Earth Science and Technology [42]

Dynamic climate models typically output a wide range of meteorological variables throughout the troposphere and into the lower stratosphere. In addition to the meteorological variables, simulations are also performed for land-surface variables (e.g. soil moisture) that influence atmospheric processes. The three-dimensional spatial structure of meteorological data, in combination with the high temporal resolution of the data sets, lead to these data sets being very large – in the order of tens of terabytes.

The model projections were conducted on the CSIR's Centre for High Performance Computer (CHPC). Utilising 36 processor clusters of 3.6 GHz each, the total simulations required 3 months of computing time to produce the complete variable set for the ensemble. The data set in its original form was about 80 TB in size. These downscalings represent the state-of-the-art in the application of a dynamic climate model, to obtain projections of future climate change. The CSIR has now commenced with the downscaling of the CGCMs for the Coordinated Regional Downscaling Experiment (CORDEX), an experiment through which the new set of CGCMs of Assessment Report 5 (AR5) of the IPCC are downscaled to a number of areas – including the African continent.

4.1 Dissemination

Most climate variables required for climate-change impact studies and application modelling are ones that describe the state of the lower troposphere. Such variables include rainfall, minimum and maximum temperature, wind speed, relative humidity, evaporation, soil moisture and run-off. Storing the surface variables requires significantly less space than archiving the full three-dimensional data sets. For example, a single projection for the African continent at a horizontal resolution of 0.5 degrees (about 60 km) and with about 20 levels in the vertical, saving output at 12-hourly intervals (1-hourly interval for the surface variables), requires about 2 TB of storage space. Storing only the key surface variables as output reduces the size of the data set to about 100 GB.

Output from the CCAM model is in the form of NetCDF files. The data available on the model's conformal cubic grid (Figure 2) and in stretched-mode the grid resolution, varies between 60 km and 80 km. What was needed for effective dissemination of this data was that the output be on a grid of a popular projection. This is necessary to allow easy display and analysis using common GIS and display packages. Conversion from CCAM native projection to uniform grid resolution was achieved through a CSIRO developed interpolation routine. This routine enables the interpolation of data from the conformal-cubic grid to a regular latitude-longitude grid for any sub-domain of choice on the globe.

An extensive communications infrastructure is necessary for the dissemination of full data sets; this is infeasible for most cases in Africa. Moreover, while access can be provided, the use of such resources may be impractical if interest is limited to only a few key variables. A host of web services have thus been developed to provide access to derived and aggregated climate-change data sets. The initial focus is on providing a visual representation of data projected over time (Web Map Service images), and data sets downloadable in NetCDF, ASCII and Microsoft™ Excel formats. Several key variables such as wind speed, sea surface temperature, air temperature and 'Very Hot Days' (i.e., maximum temperature exceeding 35° C) were selected to pilot these services and to solicit feedback from potential users on the usefulness of the user interface and data formats[1]. These variables were prepared on a uniform grid of 0.5^{o} (~60 km) on the domain 20N; -20S and -20W; 60E, and were aggregated to a monthly average of minima and maxima. This reduced file size dramatically, with data set size for a single variable of monthly average being ~28 KB. Thus CCAM outputs a value at the end of each model day. These values (31 in, for example, the month of January) are then averaged over the month of concern. This was performed for each variable over the total modelling period of 1961-2100. A similar approach is to be used for each ensemble member.

We deployed an OPeNDAP (Open Source Project for Network Data Access Protocol) [32] service (PyDAP [33]) to provide the NetCDF data to data access clients over the Internet. The OGC Web Map Service with temporal extension (WMS-Time), was deployed to render map images from spatially referenced data.

5. EO4VisTrails

Scientific workflows are becoming established as a technology for facilitating complex and distributed computational processes and are used to advance the rate of scientific endeavour [15]. Scientific workflows and workflow environments are used as means for modelling and enacting scientific experiments. They can be used to formalise the ad-hoc steps scientists follow for deriving publishable results from raw data. They enable experimentation, the exploration of ideas and production of scientific outputs that utilise timely, adequate resolution, validated data in reproducible and shareable ways [16].

We know of at least 25 different implementations of scientific workflows, most of which are focussed on specific application domains such as bioinformatics [17]. These implementations offer different functions and workflow paradigms to users. Most current technologies aim to provide a robust environment for managing access to large scale distributed high performance computing environments, disparate and distributed data sources and processing steps. Some of these environments also attempt to implement knowledge management (acquisition, publication and sharing) technologies.

We surveyed some of the most popular scientific workflow systems [17] for their ability to integrate earth observation and climate change data with distributed computing resources and analytical tools. Some of the platforms considered were Kepler [18], Taverna [19], Triana [20], Pegasus [21] and VisTrails [22]. The criteria for evaluating these environments included the efficacy of the user interface, ability to utilize web-based resources, support for access to spatial and scientific data processing libraries, provenance management, extensibility, open source nature and interoperability with external applications.

The VisTrails platform fulfilled most of these criteria, and was selected as scientific workflows platform. Initially designed for use and application in the medical domain, VisTrails is an open-source scientific workflows system that supports data exploration and visualization.

EO4VisTrails extends the VisTrails platform with functionality aimed at the spatio-temporal domain and optimized to handle earth observation data. EO4VisTrails includes functionality for access to OGC [24] standards compliant web services, interfaces to geographic information systems components from Quantum GIS [25], database extensions such as PostGIS [26], numerical analysis tools such as the NetworkX library [27], statistical programs such as R [28] and PySAL [29]. EO4Vistrails also has the capability to leverage distributed and high performance computing environments through simple parallelization and support for the composition of large numbers of parallel computations. Figure 4 shows

1 A pilot test site to these services has been set up at http://146.64.28.151/

the VisTrails interface and some of the modules implemented in EO4VisTrails, respectively.

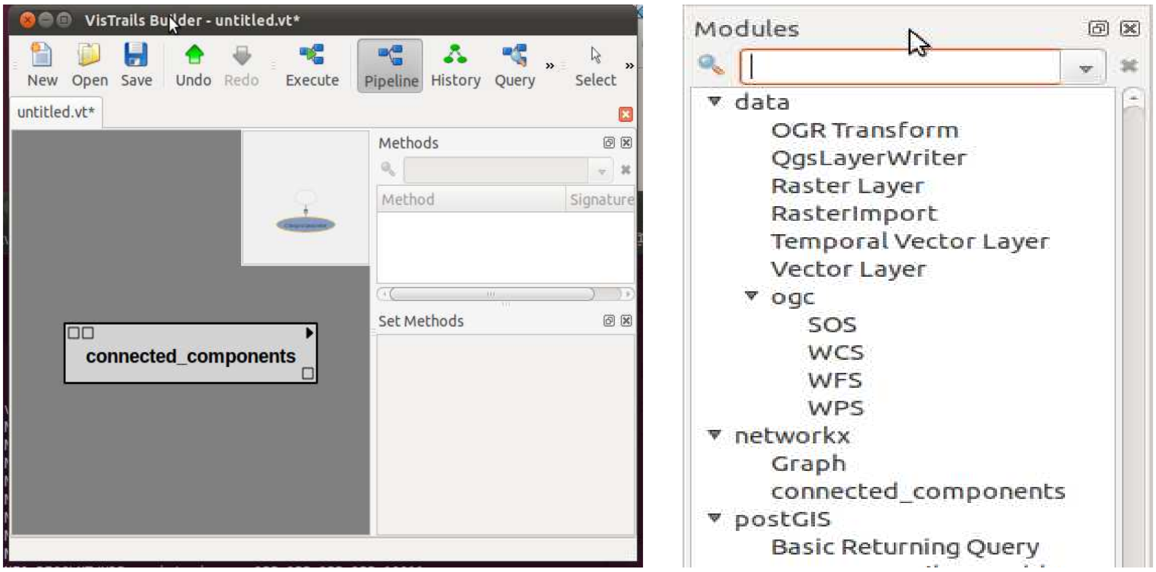

Figure 4 Left: The VisTrails interface showing the workflow canvas and mathods selector; Right: Modules in the EO4Vistrails package

The basic unit of workflow composition is a VisTrails module; modules can have a set of input ports and output ports. These ports provide connectivity to other VisTrails modules. Linked modules form an execution sequence, called a workflow. Input ports parametrise a module and can either take a value from the output port of another VisTrails module or may be set to a constant value. Developers specialise the basic Vistrails module to provide various additional modules with specific processing capabilities. All ports are typed and only input ports and output ports of the same type may be connected.

5.1 Climate change scientific workflows

A climate change scientific workflow example in EO4VisTrails is shown in Figure 5. In this instance, an EO4Vistrails module define the area of interest (Cameroon), accesses data through the PyDAP module and provides extracted data to the data cube function. From this point, two processes execute, possibly at distributed computing nodes. The left branch computes a standard histogram for the data (parameterised in some way), composes a graphic image and renders this in one of the visualization cells of EO4Vistrails. The other branch effectively computes the mean and presents both the array and the mean, using a QGIS function, in another visualization cell. This workflow can have several parameters which allow for the exploration of various scenarios. For example, results for a different area can be obtained by merely executing the workflow with different parameters for the AreaOfInterestDefiner module. This parameterisation is an example of the capabilities provided by EO4VisTrails.

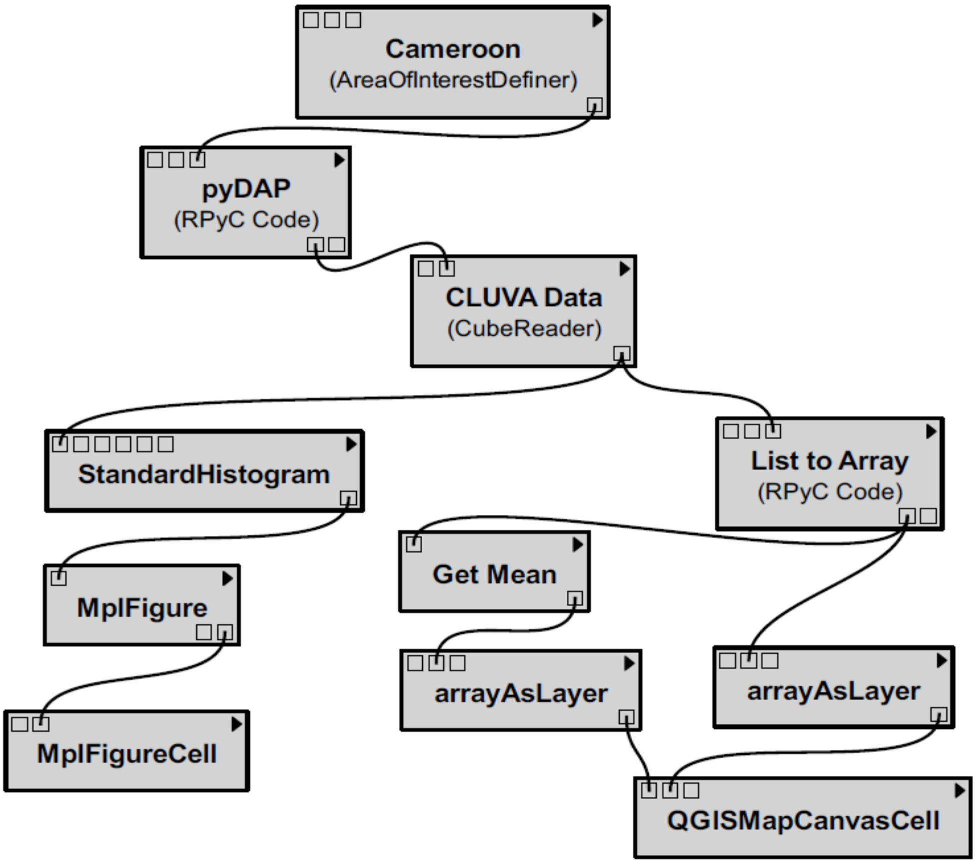

Figure 5: Example of a climate change data processing workflow

6. Summary

The climate-change data integration in cyber infrastructure research we have performed so far highlights several points. We elaborate the most important ones here, and shall provide more detailed discussions in further publications.

The VisTrails package is implemented as ‘standalone’ software and can be downloaded and easily installed on conventional personal computer systems with Microsoft™ Windows and some distributions of Linux operating systems [9]. The development of scientific workflows in VisTrails is guided by extensive user and example documentation, also available online. EO4VisTrails module code is available online [23] and can be incorporated into the standard VisTrails installation, requiring minimal IT background. Presently, EO4VisTrails is being tested and validated with the intent to release an open, operational and installable version towards the end of 2012. This version will be maintained at least for the duration of an EU FP7 funded project, until March 2014 with further support primarily dependent on funding.

Good communications bandwidth is required when accessing Web-based services, remote data or when deploying processes to remote computing processors. Data files can, of course, also be accessed at the local installation of VisTrails. For practicable download times, data files of approximately 10MB have been derived from the initial climate data sets, for selected variables and for different time scales[i].

VisTrails is well used among researchers in Environmental Science, Quantum Mechanics and the biomedical domain. One of the main reasons for choosing an open source code base, is to rally a community of practitioners around both the development, use and extension of this platform to other domains. The EO4VisTrails web site will serve as a contact interface for these purposes and for user feedback. In its present implementation, EO4Vistrails is useful for researchers wishing to explore and analyse local-scale impacts of climate change in Africa. However, this platform has generally applicability for enabling access to high-end computing and data resources. Minimal effort is required to incorporate data sets for other application domains. Some recommendations are thus for other developers to contribute to the development and extension of this platform to other domains, and for researchers to explore this platform and to provide feedback on its usefulness as a tool that can promote collaborative and shared research among African and European researchers.

We have achieved the stated objectives as follows.

- Climate change projection data sets at resolution of 0.5° have been produced for the entire African continent and we are currently generating data sets at 0.1° as well. The initial 'unprocessed' data files are approximately 80 terabyte in size and smaller files for some selected variables (e.g, air temperature, humidity, precipitation), for each model of the ensemble. These data sets are currently made available through various Web service interfaces to allow for optimal accessibility. A web site currently under development will also provide ftp access to these files at ftp://ict4eo.meraka.csir.co.za/CLUVA. This URL will become accessible externally once security concerns have been addressed. All data is accessible free of charge.
- The EO4VisTrails platform has been developed and its code base is freely available at http://code.google.com/p/eo4vistrails. The software currently includes several freeware geographic information processing tools (e.g., PostGIS and Quantum GIS), and analytical tools such as the statistical package R and Octave, the latter being a freeware equivalent of Matlab(c). [35, 36, 37, 38]

Since the climate-change unprocessed data sums up to several terabytes, broadband access is required to download this data. Clearly, the efficient use and analysis of climate-change data require powerful analytical tools and good connectivity at the end user side, or considerable pre-processing capability on the provider side. We tackled both sides by (1) extending a scientific workflow and provenance management system with particular earth observation data handling capabilities at the end-user side, enabling the research community to integrate climate change data with cyber-infrastructures and giving full control to the researcher; and (2) by deploying a number of web services with standardized interfaces on the server side that include a number of pre-processing steps and allows end users to obtain a number of key variables in various representation formats, such as maps or tables. We intend to deploy the web services operationally once they have been tested and validated fully. These services will be supplemented by more components that provide richer functionality and easy access for researchers who wish to study the impacts of climate change on Africa at a more detailed scale than what was hitherto possible.

References

[1] Global Climate Observing System, "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC", August 2010, GCOS-138, World Meteorological Organisation (WMO), Switzerland. http://www.wmo.int/pages/prog/gcos/Publications/gcos-92_GIP_ES.pdf

[2] Global Climate Observing System, UN Economic Commission for Africa and African Union Commission, "Climate Information for Development Needs: An Action Plan for Africa Report and Implementation Strategy", GCOS-108, http://www.wmo.int/pages/prog/gcos/Publications/gcos-108.pdf

[3] Intergovernmental Panel on Climate Change (IPCC), "Fourth Assessment Report: Climate Change", 2007 (AR4), IPCC Secretariat, World Meteorological Organization (WMO) Switzerland, http://www.ipcc.ch/publications_and_data/publications_and_data_reports.shtml

[4] National Center for Atmospheric Research (NCAR), Geographic Information Systems (GIS) Initiative, Boulder, Colorado, USA, http://www.gis.ucar.edu/atmo_data.jsp

[5] UK Department for Environment Food and Rural Affairs (DEFRA), "United Kingdom Climate Change Projections (UKCP09)", UK, http://ukclimateprojections.defra.gov.uk/content/view/868/531/

[6] Council for Scientific and Industrial Research, “SA Risk and Vulnerability Atlas”. Online: http://www.rvatlas.org/index.php
[7] Hulme, M., R.M. Doherty, T. Ngara, M.G. New, and D. Lister. 2001. "African climate change: 1900-2100," Climate Research, Vol. 17, No. 2, pp. 145-168.
[8] Freire, J., D. Koop, E. Santos, C. Scheidegger, C. Silva, and H.T. Vo. VisTrails, Ch 23: “The Architecture of Open Source Applications”, A. Brown and G Wilsions (Eds), 2011, Creative Commons Attribution 3.0.
[9] Scientific Computing and Imaging Institute (SCI), VisTrails wiki, University of Utah, Salt Lake City, Utah, http://www.vistrails.org/index.php/Main_Page
[10] Climate Change and Urban Vulnerability in Africa (CLUVA), Seventh Framework Programme: Environment (including climate change), Call FP7-ENV-2010, http://www.cluva.eu/
[11] McGregor J. L. (2005). C-CAM: Geometric aspects and dynamical formulation [electr. publication]. CSIRO Atmospheric Research Tech. Paper No. 70, 43 pp.
[12] Engelbrecht F.A., J.L. McGregor, and C.J. Engelbrecht. 2009. “Dynamics of the conformal-cubic atmospheric model projected climate-change signal over southern Africa” International Journal of Climatology, 29, 1013-1033.
[13 Nunez M. and J.L. McGregor. 2007. “Modelling future water environments of Tasmania, Australia”. Climate Research, 34, 25-37.
[14] Lal M., J.L. McGregor and K.C. Nguyen. 2008. “Very high resolution climate simulation over Fiji using a global variable-resolution model”. Climate Dynamics 30, 293-305.
[15] Gil, Y.; Deelman, E.; Ellisman, M.; Fahringer, T.; Fox, G.; Gannon, D.; Goble, C.; Livny, M.; Moreau, L. & Myers, J. „Examining the Challenges of Scientific Workflows“ Computer, IEEE Computer Society, 2007, 40, 24-32
[16] Taylor, I., Ewa Deelman, Dennis Gannon, and Matthew Shields (Eds.) “Workflows for e-Science” Springer, New York, Secaucus, NJ, USA, 2007.
[17] Workflow Management Research Group, Aleksander Slominski and Gregor von Laszewsky, Scientific Workflows Survey, http://www.extreme.indiana.edu/swf-survey/
[18] The Kepler Project, National Science Foundation, http://kepler.ecoinformatics.org/
[19] “Taverna Workflow Management System”, School of Computer Science, University of Manchester, UK, http://www.taverna.org.uk/
[20] Taylor, I., M. Shields, I. Wang, and A. Harrison. “The Triana Workflow Environment: Architecture and Applications”. 2007. In I. Taylor, E. Deelman, D. Gannon, and M. Shields, editors, Workflows for e-Science, pages 320-339. Springer, New York, Secaucus, NJ, US.
[21] Deelman, E., Gurmeet Singh, Mei-Hui Su, James Blythe, Yolanda Gil, Carl Kesselman, Gaurang Mehta, Karan Vahi, G. Bruce Berriman, John Good, Anastasia Laity, Joseph C. Jacob, Daniel S. Katz. “Pegasus: a Framework for Mapping Complex Scientific Workflows onto Distributed Systems”, Scientific Programming Journal, Vol 13(3), 2005, Pages 219-237.
[22] Callahan, S.P., J. Freire, E. Santos, C.E. Scheidegger, C.T. Silva, and Huy T. Vo. Managing the Evolution of Dataflows with VisTrails, In: Proc. Data Engineering Workshops, 22nd International Conference on-. 2006. 71-71, Atlanta, USA.
[23] ICT for Earth Observation Research Group, CSIR – Meraka Institute, “EO4VisTrails – Earth Observation Capabilities for VisTrails”. https://code.google.com/p/eo4vistrails/
[24] Open Geospatial Consortium, OGC. 2011. Online: http://www.opengeospatial.org
[25] Quantum GIS Development Team. 2011. Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org
[26] Holl, S., and H. Plum. “PostGIS”. GeoInformatics, Vol. 03/2009. April 2009, pp. 34-36.
[27] Hagberg, A., Schult, D., and Swart, P. “NetworkX. High productivity software for complex networks”, https://networkx.lanl.gov/
[28] Ihaka, Ross and Gentleman, Robert. "R: A Language for Data Analysis and Graphics". 1996. Journal of Computational and Graphical Statistics, 5, 3, American Statistical Association. Online http://www.jstor.org/stable/1390807
[29] Rey, Sergio J., Anselin, L., Fischer, M. M., and Getis, A. "PySAL: A Python Library of Spatial Analytical Methods", Handbook of Applied Spatial Analysis. 2010. Springer Berlin Heidelberg. 175 - 193. http://dx.doi.org/10.1007/978-3-642-03647-7_11
[30] NEPAD 2011: The New Partnership for Africa's Development. Online: http://www.nepad.org/
[31] Parry, M.L., O.F. Canziani, J.P. Palutikof, P.J. van der Linden and C.E. Hanson, (eds.) 2007. “Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change”. Cambridge University Press, Cambridge, UK
[32] Cornillon, P., and J. Gallagher and T. Sgouros. 2003. “OPeNDAP: Accessing data in a distributed, heterogeneous environment”. Data Science Journal. 2, 164-174
[33] Pydap Python Library for implementing the Data Access Protocol, Online: http://pydap.org/

[34] Engelbrecht, F.A., W.A. Landman, C.J. Engelbrecht, S. Landman, M.M. Bopape, B. Roux, J.L. McGregor and M. Thatcher, 2011. "Multi-scale climate modelling over Southern Africa using a variable-resolution global model". Water SA. 37 No.5. http://dx.doi.org/10.4314/wsa.v37i4.18.
[35] R Development Core Team, R: A language and environment for statistical computing, 2008, R Foundation for Statistical Computing, 2008, Vienna, Austria, http://www.R-project.org , Accessed August 2012
[36] Eaton, John E., 2002, GNU Octave Manual, Network Theory Limited
[37] Open Source Geospatial Foundation, The PostGIS Project, http://www.postgis.org Accessed April 2010
[38] Quantum GIS Development Team, Quantum GIS Geographic Information System, Open Source Geospatial Foundation Project, 2011, http://qgis.osgeo.org Accessed August 2011
[39] Reockner, E., G. Bauml, L. Bonaventura, R Brokpof and M. Esch, The atmospheric general circulation model ECHAM5 – Part 1, Max Plank Institute for Meteorology, http://www.mpimet.mpg.de/fileadmin/publikationen/Reports/max_scirep_349.pdf. Accessed February 2012
[40] Johns, T.C., 2006. The new Hadley Centre climate model, Journal of Climate, UK National Meteorology Laboratory, **19**, 7, 1327-1353
[41] Geophysical Fluid Dynamics Laboratory, 2007, GDFL Global Atmospheric Model, Princeton University, Princeton, USA, http://www.gfdl.noaa.gov/. Accessed August 2011
[42] Lin, J-L., Shinoda, T., Qian, T., Han, W., Roundy, P. And Zhen, Y., Intraseasonal variation of winter precipitation over the western United States simulated by 14 IPCC coupled models, J. Climate, **23**, 3094-3119, doi: http://dx.doi.org/10.1175/2009JCLI2991.1

[i] Derived data files are presently available at http://ict4eo.meraka.csir.co.za/anonftp/