Independent Review of the Major Data Sets developed or planned by the Climate Variability Group

Hadley Centre technical note 68

S F B Tett and D E Parker

10 April 2003

Hadley Centre

MS – RAND – GMR2002/4 – Product

Independent Review of the Major Data Sets developed or planned by the Climate Variability Group

S F B Tett and D E Parker
CR Division, Hadley Centre

Contract Deliverable IIIB/02/10

File: M/GMR/6/3 Annex A

Delivered from authors:	Signature	Simon Tett	Date	10th April 2003
Approved by Activity Manager:	Signature	C K Folland	Date	10th April 2003
Approved by HdGMR:	Signature		Date	
Sent to MoD	Signature		Date	

Version Number: 4.0

Number of Pages: 37 including cover.

Security Classification: Unclassified

DOCUMENT REVIEW HISTORY
Document History

Date	Version	Action/comments:	Approval
21/3/03	1.0	First Draft from David Parker	
27/3/03	2.0	Addition of responses and discussion with Vicky Pope	
2/4/03	3.0	Modification of responses following discussion with Vicky Pope and David Parker and inclusion of reviewer's comments in document	
10/4/03	4.0	Incorporation of comments from CKF/NR and DEP	

Independent Review of the Major Data Sets developed or planned by the Climate Variability Group

Synthesis

David Parker and Simon Tett

March 2003

I. Introduction

In December 2002, we requested comments from 16 independent reviewers, in the UK and worldwide, on the major data sets developed by the Climate Variability Group, and on our plans for data set development up to the year 2007. We received 11 responses. Three strands in the reviews are dominant:

The importance of our data sets and of continuing to improve many aspects of them (eg resolution, underlying data)
The need for rigorous error-estimates
The need for unhindered, near-real time availability of the data sets.

We plan substantial improvements in error-estimation, and in the availability of our data, during the next financial year and subsequently.

II. Responses to selected points made by the reviewers

Some reviewers made specific points to which we have responded.

1. **Most reviewers made the comment that we needed to improve accessibility to data sets and document the assumptions that are used in their construction.**

   We have already agreed with the Met Office that users can use the data sets via the web using a "click through" license (i.e. user agrees to license terms by clicking an OK button and then has immediate access to the data). This goes some way to improving accessibility to our data sets but probably does not go far enough in dealing with some of the issues raised by reviewers, for example the comments made by Neville Nichols and CEFAS. One extreme position would be to allow all users to access the data sets without charge. This would simplify accessibility but would likely contradict the Met Office's data policy, which seeks to recoup data costs from users. Another possibility would be to allow use by anyone for research purposes with the outcome being placed in the public domain or to provide advice to any UK government departments. Given the likely limited commercial value of these data sets and the cost of a system to utilise the data we believe there is a case for open access. There is increasing Internet competition. Thus the Kaplan (USA) sea surface temperature data set (which has some problems) is freely available and very conveniently integrated into the KNMI Explorer system. This allows a range of scientific analyses to be done very simply on line with this data set (and other data)- by anybody.

   Our current thoughts are to distribute data on the web using a "click through" license and to ask users to give us their email address and what they plan to use the data for. We have two reasons for doing this. 1) To see how many people are using our data and 2) to be able to inform users of updates, errors or changes to the data sets they are using. Documentation on the construction of the data sets will be made available on the web.

   We would appreciate a view from MOD on data distribution and licensing.

2. **John Christy made the point that data set production is a process of continuous improvement while several reviewers stressed the need for near-real-time production.**

We note Christy's point that data set production is a process of continuous improvement. Given limited resources we will need to make hard choices about where we put our effort. We will continue to be guided in our work by colleagues in the Hadley Centre and further afield and by collaboration with other groups.

So that we can inform government, the media and the public of climate events soon after they occur we need to update our major data sets in near-real-time. We will continue to distribute many of our data sets in near-real-time and educate users that the data sets are continually updated. To monitor climate and distribute data sets we need robust infrastructure. Though we understand that DEFRA are concerned only with outcome and thus give lower priority to infrastructure we believe that with reasonable investment in automated infrastructure we will be able to deliver more outcome for the same resource. Phil Jones made the point "Automation of some sort is clearly the way to go, but it does need some checking every now and then." We will include human oversight and review in our processes.

3. **Net radiation and forcing data sets (Nigel Arnell and John Christy)**

Nigel Arnell wrote "...Of these, net radiation appears currently to be the largest gap: data sets currently contain cloud cover or sunshine hours. There is a need for the construction of a consistent data base of surface net radiation (long wave and short wave)...." While John Christy wrote "...Traditionally, we think of climate datasets as datasets of response variables (temperature, wind), but it will likely unfold that climate datasets will include datasets related more to forcing (aerosols, cloud etc.)...". We note both points and agree that it would be useful to have such data sets. For example a good long-term and homogeneous short-wave data set may be able to reduce the uncertainty in the forcing from sulphate aerosols. Though some work in this area could be done in collaboration with other groups (in particular the GEBA group at ETH, Zurich), substantial progress would require extra resources or a reduction in activity in other areas.

4. **The need for high resolution SST data sets in UK waters (DEFRA)**

We did not expect comments from CEFAS and we welcome them. Should we have sufficient resources to generate a high-resolution coastal SST product then we will work closely with CEFAS to produce it. Climate change in the coastal region may be different from changes in the open ocean and as a large fraction of the UK population lives close to the coast we believe that this could be an important issue. Our belief is that, if we successfully develop a common gridding algorithm then the resources needed to produce a prototype high resolution data set would be small.

5. **Reduced priority for HadSLP and sea-ice (DEFRA)**

One important research issue is the impact of climate change on circulation changes -- HadSLP will enable us to examine this as well as "natural" climate variability. The value of this approach is shown by the first detection paper on changes of world-wide atmospheric circulation which appeared in Nature in March 2003, but there remain many important uncertainties in the observed data. In addition, early climate change detection (on means or extremes) at regional and subregional scales will require that the influences of atmospheric circulation be taken into account. Sea ice is an important feedback on climate change and current data sets contain many inhomogeneities and currently no homogeneous data set exists (see comments by Christy and Reynolds). Prior to the early 1970s our Antarctic sea-ice data set is only a climatology due to a lack of direct observations. There may be data available that would allow extension back in time of changes in Antarctic sea-ice extent.

6. The importance of a sub-surface ocean analysis (many reviewers)

Several reviewers (Christy, DEFRA, Stott , Karl, Reynolds and Trenberth) said that this was an important data set to produce. We plan to increase the priority for this project.

7. Errors will co-vary (Phil Jones and Kevin Trenberth)

Phil Jones writes "…to recognise that there needs to be consistency between the estimates on different timescales…" and Kevin Trenberth says "…many errors and data are not independent and assuming "random" is not appropriate. Temporal and spatial persistence must be factored in…". We are aware that the error estimates will co-vary. The challenging issue is how to communicate these estimates to research and, especially, policy users.

8. Use of SYNOP messages (Phil Jones)

Phil Jones wrote "…use of the SYNOP …. needs to be undertaken with extreme care" We note the point but use of such data would help improve coverage in many parts of the world (not just Europe). Our current plan is to generate monthly-means from SYNOP data and quality control them using our current process. Only if enough SYNOP data pass quality control in regions where current data sources are sparse would we consider their use.

9. Difficulties in using a common gridding algorithm (Phil Jones)

We think for gridding geographically and temporally near observations that one method will work reasonably well to generate both a best estimate and an error estimate. We believe that different variables will need different quality control approaches and covariance structures and accept that interpolation of variables will require more thought.

10. Marine data sets could be improved particularly in poorly sampled regions (Mark New).

Mark New wrote "Any improvements to Marine Temperature datasets will be of great importance, particularly over poorly sampled regions.". We could improve coverage of our marine data sets, particularly in the late 19th and early 20th centuries, if we had extra resources to digitise data held at Kew though this would also require a solution to crown copyright. Digitisation costs about £0.1/observation but focused digitisation of about a million records at a cost of about hundred thousand pounds through a contract with a specialist firm would make a real difference to our knowledge of late 19th and early 20th century climate. Resources within the existing contract would allow the management of such a contract.

Another possible, and cheaper approach, would be to work with NCDC. NCDC are well funded to digitise data. If this were possible we would need to scan some or all of the Kew archive and transfer the scans to NCDC. NCDC would then digitise the records. Crown copyright on the archives may make this harder to do than the first option and extra funding would be needed to scan the Kew archives.

11. Make use of *in situ* hourly rainfall and radar rainfall (Kevin Trenberth)

Kevin Trenberth wrote ". I urge that ... hourly data ... address frequency and intensity of precipitation and hence extremes and runoff. Also, there is a wealth of radar data on precipitation...". This is something that we had not previously considered. We have discussed the idea with the "extremes" theme manager and results would certainly be of interest. For the UK, there are hourly data in the Met Office archive and more in Water Company archives and the Met Office has been archiving "raw" radar data since 1998. More resources would be needed to utilise both the in situ hourly rainfall data and especially the radar rainfall data.

12. Datasets are produced independently and so incorporate no knowledge about the physics of the climate system. (Kevin Trenberth)

Kevin Trenberth wrote "... the datasets are largely dealt with independently. An advantage of a physically based model is that physical relationships can be exploited. A classical example is the analysis of weather maps in terms of isobars resulting in a pressure map, where winds are extensively used to set gradients, and current weather, cloud, precipitation etc is used to draw fronts etc.". This is a reasonable point. Our current view is that production of independent data sets allows exploration of these linkages. An example of the physically linked analysis of two data sets could be SLP and wind over the oceans.

III. Summary of Reviews

Remarks from all reviewers are synthesised below. Table 1 lists specific points from the reviews regarding the data sets and plans, and Table 2 is a collation of some strategic remarks. Where appropriate, sources of comments are indicated.

Table 1. Specific comments on the data sets and plans

What are the data sets used for?

- **Climate monitoring and diagnostics including extremes**
- **Input to other observational analyses, eg precipitation rates (New), hydrological studies (Arnell)**
- **Analysis and understanding of regional and subregional patterns of atmospheric and oceanic climate variability and change, with confidence intervals**
- **Boundary conditions and input data for models including (DEFRA) ecosystem models**
- **Evaluation of models including (DEFRA) ecosystem models**
- **Understanding mechanisms of climate variability and change, especially (DEFRA) thermohaline circulation**
- **Understanding climate processes**
- **Detection and attribution of climate change**
- **Seasonal to medium-term climate prediction (DEFRA, Nicholls, Sutton)**
- **Agriculture (DEFRA, Trenberth); construction (Trenberth); fisheries operations and research (DEFRA, Trenberth); flood management (DEFRA); , insurance (Karl); ship routing (Karl)**
- **IPCC Assessments and other input to policy**

What data sets will be most useful in the future?

- **SST analyses with higher spatial and temporal resolution and low bias (Reynolds). Regional high-resolution SST will benefit shelf-seas ecosystem research (DEFRA) though it requires policy justification (DEFRA)**
- **Improved SST, MAT in poorly sampled regions (New)**
- **Must maintain separate *in situ*, satellite and blended SST with bias adjustments (Jones, Nicholls)**
- **SST datasets using satellite microwave retrievals will be very useful where high resolution is less important than improved coverage (Reynolds). Cloud-immune SST analysis would help maximise return on investment in AATSR (DEFRA).**
- **Sub-surface ocean temperature analyses (Christy, DEFRA, Stott) with formal error estimates (Karl, Reynolds) and consistency with SST (Karl, Reynolds, Trenberth). Subsurface ocean temperature and salinity would help maximise return on investment in ARGO (DEFRA) and are important for understanding risk of major change to thermohaline circulation (DEFRA) and for ecosystem research (DEFRA), though it is uncertain what added value Met Office provides (DEFRA).**
- **Unbiased sea-ice analyses based on more data (Christy, Reynolds, Stott), though (DEFRA) unclear on need for the proposed improvements.**
- **NMAT will play an increasingly important role in understanding global temperature trends and in understanding heat transfer processes from the surface to the atmosphere (Christy)**
- **Land T_{max} T_{min} and humidity (Jones)**
- **Daily land temperature and precipitation observations, including UK (NCIC) over at least 50 years are essential for analysis of extremes (Arnell, DEFRA, Jones, Stott)**
- **The new 5km UK daily precipitation data set will be extremely useful (Arnell)**
- **Global daily precipitation needs resolution of 0.5° x 0.5° at the coarsest for hydrological purposes (Arnell)**
- ***In situ* and satellite land temperatures need rigorous blending with error estimates (Jones, Karl, Trenberth)**
- **Mean sea level pressure (Stott) though (DEFRA) unsure of purpose.**
- **Tropospheric and stratospheric temperature (Stott) in near-real time (Christy) with biases removed (DEFRA, Nicholls)**
- **Aerosols to accompany humidity (Christy)**
- **Need for construction of consistent data base of surface net radiation (long wave and short wave) (Arnell)**
- **Palaeodata for >2 millennia (Christy), blended consistently with instrumental data (Stott); could benefit understanding of extremes (DEFRA)**
- **High-resolution land-use (Christy)**

What are the current limitations of the datasets?

- **Rigorous sampling error, random error and bias error are needed (DEFRA, New, Reynolds, Stott, Sutton, Trenberth). This includes any gridded precipitation (New). Error estimates should include impacts on different timescales (Christy, Jones, Karl).**
- **Higher resolution needed (Christy, DEFRA, Reynolds) including temporal resolution for analysis of extreme events (Sutton). However SST resolution may not need to be finer than 1° for hydrological purposes (Arnell)**
- **Better space-time sampling of monthly and daily data needed (Karl, Stott, Trenberth)**
- **Producing datasets independently leads to inconsistencies (e.g. NMAT vs *in situ* SST vs satellite SST vs subsurface T; or surface vs upper air T, pressure vs wind, multivariate upper air) (Trenberth)**
- **Consistent integration of different observing systems' data is needed (Karl, Trenberth) including palaeo with instrumental data (Stott)**
- **Need better access and improved web pages (DEFRA, Jones, Nicholls). Need real-time updating and release to the community (New, Nicholls, Sutton). Access to researchers is hindered by formalities (DEFRA, Nicholls, Reynolds) and unclear charging policy (DEFRA).**
- **Need to coordinate with Inter-Agency Committee on Marine Science and Technology (IACMST), Global Ocean Observing System (GOOS), and the Marine Environment Data Advisory Group (MEDAG) within IACMST (DEFRA)**

What other sources of equivalent or similar data exist?

- **Marine data are available from the International Council for the Exploration of the Seas (ICES), Bundesamtes für Seeschiffahrt und Hydrographie (BSH), NOAA (DEFRA)**
- **Similar marine datasets are available from Australia, Japan, and US NOAA (Reynolds), from Kaplan (Nicholls) and from NCAR (Trenberth).**
- **The US NOAA SST datasets are available in real-time without restrictions (Reynolds)**
- **The US NOAA SST analysis uses the UK method in sea-ice zones and acknowledges its advantages (Reynolds)**

Table 2: Strategic remarks

- **Funding for data development should be long-term because data set development is progressive (Christy)**
- **Possibly foster international data exchange (Christy)**
- **Multiple observing systems and multiple data sets are necessary for climate analyses (Christy, Karl)**
- **The marine components of the current data sets are vital to a comprehensive understanding of climatic processes (DEFRA).**
- **The increasing value of long term, good quality, climate data sets to the UK cannot be overestimated, as we are increasingly able to relate the consequent observed changes in the physical and biological environment to the emerging picture of persistent climate change. This leads to greater understanding of a raft of processes and an increasing ability to predict future trends (and thus to prepare for such changes, i.e. safeguard quality of life and national wealth) (DEFRA)**
- **The undoubted quality of the Met Office data sets (world class) is vital in giving the correct insight into processes and change. The added value of the emphasis on quantified uncertainties in the data sets is both impressive and extremely useful to potential users, allowing future predictions to be associated with a level of uncertainty (DEFRA)**
- **The question of "most useful" is analogous to asking what limb is most important. You cannot understand climate unless you look at the whole picture. This proposal outlines some the highest priority state climate variables (Karl)**
- **The Met Office gridded data sets are an extremely important asset for hydrological research at national, continental and global scales (Arnell)**
- **Climate monitoring is important to understand trends and events as background to policy (DEFRA)**
- **Common gridding may be over-idealistic (Jones)**
- **The outputs are of more interest than the infrastructure (DEFRA)**

IV. Reviewers Comments

The comments are as we received them though we have carried out some minor reformatting.

Nigel Arnell

Met Office Global Gridded Climate Datasets

Professor Nigel Arnell
Department of Geography
University of Southampton

My comments come from a hydrological perspective.

Land surface data

- The variables that are particularly important for hydrological studies over a large geographic domain are temperature, humidity, windspeed, net radiation and of course precipitation. Of these, net radiation appears currently to be the largest gap: data sets currently contain cloud cover or sunshine hours. There is a need for the construction of a consistent data base of surface net radiation (long wave and short wave).

- Daily precipitation is of course extremely important in hydrological modelling. The new5km UK data set will be extremely useful. What is the resolution of the proposed new global daily precipitation data set? For hydrological purposes a resolution of O.5xO.5° at the coarsest is necessary-

- I strongly support the reconstruction of past temperature and precipitation data sets

Sea-based

- Data on sea surface temperature is important in understanding the variability of hydrological behaviour over time: there are links between Atlantic sea surface temperature and UK streamflows. It is therefore extremely important to maintain this data set, but a very fine resolution (1 guess < lxl^{O}) is arguably not necessary. What is the proposed resolution of the "high resolution" SST data set for the Eastern Atlantic and the UK?

Overall

The Met Office gridded data sets are an extremely important asset for hydrological research
at national, continental and global scales.

Consultation Exercise on Met Office Global Gridded Climate Datasets

John Christy

Overall:

This first part is more of a sermon whose themes can be interwoven into the summary.

In the US, we've recognized the need for what I call "Operational Research" which is a type of research that blends real-time monitoring and continuous reanalysis of time series to produce the best, updated products for all aspects of weather and climate. The key point here is that this is a continuing effort, thus requires permanent funding. Funding agencies must realize that today's climate quality datasets are the latest and best versions of the raw data, but are rarely "final" products. This sounds self-serving, but it is absolutely true as evidenced by climate work at Hadley and elsewhere over the past 30 years. In some way the Strategic Plan should acknowledge that this is the real picture of a good portion of climate research without sounding self-serving or frightening.

Several time series of climate variables are produced around the world, but in truth, only two "centers of excellence" exist and they bear the lion's share of responsibility for providing the wide spectrum of best-available climate data records. The first is organized by NOAA in the US and the second is the Hadley Centre. Rather than "Our datasets are key inputs to the assessments made through the IPCC" it would be more accurate to say "Our datasets, including some unique to science, are utilized to advance the understanding of climate through fundamental research world-wide and in international assessments such as the IPCC."

Some funders may complain that the US or other country produces a particular dataset, thus the Hadley Centre should not. The scientifically defensible fact is that a minimum of two independently produced versions of any dataset is necessary for climate analyses. Often, the most significant advances in improving climate datasets arise from these apparently redundant activities. (e.g. Radiosondes: LKS vs. HadRT, sub-surface temps: Levitus vs. Hadley etc.) The US (if I can speak for them) strongly urges production and continuous evaluation of these datasets by those outside of the US and recognizes that the Hadley Centre maintains the highest level of expertise to perform this.

Specifics

Pg. 1. 1. "Validation of modeled climate variability and change" is actually impossible, should be "Evaluation of modeled climate variability and change."

Pg. 2, 1. "updating and release in near real time" should capture the idea of continuous evaluation and improvement … perhaps "upgrading, updating and release in near real time"

Pg. 2. 2.1.1 NMATs are one of the unique climate datasets from Hadley (should note this) and that it will play an increasingly important role in understanding global temperature trends and in understanding heat transfer processes from the surface to the atmosphere.

Pg. 3. 2.1.2 Note that variability in humidity and aerosols confound infrared signals. The Hadley Centre at some point will need to produce aerosol datasets (troposphere and stratosphere) along with the goal of producing tropospheric humidity (2.2.2). Traditionally, we think of climate datasets as datasets of response variables (temperature, wind), but it will likely unfold that climate datasets will include datasets related more to forcing (aerosols, cloud etc.)

Pg. 4 2.1.4 Systematically-produced sea ice extent is a goal here. In the US we have different groups using different algorithms for the NH and SH separately. A global dataset produced from the Hadley Centre with a single algorithm would at least bring global, interannual consistency to the process.

Pg 3, 2.1.3 and pg. 6, 2.2.2. Fundamental aspects of the physical climate system are now gaining importance in evaluating climate model performance. Bulk quantities such as sub-surface ocean temps and tropospheric temps allow for the analysis of countable budget quantities such as joules. These are absolutely vital for climate model evaluation purposes because basic energy may be tracked through the

system in simple energy-balance models to the most sophisticated coupled GCM. It may well be that the Hadley Centre should be the first to produce time series of joules in the fluid climate components (upper ocean, tropospheric, stratospheric etc.)

Pg. 6, 2.2.2. It should be noted that a near real-time component of the upper air dataset is important and will take a little effort to perform (I see it is mentioned in general pg. 2, 1. Objectives) Thus the Hadley Centre should have a goal to produce monthly or quarterly global estimates of various bulk quantities. Also, the Hadley Centre is taking the right approach in producing a version of HadRT which is independent of other datasets and whose corrections are site-specific and level-specific. Understanding the minute changes in the vertical structure of the atmosphere requires this approach.

Pg. 7. 2.4 At a recent US meeting on paleoclimate priorities, it was noted that with current evidence of Bond-cycle type variations (1500 years) that a reconstruction back at least two millenia would provide a better context to judge current changes.

Pg. 7, 2.5 Error estimates should include a rigorous evaluation of their impact on trends (e.g. Tett et al. in preparation.)

Pg. 8. 3. Should political agreements between nations to freely share data be mentioned as one aspect of monitoring?

Non specific comments.

I did not see any reference to the possibility of generating finer-scale datasets of those now being produced (e.g. surface temps to 2.5 deg. grids) These kinds of activities need funding too if they are being contemplated.

The Met Office people sit on many committees and write many reports which evaluate climate monitoring systems, evaluate climate datasets, synthesize information for policymakers, and generally push paper. Should any of these be mentioned as part of the Dataset Development activities?

I believe the next emphasis in assessing errors of land surface temperatures will be the effect of land-use changes (both small spatial changes like urbanization to large-scale like massive agriculture and deforestation). Pielke Sr. and Tom Chase have published papers on global scale circulation (and thus temperature redistribution) changes due only to large-scale land use changes. To understand these impacts, specialized datasets of high resolution land use (photography or satellite based) will be needed. I suspect the Hadley Centre will want to be in this aspect.

DEFRA

Dr David Parker
Hadley Centre
Met Office
London Road
Bracknell
RG12 2SY

Date 13 March 2003

Dear David

Review of Dataset Development Strategy

Thank you for the opportunity to participate in your consultation exercise on climate datasets.

I have consulted colleagues in other divisions of defra, and received responses from David Richardson (Flood Management Division) and John Lock (Fisheries and Aquatic Science Unit).

David Richardson has not filled in the templates, but comments that "We would certainly welcome the work on high resolution UK rainfall data and digitisation of long term rainfall records."

John Lock has provided completed Tables 1 and 2 (attached). Table 1 largely reflects the views of the CEFAS oceanographers, and Table 2 includes views of the IACMST MEDAG chair. John also makes the following general comments:

1. The Met Office should improve communication with end-users concerning what data products can be supplied.
2. There is a need to clarify the conditions of data use, particularly for the defra agency CEFAS, which is not recognised as an academic research organisation.
3. It would be helpful if the strategy could include a statement on the Met Office's charging policy for data which are to be used for research.

We have a number of comments from Global Atmosphere Division.

2.1.1, last para – I am not clear about the policy relevance of producing a high resolution SST dataset for East Atlantic and UK coastal waters. Will this help to reduce uncertainty in future climate predictions?

2.1.2 The possibility of an SST analysis which does not suffer from cloud-clearing problems is a potentially valuable development and we would be pleased to see this. This would help maximise return on the investment defra has made in AATSR.

2.1.3 We would welcome an enhanced historical dataset of sub-surface temperature and salinity, partly because it would put ARGO data into a longer-term perspective and hence maximise the return on the investment we are making in ARGO.

Other comments have been inserted in the Table 2, attached. I hope these comments are helpful and we will be happy to discuss them if you wish.

Yours sincerely
Dr Cathy Johnson
Climate Change and Ozone Research Manager

Direct Line GTN 3533 5226

Fax 020 7944 5219

Email cathy.johnson@defra.gsi.gov.uk

Consultation Exercise on Met Office Global Gridded Climate Datasets

John Lock (Fisheries and Aquatic Science Unit).

Note to Reviewers

This template has been designed to assist you in preparing your review, and to assist us in collating and reacting to the reviews. Please use either Table 1 or Table 2: if you are a direct user of the datasets, e.g. a research scientist, Table 1 is appropriate; if you are an indirect user, e.g. a policymaker, Table 2 is more suitable. It may be appropriate to repeat a Table for different datasets, e.g. sea surface temperature, radiosonde temperatures, as these have different users and applications.

Table 1. Assessment by direct users of the datasets (eg research scientists)

Who uses the datasets?	**CEFAS-EQ Physical oceanographers/modellers CEFAS scientists studying the impact of temperature and other oceanographic changes on fish stock recruitment**
What do they use them for?	**In collaboration with scientists at the Met Office we have used the Global SST database to address particular research questions on the regional and global climate, such as: "How were SST's different during the period associated with 'Warming in the North'" and "Does the change in SST in the Nordic Seas over the last 4 decades support a slowing THC hypothesis?". Ref fish stock research, uses include development of models of fish population dynamics that incorporate processes such as water movement and temperature. Other research is aimed at gaining a better understanding of the complex interplay of environmental and biological factors affecting recruitment success.**
What datasets will be most useful to direct users them in the future?	**Apart from the continued infrequent use of the global dataset (as above) the planned high resolution SST data-set for UK waters will be of direct interest for current research in EQ, which is concerned with the ecosystem processes in the shelf sea, and presents a clear opportunity for collaboration.**
What are the current limitations of the datasets?	**For studies of ecosystem processes on the UK shelf the temporal or spatial resolution of these marine datasets, and lack of sub-surface information has meant that their direct use is limited. Lack of clarity on the issue of charges for the use of Met Office data, has meant that other sources (eg/ US, Germany) for marine and land based climate data would normally be the first resort. In addition the ease of access with web based data selection and preview plotting has made other (NOAA-CDC in particular) sources more**

	attractive in the past.
What are the applications of the datasets in their research?	**Testing hypotheses and examining the global scale forcing of regional ocean climate variability.** **Further uses would include validating and forcing ecosystem models and understanding interannual – decadal variability in the shelf seas.**
What other sources of equivalent or similar data exist?	**ICES, BODC, CRU-LINK (for climate model scenario data from the Hadley Centre), NOAA-CDC, BSH, ECMWF (from BADC), and CEFAS inshore temperature datasets.**

Table 2. Assessment by indirect users of the datasets (e.g. government departments, policymakers)

What use is made of outputs based on the data - (e.g. climate assessments such as IPCC, funded projects for specific outputs)?	**Defra has wide policy responsibilities and global grided climate datasets by many divisions, e.g. Global Atmosphere (GA), Flood Management (FM), Marine & Waterways (MWD), Fisheries (Fish III)& Agriculture.**
What relative importance is attached to these research outputs?	**GA supports the Hadley Centre and the development of a scientifically robust climate change model. Outputs are highly relevant to FM e.g. coastal squeeze, MWD health of the marine environment & Fish III because of changes in fish and shellfish stocks.** **The resolution of such data is often different for different policy customers.**
What are future needs likely to be?	1. The marine components of the current datasets are vital to a comprehensive understanding of climatic processes. 2. The increasing value of long term, good quality, climate datasets to the UK cannot be overestimated, as we are increasingly able to relate the consequent observed changes in the physical and biological environment to the emerging picture of persistent climate change. This leads to greater understanding of a raft of processes and an increasing ability to predict future trends (and thus to prepare for such changes, i.e. safeguard quality of life and national wealth). 3. The undoubted quality of the Met Office datasets (world class) is vital in giving the correct insight into processes and change. The added value of the emphasis on quantified uncertainties in the datasets is both impressive and extremely useful to potential users, allowing future predictions to be associated with a level of uncertainty.

	4. The full usefulness of the Met Office work in the climate dataset area has yet to achieve its full potential value for the UK, due to the inadequate general awareness of what is available and a lack of clarity, for potential users, as to the cost of accessing and using such datasets. **There is a need to engage with the IACMST GOOS and MEDAG groups because of their co-coordinating role.**

Table collated by DEFRA Global Atmosphere Division.

Dataset referred to in "Draft Strategy"	What use is made of outputs based on the data (eg climate assessments such as IPCC, funded projects for specific outputs)?	What relative importance is attached to these research outputs?	What are future needs likely to be?
2.1.1 Marine datasets – *In situ* surface data	IPCC assessments Data needed for detection and attribution studies – uncertainties in attribution could be reduced if datasets were better?	1	Present needs expected to continue. Becoming more important if medium-term predictions become feasible?
2.1.2 Marine datasets – Satellite sea surface temperature (SST) data	IPCC assessments Climate change detection	1	
2.1.3 Marine datasets – Sub-surface temperature and salinity	IPCC assessments Detection and attribution studies Important for understanding risk of major change to THC. Not clear what added value Met Office provides, in addition to NERC role in ARGO	1	THC likely to be a major uncertainty for some time
2.1.4 Sea-ice extent	IPCC assessments Validation of climate models Not clear what need for the proposed enhancements – we already have published comparisons of models vs obs	2	
2.2.1 Land-based in situ datasets - Surface	IPCC assessments Detection and attribution Model validation	1	
2.2.1.1 Surface climate extremes	IPCC assessments Detection and attribution Model validation	1	Need for this will continue for a long time
2.2.2 Land-based in situ datasets - Upper Air	IPCC assessments Important to resolve differences between satellite vs radiosaonde data	1	Need will continue at least until models reproduce observed variations in lapse rate
2.3.1 Blended land and marine data sets - Temperature	IPCC assessments Not clear that the blended product has a scientific, rather than presentational, value?	1	
2.3.2 Blended land and marine data sets – Mean sea level pressure	Not clear what the value of HadSLP would be. For model validation?	2	
2.4 Integration between paleo and instrumental data	Model validation Will this enable better understanding of extremes?	1	
2.5 Error estimates	Generally important for measurements of any kind	1	
3. Monitoring climate	Important to understand trends and events as background to policy	1	
4. Infrastructure	A means to an end – it's the end which interests us!	2	

1= top priority
2= lower priority

Consultation Exercise on Met Office Global Gridded Climate Datasets

Phil Jones

Bruce,

David Parker send me the draft document just before Christmas. I didn't find the template that useful for comments. I have put my comments into the attachment. If the format is a problem then I'm probably not the right person to be responding. I am a user of HC/MO products but I'm also a developer and work closely with them as you know. I would support almost all things they are doing on dataset development.

If you want to put the comments into a different form please get back to me.

Cheers
Phil

Consultation Exercise on Met Office Global Gridded Climate Datasets

Response template

Responding as a Scientist (Phil Jones at CRU). Not always easy to respond by datasets but I've tried this then decided to add comments at the end as the questions asked don't seem that appropriate to the issues I want to discuss.

Table 1. SST datasets (in situ)

Who uses the datasets?	**Scientists and combining with land temperatures**
What do they use them for?	**Monitoring climate and numerous different types of climate analyses**
What datasets will be most useful to direct users them in the future?	
What are the current limitations of the datasets?	
What are the applications of the datasets in their research?	
What other sources of equivalent or similar data exist?	**NCDC produce similar SST data in real time**

SST data - need to keep the in-situ measurements going, as these are the ground truth for satellites. We will be in a complete mess without the in-situ. Various additional data are being digitised which will improve coverage in the WW periods and in the 19th century, and at other times also.

It is vital to keep the in-situ and the satellite SST estimates separate as well as producing a blended product. It is also vital that any Satellite SST product be bias corrected so that absolute values agree with the in-situ data.

Land temperatures. **Inclusion of the max/min temperatures will prove useful, but this will require some care with the QC. Improvements to the surface humidity dataset will also be useful, but this will take time and quite a bit of effort.**

Surface Climate Extremes. **There is clearly a need to work with others that is recognised in the text. For daily precipitation it is likely that a lot of data would be available through the GPCC at DWD, but it may be difficult to get them to release the data. The key to studying changes in distributions and extremes is the observational data. The text places emphasis on HadRM3. This might be OK for the future, but analysis of the observational record is needed. The EMULATE project (which I'm co-ordinating) should develop the longest record of daily temperature and precipitation for the analysis of extremes.**

It would be good if the NCIC work can be extended back in time. A lot of daily precipitation data is digitised. Selective work should be able to augment this to enable gridded datasets to be developed (both temperature and precipitation).
A lot of related work has been undertaken in the Alps. I'm we can learn from this and undertake this work in the UK, where the orography is somewhat less demanding.

It is good to see that links are being made with groups in the UK and in Europe. I would think that emphasis should be on Europe, before extension to other parts of the world.

Blending. **The blending of temperature is a reasonable idea. This needs to be done in co-operation with CRU and NCDC. NCDC have a lot of experience in this area, but their results need rigour and documentation. There is a need to bring these diverse datasets together.**

Error Estimates. This is an important aspect to include but it needs to recognise that there needs to be consistency between the estimates on different timescales. If decadal series are plotted then the errors need to be calculatd for this timescale.

Monitoring. **Making use of the SYNOP data to produce more CLIMAT messages needs to be undertaken with extreme care. If the additional data produced just adds additional data in already dense regions of the world such as Europe and North America then I would question whether the effort was worthwhile.**

Infrastructure. **Common gridding algorithms is probably a little idealistic.**
I can't see the same method being used for temperature and precipitation, because their spatial structures are so different.

General Points. **Need to make datasets available with improved web pages. Automation of some sort is clearly the way to go, but it does need some checking every now and then.**

Consultation Exercise on Met Office Global Gridded Climate Datasets

Tom Karl

David,

Please see the attached table. All this work is very critical to both scientists and policy makers. You plans are right on target.

Tom [Karl]

****

Note to Reviewers

This template has been designed to assist you in preparing your review, and to assist us in collating and reacting to the reviews. Please use either Table 1 or Table 2: if you are a direct user of the datasets, eg a research scientist, Table 1 is appropriate; if you are an indirect user, eg a policymaker, Table 2 is more suitable. It may be appropriate to repeat a Table for different datasets, e.g. sea surface temperature, radiosonde temperatures, as these have different users and applications.

Table 1. Assessment by direct users of the datasets (eg research scientists)

Who uses the datasets?	**These data sets are used extensively by the climate research community to identify climate and global change. They are also in demand by private sector users for numerous applications**
What do they use them for?	**A few examples of private sector usage includes ship navigation and optimal routing, insurance and re-insurance analysts trying to assess natural hazards and risk. The scientific community uses these data for IPCC and other assessments and numerous evaluations of climate model simulations and improved understanding of climate processes.**
What datasets will be most useful to direct users them in the future?	**All the data sets listed are critical for understanding climate. The question of most useful is analogous to asking what limb is most important. You cannot understand climate unless you look at the whole picture. This proposal outlines some the highest priority state climate variables.**
What are the current limitations of the datasets?	**The current limitations of the data sets include spatial and temporal sampling, integration with multiple observing systems for the same parameter and assessment of time-dependent biases in a quantitative manner (and attempts to narrow these bounds where such assessments have been made).**
What are the applications of the datasets in their research?	**The applications of the data sets are numerous as indicated above.**
What other sources of equivalent or similar data	**Each data set listed has alternatives. The 11th**

exist?	**principle of climate monitoring states that multiple observing systems and multiple data sets are required to ensure adequate confidence of the results. The data sets listed by the UK Met Office are critical and the Hadley team is coordinating their work with other international teams. For example, their marine data has long been the world standard, and it is now being integrated where the US COADS data sets. The sub surface ocean data and land based data clearly require additional teams to assess error bars. At present only one US team has assessed trends in sub-surface ocean temperatures and heat content. Similar work for climate extremes is critical and the Hadley group show leadership in the international community with respect to assembling data sets.** **We at NCDC look very favourably on the work proposed and indeed very much believe that the plan laid out by the Hadley Center is outstanding and deserves highest priority.**

Table 2. Assessment by indirect users of the datasets (e.g. government departments, policymakers)

What use is made of outputs based on the data - (e.g. climate assessments such as IPCC, funded projects for specific outputs)?	
What relative importance is attached to these research outputs?	
What are future needs likely to be?	

Consultation Exercise on Met Office Global Gridded Climate Datasets

Mark New

General Comments
The work of the HC Climate Variability Data Group is crucial for climate scientists around the world, as it is one of only a few institutions with the type of sustained core funding that enables the continued collection of data, rigorous quality control, and dissemination to the wider academic community and public. Without the types of datasets produced at the HC, research into climate change and variability would be decades behind where it is today.

Table 1. Assessment by direct users of the datasets (eg research scientists): Marine Datasets, especially SST Data

Who uses the datasets?	**The datasets** are used by research scientists and research students
What do they use them for?	**1. Boundary conditions for global and regional climate model simulation aimed at understanding mechanisms of past variability and change.** **2. As input to other observational analyses – for example, to enhance statistical estimates of precipitation rates.** **3. In combination with land near-surface temperature dataset from CRU, for analysis of regional patterns of past variability and change.** **4. To place confidence intervals on estimates of regional climate change.**
What datasets will be most useful to direct users them in the future?	**1. Any improvements to Marine Temperature datasets will be of great importance, particularly over poorly sampled regions.** **2. Near real-time updating of the dataset, and release to the user community..**
What are the current limitations of the datasets?	**1. I do not believe that there are error estimates for individual grid-boxes (or if there are, are they publicly available). Research to develop error estimates would be a great enhancement.**
What are the applications of the datasets in their research?	**See "what do they use them for".**
What other sources of equivalent or similar data exist?	**Mainly the NOAA funded datasets – e.g. the Reynolds SST analysis. But I do not know of any other attempts to generate, for instance, Marine Air Temperature datasets.**

Other comments.
An attempt to create a gridded dataset of daily precipitation is welcome, *provided*, that grid estimates are accompanied by error estimates. It may be worthwhile extending the collaboration with NCDC to include the Global Precipitation Climatology Project, who have done a lot of work in this field.

Dr Mark New
Climate Research Group
Department of Geography
Oxford University

Consultation Exercise on Met Office Global Gridded Climate Datasets

Neville Nicholls

Date: Tue, 24 Dec 2002 12:08:13 +1000
From: Neville Nicholls <N.Nicholls@bom.gov.au>
Subject: Re: Review of Dataset Development Strategy

David:

I have attached a template with my comments for the review of your data products strategy. I am not sure if my comments are appropriate - I am not really sure what sort of things you want me to discuss, nor how to fit them into the template.

My overall impression is that the Hadley Centre has done a great job in producing high-quality global data sets (especially SST and adjusted upper air data) but that these are not used as widely as they might be. I think the reasons why they are not used as widely as might be expected, is that:
1. We potential users don't know how to easily access the data (compare with the simple use of the Kaplan data set in the KNMI Climate Explorer web site).
2. Confusion about need for permission. As I note in the template, some of my uses are neither bona fide academic research nor commercial applications, so I dont know if I need to seek permission.
3. Just the need to seek permission will put off some users (it certainly worries me), because it is likely to take time.
4. I am not sure if you update in near real time, or if the current data are easily available (again, I contrast this wth Kaplan's data).

Now, I may well be showing my ignorance with the above points, but I think I am trying to say that if there can be any criticism of the Hadley Centre data sets (and I think "criticism" is too strong a word here) it might be that the communication strategy is either not such as too promote the wider use of the data (ie not just in the academic community) or that potential users such as myself don't know where we can access the data easily or if we need formal permission, so we use other data sets for which these are not

problems.

I hope this is reasonably clear, and potentially useful. Can I say again that you guys have done terrific service (where would IPCC have been without you?) but I would like to see your data sets promoted more and more easily accessible by the wider scientific community (without formal agreements or collaborations etc). Of course this may be thought of as unreasonable on my part, since you have put so much work into the data sets.

Thanks for asking my opinion. And best wishes to you and all your colleagues for Christmas and New year.

Regards,
Neville

Neville Nicholls
Bureau of Meteorology Research Centre
Postal address: BMRC, PO Box 1289K, Melbourne 3001, AUSTRALIA
Street address: BMRC, Floor 13, 150 Lonsdale Street, Melbourne 3000, AUSTRALIA
Phone: 61 3 9669 4407 Fax: 61 3 9669 4660
Email: N.Nicholls@bom.gov.au
Web page: http://www.bom.gov.au/bmrc/clfor/cfstaff/neville_nicholls.htm

Consultation Exercise on Met Office Global Gridded Climate Datasets

Response template

Note to Reviewers

This template has been designed to assist you in preparing your review, and to assist us in collating and reacting to the reviews. Please use either Table 1 or Table 2: if you are a direct user of the datasets, eg a research scientist, Table 1 is appropriate; if you are an indirect user, eg a policymaker, Table 2 is more suitable. It may be appropriate to repeat a Table for different datasets, e.g. sea surface temperature, radiosonde temperatures, as these have different users and applications.

Table 1. Assessment by direct users of the datasets (eg research scientists)

Who uses the datasets?	**BMRC Climate Forecasting Group and National Climate Centre**
What do they use them for?	**1. We have used GISST to develop an operational seasonal climate forecast system for Australia, using near-global SST patterns as predictors. We use our own current SST analyses as input for the predictions – GISST was used to develop the system though (ie, to develop the statistical forecast relationships). 2. We have also started using GISST as a data set to help us do real-time analysis of Australian climate anomalies – we identify "interesting" climate anomalies and correlate them with SSTs, to determine what might be causing the climate anomalies. None of this is "commercial", nor likely ever to be.**
What datasets will be most useful to direct users them in the future?	**1. Historical SSTs will remain crucial. 2. We would also like access to the upper air data (especially for the Australian region) adjusted for the instrumental and observational inhomogenieties.**
What are the current limitations of the datasets?	**1. A limitation for us is the (apparent) lack of routine, near real-time updating. This means we tend to use other SST analyses for the current forecasting, even though we used GISST to develop the forecast equations. 2. I am confused about the purposes for which I can use Hadley Centre data sets, and what permissions I need to obtain. I recognise that you allow bona fide academic research, and that (perfectly reasonably) you would want to negotiate with anyone proposing to use the data for commercial purposes. But these two extremes do not adequately represent my uses of the data. I use SSTs to try to work out what might be causing current climate anomalies, and provide such information to politicians, media, bureaucrats – this is real-time climate monitoring and analysis, rather than academic research. It is not clear if you are happy for me to use the data in this way. I think because of this problem, that the Hadley Centre data sets are not used as widely as they deserve to be. Could I suggest that one way to enhance their use would be to place them on a ftp site and allow anyone to download them and use them (requesting appropriate citation of course) without seeking approval unless they are to be used for commercial purposes? This would be a simpler approach for potential users. 3. Because of this permission problem, and because they are not available as easily as they might be, I increasingly use other global data sets (eg., Kaplan's SSTs) instead of the Hadley**

	Centre data sets. For instance, the Hadley Centre SSTs are not available on KNMI's Climate Explorer website (except as blended with the land temperatures, which makes them unsuitable for some of my analyses), so I use Kaplan. I think the Hadley Centre data would be more widely used, if they were a bit more easily and directly accessed and it was clear that you do not need to seek permission to use them unless you have a commercial purpose in mind.
What are the applications of the datasets in their research?	**See above**
What other sources of equivalent or similar data exist?	**Mainly Kaplan for historical SSTs.**

Table 2. Assessment by indirect users of the datasets (e.g. government departments, policymakers)

What use is made of outputs based on the data - (e.g. climate assessments such as IPCC, funded projects for specific outputs)?	
What relative importance is attached to these research outputs?	
What are future needs likely to be?	

Consultation Exercise on Met Office Global Gridded Climate Datasets

Dick Reynolds

I am answering these questions for the research scientist point of view with a focus on marine datasets, only.

Table 1a. Assessment by direct users of the SST datasets (e.g., research scientists)

Who uses the datasets?	**Research Scientists**
What do they use them for?	• **Climate monitoring and diagnostics** • **Boundary conditions for atmospheric models** • **Input data for ocean models** • **Verification of coupled model SST**
What datasets will be most useful to direct users them in the future?	• **SST analyses with higher spatial and temporal resolution as long as long-term climates biases do not increase significantly** • **SST datasets with microwave satellite SSTs will be very useful for climate where high resolution is less important than the improved coverage. IR and microwave have different error characteristics which may tend to cancel**
What are the current limitations of the datasets?	• **Information on sampling error, random error and bias error are not included** • **Access to researchers is delayed because data users must first complete a UK form**
What are the applications of the datasets in their research?	• **Answered above**
What other sources of equivalent or similar data exist?	• **Similar datasets are available from Australia, Japan, and US NOAA** • **The US NOAA datasets are available in real-time without restrictions**

Table 1b. Assessment by direct users of the subsurface temperature and salinity datasets (e.g., research scientists)

Who uses the datasets?	**Research Scientists**
What do they use them for?	• **Climate monitoring and diagnostics** • **Input data for ocean models** • **Upper ocean heat content**
What datasets will be most useful to direct users them in the future?	• **Subsurface temperature data**
What are the current limitations of the datasets?	• **Information on sampling error, random error and bias error are not included** • **US NOAA dataset is not an OI procedure so errors are not computed. Also, each depth level is computed independently even though information with depth is highly correlated** • **Surface marine data are not included in the US NOAA dataset. Thus, the NOAA SST and subsurface temperature analyses don't match at the surface.** • **Access to researchers is delayed because data users must first complete a UK form** • **Sea surface salinity data is planned from satellites. These data are not considered.**
What are the applications of the datasets in their research?	• **Answered above**
What other sources of equivalent or similar data exist?	• **Similar datasets are available from Australia, Japan, and US NOAA** • **The US NOAA datasets are available without restrictions**

Table 1c. Assessment by direct users of the sea-ice extent datasets (e.g., research scientists)

Who uses the datasets?	Research Scientists
What do they use them for?	• Input data to SST analyses • Climate monitoring and diagnostics • Boundary conditions for atmospheric models
What datasets will be most useful to direct users them in the future?	• Improvements in the removal of biases, especially satellite biases • Location and possible digitization of new sea-ice datasets to improve historic accuracy
What are the current limitations of the datasets?	• Information on sampling error, random error and bias error are not included • Access to researchers is delayed because data users must first complete a UK form
What are the applications of the datasets in their research?	• Answered above
What other sources of equivalent or similar data exist?	• Similar datasets are available from Australia, Japan, and US NOAA • No other sea-ice extent dataset in the world is appropriate for climate because satellite biases are not corrected. The US NOAA SST analysis uses the UK Method and acknowledges its advantages

Consultation Exercise on Met Office Global Gridded Climate Datasets

Peter Stott

Date: Thu, 06 Mar 2003 16:25:59 +0000 (GMT)
From: Peter Stott <peter.stott@metoffice.com>
Subject: Consultation exercise on Met Office global gridded datasets
To: david.parker@metoffice.com
Cc: peter.stott@metoffice.com

David,
Here is my response to the consultation exercise,
Peter

Who uses the datasets ?

Scientists researching detection and attribution of climate change

What do they use them for ?

For determining the causes of past climate change.

What datasets will be most useful to direct users them in the future ?

Homogenous quality controlled datasets with good spatial coverage and covering at least several decades and preferably at least 50 years. Monthly mean temperature both at the surface and through the depth of the atmosphere, precipitation, PMSL, sea ice, sub-surface ocean heat content with global spatial coverage and over at least 50 years. Uncertainty estimates on these series taking account of instrumental, sampling and correction (eg bucket correction) uncertainties. Consistently merged paleo and instrumental timeseries. Daily temperature and precipitation data with, ideally, global spatial coverage and over at least 50 years.

What are the current limitations of the datasets ?

Limited spatial coverage of daily data. Limited spatial coverage further back in time of precipitation and temperature datasets. Paleo datasets not integrated with instrumental datasets. Limited uncertainty information, particularly associated with systematic errors due to corrections, eg bucket corrections.

What are the applications of the datasets in their research ?

To better understand past climate change and variability on multi-decadal time scales.

What other sources of equivalent or similar data exist ?

Temperature and precipitation datasets from CRU. Reanalysis (NCEP, ECMWF) although problems about using these in detection work. Ocean heat content from Levitus

(even though probably flawed this is the only currently available dataset). Various paleoclimate series.

Consultation Exercise on Met Office Global Gridded Climate Datasets

Rowan Sutton

Date: Fri, 07 Mar 2003 18:02:06 +0000 (GMT)
From: Rowan Sutton <rowan@met.reading.ac.uk>
Subject: Re: Review of datasets
To: David Parker <david.parker@metoffice.com>

Dear David,

Please find attached the review you requested of your draft strategy for the development of datasets. I apologise that the review is not more in depth; it has proved to be a particularly busy time.

Best regards,
Rowan.

Review of Draft Strategy for the Development of Datasets in the Met Office
---

It is hard to overstate the importance of the work done at the Met Office in the development of global gridded climate datasets. These datasets are absolutely fundamental to climate monitoring, to a wide range of research into climate processes and climate change, and to climate prediction. Moreover in many of the most important components of this activity the Met Office is unquestionably the world-leader. It is very much in the UK's interests that this leading position be maintained, and the draft strategy provides an excellent basis for ensuring this happens.

The strategy presents bold plans for the further development of existing activities and also for the initiation of new activities. The aim to provide - in time - error estimates on all datasets is important and welcome. The updating and release of datasets in near real time will also be very valuable and will undoubtedly increase the exploitation of the data by the research community. The development of higher temporal resolution datasets is an important response to the pressing need to better understand the statistics of extreme events. Overall this is clearly a strategy moving the right directions on many fronts. I trust that it will receive the full backing it deserves.

/--\
Dr. Rowan Sutton	Tel: +44 (0)118 378 8337
Royal Society Research Fellow,	Fax: +44 (0)118 378 8316
Centre for Global Atmospheric Modelling,	
Dept of Meteorology, Univ. of Reading,	email: R.Sutton@reading.ac.uk
PO Box 243, Earley Gate,	www: www.met.rdg.ac.uk/~rowan
Reading RG6 6BB, U.K.	
\--/

Consultation Exercise on Met Office Global Gridded Climate Datasets

Kevin Trenberth

Your ref D/DG(R&T)17/2/3/1 (MSG)
23 December 2002

Table 1. 2.1 Marine data sets

Who uses the datasets?	**Researchers, IPCC, many applications such as fisheries,**
What do they use them for?	**Tracking climate variability and change; model validation, tracking the ocean for fish, driving atmospheric models (SSTs)**
What datasets will be most useful to direct users them in the future?	**Complete gridded data, good error estimates, consistency among fields**
What are the current limitations of the datasets?	**Datasets produced independently, leads to inconsistencies (e.g. NMAT vs SST vs satellite vs subsurface), somewhat ad hoc corrections, error bars need improvement.**
What are the applications of the datasets in their research?	**Documenting climate variability, associations with ENSO, NAO etc, climate change, driving atmospheric models, attribution.**
What other sources of equivalent or similar data exist?	**Various datasets from NCAR and NCDC, and other sources. On surface T and SST these have been primary.**

Table 1. 2.2 Land-based in situ data sets; 2.3 Blended datasets

Who uses the datasets?	**Researchers, IPCC, many applications such as agriculture, building codes**
What do they use them for?	**Tracking climate variability and change; model validation, climate extremes**
What datasets will be most useful to direct users them in the future?	**Complete gridded data, good error estimates, consistency among fields**
What are the current limitations of the datasets?	**Datasets produced independently, leads to inconsistencies (e.g. sfc vs upper air, pressure vs wind, multivariate upper air) somewhat ad hoc corrections, error bars need improvement.**
What are the applications of the datasets in their research?	**Documenting climate variability, associations with ENSO, NAO etc, climate change. Monitoring.**
What other sources of equivalent or similar data exist?	**Various datasets from NCAR and NCDC, and other sources. On surface T these have been primary.**

Comments:

This group has done very important work in the past and has led the way in the realm of surface temperature datasets with its partnership with UEA. The current proposal extends and upgrades these efforts and is welcomed by the community. A unique facet of the approach, which is both a strength and a weakness, is the independence of “models”. In fact models of some sort have to be used and it would be argued that here they are statistical. However, it is less clear that the latest and most advanced methods will be used. A key point is that all methods involve assumptions and it should also be clear just what those assumptions are. For instance, in areas of missing data, are structures imposed based on covariability in the recent record? What does this imply for longer-term reconstructions? The independence of climate models is a strength, then, because it allows independent model validation. It is a weakness as it does not exploit multi-variate relationships.

The objectives are excellent. The main comments then are on implementation. As noted above, the datasets are largely dealt with independently. An advantage of a physically based model is that physical relationships can be exploited. A classical example is the analysis of weather maps in terms of isobars resulting in a pressure map, where winds are extensively used to set gradients, and current weather, cloud, precipitation etc is used to draw fronts etc. It is less clear that the analysis of averages results in the same product as the average of instantaneous analyses. I.e. 6 hourly isobaric analyses vs monthly mean. Corrupt or missing data can also make for physical inconsistencies, e.g. among SST and NMAT, SLP and wind, satellite vs in situ. I personally do not think any satellite product alone should be produced without full in situ data as well. It is either done implicitly or explicitly. So I suggest 2.1.2 should be a combined product. On 2.1.3 why isn’t sea level also included? Challenges exist in 2.1.4 because of changing sensors and their characteristics. In 2.2.1.1, daily precipitation is not adequate for many purposes as most of the time it does not rain. I urge that the challenge to acquire hourly data be taken on to properly address frequency and intensity of precipitation and hence extremes and runoff. Also, there is a wealth of radar data on precipitation patterns that is not exploited here. On upper air data, 2.2.2, I suspect reanalysis and multivariate approaches will be most useful and in any case result in quality flags on data that should be useful and should be exploited. Homogeneity of humidity will be difficult. There remain substantial questions about the satellite based land surface temperatures that are planned for use under 2.3.1; some of these relate to things like the diurnal cycle and also the coarse monthly time scale (vs assembling monthly means from daily). In making error estimates, many errors and data are not independent and assuming “random” is not appropriate. Temporal and spatial persistence must be factored in.

All of the statements under section 3 on the goals for monitoring are endorsed. The main questions under section 4 on infrastructure are the statistical methods to be used and whether they really are optimal (see work at Lamont by Kaplan et al. for example).

In making comments and suggestions above, the tendency is to be critical. This should not be misinterpreted to say that there are flaws in what is being done and proposed. The activities as proposed would be invaluable in any case but perhaps they could be even more so.

In my own group, our objective is to exploit datasets up to but not beyond their capabilities. We spend a lot of time evaluating datasets and getting to know their strengths and weaknesses. We wish the datasets we use to be of known quality. This does not mean they have to be perfect, instead it means we must know the weaknesses and assumptions. We will gladly use the UKMO datasets but urge that their characteristics be well documented.

Kevin Trenberth
Head

Climate Analysis Section
National Center for Atmospheric Research
Boulder CO 80307

Reviewers of Climate Variability Group Data Set Development strategy

Dr Nigel Arnell (U Southampton, UK) **nwarnell@soton.ac.uk**
*Dr John. Christy (U. Alabama, USA) **christy@atmos.uah.edu**
*Prof. Phil. Jones (CRU, UK) **p.jones@uea.ac.uk**
*Mr Tom Karl (Director, NOAA National Climate Data Center, USA) **Thomas.R.Karl@noaa.gov**
Dr Mark New (U Oxford, UK) **mark.new@geog.ox.ac.uk**
Dr Neville Nicholls (BMRC, Australia) **n.nicholls@bom.gov.au**
Dr Richard Reynolds (NOAA, USA) **Richard.W.Reynolds@noaa.gov**
Dr Ben Santer (Lawrence Livermore Lab., USA) **santer1@llnl.gov**
*Prof John Harries (Imperial College) **j.harries@ic.ac.uk**
Prof. Jagadish Shukla (COLA, USA) **shukla@cola.iges.org**
Dr Vasily Smolianitsky (Arctic and Antarctic Research Institute, St. Petersburg, Russia) **vms@aari.nw.ru**
Dr Peter Stott (Hadley Centre, Met Office, UK) **peter.stott@metoffice.com**
Dr Rowan Sutton (U Reading, UK) **rowan@met.reading.ac.uk**
*Mr Alan Thomas (WMO-GCOS) **gcosjpo@gateway.wmo.ch**
*Dr Kevin Trenberth (NCAR, USA) **trenbert@ncar.ucar.edu**
*Mr David Warrilow (DEFRA) **David.Warrilow@defra.gsi.gov.uk**

Cover email to * addressees

Dear

Please see the attached. We would be grateful if you could review, on behalf of our Ministry of Defence funders, our strategy for 2003-7 of the Climate Variability Group in developing datasets. The attached letter requests your reply by 20th February 2003 if at all possible. It would be appreciated if you can include comments on the policy relevance of our work. Please reply to me at david.parker@metoffice.com

Thanks if you are able to do this; and best wishes for 2003

David Parker

Cover email to other addressees

Dear

Please see the attached. We would be grateful if you could review, on behalf of our Ministry of Defence funders, our strategy for 2003-7 of the Climate Variability Group in developing datasets. The attached letter requests your reply by 20th February 2003 if at all possible. Please reply to me at david.parker@metoffice.com

Thanks if you are able to do this; and best wishes for 2003.

David Parker