Final session summary

Peter Thorne

Updated with notes from Kate Willett and (to be incorporated – powerpoint may be modified) Nick Rayner

Overarching principles

- Authoritative databank holding
- Multiple redundant data products generated from databank
- Common benchmarking assessment (cyclical)
- Development of tools to aid the end user
- Provision of all aspects to public as openly and transparently as feasible

Next steps

- Talks will be posted online unless there are objections
- Organisers and rapporteur cadre will write up formally for submission as report to Bulletin of American Meteorological Society and WMO Tech Doc
- Both will be distributed for limited comment period. They will contain more information than this powerpoint that exists solely as a very high level summary primarily for workshop participants although it is made generally available
- Host related materials and documents on surfacetemperatures.org, revamp frontpage and maintain blog

Building momentum

- Data offers
  - NCDC offered their three constituent databases
  - RIHMI (World Data Center B) offered to reconcile database discrepancies with NCDC
  - NCDC offered to look at boxes in their basement
  - Galaxyzoo happy to explore digitisation of existing images. Philip Brohan to help identify possibilities. NCDC to discuss and scope their substantial images archive with galaxyzoo.
  - ACRE data rescue could naturally form part of databank.
  - The major reanalyses efforts have substantial synoptic holdings. Adrian Simmons offered to be initial point of contact.
  - Steve Worley offered NCAR archives

Building momentum

- More data offers
  - UK images and archives will be scoped (contact?)
  - Albert Klein Tank offered data from his European project
  - Indonesia are collaborating with KNMI but currently only gridded data planned to be released
  - China may be able to offer 300+ stations pre-1950 recently digitised. Qingxiang Li is contact
  - Argentina may be able to offer additional data

More momentum

- Effort offers:
  - Interpolation working group effort (Tom Peterson, Adrian Simmons, Liz Kent, Tom Smith, Peter Guttorp, Michael de Podesta, Ole Tiveto, Antonio Possolo, Shawn Smith, Robert Rodhe) – will expand on this aspect in formal write ups.
  - NCDC to run algorithm on any improved databank and benchmarks to provide initial performance estimate
  - COST to provide expertise where practical?
  - NCDC offered to provide station summaries that include statistical summaries such as percentiles records etc in addition to climatologioes to countries who submit data
  - Climate code keen to help on benchmarking (could ensure double blind) and other aspects but funding dependent.

Openness and transparency

- To the extent feasible all aspects need to be open and transparent
  - Databank – provenance, version control
  - Datasets – code, intermediate products, published methodology
  - Benchmarking – delayed mode release of the analog worlds (applies to both interpolation and homogenisation benchmarking)
  - Publication culture with supporting material
  - Exceptions (e.g use of restricted data, code with restrictions) clearly outlined.

Databank

- International. Requires collaboration.
  - Central databank maintained from one location and mirrored
  - OR common format, distributed databank with common portal
  - Various views given in support of both: Agreed to let databank oversight group decide ensuring relevant voices heard (voices: Matt Menne, John Christy, Bryan Lawrence, Jay Lawrimore, Peter Guttorp, Amy Luers)
  - Need for multi-institution support for expertise and buffer factor.

Databank

- Monthly, daily and sub-daily resolution data.
- Consists of several stages (was levels in White papers but confusion with satellite levels should be avoided) from image / hardcopy to a finalized combined source
  - 0 – raw (image, hardcopy, digital count)
  - 1 – digitised
  - 2 – common data format
  - 3 – augmented (station ids, source merging et al.)
- Provenance and versioning eventual aim, but accept that must start first with something not ideal. Default unknown provenance and databank version 0 with what we have (see data offers) so we can start. Incremental procedure. IMPORTANT: Just because we may release a version doesn't imply job is done!
- Version controlling needs to be in-built from the start. First release can be what we have but must be given a version number for both release and individual databank elements.
- Metadata is key. Must be flexible but structured so as to be largely machine readable.

Databank - metadata

- Need to define what the distinct metadata elements we want are otherwise just a term with multiple interpretations.
  - Need to separate and clarify all different uses of the term metadata somehow to avoid confusion. There is raw data metadata and processing metadata and benchmarking metadata (at least!)
  - There are existing WMO usages of these terms that if possible we should leverage.
  - Metadata may consist of core (location info) and extendable (history, practices)

Data recovery

- Getting extra data 1. data that is digital but not available
  - Hearts and minds
  - Survey to ascertain reason why data can't be made available – will differ by country. Informs what will be most effective approach on a case by case basis.
  - Economic benefit must be quantified and case must be made where this is the reason
  - Provide something in return – value added data such as the NCDC station statistics that can be produced straight away and longer term any estimates following homogensiation / QC efforts
  - Utilise research relationships – students and post-doc exchanges?
  - Could have a sunset clause (10 years) after which expectation is to release – modify res. 40?

Data recovery

- Getting extra data 2. data that is not digital
  - For known unknowns at least index what we could add even if we make no prioritisation (such prioritisation would be subjective as priorities will be end-user application specific).
  - Crowdsourcing to be investigated (sitting on many images)
  - Have World Weather Records annually? Agreed to seek opinion of countries when we send the next decadal request next year.
  - WMO mechanism more successful in many, not all, cases? Use many different approaches for this problem – no one-size fits all solution will exist.

Updates

- Real time updates
  - Define and request reporting of a daily “CLIMAT”. Timing is fortunate within WMO programme. Jeremy Tandy and Stephan Bojinski agreed to act as focal points to push forwards. Jay Lawrimore and Tom Peterson to help.
  - Delayed mode transmission option over GTS?
  - Standardize exchange of sub-daily (METAR and SYNOP) and sub-hourly
  - Basic quality checks (but never delete!)
  - Metadata requirement on CLIMAT messages? Again, timing could be fortunate.

Data use

- Usage restrictions
  - Realistically we are not suddenly going to have open unrestricted access to all withheld data. In some areas this is the majority of the data.
  - Databank could have public and private areas. Trade off …
    - Implications for traceability vs …
    - Deliberately handicapping ourselves – making products that are deliberately bad by not using the information held in commercially or politically sensitive non-public facing data. Could be endemic issue in certain areas.

Data use

- Usage restrictions (cont.)
  - May mean that any dataset that leverages the private area would be encouraged (but not mandated!) to have two versions … one using the public facing data only and the other all data? [But noted this may be confusing or provide misleading indications]
    - WMO res 40 restricts to non-commercial same here? Revisit WMO res 40 so it covers climate services data?
    - What constitutes non-commercial? Climate services commercial? Consultancy by Met Services?
    - 70% of European data would be unavailable if research only clause removed
    - Provide uncertainties and summaries but not actual data?
    - Does mean cannot be fully open and transparent
  - Agreed steering committee in consultation with databank group to resolve whether public only or public + private areas and provide justification.
  - Regardless, does not mean that datasets cannot use restricted data, would just impact whether they fell under the umbrella of this exercise.

Dataset creation

- Global records will need to be automated procedures, regional not necessarily
- Need redundant, independent analyses.
- Role for national assessments – maybe with common software to assure continuity.
- Create an adjunct to the databank with these estimates and some assessment flags regarding suitability for purposes people might want to use them for.

Dataset creation

- Create portable tools to enable analyses
  - Create massively parallel suite of OS tools so that investigators can make their own based upon their preferences? Several choices for each step and tunable within each choice. Would need substantial help from software community.
- Challenge of daily and especially sub-daily is substantial, but recognition that we need to tackle these.
- Detection and adjustment not necessarily at the same temporal scale
- Attribution of causes and physical understanding preferable.
- How to deal with network wide changes is a difficult challenge.

Benchmarking

- Benchmarks
  - Note in many other areas of science this would be termed “software testing”
  - We don’t know what areas / time periods people will want to consider in creating data products from the databank so essential to subsample in space and time to databank.
  - Make simple to use (format, structure etc. identical to databank)
  - Derive from reanalyses / climate models
  - Range of assumptions from overly optimistic to overly pessimistic – test from too simple to destruction and ascertain when and why algorithms fail.
    - Don’t tune to any assumptions e.g. Use not just 20th Century forcings runs
    - Addition of sampling effects, breaks and trend-like artefacts, seasonal and diurnal cycles etc.
  - Need to really define the real-world maladies better.

Benchmarking

- Group defines benchmark specifications but third party produces the actual analog cases and performs assessment summary to ensure double blind.
- Benchmarking has a finite lifecycle (propose 3 years: Develop, implement, assess, wash-up, repeat …)
  - Wash-up includes dataset creators
  - Historical benchmark cases remain available
  - COST has done some benchmarking and need to leverage that experience
  - If eventually benchmarking becomes multi-variate it'll become much harder. Need realistic inter-variable relationships to be retained.

Benchmarking

- Climate Reference Network type measures with much better traceability provide a second hugely valuable check going forwards
- Need global CRN for next generation of researchers who will thank us
  - Adrian Simmons agreed to put on agenda of next GCOS/WCRP AOPC meeting for discussion

Assessment

- Benchmark performance
- Criteria that assess datasets for their quality for a particular purpose
- Depend on the purpose
- Include both scientific quality and other criteria
  - Ease of use, how documented, how traceable
  - Do they represent long term trends or not, regional detail, spatial correlations etc etc
  - Will depend on user needs, probably not a case of a single number

Interpolation

- Needed because we never have observed everywhere and never will.
- Like homogenisation, no one-size fits all solution
- Do uncertainty by covariances but users need something simpler – equi-probable ensemble?
  - If so need some sensible user guidance associated.
- Need to benchmark and assess just as for homogenisation
  - Group proposed to take forwards
  - Should work closely with benchmarking and assessment group.

Dataset acceptance

- Open provenance-type models - use existing formats – Invoke ISO standards for observations and measurements

- Desirable criteria to meet product standards:
  - reference source data list and rationale (two product types: use of databank alone is auditable; if non public ancillary data is used a reference must be provided)
  - description of quality control method
  - if homogenized, then benchmarking is required
  - a link to assessment report carried out by assessment team if available
  - peer reviewed publication
  - audit trail along with release of code used (but with **no** expectation of software support)

- Are these guidelines or requirements?
  - Requirements should only be those that are objective and represent best practice with other aspects being desirable?
  - Need to provide an estimated cost in terms of compliance in terms of money and time
  - Minimum requirements should be those that do not add a prohibitive overhead? Run risk of killing science otherwise ...

Governance – proposed very initial phase structure

Ad hoc steering group

- Peter Thorne (Chair)
- Blair Trewin
- Richard Chandler
- Kate Willett
- Xiaolan Wang
- Matilde Rusticucci
- Jay Lawrimore
- More TBC – gap analysis

Ad Hoc Steering Group tasks

- Write an implementation framework
  - Identify what is really new that needs support
  - Interim report of workshop outcomes and progress towards implementation by WMO Congress 16 in May 2011
  - Final report by September 2011
    - To present at the WCRP Open Science Conference in Denver, CO, USA
  - Develop communication strategy
- Dissolve itself upon creation of a more permanent group on advice of governance group

Governance team

- Individuals to frame governance
  - Peter Stott (Chair)
  - Peter Guttorp (Univ Washington, Stats)
  - Alistair Forbes (National Physical Laboratory, UK - metrology)
  - WMO Secretariat (Ghassem will help identify someone)
  - John Christy (U Alabama IN Huntsville)
  - Andreas Becker (DWD, GPCC)
  - Ms Saito (tentative – JMA)
  - Joseph Kimani (tentative – Kenya Met Dept)
  - Vyacheslav Razuvaev (RosHydroMet WDC B)

The Governance Team's TOR

- Within half a year come up with 1-2 page document
- The team recommends to the Steering Group the appropriate governance
  - Which would guide institutions that would be involved in various aspects of the project
- Probably not meet in person
- Keep in mind other teams, requirements and activities
- Pay attention to the White Paper and workshop proceedings
- Dissolve itself when Steering Group agrees that its task is done

Databank management

Who-

- People who understand the consequences of a proposed change who have some independence from the work on the databank
- Initial names: Jay Lawrimore (Chair), Matt Menne, Bryan Lawrence / Jeremy Tandy (TBD), John Christy, DWD rep (TBD) ... Group to augment and ensure truly global membership and multi-institution

- What—decision making entity that manages change to ensure transparency
  - Decisions are recorded for transparency
  - Changes are expected to be more frequent than in the past, thus an accountable authority needed
  - Establish guidelines for documenting provenance and versioning procedures
    - Working teams for provenance and versioning?
  - Assure continuity of processing

Benchmarking and assessment

- Individuals
  - Kate Willett (chair?)
  - Claude Williams
  - Ian Jolliffe
  - Others TBC
- Start from agreed outcomes of benchmarking and assessment discussions
- To be superseded / modified when governance group reports

Pilot phase?

- Closed 12 months after first benchmark lifecycle (2015)
- Focus needed – characterising centennial scale variability, trends and regional behaviour
- Still massive effort – monthly databank (political), algorithm development, benchmark definition and development, assessment cycle set up … (scientific)
- Define delivery focus
  - Define specific deliverables in the first year so this is seen to be moving forwards. Databank version 1 and analog worlds? Possibly others too?

Not just climate scientists

- Possible oversight from BIPM (metrology) + TIES (statistics) – Governance group to note.
  - Keeps on their radar
  - Assures relevant input
- Ensure relevant expertise on working groups
- Recognise harder for these folks to get funding.

Not just surface temperatures

- Marine data
  - Databanks consistent
  - ICOADS membership on databank oversight
- Reanalyses
  - Feedback files should form metadata
  - Ensure databank easily usable by reanalyses
  - Use of reanalyses in benchmarking / homogenisation?

Thankyou

- For coming
- For being active participants
- For productive discussions
- To the organizing committee
- To the Met Office and all sponsors (RMS, Exeter uni, NCDC, US GCOS Office in addition)
- To Kate … again