

Evaluation and Quality Control

20 participants from:

EUMETSAT, ECMWF, CSC, IPMA, ESA, JRC, MeteoSwiss, Croatian Met Service, Météo-France, DWD, ARPA Emilia-Romagna, GCOS, DMI, AEMET, Univ. of Reading, SMHI, Irish Met Service, Met Norway, NPL.

Scope of the CCCS

- **Regulation** = Bible, defining the scope (in few words). Importance of the conclusions from Helsinki workshop. Scope: the CCCS can't do all activities related to global projections or seasonal forecasting... Benefit from the existing and start from it. Need for a more comprehensive document: gap analysis, development of metrics...
- Definition of this scope will also **structure the EQC**.
- Could the EQC go as far as playing a role of a "**Regulatory Authority**" (this could entail liability, penalties between countries...)? But the CCCS **cannot prescribe** everything alone, there are other actors, stakeholders.

EQC: what is it?

- EQC is a platform (tool) but also an activity (people, expertise, interpretation). The term **platform maybe inappropriate**, EQC is a function.
- Much to be learnt from **existing** activities (what works?, what is to avoid?)
- Define **reference for quality of climate services** and promote best practises (at least in Europe): need to harmonize. Definition of international standards? Useful to leverage overall quality in the sector. **Transfer of existing practices/experience** (on some data/variables) **to other areas** (e.g. socio-economic: units, timescales, metadata, timeliness).
- Target/ambitions are **GCOS requirements** (they are “independent” and ambitious enough). Can help define overarching goals for the **products in the CDS**.

Internal and External QC

- We see the **need for an internal and an external** (independent) **component** to the EQC. Need to delineate the different roles. The approaches must be different.
- What is the role of an external EQC:
 - **Evaluate** the QC **processes**; evaluate the **efficiency** of the Service
 - Tap on **existing international expert communities** (example of GHRSSST, GEWEX...), providing incremental resources to allow them working on evaluating the CCCS

Internal and External QC (cont.)

Internal EQC:

- **QC of input data** is important; **homogeneisation** needed
- **Evaluate** data and geophysical model outputs ; assess the **fitness-for-purpose** of sectoral products (from impact/socio-economic model outputs...), with a feedback loop between providers and users. Establish **best practices**.
- Map **data flows** within the service: provide **traceability**, documentation
- There must be **agreed metrics** for the system/processes and for the end products. Stage 0 of the CCCS must construct these metrics.

- Information needs to be **associated with** an validated estimate of **uncertainty** (that has to be defined). There is a need to evaluate not only value, but also uncertainty/variability/spread: aim is “**best Probability Distribution Function**”; note that purpose guide specific requirements: stability, accuracy....
- Need **to collect and evaluate** evolving **user feedback and requirements** (interaction with “outreach”)
- Should address **scientific** and technical questions, in liaison with wider research activities
- **Completeness of the evaluation**: must be also cross-cutting, not only ECV by ECV (following GCOS) ; assess **interface** between climate model outputs and application models, including **downscaling** and **bias correction** aspects.

Open questions

- Defining quality for the “**attribution**” part of the service is by far **more complex** than for ECV.
- Evaluation **approaches** are largely different between e.g. reanalyses and projections.
- How to quantify **quality** from a **user’s perspective**?
- How to **assess and ensure quality of outreach** / external interface activities?
- How to **use existing** experience (eg IPCC ch 9, obs4mips...) –while it is recognised that there are specific aspects/constraints?
- The “best output” paradigm: should we be selective about what goes in the CDS: **should EQC chose and filter for “best” only** (on what criteria? For what purpose?) **or let the users decide**, based on documented quality (as far as possible) by the EQC
- How to judge when **products** reach **maturity** (and thus part of an “operational” service)?

Open questions (cont.)

- EQC: link with R&D. R&D fertilise each other. Role of existing projects. Need to make transparent the process of harmonisation and ingestion of Research results. There should be **a strong development component inside the CCCS**.
- What is **inside or outside** the CCCS (use of tiers of R&D as defined e.g. by the IG for atmosphere service)?
- Ingest **results** (data but also processors, algorithms...) **from activities funded elsewhere** (global projections, global reanalyses outside Europe, national activities) and come “for free” but are useful to include in the CDS. Sources include: CCI, SAFs, ECMWF... These are available resources. But the CCCS must add value must to them, in particular: **usability, facilitation** of access and QC.
- **Dependency on external/reference data**, in particular from *in situ* (e.g. GRUAN, TCCON...). H2020 call open to map gaps. Funding of some of these data are not sustainable. Recommendation on **required reference set**. “Falls between the cracks”. Concerns both the past (long time series) and NRT/monitoring (e.g. CALVAL).
- Some data required in the EQC process are **not freely available or accessible**; specific **data policies** exist (especially for input data, restricted to e.g. research and education applications).

Recommendations

- EQC needs to have **internal** and **external** (independent) **components**
- During **Stage 0**, an outcome should finalise a **scoping document for the CCCS**, that will include in particular:
 - gap analysis, start for the existing;
 - development of a set of metrics for the system/processes (system maturity matrix...) and for the end products (GCOS, fit-for-purpose...);
 - Definition of required reference observations (in particular *in situ*) for calibration/validation;
 - Statistical processing of this reference data (homogeneisation in time...) especially over Europe for evaluation, bias correction...
- Use the concept of tiers of R&D to guide what is inside or outside the service (<1y in; <2-3y on some aspects; >3y outside in e.g. H2020, national or international activities). There should be **a strong development component inside the CCCS**.