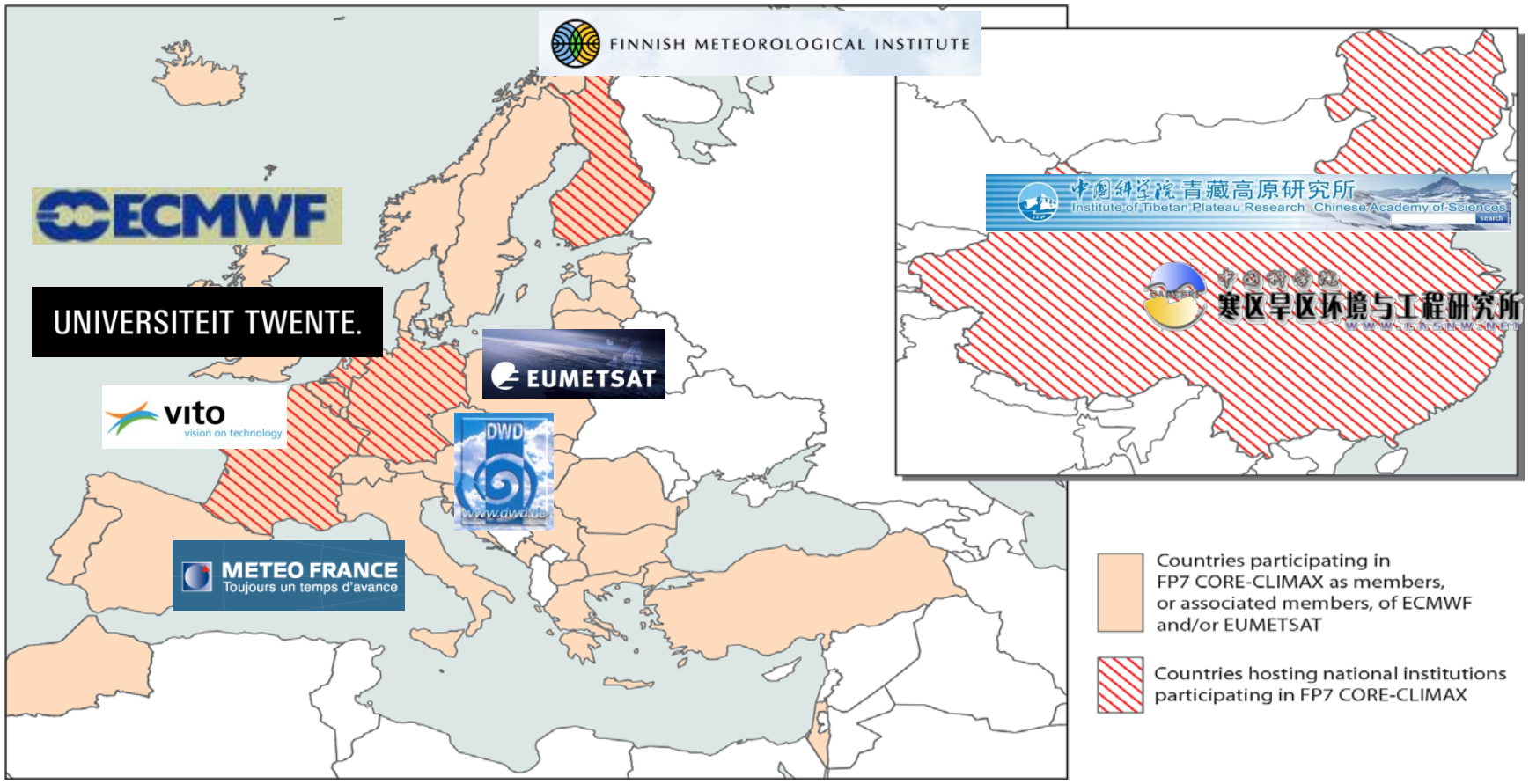
CORE-CLIMAX

COordinating Earth observation data validation for RE-analysis for CLIMAte ServiceS

Yijian Zeng on behalf of CORE-CLIMAX TEAM



CORE-CLIMAX Project Team

no.	Participant organization name	Country
1.	University of Twente, Faculty for Geo-information Science and Earth Observation (ITC)	The Netherlands
2.	European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)	
3.	European Centre for Medium-Range Weather Forecasts (ECMWF)	
4.	German Weather Service (DWD)	Germany
5.	Flemish Institute for Technological Research (VITO)	Belgium
6.	Finnish Meteorological Institute (FMI)	Finland
7.	Meteo-France (MTF)	
8.	Chinese Academy of Sciences, Institute of Tibetan Plateau Research (ITP)	China
9.	Chinese Academy of Sciences, Cold and Arid Regions Environmental and Engineering Research Institute (CAREERI)	China

Satellite data provider and producer

Reanalysis Center in Europe

ECV producers and climate service providers

Satellite data processing and validation centers

PROJECT TEAM MEMBERS

ITC/UT: Bob Su, Wim Timmermans, Yijian Zeng, Joris Timmermans, Bert Boer

EUMETSAT: Jörg Schulz, Rob Roebeling, Viju John

ECMWF: Paul Poli, David Tan

DWD: Frank Kaspar, Andrea Kaiser-Weiss, Andre Obregon

VITO: Else Swinnen, Carolien Tote, Lieven Bydekerke

FMI: Hilppa Gregow, Terhikki Manninen, Ali Nadir Arslan

MTF: Jean-Christophe Calvet, S. Lafont

ITP: Yaoming Ma, Binbin Wang

CAREERI: Jun Wen, Cai Ying, Gao Xiaoqing, Lu Shihua, Wei Zhigang, Hu Zeyong, Gao Yanhong

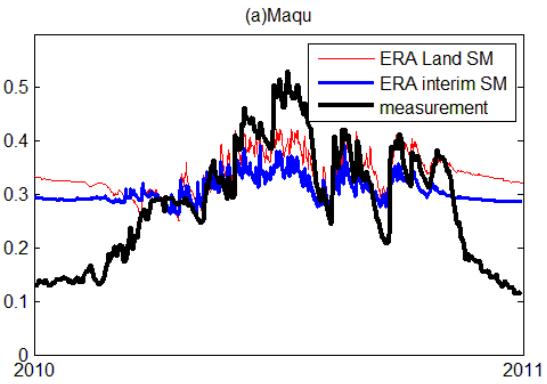
EC REA Project Officer & Advisory Board

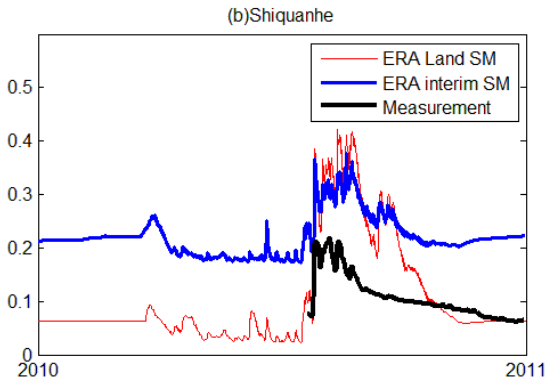
1. **REA** (Research Executive Agency): **Stijn Vermoote**

2. **Advisory Board Members**:

- **John Bates** (NOAA/NCDC, ECVs generation process and maturity index),
- **Michael Bosilovich** (NASA, reanalysis),
- **Mark Dowell** (JRC, ECVs and climate service policy requests, CEOS WG Climate),
- **Andre Jol** (EEA),
- **Steve Noyes** (EUMETNET),
- **Velina Pendolovska** (Policy Officer at DG CLIMA, email confirmation)

Core Climax



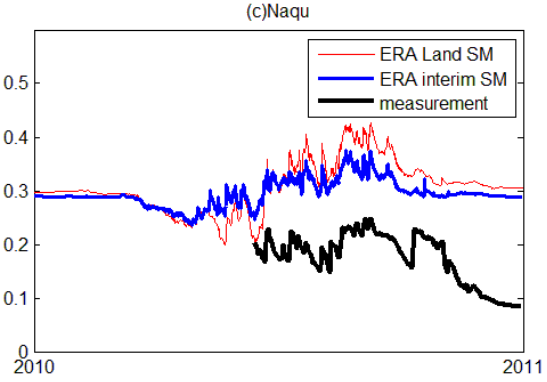


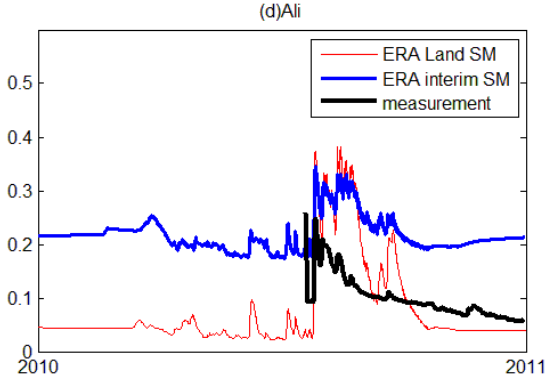
1) Coordinate with Eart

2) Propose a structured
Variables;

3) Propose a validation
climate variables;

4) Propose a feedback
process get appropr
Records;



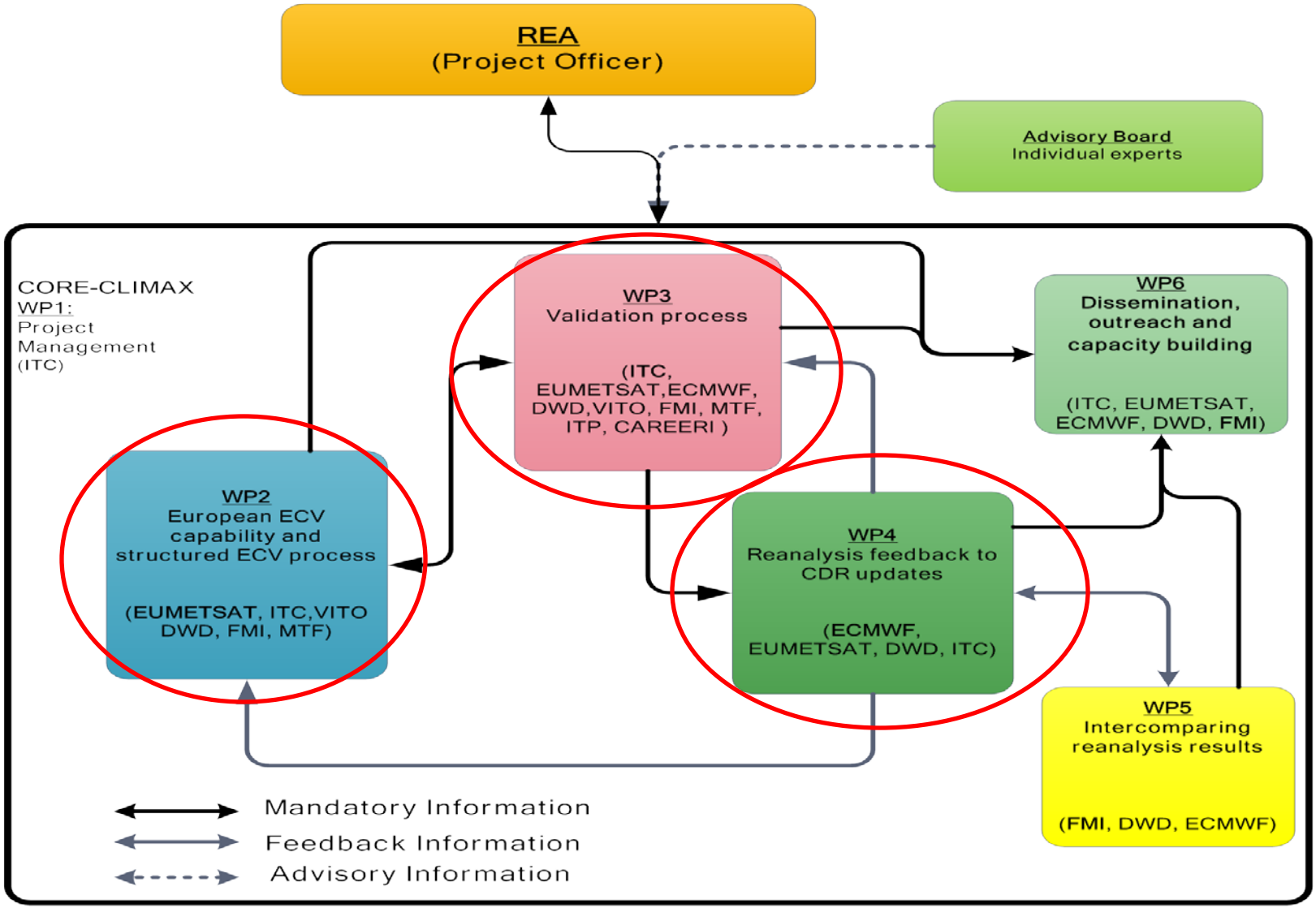


5) Propose a process to compare re-analyses.

ECV-Essential Climate Variables

CDR-Climate Data Record

CORE-CLIMAX work packages



Tools Used by CORE-CLIMAX Project

- We defined three elements for a capacity assessment:
  - Data Record Inventories that contain technical specifications and links to documented information on quality;
  - A System Maturity Matrix (SMM) that evaluates if the production of the ECV CDR follows best practices for science, engineering and utilisation;
  - An Application Performance Matrix (APM) that evaluates the performance of an ECV CDR with respect to a specific application.
- In addition User Requirements for each application, Technical Specifications and validation and/or data quality assessment results for each record are needed to ‘measure’ the performance.

Core Climax

European Commission

Copernicus
The European Earth Observation Programme

Is the software robust and maintainable?

Are the data and methods well

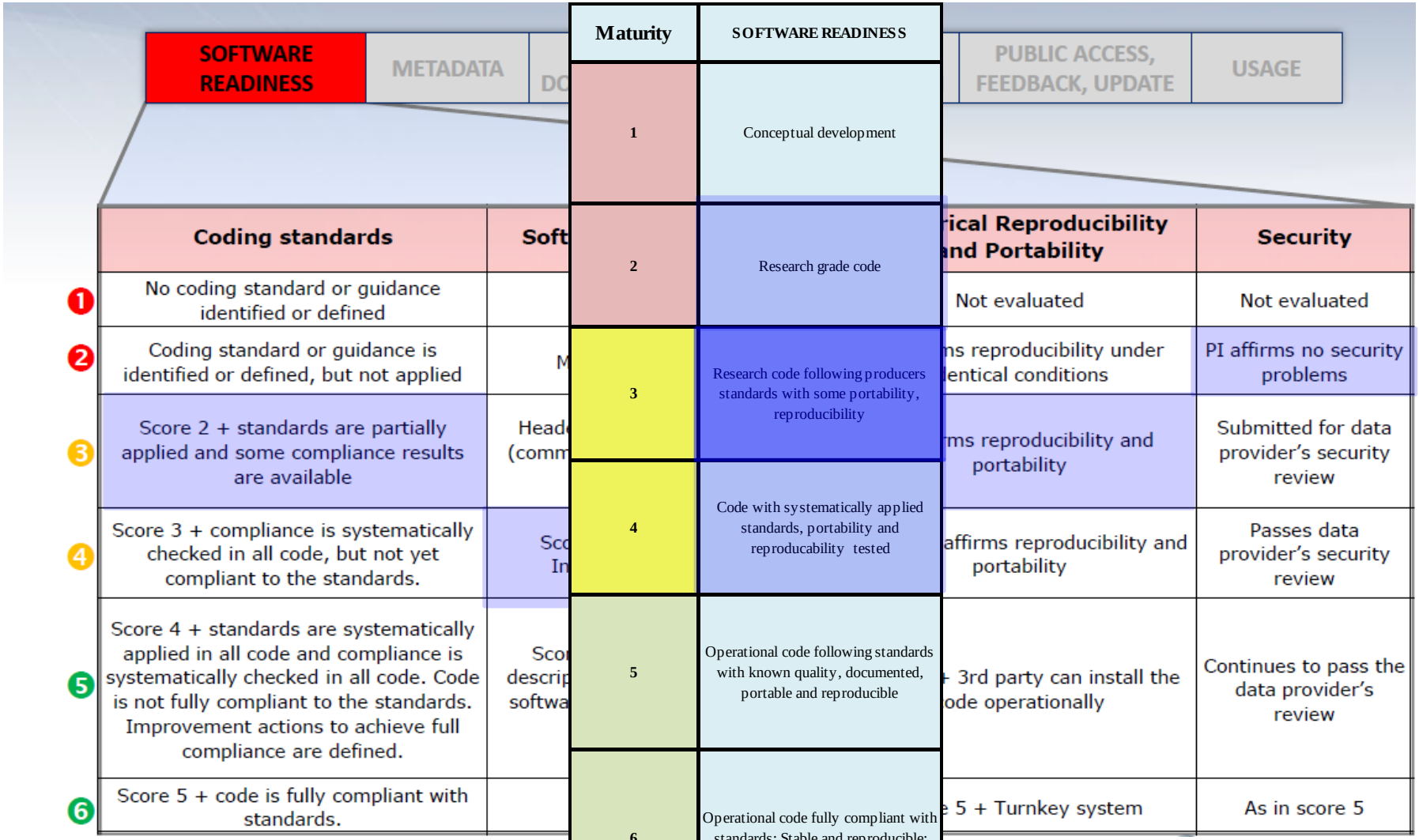
What is the trueness of the data?

Are data well used and user feedbacks taken care of?

Maturity	SOFTWARE READINESS	METADATA	USER DOCUMENTATION	UNCERTAINTY CHARACTERISATION	PUBLIC ACCESS, FEEDBACK, UPDATE	USAGE
1	Conceptual development	None	Limited scientific description of the methodology available from PI	None	Restricted availability from PI	None
2	Research grade code	Research grade	Comprehensive scientific description of the methodology, report on limited validation, and limited product user guide available from PI; paper on methodology is sumitted for peer-review	Standard uncertainty nomenclature is identified or defined; limited validation done; limited information on uncertainty available	Data available from PI, feedback through scientific exchange, irregular updates by PI	Research: Benefits for applications identified DSS: Potential benefits identified
3	Research code with partially applied standards; code contains header and comments, and a README file; PI affirms portability, numerical reproducibility and no security problems	Standards defined or identified; sufficient to use and understand the data and extract discovery metadata	Score 2 + paper on methodology published; comprehensive validation report available from PI and a paper on validation is submitted; comprehensive user guide is available from PI; Limited description of operations cocept available from PI	Score 2 + standard nomenclature applied; validation extended to full product data coverage, comprehensive information on uncertainty available; methods for automated monitoring defined	Data and documentation publically available from PI, feedback through scientific exchange, irregular updates by PI	Research: Benefits for applications demonstrated. DSS: Use occuring and benefits emerging
4	Score 3 + draft software installation/user manual available; 3rd party affirms portability and numerical reproducibility; passes data providers security review	Score 3 + standards systematically applied; meets international standards for the data set; enhanced discovery metadata; limited location level metadata	Score 3 + comprehensive scientific description available from data provider; report on inter comparison available from PI; paper on validation published; user guide available from data provider; comprehensive description of operations concept available from PI	Score 3 + procedures to establish SI traceability are defined; (inter)comparison against corresponding CDRs (other methods, models, etc); quantitative estimates of uncertainty provided within the product characterising more or less uncertain data points; automated monitoring partially implemented	Data record and documentation available from data provider and under data provider's version control; Data provider establishes feedback mechanism; regular updates by PI	Score 3 + Research: Citations on product usage in occurring DSS: societal and economical benefits discussed
5	Score 4 + operational code following standards, actions to achieve full compliance are defined; software installation/user manual complete; 3rd party installs the code operationally	Score 4+ fully compliant with standards; complete discovery metadata; complete location level metadata	Score 4 + comprehensive scientific description maintained by data provider; report on data assessment results exists; user guide is regularly updated with updates on product and validation; description on practical implementation is available from data provider	Score 4 + SI traceability partly established; data provider participated in one inter-national data assessment; comprehensive validation of the quantitative uncertainty estimates; automated quality monitoring fully implemented (all production levels)	Score 4 + soure code archived by Data Provider; feedback mechanism and international data quality assessment are considered in periodic data record updates by Data Provider	Score 4+ Research: product becomes reference for certain applications DSS: Societal and economic benefits are demonstrated
6	Score 5 + fully compliant with standards; Turnkey System	Score 5 + regularly updated	Score 5 + journal papers on product updates are and more comprehensive validation and validation of quantitative uncertainty estimates are published; operations concept regularly updated	Score 5 + SI traceability established; data provider participated in multiple inter-national data assessment and incorporating feedbacks into the product development cycle; temporal and spatial error covariance quantified; Automated monitoring in place with results fed back to other accessible information, e.g. meta data or documentation	Score 5 + source code available to the public and capability for continuous data provisions established (ICDR)	Score 5 + Research: Product and its applications becomes references in multiple research field DSS: Influence on decision and policy making demonstrated

(WP 2)

Scoring Example



Maturity	SOFTWARE READINESS
1	Conceptual development
2	Research grade code
3	Research code following producers standards with some portability, reproducibility
4	Code with systematically applied standards, portability and reproducability tested
5	Operational code following standards with known quality, documented, portable and reproducible
6	Operational code fully compliant with standards; Stable and reproducible; portable and operationally efficient

	Coding standards	Soft	ical Reproducibility nd Portability	Security
1	No coding standard or guidance identified or defined		Not evaluated	Not evaluated
2	Coding standard or guidance is identified or defined, but not applied	M	ns reproducibility under lentical conditions	PI affirms no security problems
3	Score 2 + standards are partially applied and some compliance results are available	Head (comm	ms reproducibility and portability	Submitted for data provider's security review
4	Score 3 + compliance is systematically checked in all code, but not yet compliant to the standards.	Sc In	affirms reproducibility and portability	Passes data provider's security review
5	Score 4 + standards are systematically applied in all code and compliance is systematically checked in all code. Code is not fully compliant to the standards. Improvement actions to achieve full compliance are defined.	Scor descrip softwa	+ 3rd party can install the ode operationally	Continues to pass the data provider's review
6	Score 5 + code is fully compliant with standards.		e 5 + Turnkey system	As in score 5

SOFTWARE READINESS	METADATA	USER DOCUMENTATION	UNCERTAINTY CHARATERISATION	PUBLIC ACCESS, FEEDBACK, UPDATE	USAGE

	Coding standards	Software Documentation	Numerical Reproducibility and Portability	Security
1	No coding standard or guidance identified or defined	No documentation	Not evaluated	Not evaluated
2	Coding standard or guidance is identified or defined, but not applied	Minimal documentation	PI affirms reproducibility under identical conditions	PI affirms no security problems
3	Score 2 + standards are partially applied and some compliance results are available	Header and process description (comments) in the code, README complete	PI affirms reproducibility and portability	Submitted for data provider's security review
4	Score 3 + compliance is systematically checked in all code, but not yet compliant to the standards.	Score 3 + a draft Software Installation/User Manual	3rd party affirms reproducibility and portability	Passes data provider's security review
5	Score 4 + standards are systematically applied in all code and compliance is systematically checked in all code. Code is not fully compliant to the standards. Improvement actions to achieve full compliance are defined.	Score 4 + enhanced process descriptions throughout the code; software installation/user manual complete	Score 4 + 3rd party can install the code operationally	Continues to pass the data provider's review
6	Score 5 + code is fully compliant with standards.	As in score 5	Score 5 + Turnkey system	As in score 5

What the CORE-CLIMAX Project did for the SMM (WP2)

- Made it applicable for in situ data records and other data sources such as reanalysis (we took out a lot of satellite specific language);
- Made it more easy applicable for agencies worldwide (we took out agency specific language);
- Concentrated it on the question of completeness in a sense of following best practices in science and engineering that developed over several decades;
- Tried to make the Maturity Matrix independent of individual applications;
- Accommodated many comments made by the CEOS Working Group Climate, the ESA CCI and the EUMETSAT SAFs in recent discussions of the maturity approach

CORE-CLIMAX ECV Capacity Assessment Workshop

21 - 23 January 2013
EUMETSAT, Darmstadt

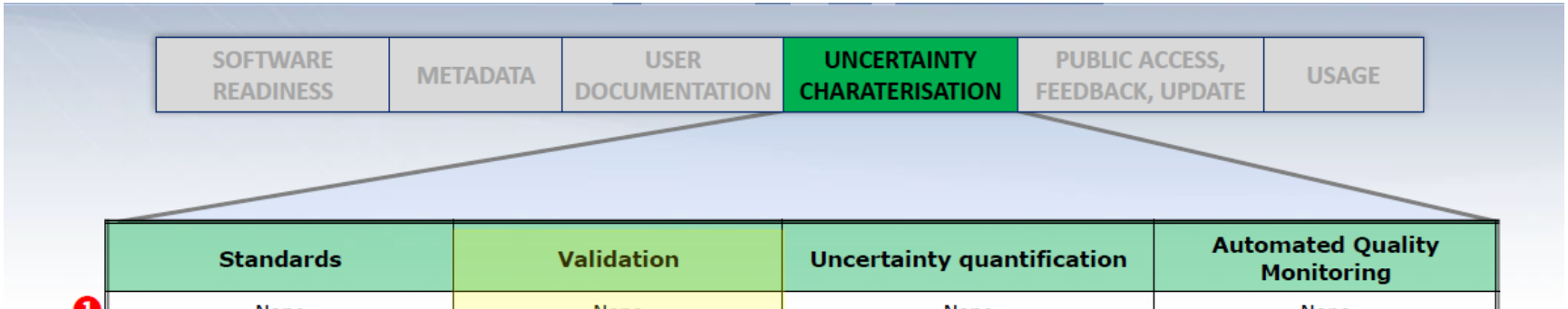
Specific Goals for the Workshop (WP2)

- Develop common understanding on the developed System Maturity Matrix (SMM); ✓
- Recommend to CORE-CLIMAX needed improvements to the SMM and instruction manual; ✓
- Discuss results of self assessment; ✓
- Discuss and agree on way forward for external/independent assessment; ✓
- Discuss value and potential of the Application Performance Matrix concept and its implementation; ✓
- Develop recommendations towards EC and other international coordination mechanisms on, e.g.:
  - Implementation of developed tools in international context;
  - Further needs for developement of requirements for applications;
  - Future assessments.

Summary

CORE-CLIMAX has proposed a structured process for assessing European capacity in delivering ECVs ;

- Using and contributing to data record inventories;
- Using an updated System Maturity Matrix approach of ‘measuring’ if data records are produced with best practises for science and engineering;
- Using a novel approach of an Application Performance Matrix to break down comprehensive information on data record quality into a performance index;
- The CORE-CLIMAX Capacity Assessment Workshop aims at using the tools to establish a first rendition of a capability data base.



	Standards	Validation	Uncertainty quantification	Automated Quality Monitoring
1	None	None	None	None
2	Standard uncertainty nomenclature is identified or defined	Validation using external reference data done for limited locations and times	Limited information on uncertainty arising from systematic and random effects in the measurement	None
3	Score 2 + Standard uncertainty nomenclature is applied	Validation using external reference data done for global and temporal representative locations and times	Comprehensive information on uncertainty arising from systematic and random effects in the measurement	Methods for automated quality monitoring defined
4	Score 3 + Procedures to establish SI traceability are defined	Score 3 + (Inter)comparison against corresponding CDRs (other methods, models, etc)	Score 3 + quantitative estimates of uncertainty provided within the product characterising more or less uncertain data points	Score 3 + automated monitoring partially implemented
5	Score 4 + SI traceability partly established	Score 4 + data provider participated in one inter-national data assessment	Score 4 + temporal and spatial error covariance quantified	Score 3 + monitoring fully implemented (all production levels)
6	Score 5 + SI traceability established	Score 4 + data provider participated in multiple inter-national data assessment and incorporating feedbacks into the product development cycle	Score 5 + comprehensive validation of the quantitative uncertainty estimates and error covariance	Score 5 + automated monitoring in place with results fed back to other accessible information, e.g. meta data or documentation

Take-home questions:

- **What are the essentials of validation in your application?** For example:

1) the generation of reference data;
2) Consistency Check (e.g. an initial analysis of physical consistency among different products that is independent from each other);
3) Definition of validation procedures/methods (e.g. direct/indirect/cross-cutting comparisons, spatial-temporal consistency analysis, large statistic, case studies, etc.);
4) Self-assessment of products;
5) Independent assessment of products;
6) External review of validation process

Yijian Zeng
y.zeng@utwente.nl

- **When is a validation process considered "complete" in your application?** (e.g. 'completeness'? or when is a data product considered "validated"?)