Electronic Supplementary Material 2
EQUIP Results Review

This EMS 2 gives the methodological underpinning of the results review. It reports the data collection method and selected data. The Results Questionnaire (Annexe 1) was filled in by all respondents (authors and reviewers) in October/November 2013 when papers were ready for submission to the journal. Authors were asked to formulate consequence statement(s) based on the outcome variable(s). These statements formed the basis of question 1 in the Results Questionnaire. In Table 1 these statements are listed together with the confidence that respondents have in the veracity of these statements. Scores were in the range of 1-5 with highest = 1.

Table 1 Authors' consequence statements (abbreviated) and respondents' confidence in consequence statements (from Results Questionnaire question 1)

Paper **Reviewed by**	**1**	**2**	**3A**	**3B**	**4**	**5**	**6**	**8**	**10**
H2013									
Predictions of all calculated indices are not reliable in UK	5	5	5	4	5	3	4	4	5
Predictions are relatively reliable in Europe & Mediterranean	4	3	4	4	3	3	3	3	3
Skill is not just related to spatial or temporal scale but is location-dependent so each index should be tested regionally	5	4	5	5	5	3	5	3	3
A warming trend is detected for most of Europe	5	4	4	4	4	3	5	4	5
O2013									
Heaviest rainfall events will be heavier & more frequent at spatial scales ≥ 50km	5	-	-	3	-	4	4	3	-
SP2013									
No evidence for increased risk of eutrophication in NE Atlantic, but the risk of eutrophication reduces if agricultural practices change to reduce nitrogen & phosphorus in rivers	3	-	-	4	5	2	-	2	4
W2013									
Impact of errors in Tmean exceeds impact of errors in Pmean	1	4	3	4	4	5	3	5	-
Impact of errors in Tsd exceeds impact of errors in Psd	3	4	1	4	4	5	2	5	-
Impact of errors in mean & sd is greatest at seasonal scale compared with subseasonal & climatological scales	3	4	2	3	4	4	3	3	-
Strengths & weaknesses of models can be compared	-	3	-	4	4	4	3	4	-

Note: Tmean = mean temperature; Tsd = standard deviation of temperature; Pmean = mean precipitation; Psd=standard deviation in precipitation

In question 3 of the Results Questionnaire we asked respondents to give the decisions that the results of the study may inform, and to rank the main sources of uncertainty in these decisions. Table 2 reports the results. Scores were in the range of 1-5 with highest = 1 (in bold). Across the four papers model inputs and model structure are assessed as the dominant

sources of uncertainty, with 11 resp. 10 scores of 1 (= highest importance) and only two scores of 5 (= lowest importance). The lowest scoring source, intrinsic variability, has one score of 1 and five of 2, and seven scores of 5. Reviewer 10 only stated the highest ranking source of uncertainty for two of the papers.

Table 2 Decisions affected by authors' consequence statements and sources of uncertainty in decisions (from Results Questionnaire question 3)

Reviewed by	**1**	**2**	**3A**	**3B**	**4**	**5**	**6**	**8**	**10**
H2013: EU level assessment for agriculture & health adaptation policies									
Model inputs	4	5	3	2	**1**	**1**	**1**	**1**	-
Model structure	**1**	**1**	4	3	3	3	3	2	-
Intrinsic variability	2	4	**1**	3	4	4	3	5	-
Parameter variability	5	3	5	**1**	4	4	2	4	**1**
Ensemble size	3	2	2	3	3	3	2	3	-
O2013: Flood defence & insurance decisions									
Model inputs	5	**1**	0	5	**1**	3	**1**	2	-
Model structure	**1**	0	0	**1**	2	**1**	**1**	**1**	-
Intrinsic variability	4	0	0	2	2	3	3	3	-
Parameter variability	0	0	0	4	2	0	2	0	-
Ensemble size	0	0	0	3	2	5	2	5	-
SP2013: Agricultural & marine management & legislation									
Model inputs	4	-	**1**	3	**1**	2	-	2	-
Model structure	**1**	-	2	**1**	2	2	-	**1**	
Intrinsic variability	3	-	5	5	5	3	-	4	-
Parameter variability	2	-	5	4	4	4	-	5	-
Ensemble size	**1**	-	**1**	2	2	3	-	3	-
W2014: Which model to use given weather input data characteristics									
Model inputs	**1**	**1**	-	2	**1**	3	**1**	2	-
Model structure	5	2	-	**1**	3	**1**	**1**	**1**	**1**
Intrinsic variability	2	4	-	5	2	5	3	5	-
Parameter variability	**1**	3	-	4	4	3	2	4	-
Ensemble size	3	5	-	3	5	3	3	3	-

Legend:

- Uncertainty in model inputs include initial conditions, boundary conditions, physical constants, driving variables
- Uncertainty in model structure include inaccuracy in model equations, spatial and temporal discretization

- Intrinsic and non-measurable stochastic variability include fundamental limits to predictability resulting from chaotic processes
- Uncertainty resulting from explicit variation of model parameters include potential over-estimation of uncertainty when producing a perturbed-parameter ensemble
- Uncertainty due to insufficient ensemble size include potential under-estimation of uncertainty due to not capturing the full range of possible model responses

EMS 2 Annexe 1 Results Questionnaire

FILLED IN BY EVERYONE

This questionnaire may be completed either individually or, where all respondents agree on responses, for a group of co-authors from the same paper.

Paper title:
Your name:

1. ***Results (this question for quantitative papers only)***
   - How confident are you in the following consequence statement, which resulted from the study (scores 1-5, where 1 is low confidence and 5 is maximum confidence)

[here the consequence statement(s) for each paper are filled in before the questionnaires are sent out]

   - Do the numerical results of this study accurately reflect the underlying uncertainties? What qualifications or adjustments would you put forward in assessing the results?
   - What qualitative information on processes and tradeoffs results from this study?

2. ***Users (this question for quantitative papers only)***
   - Who do you foresee using these results?
   - Was this presentation of uncertainty targeted towards any particular user group? If so, how?
   - Are there further steps that could be taken to present the uncertainty in this study to a user group?

3. ***Decisions (this question for quantitative papers only)***
   - Which decisions might these results inform?
   - Rank the sources of uncertainty below (which are taken from the common uncertainty reporting format) in order of their relative contribution to the total uncertainty in each decision
     1. Uncertainty in model inputs (e.g. initial conditions, boundary conditions, physical constants, driving variables)
     2. Uncertainty in model structure (e.g. inaccuracy in model equations, spatial and temporal discretization)
     3. Intrinsic and non-measurable stochastic variability (e.g. fundamental limits to predictability resulting from chaotic processes)
     4. Uncertainty resulting from explicit variation of model parameters (i.e. potential over-estimation of uncertainty when producing a perturbed-parameter ensemble)

5. Uncertainty due to insufficient ensemble size (i.e. potential under-estimation of uncertainty due to not capturing the full range of possible model responses)

4. Lessons learnt (this question for all papers)

- What are the implications of this study for how uncertainty should be dealt with in other studies? e.g. key information regarding quantification and/or management of uncertainty
- Are there any other generalisable elements to the results of this study?
- What questions about uncertainty / areas for further work arose from this study?
- For the topic(s) of this paper, where should investment now be prioritised: reducing uncertainty, or Better quantification of uncertainty? Please state your reasoning.
- What form would such investment take (e.g. more precise or densely distributed observations, reanalysis; development of specific models, greater computing power, running multi-model ensembles, improving multi-model techniques; research into fundamental processes)?