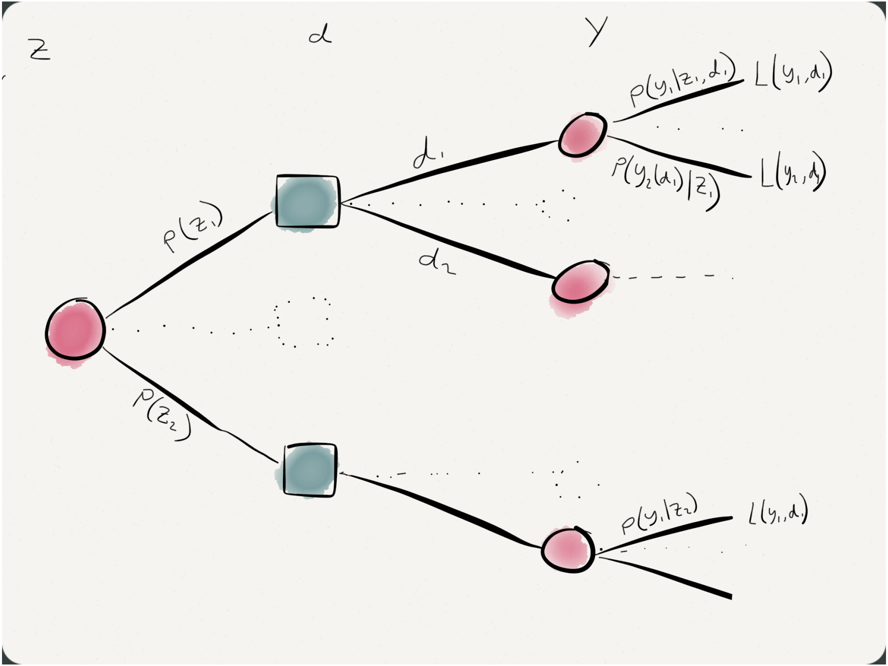
Probability for future weather: how close are we to decision actionable expert judgement?

Danny Williamson

University of Exeter

April 16, 2015

Climate: probabilities for future weather



Z
d
Y
p(z1)
p(z2)
d1
d2
p(y1|z1,d1)
L(y1,d1)
p(y2(d1)|z1)
L(y2,d1)
p(y1|z2)
L(y1,d1)

Reality, Data, Models and Loss

Two principal challenges involve specifying L and assessing $\mathrm{P}(Y_{1:t}(d))$.

Reality, Data, Models and Loss

Two principal challenges involve specifying L and assessing $\mathrm{P}(Y_{1:t}(d))$. We use models to assist with both.

- Simplifying, let $Y = (Y_H, Y_F)'$.
- We observe climate with error $Z_H = Y_H + e_H$
- If we can find $\mathrm{P}(Y)$ and we know the distribution of e_H, we can easily derive $\mathrm{P}(Y_F|Z_H)$.

Reality, Data, Models and Loss

Two principal challenges involve specifying L and assessing $\mathrm{P}(Y_{1:t}(d))$. We use models to assist with both.

- Simplifying, let $Y = (Y_H, Y_F)'$.
- We observe climate with error $Z_H = Y_H + e_H$
- If we can find $\mathrm{P}(Y)$ and we know the distribution of e_H, we can easily derive $\mathrm{P}(Y_F|Z_H)$.
- Enter climate models.
- We have a selection of climate models $f_i(x_{[i]}, \theta)$ used to try to predict Y under forcing θ.
- How can information from the f_i's get us to $\mathrm{P}(Y)$ (or $\mathrm{P}(Y_F|Z_H)$)?

Spartacus #1: One Climate model

One model approach:

- Each model is informative for $Y(\theta)$, but there is structural discrepancy left over:

$$Y(\theta) = f_i(x^*_{[i]}, \theta) + \eta_i(\theta)$$

Spartacus #1: One Climate model

One model approach:

- Each model is informative for $Y(\theta)$, but there is structural discrepancy left over:

$$Y(\theta) = f_i(x^*_{[i]}, \theta) + \eta_i(\theta)$$

- This is the approach used in the UK climate projections

Spartacus #1: One Climate model

One model approach:

- Each model is informative for $Y(\theta)$, but there is structural discrepancy left over:

$$Y(\theta) = f_i(x^*_{[i]}, \theta) + \eta_i(\theta)$$

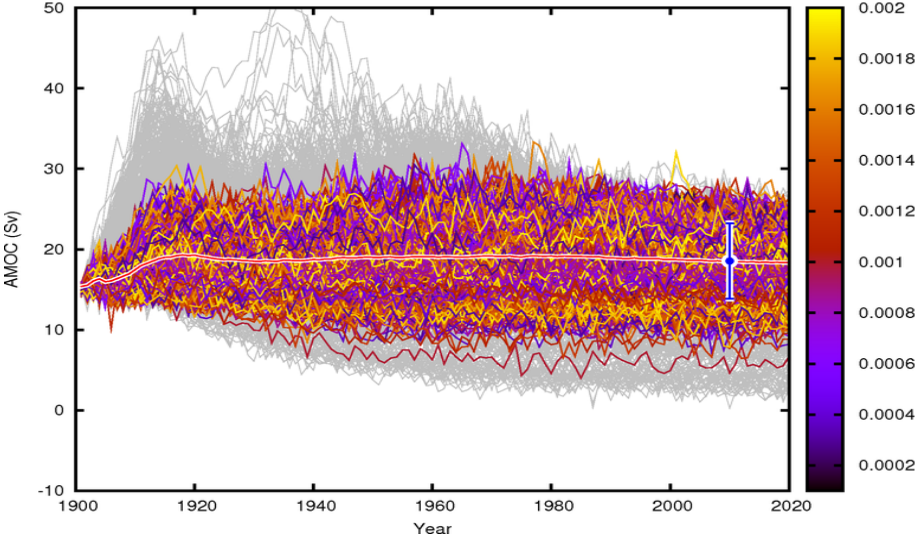
- This is the approach used in the UK climate projections
- We can get Monte Carlo samples from $\mathrm{P}(Y(\theta))$ if we can sample from

$$\mathrm{P}(f_i(x^*_{[i]}, \theta) | x^*_{[i]}) \mathrm{P}(x^*_{[i]}) \mathrm{P}(\eta_i(\theta))$$

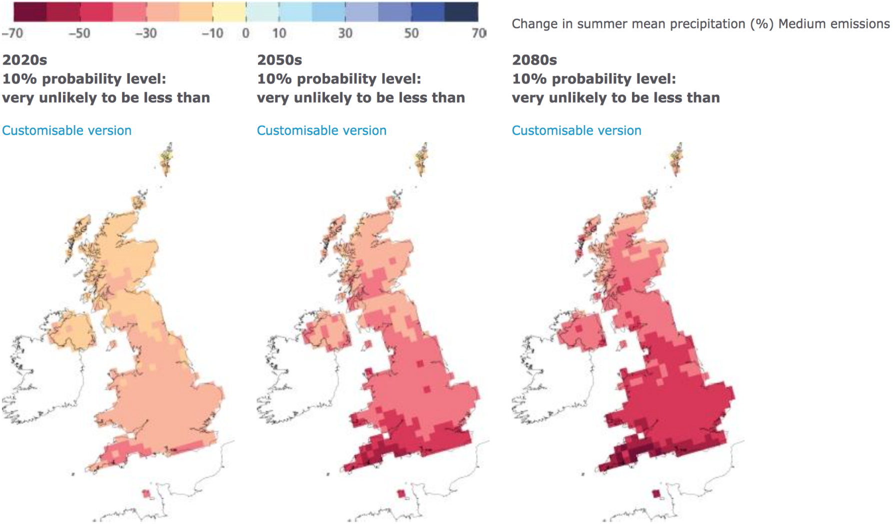
- In practice, the joint distribution of $\{f_i(\cdot, \theta), x^*_{[i]}, \eta_i(\theta)\}$ is conditioned on Z_H.

The UK climate projections

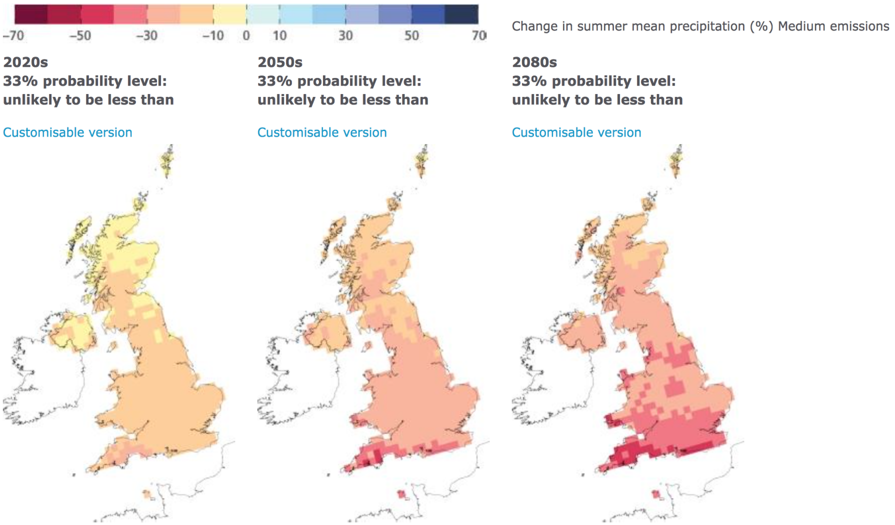
- The UK Climate Projections use an ensemble of runs on one model and the above framework to get "probabilities" for 3 scenarios.



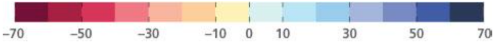
The UK climate projections



The UK climate projections



The UK climate projections



Change in summer mean precipitation (%) Medium emissions

2020s
50% probability level:
central estimate

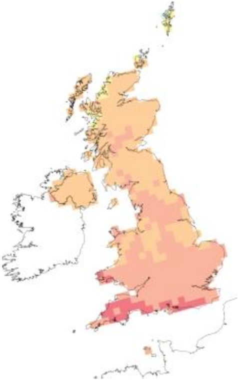
Customisable version

2050s
50% probability level:
central estimate

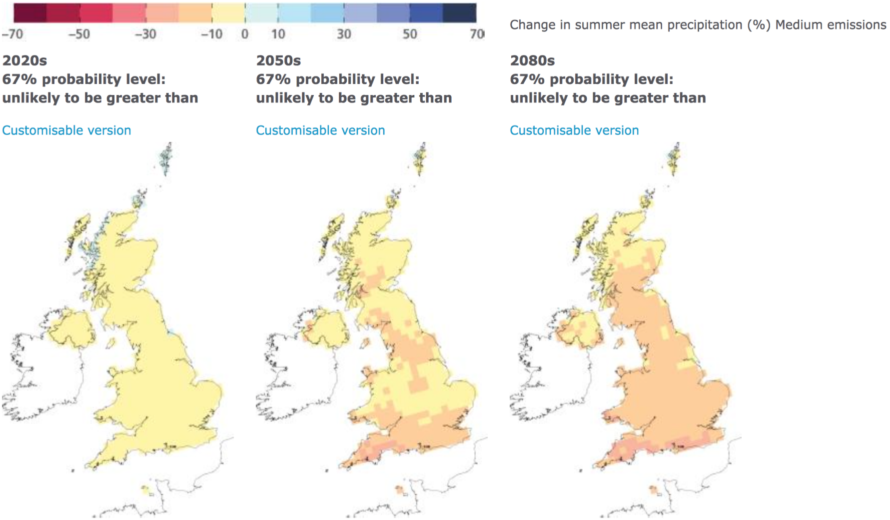
Customisable version

2080s
50% probability level:
central estimate

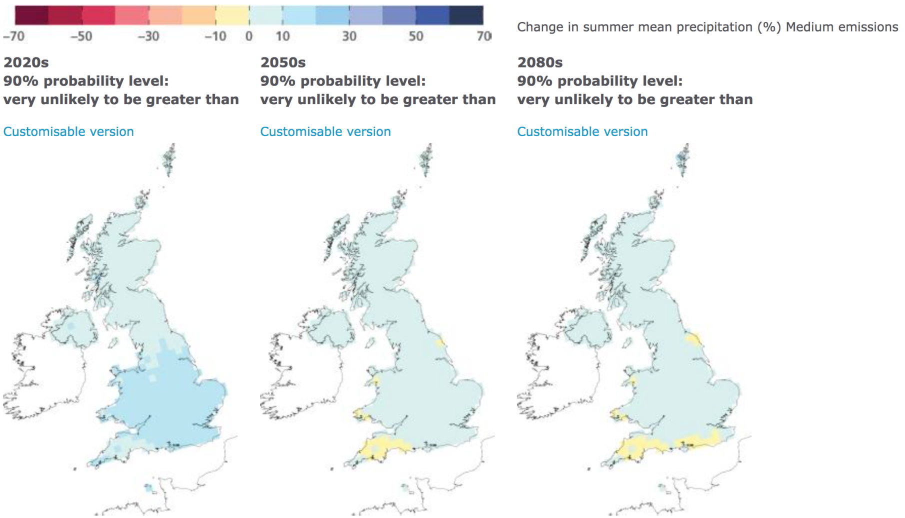
Customisable version



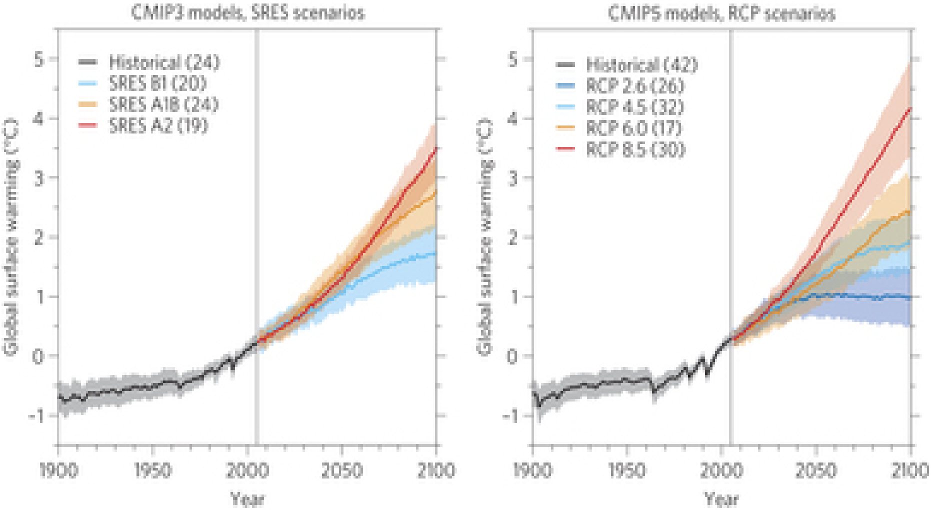
The UK climate projections



The UK climate projections



Spartacus #2: Multi-model ensembles - CMIP-X



Statistical modelling

Multi-model approach:

- The models are exchangeable and $Y(\theta)$ relates to the collection: E.g.

$$f_i(x^*_{[i]}, \theta) = \mathcal{M}(\theta) + R_i(\theta); \qquad Y(\theta) = \alpha\mathcal{M}(\theta) + U(\theta)$$

Statistical modelling

Multi-model approach:

- The models are exchangeable and $Y(\theta)$ relates to the collection: E.g.

$$f_i(x^*_{[i]}, \theta) = \mathcal{M}(\theta) + R_i(\theta); \qquad Y(\theta) = \alpha\mathcal{M}(\theta) + U(\theta)$$

- We observe $f_1(x^t_{[1]}), \ldots, f_n(x^t_{[n]})$ and we can get Monte Carlo samples from $\mathrm{P}(Y(\theta))$ if we can sample from

$$\mathrm{P}(U(\theta))\mathrm{P}(\alpha, \mathcal{M}(\theta)) \prod_{i=1}^{k} \mathrm{P}(f_i(x^*_{[i]}) | f_i(x^t_{[i]}), x^*_{[i]}, \mathcal{M}(\theta))\mathrm{P}(x^*_{[i]})$$

Current practice: What lurks in the conditioning?

Often, uncertainties are ignored instead of quantified. This does not reduce uncertainty, it removes problems to the conditioning...

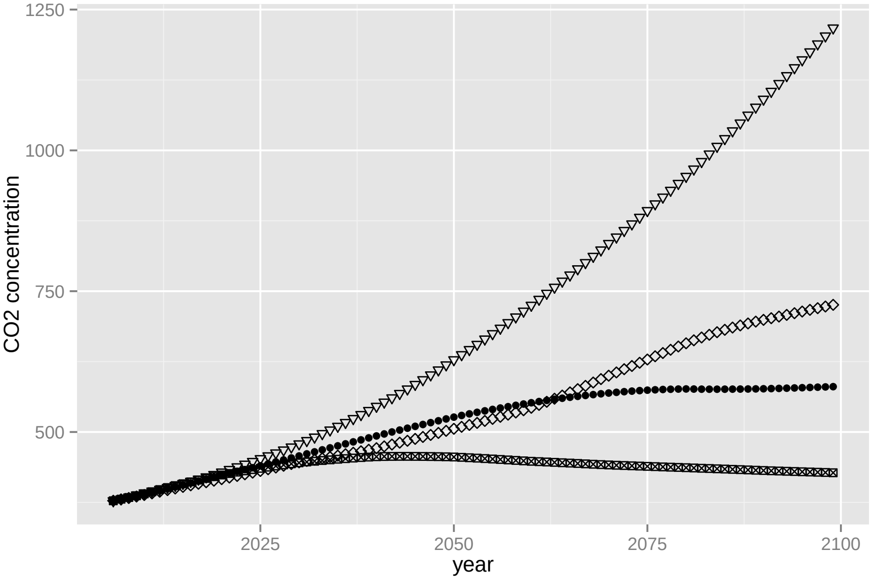
Current practice: What lurks in the conditioning?

Often, uncertainties are ignored instead of quantified. This does not reduce uncertainty, it removes problems to the conditioning...

Example

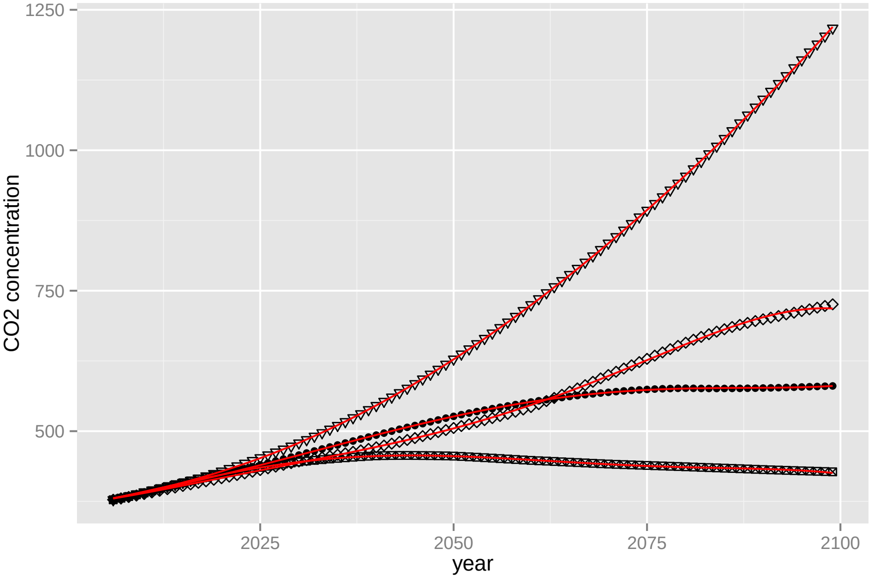
- The CMIP GCMs are run at $x^t_{[i]} \neq x^*_{[i]}$. I.e. they are not optimally tuned.
- But this is not addressed. In fact, we act as if $x^t_{[i]} = x^*_{[i]}$.
- Now $\mathrm{P}(x^*_{[i]})$ is gone and $\mathrm{P}(f_i(x^t_{[i]}, \theta))$, has no code uncertainty!
- Hence we obtain samples from internal variability only and can get to $\mathrm{P}(Y(\theta) | x^*_{[i]} = x^t_{[i]})$.

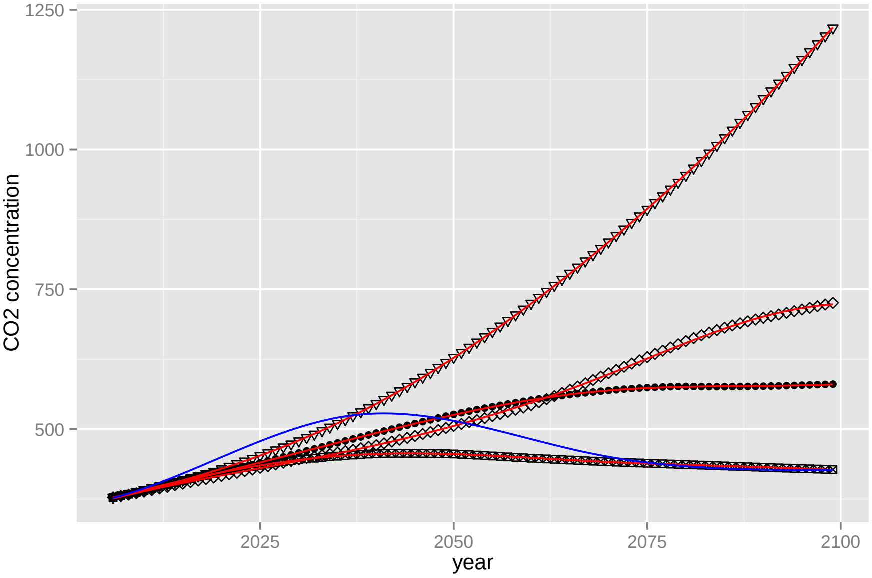
What are the scenarios?



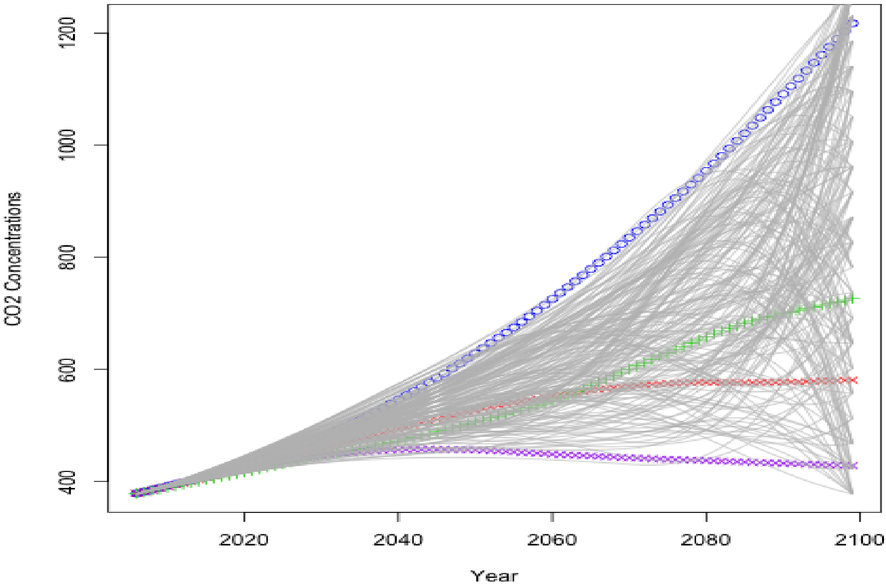
Policy support: Beyond Scenario Analysis

- $\mathrm{P}(Y(\theta)) = \mathrm{P}(Y|\theta)$.
- Can we get to $\mathrm{P}(Y)$ or $\mathrm{P}(Y|\theta^*)$?
- If policy makers really wanted it, we could make inference and provide decision support beyond the RCPs/SSPs.
- All this would take would be a little creative statistical modelling and better ensemble design!


1250
1000
750
500
CO2 concentration
2025
2050
2075
2100
year



1250
1000
750
500
CO2 concentration
2025
2050
2075
2100
year



CO2 Concentrations
1200
1000
800
600
400
2020
2040
2060
2080
2100
Year

Should we expect probability any time soon?

- There is a wealth of data and sophisticated models in climate science.

Should we expect probability any time soon?

- There is a wealth of data and sophisticated models in climate science.
- However, the quest for **a probability distribution** for future climate is not going well.

Should we expect probability any time soon?

- There is a wealth of data and sophisticated models in climate science.
- However, the quest for **a probability distribution** for future climate is not going well.
- 'Probabilistic' analyses are abstracted from the real world by scenarios.

Should we expect probability any time soon?

- There is a wealth of data and sophisticated models in climate science.
- However, the quest for **a probability distribution** for future climate is not going well.
- 'Probabilistic' analyses are abstracted from the real world by scenarios.
- Key uncertainties remain unquantified and ignored.

Should we expect probability any time soon?

- There is a wealth of data and sophisticated models in climate science.
- However, the quest for **a probability distribution** for future climate is not going well.
- 'Probabilistic' analyses are abstracted from the real world by scenarios.
- Key uncertainties remain unquantified and ignored.
- The field seems to view model resolution as a panacea.

Should we expect probability any time soon?

- There is a wealth of data and sophisticated models in climate science.
- However, the quest for **a probability distribution** for future climate is not going well.
- 'Probabilistic' analyses are abstracted from the real world by scenarios.
- Key uncertainties remain unquantified and ignored.
- The field seems to view model resolution as a panacea.
- Expert judgement is arguably seen as undesirable and even dangerous.

Should we expect probability any time soon?

- There is a wealth of data and sophisticated models in climate science.
- However, the quest for **a probability distribution** for future climate is not going well.
- 'Probabilistic' analyses are abstracted from the real world by scenarios.
- Key uncertainties remain unquantified and ignored.
- The field seems to view model resolution as a panacea.
- Expert judgement is arguably seen as undesirable and even dangerous.
- But even if we eventually fix all of these issues, could we ever get pdfs that we really believe?

Should we expect probability any time soon?

- There is a wealth of data and sophisticated models in climate science.
- However, the quest for **a probability distribution** for future climate is not going well.
- 'Probabilistic' analyses are abstracted from the real world by scenarios.
- Key uncertainties remain unquantified and ignored.
- The field seems to view model resolution as a panacea.
- Expert judgement is arguably seen as undesirable and even dangerous.
- But even if we eventually fix all of these issues, could we ever get pdfs that we really believe?
- Is the quest misguided?

Posterior belief assessment

Not just a problem for climate science

- One school of thought might be that we need better, more careful judgements from the scientific community.
- Then our statistical models and the Bayesian machinery will lead us to decision actionable probability.

Not just a problem for climate science

- One school of thought might be that we need better, more careful judgements from the scientific community.
- Then our statistical models and the Bayesian machinery will lead us to decision actionable probability.

Argument:

- For complex Bayesian models applied in any scientific discipline, we never believe all of the judgements in the prior and likelihood.
- Our posterior samples are not draws from anyones “probability distribution”.
- What makes these probabilities ‘decision actionable’, but those derived by, say, UKCP09, not?

Posterior belief assessment

- **Basic idea**: Post process the results of imperfect Bayesian analyses to extract judgments that are decision actionable.

Posterior belief assessment

- **Basic idea**: Post process the results of imperfect Bayesian analyses to extract judgments that are decision actionable.
- **Key premise:** We do not require the entire posterior, only a handful of key expectations (*prevision*).

Posterior belief assessment

- **Basic idea**: Post process the results of imperfect Bayesian analyses to extract judgments that are decision actionable.
- **Key premise:** We do not require the entire posterior, only a handful of key expectations (*prevision*).
- Key policy relevant previsions might be:
  - Expected losses
  - Threshold probabilities (e.g. $\mathrm{E}\left[\mathbb{I}\{Y_F - \bar{Y}_H > 2^o\}|Z_H\right]$)
  - Variances (risk profiles)

Posterior belief assessment

- **Basic idea**: Post process the results of imperfect Bayesian analyses to extract judgments that are decision actionable.
- **Key premise:** We do not require the entire posterior, only a handful of key expectations (*prevision*).
- Key policy relevant previsions might be:
    - Expected losses
    - Threshold probabilities (e.g. $\mathrm{E}\left[\mathbb{I}\{Y_F - \bar{Y}_H > 2^o\}|Z_H\right]$)
    - Variances (risk profiles)
- Our view is that most of the time we use probability, we are using it as a modelling language rather than a measure of our actual subjective beliefs.
- **Posterior belief assessment** is a method for using that rich and powerful modelling language to reach a handful of statements we are prepared to adopt as our judgments.

Collections of pre-posterior judgements

- The Bayesian analysis we produce is based on a collection of judgements J_0.
  - Statistical modelling
  - Prior distributions
  - Hyper-parameters
  - Computational issues (e.g. MCMC proposals/convergence)
  - Model fitting

Collections of pre-posterior judgements

- The Bayesian analysis we produce is based on a collection of judgements J_0.
    - Statistical modelling
    - Prior distributions
    - Hyper-parameters
    - Computational issues (e.g. MCMC proposals/convergence)
    - Model fitting
- A subset might represent what you and the expert really believe.

Collections of pre-posterior judgements

- The Bayesian analysis we produce is based on a collection of judgements J_0.
    - Statistical modelling
    - Prior distributions
    - Hyper-parameters
    - Computational issues (e.g. MCMC proposals/convergence)
    - Model fitting
- A subset might represent what you and the expert really believe.
- Many though are influenced by pragmatism, inertia, time constraints, limited access to the expert, potentially robust choices.

Collections of pre-posterior judgements

- The Bayesian analysis we produce is based on a collection of judgements J_0.
  - Statistical modelling
  - Prior distributions
  - Hyper-parameters
  - Computational issues (e.g. MCMC proposals/convergence)
  - Model fitting
- A subset might represent what you and the expert really believe.
- Many though are influenced by pragmatism, inertia, time constraints, limited access to the expert, potentially robust choices.
- We claim that there exist alternative judgements $J_1, J_2, \ldots$ that "could" be better representations of your beliefs.

What do our judgements mean?

- We are interested in certain key conditional expectation judgements following our Bayesian analysis.

What do our judgements mean?

- We are interested in certain key conditional expectation judgements following our Bayesian analysis.
- We can compute $\mathrm{E}\left[y|z;J_0\right]$ from our posterior distribution.

What do our judgements mean?

- We are interested in certain key conditional expectation judgements following our Bayesian analysis.
- We can compute $\mathrm{E}[y|z; J_0]$ from our posterior distribution.
- If $\mathrm{E}[y|z; J_0]$ were our prevision for y having seen z (our actual judgement), we prefer the random penalty $K(y - \mathrm{E}[y|z; J_0])^2$ to any random penalty $K(y - A(z))^2$ for constant K.
- I.e. $\mathrm{E}[y|z; J_0] = \mathrm{argmin}_z \mathrm{E}\left[(y - A(z))^2\right]$

What do our judgements mean?

- We are interested in certain key conditional expectation judgements following our Bayesian analysis.
- We can compute $\mathrm{E}[y|z; J_0]$ from our posterior distribution.
- If $\mathrm{E}[y|z; J_0]$ were our prevision for y having seen z (our actual judgement), we prefer the random penalty $K(y - \mathrm{E}[y|z; J_0])^2$ to any random penalty $K(y - A(z))^2$ for constant K.
- I.e. $\mathrm{E}[y|z; J_0] = \mathrm{argmin}_z \mathrm{E}\left[(y - A(z))^2\right]$
- We might argue that, as we adopted J_0, we do currently prefer $K(y - \mathrm{E}[y|z; J_0])^2$ over any $K(y - \mathrm{E}[y|z; J_k])^2$ with $k \neq 0$.

What do our judgements mean?

- We are interested in certain key conditional expectation judgements following our Bayesian analysis.
- We can compute $\mathrm{E}[y|z;J_0]$ from our posterior distribution.
- If $\mathrm{E}[y|z;J_0]$ were our prevision for y having seen z (our actual judgement), we prefer the random penalty $K(y-\mathrm{E}[y|z;J_0])^2$ to any random penalty $K(y-A(z))^2$ for constant K.
- I.e. $\mathrm{E}[y|z;J_0] = \mathrm{argmin}_z \mathrm{E}\left[(y-A(z))^2\right]$
- We might argue that, as we adopted J_0, we do currently prefer $K(y-\mathrm{E}[y|z;J_0])^2$ over any $K(y-\mathrm{E}[y|z;J_k])^2$ with $k \neq 0$.
- Though there may be some parts of J_0, made for pragmatism only, that may cloud this.

Posterior belief assessment

- Suppose we consider the random penalty

$$K\left(y - \sum_{i=0} \alpha_i \mathcal{G}_i(\mathrm{E}[y|z;J_0], \mathrm{E}[y|z;J_1], .., \mathrm{E}[y|z;J_{n_k}])\right)^2$$

  with $\mathcal{G}(\cdot)$ a vector containing specified functionals of a finite number, $n_k + 1$, of conditional expectations as calculated using the same Bayesian machinery, with different collections of judgements $J_0, J_1, \ldots, J_{n_k}$.
- Then, our Prevision for y is

$$\sum_{i=0} \hat{\alpha}_i \mathcal{G}_i(\mathrm{E}[y|z;J_0], \mathrm{E}[y|z;J_1], .., \mathrm{E}[y|z;J_{n_k}])$$

  with $\hat{\alpha}$ chosen to minimise the expectation of the given random penalty.

Posterior belief assessment

Define $P_t(y)$ to be an actual posterior prevision that we would make at time t after seeing data z.

Theorem

Let

$$\mathrm{E}_{\mathcal{G}}[y] = \mathrm{E}[y] + \mathrm{Cov}[y, \mathcal{G}]\,\mathrm{Var}[\mathcal{G}]^{-1}(\mathcal{G} - \mathrm{E}[\mathcal{G}]). \tag{1}$$

Then

(i) $\mathrm{E}_{\mathcal{G}}[y]$ *is at least as close to* y *as* $\mathrm{E}[y|z;J_0]$. *Equivalently, for each* i,

$$\mathrm{E}\left[(y_i - \mathrm{E}_{\mathcal{G}}[y_i])^2\right] \leq \mathrm{E}\left[(y_i - \mathrm{E}[y_i|z;J_0])^2\right].$$

where $\mathrm{E}_{\mathcal{G}}[y_i]$ *is the* i*th component of* $\mathrm{E}_{\mathcal{G}}[y]$.

(ii) $\mathrm{E}_{\mathcal{G}}[y]$ *is at least as close to* $P_t(y)$ *as* $\mathrm{E}[y|z;J_0]$. *Equivalently, for each* i,

$$\mathrm{E}\left[(P_t(y_i) - \mathrm{E}_{\mathcal{G}}[y_i])^2\right] \leq \mathrm{E}\left[(P_t(y_i) - \mathrm{E}[y_i|z;J_0])^2\right].$$

Practical posterior belief assessment

- The theorem shows that $\mathrm{E}_{\mathcal{G}}[y]$, if we can compute it, is closer to our prevision (what we really believe) than our preferred Bayesian analysis $\mathrm{E}[y|z, J_0]$.
- Our method allows for an infinite set of possible alternative judgements $J_1, J_2, \ldots$ and for a carefully chosen sample of alternative Bayesian analyses from this set to be completed and used to obtain $\mathcal{G}$.
- We describe a sampling method for computing $\mathrm{E}[y]$, $\mathrm{Cov}[y, \mathcal{G}]$, $\mathrm{Var}[\mathcal{G}]$, and $\mathrm{E}[\mathcal{G}]$.
- Details, examples and a proof for the theorem can be found in

Williamson, D., Goldstein, M. (2014),
Posterior belief assessment: extracting meaningful subjective judgements from Bayesian analyses with complex statistical models, Bayesian Analysis, In revision.

Posterior belief assessment and decision support for future climate

Focussing on decision support

- Decision support only ever requires a few key expectations.

Focussing on decision support

- Decision support only ever requires a few key expectations.
- Focussing on making these previsions using the information from climate models and data may be a better use of time.

Focussing on decision support

- Decision support only ever requires a few key expectations.
- Focussing on making these previsions using the information from climate models and data may be a better use of time.
- Posterior belief assessment may be a tool that would allow us to do this using the data and model output we have now.
- Specifically, the models/data/frameworks/priors, even individual projections could collectively form a judgement set J_k.
- We don't believe J_k, but we can use it to compute $\mathrm{E}[y|z; J_k]$.

Focussing on decision support

- Decision support only ever requires a few key expectations.
- Focussing on making these previsions using the information from climate models and data may be a better use of time.
- Posterior belief assessment may be a tool that would allow us to do this using the data and model output we have now.
- Specifically, the models/data/frameworks/priors, even individual projections could collectively form a judgement set J_k.
- We don't believe J_k, but we can use it to compute $\mathrm{E}[y|z; J_k]$.
- By carefully thinking about alternative judgements $J_1, J_2, \ldots$, we can reframe decision support as a posterior belief assessment.
- We can/should do this working directly with the decision makers, rather than within climate science.