A framework for evaluating uncertainty in crop model predictions

Daniel Wallach, Peter Thorburn, Senthold Asseng, Andrew J. Challinor, Frank Ewert, James W. Jones, Reimund P. Rötter, Alex Ruane

iCROPM2016

Most forecasters

recognize the importance of providing interval forecasts as well as (or instead of) point forecasts.

- to understand limitations of forecast
- to compare forecasting methods

- We need to associate a realistic measure of quality with crop model forecasts

Vocabulary

- Error versus uncertainty:
  - Error is a number (true – simulated)
  - Uncertainty is a distribution
  - But don't rely on English, rely on equations

Define prediction uncertainty

- Prediction uncertainty is the distribution of

  $Y\text{-}f(X;\theta)$

  – where Y is the true value (e.g. yield)

  – $f(X; \theta)$ is the predictor (our model)

    - f is model structure
    - X is input vector
    - θ is parameter vector

The predictor is completely determined by these 3 elements

MSEP

- It's not convenient to work with a distribution.
- A simple summary of prediction uncertainty is the mean squared error of prediction.

This means that the value of X (weather, soil, etc) is fixed

- For some specific X

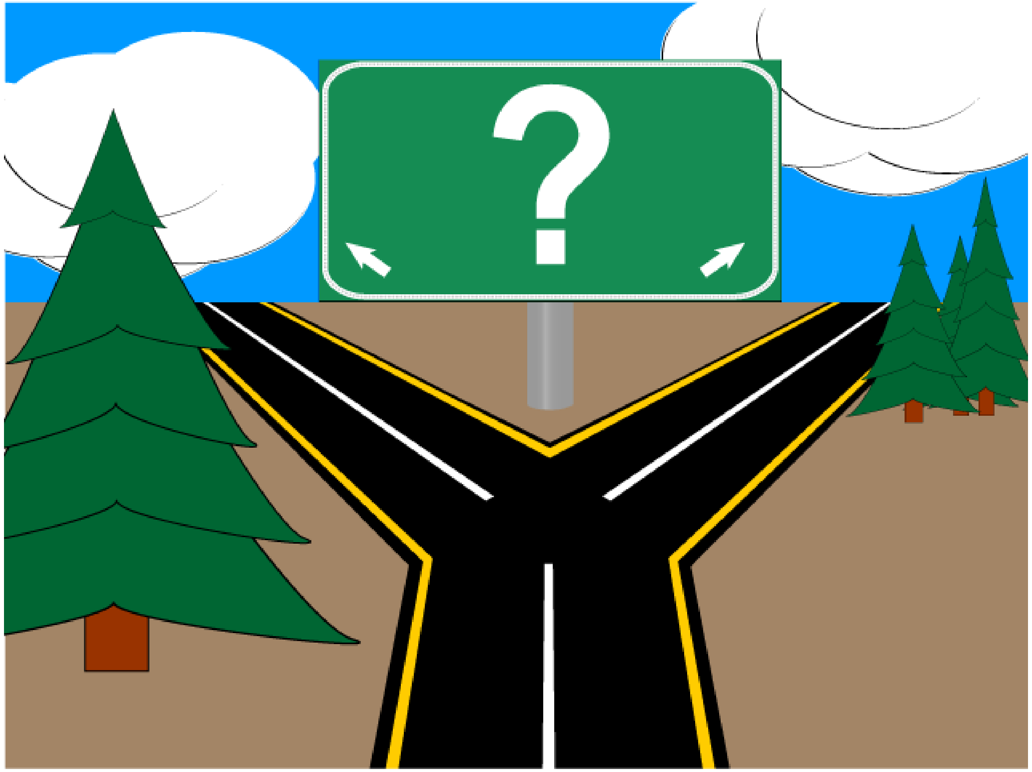
$$MSEP(X)=E\{[Y-f(X;\theta)]^2|X\}$$

Our criterion of prediction uncertainty

$$MSEP(X)=E\{[Y-f(X;\theta)]^2|X\}$$

- The output Y is a random variable
  - Y can have a range of values, even once X is given
  - Because input variables don't explain everything
- But what about the model? Fixed or random?

Fixed or random model?



Fixed or random model?

- That is the main topic of this talk
  - How to estimate MSEP in each case
  - Advantages and drawbacks of each
- The two possibilities correspond to very different ways of thinking about, and estimating, prediction uncertainty.
  - So the choice is important.
- All work related to prediction uncertainty is based on one or the other
  - Even though it isn't said like that

Fixed predictor f(X; θ)

- We have one specific model (fixed)
- With given parameter values (fixed)
- Assume that the inputs are known without uncertainty

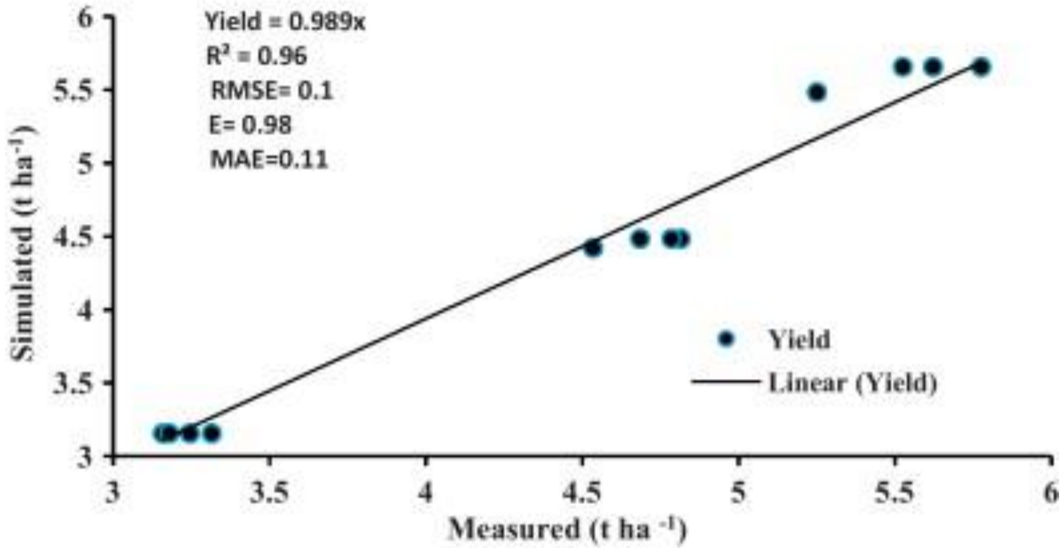
Estimation of MSEP

- Compare hindcasts with observations
- Can't estimate for each X (only have y_i for a few X)
- So estimate average over all predictions
- Estimator is $MSE=(1/n)\sum[y_i-f(X_i;\theta)]^2$
- Estimates $MSEP_{fixed}= E[MSEP_{fixed}(X)]$

Fixed model, in practice

- MSE (or RMSE) is most common criterion in model "evaluation" or "model validation" or "model performance".
- This estimates $MSEP_{fixed}$ (assumes fixed model)
  - There are other distance measures based on hindcasts, but they also estimate properties of prediction uncertainty assuming fixed model

One example (from very many)



	RMSE	E	MAE	R2
Grain yield (t ha−1)	0.1	0.98	0.11	0.96
Biomass (t ha−1)	0.75	0.95	1.08	0.9
WP (kg ha−1 mm−1)	1.2	0.74	0.71	0.77

Abedinpour et al. 2012.

Features of $MSEP_{fixed}$

- Assumption (often implicit) is that past errors are representative of future errors
  - So this is an estimate of prediction uncertainty
- Only calculate average MSEP
  - One value for all predictions
  - **Average** of past errors is representative of **average** of future errors
- For MSE to be unbiased estimate of $MSEP_{fixed}$, requires independence between evaluation and calibration errors.
  - Not easy to insure
  - In example, calibrate in 2009, evaluate for same field in 2010
  - Prediction error for other sites?

Random predictor $f(X;\theta)$

- There are multiple alternative models.
  - Distribution: Simplest assumption - all equally probable
- Input variables measured or estimated with error
  - Distribution: from measurement sd or literature or multiple models (future climate)
- Parameter vector
  - From literature, from calibration (frequentist or Bayesian)

$MSEP_{random}(X)$ is sum of two terms

- bias: error, averaged over X, of predictor averaged over equations, inputs, parameters,
- variance: uncertainty in predictor due to uncertainties in equations, inputs, parameters

bias term

variance term

$$MSEP_{random}(X) = E\left\{\left[\left(y - E[f(X;\theta) \mid X]\right)^2\right] \mid X\right\} + \mathrm{var}[f(X;\mid\theta) \mid X]$$

Estimation, random model

- Variance term
  - Do a computer experiment
  - This is specific for each X
  - Calculate variance of simulated values

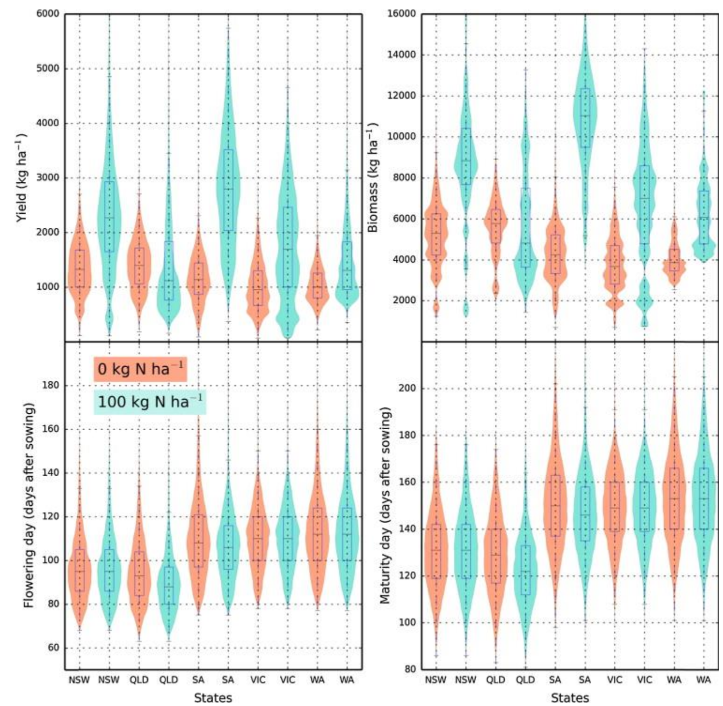
structure	ϴ	X	y(X)
model 1	θ_{11}, θ_{12},…	X_1, X_2	Y_{111},y_{121},…
:			
model M	θ_{M1}, θ_{M2},…	X_1, X_2	Y_{M11},y_{M21},…

calculate variance

- Bias term
  - Use hindcasts, compared to data
  - This estimates average over X

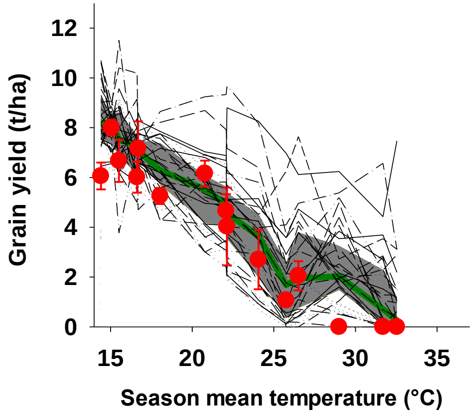
Random model in practice

- Several studies with parameter and/or input uncertainty
  - These studies estimate only the variance term.
    - Ignore structure uncertainty
    - Ignores bias term



Zhao et al. 2014.

- Many recent studies with model structure uncertainty (multi-model ensembles)
  - Estimate variance term
    - Ignore uncertainty in inputs and parameters
  - Error of mean of models estimates bias term
    - Small compared to variance term



Asseng et al. 2014

Features of $MSEP_{random}(X)$

- This is mean squared error for a specific X
- But averaged over distribution of predictors
- That is the trade-off
- Note that standard statistical prediction intervals in regression treat model parameters as random. Bayesian credible intervals also.

An example of the difference between fixed and random models

	AR 2009	AR 2010
$MSEP_{fixed}$		
model 1. MSEP	0.19	0.19
model 2. MSEP	2.02	2.02
$MSEP_{random}(X)$		
bias term	0.23	0.23
variance term	1.55	2.77
sum=$MSEP_{random}(X)$	1.78	3.00

Asseng, 2013

Conclusions

- $MSEP_{random}$ has both a variance term and a bias term, need to estimate both
  - Usual studies don't add bias
  - And only look at part of variance
- $MSEP_{random}(X)$ has important advantages
  - It shows how prediction uncertainty varies with the prediction situation (with X)
  - $MSEP_{random}$ allows separate estimation of effects of structure, input and parameter uncertainty

Going forward

- Estimate $MSEP_{random}(X)$ systematically
  - Can help answer question: is model good enough for this specific application
  - Useful even if only part of uncertainty taken into account
  - Can be compared with $MSEP_{fixed}$
- Further work needed
  - More experience with size of bias term relative to variance
  - A major problem is estimating parameter uncertainty. This may be very important

k22329576 fotosearch.com ©