

# Evaluating IAMs

Charlie Wilson

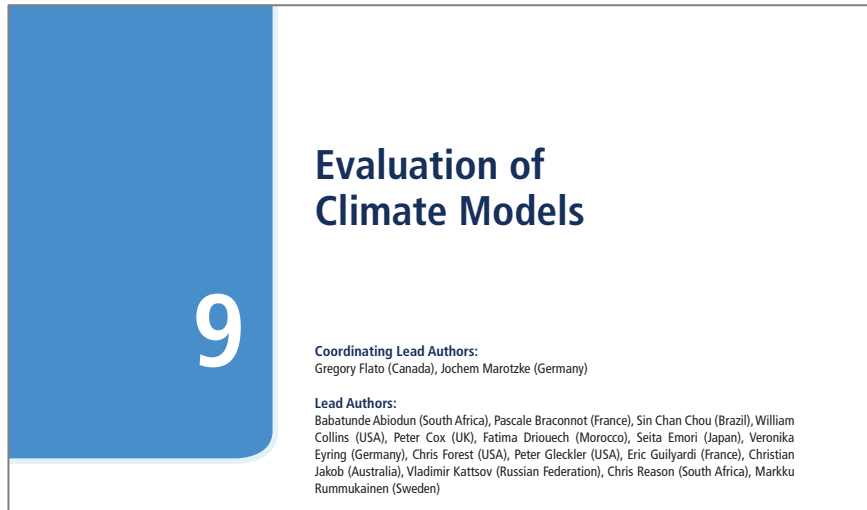
+ Celine Guivarch, Elmar Kriegler, Bas van Ruijven, Detlef van Vuuren,  
Volker Krey, Jana Schwanitz, Erica Thompson

*May 2020*

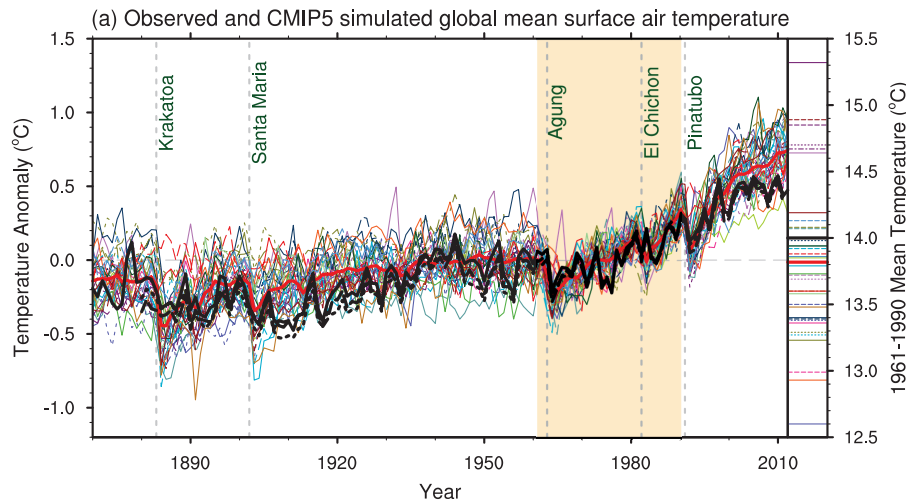
Expert & Stakeholder Workshop on:  
Robustness and legitimacy of models for climate policy assessment

# IAMs are increasingly influential on the real world ... but their fitness for purpose is not *visibly* evaluated

## Climate Models



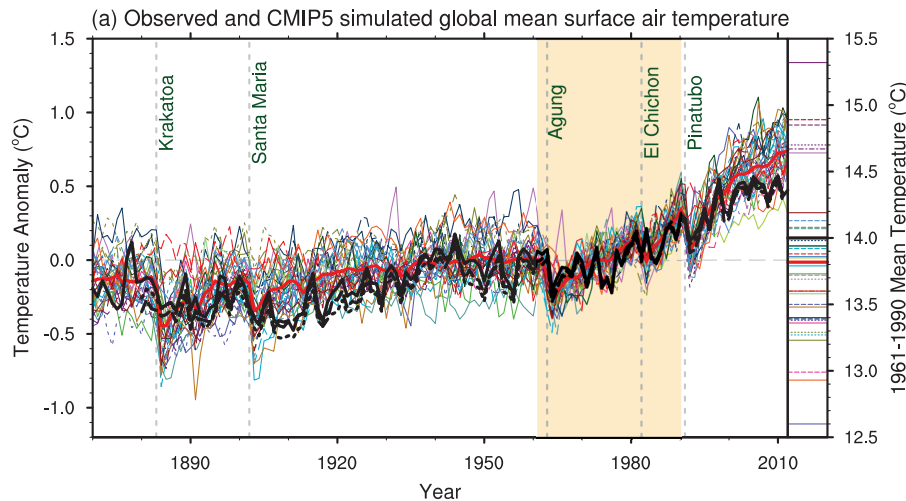
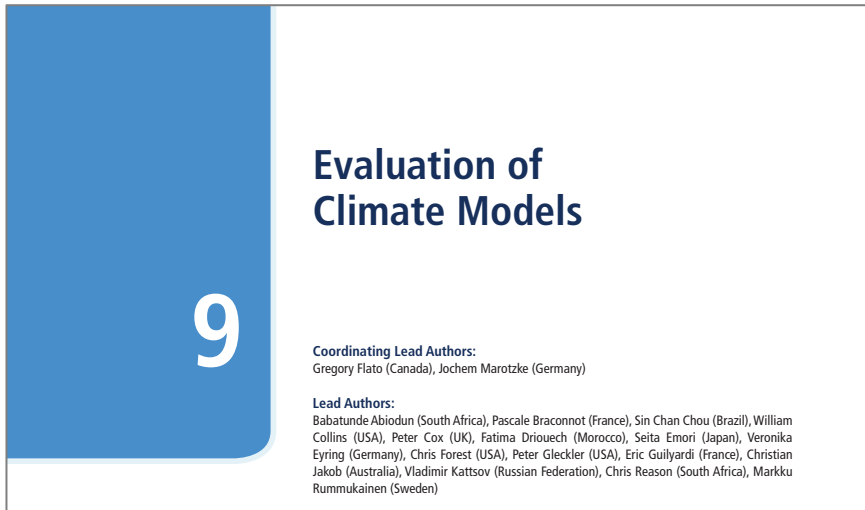
## Integrated Assessment Models



Flato et al. (2013). Evaluation of Climate Models. WG1 Contribution to IPCC Fifth Assessment Report.

# IAMs are increasingly influential on the real world ... but their fitness for purpose is not *visibly* evaluated

## Climate Models



Flato et al. (2013). Evaluation of Climate Models. WG1 Contribution to IPCC Fifth Assessment Report.

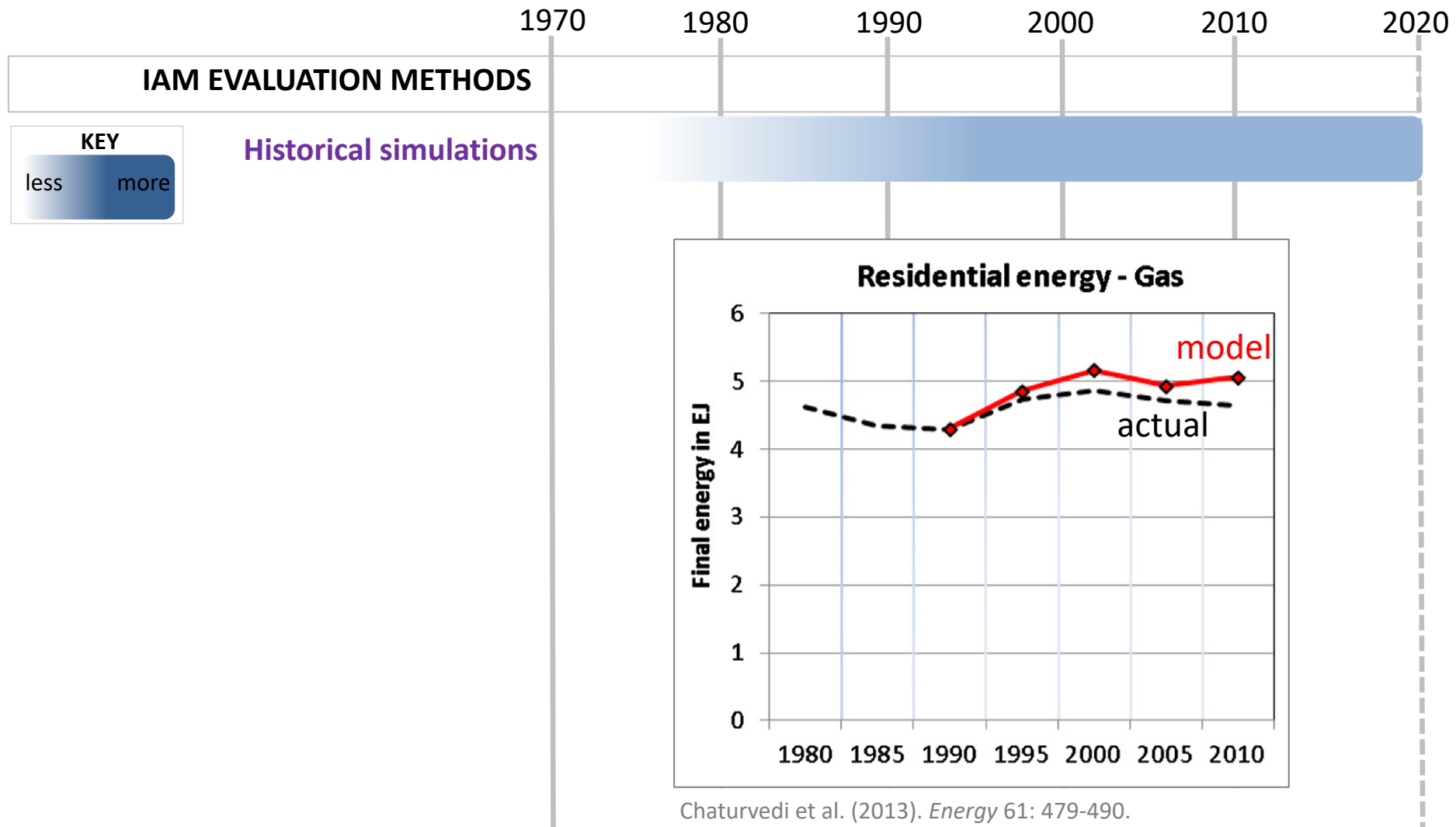
## Integrated Assessment Models

**evaluation =**  
*an open-ended process of testing and improving IAMs' usefulness as scientific tools for informing policy*

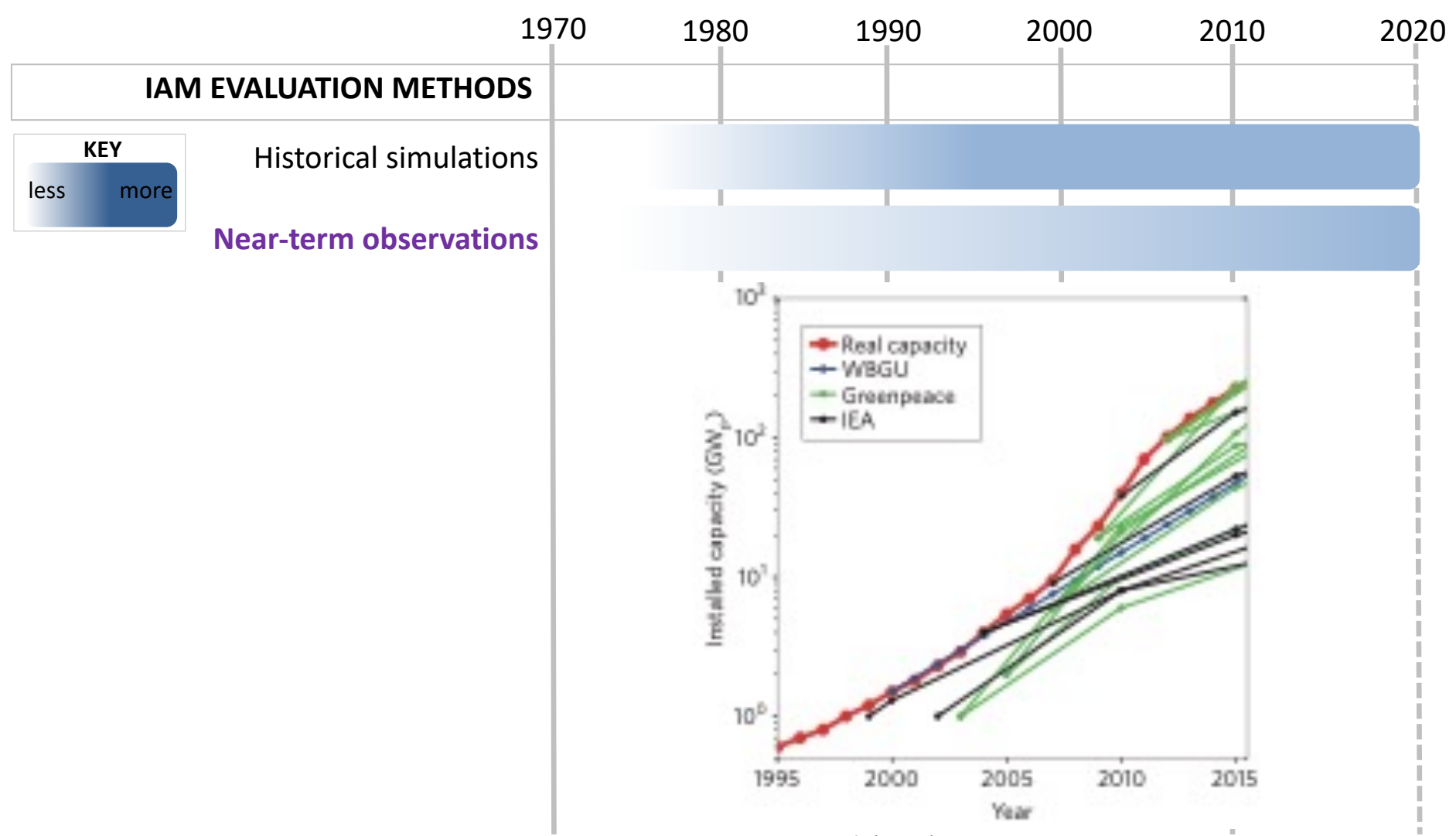
**structural validity =**  
*model is an accurate representation of the system response being modelled*

**behavioural validity =**  
*model outputs are consistent with observational data*

# IAM evaluation methods use observational data ...

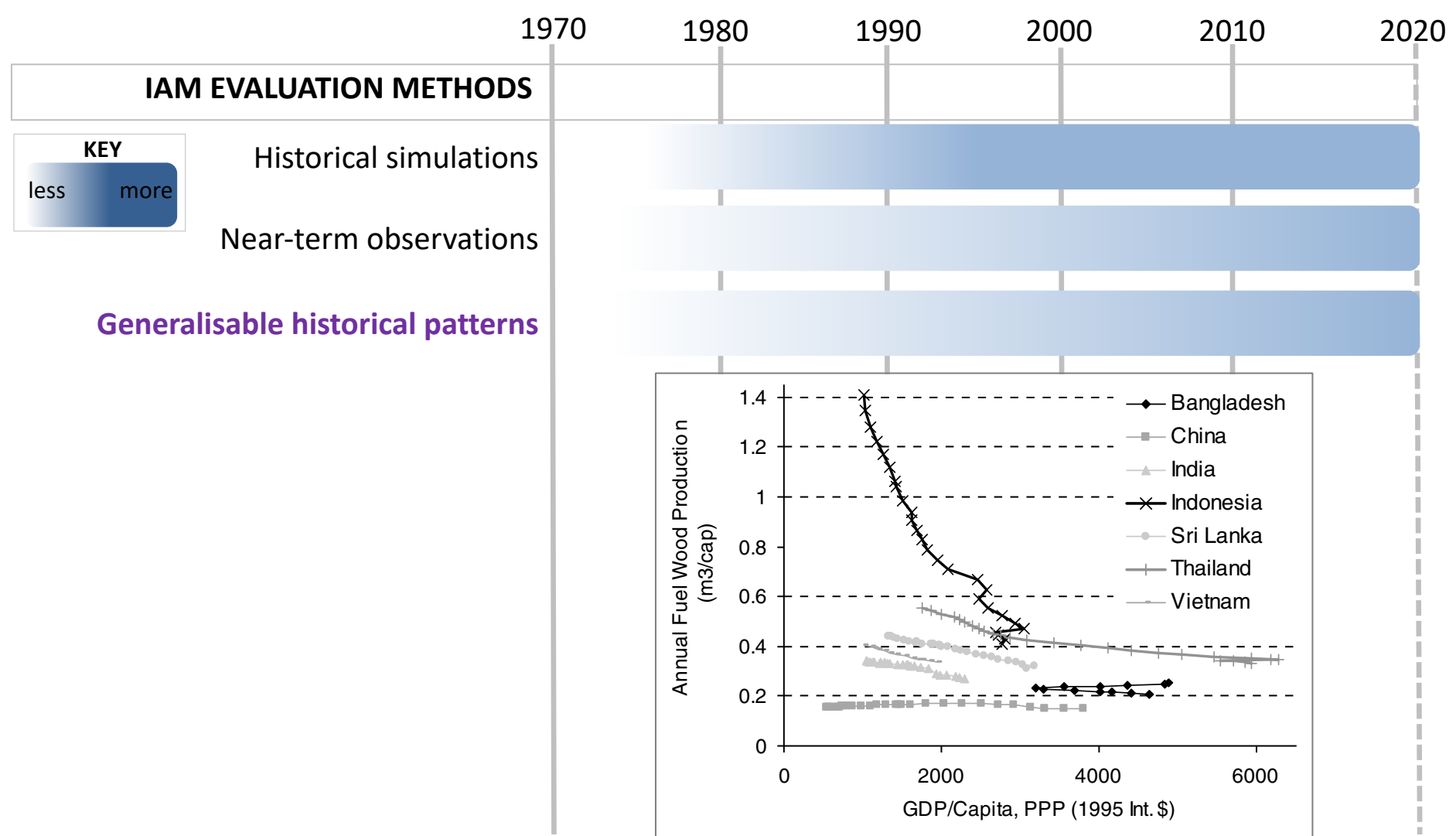


# IAM evaluation methods use observational data ...



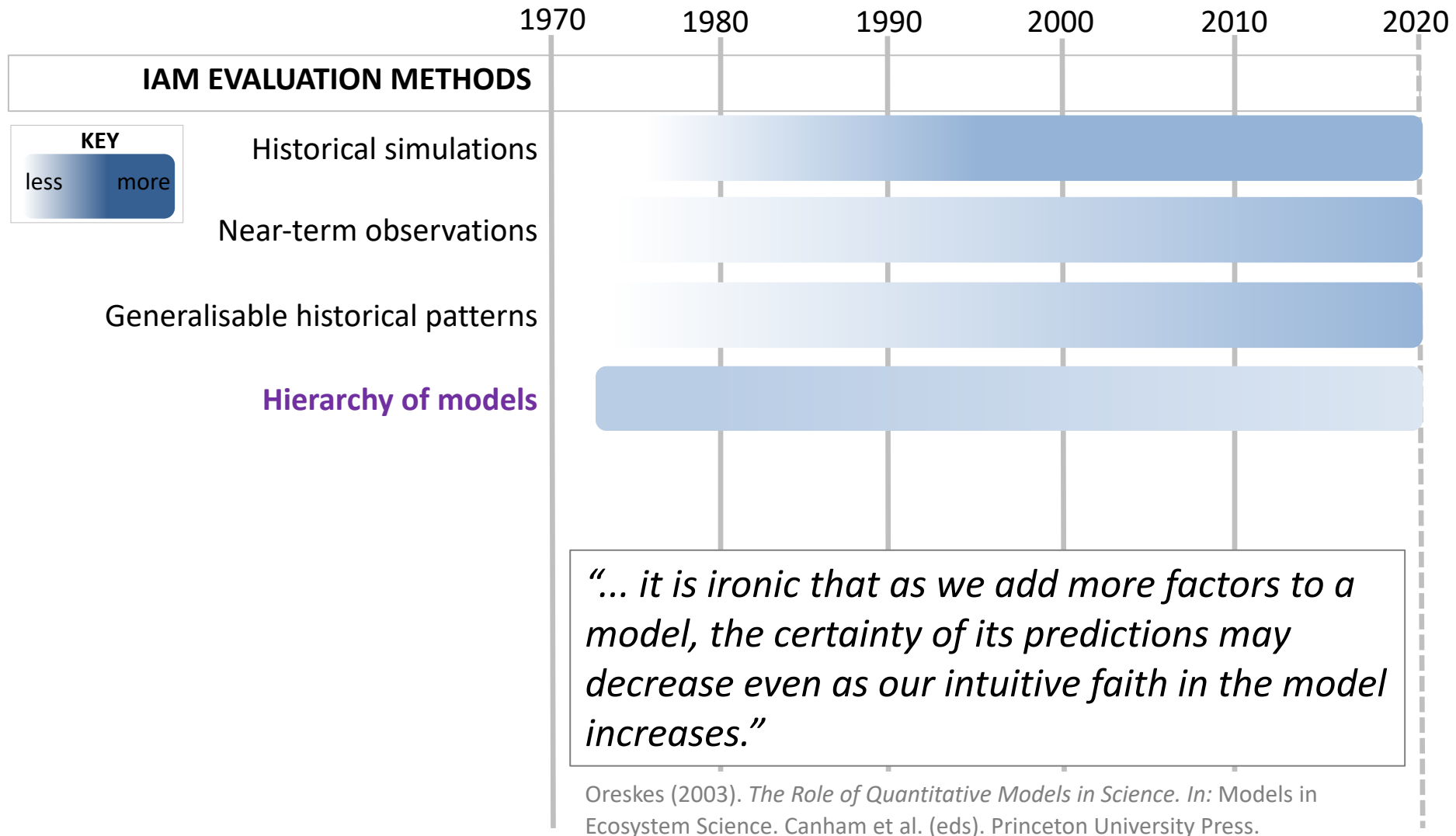
Creutzig et al. (2017). *Nature Energy* 2: 17140.

# IAM evaluation methods use observational data ...

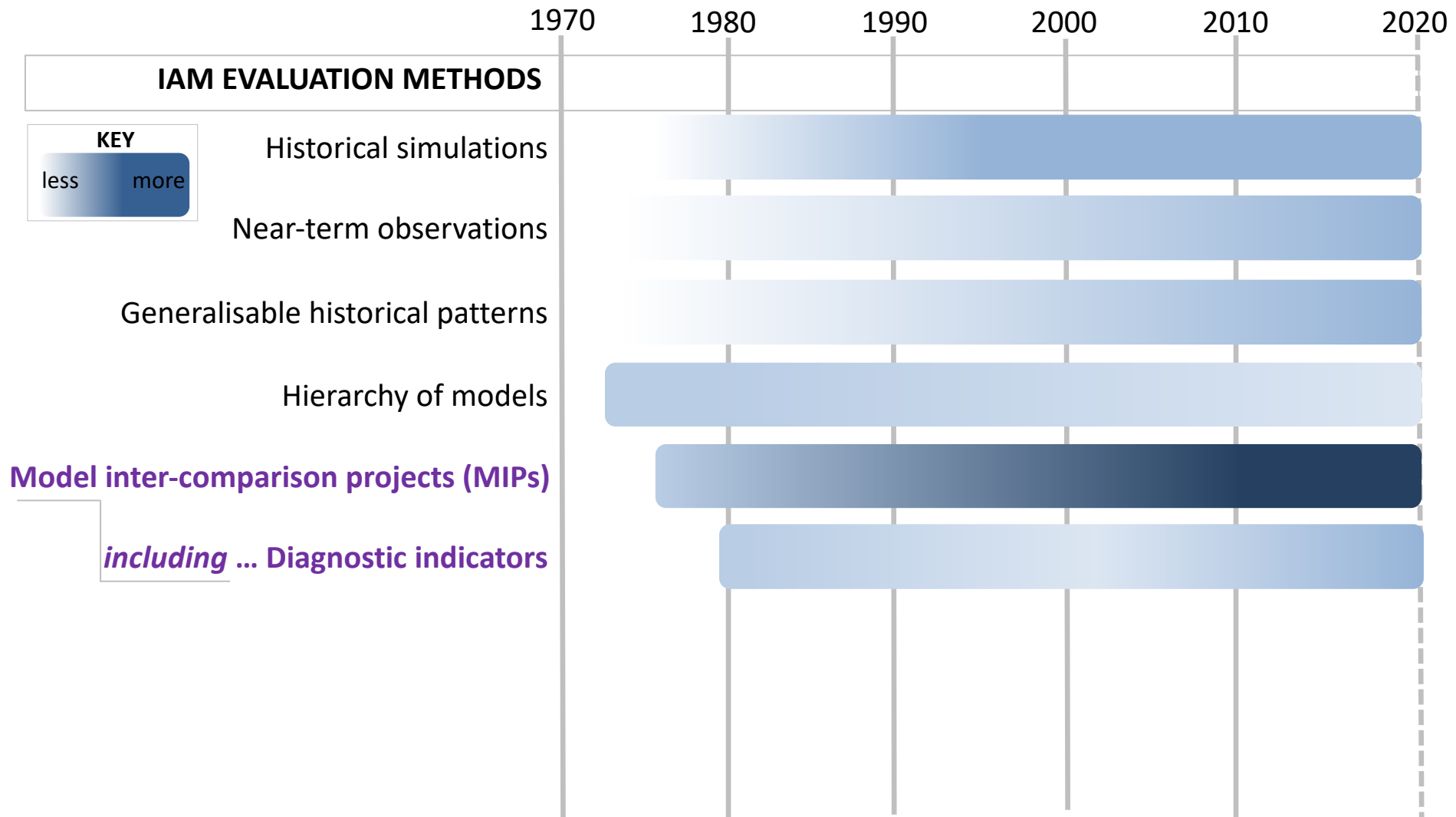


van Ruijven et al. (2008).  
*World Development* 36: 2801-2821.

# IAM evaluation methods use observational data, other models ...

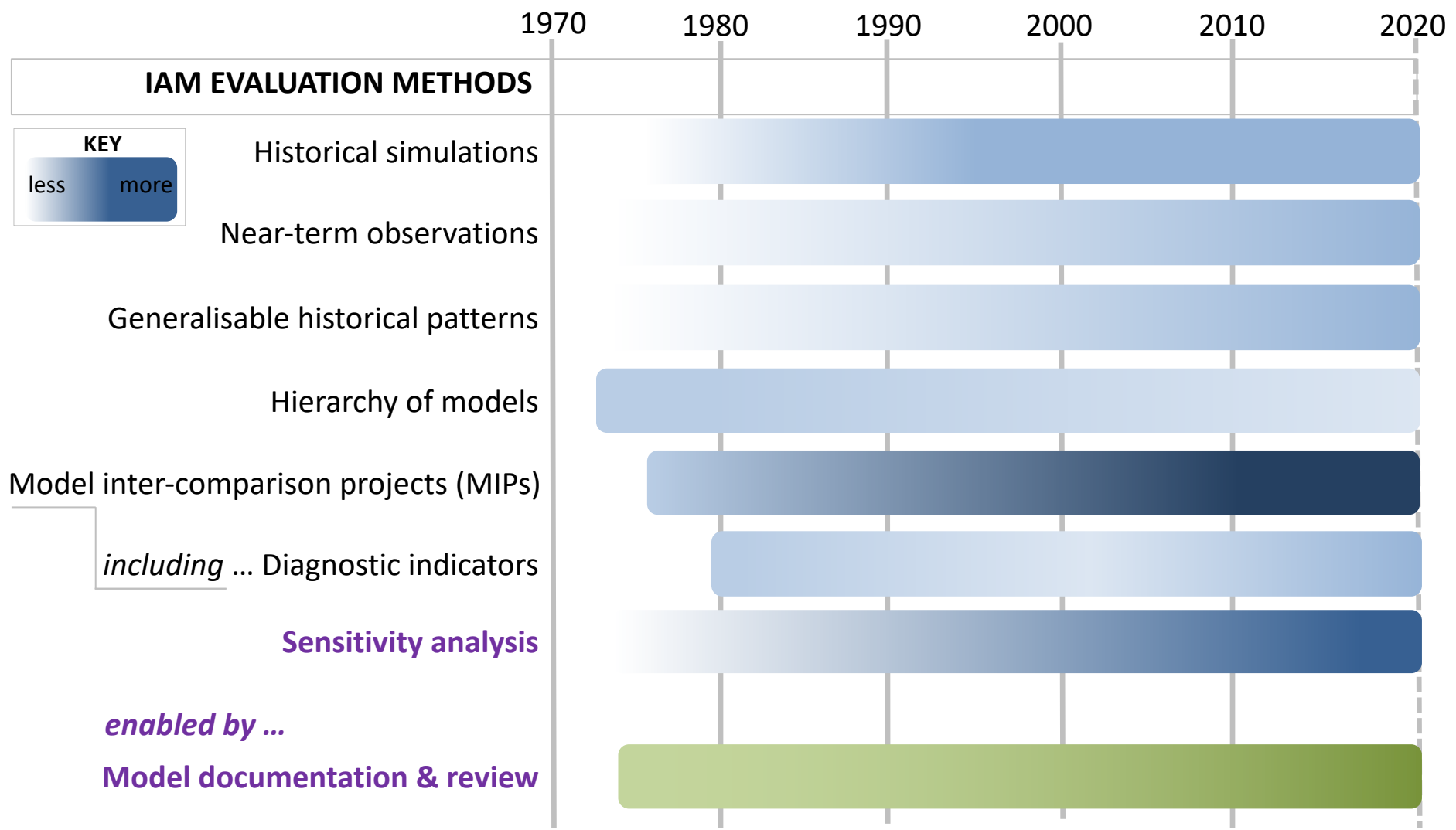


# IAM evaluation methods use observational data, other models ...

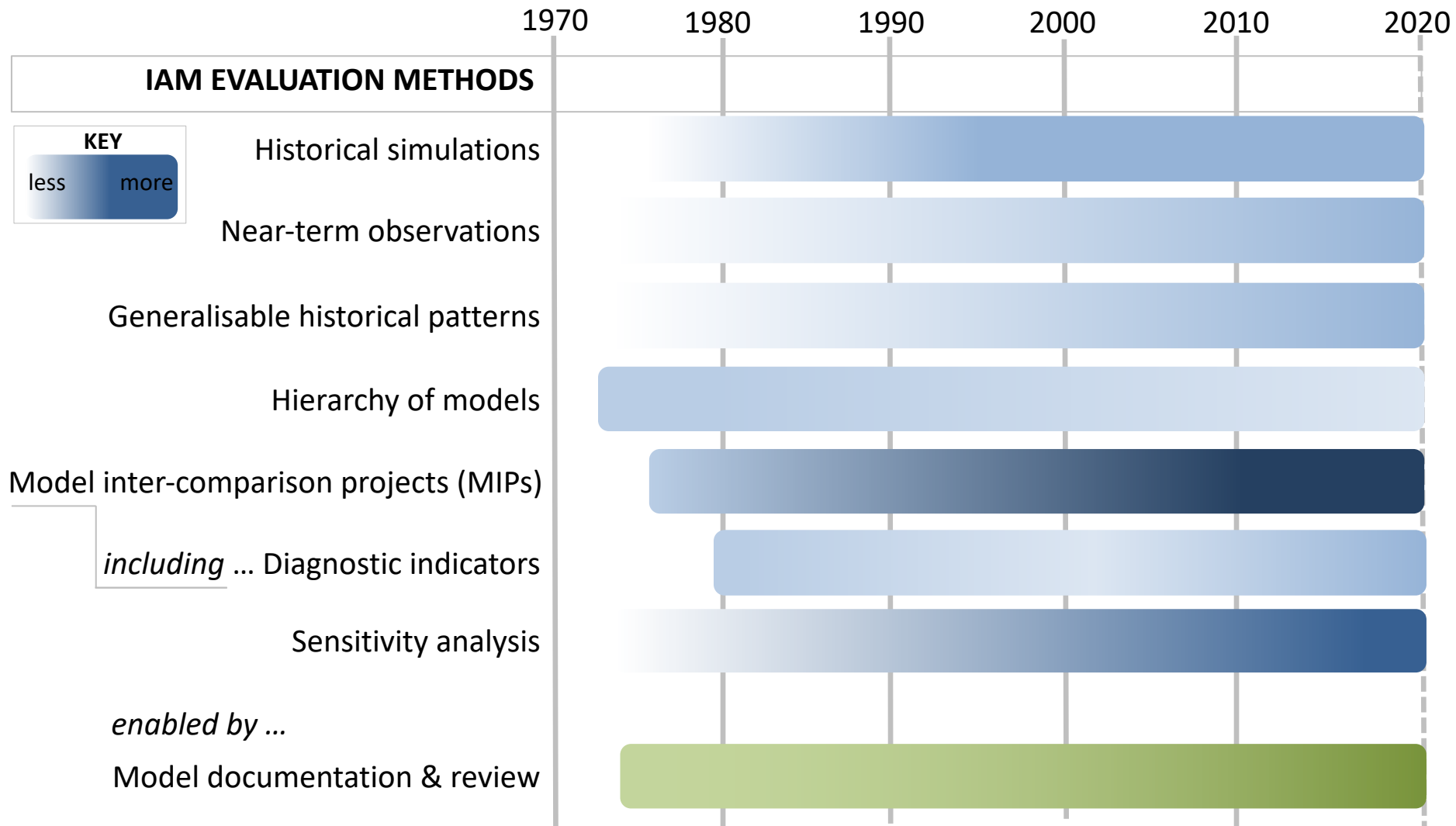




# IAM evaluation methods use observational data, other models, and internal checks



# IAM evaluation methods have been **applied with varying ‘intensity’** both historically and currently



# IAM evaluation methods have different and **complementary strengths** as well as limitations

## IAM EVALUATION METHODS

Historical simulations

Near-term observations

Generalisable historical patterns

Hierarchy of models

Model inter-comparison projects (MIPs)

*including* ... Diagnostic indicators

Sensitivity analysis

*enabled by ...*

Model documentation & review

## strengths

based on  
observations of how  
the real world works

## limitations

internal to the  
modelled world

# IAM evaluation methods have different and **complementary strengths** as well as limitations

## IAM EVALUATION METHODS

Historical simulations

Near-term observations

Generalisable historical patterns

Hierarchy of models

Model inter-comparison projects (MIPs)

*including* ... Diagnostic indicators

Sensitivity analysis

*enabled by ...*

Model documentation & review

## strengths

tests overall model  
performance

tests overall model  
performance

## limitations

only tests subset of  
causal mechanisms  
or spatial scales

# IAM evaluation methods have different and **complementary strengths** as well as limitations

## IAM EVALUATION METHODS

Historical simulations

Near-term observations

Generalisable historical patterns

Hierarchy of models

Model inter-comparison projects (MIPs)

*including* ... Diagnostic indicators

Sensitivity analysis

*enabled by* ...

Model documentation & review

## strengths

clear interpretability  
and attribution of  
results

clear interpretability  
and attribution of  
results

## limitations

hard to translate  
findings into model  
improvements

hard to translate  
findings into model  
improvements

# IAM evaluation methods have different and **complementary strengths** as well as limitations

## IAM EVALUATION METHODS

Historical simulations

Near-term observations

Generalisable historical patterns

Hierarchy of models

Model inter-comparison projects (MIPs)

*including* ... Diagnostic indicators

Sensitivity analysis

*enabled by ...*

Model documentation & review

## strengths

opens up evaluation  
to 3<sup>rd</sup> parties

opens up evaluation  
to 3<sup>rd</sup> parties

opens up evaluation  
to 3<sup>rd</sup> parties

## limitations

tacit learning within  
modelling teams

# Evaluation is essential for strengthening the **interpretability, credibility** and **relevance** of IAMs

## IAM EVALUATION METHODS

Historical simulations

Near-term observations

Generalisable historical patterns

Hierarchy of models

Model inter-comparison projects (MIPs)

*including* ... Diagnostic indicators

Sensitivity analysis

*enabled by ...*

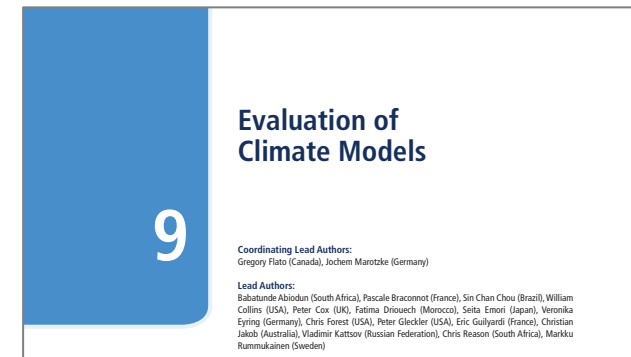
Model documentation & review

## Strengthening IAM evaluation:

**concerted:** from single to multiple methods in combination

**systematic:** from ad hoc to synthesis, protocols, reporting

**visible:**



Wilson et al. (2017). Evaluating Process-Based Integrated Assessment Models of Climate Change Mitigation.

IIASA Working Papers. Laxenburg, Austria, International Institute for Applied Systems Analysis (IIASA).

# Evaluating IAMs

Charlie Wilson

+ Celine Guivarch, Elmar Kriegler, Bas van Ruijven, Detlef van Vuuren,  
Volker Krey, Jana Schwanitz, Erica Thompson

*May 2020*

Expert & Stakeholder Workshop on:  
Robustness and legitimacy of models for climate policy assessment



# Extra Slides

**Table 2. IAM evaluation criteria and methods.**

Correspondence between criteria and methods is subjectively labelled as strong (✓), partial (~), or weak/none (✗) based on evidence presented in the text.

IAM evaluation methods   <	
---	--