Evaluating Model Analysis of Climate Change Mitigation

Charlie Wilson
November 2015, IAMC Conference

(Elmar Kriegler, Jana Schwanitz
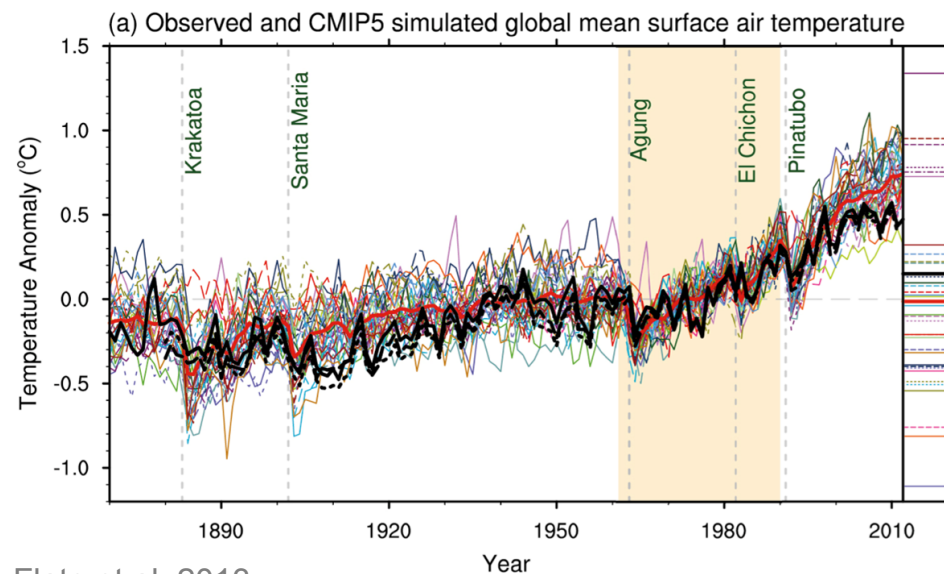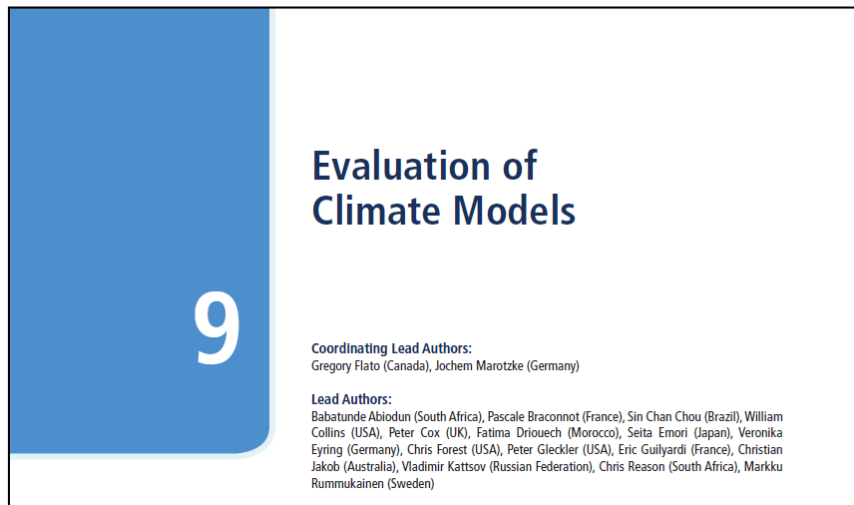Celine Guivarch, Volker Krey, Keywan Riahi, Detlef van Vuuren)

Tyndall°Centre®
for Climate Change Research

How are IAMs evaluated? **To what end?** Why is IAM evaluation less **visible** than climate model evaluation?
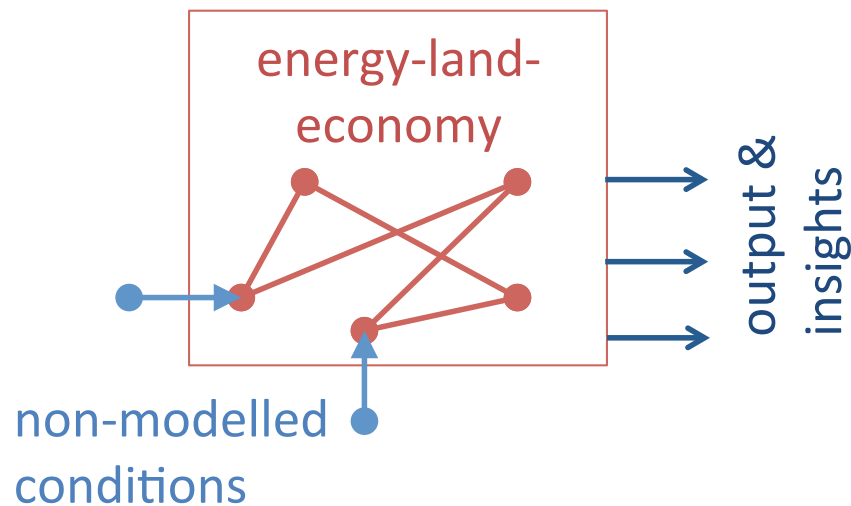
GCMs

IAMs

Flato et al. 2013.

Tyndall°Centre®
for Climate Change Research

Evaluation is about whether models generate the "**right behaviour** for the **right reasons**"
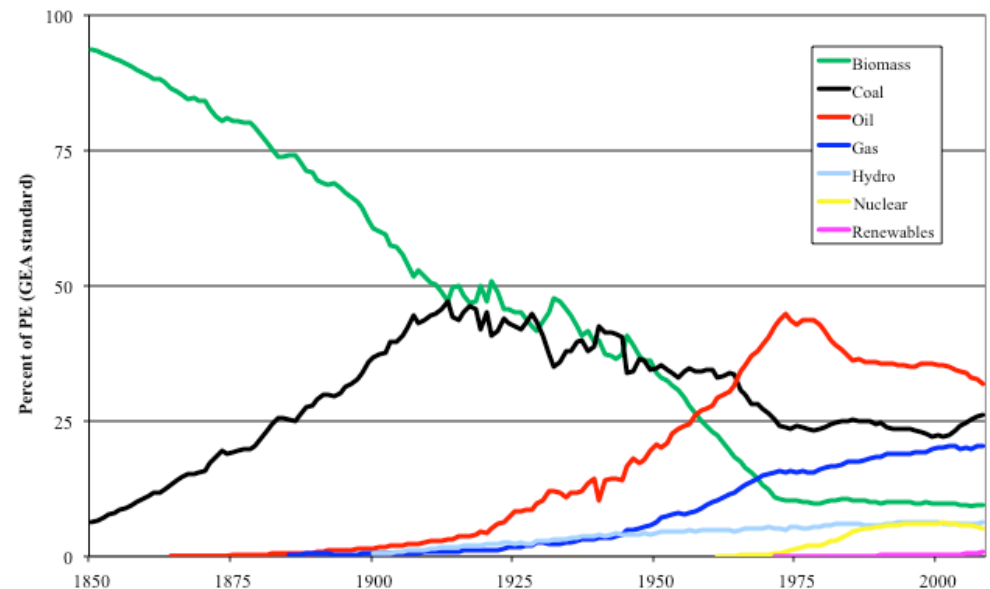
structural validity

model is an accurate representation of the system response being modelled
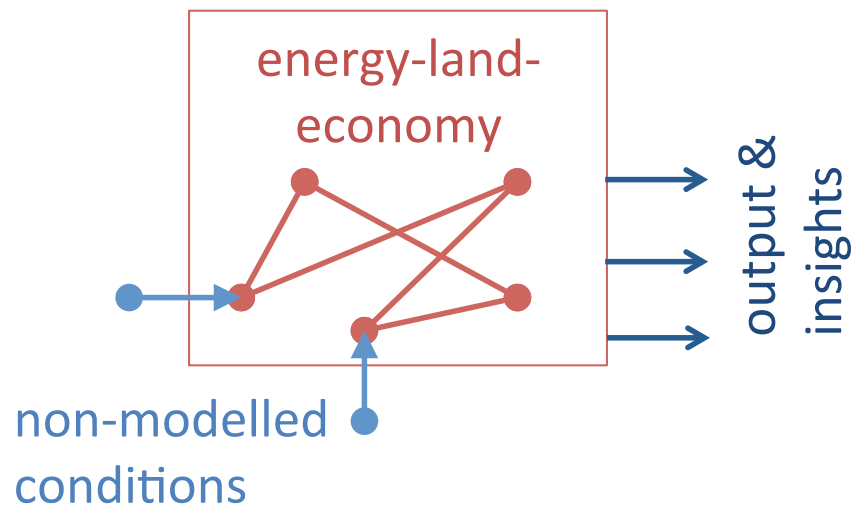
behavioural validity

model predictions are consistent with observational data

Oreskes et al. 1994. | Oreskes 1998. | Barlas 1996.

Behavioural validity can not be demonstrated for simulation models of **dynamic**, complex systems

structural validity

model is an accurate representation of the system response being modelled

energy-land-economy

output & insights

non-modelled conditions

behavioural validity

model predictions are consistent with observational data

✖ over-tuning, non-uniqueness

✖ limited to historical conditions

Oreskes et al. 1994 | Oreskes & Berlitz 2001 | Tebaldi & Knutti 2007 | Schindler & Hilborn 2015

Tyndall°Centre®
for Climate Change Research

Structural validity can not be demonstrated for simulation models of dynamic, **complex** systems
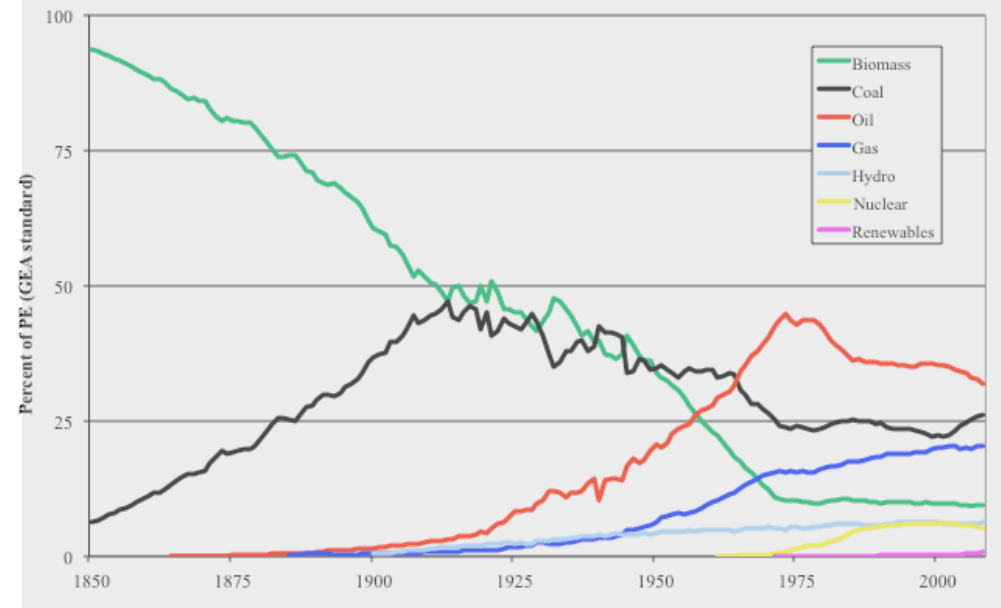
structural validity

model is an accurate representation of the system response being modelled

behavioural validity

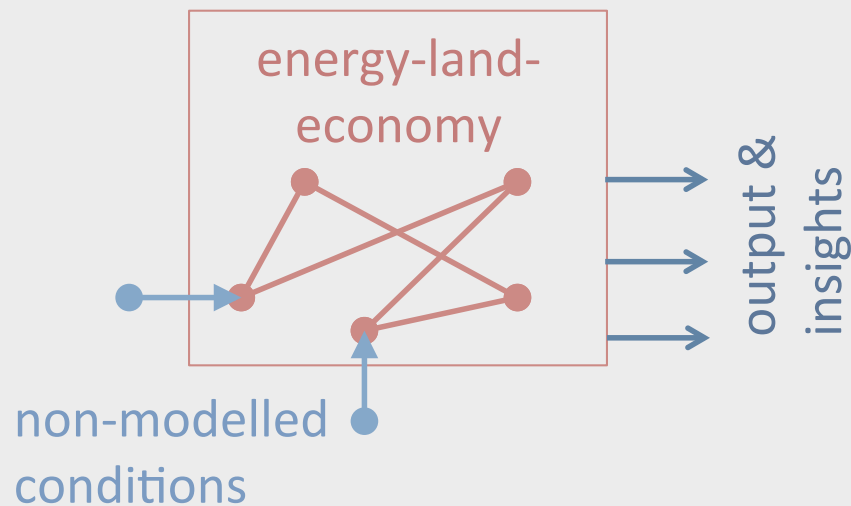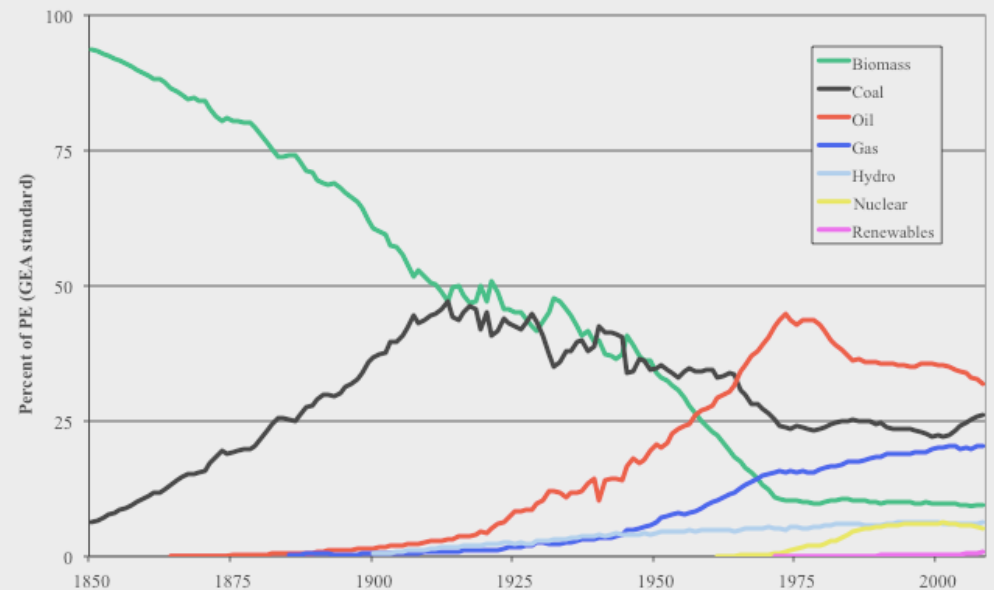*model predictions are consistent with observational data*

✖ irreducible uncertainties (data, parametric, structural)

✖ necessary simplifications

Oreskes 1998 | Schindler & Hilborn 2015

Tyndall°Centre®
for Climate Change Research

IAM evaluation is an **open-ended process** of testing, learning & improving a model and its performance

Evaluation criteria for IAMs

appropriateness — is model purpose and design consistent with the research question?

interpretability — are model results clearly interpretable in light of model structure and parameterisation?

verifiability — are model results repeatable or is model structure accessible to 3rd parties?

credibility — is model seen as good enough for its intended purpose by both users and modellers?

usefulness — do model insights help understand uncertainties, trade-offs, alternatives?

Tyndall°Centre®
for Climate Change Research

Different **evaluation methods** are used with IAMs, particularly to test structural validity
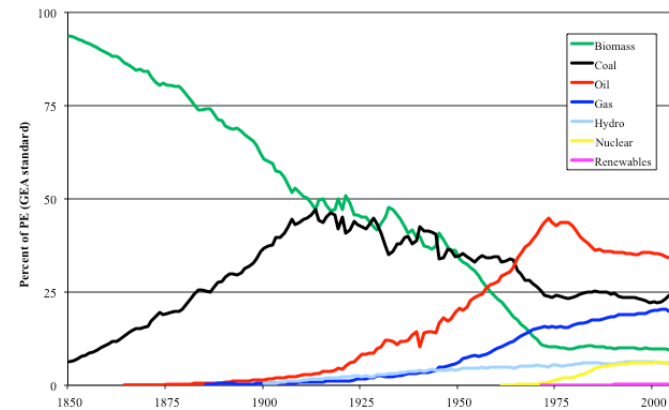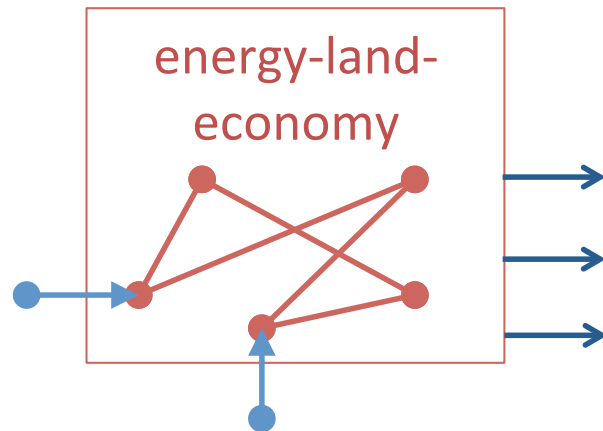
behavioural validity

structural validity

model checks

model inter-comparisons

historical trends

historical simulations

transparent documentation

diagnostic indicators

generalisable historical patterns

expert review

sensitivity analysis

simple models

Tyndall°Centre®
for Climate Change Research

"Climate models ... reproduce many important aspects of **observed climate** ..." [IPCC AR1 - AR5]
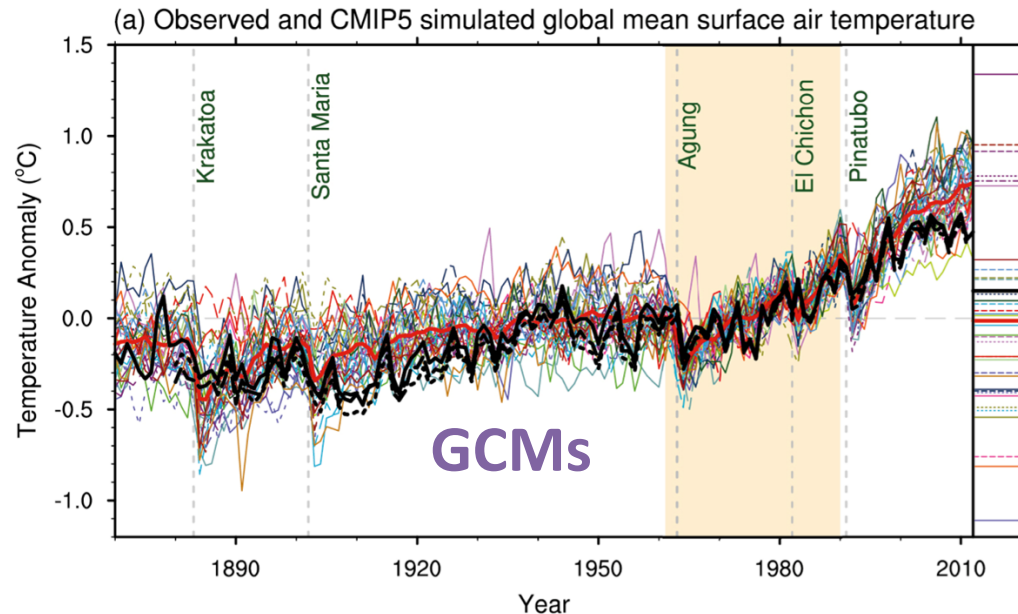
(a) Observed and CMIP5 simulated global mean surface air temperature

GCMs

- many simulated (un-tuned) quantities for different processes & scales
- statistical measures of performance

Flato et al. 2013.

Tyndall°Centre®
for Climate Change Research

Historical simulations are **not a common feature** of IAM evaluation

(a) Observed and CMIP5 simulated global mean surface air temperature

GCMs

IAMs
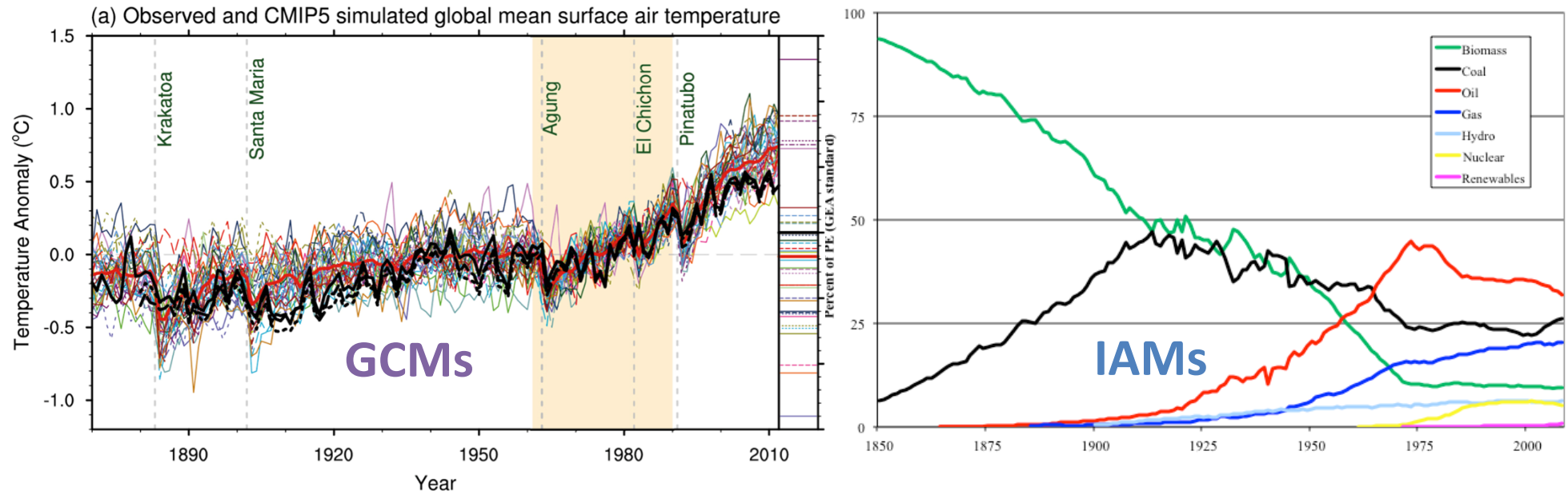
- many simulated (un-tuned) quantities for different processes & scales
- statistical measures of performance

- no long-run simulations of aggregate system variables
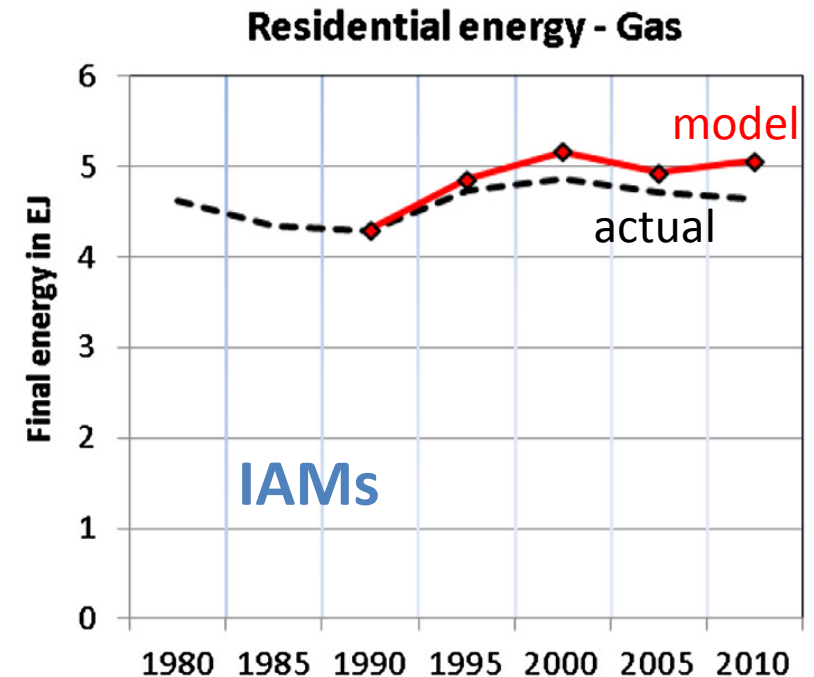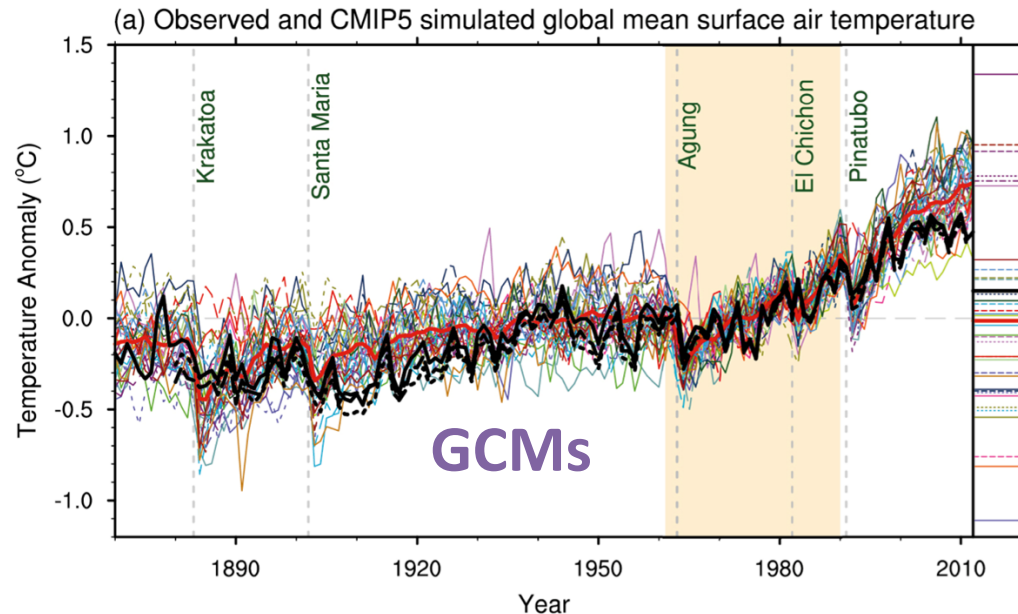
Chaturvedi et al. 2013 | van Ruijven et al. 2010 | Guivarch et al. 2009 | Beugin & Jaccard 2012

Tyndall°Centre®
for Climate Change Research

Historical simulations are **not a common feature** of IAM evaluation

(a) Observed and CMIP5 simulated global mean surface air temperature

GCMs

Residential energy - Gas

IAMs

model

actual

- many simulated (un-tuned) quantities for different processes & scales
- statistical measures of performance

- very limited in scope (process, time horizon)
- divergence -> Δ parameter

Chaturvedi et al. 2013 | van Ruijven et al. 2010 | Guivarch et al. 2009 | Beugin & Jaccard 2012

Tyndall°Centre®
for Climate Change Research

Historical simulations are **not a common feature** of IAM evaluation

(a) Observed and CMIP5 simulated global mean surface air temperature

GCMs

Temperature Anomaly (°C) vs Year — Krakatoa, Santa Maria, Agung, El Chichon, Pinatubo

Residential energy - Gas

model

actual

IAMs

Final energy in EJ vs year (1980–2010)

- many simulated (un-tuned) quantities for different processes & scales
- statistical measures of performance

- very limited in scope (process, time horizon)
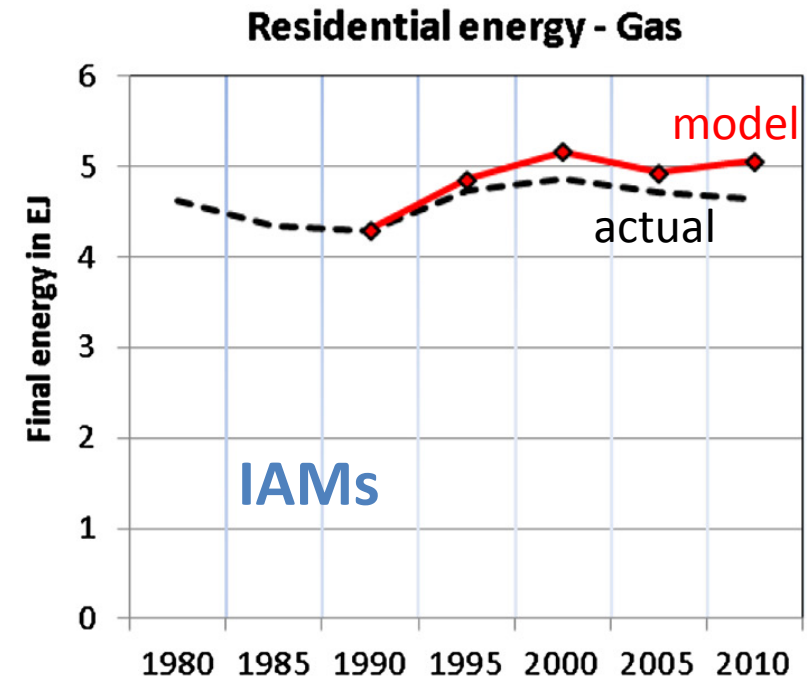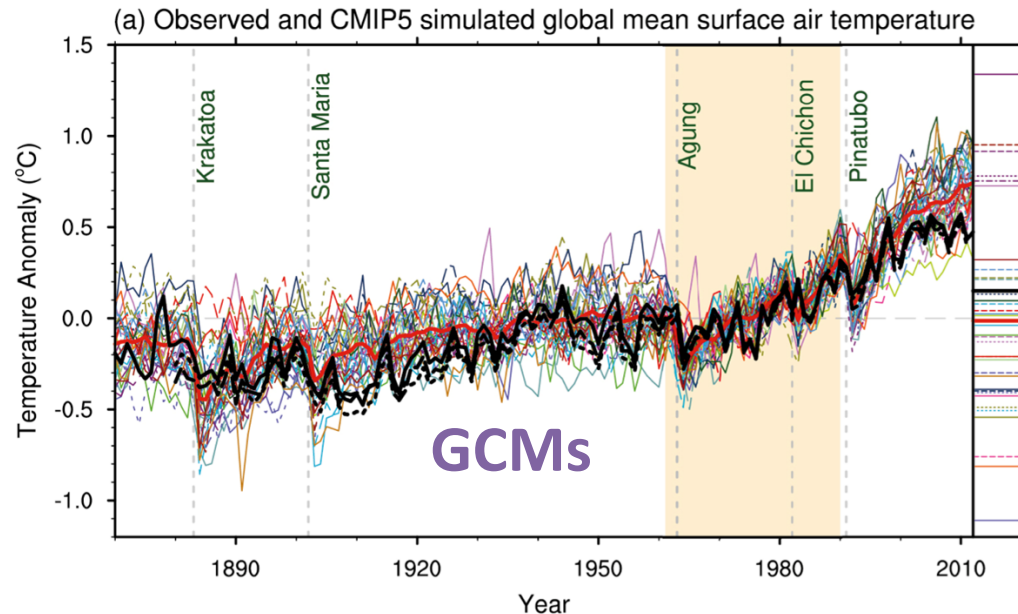- divergence -> Δ parameter

similar issues with behavioural validity testing

over-tuning, non-uniqueness
limited historical conditions

Historical simulations are **not a common feature** of IAM evaluation

Residential energy - Gas

model

actual

IAMs

- very limited in scope (process, time horizon)
- divergence -> Δ parameter
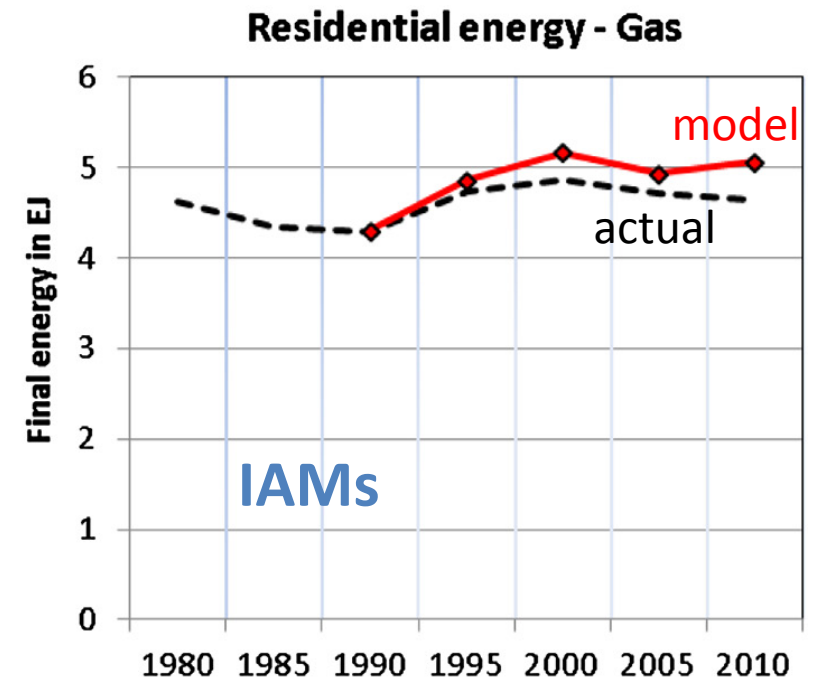
lack of good observational data heterogeneous causal processes (normative design)

issues unique to IAMs?

Chaturvedi et al. 2013 | van Ruijven et al. 2011 | Schwanitz 2013 | van Vuuren et al. 2010

Tyndall°Centre®
for Climate Change Research

The **historical record** can be used for IAM evaluation in other ways - **generalisable historical patterns**
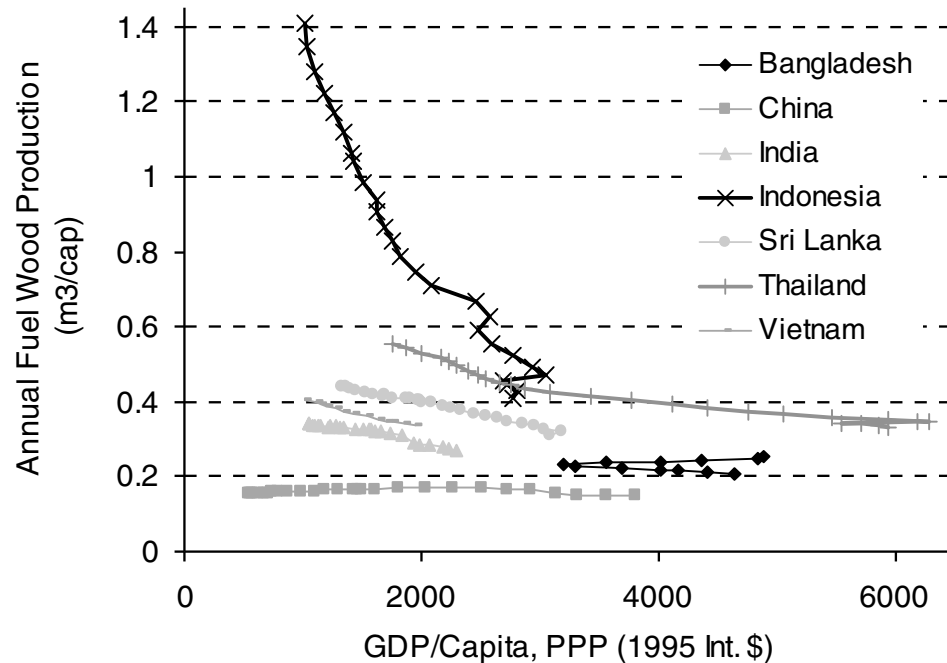
- useful learning exercise ... but what then?
- no clear methodology or metrics
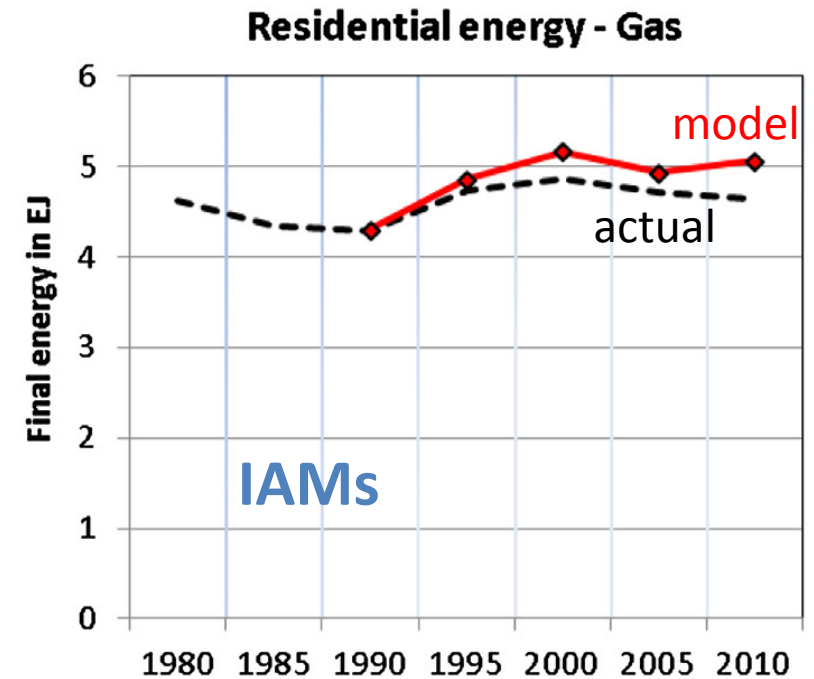
- very limited in scope (process, time horizon)
- divergence -> Δ parameter

van Ruijven et al. 2008 | Wilson et al. 2012 | Schwanitz 2013 | Leimbach et al. 2015

Tyndall°Centre®
for Climate Change Research

Simple models help understand representations of key processes embedded in more complex models
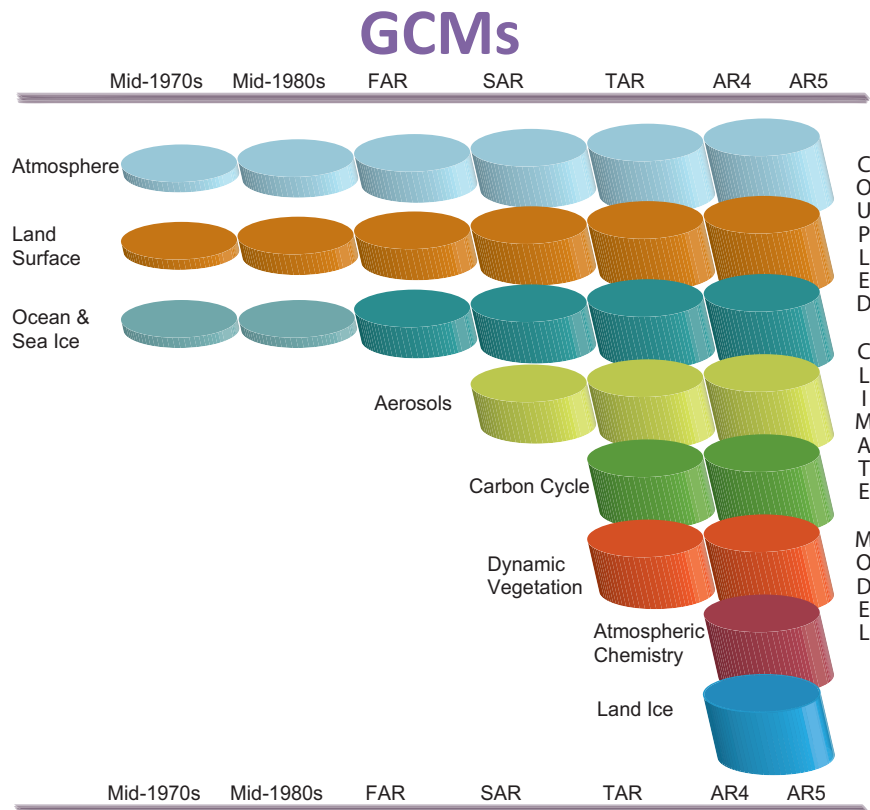

GCMs

	Mid-1970s	Mid-1980s	FAR	SAR	TAR	AR4	AR5
Atmosphere							
Land Surface							
Ocean & Sea Ice							
Aerosols							
Carbon Cycle							
Dynamic Vegetation							
Atmospheric Chemistry							
Land Ice							

COUPLED CLIMATE MODEL

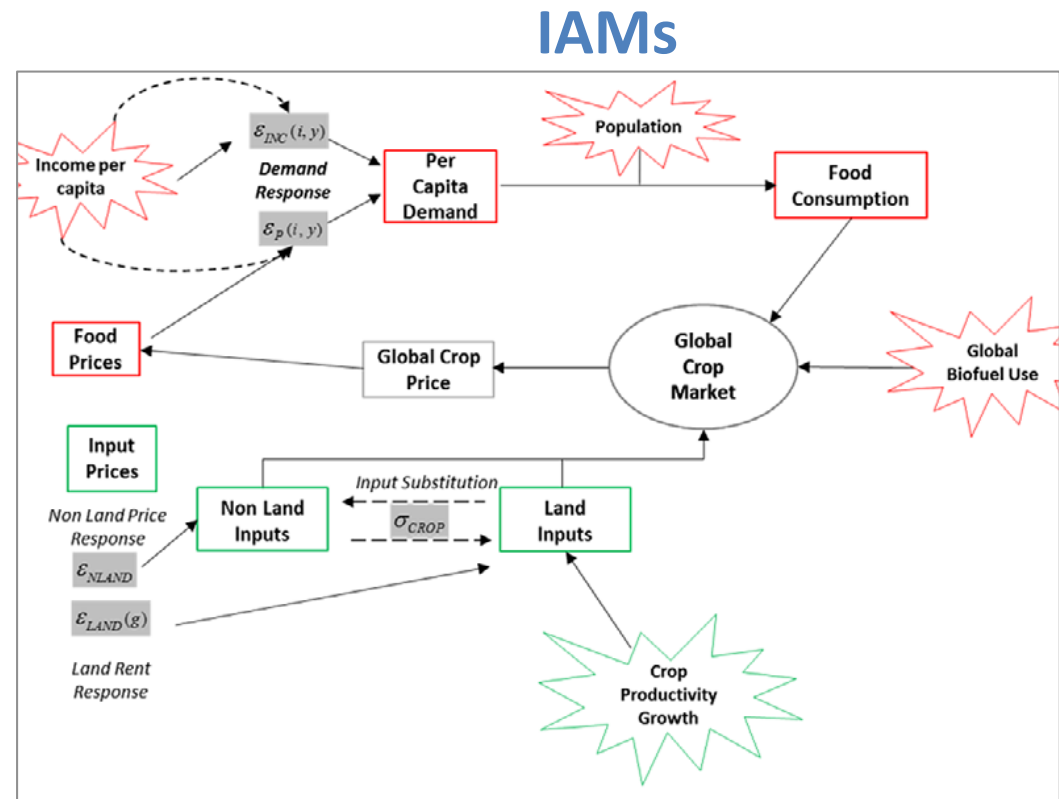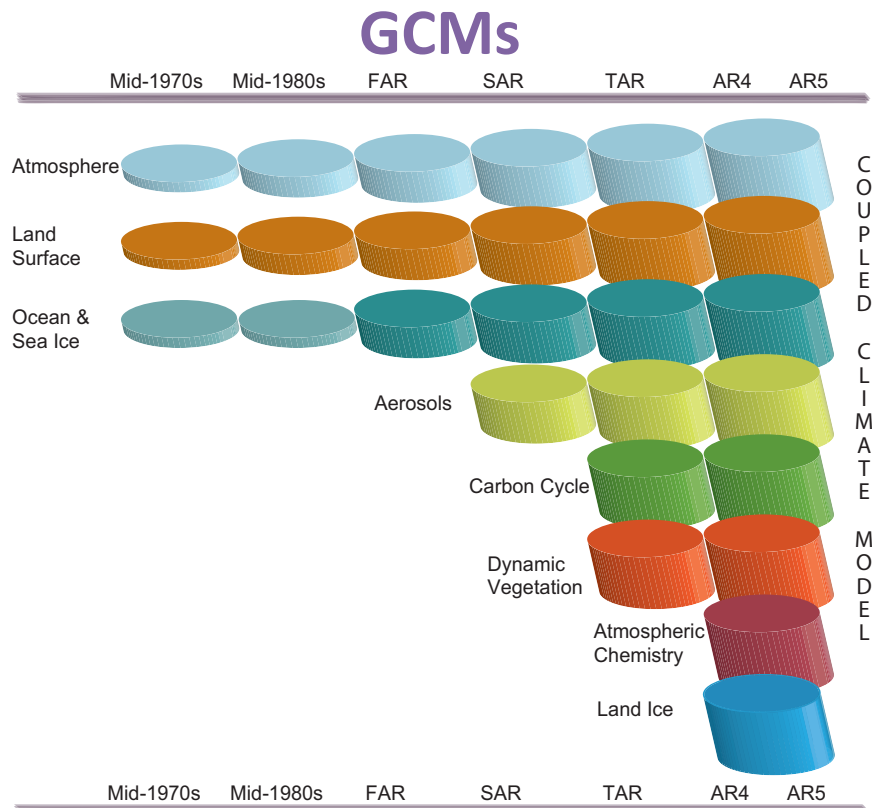*"A complex model may be more realistic, yet … as we add more factors to a model, the certainty of its predictions may decrease even as our intuitive faith in the model increases."*

- elegance vs. elaboration
- simpler models preserved in a 'hierarchy of models'

Held 2005 | Frame et al. 2005 | Stocker et al. 2013 | Stainforth et al. 2007 | Oreskes 2003

Tyndall°Centre®
for Climate Change Research

Simple models help understand representations of key processes … but are not common in IAMs

GCMs

IAMs

- elegance vs. elaboration
- simpler models preserved in a 'hierarchy of models'

- 'SIMPLE' global agriculture model - biophysical, economic
- historical simulations (1961-2000)

Baldos & Hertel 2013

Tyndall°Centre®
for Climate Change Research

Model inter-comparison projects explore structural uncertainty (across different model representations)

GCMs e.g., CMIP5

Snow cover extent change

- harmonised experiments & results
- model performance metrics

Collins et al. 2013 | Flato et al. 2013 | Taylor et al. 2012 | Parker 2013 | Pirtle et al. 2010

Tyndall°Centre®
for Climate Change Research
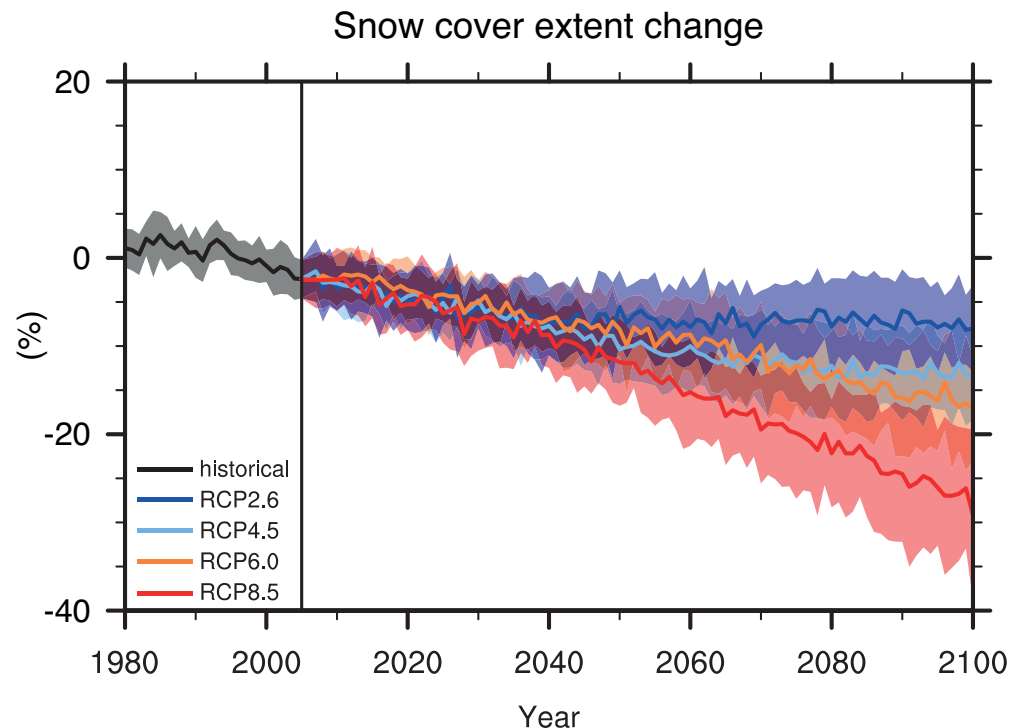
Model inter-comparison is a long tradition for IAMs (9 MIPs contributed 95% of AR5 mitigation scenarios)
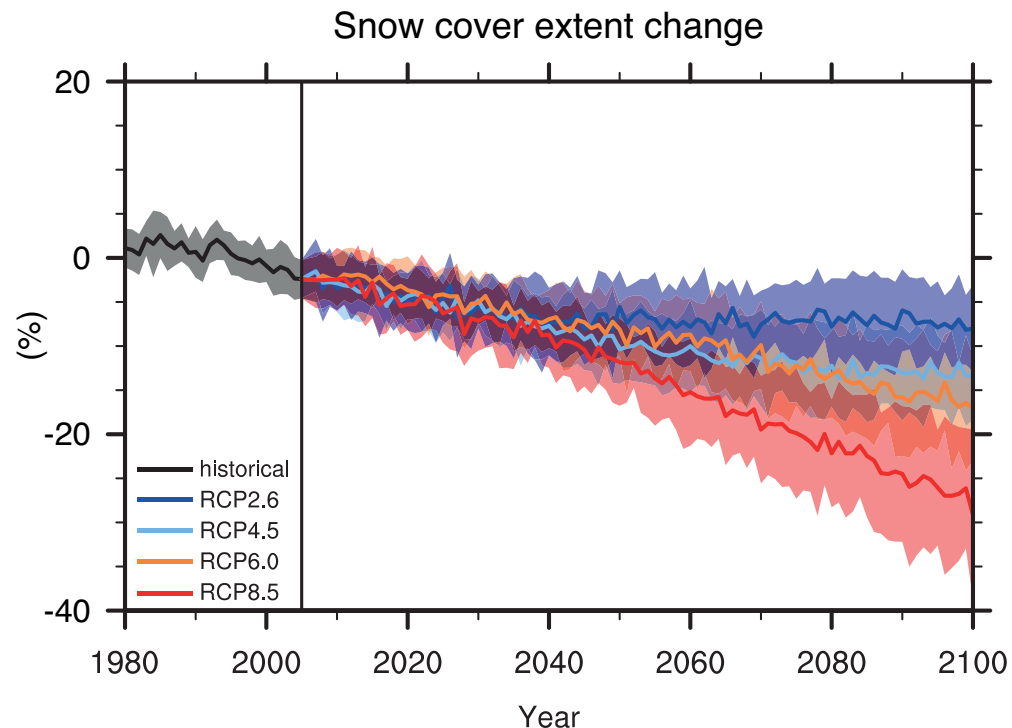
GCMs e.g., CMIP5

Snow cover extent change

IAMs

$50 carbon tax (2010), increasing 4% per year−−World

- harmonised experiments & results
- model performance metrics
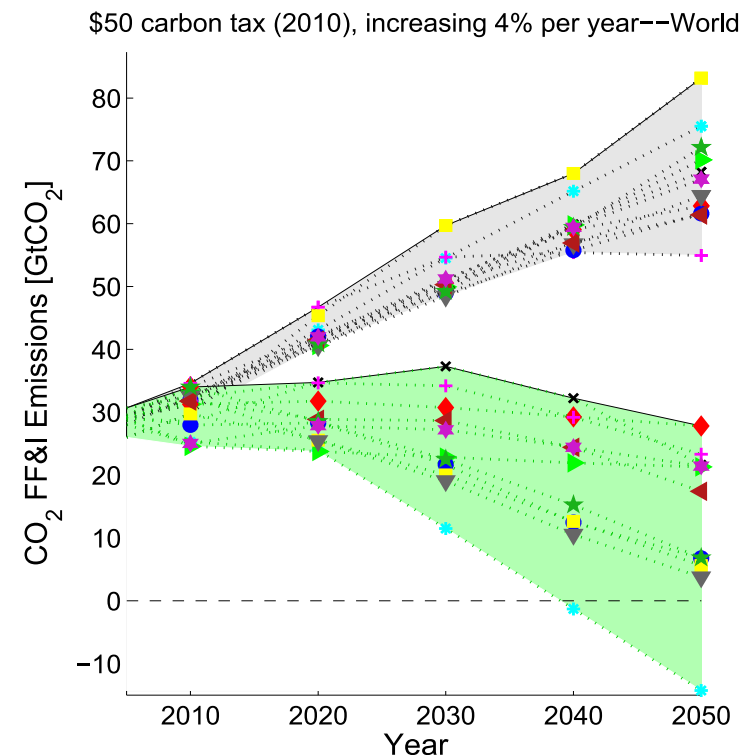
- emphasis on robust results
- diagnostic indicators (recent)
- *link structure <-> behaviour*

Kriegler et al. 2015. | Clarke et al. 2014 | Riahi et al. 2015 | Smith et al. 2015

Tyndall°Centre®
for Climate Change Research

Evaluation research for GCMs is generally more **developed** and **prominently reported** than for IAMs

evaluation method

GCMs vs. IAMs

GCM vs. IAM differences

historical
simulations

generalisable
historical patterns

simple
models

model inter-
comparisons

(1) modelled system

- underlying principles

- observational data

(2) domain of application

- uniqueness of insights

- expertise of policy users

+ sensitivity analysis
+ expert review
+ documentation ...

Tyndall°Centre®
for Climate Change Research

Each evaluation method has characteristic **strengths** and **weaknesses**

evaluation method	strengths	weaknesses
historical simulations	e.g., use of observations	e.g., limited applicability (time horizon, processes)
generalisable historical patterns	e.g., use of observed dynamics	e.g., unclear implications for structural validity
simple models	e.g., insights robust to structural uncertainty	e.g., attribution of divergence to model differences
model inter-comparisons	e.g., understanding of key system processes	e.g., lack of realism

+ sensitivity analysis
+ expert review
+ documentation

Each evaluation method has **strengths** & **weaknesses**
... and contributes more to certain **evaluation criteria**

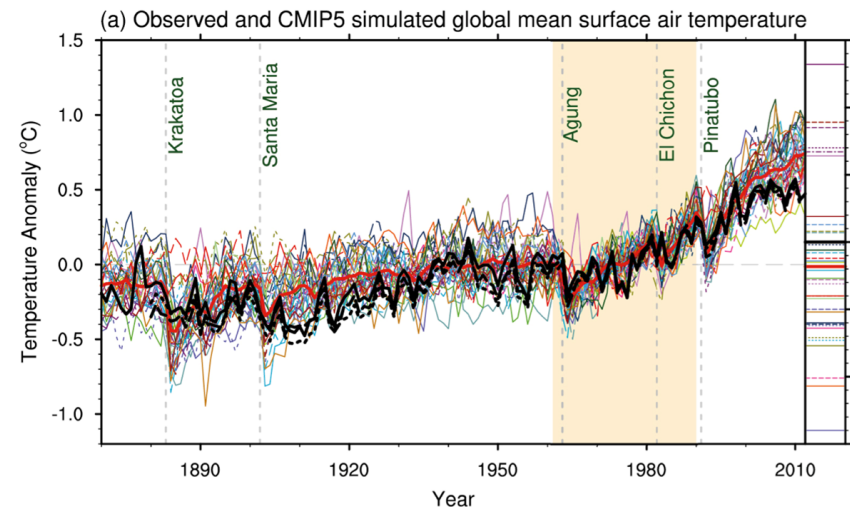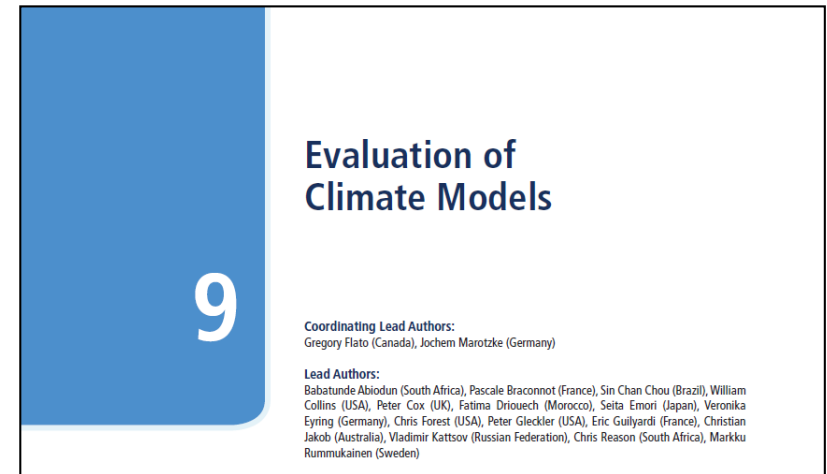|	appropriate-ness	interpret-ability	verifiability	credibility	usefulness
evaluation method			**evaluation criteria**		
historical simulations	~	~		✔	
generalisable historical patterns		~		~	
simple models	✔	✔	~	~	
model inter-comparisons		✔		✔	✔

+ sensitivity analysis
+ expert review
+ documentation ...

Tyndall°Centre®
for Climate Change Research

Conclusion: **Systematic** & more **prominent** evaluation effort to strengthen and maintain **confidence** in IAMs

- **systematic**:
  multiple methods concurrently

- **prominent**:
  concerted, synthesis products

- **learning**:
  insights from GCMs

*ongoing articulation of the grounds
on which IAMs can be declared
good enough for their intended uses*

Evaluation of Climate Models

9

Coordinating Lead Authors:
Gregory Flato (Canada), Jochem Marotzke (Germany)

Lead Authors:
Babatunde Abiodun (South Africa), Pascale Braconnot (France), Sin Chan Chou (Brazil), William Collins (USA), Peter Cox (UK), Fatima Driouech (Morocco), Seita Emori (Japan), Veronika Eyring (Germany), Chris Forest (USA), Peter Gleckler (USA), Eric Guilyardi (France), Christian Jakob (Australia), Vladimir Kattsov (Russian Federation), Chris Reason (South Africa), Markku Rummukainen (Sweden)

(a) Observed and CMIP5 simulated global mean surface air temperature

Evaluating Model Analysis of Climate Change Mitigation

Charlie Wilson
November 2015, IAMC Conference

(Elmar Kriegler, Jana Schwanitz
Celine Guivarch, Volker Krey, Keywan Riahi, Detlef van Vuuren)

Tyndall°Centre®
for Climate Change Research