When Forecasts Meet Reality: Assessing Climate Damage Functions

Johannes Emmerling[1,2] Paul Waidelich[3] Matthieu Bellon[4] Emanuele Massetti[1,2]

[1]CMCC Foundation - Euro-Mediterranean Center on Climate Change
[2]RFF-CMCC European Institute on Economics and the Environment
[3]ETH Zurich, Energy Technology and Policy Group
[4]International Monetary Fund

ECEMP Conference 2025
Brussels, October 16, 2025

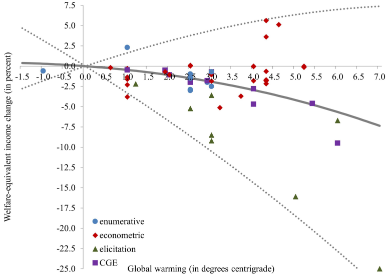
The Climate Damage Function Challenge

Growing divergence in climate impact estimates:

- Dell et al. (2012), Burke et al. (2015), Kotz et al. (2024), Bilal & Känzig (2024)
- Social Cost of Carbon: \$50 to \$2,500 per tCO2
- Order of magnitude higher than bottom-up approach

Three fundamental challenges:

1. True counterfactual GDP growth path unknown
2. How do damage functions compare?
3. What are the empirical/implicit climate damages today?



Source: Tol (2024)

Our Approach: Using IMF Forecasts as Climate-Naive Counterfactuals

Key insight:

- IMF World Economic Outlook: extensive economic fundamentals
- No systematic climate projections (especially pre-2018)
- Natural disasters included via historical probabilities only
- $\Rightarrow$ Ideal "climate-naive" counterfactuals

Our methodology addresses three challenges:

1. **Identification:** Placebo test validates that climate impacts are causal
2. **Validation:** Which damage functions explain forecast errors?
3. **Nowcasting:** What are historical climate impacts by country?

Data: IMF WEO Forecasts 1990-2023

IMF WEO Database:

- 180+ countries, biannual forecasts
- Up to 5 years ahead
- Country-specific modeling of varying complexity
- Desk economists + global model

Climate Data (Gortan et al. 2024):

- Population-weighted temperature (ERA5)
- Precipitation, extreme indices
- Global mean surface temperature (GMST)

Focus period: 2011-2019

- Excludes financial crisis and COVID

Variable	Mean	SD
Temperature (°C)	19.9	7.1
Precipitation (mm)	1,239	844
GMST change (°C)	0.98	0.24
Real GDP growth	3.1%	6.1%
IMF forecast (1y)	4.2%	4.0%

Challenge 1: Are Climate-Growth Relationships Causal?

Identification concern: Temperature-growth correlations might reflect omitted country-specific trends

Standard approach (potentially biased):

$$g_{it} = f(C_{it}) + \gamma_i + \tau_t + \underbrace{\varepsilon_{it}}_{\text{includes omitted } X_{it}}$$

If $Cov(C_{it}, X_{it}) \neq 0 \Rightarrow$ biased estimates

Challenge 1: Are Climate-Growth Relationships Causal?

Identification concern: Temperature-growth correlations might reflect omitted country-specific trends

IMF forecasts capture economic fundamentals but not climate:

$$\hat{g}_{it}^{h} = g(X_{it}) + \gamma_i + \tau_t + \eta_{it}$$

Forecast error isolates climate component:

$$g_{it} - \hat{g}_{it}^{h} = f(C_{it}) + (\varepsilon_{it} - \eta_{it})$$

No longer contains omitted X_{it}!

Challenge 1: Are Climate-Growth Relationships Causal?

Identification concern: Temperature-growth correlations might reflect omitted country-specific trends

Placebo test: If forecasts are climate-naive, climate effects should vanish:

$$\hat{g}_{it}^{h} = f(C_{it}) + \gamma_i + \tau_t + \nu_{it} \quad \Rightarrow \quad f(C_{it}) = 0$$

Empirical Result

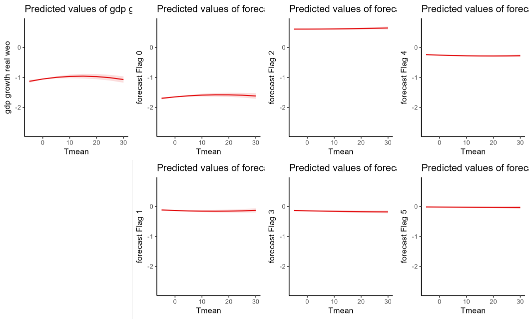
Temperature-growth relationship completely disappears when using IMF forecasts as dependent variable

Implications:

- Validates IMF forecasts as climate-naive
- Supports causal interpretation of climate-growth estimates

Placebo Test Results: Climate Effects Vanish in Forecasts

	Actual Growth	Forecast 0y	Forecast 1y	Forecast 2y	Forecast 3y
Temperature	0.0176***	0.0076**	-0.0035	0.00001	-0.0027
	(0.0063)	(0.0031)	(0.0027)	(0.0020)	(0.0024)
Temperature^2	-0.0006***	-0.0002**	0.0001	0.00004	0.00005
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Within R^2	0.005	0.001	0.001	0.003	0.002



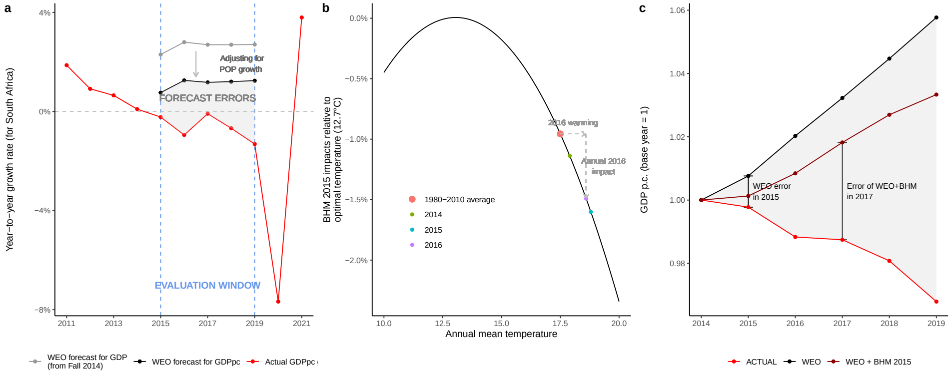
Damage Functions Implemented

Specification	Impact on	Driver	Reference
ABM 2024	Growth	Heat, drought, temp	Akyapı et al. (2025)
BHM 2015	Growth	Mean temp	Burke et al. (2015)
KMNPRY 2021	Growth	Mean temp	Kahn et al. (2021)
KW 2020	Growth	Mean temp	Kalkuhl & Wenz (2020)
KLW 2022	Growth	Mean temp	Kotz et al. (2022)
KLW 2024	Growth	Mean temp	Kotz et al. (2024)
BK 2024	Growth	GMST	Bilal & Känzig (2024)
COACCH	Level	GMST	van der Wijst et al. (2023)
DICE 2016	Level	GMST	Nordhaus (2019)

Implementation:

- Apply coefficients to consistent climate dataset
- Add impacts to IMF forecasts: $\hat{y}^s_{it} = \hat{y}_{it} + \Delta^s_{it}$
- Compare forecast errors with and without climate impacts

Methodology Overview



- **Step 1:** Calculate climate impacts using existing damage functions (2011-2014)
- **Step 2:** Add impacts to 2014 IMF forecasts for 2015-2019
- **Step 3:** Compare forecast accuracy: IMF vs. IMF + climate impacts

Challenge 2: Evaluation Metrics

Three complementary approaches:

1. **Sign matching:** Do climate impacts move forecasts in correct direction?

$$\text{SignMatch}^s = \frac{1}{NT}\sum_{i,t} 1\{\text{sgn}(\Delta_{it}^s) = \text{sgn}(e_{it})\}$$

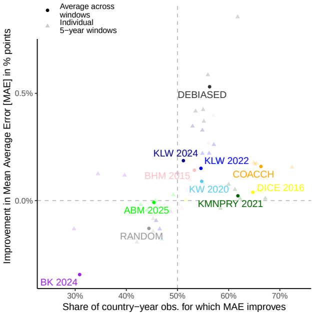
2. **Forecast accuracy:** How much do damage functions reduce errors?

$$\Delta\text{MAE}^s = \text{MAE}^{\text{baseline}} - \text{MAE}^s$$

3. **Error decomposition:** Global average vs. country heterogeneity

$$\Delta\text{MAE}^s_{\text{GLOBAL}} = \text{improvement in global growth}$$
$$\Delta\text{MAE}^s_{\text{COUNTRY}} = \text{improvement in regional patterns}$$

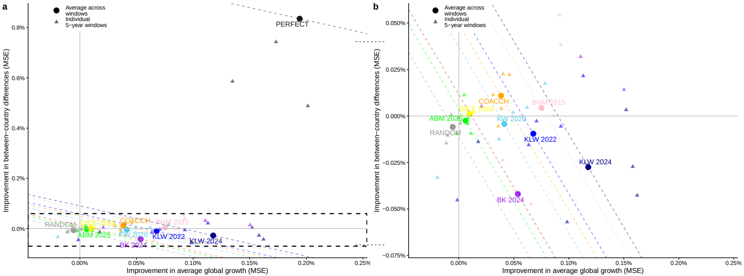
Challenge 2: How do the damage functions compare?



Key findings:

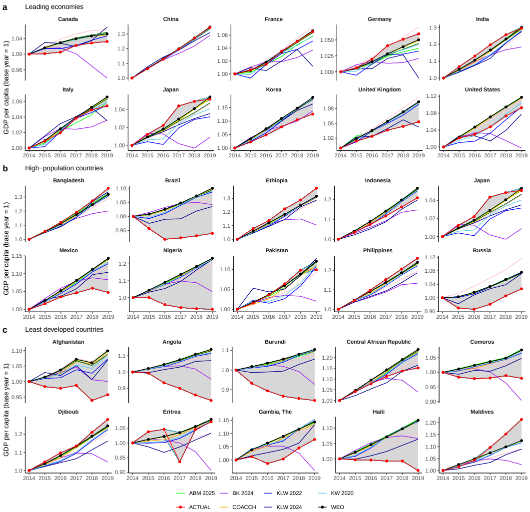
- Most functions reduce MAE by 0.1-0.2 pp (7-10% of baseline error)
- COACCH shows most consistent performance
- BK 2024 deteriorates forecast accuracy

Decomposing Performance: Global vs. Regional

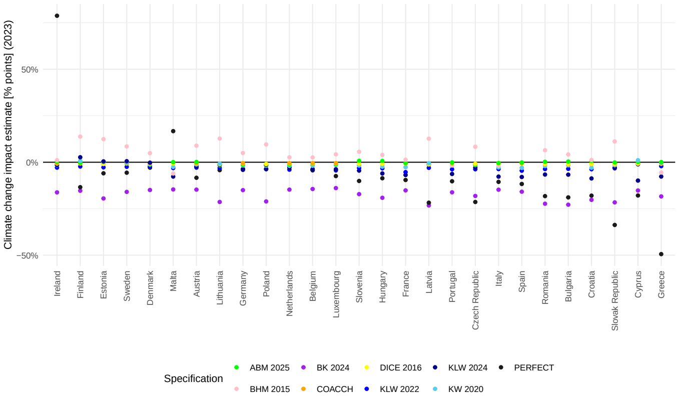


- Most functions improve both global growth and country heterogeneity
- COACCH, KLW 2024: best balance across both dimensions
- BK 2024: worsens global growth forecast

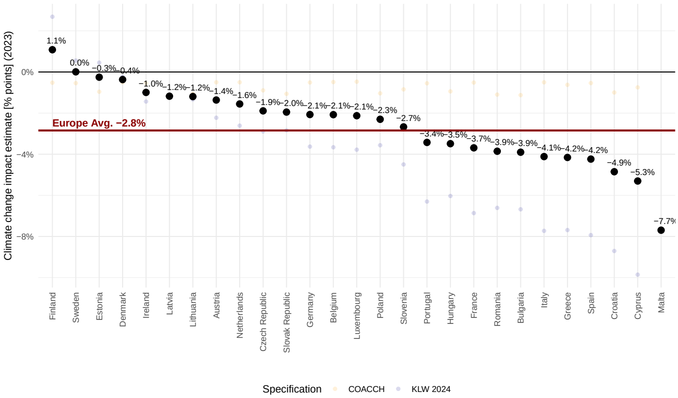
Challenge 3: Nowcasting and Reality Check for High-Impact Functions



Nowcasting Historical Impacts: Europe 2023



Nowcasting Historical Impacts: Europe 2023



Best-performing functions (COACCH + KLW 2024):

- European average: -2.8% GDP impact
- Clear gradient: +1.1% (Finland) to -7.7% (Malta)
- Southern Europe: -3 to -7%
- Pattern consistent, e.g. with PESETA V

Key Takeaways

Methodological contributions:

1. Professional forecasts provide valuable validation tool
2. Placebo tests support causal interpretation of climate-growth estimates
3. Approach complements traditional cross-validation methods

Substantive findings:

1. Climate impacts explain detectable (7-10%) but modest portion of forecast errors
2. Bottom-up functions (COACCH) show most consistent performance
3. Most severe functions overstate near-term impacts in rich countries

Policy implications:

- Climate impacts are real and detectable in economic data
- Extreme SCC estimates ($1,000+) lack near-term validation
- More moderate functions provide reliable foundation for policy analysis

Limitations and Future Work

Important caveats:

- Validation period: modest warming (0.2°C, 2011-2019)
- Cannot perfectly decompose forecast errors
- Excludes crisis years (financial crisis, COVID)
- Focus on GDP growth, not comprehensive impacts

Future research:

- Extend to longer periods with greater warming
- Apply to other forecast sources (OECD, World Bank)
- Validate sector-specific and regional models
- Investigate mechanisms behind performance differences

Thank you!

Questions and Comments?

Johannes Emmerling
johannes.emmerling@cmcc.it

Funding: EU Horizon Europe PRISMA (101081604) and SPARCCLE (101081369), Swiss SERI (22.00541)

Paper available soon.

Backup: IMF Forecasting Process

WEO methodology:

- "Top-down": Global econometric model
- "Bottom-up": Country desk economists with local expertise
- Inputs: Country authorities, private forecasts, structural models
- Methods vary by data availability and country complexity

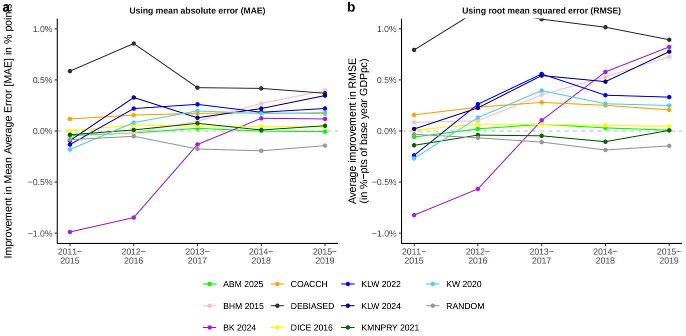
Climate treatment (per IMF 2016 guidance):

- Natural disasters: Historical probability $\times$ expected impact
- **No systematic use of climate projections** (especially pre-2018)
- Weather shocks: Not systematically incorporated
- 2022 staff note: No climate variable projections referenced before 2018

Known forecast properties:

- Similar accuracy to private sector
- Modest optimism bias (over-predict growth)

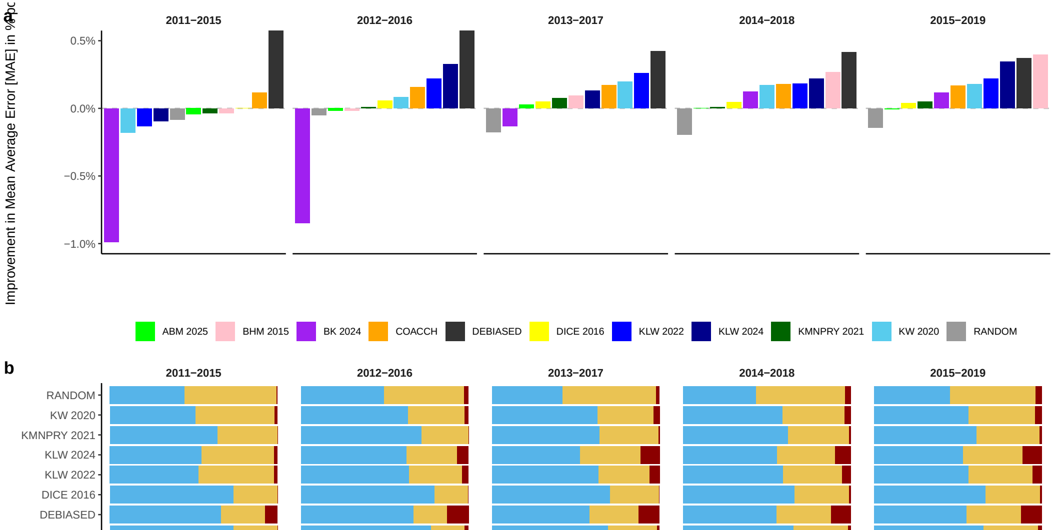
Performance Varies Across Evaluation Windows



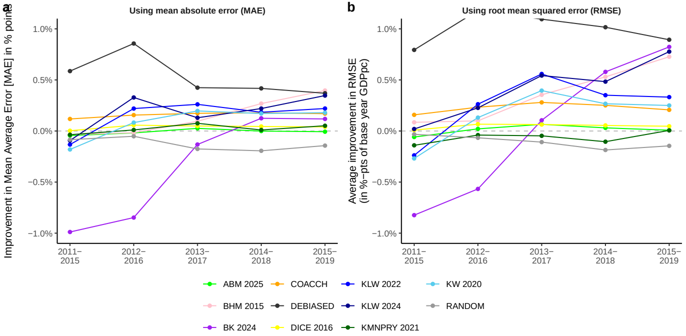
Patterns:

- No improvement for 2011-2015 period
- BHM performs best for 2014-2018, 2015-2019
- COACCH: only function with consistent improvement across all windows

Backup: Performance by Country and Window



Backup: Why MAE Rather Than RMSE?



Reasoning:

- RMSE heavily influenced by outlier countries
- Some extreme functions achieve good RMSE via large corrections for outliers
- MAE provides more balanced assessment across all countries