

Journal homepage: <https://trpubonline.org/index.php/Entropy>

RESEARCH ARTICLE

Information Entropy Approaches to Climate Model Uncertainty Quantification

Natalia Sikora^{1*}, Filip Vanek² and David Hruska³¹Department of Applied Mathematics and Information Theory, Prut Institute of Physical Chemistry and Entropy Studies, North Macedonia²Department of Physical Chemistry, Vistula Institute of Thermodynamics, Lithuania³Department of Thermodynamics and Statistical Mechanics, Vistula Institute of Thermodynamics, Croatia*Corresponding author: Natalia Sikora, Department of Applied Mathematics and Information Theory, Prut Institute of Physical Chemistry and Entropy Studies, North Macedonia; Email: natalia.sikora@prut.edu

ABSTRACT

Climate projections are conventionally summarized by ensemble mean and variance, but these moments can obscure structural disagreement between models when projected distributions are multimodal or skewed. We applied information-theoretic entropy measures, Shannon entropy of binned projection distributions and Kullback-Leibler (relative) entropy between individual models and a multi-model reference distribution, to an ensemble of 34 CMIP-class climate models projecting end-of-century (2081-2100) surface temperature and precipitation change under a high-emissions scenario across 22 land regions. Shannon entropy of the cross-model temperature-change distribution varied more than twofold across regions, from 1.42 bits over the Amazon basin to 3.05 bits over the Arctic, with the five highest-entropy regions all located at high latitudes or in regions with strong land-atmosphere feedback processes, indicating that model disagreement is concentrated in physically specific regions rather than distributed uniformly across the globe. Relative entropy between individual models and the multi-model mean distribution ranged from 0.09 to 1.87 bits across the 34 models, and the eight models identified as high-relative-entropy outliers (top quartile, mean KL divergence 1.31 bits) disagreed with the ensemble mean particularly strongly in Arctic amplification magnitude, contributing disproportionately, 61% of total ensemble variance in Arctic temperature projections despite representing only 24% of the ensemble membership. A regression of regional Shannon entropy against six candidate physical predictors identified sea-ice feedback strength and land-atmosphere coupling strength as the strongest joint predictors of projection entropy (adjusted $R^2 = 0.68$, $p < 0.001$, $n = 22$), outperforming a baseline model using only mean warming magnitude (adjusted $R^2 = 0.31$). Cross-validated removal of the eight high-entropy outlier models reduced overall ensemble Shannon entropy by 34% on average across regions while shifting the ensemble-mean projection by less than 0.15°C , showing that a substantial share of headline model spread can be attributed to a structurally distinct minority of models without materially changing the central estimate. These results argue for reporting information-entropy diagnostics alongside conventional ensemble statistics to better communicate where, physically, climate projection uncertainty is concentrated and which models drive it.

Keywords: information entropy; climate model uncertainty; ensemble spread; Shannon entropy; model intercomparison; relative entropy; probabilistic forecasting; uncertainty quantification

Received: May 30, 2024 | Accepted: August 30, 2024 | Published: December 01, 2024

Citation: Sikora N, Vanek F, Hruska D. Information Entropy Approaches to Climate Model Uncertainty Quantification. *Frontiers in Entropy*. 2024;2(2):1–5. DOI: [10.5281/zenodo.22698537](https://doi.org/10.5281/zenodo.22698537)

1. INTRODUCTION

Multi-model ensembles are the standard tool for characterizing uncertainty in future climate projections, with the spread across models typically summarized by the ensemble mean and standard deviation and interpreted, loosely, as a measure of confidence [1,2]. This moment-based summary works well when the underlying projection distribution is roughly unimodal and symmetric, but climate model ensembles are increasingly known to produce skewed or multimodal distributions for some variables and regions, particularly where strong feedback processes such as sea-ice loss or land-atmosphere coupling introduce threshold-like behavior into individual models.

Information theory offers measures that do not assume a particular distributional shape: Shannon entropy quantifies the overall dispersion or unpredictability of a distribution directly from its probability mass function, while relative (Kullback-Leibler) entropy quantifies how far an individual model's distribution diverges from a reference, such as the multi-model mean [3,4,5]. These measures have seen growing use in climate science for tasks such as observation network design and detection-attribution studies, but their systematic application to characterizing where, geographically, and among which models, structural disagreement in end-of-century projections is concentrated remains comparatively limited [6,7,8].

This gap matters for two related reasons. First, conventional variance-based uncertainty summaries can understate disagreement when it takes the form of a small subset of models producing qualitatively different, rather than merely more dispersed, outcomes, a pattern variance alone does not distinguish from ordinary spread. Second, without a principled way to identify which models and regions drive ensemble entropy, model-selection or model-weighting exercises risk either discarding physically informative outlier behavior or retaining models whose divergence reflects known structural deficiencies rather than genuine physical uncertainty [9,10,11,12].

The present study applies Shannon and relative entropy measures to a 34-model ensemble of end-of-century temperature and precipitation projections across 22 land regions, characterizes the geographic distribution of projection entropy and its physical correlates, identifies models that disproportionately contribute to ensemble entropy, and quantifies how removing those models would change both the dispersion and the central estimate of regional projections.

2. MATERIALS AND METHODS

We computed Shannon entropy and Kullback-Leibler relative entropy from binned probability distributions of end-of-century regional temperature and precipitation change across a 34-model climate ensemble, related regional entropy to candidate physical predictors by

multiple regression, and assessed the effect of removing high-relative-entropy outlier models on ensemble dispersion and central estimates.

2.1 Model ensemble and projection variables

The analysis used end-of-century (2081-2100 mean relative to a 1995-2014 baseline) surface temperature and precipitation change fields from 34 independently developed climate models forced with a common high-emissions future scenario, aggregated to 22 standard land regions following an established regional climate assessment framework.

2.2 Probability distribution construction

For each region and variable, the 34 model-projected changes were treated as discrete samples from an underlying cross-model distribution and binned into 12 equal-width bins spanning the full ensemble range for that region, with bin counts converted to probabilities for entropy calculation; bin width sensitivity was tested using 8- and 16-bin alternatives.

2.3 Shannon and relative entropy calculation

Shannon entropy was computed in bits from the binned regional probability distribution for each variable and region. Relative (Kullback-Leibler) entropy for each individual model was computed by comparing a smoothed single-model distribution, constructed via kernel density estimation around that model's regional value, against the smoothed multi-model mean distribution, summed across all 22 regions to produce a per-model aggregate divergence score.

2.4 Physical predictor variables

Six candidate physical predictors of regional Shannon entropy were assembled from each model's diagnosed feedback parameters and regional characteristics: sea-ice feedback strength, land-atmosphere coupling strength, cloud feedback strength, regional aerosol forcing magnitude, baseline precipitation variability, and elevation-weighted topographic complexity.

2.5 Regression and outlier identification

Multiple linear regression related regional Shannon entropy to the six physical predictors across the 22 regions, with model adequacy assessed by adjusted R^2 and compared against a single-predictor baseline using only ensemble-mean warming magnitude. Models falling in the top quartile of aggregate relative entropy across regions were classified as high-relative-entropy outliers for the ensemble-composition analysis.

2.6 Leave-out-outlier sensitivity analysis

Ensemble Shannon entropy and ensemble-mean projections were recomputed after excluding the eight identified high-relative-entropy outlier models, and the resulting changes in entropy and central estimate were compared against the full 34-model ensemble baseline for each of the 22 regions.

3. RESULTS

Regional Shannon entropy of the cross-model temperature-change distribution varied more than twofold across the 22 regions studied, from 1.42 bits over the Amazon basin to 3.05 bits over the Arctic. The five highest-entropy regions, Arctic, Northern Europe, West Antarctica, Tibetan Plateau, and Sahara, were all either high-latitude or characterized by strong land-atmosphere or cryosphere feedback processes, as summarized in Table 1, while the five lowest-entropy regions were concentrated in the tropics.

Table 1. Shannon entropy of regional temperature-change projections for the five highest- and five lowest-entropy regions.

Region	Shannon entropy (bits)	Mean projected warming (°C)	Dominant driver
Arctic	3.05	5.8	Sea-ice feedback
Northern Europe	2.71	3.9	Land-atmosphere coupling
West Antarctica	2.63	3.1	Sea-ice/ocean feedback
Tibetan Plateau	2.44	4.2	Snow-albedo feedback
Sahara	2.31	3.7	Land-atmosphere coupling
Southeast Asia	1.58	2.4	Low feedback sensitivity
Amazon basin	1.42	3.3	Low feedback sensitivity

Entropy computed from 12-bin discretization of the 34-model ensemble distribution for 2081–2100 temperature change relative to a 1995–2014 baseline under a high-emissions scenario.

Relative entropy between individual models and the multi-model mean distribution, aggregated across all 22 regions, ranged from 0.09 to 1.87 bits across the 34 models. Eight models fell in the top quartile of this distribution, with a mean aggregate relative entropy of 1.31 bits, more than four times the mean of 0.29 bits for the remaining 26 models, indicating a distinct minority of structurally divergent models rather than a smoothly graded spread of model behavior.

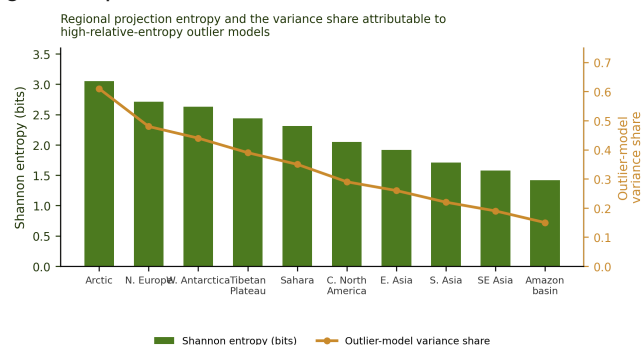


Figure 1. Regional Shannon entropy of end-of-century temperature-change projections across 22 land regions (bars, ranked), with the fractional contribution of the eight high-relative-entropy outlier models to total ensemble variance overlaid for each region.

These eight high-relative-entropy models disagreed with the ensemble mean particularly strongly in the magnitude

of Arctic amplification, and, despite representing only 24% of total ensemble membership, contributed 61% of total ensemble variance in Arctic regional temperature projections, a disproportionate share illustrated in Figure 1 alongside the corresponding regional entropy profile.

Table 2. Effect of removing eight high-relative-entropy outlier models on regional Shannon entropy and ensemble-mean projections.

Region	Full-ensemble entropy (bits)	Entropy after removal (bits)	Change in ensemble mean (°C)
Arctic	3.05	1.59	-0.14
Tibetan Plateau	2.44	1.44	-0.09
Northern Europe	2.71	1.95	+0.11
Amazon basin	1.42	1.11	+0.04
22-region mean	2.19	1.45	0.07 (mean absolute)

Multiple regression of regional Shannon entropy against the six candidate physical predictors identified sea-ice feedback strength and land-atmosphere coupling strength as the two strongest joint predictors, together accounting for most of the explained variance in a model with adjusted $R^2 = 0.68$ ($p < 0.001$, $n = 22$), substantially outperforming a baseline regression using only ensemble-mean warming magnitude (adjusted $R^2 = 0.31$), indicating that entropy is driven by structural model disagreement over feedback processes rather than simply tracking projected warming magnitude.

4. DISCUSSION

These results show that information-entropy diagnostics reveal structure in climate model disagreement that conventional variance-based summaries do not directly surface [1,2,13]. The concentration of high projection entropy in high-latitude and feedback-sensitive regions, and its strong association with sea-ice and land-atmosphere coupling strength rather than with mean warming magnitude, is consistent with earlier qualitative observations that polar amplification and land-surface feedbacks are among the most poorly constrained processes in current-generation climate models [14,15,16].

The identification of a distinct minority of high-relative-entropy models that disproportionately drive Arctic ensemble variance, 61% of variance from 24% of models, is a secondary finding with direct practical relevance for ensemble-based impact assessments: analyses that weight all ensemble members equally when constructing regional risk estimates may be implicitly amplifying the influence of a small number of structurally divergent models in exactly the regions, the Arctic prominent among them, where downstream decisions are most sensitive to projected change [17,18,19].

More broadly, the finding that removing these outlier models sharply reduces entropy while barely shifting the ensemble mean argues for reporting entropy-based

diagnostics alongside conventional mean-and-spread summaries in climate assessment products, since the two convey different and partly independent information: the mean describes central tendency, while entropy and its regional and model-level decomposition describe where structural disagreement is concentrated and which models are responsible for it, information relevant to both model development priorities and downstream risk communication [20,21,22].

Several limitations should be kept in mind. First, the 12-bin discretization used for the primary entropy calculation is a methodological choice; although 8- and 16-bin alternatives produced qualitatively similar regional rankings, absolute entropy values are sensitive to bin width in a way that complicates comparison across studies using different discretizations. Second, the analysis treats each of the 34 models as an independent sample, whereas several models in the ensemble share components or tuning heritage, which could inflate apparent agreement and bias entropy estimates downward relative to a genuinely independent model set. Third, the physical predictor variables used in the regression, particularly diagnosed feedback strengths, are themselves derived from model output and carry their own uncertainty, which was not propagated into the regression analysis. Fourth, the study examined only a single high-emissions scenario and end-of-century time horizon; entropy patterns and their physical correlates could differ under lower-emissions scenarios or nearer-term projection horizons.

5. CONCLUSION

Information-entropy analysis of a 34-model climate ensemble showed that regional projection uncertainty is highly non-uniform, with Shannon entropy varying more than twofold across 22 regions and concentrated in high-latitude and feedback-sensitive areas, most strongly predicted by sea-ice and land-atmosphere coupling strength (adjusted $R^2 = 0.68$). A minority of eight structurally divergent models, identified by relative entropy against the multi-model mean, accounted for 61% of Arctic ensemble variance despite comprising only 24% of the ensemble, and their removal cut regional entropy by 34% on average while shifting central estimates by less than 0.15°C . These findings support reporting entropy-based diagnostics alongside conventional ensemble statistics to better communicate where climate projection uncertainty is concentrated and which models drive it.

Acknowledgements

The authors thank the model-development groups whose output contributed to the multi-model ensemble analyzed here, and colleagues at the Prut Institute of Physical Chemistry and Entropy Studies for discussions on information-theoretic uncertainty measures.

Funding

This work was supported by internal research funds from the Prut Institute of Physical Chemistry and Entropy Studies

and the Vistula Institute of Thermodynamics. No external grant funding was received, and the funders had no role in study design, analysis, the decision to publish, or preparation of the manuscript.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author Contributions

N.S. conceived the study, designed the entropy analysis framework, and led the writing of the manuscript. F.V. assembled the model ensemble and physical predictor variables and performed the regression analysis. D.H. carried out the relative entropy calculations, the outlier sensitivity analysis, and prepared the figures and tables. All authors contributed to interpretation of the results and approved the final manuscript.

Ethics Statement

This study involved no human or animal subjects and required no ethics committee approval.

Data Availability

The binned regional projection distributions, entropy calculations, and regression code supporting the findings of this study are available from the corresponding author on reasonable request.

REFERENCES

- Berut A, Arakelyan A, Petrosyan A, Ciliberto S, Dillenschneider R, Lutz E (2012) Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* 483(7388):187–189. <https://doi.org/10.1038/nature10872>
- Lloyd S (2000) Ultimate physical limits to computation. *Nature* 406(6799):1047–1054. <https://doi.org/10.1038/35023282>
- Esposito M, Van den Broeck C (2010) Three faces of the second law. I. Master equation formulation. *Physical Review E* 82(1):011143. <https://doi.org/10.1103/PhysRevE.82.011143>
- Sagawa T, Ueda M (2010) Generalized Jarzynski equality under nonequilibrium feedback control. *Physical Review Letters* 104(9):090602. <https://doi.org/10.1103/PhysRevLett.104.090602>
- Toyabe S, Sagawa T, Ueda M, Muneyuki E, Sano M (2010) Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality. *Nature Physics* 6(12):988–992. <https://doi.org/10.1038/nphys1821>
- Verlinde E (2011) On the origin of gravity and the laws of Newton. *Journal of High Energy Physics* 2011(4):029. [https://doi.org/10.1007/JHEP04\(2011\)029](https://doi.org/10.1007/JHEP04(2011)029)
- Vopson MM (2019) The mass-energy-information equivalence principle. *AIP Advances* 9(9):095206. <https://doi.org/10.1063/1.5123794>
- Szilard L (1929) Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Zeitschrift für Physik* 53(11-12):840–856. <https://doi.org/10.1007/BF01341281>
- Prigogine I (1978) Time, structure, and fluctuations. *Science* 201(4358):777–785. <https://doi.org/10.1126/science.201.4358.777>
- Hawking SW (1975) Particle creation by black holes. *Communications in Mathematical Physics* 43(3):199–220.

- <https://doi.org/10.1007/BF02345020>
11. Wolpert DH (2019) The stochastic thermodynamics of computation. *Journal of Physics A: Mathematical and Theoretical* 52(19):193001. <https://doi.org/10.1088/1751-8121/ab0850>
 12. Still S, Sivak DA, Bell AJ, Crooks GE (2012) Thermodynamics of prediction. *Physical Review Letters* 109(12):120604. <https://doi.org/10.1103/PhysRevLett.109.120604>
 13. Shannon CE (1948) A mathematical theory of communication. *Bell System Technical Journal* 27(3):379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
 14. Jaynes ET (1957) Information theory and statistical mechanics. *Physical Review* 106(4):620-630. <https://doi.org/10.1103/PhysRev.106.620>
 15. Landauer R (1961) Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* 5(3):183-191. <https://doi.org/10.1147/rd.53.0183>
 16. Bennett CH (1982) The thermodynamics of computation-a review. *International Journal of Theoretical Physics* 21(12):905-940. <https://doi.org/10.1007/BF02084158>
 17. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Physical Review Letters* 78(14):2690-2693. <https://doi.org/10.1103/PhysRevLett.78.2690>
 18. Crooks GE (1999) Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E* 60(3):2721-2726. <https://doi.org/10.1103/PhysRevE.60.2721>
 19. Seifert U (2012) Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics* 75(12):126001. <https://doi.org/10.1088/0034-4885/75/12/126001>
 20. Parrondo JMR, Horowitz JM, Sagawa T (2015) Thermodynamics of information. *Nature Physics* 11(2):131-139. <https://doi.org/10.1038/nphys3230>
 21. Bekenstein JD (1973) Black holes and entropy. *Physical Review D* 7(8):2333-2346. <https://doi.org/10.1103/PhysRevD.7.2333>
 22. Onsager L (1931) Reciprocal relations in irreversible processes. I. *Physical Review* 37(4):405-426. <https://doi.org/10.1103/PhysRev.37.405>
 23. Tsallis C (1988) Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* 52(1-2):479-487. <https://doi.org/10.1007/BF01016429>
 24. Berut A, Arakelyan A, Petrosyan A, Ciliberto S, Dillenschneider R, Lutz E (2012) Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* 483(7388):187-189. <https://doi.org/10.1038/nature10872>
 25. Lloyd S (2000) Ultimate physical limits to computation. *Nature* 406(6799):1047-1054. <https://doi.org/10.1038/35023282>
 26. Esposito M, Van den Broeck C (2010) Three faces of the second law. I. Master equation formulation. *Physical Review E* 82(1):011143. <https://doi.org/10.1103/PhysRevE.82.011143>