David Hagen

Prof. Demetris Koutsoyiannis has quantified how the Temperature Hen PRECEDES the CO2 Egg in numerous scientific publications posted at most of his 79 publications with the ITIA group at the National Technical University of Athens as posted at https://www.itia.ntua.gr/en/search/?authors=koutsoyiannis&tags=climate.
Koutsoyiannis' quantitative scientific evidence nullifies the very premise of California's climate laws and regulations. I recommend focusing on lowest life cycle costs working towards long term sustainable energy and fuel use for the common good. Using geologically stored solar energy provided by our Creator enables rapid inexpensive rise out of poverty into sustainable society (aka natural gas, oil and coal.) Attached is a list of Koutsoyiannis' publications on Climate and four representative papers quantifying that global Temperature change precedes CO2 change.

Demetris Koutsoyianns' Publications on Climate

Posted at "https://www.itia.ntua.gr/en/search/?authors=koutsoyiannis&tags=climate

Publications in scientific journals

1. D. Koutsoyiannis, and G. Tsakalias, Unsettling the settled: Simple musings on the complex climatic system, *Frontiers in Complex Systems*, 3, 1617092, doi:10.3389/fcpxs.2025.1617092, 2025.

[doc_id=2537] English More information and full text

2. D. Koutsoyiannis, When are models useful? Revisiting the quantification of reality checks, *Water*, 17 (2), 264, doi:10.3390/w17020264, 2025.

[doc_id=2522] English More information and full text

3. D. Koutsoyiannis, The relationship between atmospheric temperature and carbon dioxide concentration, *Science of Climate Change*, 4 (3), 39–59, doi:10.53234/scc202412/15, 2024.

[doc_id=2517] English More information and full text

4. D. Koutsoyiannis, Relative importance of carbon dioxide and water in the greenhouse effect: Does the tail wag the dog?, *Science of Climate Change*, 4 (2), 36–78, doi:10.53234/scc202411/01, 2024.

[doc_id=2501] English More information and full text

5. D. Koutsoyiannis, Definite change since the formation of the Earth [Reply to Kleber, A. Comment on "Koutsoyiannis, D. Net isotopic signature of atmospheric CO2 sources and sinks: No change since the Little Ice Age. Sci 2024, 6, 17"], *Sci*, 6 (4), 63, doi:10.3390/sci6040063, 2024.

[doc_id=2498] English More information and full text

6. D. Koutsoyiannis, Refined reservoir routing (RRR) and its application to atmospheric carbon dioxide balance, *Water*, 16 (17), 2402, doi:10.3390/w16172402, 2024.

[doc_id=2474] English More information and full text

7. D. Koutsoyiannis, Stochastic assessment of temperature – CO_2 causal relationship in climate from the Phanerozoic through modern times, *Mathematical Biosciences and Engineering*, 21 (7), 6560–6602, doi:10.3934/mbe.2024287, 2024.

[doc_id=2467] English More information and full text

8. D. Koutsoyiannis, Net isotopic signature of atmospheric CO_2 sources and sinks: No change since the Little Ice Age, *Sci*, 6 (1), 17, doi:10.3390/sci6010017, 2024.

[doc_id=2444] English More information and full text

9. D. Koutsoyiannis, and C. Vournas, Revisiting the greenhouse effect—a hydrological perspective, *Hydrological Sciences Journal*, 69 (2), 151–164, doi:10.1080/02626667.2023.2287047, 2024.

[doc_id=2371] English More information and full text

10. D. Koutsoyiannis, C. Onof, Z. W. Kundzewicz, and A. Christofides, On hens, eggs, temperatures and CO_2: Causal links in Earth's atmosphere, *Sci*, 5 (3), 35, doi:10.3390/sci5030035, 2023.

[doc_id=2342] English More information and full text

11. D. Koutsoyiannis, T. Iliopoulou, A. Koukouvinos, N. Malamos, N. Mamassis, P. Dimitriadis, N. Tepetidis, and D. Markantonis, In search of climate crisis in Greece using hydrological data: 404 Not Found, *Water*, 15 (9), 1711, doi:10.3390/w15091711, 2023.

[doc_id=2287] English More information and full text

12. D. Koutsoyiannis, C. Onof, A. Christofides, and Z. W. Kundzewicz, Revisiting causality using stochastics: 2. Applications, *Proceedings of The Royal Society A*, 478 (2261), 20210836, doi:10.1098/rspa.2021.0836, 2022.

[doc_id=2194] English More information and full text

13. D. Koutsoyiannis, C. Onof, A. Christofides, and Z. W. Kundzewicz, Revisiting causality using stochastics: 1.Theory, *Proceedings of The Royal Society A*, 478 (2261), 20210835, doi:10.1098/rspa.2021.0835, 2022.

[doc_id=2193] English More information and full text

14. D. Koutsoyiannis, and A. Montanari, Climate extrapolations in hydrology: The expanded Bluecat methodology, *Hydrology*, 9, 86, doi:10.3390/hydrology9050086, 2022.

[doc_id=2192] English More information and full text

15. D. Koutsoyiannis, Rethinking climate, climate change, and their relationship with water, *Water*, 13 (6), 849, doi:10.3390/w13060849, 2021.

[doc_id=2098] English More information and full text

16. D. Koutsoyiannis, and Z. W. Kundzewicz, Atmospheric temperature and CO_2: Hen-or-egg causality?, *Sci*, 2 (4), 83, doi:10.3390/sci2040083, 2020.

[doc_id=2064] English More information and full text

17. Z. W. Kundzewicz, I. Pińskwar, and D. Koutsoyiannis, Variability of global mean annual temperature is significantly influenced by the rhythm of ocean-atmosphere oscillations, *Science of the Total Environment*, 747, 141256, doi:10.1016/j.scitotenv.2020.141256, 2020.

[doc_id=2055] English More information and full text

18. T. Iliopoulou, and D. Koutsoyiannis, Projecting the future of rainfall extremes: better classic than trendy, *Journal of Hydrology*, 588, doi:10.1016/j.jhydrol.2020.125005, 2020.

[doc_id=2045] English More information and full text

19. D. Koutsoyiannis, Revisiting the global hydrological cycle: is it intensifying?, *Hydrology and Earth System Sciences*, 24, 3899–3932, doi:10.5194/hess-24-3899-2020, 2020.

[doc_id=2042] English More information and full text

20. D. Koutsoyiannis, Time's arrow in stochastic characterization and simulation of atmospheric and hydrological processes, *Hydrological Sciences Journal*, 64 (9), 1013–1037, doi:10.1080/02626667.2019.1600700, 2019.

[doc_id=1944] English More information and full text

21. H. Tyralis, and D. Koutsoyiannis, On the prediction of persistent processes using the output of deterministic models, *Hydrological Sciences Journal*, 62 (13), 2083–2102, doi:10.1080/02626667.2017.1361535, 2017.

[doc_id=1727] English More information and full text

22. Y. Markonis, A. N. Angelakis, J. Christy, and D. Koutsoyiannis, Climatic variability and the evolution of water technologies in Crete, Hellas, *Water History*, 8 (2), 137–157, doi:10.1007/s12685-016-0159-9, 2016.

[doc_id=1623] English More information and full text

23. Y. Markonis, and D. Koutsoyiannis, Scale-dependence of persistence in precipitation records, *Nature Climate Change*, 6, 399–401, doi:10.1038/nclimate2894, 2016.

[doc_id=1594] English More information

24. D. Koutsoyiannis, and A. Montanari, Negligent killing of scientific concepts: the stationarity case, *Hydrological Sciences Journal*, 60 (7-8), 1174–1183, doi:10.1080/02626667.2014.959959, 2015.

[doc_id=1485] English More information and full text

25. A. Montanari, and D. Koutsoyiannis, Modeling and mitigating natural hazards: Stationarity is immortal!, *Water Resources Research*, 50 (12), 9748–9756, doi:10.1002/2014WR016092, 2014.

[doc_id=1511] English More information and full text

26. D. Koutsoyiannis, Social vs. scientific perception of change in hydrology and climate — Reply to the Discussion by Arie Ben-Zvi on the Opinion Paper "Hydrology and Change", *Hydrological Sciences Journal*, 59 (8), 1625–1626, doi:10.1080/02626667.2014.935382, 2014.

[doc_id=1465] English More information and full text

27. H. Tyralis, and D. Koutsoyiannis, A Bayesian statistical model for deriving the predictive distribution of hydroclimatic variables, *Climate Dynamics*, 42 (11-12), 2867–2883, doi:10.1007/s00382-013-1804-y, 2014.

[doc_id=1361] English More information and full text

28. D. Koutsoyiannis, Hydrology and Change, *Hydrological Sciences Journal*, 58 (6), 1177–1197, doi:10.1080/02626667.2013.804626, 2013.

[doc_id=1351] English More information and full text

29. Y. Markonis, and D. Koutsoyiannis, Climatic variability over time scales spanning nine orders of magnitude: Connecting Milankovitch cycles with Hurst–Kolmogorov dynamics, *Surveys in Geophysics*, 34 (2), 181–207, doi:10.1007/s10712-012-9208-9, 2013.

[doc_id=1297] English More information and full text

30. D. Koutsoyiannis, A. Christofides, A. Efstratiadis, G. G. Anagnostopoulos, and N. Mamassis, Scientific dialogue on climate: is it giving black eyes or opening closed eyes? Reply to “A black eye for the Hydrological Sciences Journal” by D. Huard, *Hydrological Sciences Journal*, 56 (7), 1334–1339, doi:10.1080/02626667.2011.610759, 2011.

[doc_id=1140] English More information and full text

31. D. Koutsoyiannis, Hurst-Kolmogorov dynamics as a result of extremal entropy production, *Physica A: Statistical Mechanics and its Applications*, 390 (8), 1424–1432, doi:10.1016/j.physa.2010.12.035, 2011.

[doc_id=1102] English More information and full text

32. D. Koutsoyiannis, Hurst-Kolmogorov dynamics and uncertainty, *Journal of the American Water Resources Association*, 47 (3), 481–495, doi:10.1111/j.1752-1688.2011.00543.x, 2011.

[doc_id=1001] English More information and full text

33. G. Di Baldassarre, A. Montanari, H. F. Lins, D. Koutsoyiannis, L. Brandimarte, and G. Blöschl, Flood fatalities in Africa: from diagnosis to mitigation, *Geophysical Research Letters*, 37, L22402, doi:10.1029/2010GL045467, 2010.

[doc_id=1012] English More information and full text

34. G. G. Anagnostopoulos, D. Koutsoyiannis, A. Christofides, A. Efstratiadis, and N. Mamassis, A comparison of local and aggregated climate model outputs with observed data, *Hydrological Sciences Journal*, 55 (7), 1094–1110, doi:10.1080/02626667.2010.513518, 2010.

[doc_id=978] English More information and full text

35. D. Koutsoyiannis, A random walk on water, *Hydrology and Earth System Sciences*, 14, 585–601, doi:10.5194/hess-14-585-2010, 2010.

[doc_id=923] English [More information and full text]

36. D. Koutsoyiannis, A. Montanari, H. F. Lins, and T.A. Cohn, Climate, hydrology and freshwater: towards an interactive incorporation of hydrological experience into climate research—DISCUSSION of "The implications of projected climate change for freshwater resources and their management", *Hydrological Sciences Journal*, 54 (2), 394–405, doi:10.1623/hysj.54.2.394, 2009.

[doc_id=907] English [More information and full text]

37. D. Koutsoyiannis, C. Makropoulos, A. Langousis, S. Baki, A. Efstratiadis, A. Christofides, G. Karavokiros, and N. Mamassis, Climate, hydrology, energy, water: recognizing uncertainty and seeking sustainability, *Hydrology and Earth System Sciences*, 13, 247–257, doi:10.5194/hess-13-247-2009, 2009.

[doc_id=878] English [More information and full text]

38. D. Koutsoyiannis, A. Efstratiadis, N. Mamassis, and A. Christofides, On the credibility of climate predictions, *Hydrological Sciences Journal*, 53 (4), 671–684, doi:10.1623/hysj.53.4.671, 2008.

[doc_id=864] English [More information and full text]

39. C. Cudennec, C. Leduc, and D. Koutsoyiannis, Dryland hydrology in Mediterranean regions -- a review, *Hydrological Sciences Journal*, 52 (6), 1077–1087, doi:10.1623/hysj.52.6.1077, 2007.

[doc_id=837] English [More information and full text]

40. D. Koutsoyiannis, and A. Montanari, Statistical analysis of hydroclimatic time series: Uncertainty and insights, *Water Resources Research*, 43 (5), W05429, doi:10.1029/2006WR005592, 2007.

[doc_id=781] English [More information and full text]

41. D. Koutsoyiannis, A. Efstratiadis, and K. Georgakakos, Uncertainty assessment of future hydroclimatic predictions: A comparison of probabilistic and scenario-based approaches, *Journal of Hydrometeorology*, 8 (3), 261–281, doi:10.1175/JHM576.1, 2007.

[doc_id=728] English [More information and full text]

42. D. Koutsoyiannis, Nonstationarity versus scaling in hydrology, *Journal of Hydrology*, 324, 239–254, doi:10.1016/j.jhydrol.2005.09.022, 2006.

[doc_id=673] English [More information and full text]

43. D. Koutsoyiannis, A toy model of climatic variability with scaling behaviour, *Journal of Hydrology*, 322, 25–48, doi:10.1016/j.jhydrol.2005.02.030, 2006.

[doc_id=648] English More information and full text

44. D. Koutsoyiannis, Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, 48 (1), 3–24, doi:10.1623/hysj.48.1.3.43481, 2003.

[doc_id=537] English More information and full text

Book chapters and fully evaluated conference publications

45. D. Koutsoyiannis, Climate change impacts on hydrological science: How the climate change agenda has lowered the scientific level of hydrology (Plenary talk), *13th International Conference on Hydroinformatics (HIC 2018)*, Palermo, Italy, doi:10.13140/RG.2.2.12249.42084, 2018.

[doc_id=1847] English More information and full text

Conference publications and presentations with evaluation of abstract

46. D. Koutsoyiannis, The relationship between atmospheric temperature and carbon dioxide concentration (Presentation), *Two-day Clintel Conference in the Parliament of Prague: Climate Change, Facts and Myths in the Light of Science*, Prague, 2024.

[doc_id=2502] English More information and full text

47. D. Koutsoyiannis, Hydrology, society, change and uncertainty (invited talk), *European Geosciences Union General Assembly 2014, Geophysical Research Abstracts, Vol. 16*, Vienna, EGU2014-4243, doi:10.13140/RG.2.2.15432.93441, European Geosciences Union, 2014.

[doc_id=1441] English More information and full text

48. E. Steirou, and D. Koutsoyiannis, Investigation of methods for hydroclimatic data homogenization, *European Geosciences Union General Assembly 2012, Geophysical Research Abstracts, Vol. 14*, Vienna, 956-1, doi:10.13140/RG.2.2.23854.31046, European Geosciences Union, 2012.

[doc_id=1212] English More information and full text

49. D. Koutsoyiannis, Memory in climate and things not to be forgotten (Invited talk), *11th International Meeting on Statistical Climatology*, Edinburgh, doi:10.13140/RG.2.2.17890.53445, International Meetings on Statistical Climatology, University of Edinburgh, 2010.

[doc_id=991] English More information and full text

50. D. Koutsoyiannis, A random walk on water (Henry Darcy Medal Lecture), *European Geosciences Union General Assembly 2009, Geophysical Research Abstracts, Vol. 11*, Vienna, 14033, doi:10.13140/RG.2.1.2139.4800, European Geosciences Union, 2009.

[doc_id=896] English More information and full text

51. D. Koutsoyiannis, From climate certainties to climate stochastics (Opening Lecture), *IHP 2008 Capri Symposium: "The Role of Hydrology in Water Resources Management"*, Capri, Italy, doi:10.13140/RG.2.2.28481.15205/1, UNESCO, International Association of Hydrological Sciences, 2008.

[doc_id=880] English More information and full text

52. D. Koutsoyiannis, and T.A. Cohn, The Hurst phenomenon and climate (solicited), *European Geosciences Union General Assembly 2008, Geophysical Research Abstracts, Vol. 10*, Vienna, 11804, doi:10.13140/RG.2.2.13303.01447, European Geosciences Union, 2008.

[doc_id=849] English More information and full text

Presentations and publications in workshops

53. D. Koutsoyiannis, H_2O, CO_2, Climate Change: A holistic refutation of "climate science", *International conference: Beyond the Climate Change Consensus*, Budapest, Hungary, Hungarian Academy of Sciences, 2025.

[doc_id=2573] English More information and full text

54. D. Koutsoyiannis, Climate crisis: natural or political event?, *Natural & Biological Sciences Seminar, Free University*, Parnassos Literary Society, Athens, 2025.

[doc_id=2541] Greek More information and full text

55. D. Koutsoyiannis, The contradictions of the settled "climate science", *International Hellenic Association (IHA) Lectures*, doi:10.13140/RG.2.2.14899.13605, 2025.

[doc_id=2528] Greek More information and full text

56. D. Koutsoyiannis, Fundamental ideas in climate research: How they evolved and how correct they are, *Clintel Workshop on Recent Research Developments on Atmospheric Temperature, Carbon Dioxide, and Their Relationship*, National Technical University of Athens, Greece, 2024.

[doc_id=2475] English More information and full text

57. D. Koutsoyiannis, Climate change impacts on hydrological science: How the climate change agenda has lowered the scientific level of hydrology, *School for Young Scientists "Modelling and forecasting of river flows and managing hydrological risks: Towards a new generation of methods" (2018)*, doi:10.13140/RG.2.2.11110.06727, Russian Academy of Sciences, Lomonosov Moscow State University, 2018.

[doc_id=1901]	English	More information and full text

58. D. Koutsoyiannis, Saving the world from climate threats vs. dispelling climate myths and fears, *Invited Seminar*, Lunz am See, Austria, doi:10.13140/RG.2.2.34278.42565, WasserCluster Lunz – Biologische Station GmbH, 2017.

[doc_id=1706]	English	More information and full text

59. D. Koutsoyiannis, Climate is changing ... since 4.5 billion years ago, *Climate change: natural or human-induced*, Athens, doi:10.13140/RG.2.2.24054.19524, Massachusetts Institute of Technology Alumni, University of Michigan Alumni, Athens, 2011.

[doc_id=1181]	Greek	More information and full text

60. A. Christofides, and D. Koutsoyiannis, God and the arrogant species: Contrasting nature's intrinsic uncertainty with our climate simulating supercomputers, *104th Annual Conference & Exhibition*, Orlando, Florida, Air & Waste Management Association, 2011.

[doc_id=1153]	English	More information and full text

61. D. Koutsoyiannis, Climate change as a scapegoat in water science, technology and management, *EUREAU Workshop on Climate Changes Impact on Water Resources with Emphasis on Potable Water*, Chania, doi:10.13140/RG.2.2.35519.71843, European Association of Water and Wastewater Services, Hellenic Union of Water and Wastewater Enterprises, 2008.

[doc_id=865]	English	More information and full text

62. D. Koutsoyiannis, Climatic uncertainty, the Joseph effect and the water resource management, *Man and environment in the 21st century - The crucial problems - Atmosphere and climate*, Athens, doi:10.13140/RG.2.2.22533.76008, Goulandris Natural History Museum, 2005.

[doc_id=662]	Greek	More information and full text

Various publications

63. D. Koutsoyiannis, and T. Iliopoulou, A cool look at hydroclimatic risk, *Natl Institute of Natural Hazards & Chinese Academy of Sciences*, doi:10.13140/RG.2.2.24705.36964, Beijing, China, July 2025.

[doc_id=2546]	English	More information and full text

64. D. Koutsoyiannis, The superiority of refined reservoir routing (RRR) in modelling atmospheric carbon dioxide, doi:10.13140/RG.2.2.32285.32489, ResearchGate, 2024.

[doc_id=2505]	English	More information and full text

65. D. Koutsoyiannis, From hen's egg to serpent's egg: Peer reviews and other attacks on science for silencing voices opposing the "climate crisis" narrative, NTUA, Athens, 2024.

[doc_id=2468]   English   More information and full text

66. A. Christofides, D. Koutsoyiannis, C. Onof, and Z. W. Kundzewicz, Causality, Climate, Etc., doi:10.13140/RG.2.2.21608.44803, Climate Etc. (Judith Curry's blog), 2023.

[doc_id=2353]   English   More information and full text

67. F. Battaglia, and D. Koutsoyiannis, Interview with Demetris Koutsoyiannis, doi:10.13140/RG.2.2.20246.93767, Newspaper La Verità, 27 October 2023.

[doc_id=2349]   English   More information and full text

68. D. Koutsoyiannis, A voyage in climate, hydrology and life on a 4.5-billion-years old planet, *Self-organized lecture*, doi:10.13140/RG.2.2.27000.26883, School of Civil Engineering – National Technical University of Athens, Athens, 20 July 2020.

[doc_id=2036]   English   More information and full text

69. D. Koutsoyiannis, The political origin of the climate change agenda, *Self-organized lecture*, doi:10.13140/RG.2.2.10223.05283, School of Civil Engineering – National Technical University of Athens, Athens, 14 April 2020.

[doc_id=2035]   English   More information and full text

70. D. Koutsoyiannis, Personal knowable moments (DK-moments) for high-order characterization of coincidence in totalitarianism, *Self-organized lecture*, doi:10.13140/RG.2.2.23117.38885/1, Bologna, Italy, 17 December 2019.

[doc_id=2015]   English   More information and full text

71. D. Tsaknias, D. Bouziotas, and D. Koutsoyiannis, Statistical comparison of observed temperature and rainfall extremes with climate model outputs in the Mediterranean region, *ResearchGate*, doi:10.13140/RG.2.2.11993.93281, 2016.

[doc_id=1650]   English   More information and full text

72. D. Koutsoyiannis, LTP: Looking Trendy—Persistently, *Climate Dialogue*, doi:10.13140/RG.2.2.13070.36169, 2013.

[doc_id=1348]   English   More information and full text

73. D. Koutsoyiannis, Beware saviors!, *Climate Science (by Roger Pielke Sr.)*, 2 pages, doi:10.13140/RG.2.2.23765.83688, 2009.

[doc_id=936]   English   More information and full text

74. Th. Xanthopoulos, and D. Koutsoyiannis, Climate worsening: Inherent weaknesses in reliable prediction, and unjustified doomsaying, *Bulletin of the National Chamber of Greece*, 144–146, 8 July 2002.

[doc_id=544] Greek More information and full text

75. Th. Xanthopoulos, and D. Koutsoyiannis, Prediction of climate: Scientific evidence, historical experience and the truth, *Newspaper "To Vima"*, A10–A11, 17 September 2000.

[doc_id=174] Greek More information and full text

76. D. Koutsoyiannis, Climate change: Myths and reality, *New Ecology*, 151, 27–28, May 1997.

[doc_id=197] Greek More information and full text

Books

77. T. Iliopoulou, and D. Koutsoyiannis, *Have Rainfall Patterns Changed? A Global Analysis of Long-Term Rainfall Records and Re-Analysis Data*, 47 pages, SR 306, The Heritage Foundation, Washington, DC, USA, 2025.

[doc_id=2521] English More information and full text

78. D. Koutsoyiannis, *Stochastics of Hydroclimatic Extremes - A Cool Look at Risk*, Edition 5, *ISBN: 978-618-85370-0-2*, 420 pages, doi:10.57713/kallipos-1, Kallipos Open Academic Editions, Athens, 2025.

[doc_id=2000] English More information and full text

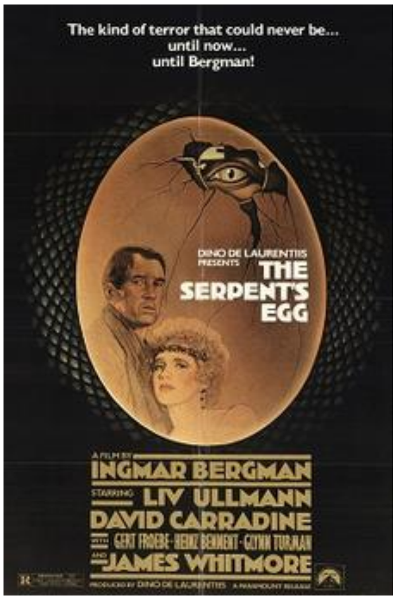
79. D. Koutsoyiannis, and T. Iliopoulou, *Understanding Climate: Gifts from the Nile*, 60 pages, SR 301, The Heritage Foundation, Washington, DC, USA, 2024.

[doc_id=2510] English More information and full text

From hen's egg to serpent's egg: Peer reviews and other attacks on science for silencing voices opposing the "climate crisis" narrative

by Demetris Koutsoyiannis

July 2024



Sources of images: (**left**) Flyte so Fancy[1]; (**right**) Wikipedia[2]

Crime begins with propaganda, even if such propaganda is for a good cause (Hans Fritzsche[3])

In the four-year period 2020-2023, I published a dozen of papers on climate. Four of them, prepared jointly with other colleagues, investigated causality in climate. Two of these four were published in the MDPI journal *Sci* and two in the *Proceedings of The Royal Society A – Mathematical, Physical and Engineering Studies*. Their details are:

1. D. Koutsoyiannis, and Z. W. Kundzewicz, 2020. Atmospheric temperature and CO_2: Hen-or-egg causality?. *Sci*, 2 (4), 83, doi: 10.3390/sci2040083.
2. D. Koutsoyiannis, C. Onof, A. Christofides, and Z. W. Kundzewicz, 2022. Revisiting causality using stochastics: 1.Theory. *Proceedings of The Royal Society A*, 478 (2261), 20210835, doi: 10.1098/rspa.2021.0835.
3. D. Koutsoyiannis, C. Onof, A. Christofides, and Z. W. Kundzewicz, 2022. Revisiting causality using stochastics: 2. Applications. *Proceedings of The Royal Society A*, 478 (2261), 20210836, doi: 10.1098/rspa.2021.0836.
4. D. Koutsoyiannis, C. Onof, Z. W. Kundzewicz, and A. Christofides, 2023. On hens, eggs, temperatures and CO_2: Causal links in Earth's atmosphere. *Sci*, 5 (3), 35, doi: 10.3390/sci5030035.

---

[1] *How many eggs can a hen lay - The Lifecycle of Laying Hens*, https://www.flytesofancy.co.uk/blogs/information-centre/how-many-eggs-can-a-hen-lay.

[2] *The Serpent's Egg*, https://en.wikipedia.org/wiki/The_Serpent%27s_Egg_(film)

[3] The quoted phrase by Hans Fritzsche (of the Reich Ministry of Public Enlightenment and Propaganda in Nazi Germany; https://en.wikipedia.org/wiki/Hans_Fritzsche) is taken from: L. Goldensohn, Hans Fritzsche interview, in *The Nuremberg Interviews*, Ed. by R. Gellately, Vintage Books, New York, 2005.

In these we examined the potentially causal relationship between the atmospheric carbon dioxide concentration ([CO_2]) and atmospheric temperature (T). While the established narrative, supported by conventional wisdom, is that increased [CO_2] causes increase in T, in paper #1 we questioned this conviction and put the relationship of the two in a Plutarchean hen-or-egg framework.

In paper #2 we developed an advanced stochastic methodology for identifying potential causality, which in paper #3 we applied to several problems, including the T – [CO_2] relationship. We concluded that there is a unidirectional, potentially causal link between T as the cause and [CO_2] as the effect. The reverse relationship (i.e., that promoted by the established narrative) was excluded as violating a necessary condition of causality.

In paper #4 we provided more detailed analyses and additional evidence enhancing the validity of the results of paper #3.

I have now published one more paper using several proxy series, extending over the entire Phanerozoic or over parts of it, as well as instrumental data of the modern period. The details are:

5. D. Koutsoyiannis, 2024. Stochastic assessment of temperature – CO_2 causal relationship in climate from the Phanerozoic through modern times. *Mathematical Biosciences and Engineering*, 21 (7), 6560–6602, doi: 10.3934/mbe.2024287.

The extensive analyses confirm the findings of the earlier papers for the modern period and expand them for 500 million years in the past. The results converge to the single inference that change in temperature leads and that in carbon dioxide concentration lags.

Publishing papers that challenge conventional wisdom is not easy at all. Indeed, I struggled to publish each one of the papers that contradict the established climate narrative[4]. I still struggle to publish others which are being reviewed or have been rejected.

The attacks continued after publication of these papers. With “attack” I do not mean criticism, which is healthy and welcomed, as it contributes to improvement of papers and correction of possible errors. What I mean by “attack” is an attempt to block publication of a paper and to silence a voice that does not comply with the established narrative. I also include in “attack” an attempt to force a publisher to retract a published paper.

Such attacks on our papers are mentioned in a long discussion in Judith Curry’s blog[5]. However, these were the exceptions, as most of the criticisms of the papers in the blog were healthy. The extent of the discussion suggests that the papers raised wide interest. It can be regarded as an interesting case of post-publication crowd-reviewing, which our work withstood well. In addition, the discussion offered independent confirmation of our results,

---

[4] My papers and other documents related to climate can be accessed from my web site, https://www.itia.ntua.gr/en/search/?authors=koutsoyiannis&tags=climate. The full list of my journal papers can be accessed at https://www.itia.ntua.gr/en/byauthor/Koutsoyiannis/0/. The full list of my works can be accessed at http://www.itia.ntua.gr/en/search/?title=&authors=koutsoyiannis.

[5] https://judithcurry.com/2023/09/26/causality-and-climate/.

based on a different methodology, and shed light on the physical processes that justify the causality direction found.

I have gathered all comments (about 1000, 18% of which were my own replies) into the following 370-page "book":

6. A. Christofides, D. Koutsoyiannis, C. Onof, and Z. W. Kundzewicz, 2023. *Causality, Climate, Etc.* Climate Etc. (Judith Curry's blog), [doi: 10.13140/RG.2.2.21608.44803](https://doi.org/10.13140/RG.2.2.21608.44803).

What follows is another example of an attack. It refers to the review of paper #5. The first version of the paper was submitted on 29 March 2024, after an invitation by the journal. It received three constructive reviews favouring its publication. The editor's decision was major revision. I addressed all review comments and submitted a revised version on 25 May 2024, along with a detailed report with replies to review comments. All three original reviewers were satisfied with the way I addressed their comments and they recommended publication of the revised manuscript.

However, something unusual happened as two additional reviewers were involved, who tried to block the publication of my paper. Interestingly, the comments of these additional reviewers were focused on papers #1 – #4 and not on #5, the one under review, despite the fact that their comments were purported to be about paper #5. Using material from online attacks on the earlier papers, these hostile reviewers adopted the same tactics claiming that our methodology was inadequate or that our work contained errors.

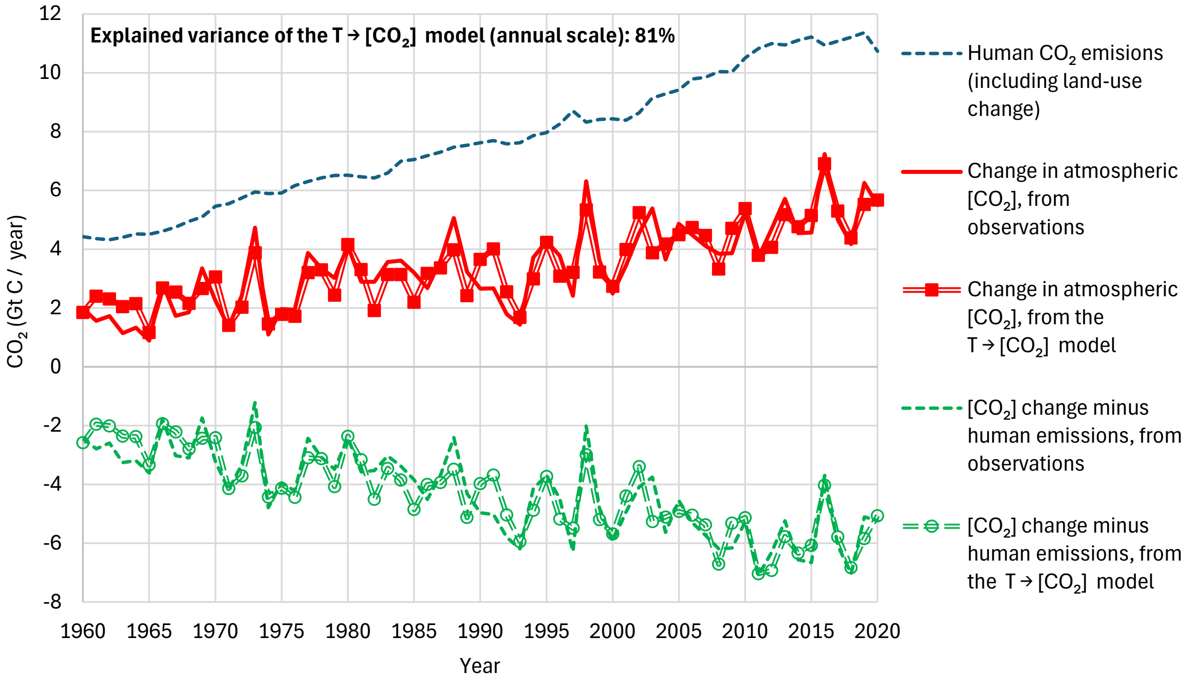
However, if errors were found, then we would either correct them or retract the papers. But not a single error was found. And none of the papers were retracted, despite the efforts of the attackers.

Eventually, paper #5 was accepted for publication, without further changes, based on the rebuttal of these review comments, which I copy in the next pages. In other words, the attacks failed. As explained in the next pages, normally I would not respond to these comments, but once I was forced to do, I thought that it would be fun to make this rebuttal public, as it also responds to earlier attacks.

The only notable change I made to the final paper is that I deleted the motto in its beginning. I explain why in the rebuttal. Yet, I used that motto in the present document as I think it is more relevant to this.

Thus, the next 41 pages is the part of my rebuttal report containing my responses to these two additional reviewers.

A side product of this rebuttal is the following graph, whose initial version is contained in the rebuttal report. The version immediately following below is somewhat improved and contains a detailed description of the graph, its use in the attacks, and its inappropriateness to assess the causality direction. The rebuttal contains detailed instructions how to reproduce the graph.



This graph was used in several versions—but, of course, without the $T \rightarrow [CO_2]$ model curves—as the main weapon (sort of "Scientific Sword Excalibur") in the attacks to our causality papers.

It seems to have first appeared (for years 1960-2005) in Cawley's (2011) paper (in *Energy and Fuels*).

Cawley also used his version in his attack in *PubPeer*, with the (admitted) purpose to force retraction of our papers in *Proceedings of The Royal Society A*. It was also used in a blog with the same purpose.

Subsequently, Engelbeen used his own version to dispute our papers in Judith Curry's blog (Climate Etc.). He uploaded/linked his version eight times in the same blog discussion.

More recently, an anonymous reviewer of my 2024 paper in *Mathematical Biosciences and Engineering* used Cawley's original version to attack the paper under review, aiming at its rejection.

The above version of the graph (initially produced for my rebuttal to that review) contains also the results of our $T \rightarrow [CO_2]$ toy model—Equation (10) in our *Sci* (2023) paper. Notice the impressive percentage of model's explained variance (81% on annual scale, becoming 99.9% on monthly scale).

Despite being used as "Sword Excalibur", this graph, in its original versions, says nothing about causality. It does not contain any information about time precedence. Only my above version says something: that the $T \rightarrow [CO_2]$ potential causality is consistent with the data.

Produced by Demetris Koutsoyiannis

Key for the pages that follow:

Review comment.

Response.

Quotation from manuscript.

Quotation from other documents.

Appendix B: Response to additional reviewers' comments of Round 2 on "Stochastic assessment of temperature – CO_2 causal relationship in climate from the Phanerozoic through modern times"

by Demetris Koutsoyiannis

Reviewer 4

R2-4.1. The author proposes a method for studying causality (X causes Y) based on the regression of one variable (Y) on lagged values of the other (X). On page 22, the author acknowledges that there are other methods for studying causality, such as "Granger causality."

Thanks for the summary. However, it is not accurate. I clearly state that I study the causality between processes, not variables. I start Section 3.2, *Stochastic methodology*, as follows:

> The stochastic methodology used here for identifying potential causal links was developed in [1,2,3] and is based on the impulse response function (IRF) between two stochastic processes $\underline{x}(t), \underline{y}(t)$, denoted as $g(h)$ where h denotes time lag, based on the convolution…

This simple statement clarifies that: (a) I use a new methodology, based on stochastic processes, and (b) that the methodology was already peer-reviewed and published, as seen in the above three references. It was also extensively reviewed at a post-publication phase [4]. Quoting from an even newer paper (Koutsoyiannis, 2024 [5]):

> The latter study [3] raised wide interest and was subsequently discussed in several forums, among which most representative is Judith Curry's blog [6]. With its about 1000 comments, 18% of which were replies by the principal author, this extended discussion, equivalent in length to a book of 370 pages [4], can be regarded as an interesting case of post-publication crowd reviewing, in which the study withstood.

In addition, this post-publication crowd reviewing offered independent confirmation of our results based on a different methodology (using cross-spectral analysis). Based on these facts, I think that it is justified not to repeat in the present paper the full scientific details of the method, the mathematical and factual proofs, and the comparisons with other methods. Any reviewer or reader can find all those in the given references.

As the reviewer correctly notes, I briefly refer to other methods in the following statement:

> As detailed in [2,3], there exist a variety of other methods for estimating IRF and for inferring causality but our method differs conceptually and computationally from them, including from the so-called "Granger causality" [7,8] and the framework proposed by Pearl and collaborators [9-11]

I clarify that this statement refers to (a) other methods for estimating IRF, (b) the so-called "Granger causality" (which, despite its name, does not identify causality but potential for prediction) and (c) the framework proposed by Pearl and collaborators. All these are out of the scope of the present paper, but the interested reviewer or reader may find any detail of these methods in the 2+2 references given above or may refer to our own papers

(Koutsoyiannis et al. [1-3]), as well as to our earlier paper, Koutsoyiannis and Kundzewicz [12], also referred to in my new paper.

> **R2-4.2.** I am more familiar with the latter option than with the method proposed by the author. In Granger causality, two models are studied: one involving the regression of y on its past values, and another involving the regression of y on its past values as well as the past values of X. Subsequently, the significance of the coefficients associated with the lagged values of X is examined. Specifically, it is tested whether including X improves the prediction of y or if Y's past values alone are sufficient for prediction.

I understand that the reviewer is more familiar with Granger's method, as it is an old one (from 1969) while my colleagues' and my methodology is new. I agree with the reviewer when she/he says "*it is tested whether including X improves the prediction of y or if Y's past values alone are sufficient for prediction*". Our framework has many differences with Granger's method, including this one (copied from Koutsoyiannis et al. [1]):

> A second difference is that our focus is upon maximizing not the predictability per se, but the lucidity in identifying the (potentially causal) relationship between two processes $\underline{x}_\tau$ and $\underline{y}_\tau$. This can be seen by comparing Granger's expression in equation (2.11) with our expression in equation (3.20). To estimate y_τ, the former includes terms y_i for times earlier than τ while the second does not. Such terms may increase predictability but say nothing about a potentially causal relationship between the two processes $\underline{x}_\tau$ and $\underline{y}_\tau$; rather, they may obscure that relationship, as autocorrelation is by definition symmetric in time.

In other words, as identification of causality (instead of improvement of predictability) is concerned, the inclusion of term y_i for earlier times obscures (rather than improves) the performance.

I clarify that I appreciate Granger's method (originally developed for econometrics) and by no means do I want to devalue it. It has been successfully applied in several fields but also misused. Granger himself was aware of its misuses as in his Nobel Lecture [13] he stated: "*Of course, many ridiculous papers appeared.*"

> **R2-4.3.** This methodology also has its weaknesses, such as determining the appropriate lags for the regression and including lags for variable X that differ from those in the regression of Y on itself.

I agree. Our method is much stronger in determining lags.

> **R2-4.4.** Another approach involves modeling each series using a model that, when filtering the series, produces another with white noise structure, and calculating the cross-correlation between the two residual series. Depending on the location (right or left of 0) where significant cross-correlation values appear, causality in one direction or the other is indicated.

We had applied the method that seems to be preferred by the reviewer in our earlier paper (Koutsoyiannis and Kundzewicz [12]). Copying from Section 4.1 of that paper:

Yet, we can define a dominant direction of causality based on the time lag η_1 maximizing cross-correlation. Formally, η_1 is defined for a specified ν as

$$\eta_1 \coloneqq \arg\max_{\eta} \left| r_{\tilde{x}\tilde{y}}(\nu, \eta) \right| \tag{13}$$

We can thus distinguish the following three cases:

If $\eta_1 = 0$, then there is no dominant direction.

If $\eta_1 > 0$, then the dominant direction is $\underline{x}_\tau \to \underline{y}_\tau$.

If $\eta_1 < 0$, then the dominant direction is $\underline{y}_\tau \to \underline{x}_\tau$.

Justification and further explanations of these conditions are provided in Appendix A.3.

The interested reviewer or reader may see Appendix A.3 of that paper (Koutsoyiannis and Kundzewicz [12]) for details.

The above approach, which seems to be preferable by the reviewer, was our initial attempt in trying to identify causality. In our later works (Koutsoyiannis et al. [1-3]) we substantially advanced the initial methodology and in the present paper I use the newer and more advanced methodology.

The reviewer may infer the superiority of the new methodology from the fact the old methodology is merely a (non-realistic) special case of the new methodology, in which the impulse response function is reduced to a Dirac delta function. This is clearly stated in the beginning of Section 3.2:

The stochastic methodology used here for identifying potential causal links was developed in [1,2,3] and is based on the impulse response function (IRF) between two stochastic processes $\underline{x}(t), \underline{y}(t)$, denoted as $g(h)$ where h denotes time lag, based on the convolution:

$$\underline{y}(t) = \int_{-\infty}^{\infty} g(h)\underline{x}(t-h)\mathrm{d}h + \underline{v}(t) \tag{3}$$

where $\underline{v}(t)$ is another stochastic process representing the part that is not explained by the causal link. Notice that we use the Dutch notational convention, in which stochastic variables and processes are underlined, while common variables and functions are not.

To see that the function $g(h)$ is the impulse response function (IRF) of the system $(\underline{x}(t), \underline{y}(t))$, we set $\underline{v}(t) \equiv 0$ and $\underline{x}(t) = \delta(t)$ (the Dirac delta function, representing an impulse of infinite amplitude at $t = 0$ and attaining the value 0 for $t \neq 0$), and we readily get $\underline{y}(t) = g(t)$. On the other hand, if we set $g(h) = b\,\delta(h - h_0)$ (with constant b and h_0), which means that the IRF is zero for every lag except for the specific lag h_0, then equation (1) becomes $\underline{y}(t) = b\underline{x}(t - h_0) + \underline{v}(t)$. This special case is equivalent to simply correlating $\underline{y}(t)$ with $\underline{x}(t - h_0)$ at any time instance t. It is easy to find (cf. linear regression) that in this case the multiplicative constant b is the correlation coefficient of $\underline{y}(t)$ and $\underline{x}(t - h_0)$ multiplied by the ratio of the standard deviations of the two processes. In general, however, we expect that the actual $g(h)$ is not a Dirac delta function but a continuous one over some domain. Thus, the IRF is a much more powerful tool than correlation, as it integrates the correlations in the entire spectrum of lags.

R2-4.5. I have serious doubts that the proposed method detects causality in the aforementioned sense.

I believe it is healthy (and should be the rule in science) to have doubts and express them. Since we have published our methodology in three peer-reviewed journal papers, and also presented it in conferences, I invite the reviewer to peruse the methodology and independently test it, based on her/his doubts.

R2-4.6. The results presented are quantified numerically without a study of the significance of these values. For example, equation (5) clearly quantifies whether the regression on X eliminates information about Y or not. If the two variances are similar (coefficient e close to 0), it indicates that X does not cause Y. However, a value of coefficient e close to 1 indicates the opposite. No test is proposed to contrast these hypotheses. In this sense, the values that appear, for example, in Table 1 are not informative. In the "Explained Variance Causal" columna; is the value 0.62 different from zero or is the value 0.11 different from zero, or are both different from zero?

The reviewer is right that a value of e close to 0 indicates nonexistence of causality and a value close to 1 indicates the opposite (i.e. potential causality). Apparently, a value 0.62 provides more evidence than a value of 0.11. This is something common in statistical methods of inference by induction, as opposite to deduction. Giving a significance level of how this value differs from zero, does not make the method deductive. It remains inductive and inferior than before, because it became affected by an arbitrary choice of a significance level, as well as by several assumptions underlying the hypothesis. For example, in her/his comment R2-4.8 below, the reviewer correctly points out the effect of intrinsic correlation, which most statistical tests disregard, thus making wrong inference.

I do not know how familiar the reviewer is with new developments casting doubts on the usefulness of significance testing of “difference from zero”. The following extract from Iliopoulou and Koutsoyiannis [14] provides a summary; it refers to statistical testing of trends in hydrology, but it is also relevant to other geophysical disciplines and other types of tests:

> The most established technique to evaluate fitted trends is statistical hypothesis testing, i.e. a statistical inference technique that estimates the probability of an outcome as far from what is expected as the observed under the assumption that the null hypothesis is true (Gauch, 2003 [15]). The latter is known as the p-value and is compared to predefined significance levels, in order to reject or not the null hypothesis. This is a scientific method for model evaluation, which has been in part misused. For instance, its misuse in hydrology has been showcased by seminal studies (e.g., Cohn and Lins, 2005 [16]; Koutsoyiannis and Montanari, 2007 [17]; Serinaldi et al., 2018 [18]) which have established the fact that for hydrological, non i.i.d. data the null hypothesis, which tacitly contains independence, is a priori wrong, and its rejection, if correctly interpreted,

should point out to the wrong independence assumption. Still, the common practice has been to misinterpret outcomes in favour of trends. Part of the statistician community argues against the concept of significance testing (Nuzzo, 2014 [19]; Wasserstein and Lazar, 2016 [20]; Amrhein and Greenland, 2018 [21]; Trafimow et al., 2018 [22]; Wasserstein et al., 2019 [23]), with the main critique summarized in the statement of the American Statistical Association that "the widespread use of 'statistical significance' (generally interpreted as '$p \leq 0.05$') as a license for making a claim of a scientific finding (or implied truth) leads to considerable distortion of the scientific process" (Wasserstein and Lazar, 2016 [20]).

In particular, the seminal paper Cohn and Lins [16] explains the dramatic impact on statistical inference, based on significance testing, of the model assumed for the process studied (or the inappropriateness thereof, e.g. the neglect of stochastic dynamics such as Long-Term Persistence). I highlight the following phrase from this paper: "*In changing from one test to another, 25 orders of magnitude of significance vanished.*" Additional information can be found in my recent book "*Stochastics of Hydroclimatic Extremes*" [24], where I explain theoretically and with examples that a time series is not a sample and that statistical tests and their related significance are not valid when dealing with time series. Instead, we need advanced Monte Carlo techniques, information of which I give in chapter 7 of the book. I have tried to make this clear more than 20 years ago (Koutsoyiannis, 2003 [25]).

That said, it is self-evident that if the reviewer or any interested reader believes that a statistical test is relevant, she/he may feel free to develop one and publish it. My colleagues and I have published all mathematical details of the methodology in Koutsoyiannis et al. [1-3] and any interested reader may retrieve them from there in order to develop her/his test.

But clearly this is totally out of the scope of the current paper. I have done a lot of work in the present paper—but on another scope as seen in its title. Its length is already 50 pages. I hope the reviewer and any reader would accept the fact that the scope of the paper is that described by the title, the abstract and the introduction of the paper and not any other stuff that I have studied in earlier papers (jointly with other coauthors) or I do not think it deserves studying.

R2-4.7. The restriction (6), where all g_i are positive, is very restrictive.

Again this is something that has been explained in my earlier publications. Quoting from Section 3.3 of Koutsoyiannis et al. [1]:

In contrast to Granger's analysis of causality (…), which treats the processes in discrete time by definition, here we treat them in continuous (i.e. natural) time, and we only convert them to discrete time for estimation purposes. If we think of the processes in natural time, we understand that a causality relationship is not an instantaneous one. In other words, if $\underline{x}(t')$ affects $\underline{y}(t)$, where $t' < t$, it is reasonable to assume that, for small h, $\underline{x}(t' \pm h)$ will also affect $\underline{y}(t)$. Therefore, the IRF, $g(h)$, is not a Dirac delta function, but one with some domain, $\mathbb{h} \subseteq \mathbb{R}$, of nonzero (and potentially infinite) measure, where $g(h) \neq 0$ for $h \in \mathbb{h}$. It is also reasonable to assume that $g(h)$ is a continuous function and has the same sign for all $h \in \mathbb{h}$. The latter can be justified as follows. If $\underline{x}(t')$ is positively correlated with $\underline{y}(t)$, then it is reasonable that $\underline{x}(t' \pm h)$ are also positively correlated with $\underline{y}(t)$. Without loss of generality, in what follows we will assume that $g(h) \geq 0$ for $h \in \mathbb{h}$

(if it were $g(h) \leq 0$, we would reflect $\underline{x}(t)$, i.e. replace it with $-\underline{x}(t)$, and hence $g(h)$ would also be reflected becoming nonnegative).

Here we clarify that the problem of identifying causality is different from that of recovering the full system dynamics. The former and not the latter, is the scope of our study. We note that, while there exist oscillatory nonlinear systems, in which the sign of $g(h)$ could alternate, we avoid subsuming them under the causality notion, particularly when causality is inferred from data in an inductive manner. This choice is consistent with Cox's (1992 [26]) conditions for causality, according to which the effect *"shows a monotone relation with 'dose'"* of the cause. Here we note that in our framework the "dose" is not regarded as an instantaneous event, but one with some time span (see details in Supplementary Information, section SI1.2).

R2-4.8. The method for determining h_c on page 21 is based on the calculation of cross-correlation on the series X and Y. This cross-correlation is affected by the intrinsic correlation of each series, and the significance study of each value at each lag can be erroneous.

I agree. The significance of this cross-correlation value can be erroneous. That is the reason why I do not calculate at all and did not give the significance level in the paper (see also my reply to comment R2-4.6 above). For the same reason, I give the emphasis on the other two indices, the mean (time average), μ_h, and the median, $h_{1/2}$, of the sequence g_j. Yet I additionally give the value of the cross-correlation (without its significance), for the completeness of the presentation. As mentioned in section 3.2 for μ_h and $h_{1/2}$:

extensive analyses in [1] showed that their estimation is quite robust.

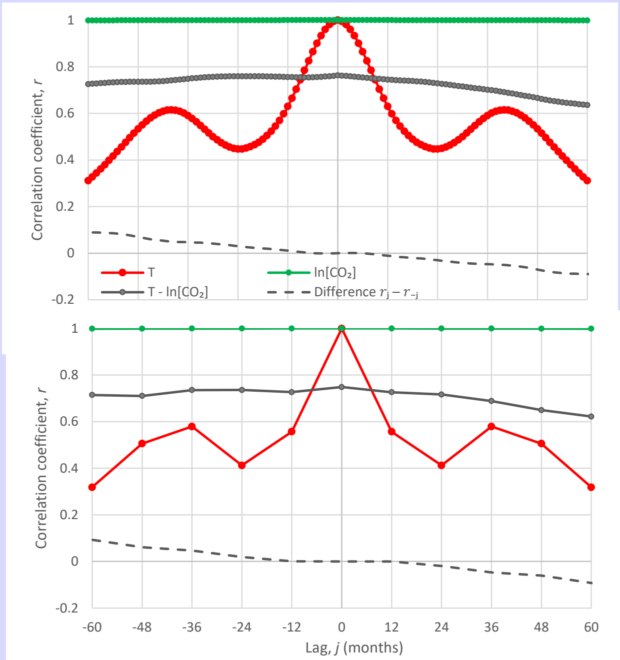
R2-4.9. Apply the "Granger Causality" method to the data and compare it with your method.

I must thank the reviewer for reminding me of my youth, some 40 years ago, when I was a seaman in the Greek Navy for 25 months. This is because this last comment sounds like a command that I must execute.

So: **Yes Sir!** I have already executed it four years ago. Quoting from Section 5.1 of Koutsoyiannis and Kundzewicz [12]:

Somewhat more informative is Figure 9, which depicts lagged cross-correlations of the two processes, based on the methodology in Section 4.1 but without differencing the processes. Specifically, Figure 9 shows the cross-correlogram between UAH temperature and Mauna Loa $\ln[CO_2]$ at monthly and annual scales; the autocorrelograms of the two processes are also plotted for comparison. In both time scales, the cross-correlogram shows high correlations at all lags, with the maximum attained at lag zero. This does not hint at a direction. However, the cross-correlations for negative lags are slightly greater than those in the positive lags. Notice that to make this clearer, we have also plotted the differences $r_j - r_{-j}$ in the graph. This behaviour could be interpreted as supporting the causality direction $[CO_2] \rightarrow T$. However, we deem that the entire picture is spurious as it is heavily affected by the fact that the autocorrelations are very high and, in particular, those of $\ln[CO_2]$ are very close to 1 for all lags shown in the figure.

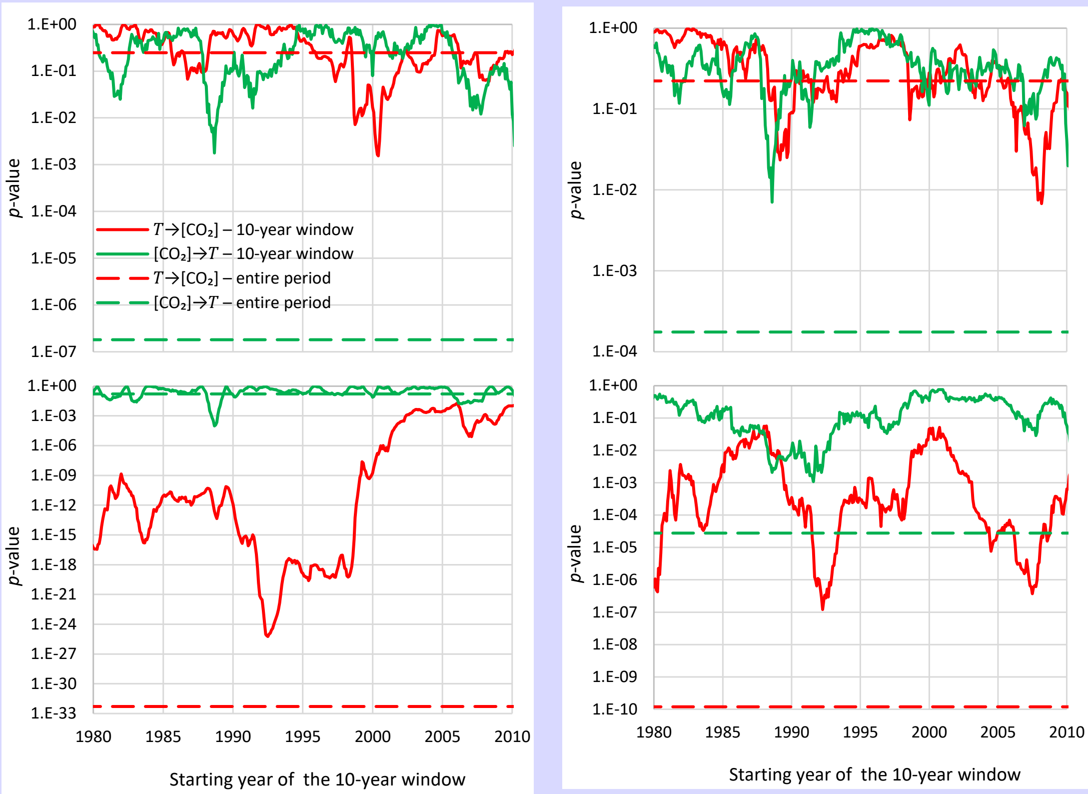
In our investigation, we also applied the Granger test on these two time series in both time directions. To calculate the p-value of the Granger test, we used free software (namely the function GRANGER_TEST [27,28]). It appears that in the causality direction $[CO_2] \rightarrow T$, the null hypothesis is rejected at all usual significance levels. The attained p-value of the test is 1.8×10^{-7} for one regression lag ($\eta = 1$), 1.8×10^{-4} for $\eta = 2$, and remains below 0.01 for subsequent η. By contrast, in the direction $T \rightarrow [CO_2]$, the null hypothesis is not rejected at all usual significance levels. The attained p-value of the test is 0.25 for $\eta = 1$, 0.22 for $\eta = 2$, and remains above 0.1 for subsequent η.



[Figure R2-4.1; reproduced from Koutsoyiannis and Kundzewicz [12]; original caption follows]
Figure 9. Auto- and cross-correlograms of the time series of UAH temperature and logarithm of CO_2 concentration at Mauna Loa.

Therefore, one could directly interpret these results as unambiguously showing one-way causality between the total greenhouse gases and temperature and, hence, validating the consensus view that human activity is responsible for the observed rise in global temperature. However, these results are certainly not unambiguous and, most probably, they are spurious. To demonstrate that they are not unambiguous, we have plotted, as shown in the upper panels of Figure 10, the p-values of the Granger test for moving windows with a size of 10 years for number of lags $\eta = 1$ and 2. The values for the entire length of time series, as given above, are also shown as dashed lines. Now the picture is quite different: each of the two directions appear dominating (meaning that the attained significance level is lower in one over the other) in about equal portions of the time. For example, for $\eta = 2$, the $T \rightarrow [CO_2]$ dominates over $[CO_2] \rightarrow T$ for 58% of the time. The attained p-value for direction $T \rightarrow [CO_2]$ is lower than 1% for 1.4% of the time, much higher than in the opposite direction (0.3% of the time). All of these observations favour the $T \rightarrow [CO_2]$ direction.

To show that the results are spurious and, in particular, affected by the very high autocorrelations of ln [CO_2] and, more importantly, by its annual cyclicity, we have "removed" the latter by averaging over the previous 12 months. We did that for both series and plotted the results in the lower panels of Figure 10. Here, the results are stunning. For both lags η = 1 and 2 and for the entire period (or almost), $T \rightarrow [CO_2]$ dominates, attaining p-values as low as in the order of 10^{-33}. However, we will avoid interpreting these results as unambiguous evidence that the consensus view (i.e., human activity is responsible for the observed warming) is wrong. Rather, what we want to stress is that it is inappropriate to draw conclusions from a methodology which is demonstrated to be so sensitive to the used time windows and data processing assumptions. In this respect, we have included this analyses in our study only (a) to show its weaknesses (which, for the reasons we explained in Section 4.2, we believe would not change if we used different statistics or different time series) and (b) to connect our study to earlier ones. For the sake of drawing conclusions, we contend that our full methodology in Sections 4.1 and 4.3 is more appropriate. We apply this methodology in Section 5.2.



[Figure R2-4.2; reproduced from Koutsoyiannis and Kundzewicz [12]; original caption follows]
Figure 10. Plots of p-values of the Granger test for 10-year-long moving windows for the monthly time series of UAH temperature and logarithm of CO_2 concentration at Mauna Loa for number of lags (**left**) η = 1 and (**right**) η = 2. The time series used are (**upper**) the original and (**lower**) that obtained after "removing" the periodicity by averaging over the previous 12 months.

Also, quoting from Section 5.2 of the same paper (Koutsoyiannis and Kundzewicz [12]), referring to the differenced time series:

> While, as explained in Sections 4.2 and 5.1, the Granger test has weaknesses that may not help in drawing conclusions, for completeness and as a confirmation, we list its results here:
>
> - For the monthly scale and the causality direction $[CO_2] \rightarrow T$, the null hypothesis is not rejected at all usual significance levels for lag $\eta = 1$ and is rejected for significance level 1% for $\eta = 2$–8, with minimum attained p-value 1.8×10^{-4} for $\eta = 6$.
> - For the monthly scale and the causality direction $T \rightarrow [CO_2]$, the null hypothesis is rejected at all usual significance levels for all lags η, with minimum attained p-value 2.1×10^{-8} for $\eta = 7$.
> - For the monthly scale, the attained p-values in the direction $T \rightarrow [CO_2]$ are always smaller than in direction $[CO_2] \rightarrow T$ by about 4 to 5 orders of magnitude, thus clearly supporting $T \rightarrow [CO_2]$ as dominant direction.
> - For the annual scale with fixed year specification and the causality direction $[CO_2] \rightarrow T$, the null hypothesis is not rejected at all usual significance levels for any lag η, thus indicating that this causality direction does not exist.
> - For the annual scale with fixed year specification and the causality direction $T \rightarrow [CO_2]$, the null hypothesis is not rejected at significance level 1% for all lags $\eta = 1$–6, with minimum attained p-value 5% for lag $\eta = 2$, thus supporting this causality direction at this significance level.
> - For the annual scale with fixed year specification, the attained p-values in the direction $T \rightarrow [CO_2]$ are always smaller than in direction $[CO_2] \rightarrow T$, again clearly supporting $T \rightarrow [CO_2]$ as the dominant direction.
>
> We note that the test cannot be applied for the sliding time window case and, hence, we cannot provide results for this case.
>
> […]
>
> In brief, all above confirm the results of our methodology that the dominant direction of causality is $T \rightarrow [CO_2]$.

Reviewer 5

Note: The reviewer's citations, originally denoted as [1] etc., have been changed to [[1]] etc. so as to be distinguished from the citations of my rebuttal report. Those citations that refer to my own paper, which the reviewer also denotes as [1] etc., have been changed to {1} etc., again to be distinguished from the citations of my rebuttal report.

> **R2-5.1.** This paper investigates the use of an existing statistical method for causal inference to investigate the causality linking CO2 and temperature. Sadly the inference that for the instrumental data T causes CO2 is demonstrably incorrect. The error is due to the differencing of the time series, which decouples the long term linear trend entirely from the analysis. The finding that T causes CO2 in paleoclimate is overly simplistic, but entirely

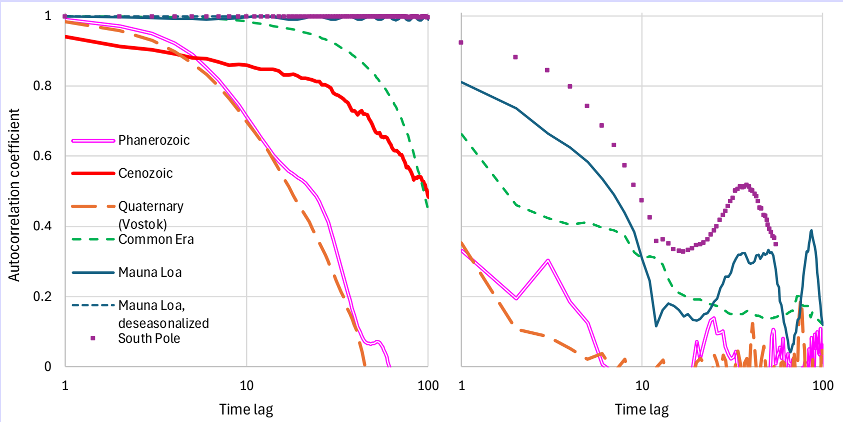
uncontraversial. There is no clear statement of the mathematical novelty of this paper (the moving average analsysis of the instrumental data? However that is a relatively minor contribution and the results are unreliable because of the differencing). I see no scope for the paper being modifiable to a state where it warrants publication.

I understand that the reviewer does not want my paper to be published. This is clear in the last sentence, "*I see no scope for the paper being modifiable to a state where it warrants publication*".

But it is also clear that the reviewer's critiques do not apply to the present paper. They apply to papers that have already published (Koutsoyiannis and Kundzewicz [12]; Koutsoyiannis et al. [1-3]). For "*the differencing of the time series*" which the reviewer regards to be an error has been thoroughly explained and justified in all four papers. In the present paper I also apply the method with non-differenced series where applicable. Apparently, the reviewer has not seen it in the paper as she/he seems to have consulted several texts that (unsuccessfully) have attacked our earlier papers (not the present one).

The reviewer's claim that "*the results are unreliable because of the differencing*" is totally wrong. Only its negation is right, i.e. "*the results are **reliable** because of the differencing*" or "the results *would be unreliable **without** the differencing*". As shown in Koutsoyiannis and Kundzewicz [12] and reproduced above in my reply to comment R2-4.9 (by Reviewer #4), when the autocorrelations are close to 1 for the lags of interest, without differencing the "*results are certainly not unambiguous and, most probably, they are spurious*". In the present paper, the autocorrelations are close to 1 at all lags of interest for the instrumental data, as seen in the following part of section 4.1 and the Figure 9 of the paper:

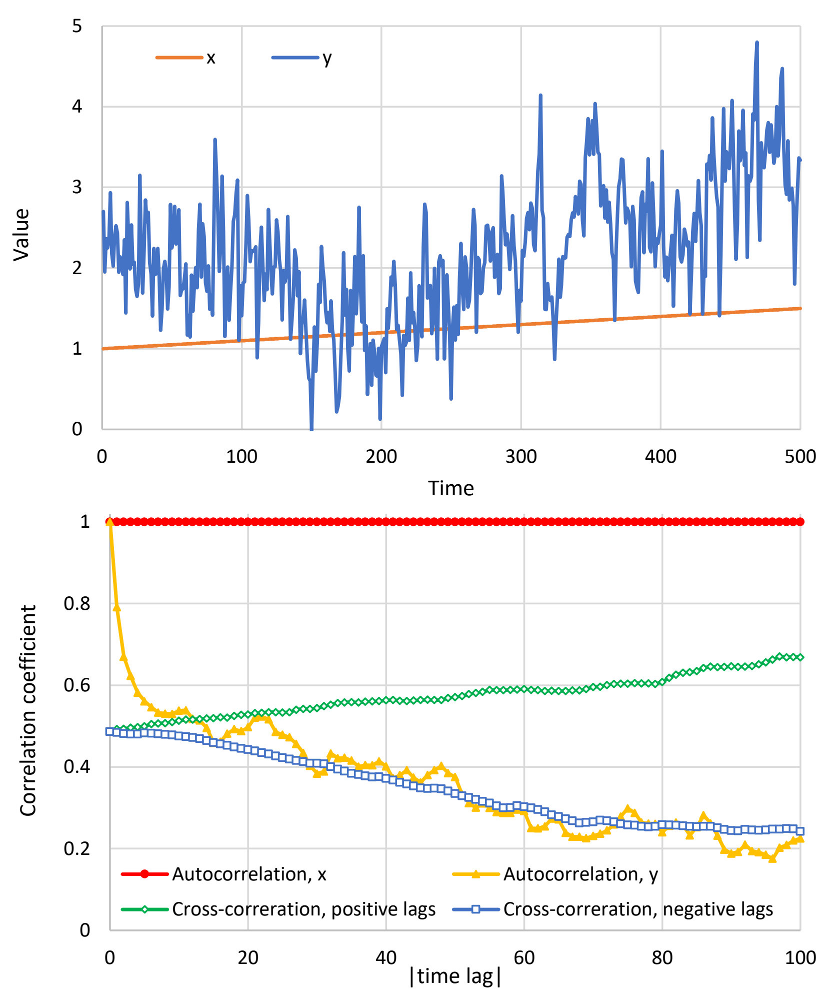
> Figure 11 shows the empirical autocorrelation functions of the [CO_2] series, original and differenced. In the instrumental series, the autocorrelations are almost 1, even for lags as high as 100. This prohibits any inference from the original series. However, their differenced series have reasonable positive autocorrelations, which make inference possible. The proxy series have high autocorrelations at small lags, but reasonable ones at large lags. For those we examined both the original and the differenced series, provided that the latter are positive.



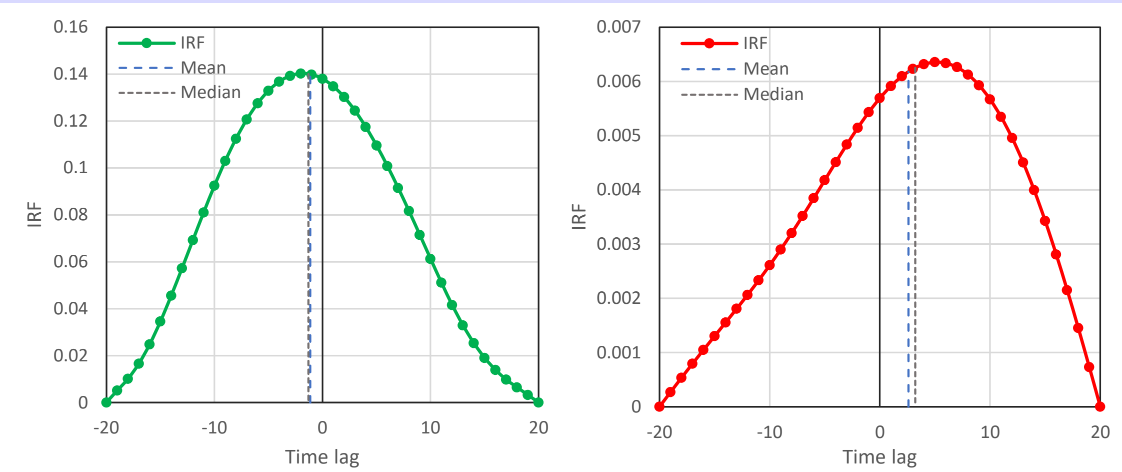
[Figure R2-5.1; original caption follows]
Figure 11 Autocorrelation functions of [CO_2] series: (**left**) original; (**right**) differenced. The differenced series autocorrelation for Cenozoic is not plotted as it is mostly negative. The time lag is in discrete time j, i.e. dimensionless, and, to make it dimensional, we should multiply by the time step Δ of each series ($h = j\Delta$).

The reason that autocorrelations close to 1 do not enable inference on causality, has been explained in several places of the earlier works. As an example, below I am reproducing the section SI2.2 from the Supplementary Information of Koutsoyiannis et al. [2]:

> **SI2.2 On high autocorrelations and spurious IRF estimates**
>
> As stated in the main papers (Koutsoyiannis et al., 2022a,b [1,2]), high autocorrelation results in increased estimation uncertainty and may even result in spurious causality claims. To illustrate this, we devise a synthetic example, in which the processes $\underline{x}_\tau$ and $\underline{y}_\tau$ are, by construction, independent of each other and with high autocorrelation.
>
> Specifically, two time series x_τ and y_τ, each of length 500, are generated independently from each other. The time series x_τ is constructed by the deterministic rule $x_\tau = 1 + 0.001\tau$. If its values are treated statistically, the resulting autocorrelation estimate is constant for all lags, $\hat{r}_{xx}(h) = 1$. The time series y_τ is generated from a Hurst-Kolmogorov process with Gaussian distribution and with a high Hurst parameter, $H = 0.95$, reflecting strong long-range dependence (LRD). By now, it is well known (e.g. Cohn and Lins, 2005 [16]; Koutsoyiannis, 2013 [29]) that realizations of processes with LRD look "trendy" even though the processes are stationary. This is evident in Figure SI2.5 (upper), which depicts both time series. Their auto- and cross-correlations, estimated using standard statistical estimators, are shown in Figure SI2.5 (lower). Interestingly, while by construction the cross correlations are $r_{yx}(h) = 0$ for any lag h, their estimates $\hat{r}_{yx}(h)$ appear very high, i.e., 0.46 ± 0.21 in the plotted interval of lag h, $(-100,100)$.
>
> Because of the high cross-correlations, if we estimate the IRF with the proposed method, as seen in Figure SI2.4, spurious Hen-or-Egg (HOE) causality is identified in both directions $\underline{x} \rightarrow \underline{y}$ and $\underline{y} \rightarrow \underline{x}$. The dominant causality direction appears to be $\underline{y} \rightarrow \underline{x}$ with mean lag $\mu_h = 2.6$, median lag $h_{1/2} = 3.2$ and explained variance ratio $e = 0.47$. All these are obviously invalid estimates (as there are no true lags in this case and the true value of the explained variance ratio is $e = 0$), even though the calculations are correct.
>
> Naturally, a remedy in such spurious cases is to reduce the autocorrelations. This becomes possible if instead of the time series x_τ and y_τ we study the differenced time series $\Delta x_\tau \coloneqq x_\tau - x_{\tau-1}$ and $\Delta y_\tau \coloneqq y_\tau - y_{\tau-1}$. Taking the differences is definitely reasonable: if x_τ causes y_τ, then a change in x_τ should cause a change in y_τ. In our example, we will have constant $\Delta x_\tau = 0.001$ and hence the cross-covariances would be zero, which will exclude any causality claim.



[Figure R2-5.2; reproduced from Koutsoyiannis et al. [3]; original caption follows]
Figure SI2.3 (**upper**) Time series of the synthetic example described in the text. (**lower**) Auto- and cross-correlation function estimates for the two time series.



[Figure R2-5.3; reproduced from Koutsoyiannis et al. [3]; original caption follows]
Figure SI2.4 IRFs for the synthetic example of spurious IRF estimation due to high autocorrelation of Figure SI2.3 for causality directions (left) $\underline{x} \rightarrow \underline{y}$ and (right) $\underline{y} \rightarrow \underline{x}$. For the estimated IRFs the number of weights is $2J + 1$ with $J = 20$.

If the reviewer thinks that the above four papers, which were prepared jointly with other colleagues, are "*demonstrably incorrect*" then I invite her/him to demonstrate that with mathematical and logical arguments, and publish her/his demonstration in a scientific journal. This would be more constructive than trying to block the publication of my new paper as an anonymous reviewer by repeating what she/he saw in attacks of earlier papers on the internet (some of which he also cites; see comment R2-5.14 below).

In addition, publishing her/his demonstrations would give the opportunity of a public dialogue.

> **R2-5.2.** The paper is perhaps overlong, but the formatting, use of language and structure are all satisfactory.

I welcome this remark.

> **R2-5.3. Abstract**
>
> The fundamental problem with this paper is evident in the final sentence of the abstract:
>
> > "These results contradict the conventional wisdom, according to which the temperature rise is caused by [CO_2] increase."
>
> Firstly, this is not merely "conventional wisdom", it is the conclusion of over a century of scientific research (known as the ``enhanced greenhouse effect" - EGHE), including a well-understood causal mechanism, and comprelling experimental and observational evidence of the process as a whole, and it's component parts. If a physics-free, purely statistical method, contradicts such strong scientific finding, then it would be prudent to be deeply skeptical of the statistical method (which does not have the benefit of a causal mechansim or evidence for the component parts, or consilience with other scientific findings). In this case, the flaw in the statistical method is easily identified (differencing of the time series), it is an error that has been made on numerous occasion, e.g. Salby and Humlum et al. etc., it has also previously been drawn to the author's attention prior to the submission of the current paper.
>
> From research on the EGHE and on the carbon cycle, we know that there are causal mechanisms in **both** directions: The EGHE means that as CO2 levels rise, all things being otherwise equal, there will be a rise in global temperature, modified by climate feedbacks, such as water vapour feedback. However, multiple carbon cycle feedback mechanisms mean that changes in global temperatures will affect the level of atmospheric CO2. Like the EGHE, there is also considerable research on carbon cycle feedback. It is the interplay between these forces that results in a dynamic equilibrium, that has kept the Earth's temperature and it's atmospheric CO2 reasonably stable until large scale anthropogenic emissions disturbed this equilibrium, coinciding with the industrial revolution. Any model of the relationship between global temperatures and atmospheric CO2 that does not include these feedbacks is at variance with observed reality and should be discarded - it doesn't match what we observe to be the case.

Yes, "conventional wisdom" could be the result of over-century research. Longevity does not make it unconventional. Neither makes it correct. It is well known that Aristotle's geocentric system was conventional wisdom for about 18 centuries, on which conventional research of its time was based. At the same time, Arisctarchus's heliocentric system has been rejected for 18 centuries. On the other hand, Aristotle's correct explanation of the Nile's floods was rejected for 21 centuries (Koutsoyiannis, and Mamassis [30]).

Given that conventional wisdom is not necessarily correct, I think I have the right to challenge it in a scientific context.

The statement cited by the reviewer from the abstract reflects the content of the paper and is 100% accurate. No modification is required.

Note that the statement does not say that conventional wisdom is wrong. It says that our results contradict conventional wisdom. If conventional wisdom is right, then we are wrong. This is likely but must be proven with mathematical and logical arguments using the scientific method. None of the attacks made so far, including the one by Reviewer #5, has any of these characteristics.

There may be "*physics-free, purely statistical methods*", but when applied to physical problems, statistical methods become parts of physics. Clockwise physics, without using probability and statistics, has been conventional wisdom for a couple of centuries but has proved to be weak and inadequate. Hence, stochastics has long ago been incorporated into physics. This occurred one century and a half ago, but admittedly, many of us, including this reviewer, are not updated on this fact yet and continue to contrast physics and statistics. Therefore, I am providing the following information in bulleted form (along with my apology for being didactic):

- Statistical physics (cf. Boltzmann, Gibbs, Planck) used the probabilistic concept of entropy (which is nothing other than a quantified measure of uncertainty defined within the probability theory) to explain fundamental physical laws (most notably the Second Law of thermodynamics), thus leading to a new understanding of natural behaviors and to powerful predictions of macroscopic phenomena. Atmospheric processes are explained by statistical physics in all respects (thermodynamic equilibrium, blackbody radiation, transport processes)

- Quantum theory (cf. Heisenberg) has emphasized the intrinsic character of uncertainty and the necessity of probability in the description of nature.

- Developments in numerical mathematics for applications in physics (cf. Metropolis) highlighted the effectiveness of stochastic methods in solving physical problems that are even purely deterministic, such as numerical integration in high-dimensional spaces and global optimization of non-convex functions (where stochastic techniques, e.g., stochastics-based evolutionary algorithms and simulated annealing, are in effect the only feasible solution in complex problems that involve many local optima).

This extends even beyond physics. Thus,

- Genetics (cf. Mendel) and evolutionary biology have emphasized the importance of stochasticity (e.g., in gametes fusion, selection and mutation procedures, and environmental changes) as a driver of evolution.

- Developments in mathematical logic, and particularly Gödel's incompleteness theorem, challenged the almightiness of deduction (inference by mathematical proof). This necessitates the use of induction in physical problems, whose theoretical basis is offered by the field of stochastics

It is completely untrue that in our papers we do not "*benefit of a causal mechanism or evidence for the component parts, or consilience with other scientific findings*". Here I repeat the related parts from our earlier papers.

From Koutsoyiannis and Kundzewicz [12], Section 6:

> The omnipresence of positive lags on both monthly and annual time scales and the confirmation by Granger tests reduce the likelihood that our results are statistical artefacts. Still, our results require physical interpretation which we seek in the natural process of soil respiration.
>
> Soil respiration, R_s, defined to be the flux of microbially and plant-respired CO_2, clearly increases with temperature. It is known to have increased in the recent years [31,32]. Observational data of R_s (e.g., [33,34]; see also [35]) show that the process intensity increases with temperature. Rate of chemical reactions, metabolic rate, as well as microorganism activity, generally increase with temperature. This has been known for more than 70 years (Pomeroy and Bowlus [36]) and is routinely used in engineering design.
>
> The Figure 6.1 of the latest report of the IPCC [32] provides a quantification of the mass balance of the carbon cycle in the atmosphere that is representative of recent years. The soil respiration, assumed to be the sum of respiration (plants) and decay (microbes), is 113.7 Gt C/year (IPCC gives a value of 118.7 including fire, which along with biomass burning, is estimated to be 5 Gt C/year by Green and Byrne [37]).
>
> We can expect that sea respiration would also have increased. Moreover, outgassing from the oceans must also have increased as the solubility of CO_2 in water decreases with increasing temperature [38,39]. In addition, photosynthesis must have increased, as in the 21st century the Earth has been greening, mostly due to CO_2 fertilization effects [40] and human land-use management [41]. Specifically, satellite data show a net increase in leaf area of 2.3% per decade [41]. The sums of carbon outflows from the atmosphere (terrestrial and maritime photosynthesis as well as maritime absorption) amount to 203 Gt C/year. The carbon inflows to the atmosphere amount to 207.4 Gt C/year and include natural terrestrial processes (respiration, decay, fire, freshwater outgassing as well as volcanism and weathering), natural maritime processes (respiration) as well as anthropogenic processes. The latter comprise human CO_2 emissions related to fossil fuels and cement production as well as land-use change, and amount to 7.7 and 1.1 Gt C/year, respectively. The change in carbon fluxes due to natural processes is likely to exceed the change due to anthropogenic CO_2 emissions, even though the latter are generally regarded as responsible for the imbalance of carbon in the atmosphere.

From Koutsoyiannis et al. [2], Section 3:

> In other words, it is the increase of temperature that caused increased CO_2 concentration. Though this conclusion may sound counterintuitive at first glance, because it contradicts common perception (and for this reason we have assessed the case with an alternative parametric methodology in the Supplementary Information, section SI2.4, with results confirming those presented here), in fact it is reasonable. The temperature increase began at the end of the Little Ice Period, in the early 19th century, when human CO_2 emissions were negligible; hence other factors, such as the solar activity (measured by sunspot numbers), as well as internal long-range mechanisms of the complex climatic systems had to play their roles.
>
> A possible physical mechanism for the [CO_2] increase, as a result of temperature increase, was proposed by Koutsoyiannis and Kundzewicz [12] and involves biochemical reactions, as, at higher temperatures, soil respiration, and hence natural CO_2 emissions, are increasing. In addition, as pointed out by Liu et al. [42] the influence of El Niño on climate is accompanied by large changes to the carbon cycle, with the pantropical biosphere releasing much more carbon into the atmosphere during large El Niño occurrences. Noticeably, in a very recent paper, Goulet Coulombe and Göbel [43] seem to confirm the finding by Koutsoyiannis and Kundzewicz [12], yet they deem it an "*apparently counterintuitive finding that GMTA* [global mean surface temperature anomalies] *explains a larger portion of the forecast error variance of CO_2 than vice versa*". To "*resolve*" it, they "*explore a last avenue, that of using annual CO_2 emissions*". However, using anthropogenic CO_2 emissions, which contribute only a small portion (3.8%) to the global carbon cycle (Koutsoyiannis [44]), as a principal variable is definitely less meaningful than using the atmospheric CO_2 concentration.
>
> We believe that counterintuitive results, such as those about the causal link between temperature and CO_2 concentration conveyed in this paper, can indeed open up avenues of research. However, these avenues of research might not resolve the issue in a way compatible with what intuition dictates. In the history of science, such avenues were often created when established ideas were overturned by new findings.

The following extract from Koutsoyiannis et al. [3], Section 9 provides additional information and also disputes the Goulet Coulombe and Göbel [43] "avenue":

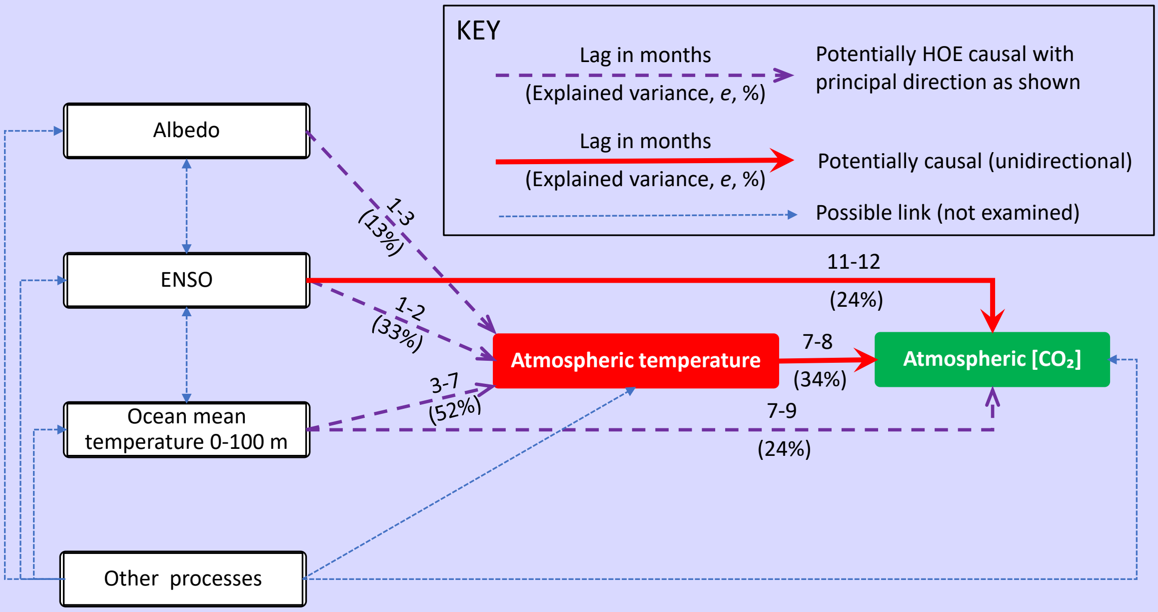
> In terms of the carbon cycle […], several physical, chemical, biochemical and human processes are involved in it. The human CO_2 emissions due to the burning of fossil fuels have largely increased since the beginning of the industrial age. However, the global temperature increase began succeeding the Little Ice Period, at a time when human CO_2 emissions were very low. To cast light on the problem, we examine the issue of CO_2 emissions vs. atmospheric temperature further in the Supplementary Information, where we provide evidence that they are not correlated with each other. The outgassing from the sea is also highlighted sometimes in the literature among the climate-related mechanisms. On the other hand, the role of the biosphere and biochemical reactions is often downplayed, along with the existence of complex interactions and feedback. This role can be summarized in the following points, examined in detail and quantified in Appendix A1.

- Terrestrial and maritime respiration and decay are responsible for the vast majority of CO_2 emissions [45, Figure 5.12].
- Overall, natural processes of the biosphere contribute 96% to the global carbon cycle, the rest, 4%, being human emissions (which were even lower in the past [44]).
- The biosphere is more productive at higher temperatures, as the rates of biochemical reactions increase with temperature, which leads to increasing natural CO_2 emission [12].
- Additionally, a higher atmospheric CO_2 concentration makes the biosphere more productive via the so-called carbon fertilization effect, thus resulting in greening of the Earth [40,46], i.e., amplification of the carbon cycle, to which humans also contribute through crops and land-use management [41].

In addition to the biosphere, there are other factors that drive the Earth's climate in periodic and non-periodic way. Orbital parameters of Earth's revolution change quasi-cyclically in a multi-millennial scale (variations in eccentricity, axial tilt, and precession of Earth's orbit), as interpreted by Milanković [47-51], and changes in the orbit geometry influence the amount of insolation. The non-periodic drivers of the Earth's climate variability include volcanic eruptions and collisions with large extraterrestrial objects, e.g., asteroids. An important climate driver is water in its three phases [44]. Another apparent factor is solar activity (including solar cycles) and the solar radiation (im)balance on Earth (e.g., albedo changes; see [44] and Appendix A2). Notably, a recent study [52], by assessing 20 years of direct observations of energy imbalance from Earth-orbiting satellites, showed that the global changes observed appear largely from reductions in the amount of sunlight scattered by Earth's atmosphere.

ENSO and ocean heating, both of which affect temperature, are examined in Appendices A3 and A4, respectively. The results of Appendices A2–A4 are summarized in the schematic of Figure 13. Changes in all three examined processes, albedo, ENSO and the upper ocean heat, precede in time the changes in temperature and even more so those in [CO_2]. Generally, the time lags shown in Figure 13 complete a consistent picture of potential causality links among climate processes and always confirm the $T \rightarrow [\mathrm{CO_2}]$ direction.

The examined processes in the Appendices are internal to the climatic system. Other processes affecting T, not examined here, could also be external (e.g., solar and astronomical [53,54] and geological [55-59]). Generally, in complex systems, an identified causal link, even though it gives some explanation of a phenomenon, raises additional questions, e.g., what caused the change in the identified cause, etc. In turn, causal links in complex systems may form endless sequences. For this reason, it is naïve to expect complete answers to problems related to complex systems or to assume that a complex system is in permanent equilibrium and that an external agent is needed to "kick" it out of the equilibrium and produce change. Yet the investigation of a single causal link between two processes, as is the focus of this paper, provides useful information, with possible significant scientific, technical, practical, epistemological and philosophical implications. These are not covered in this paper. Readers interested in epistemological and philosophical aspects of causality are referred to Koutsoyiannis et al. [1], while those interested in the perennial changes in complex systems are referred to Koutsoyiannis [60,61].



[Figure R2-5.4; reproduced from Koutsoyiannis et al. [3]; original caption follows]
Figure 13. Schematic of the examined possible causal links in the climatic system, with noted types of potential causality, HOE or unidirectional, and its direction. Other processes, not examined here, could be internal of the climatic system or external.

Finally, Appendix A.1 from Koutsoyiannis et al. [3] provides quantification of changes in natural CO_2 emissions due to temperature changes:

The greatest part of the inflows is due to the respiration of the biosphere, i.e., the biochemical reaction whereby living organisms convert organic matter (e.g., glucose) to CO_2, releasing energy and consuming molecular oxygen [62]. As seen in Figure A1 (and in several publications, e.g., [31]), respiration has increased in recent years, the main reason for this being the increased temperature. Photosynthesis, the biochemical process that removes CO_2 from the atmosphere, producing carbohydrates in plants, algae and bacteria using the energy of light [62], has also increased, resulting in the greening of Earth [40,41] due to the increased atmospheric concentration of CO_2, which is plants' food.

It is not difficult to quantify the increase in respiration due to the temperature rise. The mechanism has been known in chemistry for more than a century. The rate of a chemical reaction k_T at temperature T is an increasing function of T, given by the Arrhenius equation [63]:

$$k_T = A \exp\left(-\frac{a}{R_* T}\right) \tag{A1}$$

where A and a are free parameters and R_* is the universal gas constant. Typically, the rate is measured in moles per unit volume, but it can readily be expressed in mass units. Expressing the relationship at a reference temperature T_0 and dividing with (A1), we obtain:

$$\frac{k_T}{k_0} = \exp\left(-\frac{a}{R_*}\left(\frac{1}{T} - \frac{1}{T_0}\right)\right) \tag{A2}$$

Taking the logarithms and setting $\Delta T \coloneqq T - T_0$ we find

$$\begin{aligned} \ln\left(\frac{k_T}{k_0}\right) &= -\frac{a}{R_*}\left(\frac{1}{T} - \frac{1}{T_0}\right) = \frac{a}{R_* T_0}\left(1 - \frac{T_0}{T}\right) = \frac{a}{R_* T_0}\left(\frac{\Delta T}{T_0 + \Delta T}\right) \\ &= \frac{a}{R_* T_0}\left(\frac{\Delta T}{T_0} - \left(\frac{\Delta T}{T_0}\right)^2 + \left(\frac{\Delta T}{T_0}\right)^3 - \cdots\right) \end{aligned} \tag{A3}$$

and assuming that $\Delta T / T_0$ is small (nb., T_0 is of the order of 300 K, while typical values of ΔT is of the order of 1–10 K). We can neglect all terms beyond first order and find:

$$\frac{k_T}{k_{T_0}} = \exp\left(\frac{a}{R_* T_0^2}\,\Delta T\right) = \left(\exp\left(\frac{a}{R_* T_0^2}\right)\right)^{\Delta T} = Q_1^{\Delta T} = Q_{10}^{\Delta T/10} \tag{A4}$$

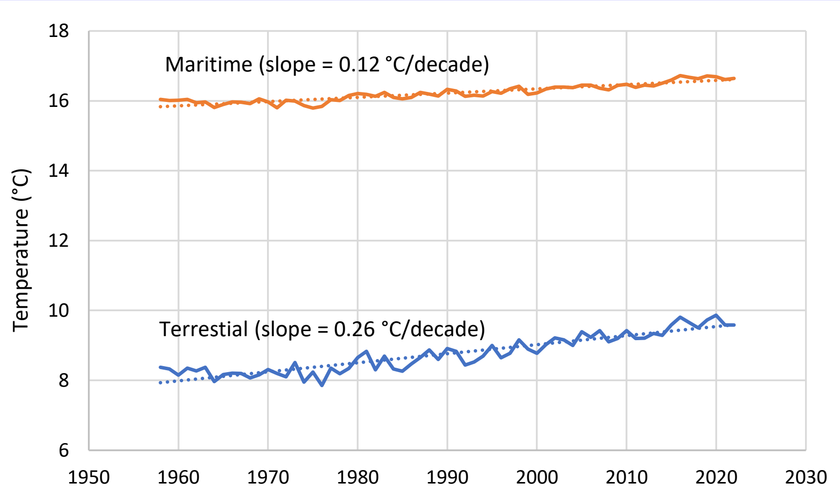
where

$$Q_1 \coloneqq \exp\left(\frac{a}{R_* T_0^2}\right), \qquad Q_{10} \coloneqq Q_1^{10} \tag{A5}$$

Notice that both Q_1 and Q_{10} are dimensionless numbers > 1. The exponential expression in which Q_{10} is raised to power $\Delta T/10$ is known as the Q_{10} model [64].

The exponential increase of the process rate with temperature is a general chemical behavior, also extending to biochemical reactions. This is not a hypothesis or speculation but a proven fact that is widely used in engineering. For example, the metabolic rate in wastewater and sewer systems is expressed by the so-called effective BOD (EBOD, with BOD standing for biochemical oxygen demand). It has been known for more than 75 years that the metabolic rate increases with temperature, as microorganism activity generally increases with temperature. The relationship of EBOD with temperature has been expressed by Pomeroy and Bowlus [36] as $[\mathrm{EBOD}] = [\mathrm{BOD}]\,(1.07)^{T-20}$, which is similar to (A4), where the reference temperature is $T_0 = 20$ °C and $Q_1 = 1.07$ ($Q_{10} = 2.0$). The latter relationship is the standard of engineering design in sewer systems.

To apply this framework to find the increase of respiration in the last 65-year period investigated in our study, we first retrieved the global temperature separately for land and sea from the NCEP/NCAR Reanalysis data set. These are shown on an annual timescale in Figure A2. The resulting linear trends, also shown in Figure A2, yield an increase $\Delta T = 1.69$ °C for the terrestrial and 0.78 °C for the maritime part for the 65-year period.



[Figure R2-5.5; reproduced from Koutsoyiannis et al. [3]; original caption follows]
Figure A2. Evolution of global land (terrestrial) and sea (maritime) temperature at 2 m from the NCEP/NCAR Reanalysis data set, retrieved from the ClimExp platform, and resulting slopes of linear trends.

Now the literature gives representative average Q_{10} values of 3.05 for terrestrial respiration [64] and 4.07 for maritime respiration [65]. If R_B and R_E denote the respiration rate at the beginning and the end of the 65-year period, and $\Delta R \coloneqq R_\mathrm{E} - R_\mathrm{B}$, then according to (A4),

$$\frac{R_\mathrm{E}}{R_\mathrm{B}} = Q_{10}^{\Delta T/10} \tag{A6}$$

and hence

$$\Delta R = R_\mathrm{E}\left(1 - \frac{1}{Q_{10}^{\Delta T/10}}\right) \tag{A7}$$

For the above given values of Q_{10} and ΔT, the expression in parentheses becomes 0.172 for the terrestrial part and 0.104 for the maritime part. Multiplying these by the R_E values shown in Figure A1, i.e., 136.7 and 77.6 Gt C/year, respectively, we find $\Delta R = 23.5$ and 8.1 Gt C/year, respectively, i.e., a total global increase in the respiration rate of $\Delta R = 31.6$ Gt C/year. This rate, which is a result of natural processes, is 3.4 times greater than the CO_2 emission by fossil fuel combustion (9.4 Gt C /year including cement production).

R2-5.4. Returning to the abstract:

> "Its application to instrumental measurements of temperature (T) and atmospheric carbon dioxide concentration ([CO_2]) over the last seven decades provided evidence for a unidirectional, potentially causal link between T as the cause and [CO_2] as the effect. "

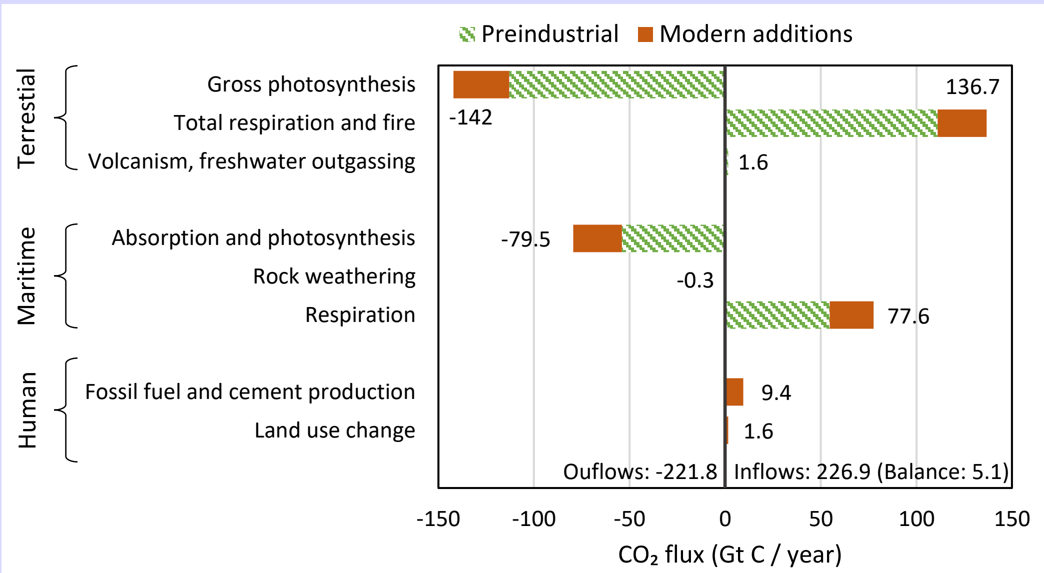
This is easily shown to be an incorrect inference - it is directly refuted by the Earth's "carbon budget" data (see e.g. [[1]]).

My statement is 100% correct. It follows the statement "*As a* ***result of recent research****, a new stochastic methodology of assessing causality was developed.*" And this stochastic methodology, described in Koutsoyiannis et al. [1-3] gave exactly the results summarized in the next statement, quoted by the reviewer.

Furthermore, these findings are 100% consistent with Earth's carbon budget data. The reviewer recommends citing the reference to Cawley (2011) [66]. However, we have used the newer data of the most recent (2021) IPCC's Assessment Report (AR6) [32]. And our findings are fully consistent with both the IPCC and Cawley data.

Most probably, the reviewer has not seen the part of the paper Koutsoyiannis et al. [3], where we use the IPCC carbon budget data. Hence, I am reproducing this part from Appendix A.1 here:

> Here we follow the IPCC's [32] account in its recent (2021) Assessment Report (AR6). Its schematic (Figure 5.12 in that Report) does not hide (a) the imbalances in the different parts of Earth and (b) the fact that the natural carbon inputs and outputs in the atmosphere change over time—even though the IPCC's schematic implicitly assumes that "natural" is the budget that occurred in the preindustrial age (1750) and that any change that occurred since is anthropogenic. Interestingly, in an alternative view by Hansen et al. [67], civilization always produced greenhouse gases and aerosols, and humans likely contributed to the increase of both in the past 6000 years, thus resulting in climate forcings.
>
> Based on the IPCC's representation, we have summarized in Figure A1 the information given in IPCC's schematic, in terms of annual carbon balance. When seen in the entire picture, the human emissions due to fossil fuel combustion (9.4 Gt C/year including cement production) is a small part (4%) of the total CO_2 inflows to the atmosphere.



[Figure R2-5.6; reproduced from Koutsoyiannis et al. [3]; original caption follows]
Figure A1. Annual carbon balance in the Earth's atmosphere, in Gt C/year, based on the IPCC [32] estimates. The balance of 5.1 Gt C/year is the annual accumulation of carbon (in the form of CO_2) in the atmosphere.

As per the graph contained in Cawley (2011) [66], namely its Figure 2, which the reviewer reproduces below, I show in my reply to her/his next comment (R2-5.5) that it is fully consistent with our findings.

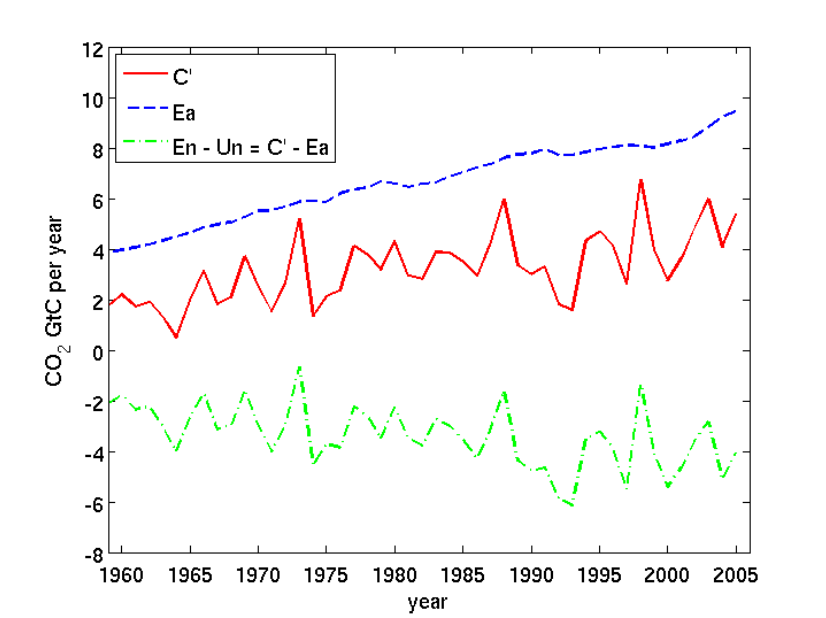
R2-5.5. Briefly stated, the carbon cycle obeys conservation of mass, so the annual change in atmospheric CO2, C', is equal to the difference in total emissions into the atmosphere, and total update from the atmosphere, i.e.

C' = Ea + En - Un

Where Ea is total emissions from anthropogenic sources, En is total emissions from all natural sources and Un it total uptake from the atmosphere by all natural sinks (note there is no non-negligible anthropogenic uptake, there is no significant amount of carbon capture and storage currently in action). Note that natural sources do not differentiate between CO2 from natural and anthropogenic, so Un is a combination of both in proportion to their atmospheric mixing ratio.
Via basic algebra, we obtain

C' - Ea = En - Un

In other words the difference between natural emissions and uptake by natural sinks (i.e. the net action of the natural carbon cycle) is equal to the difference between the annual rise in CO2 and the level of anthropogenic emissions, both of which are reliably known - C' from e.g. Mauna Loa Observatory (MLO) data and Ea from givenrment/commercial records.



[Figure R2-5.7 (by Reviewer #5 without a caption]
If we look at the carbon budget data, we see that every year since 1960 (when the MLO data became available), the annual change in atmospheric CO2 has been less (about half) the level of anthropogenic emissions, which means that the natural environment must necessarily have been a net carbon sink. The natural environment has been **opposing** the increase, and we can see that this opposition has been increasing with time.

This rules out a change in global temperature being the cause of the rise in atmospheric CO2 through some response of the natural carbon cycle.

This comment is most interesting and I will reply to it in full detail. Generally, while Reviewer's #5 comments do not seem to be plagiarized in the sense recently revealed and studied by Piniewski et al. [68], this particular comment is a paraphrasis of Cawley's [69] comment in the pubpeer site, created by Rice [70]. The notations, equations and the figure are precisely the same (without paraphrasis) as in Cawley's [69] comment.

Clearly then, this comment, which was posted in 2022, refers to the paper Koutsoyiannis et al. (2022) [2]. The reason why Rice [70] and Cawley [69] created the pubpeer site and posted their comments are revealed by Rice [71] in his blog. He says:

> I thought I might simply highlight that I started a PubPeer thread about this paper and Gavin Cawley has already posted a couple of useful comments. A PubPeer thread about the Zharkova et al. paper produced quite an extensive comment thread and probably played a role in it being retracted. You might argue that a paper shouldn't be retracted just because one of the case studies produces results that are almost certainly wrong. On the other hand, you might also argue that if one of their case studies produces such results that it rather undermines their whole method. You might also argue that it's rather embarassing that one of the Royal Society's journals could publish a paper with what is, these days, a very obviously wrong result.

In other words, these two gentlemen are the supreme judges of what is scientifically wrong and right—as well as able to dictate to scientific organizations, including the Royal Society, what they are allowed to publish and what they should retract.

In a later post, Rice [72] uncovers two things: (a) they emailed to the Royal Society to retract our papers, and (b) that the Royal Society responded negatively (how dared they!). Specifically, Rice [72] writes:

> When the Proceedings of the Royal Society A paper came out last year, I emailed the editor to point out that they'd published a paper with a rather nonsensical result. I didn't get a response. However, I was cc'd into a response to someone else who had also complained. This response was, unfortunately, rather dismissive and somewhat insulting. The response said that the criticism had been discussed with the board member and subject editor who handled the paper. According to them, the criticism misinterpreted what was in the paper and was not well-founded. Apparently, the result was scientifically intriguing and would be of interest to many of their readers.

While my coauthors and I are generally responsive to comments and criticisms by anyone interested (and we have stated in our papers that we welcome dialogue), we had decided not to respond to the comments of this group in pubpeer and in their blog for two reasons: (a) because their purpose was not the scientific dialogue, but our silencing and (b) because of the low level of their criticism.

I am very much disappointed that, because of a surprising editorial decision to invite additional reviewers, I am forced now, two years after, to reply to this old comment, referring to a paper of 2022. And this I must do in the framework of the peer review of my new paper in 2024!

So here is my reply, which, however, I do not include in the present paper, because it is totally out of its scope. It refers to my earlier papers. My reply, in full detail, to this particular comment is only part of this report.

The equations and the graph copied from Cawley [66,69] by the reviewer, as well as the interpretation given is totally irrelevant to causality. A necessary condition to assess causality is time precedence and there is nothing in the Reviewer's #5 and Cawley [69] comments related to time lead or lag.

The equations and the graph are fully consistent with our approach in Koutsoyiannis et al. [1-3], and with their finding that the temperature change leads and the [CO_2] change lags, which excludes the possibility that [CO_2] change is a cause of temperature change.

Initially I include the following quotation from Koutsoyiannis et al. [3], Section 9, which contains a simple and brief model of the covariation of the two variables:

> As already clarified, the scope of our work is not to provide detailed modeling of the processes studied but to check causality conditions. We highlight the fact that the relationship we established explains only about 1/3 of the actual variance of $\Delta\ln[\mathrm{CO_2}]$. This is not negligible for investigating causality, but also leaves a margin for many other climatic factors to act.
>
> Nonetheless, our results can certainly be improved if we change our scope to more detailed modeling. To illustrate this, we provide the following toy model. Based on our result that the T-[CO_2] system is potentially causal with direction $\Delta T \rightarrow \Delta\ln[\mathrm{CO_2}]$, we estimate $\Delta\ln[\mathrm{CO_2}]$ as
>
> $$\Delta\ln[\mathrm{CO_2}] = \sum_{j=0}^{20} g_j \Delta T_{\tau-j} + \mu_v \quad (8)$$
>
> and we proceed a step further, assuming that the mean μ_v also depends on past temperature, averaged at timescale m, i.e.,
>
> $$\mu_v = \alpha(T_m - T_0) \quad (9)$$
>
> where T_m is the average temperature of the previous m years, and α and T_0 are constants (parameters). Such a simple linear relationship is supported by the above-listed points related to the productivity of the biosphere. Equation (9) will result in negative values μ_v if $T_m < T_0$ and positive otherwise.
>
> By re-evaluating the IRF coordinates g_j simultaneously with the parameters of Equation (9), we find the modified version of the IRF plotted in Figure 14. The optimized additional parameters are $m = 4$ (years), $\alpha = 0.0034, T_0 = 285.84$ K. Similarly to [1], we used a common spreadsheet software solver to perform the optimization, adding the two parameters α and T_0 to the unknown coordinates g_j of the IRF and performing the (nonlinear) optimization for all unknowns (m was found by trial-and-error). A graphical

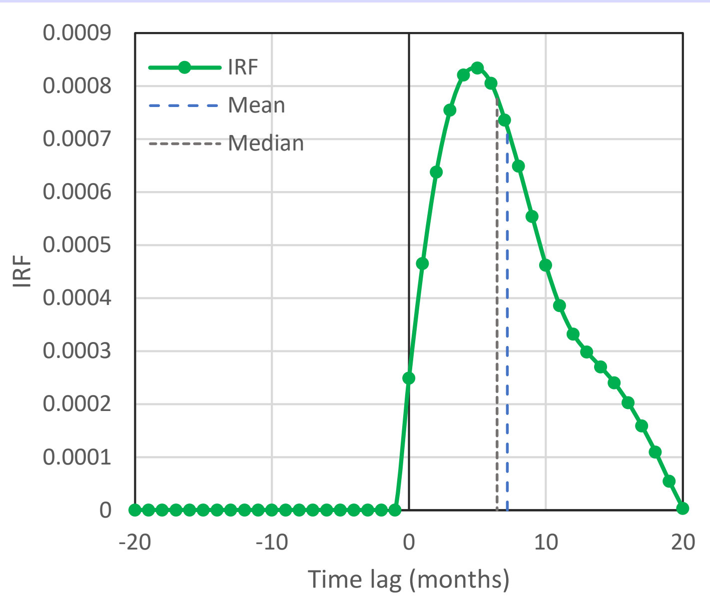
comparison of the actual $\Delta\ln[\mathrm{CO_2}]$ and $[\mathrm{CO_2}]$ with those simulated by the model of Equations (8) and (9) is given in Figure 15. The explained variance for $\Delta\ln[\mathrm{CO_2}]$ was drastically increased from 34% to 55.5% and that for $[\mathrm{CO_2}]$ is an impressive 99.9%.

For the convenience of the readers who are interested in repeating the calculations, we also give a parametric expression of IRF and summarize the toy model as:

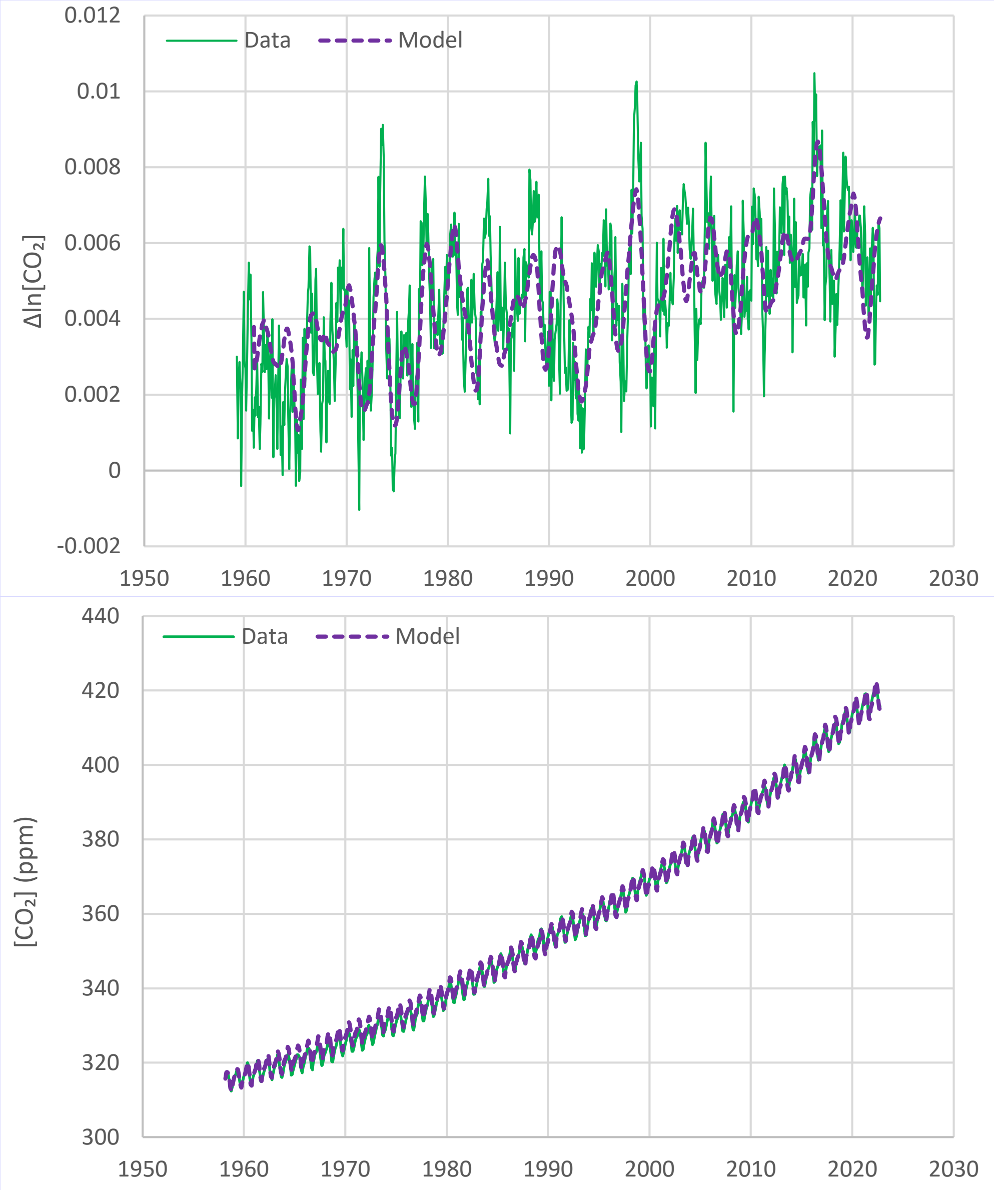
$$\Delta\ln[\mathrm{CO_2}] = \sum_{j=0}^{20} g_j \Delta T_{\tau-j} + \mu_v, \tag{10}$$
$$g_j = 0.00076\, j^{0.67} e^{-0.2j}/\mathrm{K}, \qquad \mu_v = 0.0034\,(T_4/\mathrm{K} - 285.84)$$

(where K is the unit of kelvin).

We emphasize, however, that we do not exploit the impressive result of explained variance of 99.9% to assert that we have built a decent model, even though this toy model is both accurate (in the lower panel of Figure 15, the simulated curve is indistinguishable from the actual) and parsimonious (the model expression in (10) contains only 5 fitted parameters). We prefer to highlight the fact that the hugely complex climate system entails high uncertainty, and its study needs reliable data that provide the basis for the modeling and testing of hypotheses.



[Figure R2-5.8; reproduced from Koutsoyiannis et al. [3]; original caption follows]
Figure 14. Modified IRF for temperature–CO_2 concentration based on the NCEP/NCAR Reanalysis temperature at 2 m and Mauna Loa time series, respectively, similar to Figure 2 but with IRF coordinates simultaneously optimized with the parameters of Equation (9).



[Figure R2-5.9; reproduced from Koutsoyiannis et al. [3]; original caption follows]
Figure 15. Comparison of the actual $\Delta\ln[CO_2]$ (**upper**) and $[CO_2]$ (**lower**) with those simulated by the model of Equations (8) and (9).

Now I use this toy model to reproduce Figure R2-5.7, i.e. Cawley's [66,69] and Reviewer's #5 graph. The resulting graph is seen in Figure R2-5.10 below. This is the original Figure R2-5.7 on which I overlay the results of our model in equation (10) in Koutsoyiannis et al. [3], also reproduced above. If the reviewer wants to test and reproduce it independently, here are the steps:

1. I digitized the "Ea" curve from Figure R2-5.7, i.e. Cawley's [66,69] and Reviewer's #5 graph.

2. I took the monthly global temperatures from the NCEP/NCAR reanalysis, as described in Koutsoyiannis et al. [3].

3. I applied the model of equation (10) to produce monthly values of $\Delta\ln[CO_2]$.

4. I aggregated the values of $\Delta\ln[CO_2]$ to find monthly values of $\ln[CO_2]$ and then exponentiated them to find monthly values of $[CO_2]$.

5. I aggregated from monthly scale to annual scale, and I took the annual differences of $[CO_2]$, i.e. C'.

6. I calculated the differences $C' - \mathrm{Ea}$

7. I plotted the annual series of C' and $C' - \mathrm{Ea}$.

It is seen that the model results, which are clearly based on the finding that temperature leads and $[CO_2]$ lags, is fully consistent with the Cawley's [66,69] and Reviewer's #5 graph.

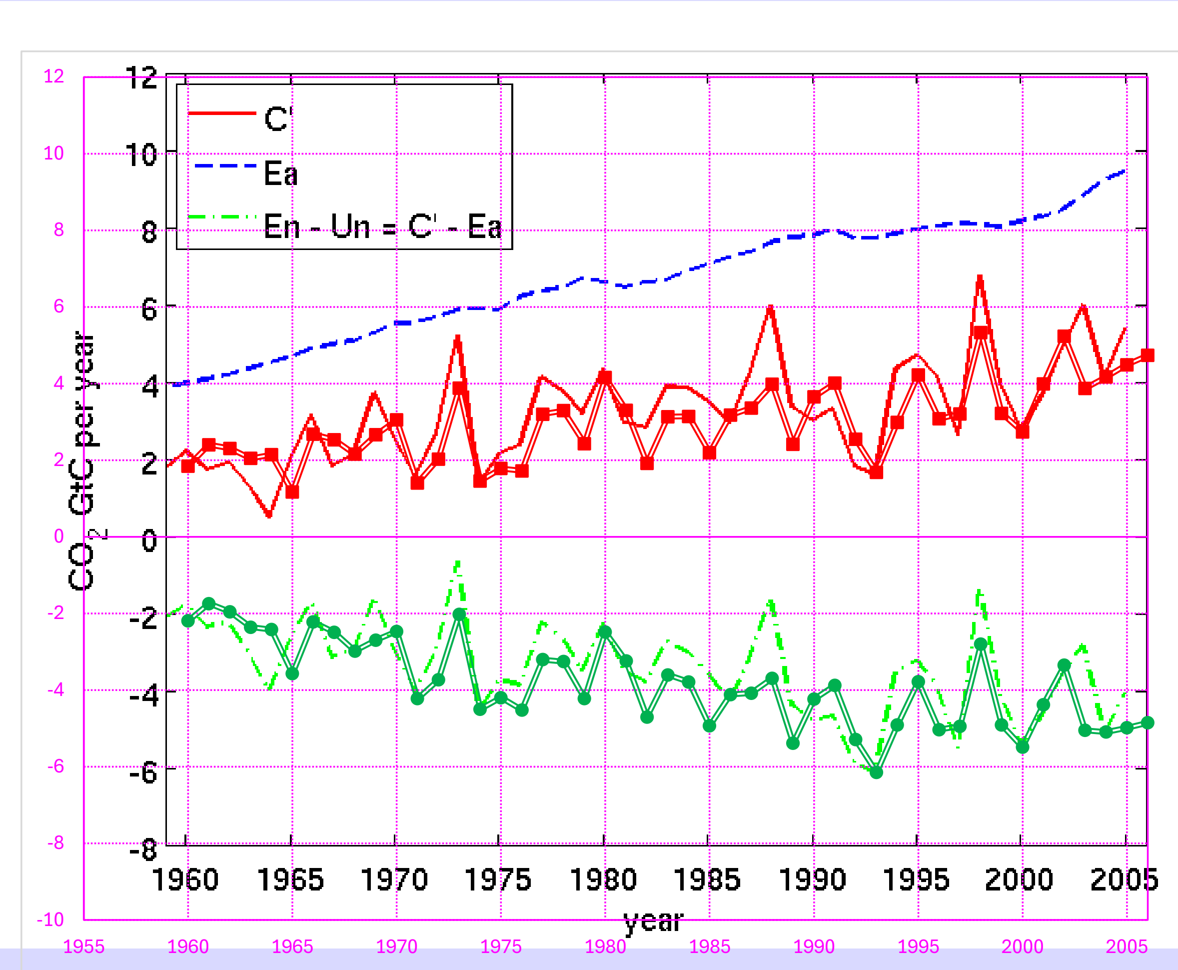


Figure R2-5.10 (new). Plots of the results of model of Equation (10), overlayed to the graph of Cawley [66,69] and Reviewer #5, seen in Figure R2-5.7.

One may notice that Cawley's [66,69] and Reviewer's #5 graph stops at year 2005. Since I took this effort, I expanded the graph up to year 2020, also using the most recent time series of emissions and of $[CO_2]$ as described in Koutsoyiannis et al. [3]. This is seen in Figure R2-5.11 (new). The model performance is excellent, explaining 81% of the annual variability.

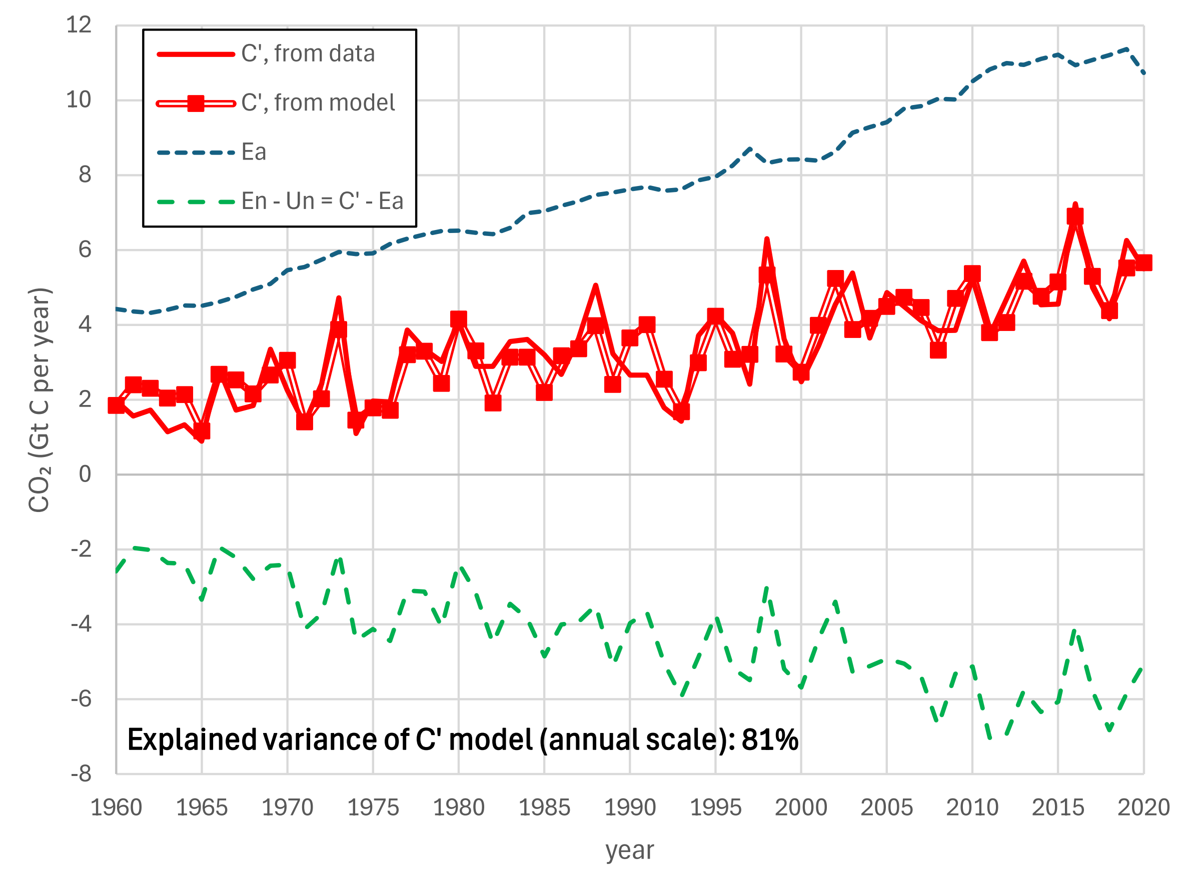


Figure R2-5.11 (new). Update of Cawley's [66,69] and Reviewer's #5 graph (seen in Figure R2-5.7) using the most recent CO_2 emission and concentration time series, and expansion up to year 2020. The results of model of Equation (10) are also shown.

As seen in Koutsoyiannis et al. [3], at the monthly scale the performance is even better, with an explained variance of 99.9%. Despite that, I wish to repeat again the following extract from Koutsoyiannis et al. [3]:

> We emphasize, however, that we do not exploit the impressive result of explained variance of 99.9% to assert that we have built a decent model […]. We prefer to highlight the fact that the hugely complex climate system entails high uncertainty, and its study needs reliable data that provide the basis for the modeling and testing of hypotheses.

After all this analysis, I wish to address an invitation to the reviewer to use my updated graph in Figure R2-5.11 for the next time she/he wants to present this stuff. That would indeed be preferrable because Figure R2-5.11 is updated and expanded, as well as for the additional reason that it also includes our model results, which agree very well with the observational data.

R2-5.6. This is not the only line of evidence that contradicts the finding of the current paper. There is also the timing of the increase in atmospheric CO2 matching the increase in anthropogenic emissions; the ratio of the atmospheric rise and anthropogenic emissions being approximately constant; declining 14C ratio; declining 13C ratio; declining; increasing oceanic CO2 etc. This is all well known - see e.g. IPCC WG1 reports.

Apparently, the reviewer is not aware of my recent publication [5] entitled "*Net isotopic signature of atmospheric CO_2 sources and sinks: No change since the Little Ice Age*". I am quoting here the abstract and the graphical abstract, and invite the reviewer to see the entire paper, which refutes her/his above claims.

> **Abstract** Recent studies have provided evidence, based on analyses of instrumental measurements of the last seven decades, for a unidirectional, potentially causal link between temperature as the cause and carbon dioxide concentration ([CO_2]) as the effect. In the most recent study, this finding was supported by analysing the carbon cycle and showing that the natural [CO_2] changes due to temperature rise are far larger (by a factor > 3) than human emissions, while the latter are no larger than 4% of the total. Here, we provide additional support for these findings by examining the signatures of the stable carbon isotopes, 12 and 13. Examining isotopic data in four important observation sites, we show that the standard metric $\delta^{13}C$ is consistent with an input isotopic signature that is stable over the entire period of observations (>40 years), i.e., not affected by increases in human CO_2 emissions. In addition, proxy data covering the period after 1500 AD also show stable behaviour. These findings confirm the major role of the biosphere in the carbon cycle and a non-discernible signature of humans.

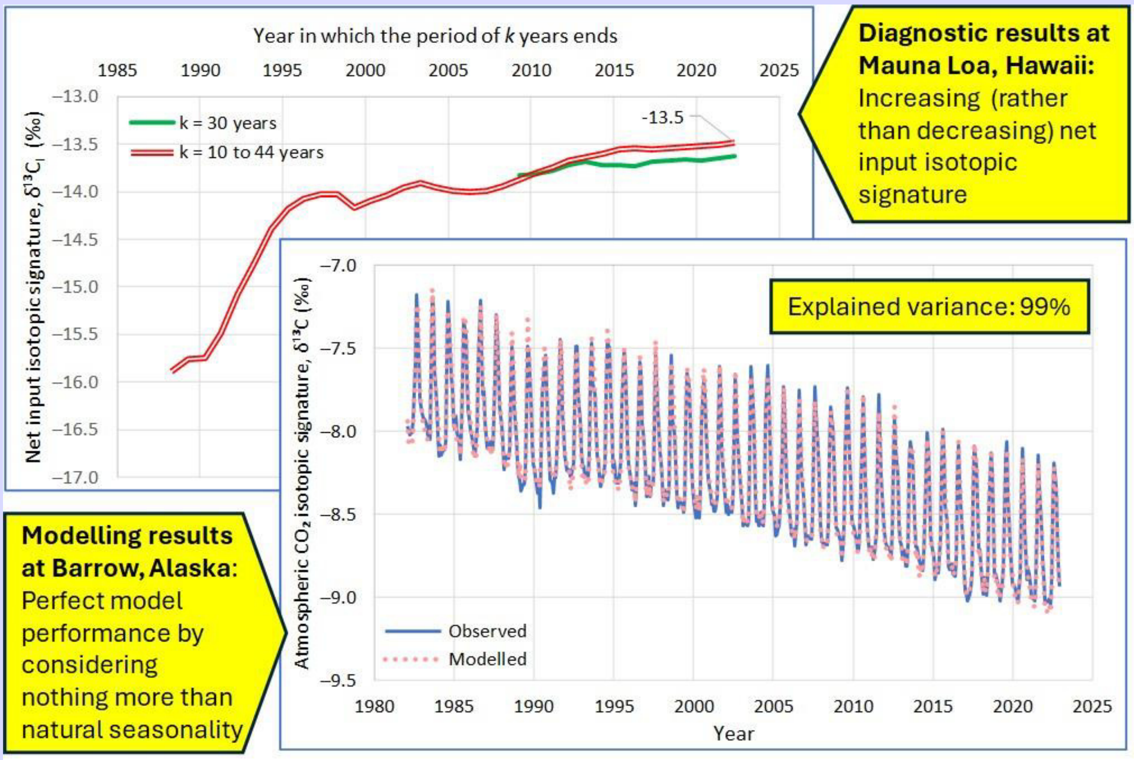


Figure R2-5.12. Graphical abstract of Koutsoyiannis [5].

R2-5.7. Returning to the abstract once more, the next claim is:

> "Several proxy series, extending over the entire Phanerozoic or parts of it, ... , are compiled, paired and analyzed. The extensive analyses made converge to the single inference that change in temperature leads, and that in carbon dioxide concertation lags."

This more moderate claim is essentially correct, but **entirely** uncontraversial. Temperatures leading and CO2 lagging does not imply that the causal relationship is exclusively from

temperature to CO2, because the same thing can result from a system in dynamic equilibrium with causal relationships in both directions. Throughout most of the Phanerozoic, the carbon cycle has mostly acted as a feedback mechansim, that on some timescales amplifies changes in temperature and on other timescales damping them down. For instance carbon cycle feedback amplified the very small changes in solar forcing from Milankovic cycles, giving rise to large temperature swings between glacial and interglacial periods in over the last 800,000 years seen in the Vostok ice core. The change in solar forcing is much too small to explain the temperature change in isolation. In this case, as the carbon cycle is acting as a positive feedback, it is not surprising that CO2 lags temperature (it takes time for thermal inetria of the oceans to be gradually overcome and so there is a delay in the ougassing) - see [[4]] for a basic explanation. On longer timescales, the "weathering thermostat" kept planetary temperature fairly stable - the chemical weathinging of silicate rocks is temperature dependent. If temperatures are high, this cases greated weathering and more CO2 is taken out of the atmosphere. This reduces the GHE, which tends to lower temperatures again, over many thousands of years, and the dynamic equilibrium is restored. Again, temperature leads and CO2 lags. None of this is obscure knowledge - it can easily be found in public understanding of science texts, such as Archer [[2]] or Volk [[3]]. As we know this from existing studies, informed by physics, a statistical model sheds no real light on the issue and does not justify publication.
However, on geological timescales, the carbon cycle can also act as a forcing. For example, the most plausible explanation for the emergence from the "Snowball Earth" conditions during the Cryogenian, Late Ordovician and Silurian is that the weathering thermostat was "switched off" by the ice coverage, which allowed GHG emissions from volcanic activity to gradually accumulate in the atmosphere, until the EGHE was sufficient for temperatures to rise to the point where the ice melted. In that case, CO2 led and temperatures lagged. Another example would be changes in the weathering thermostat due to changes in the position of the continents (see e.g. [[5]]) or the creation of mountain ranges (see e.g. [[6]]). In those cases, CO2 would also lead and temperatures lag because there the CO2 is acting as a forcing rather than a feedback. Again, this information has been discussed in the public debate on the science of climate.

The entire comment is irrelevant. The entire abstract, from which the reviewer quotes a couple of sentences is:

> As a result of recent research, a new stochastic methodology of assessing causality was developed. Its application to instrumental measurements of temperature (T) and atmospheric carbon dioxide concentration ([CO_2]) over the last seven decades provided evidence for a unidirectional, potentially causal link between T as the cause and [CO_2] as the effect. Here we refine and extend this methodology, and apply it to both paleoclimatic proxy data and instrumental data of T and [CO_2]. Several proxy series, extending over the entire Phanerozoic or parts of it, gradually improving in accuracy and temporal resolution up to the modern period of accurate records, are compiled, paired and analyzed. The extensive analyses made converge to the single inference that change in temperature leads, and that in carbon dioxide concentration lags. This conclusion is valid for both proxy and instrumental data in all time scales and time spans. The time scales examined start from annual and decadal for the modern period (instrumental data) and the last two millennia (proxy data), and reach one million years

> for the most sparse time series for the entire Phanerozoic. The type of causality appears to be unidirectional, $T \rightarrow [CO_2]$, as in earlier studies. The time lags found depend on the time span and time scale and are of the same order of magnitude as the latter. These results contradict the conventional wisdom, according to which the temperature rise is caused by $[CO_2]$ increase.

This abstract summarizes the content of the paper and is fully consistent with that content. It does not reflect the other studies which the reviewer likes and refers to. It reflects the analyses contained in my paper. In this respect, the sentences quoted by the reviewer are 100% accurate and no change is required.

> **R2-5.8.** In summary, the abstract contains nothing that is both correct and novel. The conclusion is naive, ignores a substantial amount of well established scientific work, reflecting a lack of adequate scholarship.

As I already stated, the abstract contains a summary of the paper. The analyses and results presented in the paper are correct and novel, as also recognized by the three original reviewers.

I did not invite the reviewer to assess whether my scholarship is adequate or not. Of course he has the right to do that without my invitation, but she/he may keep his assessment for her/him self.

> **R2-5.9. Introduction**
>
> Crime begins with propaganda, even if such propaganda is for a good cause Hans Fritzsche {1}
>
> This sort of rhetoric is simply unacceptable in a scientific paper and oonly undermines the authors work. It seems to imply that the mainstream scientific position on CO2 is (well intentioned) "propaganda" rather than the result of dispassionate scientific research, which is simply not the case.

I put this as a motto relevant to the paper. I could argue about its relevance, but the discussion would be too long and would distract the interest from the focus of the paper. Therefore, I preferred to delete the motto in the re-revised manuscript.

> **R2-5.10.** One of the most controversial issues of our time is the causal relationship between atmospheric temperature (T) and carbon dioxide (CO2) concentration ([CO2]) in Earth's climate.
>
> There is no genuine *scientific* controversy on the existence and mechanism of the EGHE, which has been well understood for over a century, nor on the basic science of the carbon cycle or the evidence that demonstrates that the rise in atmospheric CO2 is due to anthropogenic emissions (the IPCC WG1 reports have long had an explicit section on that issue). There is controversy in the public debate on climate, but it is largely due to misinformation promulgated on climate blogs and media. There are occasional failures of peer review that result in the publication of fundamentally flawed papers, but in the long run science is self-correcting and robust to these problems.

What is *misinformation*, *propaganda*, *censoring* and *silencing* is a tough issue to discuss and, again, the discussion would be a distraction from the focus of the paper. I invite the reviewer to see another paper of mine, [44], and in particular its section 6, where I discuss the political origin of the climate change agenda.

To address this comment, I changed the quoted phrase, which now reads:

> Some of the most controversial issues of our time are related to Earth's climate, not excluding the causal relationship between atmospheric temperature (*T*) and carbon dioxide (CO_2) concentration ([CO_2]).

R2-5.11. It is also accepted that the cosmic ray flux has a large effect on the climate and this flux had variations, including a cycle with a period of about 145 million years, corresponding to the passage of the solar system through one of the two sets of spiral arms of the Milky Way {16}.

While time constraints means that I can't check on all of the references cited, this is something that most certainly is not accepted by the scientific community (although it has been widely discussed on climate skeptic blogs). From the IPCC AR6 WG1 report (page 958)

> AR5 concluded that the GCR effect on CCN is too weak to have any detectable effect on climate and no robust association was found between GCR and cloudiness (Boucher et al., 2013). Published literature since AR5 robustly supports these conclusions with key laboratory, theoretical and observational evidence. There is high confidence that GCRs contribute a negligible ERF over the period 1750–2019.

The section makes reference to the CLOUD project at CERN that was set up to test the the proposed causal mechanism (an increase in cloud consensation nucleii) and found that the effect was very small. An effect on geological timescales has not been completely ruled out as far as I am aware, but there is very little beyond a statistical correlation as evidence.

I prefer not to include this information in the paper, as it is another political issue. We should not forget that IPCC is a political organization. Perhaps the reviewer is not aware of the developments related to the political aspects of the CLOUD project at CERN. Here I quote a part from an interview from 2011 of the then CERN's Director General Rolf-Dieter Heuer in the German newspaper Die Welt [73] using machine translation from German to English (my emphasis):

> **Heuer**: This is actually about understanding cloud formation better. There are many parameters in nature that influence this - including temperature, humidity, impurities and cosmic radiation. The "Cloud" experiment is about investigating the influence of cosmic radiation on cloud formation. The radiation used for this comes from the accelerator. And in an experimental chamber, it is possible to research under controlled conditions how droplet formation depends on radiation and suspended particles. The results will be published shortly. **I asked my colleagues to make the results clear but not to interpret them. This would immediately enter the highly political arena of the climate change discussion.** It must be clear that cosmic radiation is only one of many parameters.

What I write is consistent with the last sentence in the above quotation, that that cosmic radiation is only one of many parameters.

To address the reviewer's comment, in the revised manuscript and in the phrase quoted by the her/him, I have changed the verb "accepted" to "asserted" (which was actually what I meant but perhaps I made a typing error and the automatic speller changed it).

> **R2-5.12.** "According to the established narrative, which is simplistic and negligent of the huge complexity of the climatic system, ..."
>
> Again this is rhetoric unbecoming of a scientific paper and devalues the authors argument. It isn't an "established narrative", it is a century or more of diligent scientific research and is not simplistic (demonstrated by the page count of the IPCC report) not does it neglect complexity. I will not comment on further uses of rhetoric in this review, as it would be an egregious waste of my time as a reviewer. It is not to the authors advantage in communicating his findings. That ought to be sufficient.

I did not invite the reviewer to advise me regarding my style of writing. I am 69, I have published more than 250 journal articles and more than 1100 documents in total, including editorial notes and books and I feel I have the right to express myself according to the experience I developed through the years.

That said, in the phrase quoted by the reviewer, I changed "established" to "this".

> **R2-5.13.** It is concerning that the author cites one of their previous papers on this topic (reference 22 in the paper) but does not appear to mention the critical comment paper that was published explaining why the findings of that paper were incorrect. Unfortunately the same error is made in the current paper. Also cited is the paper by Humlum et al. (reference 34), which makes essentially the same mistake as this paper (differencing of the time series data deletes the contribution of anthropogenic emissions), but none of the papers that explained the error. This is not an acceptable practice.

I do not understand which error and which "critical comment paper that was published" the reviewer refers. There is no error. I do not understand why it is concerning to the reviewer that I cite my previous works. Should I repeat in the present work what I have already published?

See additional information in my reply to the next comment.

> **R2-5.14. Methodology**
>
> One important issue that should be kept in mind is related to the very high autocorrelations, which appear particularly in [CO 2] series. As high autocorrelation increases uncertainty in the long term, this is a major case leading to false identification of potential causality. This problem was discussed in {22} and illustrated in the electronic supplementary material of this publication, along with the technique to handle such situations and avoid false conclusions. Specifically, the appropriate technique is to difference the time series, so as to investigate the changes of the related processes, rather than the processes per se. Differencing reduces the autocorrelations substantially and thus avoids spurious results but has the disadvantage of reducing the explained variance.

This, as pointed out by Asbrink [[7]] and in the pubpeer discussion of the Author's previous work [[8]], is a fundamental error. Differencing the time series causes the long term linear trend component of the time series to become an additive constant in the differenced time series. If the method of analysis is insensitve to additive constants, then there is no mathematical link between the findings of the analysis and the long term increase. This is problematic in the case of the carbon cycle because the cause of the long term increase in both series (fossil fuel emissions and the EGHE) is not the same as the causes of the short term variability around the trend namely seasonality and the effects of ENSO (as noted by Bacastow, see [[7]]). Seasonality can be dealt with via the running mean used in the paper, or use of anomalies, which would be a more usual approach. In the case of ENSO, it affects temperature directly, through changes in ocean surface temperature, but the effect on CO2 is not via temperature, but via changes in precipitation leading to variation in growth and decay in the terrestrial biosphere. It is a confounding variable that is missing from the analysis. Unfortunately, this means the analysis will incorrectly attribute the cause of the long term change to the cause of the short term variability, which is essentially unrelated. This is the same error made by Humlum et al., as well as several others, including Murry Salby and the error has been repeatedly pointed out. This primarily affects the instrumental data, but similar problems may also be present in the paleoclimate studies.

First off, we should distinguish Åsbrink's [74] Commentary from the "*pubpeer discussion*". The former is a scientific work, formally published in a journal. The latter is a failed action to force retraction of our paper. As the "*pubpeer discussion*" is already discussed above (comment R2-5.5), I am replying here to the part of the reviewer's comment that is related to Åsbrink's Comment. I admit that it was my omission not to cite Åsbrink, which I now corrected.

Åsbrink did not challenge our methodology nor our results. His comment does not contain anything related to our mathematical part. He simply expressed disagreements with some of our formulations and interpretations. These disagreements allowed us to confirm our results in the original papers [1,2] and to shed further light on why our original formulations stand. As our new analyses exceeded the length of a typical reply to a Commentary, we published them in a stand-alone new paper [3]. In this, we responded to all Åsbrink's disagreements and we further developed our method. We let both the Royal Society and Åsbrink know that they can find our complete set of replies to our newer paper [3]. This triggered a constructive email exchange with Åsbrink.

In particular, with respect to Åsbrink claim that

> *'Common perception' is different for different timescales. It is a well-known fact that the fluctuations of [CO_2] in a timescale of a couple of years are caused by temperature variations.*

and

> *Hence, the common perception that increasing [CO_2] causes increased T seems likely.*

in our paper [3] we proved that this common perception cannot stand. Already in our paper [2], our Figure 15, whose time lags span 200 months (≈ 17 years) confirms the validity of our results for timescales larger than decadal. In addition to what is reported in our paper [2], in [3] we tested larger timescales, using longer temperature series along with the Mauna Loa [CO_2] measurements, which began in 1958. These analyses gave essentially the same results as in the case study presented in [2], suggesting a potentially causal system with the directionality being $\Delta T \rightarrow \Delta\ln[CO_2]$, with even better characteristics (higher cross correlation and explained variance, and time lags greater than or equal to those in [2]). In addition, we increased the time interval of differencing from one year to more than a decade. The results were again essentially the same. As per the Åsbrink's suggestion that "*For the annual cycle, one should actually look at the two hemispheres separately*", in our paper [3] we conducted an additional analysis with the South Pole $[CO_2]$ measurements, with the results being very similar: potentially causal system with causality direction $\Delta T \rightarrow \Delta\ln[CO_2]$, and lags close to 10 months. In all our case studies the possible causality $\Delta\ln[CO_2] \rightarrow \Delta T$ has been excluded as not satisfying the necessary condition of time precedence.

Concerning the reviewer's remark about the effect that differencing the time series may have on the long-term trend, our analysis in Section 9 of our paper [3] shows that clearly this is not the case. My reproduction of the reviewer's graph in her/his comment R2-5.5 clearly confirms that after aggregating the differenced results, our model, fully based on causality direction $T \rightarrow [CO_2]$, has even better performance than in the differenced case: explained variance 81% at annual scale and 99.9% on monthly scale.

Other issues mentioned by the reviewer such as ENSO and the role of the biosphere are fully analyzed in our paper [3]—see Section 9 and Appendices A.1 and A.3.

R2-5.15. Conclusions

> Is there a mechanism that, at the recent period has, due to human or presumed 'unnatural' actions, reversed directionality, as the popular claim is? Perhaps, but no analysis based on observational data has shown that.

Yes. Extracting fossil carbon from the lithosphere and injecting it into the atmosphere is a mechansim that will increase atmospheric CO2; the extended greenhouse effect means that this will result in an increase in global mean surface temperature (all things being otherwise equal. The carbon budget/mass balance analysis given earlier (see also [[1]]) establishes the former; a variety of evidence confirms the latter, e.g. [[9]] but see also the IPCC WG1 report

I have changed the quoted phrase, which now reads:

> Did human actions, such as fossil fuel combustion and other presumed 'unnatural' actions, reverse directionality, as the popular claim is?

R2-5.16. Rather, such claims are based on imagination and climatic models full of assumptions.

All models are full of assumptions, including the statistical models used in the paper. Again, this is rhetoric that does not belong in a balanced, objective scientific paper.

Indeed, models are based on assumptions, but not all models are full of assumptions. The stochastic methodology on which the paper is based is extraordinarily parsimonious in terms of assumptions, as seen in Section 3.2.

I have many papers showing the poor performance of climate models in representing reality, such as [75-80], while, most importantly, our recent paper, Koutsoyiannis et. al. [3] shows that the causality direction in climate models, identified by the same methodology, is opposite to that identified from real-world data. However, I prefer not to expand the paper about this, as it would be a distraction.

In terms of my style of writing (the "*rhetoric*" according to the reviewer), please see my reply to comment R2-5.12.

R2-5.17 References [in Reviewer's #5 review]

[[1]] Gavin C. Cawley, On the atmospheric residence time of anthropogenically sourced carbon dioxide, *Energy & Fuels*, volume 25, number 11, pages 5503–5513, September 2011.
[[2]] David Archer, "The Global Carbon Cycle", Princeton Primers in Climate, 2010. ISBN: 978-0691144146
[[3]] Tyler Volk, "CO2 Rising", MIT Press, 2008. ISBN: 978-0262515214
[[4]] https://www.carbonbrief.org/explainer-how-the-rise-and-fall-of-co2-levels-influenced-the-ice-ages/
[[5]] Brune, S., Williams, S.E. & Müller, R.D. Potential links between continental rifting, CO2 degassing and climate change through time. *Nature Geosci* **10**, 941–946 (2017). https://doi.org/10.1038/s41561-017-0003-6
[[6]] Raymo, M., Ruddiman, W. Tectonic forcing of late Cenozoic climate. *Nature* **359**, 117–122 (1992). https://doi.org/10.1038/359117a0
[[7]] Asbrink, Leif (2023), Revisiting causality using stochastics on atmospheric temperature and CO2 concentration, Proc. R. Soc. A.47920220529http://doi.org/10.1098/rspa.2022.0529
[[8]] https://pubpeer.com/publications/7828A34E1F905217D557E4F8E93CC1
[[9]] Kramer, R. J., He, H., Soden, B. J., Oreopoulos, L., Myhre, G., Forster, P. M., & Smith, C. J. (2021). Observational evidence of increasing global radiative forcing. *Geophysical Research Letters*, 48, e2020GL091585. https://doi.org/10.1029/2020GL091585

References [of Appendix B of this report]

1. D. Koutsoyiannis, C. Onof, A. Christofides, Z.W. Kundzewicz, Revisiting causality using stochastics: 1. Theory, *Proc. R. Soc. A*, 478 (2022), 20210836. doi: 10.1098/rspa.2021.0836
2. D. Koutsoyiannis, C. Onof, A. Christofides, Z. W. Kundzewicz, Revisiting causality using stochastics: 2. Applications, *Proc. R. Soc. A*, 478 (2022), 20210835. doi: 10.1098/rspa.2021.0835
3. D. Koutsoyiannis, C. Onof, Z. W. Kundzewicz, A. Christofides, On hens, eggs, temperatures and CO_2: Causal links in Earth's atmosphere, *Sci*, 5 (2023), 35. doi:10.3390/sci5030035
4. A. Christofides, D. Koutsoyiannis, C. Onof, and Z. W. Kundzewicz, Causality, Climate, Etc., , Climate Etc. (Judith Curry's blog), 2023. doi: 10.13140/RG.2.2.21608.44803
5. D. Koutsoyiannis, Net isotopic signature of atmospheric CO_2 sources and sinks: No change since the Little Ice Age, *Sci*, 6 (1) (2024), 17. doi:10.3390/sci6010017
6. Climate Etc. (Judith Curry's blog), Causality and climate, 2023, Available online, https://judithcurry.com/2023/09/26/causality-and-climate/.
7. C. W. Granger, Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, 37 (1969), 424-438.
8. C. W. Granger, Testing for causality: a personal viewpoint, *J. Econ. Dynamics and Control*, 2 (1980), 329-352.

9. J. Pearl, Causal inference in statistics: An overview, *Statistics Surveys*, 3 (2009), 96-146. doi: 10.1214/09-SS057
10. J. Pearl, M. Glymour, N.P. Jewell, *Causal Inference in Statistics: A Primer*, Wiley, Chichester, UK, 2016.
11. A. Hannart, J. Pearl, F .E. L. Otto, P. Naveau, M. Ghil, Causal counterfactual theory for the attribution of weather and climate-related events. *Bull. Amer. Met. Soc.*, 97 (2016), 99-110.
12. D. Koutsoyiannis, Z. W. Kundzewicz, Atmospheric temperature and CO_2: Hen-or-egg causality?, *Sci*, 2 (2020), 83. doi: 10.3390/sci2040083
13. C. W. J. Granger, Time series analysis, cointegration, and applications. *American Economic Review*, 94(3) (2004), 421-425.
14. T. Iliopoulou and D. Koutsoyiannis, Projecting the future of rainfall extremes: Better classic than trendy. *Journal of Hydrology*, 588 (2020), 125005. doi: 10.1016/j.jhydrol.2020.125005
15. H. G. Gauch Jr, *Scientific Method in Practice*. Cambridge University Press, 2003.
16. T. A. Cohn and H. F. Lins, Nature's style: naturally trendy. *Geophys. Res. Lett.*, 32 (2005).
17. D. Koutsoyiannis and A. Montanari, Statistical analysis of hydroclimatic time series: uncertainty and insights. *Water Resour. Res.* 43 (2007).
18. F. Serinaldi and C. G. Kilsby, Unsurprising surprises: the frequency of record-breaking and overthreshold hydrological extremes under spatial and temporal dependence. *Water Resour. Res.* 54 (2018), 6460–6487.
19. R. Nuzzo, Scientific method: statistical errors. *Nature News* 506 (2014), 150.
20. R. L. Wasserstein and N. A. Lazar, The ASA statement on p-values: context, process, and purpose. *Am. Statist.* 70 (2016), 129–133. doi: 10.1080/00031305.2016.1154108
21. V. Amrhein and S. Greenland, Remove, rather than redefine, statistical significance. *Nat. Hum. Behav.* 2 (2018) 4.
22. D. Trafimow, V. Amrhein, C. N. Areshenkoff, C. J. Barrera-Causil, E. J. Beh, Y. K. Bilgiç, R. Bono, M. T. Bradley, W. M. Briggs and H. A. Cepeda-Freyre, Manipulating the alpha level cannot cure significance testing. *Front. Psychol.* 9 (2018).
23. R. L. Wasserstein, A. L. Schirm and N. A. Lazar, Moving to a World beyond "$p < 0.05$". *The American Statistician*, 73(sup1) (2019), 1-19.
24. D. Koutsoyiannis, *Stochastics of Hydroclimatic Extremes - A Cool Look at Risk*, Edition 3, 2023, ISBN: 978-618-85370-0-2, 391 pages, Kallipos Open Academic Editions, Athens. doi: 10.57713/kallipos-1
25. D. Koutsoyiannis, Climate change, the Hurst phenomenon, and hydrological statistics, *Hydrological Sciences Journal*, 48 (1) (2003), 3–24, doi: 10.1623/hysj.48.1.3.43481
26. D. R. Cox Causality: Some statistical aspects. *J. Roy. Stat. Soc. A* 155 (1992), 291–301. doi: 10.2307/2982962
27. C. Zaiontz, Real Statistics Using Excel. http://www.realstatistics.com/; Accessed Sep. 2020.
28. C. Zaiontz, Real Statistics Examples Workbooks. http://www.real-statistics.com/free-download/real-statistics-examples-workbook/, Accessed Sep. 2020.
29. Koutsoyiannis, D., 2013. LTP: Looking Trendy—Persistently, *Climate Dialogue*, doi: 10.13140/RG.2.2.13070.36169.
30. D. Koutsoyiannis and N. Mamassis, From mythology to science: the development of scientific hydrological concepts in the Greek antiquity and its relevance to modern hydrology, *Hydrology and Earth System Sciences*, 25 (2021), 2419–2444. doi: 10.5194/hess-25-2419-2021

31. B. Bond-Lamberty and A. Thomson, Temperature-associated increases in the global soil respiration record. *Nature*, 464 (2010), 579.
32. IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, 1535 pp. http://www.climatechange2013.org/report/ (accessed 2020-02-14), 2013.
33. J. W. Raich and W. H. Schlesinger, The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. *Tellus B Chem. Phys. Meteorol.* 44 (1992), 81–99.
34. N. Makita, Y. Kosugi, A. Sakabe, A. Kanazawa, S. Ohkubo and M. Tani, Seasonal and diurnal patterns of soil respiration in an evergreen coniferous forest: Evidence from six years of observation with automatic chambers. *PLoS ONE* 13 (2018), e0192622, doi: 10.1371/journal.pone.0192622
35. C. ÓhAiseadha, G. Quinn, R. Connolly, M. Connolly and W. Soon, Energy and climate policy—an evaluation of global climate change expenditure 2011–2018. *Energies* 13 (2020), 4839.
36. R. Pomeroy and F. D. Bowlus, Progress report on sulfide control research. *Sewage Works Journal* 18 (4) (1946), 597-640.
37. C. Green and K. A. Byrne, Biomass: Impact on carbon cycle and greenhouse gas emissions. In *Encyclopedia of Energy*, Ed. by C. J. Cleveland, Elsevier, 223-236, 2004. doi: 10.1016/B0-12-176480-X/00418-6
38. R. Connolly, Review of "Atmospheric temperature and CO2: Hen-or-egg causality?" (Version 1). Sci 2020, https://www.mdpi.com/2413-4155/2/3/72 (posted and accessed on 09 October 2020)
39. R. Weiss, Carbon dioxide in water and seawater: the solubility of a non-ideal gas. *Marine Chemistry* 2 (3) (1974), 203-215.
40. Z. Zhu, S. Piao, R. B. Myneni, M. Huang, Z. Zeng, J. G. Canadell, P. Ciais, S. Sitch, P. Friedlingstein, A. Arneth et al. Greening of the earth and its drivers. *Nat. Clim. Chang.* 6 (2016), 791–795.
41. C. Chen, T. Park, X. Wang, S. Piao, B. Xu, R. K. Chaturvedi, R. Fuchs, V.Brovkin, P. Ciais, R.Fensholt, et al. China and India lead in greening of the world through land-use management. *Nat. Sustain.* 2 (2019), 122–129.
42. J. Liu, K. W. Bowman, D. S. Schimel, N. C. Parazoo, Z. Jiang, M. Lee, A. A. Bloom, D. Wunch, C. Frankenberg, Y. Sun and C. W. O'Dell, Contrasting carbon cycle responses of the tropical continents to the 2015–2016 El Niño. *Science* 358(6360) (2017), doi: 10.1126/science.aam5690.
43. P. Goulet Coulombe and M. Göbel, On spurious causality, CO2, and global temperature. *Econometrics* 9(3) (2021), 33. doi: 10.3390/econometrics9030033
44. D. Koutsoyiannis, Rethinking climate, climate change, and their relationship with water. *Water* 13(6) (2021), 849. doi:10.3390/w13060849
45. IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2391 pp. doi:10.1017/9781009157896.

46. Y. Li, Z. L. Li, H. Wu, C. Zhou, X. Liu, P. Leng, P. Yang, W. Wu, R. Tang, G. F. Shang and L. Ma, Biophysical impacts of earth greening can substantially mitigate regional land surface temperature warming. *Nat. Commun.* 14 (2023), 121. doi: 10.1038/s41467-023-35799-4
47. M. Milanković, , *Nebeska Mehanika*, Beograd, 1935.
48. M. Milanković, *Kanon der Erdbestrahlung und seine Anwendung auf das Eiszeitenproblem*, Koniglich Serbische Akademice, Beograd, 1941.
49. M. Milanković, *Canon of Insolation and the Ice-Age Problem*, Agency for Textbooks, Belgrade, 1998.
50. G. Roe, In defense of Milankovitch. *Geophys. Res. Lett.* 33 (2006). doi: 10.1029/2006GL027817
51. Y. Markonis and D. Koutsoyiannis, Climatic variability over time scales spanning nine orders of magnitude: Connecting Milankovitch cycles with Hurst–Kolmogorov dynamics. *Surveys in Geophysics* 34 (2) (2013), 181–207. doi: 10.1007/s10712-012-9208-9
52. G. L. Stephens, M. Z. Hakuba, S. Kato, A. Gettelman, J.-L. Dufresne, T. Andrews, J. N. S. Cole, U. Willen and T. Mauritsen, The changing nature of Earth's reflected sunlight. *Proc. R. Soc. A.* (2022), 4782022005320220053. doi: 10.1098/rspa.2022.0053
53. R. Connolly, W. Soon, M. Connolly, S. Baliunas, J. Berglund, C. J. Butler, R. G. Cionco, A. G. Elias, V. M. Fedorov, H. Harde, G. W. Henry, et al., How much has the Sun influenced Northern Hemisphere temperature trends? An ongoing debate. *Research in Astronomy and Astrophysics*, 21(6) (2021), 131, 1-68. doi: 10.1088/1674-4527/21/6/131
54. N. Scafetta and A. Bianchini, The planetary theory of solar activity variability: A review. *Front. Astron. Space Sci.* 9 (2022), 937930. doi: 10.3389/fspas.2022.937930
55. J. E. Kamis, *The Plate Climatology Theory: How Geological Forces Influence, Alter, or Control Earth's Climate and Climate Related Events*. https://books.google.gr/books/?id=7lRqzgEACAAJ (accessed 10 March 2023).
56. D. Chakrabarty, *The Climate of History in a Planetary Age*. University of Chicago Press, 2021. https://books.google.gr/books?id=ETQXEAAAQBAJ (accessed 10 March 2023).
57. E. Davis, K. Becker, R. Dziak, J. Cassidy, K. Wang and M. Lilley, Hydrological response to a seafloor spreading episode on the Juan de Fuca ridge. *Nature* 430 (6997) (2004), 335-338.
58. L. S. Urakawa, and H. Hasumi, A remote effect of geothermal heat on the global thermohaline circulation. *Journal of Geophysical Research: Oceans*, 114 (2009), C07016. doi: 10.1029/2008JC005192
59. L. Patara and C. W. Böning, Abyssal ocean warming around Antarctica strengthens the Atlantic overturning circulation. *Geophysical Research Letters*, 41 (11) (2014), 3972-3978.
60. D. Koutsoyiannis, A random walk on water, *Hydrology and Earth System Sciences*, 14 (2010), 585–601. doi:10.5194/hess-14-585-2010
61. D. Koutsoyiannis, Hydrology and change, *Hydrological Sciences Journal*, 58 (6) (2013), 1177–1197. doi:10.1080/02626667.2013.804626
62. V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, et al. (Eds.) IPCC, Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2021; 2391p. doi: 10.1017/9781009157896

63. S. A. Arrhenius, Über die Dissociationswärme und den Einfluß der Temperatur auf den Dissociationsgrad der Elektrolyte. Z. *Phys. Chem.* 4 (1889), 96–116. doi:10.1515/zpch-1889-0408
64. K. F. Patel, B. Bond-Lamberty, J. Jian, K. A. Morris, S. A. McKever, C. G. Norris, J. Zheng and V. L. Bailey, Carbon flux estimates are sensitive to data source: a comparison of field and lab temperature sensitivity data. *Environmental Research Letters* 17 (11) (2022), 113003.
65. C. Robinson, Microbial respiration, the engine of ocean deoxygenation. *Frontiers in Marine Science* 5 (2019), 533.
66. G. C. Cawley, On the atmospheric residence time of anthropogenically sourced carbon dioxide. *Energy and Fuels* 25 (11) (2011), 5503-5513.
67. J. E. Hansen, M. Sato, L. Simons, L. S. Nazarenko, K. von Schuckmann, N. G. Loeb, M. B. Osman, P. Kharecha, Q. Jin, G. Tselioudis and A. Lacis, Global warming in the pipeline. arXiv preprint, arXiv:2212.04474, 2022.
68. M. Piniewski, I. Jarić, D. Koutsoyiannis, and Z. W. Kundzewicz, Emerging plagiarism in peer-review evaluation reports: a tip of the iceberg?, *Scientometrics* (2024), doi:10.1007/s11192-024-04960-1.
69. G. C. Cawley, comment #3 in pubpeer site on "Revisiting causality using stochastics: 2. Applications" (2022), https://pubpeer.com/publications/7828A34E1F905217D557E4F8E93CC1#
70. K. Rice, comment #1 in pubpeer site on "Revisiting causality using stochastics: 2. Applications" (2022), https://pubpeer.com/publications/7828A34E1F905217D557E4F8E93CC1#3
71. …and Then There's Physics (blog by K. Rice), Revisiting causality using stochastics (2022), https://andthentheresphysics.wordpress.com/2022/07/27/revisiting-causality-using-stochastics/
72. …and Then There's Physics (blog by K. Rice), Scientifically intriguing? (2023), https://andthentheresphysics.wordpress.com/2023/10/01/scientifically-intriguing/
73. WELT, Antimaterie: Wie "Illuminati" den Cern-Forschern geholfen hat (2011), https://www.welt.de/wissenschaft/article13488331/Wie-Illuminati-den-Cern-Forschern-geholfen-hat.html
74. L. Åsbrink, Revisiting causality using stochastics on atmospheric temperature and CO2 concentration. *Proceedings of the Royal Society A*, 479(2269) (2023), 20220529.
75. D. Koutsoyiannis, A. Efstratiadis, and K. Georgakakos, Uncertainty assessment of future hydroclimatic predictions: A comparison of probabilistic and scenario-based approaches, *Journal of Hydrometeorology*, 8 (3) (2007), 261–281. doi:10.1175/JHM576.1
76. D. Koutsoyiannis, A. Efstratiadis, N. Mamassis, and A. Christofides, On the credibility of climate predictions, *Hydrological Sciences Journal*, 53 (4) (2008), 671–684 doi: 10.1623/hysj.53.4.671
77. G. G. Anagnostopoulos, D. Koutsoyiannis, A. Christofides, A. Efstratiadis, and N. Mamassis, A comparison of local and aggregated climate model outputs with observed data, *Hydrological Sciences Journal*, 55 (7) (2010), 1094–1110. doi: 10.1080/02626667.2010.513518
78. D. Koutsoyiannis, A. Christofides, A. Efstratiadis, G. G. Anagnostopoulos, and N. Mamassis, Scientific dialogue on climate: is it giving black eyes or opening closed eyes? Reply to "A black eye for the Hydrological Sciences Journal" by D. Huard, *Hydrological Sciences Journal*, 56 (7) (2011), 1334–1339. doi: 10.1080/02626667.2011.610759

79. H. Tyralis, and D. Koutsoyiannis, On the prediction of persistent processes using the output of deterministic models, *Hydrological Sciences Journal*, 62 (13) (2017), 2083–2102/ doi: 10.1080/02626667.2017.1361535
80. D. Koutsoyiannis, Revisiting the global hydrological cycle: is it intensifying?, *Hydrology and Earth System Sciences* 24 (2020), 3899–3932, doi: 10.5194/hess-24-3899-2020

Sci

MDPI

Article

Atmospheric Temperature and CO_2: Hen-Or-Egg Causality?

Demetris Koutsoyiannis [1,*] and Zbigniew W. Kundzewicz [2]

1 Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, 157 80 Athens, Greece
2 Institute for Agricultural and Forest Environment, Polish Academy of Sciences, 60-809 Poznań, Poland; kundzewicz@yahoo.com
* Correspondence: dk@itia.ntua.gr

Received: 7 September 2020; Accepted: 16 November 2020; Published: 25 November 2020

check for updates

Abstract: It is common knowledge that increasing CO_2 concentration plays a major role in enhancement of the greenhouse effect and contributes to global warming. The purpose of this study is to complement the conventional and established theory, that increased CO_2 concentration due to human emissions causes an increase in temperature, by considering the reverse causality. Since increased temperature causes an increase in CO_2 concentration, the relationship of atmospheric CO_2 and temperature may qualify as belonging to the category of "hen-or-egg" problems, where it is not always clear which of two interrelated events is the cause and which the effect. We examine the relationship of global temperature and atmospheric carbon dioxide concentration in monthly time steps, covering the time interval 1980–2019 during which reliable instrumental measurements are available. While both causality directions exist, the results of our study support the hypothesis that the dominant direction is $T \rightarrow CO_2$. Changes in CO_2 follow changes in T by about six months on a monthly scale, or about one year on an annual scale. We attempt to interpret this mechanism by involving biochemical reactions as at higher temperatures, soil respiration and, hence, CO_2 emissions, are increasing.

Keywords: temperature; global warming; greenhouse gases; atmospheric CO_2 concentration

---

Πότερον ἡ ὄρνις πρότερον ἢ τὸ ᾠὸν ἐγένετο (Which of the two came first, the hen or the egg?).

Πλούταρχος, Ηθικά, Συμποσιακὰ Β, Πρόβλημα Γ (Plutarch, Moralia, Quaestiones convivales, B, Question III).

1. Introduction

The phrase "hen-or-egg" is a metaphor describing situations where it is not clear which of two interrelated events or processes is the cause and which the effect. Plutarch was the first to pose this type of causality as a philosophical problem using the example of the hen and the egg, as indicated in the motto above. We note that in the original Greek text, "ἡ ὄρνις" is feminine (article and noun), meaning the hen rather than the chicken. Therefore, here, we preferred the form "hen-or-egg" over "chicken-or-egg", which is more common in English (Very often, in online Greek texts, e.g., [1,2], "ἡ ὄρνις" appears as "ἡ ἄρνις". After extended searching, we contend that this must be an error, either an old one in copying of manuscripts, e.g., by monks in monasteries, or a modern one, e.g., in OCR. We are confident that the correct word is "ὄρνις").

The objective of the paper is to demonstrate that the relationship of atmospheric CO_2 and temperature may qualify as belonging to the category of "hen-or-egg" problems. First, we discuss the relationship between temperature and CO_2 concentration by revisiting intriguing results from proxy data-based palaeoclimatic studies, where the change in temperature leads and the change in CO_2 concentration follows. Next, we discuss the databases of modern (instrumental) measurements related to global temperature and atmospheric CO_2 concentration and introduce a methodology to analyse them. We develop a stochastic framework, introducing useful notions of time irreversibility and system causality while we discuss the logical and technical complications in identifying causality, which prompt us to seek just necessary, rather than sufficient, causality conditions. In the Results section, we examine the relationship of these two quantities using the modern data, available at the monthly time step. We juxtapose time series of global temperature and atmospheric CO_2 concentration from several sources, covering the common time interval 1980–2019. In our methodology, it is the timing rather than the magnitude of changes that is important, being the determinant of causality. While logical, physically based arguments support the "hen-or-egg" hypothesis, indicating that both causality directions exist, interpretation of cross-correlations of time series of global temperature and atmospheric CO_2 suggests that the dominant direction is $T \rightarrow CO_2$, i.e., the change in temperature leads and the change in CO_2 concentration follows. We attempt to interpret this latter mechanism by noting the positive feedback loop—higher temperatures increase soil respiration and, hence, CO_2 emissions.

The analysis reported in this paper was prompted by observation of an unexpected (and unfortunate) real-world experiment: during the COVID-19 lockdown in 2020, despite the unprecedented decrease in carbon emissions (Figure 1), there was an increase in atmospheric CO_2 concentration, which followed a pattern similar to previous years (Figure 2). Indeed, according to the International Energy Agency (IEA) [3], global CO_2 emissions were over 5% lower in the first quarter of 2020 than in that of 2019, mainly due to an 8% decline in emissions from coal, 4.5% from oil, and 2.3% from natural gas. According to other estimates [4], the decrease is even higher: the daily global CO_2 emissions decreased by 17% by early April 2020 compared with the mean 2019 levels, while for the whole 2020, a decrease between 4% and 7% is predicted. Despite that, as seen in Figure 2, the normal pattern of atmospheric CO_2 concentration (increase until May and decrease in June and July) did not change. Similar was the behaviour after the 2008–2009 financial crisis, but the most recent situation is more characteristic because the COVID-19 decline in 2020 is the most severe ever, even when considering the periods corresponding to World Wars. It is also noteworthy that, as shown in Figure 1, there are three consecutive years in the 2010s where there are no major increases, in emissions while there was an increase in CO_2 concentration. (At first glance, this does not sound reasonable and we have therefore cross-checked the data with other sources, namely the Global Carbon Atlas [5], and the database of Our World In Data [6]; we found only slight differences.) Interestingly, Figure 1 also shows a rapid growth in emissions after the 2008–2009 global financial crisis, which is in agreement with Peters et al. [7].

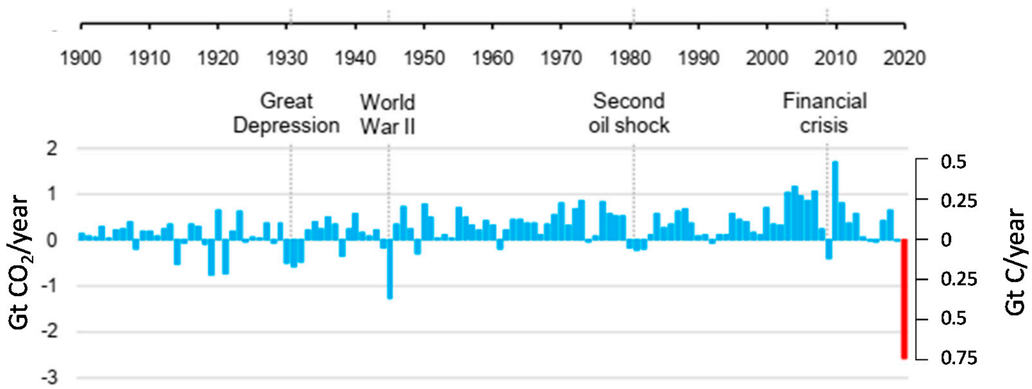


Figure 1. Annual change in global energy-related CO_2 emissions (adapted from International Energy Agency (IEA) [3]).

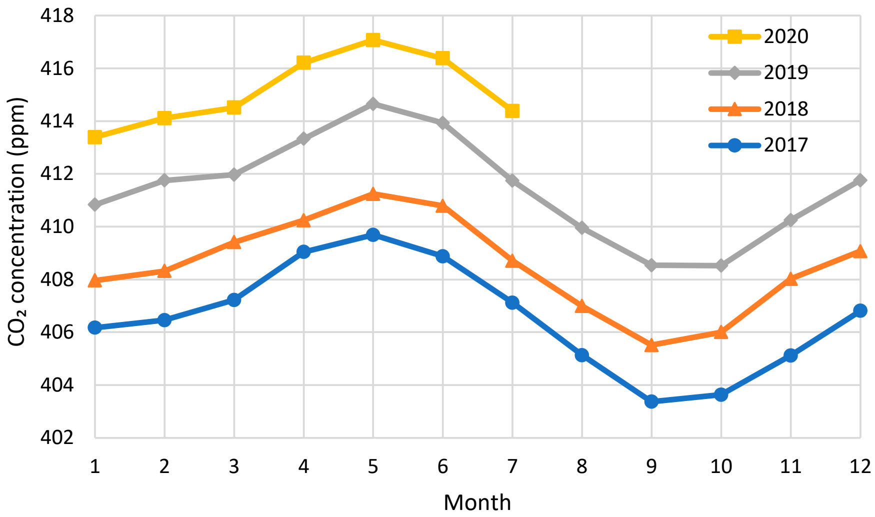


Figure 2. Atmospheric CO_2 concentration measured in Mauna Loa, Hawaii, USA, in the last four years.

2. Temperature and Carbon Dioxide—From Arrhenius and Palaeo-Proxies to Instrumental Data

Does the relationship of atmospheric carbon dioxide (CO_2) and temperature classify as a "hen-or-egg"-type causality? If we look at the first steps of studying the link between the two, the reply is clearly negative. Arrhenius (1896, [8]), the scientist most renowned for studying the causal relationship between two quantities, regarded the changes in atmospheric carbon dioxide concentration as the cause and the changes in temperature as the effect. Specifically, he stated:

> *Conversations with my friend and colleague Professor Högbom together with the discussions above referred to, led me to make a preliminary estimate of the probable effect of a variation of the atmospheric carbonic acid* [meaning CO_2] *on the temperature of the earth. As this estimation led to the belief that one might in this way probably find an explanation for temperature variations of 5–10 °C, I worked out the calculation more in detail and lay it now before the public and the critics.*

Furthermore, following the Italian meteorologist De Marchi (1895, [9]), whom he cited, he rejected what we call today *Milanković cycles* as possible causes of the glacial periods. In addition, he substantially overestimated the role of CO_2 in the greenhouse effect of the Earth's atmosphere. He calculated the relative weights of absorption of CO_2 and water vapour as 1.5 and 0.88, respectively, or a ratio of 1:0.6.

Arrhenius [8] also stated that "if the quantity of carbonic acid increases in geometric progression, the augmentation of the temperature will increase nearly in arithmetic progression". This Arrhenius's "rule" (which is still in use today) is mathematically expressed as

$$T - T_0 = \alpha \ln\left(\frac{[CO_2]}{[CO_2]_0}\right) \tag{1}$$

where T and $[CO_2]$ denote temperature and CO_2 concentration, respectively, T_0 and $[CO_2]_0$ represent reference states, and α is a constant.

Here, it is useful to note that Arrhenius's studies were not the first on the subject. Arrhenius [8] cites several other authors, among whom Tyndall (1865, [10]) for pointing out the enormous importance of atmospheric absorption of radiation and for having the opinion that water vapour has the greatest influence on this. Interestingly, it appears that the first experiments on the ability of water vapour and carbon dioxide to absorb heat were undertaken even earlier by Eunice Newton Foote (1856, [11]), even though she did not recognize the underlying mechanisms or even the distinction of short- and long-wave radiation [12–14]).

While the fact that the two variables are tightly connected is beyond doubt, the direction of the simple causal relationship needs to be studied further. Today, additional knowledge has been accumulated, particularly from palaeoclimatic studies, which allow us to examine Arrhenius's hypotheses on a sounder basis. In brief, we can state the following:

- Indeed, CO_2 plays a substantial role as a greenhouse gas. However, modern estimates of the contribution of CO_2 to the greenhouse effect differ largely from Arrhenius's results, attributing 19% of the long-wave radiation absorption to CO_2 against 75% of water vapour and clouds (Schmidt et al. [15]), or a ratio of 1:4.
- During the Phanerozoic Eon, Earth's temperature varied by even more than 5–10 °C, which was postulated by Arrhenius—see Figure 3. Even though the link of temperature and CO_2 is beyond doubt, this is not clear in Figure 3, where it is seen that the CO_2 concentration has varied by about two orders of magnitude and does not always synchronize with the temperature variation. Other factors may become more important at such huge time scales. Thus, an alternative hypothesis of the galactic cosmic ray flux as a climate driver via solar wind modulation has been suggested [16,17], which has triggered discussion or dispute [14,18–23]. The T–CO_2 relationship becomes more legible and rather indisputable in proxy data of the Quaternary (see Figure 4). It has been demonstrated in a persuasive manner by Roe [24] that in the Quaternary, it is the effect of Milanković cycles (variations in eccentricity, axial tilt, and precession of Earth's orbit), rather than of atmospheric CO_2 concentration, that explains the glaciation process. Specifically (quoting Roe [24]),

  > *variations in atmospheric CO_2 appear to lag the rate of change of global ice volume. This implies only a secondary role for CO_2—variations in which produce a weaker radiative forcing than the orbitally-induced changes in summertime insolation—in driving changes in global ice volume.*

Despite falsification of some of Arrhenius's hypotheses, his line of thought remained dominant. Yet, there have been some important studies, based on palaeoclimatological reconstructions (mostly the Vostok ice cores [25,26]), which have pointed to the opposite direction of causality, i.e., the change in temperature as the cause and that of CO_2 concentration as the effect. Such claims have been based on the fact that temperature change leads and CO_2 concentration change follows. In agreement with Roe [24], several papers have found the time lag to be positive, with estimates varying from 50 to 1000 years or more, depending on the time period and the particular study [27–32]. Claims that CO_2 concentration leads (i.e., a negative lag) have not generally been made in these studies. At most, a synchrony claim has been sought on the basis that the estimated positive lags are often within the 95% uncertainty range [31], while in one publication [29], it has been asserted that a "short lead of CO_2 over temperature cannot be excluded". With respect to the last deglacial warming, Liu et al. [32], using breakpoint lead–lag analysis, again find positive lags and conclude that the CO_2 is an internal feedback in Earth's climate system rather than an initial trigger.

Since palaeoclimatic data suggest a direction opposite to that assumed by Arrhenius, Koutsoyiannis [30], using palaeoclimatic data from the Vostok ice cores at a time resolution of 1000 years and a stochastic framework similar to that of the present study (see Section 4.1), concluded that a change in temperature precedes that of CO_2 by one time step (1000 years), as illustrated in Figure 4. He also noted that this causality condition holds for a wide range of time lags, up to 26,000 years, and, hence, the time lag is positive and most likely real. He asserted that the problem is obviously more complex than that of exclusive roles of cause and effect, classifying it as a "hen-or-egg" causality problem. Obviously, however, the proxy character of these data and the overly large time step of the time series reduce the reliability and accuracy of the results.

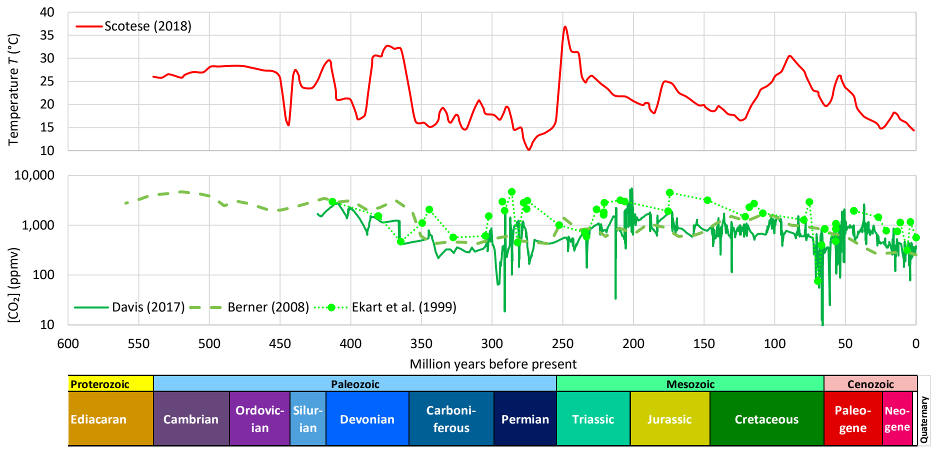


Figure 3. Proxy-based reconstructions of global mean temperature and CO_2 concentration during the Phanerozoic. The temperature reconstruction by Scotese [33] was mainly based on proxies from [21,34–36], while the CO_2 concentration proxies have been taken from Davis [37], Berner [38], and Ekart et al. [39]; all original figures were digitized in this study. The chronologies of geologic eras shown in the bottom of the figure have been taken from the International Commission on Stratigraphy [40].

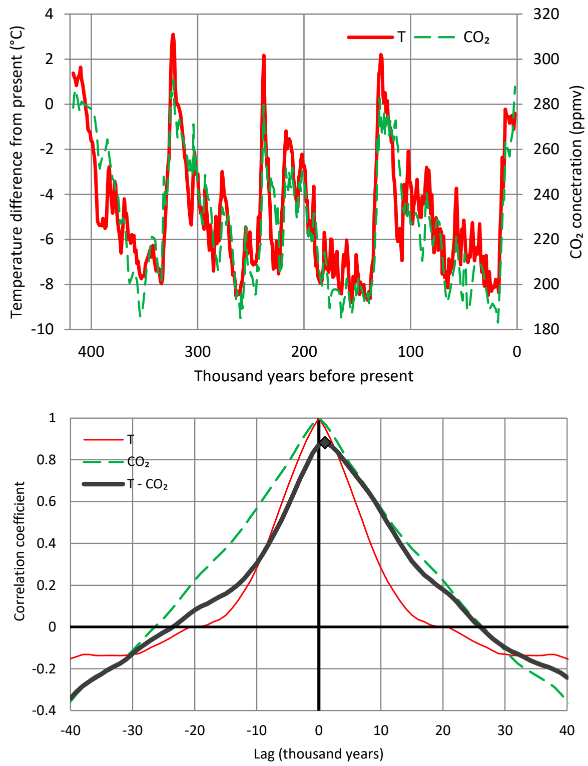


Figure 4. (**upper**) Time series of temperature and CO_2 concentration from the Vostok ice core, covering part of the Quaternary (420,000 years) with time step of 1000 years. (**lower**) Auto- and cross-correlograms of the two time series. The maximum value of the cross-correlation coefficient, marked as ◆, is 0.88 and appears at lag 1 (thousand years) (adapted from Koutsoyiannis [30]).

Studies exploring the rich body of modern datasets have also been published. Most of the studies have been based on the so-called "Granger causality test" (see Section 4.2). To mention a few, Kodra et al. [41], after testing several combinations and lags within the Granger framework, did not find any statistically significant results at the usual 5% significance level (they only found 2 cases at the

10% significance level; see their Tables 2 and 3). Stern and Kaufmann [42] studied, again within the Granger framework, the causality between radiative forcing and temperature, and found that both natural and anthropogenic forcings cause temperature change, and also that the inverse is true, i.e., temperature causes greenhouse gas concentration changes. They concluded that their results

> *show that properly specified tests of Ganger* [sic] *causality validate the consensus that human activity is partially responsible for the observed rise in global temperature and that this rise in temperature also has an effect on the global carbon cycle.*

By contrast, Stips et al. [43] used a different method [44] to investigate the causal structure and concluded that their

> *study unambiguously shows one-way causality between the total Greenhouse Gases and GMTA* [global mean surface temperature anomalies]. *Specifically, it is confirmed that the former, especially CO_2, are the main causal drivers of the recent warming.*

Here, we use a different path to study the causal relation between temperature and CO_2 concentration with the emphasis given on the exploratory and explanatory aspect of our analyses. While we occasionally use the Granger statistical test, this is not central in our approach. Rather, we place the emphasis on time directionality in the relationship, which we try to identify in the simplest possible manner, i.e., by finding the lag, positive or negative, which maximizes the cross-correlation between the two processes (see Section 4.1). We visualize our results by plots, so as to be simple, transparent, intuitive, readily understandable by the reader, and hopefully persuading. For the algorithmic-friendly reader, we also provide statistical testing results which just confirm what is directly seen in the graphs.

Another difference of our study, from most of the earlier ones, is our focus on changes, rather than current states, in the processes we investigate. This puts the technique of process differencing in central place in our analyses. This technique is quite natural and also powerful for studying time directionality [30]. We note that differencing has also been used in a study by Humlum et al. [45], which has several similarities with our study, even though it is not posed in a formal causality context, as well as in the study by Kodra et al. [41]. However, differencing has been criticized for potentially eliminating long-run effects and, hence, providing information only on short-run effects [42,46]. Even if this speculation were valid, it would not invalidate the differencing technique for the following reasons:

- The short-term effects deserve to be studied, as well as the long-term ones.
- The modern instrumental records are short themselves and only allow the short-term effects to be studied.
- For the long-term effects, the palaeo-proxies provide better indications, as already discussed above.

3. Data

Our investigation of the relationship of temperature with concentration of carbon dioxide in the atmosphere is based on two time series of the former process and four of the latter. Specifically, the temperature data are of two origins, satellite and ground-based. The satellite dataset, developed at the University of Alabama in Huntsville (UAH), infers the temperature, T, of three broad levels of the atmosphere from satellite measurements of the oxygen radiance in the microwave band using advanced (passive) microwave sounding units on NOAA and NASA satellites [47,48]. The data are publicly available on the monthly scale in the forms of time series of "anomalies" (defined as differences from long-term means) for several parts of earth as well as in maps. Here, we use only the global average on monthly scale for the lowest level, referred to as the lower troposphere. The ground-based data series we use is the CRUTEM.4.6.0.0 global T2m land temperature [49]. This originates from a gridded dataset of historical near-surface air temperature anomalies over land. Data are available for each month from January 1850 to the present. The dataset is a collaborative product of the Met Office

Hadley Centre and the Climatic Research Unit at the University of East Anglia. We note that both sources of information, UAH and CRUTEM, provide time series over the globe, land, and oceans; here, we deliberately use one source for the globe and one for the land.

The two temperature series used in the study are depicted in Figure 5. They are consistent with each other (and correlated, $r = 0.8$), though the CRUTEM4 series shows a larger increasing trend than the UAH series. The differences can be explained by three reasons: (a) the UAH series includes both land and sea, while the chosen CRUTEM4 series is for land only, in which the increasing trend is substantially higher than in sea; (b) the UAH series refers to some high altitude in the troposphere (see details in Koutsoyiannis [50]), while the CRUTEM4 series refers to the ground level; and (c) the ground-based CRUTEM4 series might be affected by urbanization (many ground stations are located in urban areas). In any case, the difference in the increasing trends is irrelevant for the current study, as the timing, rather than the magnitude, of changes is the determinant of causality. This will manifest in our results.

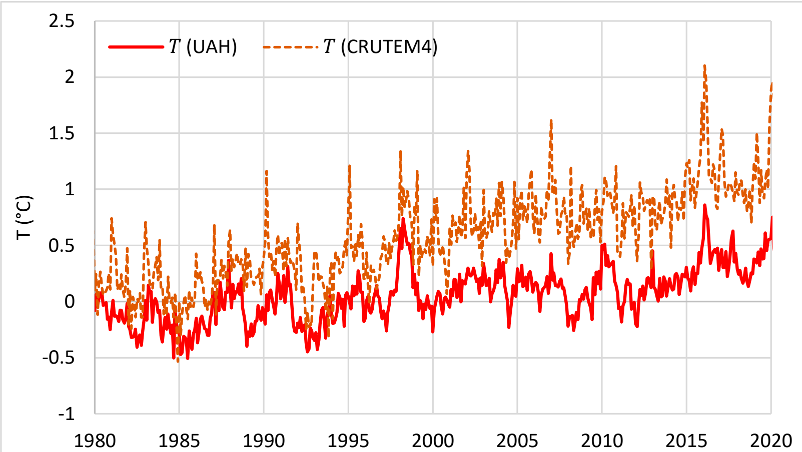


Figure 5. Plots of the data series of global temperature "anomalies" since 1980, as used in the study, from satellite measurements over the globe (UAH) and from ground measurements over land (CRUTEM4).

The most famous CO_2 dataset is that of Mauna Loa Observatory [51]. The Observatory, located on the north flank of Mauna Loa Volcano on the Big Island of Hawaii, USA, at an elevation of 3397 m above sea level, is a premier atmospheric research facility that has been continuously monitoring and collecting data related to the atmosphere since the 1950s. The NOAA also has other stations that systematically measure atmospheric CO_2 concentration, namely at Barrow, Alaska, USA and at South Pole. The NOAA's Global Monitoring Laboratory Carbon Cycle Group also computes global mean surface values of CO_2 concentration using measurements of weekly air samples from the Cooperative Global Air Sampling Network. The global estimate is based on measurements from a subset of network sites. Only sites where samples are predominantly of well-mixed marine boundary layer air, representative of a large volume of the atmosphere, are considered (typically at remote marine sea level locations with prevailing onshore winds). Measurements from sites at high altitude (such as Mauna Loa) and from sites close to anthropogenic and natural sources and sinks are excluded from the global estimate. (Details about this dataset are provided in [52]).

The period of data coverage varies, but all series cover the common 40-year period 1980–2019 which, hence, constitute the time reference of all our analyses. As a slight exception, the Barrow (Alaska) and South Pole measurements have not yet been available in final form for 2019 and, thus, this year was not included in our analyses of these two time series. The data of the latter two stations are given in irregular-step time series, which was regularized (by interpolation) to monthly in this study. All other data series have already been available on a monthly scale. While some of the earlier studies refer to a longer time span (e.g., [42,43] which start from 1850s), here, we avoid using non-systematic data earlier than 1980 due to their low reliability and bypass the raised controversies explained in Appendix A.1.

All four CO_2 time series used in the study are depicted in Figure 6. They show a superposition of increasing trends and annual cycles whose amplitudes increase as we head from the South to the North Pole. The South Pole series has opposite phase of oscillation compared to the other three.

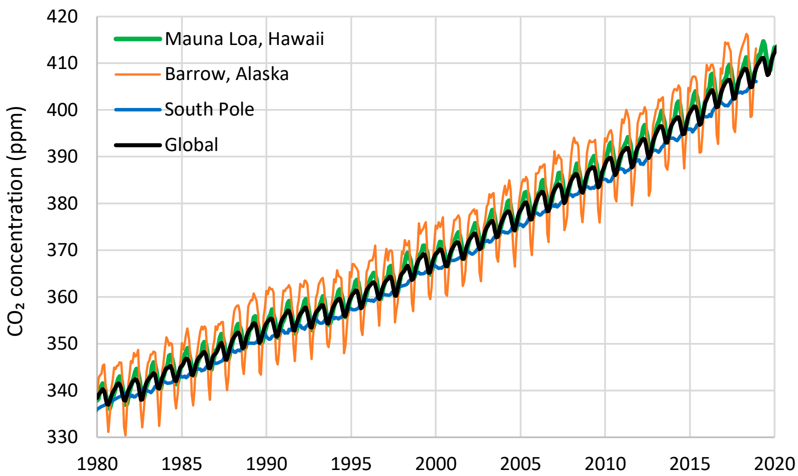


Figure 6. Plots of the data series of atmospheric CO_2 concentration measured in Mauna Loa (Hawaii, USA), Barrow (Alaska, USA), and South Pole, and the global average.

The annual cycle is better seen in Figure 7, where we have removed the trend with standardization, namely by dividing each monthly value by the geometric average of the preceding 5-year period. The reason why we used division rather than subtraction and geometric rather than arithmetic average (being thus equivalent to subtracting or averaging the logarithms of CO_2 concentration) will become evident in Section 5. In the right panel of Figure 7, which depicts monthly statistics of the time series of the left panel, it is seen that in all sites but the South Pole, the annual maximum occurs in May; that of the South Pole occurs in September.

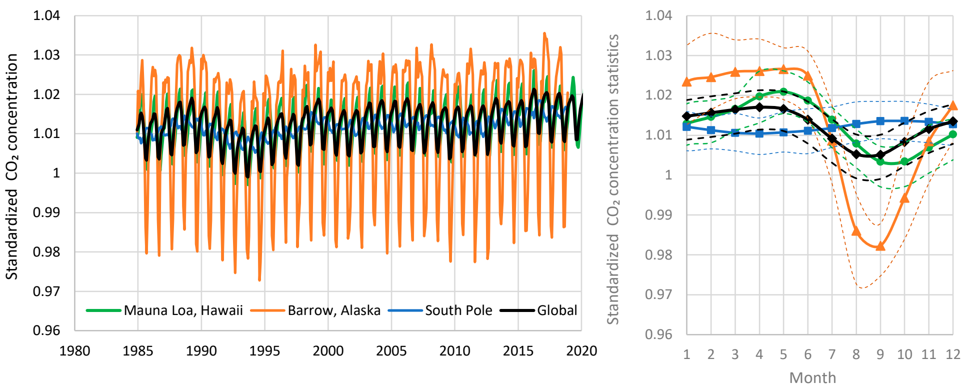


Figure 7. Plots of atmospheric CO_2 concentration after standardization: (**left**) each monthly value is standardized by dividing with the geometric average of the 5-year period before it. (**right**) Monthly statistics of the values of the left panel; for each month, the average is shown in continuous line and the minimum and maximum in thin dashed lines of the same colour as the average.

4. Methods

4.1. Stochastic Framework

A recent study [30] investigated time irreversibility in hydrometeorological processes and developed a theoretical framework in stochastic terms. It also studied necessary conditions for causality, which is tightly linked to time irreversibility. A simple definition of time reversibility within stochastics is the following, where underlined symbols denote stochastic (random) variables and non-underlined ones denote values thereof or regular variables.

A stochastic process $\underline{x}(t)$ at continuous time t, with nth order distribution function:

$$F(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) := P\Big\{\underline{x}(t_1) \le x_1, \underline{x}(t_2) \le x_2, \ldots, \underline{x}(t_n) \le x_n\Big\} \tag{2}$$

is time-symmetric or time-reversible if its joint distribution does not change after reflection of time about the origin, i.e., if for any $n, t_1, t_2, \ldots, t_n$,

$$F(x_1, x_2, \ldots, x_n; t_1, t_2, \ldots, t_n) = F(x_1, x_2, \ldots, x_n; -t_1, -t_2, \ldots, -t_n). \tag{3}$$

If times t_i are equidistant, i.e., $t_i - t_{i-1} = D$, the definition can be also written by reflecting the order of points in time, i.e.,

$$F(x_1, x_2, \ldots, x_{n-1}, x_n; t_1, t_2, \ldots, t_{n-1}, t_n) = F(x_1, x_2, \ldots, x_{n-1}, x_n; t_n, t_{n-1}, \ldots, t_2, t_1). \tag{4}$$

A process that is not time-reversible is called time-asymmetric, time-irreversible, or time-directional. Important results related to time (ir)reversibility are the following:

- A time reversible process is also stationary (Lawrance [53]).
- If a scalar process $\underline{x}(t)$ is Gaussian (i.e., all its finite dimensional distributions are multivariate normal) then it is reversible (Weiss [54]). The consequences are (a) a directional process cannot be Gaussian; (b) a discrete-time ARMA process (and a continuous-time Markov process) is reversible if and only if it is Gaussian.
- However, a vector (multivariate) process can be Gaussian and irreversible at the same time. A multivariate Gaussian linear process is reversible if and only if its autocovariance matrices are all symmetric (Tong and Zhang [55]).

Time asymmetry of a process can be studied more conveniently (or even exclusively in a scalar process) through the differenced process, i.e.,

$$\underline{\tilde{x}}_{\tau,\nu} := \underline{x}_{\tau+\nu} - \underline{x}_\tau \tag{5}$$

for an appropriate time-step ν of differencing. The differenced process represents change of the original process within a time period of length ν. We further define the cumulative process of $\underline{x}_\tau$ for discrete time κ as

$$\underline{X}_\kappa := \underline{x}_1 + \underline{x}_2 + \ldots + \underline{x}_\kappa. \tag{6}$$

Through this, we find that the time average of the original process $\underline{x}_\tau$ for discrete time scale κ is

$$\underline{x}_\tau^{(\kappa)} := \frac{\underline{x}_{(\tau-1)\kappa+1} + \underline{x}_{(\tau-1)\kappa+2} + \ldots + \underline{x}_{\tau\kappa}}{\kappa} = \frac{\underline{X}_{\tau\kappa} - \underline{X}_{(\tau-1)\kappa}}{\kappa}. \tag{7}$$

Similar equations for the cumulative and averaged processes for the differenced process $\underline{\tilde{x}}_{\tau,\nu}$ are given in Appendix A.2.

The variance of the process $\underline{x}_\tau^{(\kappa)}$ is a function of the time scale κ, which is termed the climacogram of the process:

$$\gamma_\kappa := \mathrm{var}[\underline{x}_\tau^{(\kappa)}]. \tag{8}$$

The autocovariance function for time lag η is derived from the climacogram through the relationship [56]

$$c_\eta = \frac{(\eta+1)^2\gamma_{|\eta+1|} + (\eta-1)^2\gamma_{|\eta-1|}}{2} - \eta^2\gamma_{|\eta|}. \quad (9)$$

For sufficiently large κ (theoretically as $\kappa \to \infty$), we may approximate the climacogram as

$$\gamma_\kappa \propto \kappa^{2H-2} \quad (10)$$

where H is termed the *Hurst parameter*. The theoretical validity of such (power-type) behaviour of a process was implied by Kolmogorov (1940 [57]). The quantity $2H$—2 is visualized as the slope of the double logarithmic plot of the climacogram for large time scales. In a random process, $H = 1/2$, while in most natural processes, $1/2 \leq H \leq 1$, as first observed by Hurst (1951 [58]). This natural behaviour is known as (long-term) *persistence* or *Hurst–Kolmogorov* (HK) *dynamics*. A high value of H (approaching 1) indicates enhanced presence of patterns, enhanced change, and enhanced uncertainty (e.g., in future predictions). A low value of H (approaching 0) indicates enhanced fluctuation or *antipersistence* (sometimes misnamed as quasi-periodicity as the period is not constant).

For a stationary stochastic process $\underline{x}_\tau$, the differenced process $\underline{\tilde{x}}_\tau$ has mean zero and variance:

$$\tilde{\gamma}_{\nu,1} := \mathrm{var}[\underline{\tilde{x}}_{\tau,\nu}] = \mathrm{var}[\underline{x}_{\kappa+\nu}] + \mathrm{var}[\underline{x}_\tau] - 2\mathrm{cov}[\underline{x}_{\tau+\nu}, \underline{x}_\tau] = 2(\gamma_1 - c_\nu) \quad (11)$$

where γ_1 and c_ν are the variance and lag ν autocovariance, respectively, of $\underline{x}_\tau$. Furthermore, it has been demonstrated [30] that the Hurst coefficient of the differenced process $\underline{\tilde{x}}_\tau$ precisely equals zero, which means that $\underline{\tilde{x}}_\tau$ is completely antipersistent, irrespective of γ_κ.

To study irreversibility in vector processes, we can use second-order moments and, in particular, cross-covariances among the different components of the vector. In particular (adapting and simplifying the analyses and results in Koutsoyiannis, [30]), given two processes $\underline{x}_\tau$ and $\underline{y}_\tau$, we could study the cross-correlations:

$$r_{\tilde{x}\tilde{y}}[\nu,\eta] := \mathrm{corr}\Big[\underline{\tilde{x}}_{\tau,\nu}, \underline{\tilde{y}}_{\tau+\eta,\nu}\Big]. \quad (12)$$

Time (ir)reversibility could then be characterized by studying the properties of symmetry or asymmetry of $r_{\tilde{x}\tilde{y}}(\nu,\eta)$ as a function of the time lag η. In a symmetric bivariate process, $r_{\tilde{x}\tilde{y}}[\nu,\eta] = r_{\tilde{x}\tilde{y}}[\nu,-\eta]$, and if the two components are positively correlated, the maximum of $r_{\tilde{x}\tilde{y}}[\nu,\eta]$ will appear at lag $\eta = 0$. If the bivariate process is irreversible, this maximum will appear at a lag $\eta_1 \neq 0$ and its value will be $r_{\tilde{x}\tilde{y}}[\nu,\eta_1]$.

Time asymmetry is closely related to causality, which presupposes irreversibility. Thus, "no causal process (i.e., such that of two consecutive phases, one is always the cause of the other) can be reversible" (Heller, [59]; see also [60]). In probabilistic definitions of causality, time asymmetry is determinant. Thus, Suppes [61] defines causation thus: "An event $B_{t'}$ [occurring at time t'] is a prima facie cause of the event A_t [occurring at time t] if and only if (i) $t' < t$, (ii) $P\{B_{t'}\} > 0$, (iii) $P(A_t|B_{t'}) > P(A_t)$". In addition, Granger's [62] first axiom in defining causality reads, "The past and present may cause the future, but the future cannot".

Consequently, in simple causal systems, in which the process component $\underline{x}_\tau$ is the cause of $\underline{y}_\tau$ (like in the clear case of rainfall and runoff, respectively), it is reasonable to expect $r_{\tilde{x}\tilde{y}}[\nu,\eta] \geq 0$ for any $\eta \geq 0$, while $r_{\tilde{x}\tilde{y}}[\nu,\eta] = 0$ for any $\eta = 0$. However, in "hen-or-egg" causal systems, this will not be the case, and we reasonably expect $r_{\tilde{x}\tilde{y}}[\nu,\eta] \neq 0$ for any η. Yet, we can define a dominant direction of causality based on the time lag η_1 maximizing cross-correlation. Formally, η_1 is defined for a specified ν as

$$\eta_1 := \underset{\eta}{\mathrm{argmax}}\big|r_{\tilde{x}\tilde{y}}(\nu,\eta)\big|. \quad (13)$$

We can thus distinguish the following three cases:

- If $\eta_1 = 0$, then there is no dominant direction.

- If $\eta_1 > 0$, then the dominant direction is $\underline{x}_\tau \rightarrow \underline{y}_\tau$.
- If $\eta_1 < 0$, then the dominant direction is $\underline{y}_\tau \rightarrow \underline{x}_\tau$.

Justification and further explanations of these conditions are provided in Appendix A.3.

4.2. Complications in Seeking Causality

It must be stressed that the above conditions are considered as necessary and not sufficient conditions for a causative relationship between the processes $\underline{x}_\tau$ and $\underline{y}_\tau$. Following Koutsoyiannis [30] (where additional necessary conditions are discussed), we avoid seeking sufficient conditions, a task that would be too difficult or impossible due to its deep philosophical complications as well as the logical and technical ones.

Specifically, it is widely known that correlation is not causation. As Granger [62] puts it,

> *when discussing the interpretation of a correlation coefficient or a regression, most textbooks warn that an observed relationship does not allow one to say anything about causation between the variables.*

Perhaps that is the reason why Suppes [61] uses the term "prima facie cause" in his definition given above which, however, he does not explain, apart for attributing "prima facie" to Jaakko Hintikka. Furthermore, Suppes discusses *spurious causes* and eventually defines the *genuine cause* as a "prima facie cause that is not spurious"; he also discusses the very existence of genuine causes which under certain conditions (e.g., in a Laplacean universe) seems doubtful.

Granger himself also uses the term "prima facie cause", while Granger and Newbold [63] note that a cause satisfying a causality test still remains prima facie because it is always possible that, if a different information set were used, it would fail the new test. Despite the caution issued by its pioneers, including Granger, through the years, the term "Granger causality" has become popular (particularly in the so-called "Granger causality test", e.g., [64]). Probably because of that misleading term, the technique is sometimes thought of as one that establishes causality, thus resolving or overcoming the "correlation is not causation" problem. In general, it has rarely been understood that identifying genuine causality is not a problem of choosing the best algorithm to establish a statistical relationship (including its directionality) between two variables. As an example of misrepresentation of the actual problems, see [65], which contains the statement:

> *Determining true causality requires not only the establishment of a relationship between two variables, but also the far more difficult task of determining a direction of causality.*

In essence, the "Granger causality test" studies the improvement of prediction of a process $\underline{y}_\tau$ by considering the influence of a "causing" process $\underline{x}_\tau$ through the Granger regression model:

$$\underline{y}_\tau = \sum_{j=1}^{\eta} a_j \underline{y}_{\tau-j} + \sum_{j=1}^{\eta} b_j \underline{x}_{\tau-j} + \varepsilon_\tau \tag{14}$$

where a_j and b_j are the regression coefficients and ε_τ is an error term. The test is based on the null hypothesis that the process $\underline{x}_\tau$ is not actually causing $\underline{y}_\tau$, formally expressed as

$$H_0 : b_1 = b_2 = \ldots = b_\eta = 0 \tag{15}$$

Algorithmic details of the test are given in [64], among others. The rejection of the null hypothesis is commonly interpreted in the literature with a statement that $\underline{x}_\tau$ "Granger-causes" $\underline{y}_\tau$.

This is clearly a misstatement and, in fact, the entire test is based on correlation matrices. Thus, it again reflects correlation rather than causation. The rejection of the null hypothesis signifies improvement of prediction and this does not mean causation. To make this clearer, let us consider the following example: people sweat when the atmospheric temperature is high and also wear light clothes. Thus, it is reasonably expected that in the prediction of sweat quantity, temperature matters. In the absence of temperature measurements (e.g., when we have only visual information, like when watching a video), the weight of the clothes algorithmically improves the prediction of sweat quantity. However, we could not say that the decrease in clothes weight causes an increase in sweat (the opposite is more reasonable and becomes evident in a three-variable regression, temperature–clothes weight–sweat, as further detailed in Appendix A.4).

Cohen [66] suggested replacing the term "Granger causality" with "Granger prediction" after correctly pointing out this:

> *Results from Granger causality analyses neither establish nor require causality. Granger causality results do not reveal causal interactions, although they can provide evidence in support of a hypothesis about causal interactions.*

To avoid such philosophical and logical complications, here, we replace the "prima facie" or "Granger" characterization of a cause and, as we already explained, we abandon seeking for genuine causes by using the notion of *necessary conditions* for causality. One could say that if two processes satisfy the necessary conditions, then they define a prima facie causality, but we avoid stressing that as we deem it unnecessary. Furthermore, we drop "causality" from "Granger causality test", thus hereinafter calling it "Granger test".

Some have thought they can approach genuine causes and get rid of the caution "correlation is not causation" by replacing the correlation with other statistics in the mathematical description of causality. For example, Liang [44] uses the concept of information (or entropy) flow (or transfer) between two processes; this method has been called "Liang causality" in the already cited work he co-authors [43]. The usefulness of such endeavours is not questioned, yet their vanity to determine genuine causality is easy to infer: It suffices to consider the case where the two processes, for which causality is studied, are jointly Gaussian. It is well known that in any multivariate Gaussian process, the covariance matrix (or the correlation matrix along with the variances) fully determines all properties of the multivariate distribution of any order. For example, the mutual information in a bivariate Gaussian process is (Papoulis, [67])

$$H[\underline{y}|x] = \ln \sigma_y \sqrt{2\pi e(1 - r^2)} \tag{16}$$

where σ and r denote standard deviation and correlation, respectively. Thus, using any quantity related to entropy (equivalently, information), is virtually identical to using correlation. Furthermore, in Gaussian processes, whatever statistic is used in describing causality, it is readily reduced to correlation. This is evident even in Liang [44], where, e.g., in his Equation (102), the information flow turns out to be the correlation coefficient multiplied by a constant. In other words, the big philosophical problem of causality cannot be resolved by technical tricks.

From what was exposed above (Section 4.1), the time irreversibility (or directionality) is most important in seeking causality. In this respect, we certainly embrace Suppes's condition (i) and Granger's first axiom, as stated above. Furthermore, we believe there is no meaning in refusing that axiom and continuing to speak about causality. We note though that there have been recent attempts to show that

> *coupled chaotic dynamical systems violate the first principle of Granger causality that the cause precedes the effect*. [68]

Apparently, however, the particular simulation experiment performed in the latter work which, notably, is not even accompanied by any attempt for deduction based on stochastics, cannot show any violation. In our view, such a violation, if it indeed happened, would be violation of logic and perhaps of common sense.

Additional notes for other procedures detecting causality, which are not included in the focus of our study, are given in Appendix A.4.

4.3. Additional Clarifications of Our Approach

After the above theoretical and methodological discourse, we can clarify our methodological approach by emphasizing the following points.

1. To make our assertions and, in particular, to use the "hen-or-egg" metaphor, we do not rely on merely statistical arguments. If we did that, based on our results presented in the next section, we would conclude that only the causality direction $T \rightarrow [CO_2]$ exists. However, one may perform a thought experiment of instantly adding a big quantity of CO_2 to the atmosphere. Would the temperature not increase? We believe it would, as CO_2 is known to be a greenhouse gas. The causation in the opposite direction is also valid, as will be discussed in Section 6, "Physical Interpretation". Therefore, we assert that both causality directions exist, and we are looking for the dominant one under the current climate conditions (those manifest in the datasets we use) instead of trying to make assertions of an exclusive causality direction.
2. While we occasionally use statistical tests (namely, the Granger test, Equations (14) and (15)), we opt to use, as the central point of our analyses, Equation (13) (and the conditions below it) because it is more intuitive and robust, fully reflects the basic causality axiom of time precedence, and is more straightforward, transparent (free of algorithmic manipulations), and easily reproducible (without the need for specialized software).
3. For simplicity, we do not use any statistic other than correlation here. We stress that the system we are examining is indeed classified as Gaussian and, thus, it is totally unnecessary to examine any statistic in addition to correlation. The evidence of Gaussianity is provided by Figures A1 and A2 in Appendix A.5, in terms of marginal distributions of the processes examined and in terms of their relationship. In particular, Figure A2 suggests a typical linear relationship for the bivariate process. We note that the linearity here is not a simplifying assumption or a coincidence as there are theoretical reasons implying it, which are related to the principle of maximum entropy [67,69].
4. All in all, we adhere to simplicity and transparency and, in this respect, we illustrate our results graphically, so they are easily understandable, intuitive, and persuasive. Indeed, our findings are easily verifiable even from simple synchronous plots of time series, yet we also include plots of autocorrelations and lagged cross-correlation, which are also most informative in terms of time directionality.

5. Results

5.1. Original Time Series

Here, we examine the relationship of atmospheric temperature and carbon dioxide concentration using the available modern data (observations rather than proxies) in monthly time steps, as described in Section 3. To apply our stochastic framework, we must first make the two time series linearly compatible. Specifically, based on Arrhenius's rule (Equation (1)), we take the logarithms of CO_2 concentration while we keep T untransformed. Such a transformation has also been performed in previous studies, which consider the logarithm of CO_2 concentration as a proxy of total radiative forcing (e.g., [41]). However, by calling this quantity "forcing", we indirectly give it, a priori (i.e., before investigating causation), the role of being the cause. Therefore, here, we avoid such interpretations; we simply call this variable the logarithm of carbon dioxide concentration and denote it as $\ln[CO_2]$.

A synchronous plot of the two processes (specifically, UAH temperature and $\ln[CO_2]$ at Mauna Loa) is depicted in Figure 8. Very little can be inferred from this figure alone. Both processes show increasing trends and thus appear as positively correlated. On the other hand, the two processes appear to have different behaviours. Temperature shows an erratic behaviour while $\ln[CO_2]$ has a smooth evolution marked by the annual periodicity. It looks impossible to infer causality from that graph alone.

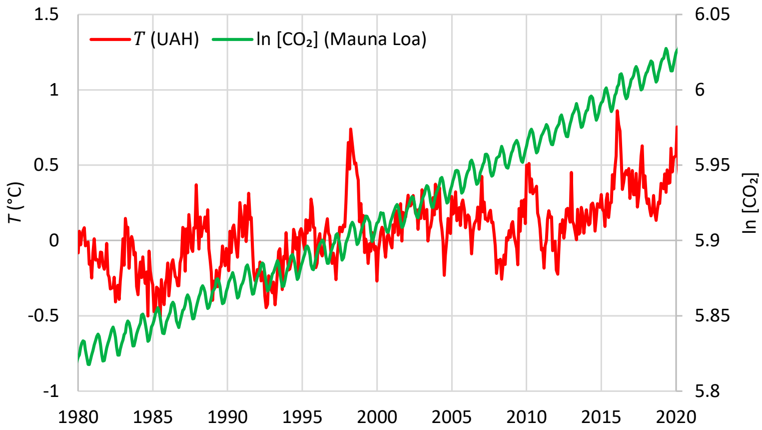


Figure 8. Synchronous plots of the time series of UAH temperature and logarithm of CO_2 concentration at Mauna Loa at monthly scale.

Somewhat more informative is Figure 9, which depicts lagged cross-correlations of the two processes, based on the methodology in Section 4.1 but without differencing the processes. Specifically, Figure 9 shows the cross-correlogram between UAH temperature and Mauna Loa $\ln[CO_2]$ at monthly and annual scales; the autocorrelograms of the two processes are also plotted for comparison. In both time scales, the cross-correlogram shows high correlations at all lags, with the maximum attained at lag zero. This does not hint at a direction. However, the cross-correlations for negative lags are slightly greater than those in the positive lags. Notice that to make this clearer, we have also plotted the differences $r_j - r_{-j}$ in the graph. This behaviour could be interpreted as supporting the causality direction $[CO_2] \rightarrow T$. However, we deem that the entire picture is spurious as it is heavily affected by the fact that the autocorrelations are very high and, in particular, those of $\ln[CO_2]$ are very close to 1 for all lags shown in the figure.

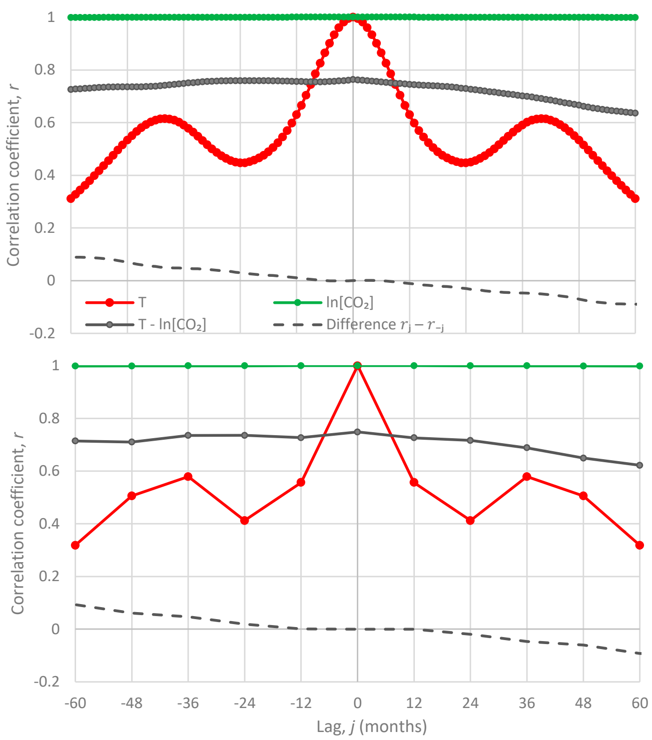


Figure 9. Auto- and cross-correlograms of the time series of UAH temperature and logarithm of CO_2 concentration at Mauna Loa.

In our investigation, we also applied the Granger test on these two time series in both time directions. To calculate the p-value of the Granger test, we used free software (namely the function GRANGER_TEST [70,71]). It appears that in the causality direction $[CO_2] \rightarrow T$, the null hypothesis is rejected at all usual significance levels. The attained p-value of the test is 1.8×10^{-7} for one regression lag ($\eta = 1$), 1.8×10^{-4} for $\eta = 2$, and remains below 0.01 for subsequent η. By contrast, in the direction $T \rightarrow [CO_2]$, the null hypothesis is not rejected at all usual significance levels. The attained p-value of the test is 0.25 for $\eta = 1$, 0.22 for $\eta = 2$, and remains above 0.1 for subsequent η.

Therefore, one could directly interpret these results as unambiguously showing one-way causality between the total greenhouse gases and temperature and, hence, validating the consensus view that human activity is responsible for the observed rise in global temperature. However, these results are certainly not unambiguous and, most probably, they are spurious. To demonstrate that they are not unambiguous, we have plotted, as shown in the upper panels of Figure 10, the p-values of the Granger test for moving windows with a size of 10 years for number of lags $\eta = 1$ and 2. The values for the entire length of time series, as given above, are also shown as dashed lines. Now the picture is quite different: each of the two directions appear dominating (meaning that the attained significance level is lower in one over the other) in about equal portions of the time. For example, for $\eta = 2$, the $T \rightarrow [CO_2]$ dominates over $[CO_2] \rightarrow T$ for 58% of the time. The attained p-value for direction $T \rightarrow [CO_2]$ is lower than 1% for 1.4% of the time, much higher than in the opposite direction (0.3% of the time). All of these observations favour the $T \rightarrow [CO_2]$ direction.

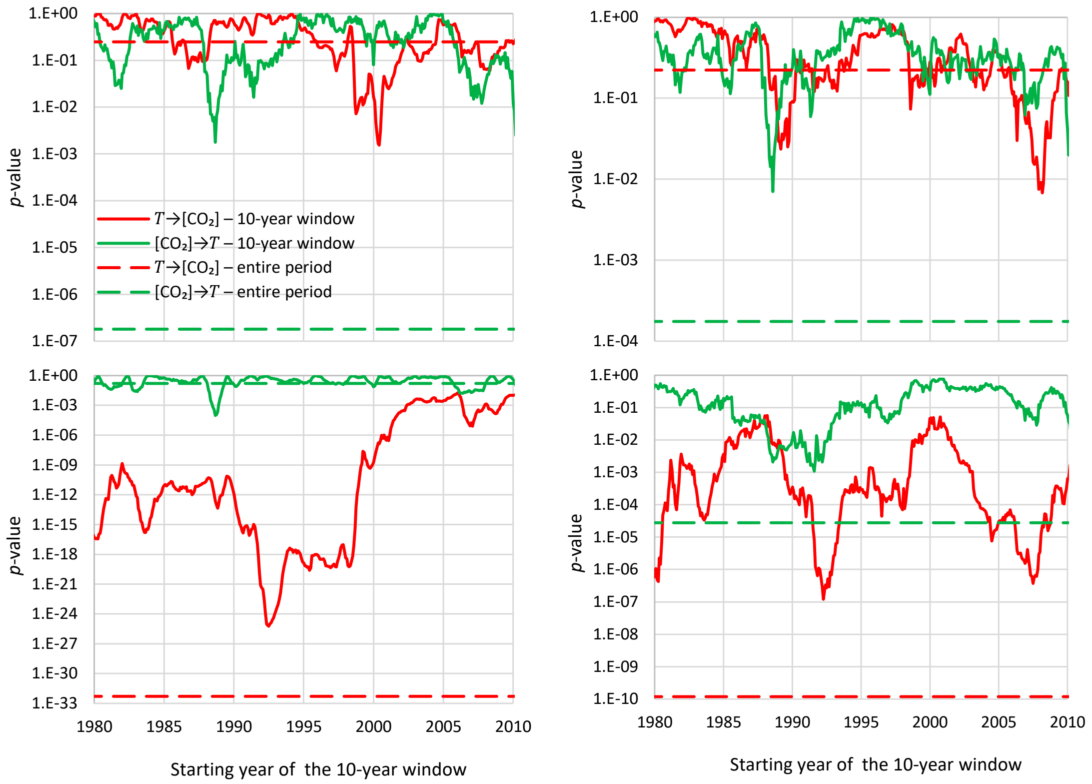


Figure 10. Plots of p-values of the Granger test for 10-year-long moving windows for the monthly time series of UAH temperature and logarithm of CO_2 concentration at Mauna Loa for number of lags (**left**) $\eta = 1$ and (**right**) $\eta = 2$. The time series used are (**upper**) the original and (**lower**) that obtained after "removing" the periodicity by averaging over the previous 12 months.

To show that the results are spurious and, in particular, affected by the very high autocorrelations of $\ln[CO_2]$ and, more importantly, by its annual cyclicity, we have "removed" the latter by averaging over the previous 12 months. We did that for both series and plotted the results in the lower panels of Figure 10. Here, the results are stunning. For both lags $\eta = 1$ and 2 and for the entire period (or almost), $T \to [CO_2]$ dominates, attaining p-values as low as in the order of 10^{-33}. However, we will avoid interpreting these results as unambiguous evidence that the consensus view (i.e., human activity is responsible for the observed warming) is wrong. Rather, what we want to stress is that it is inappropriate to draw conclusions from a methodology which is demonstrated to be so sensitive to the used time windows and data processing assumptions. In this respect, we have included this analyses in our study only (a) to show its weaknesses (which, for the reasons we explained in Section 4.2, we believe would not change if we used different statistics or different time series) and (b) to connect our study to earlier ones. For the sake of drawing conclusions, we contend that our full methodology in Sections 4.1 and 4.3 is more appropriate. We apply this methodology in Section 5.2.

5.2. Differenced Time Series

We have already explained the advantages of investigating the differenced processes, which quantify changes from a mathematical and logical point of view. In our case, taking differences is also physically meaningful as both CO_2 concentration and temperature (equivalent to thermal energy) represent "stocks", i.e., stored quantities and, thus, the mass and energy fluxes are indeed represented by differences.

The chosen time step of differencing is equal to one year ($\nu = 12$ for the monthly time step of the time series). For instance, from the value of January of a certain year, we subtract the value of January of the previous year and so forth. A first reason for this choice is that it almost eliminates the effect of the annual cycle (periodicity). A second reason is that the temperature data are given in terms of "anomalies", i.e., differences from an average which varies from month to month. By taking $\nu = 12$, the varying means are eliminated, and "anomalies" are effectively replaced by the actual processes (as the differences in the actual values equal the differences of "anomalies").

We perform all analyses on both monthly and annual time scales. Figure 11 shows the differenced time series for the UAH temperature and Mauna Loa CO_2 concentration at monthly scale; the symbols ΔT and $\Delta \ln[CO_2]$ are used interchangeably with $\underline{\tilde{x}}_{\tau,12}$ and $\underline{\tilde{y}}_{\tau,12}$, respectively.

Comparing Figure 8 (undifferenced series) with Figure 11 (differenced series), one can verify that the latter is much more informative in terms of the directionality of the relationship of the two processes. While Figure 8 did not provide any relevant hints, Figure 11 clearly shows that, most often, the temperature curve leads and that of CO_2 follows. However, there are cases where the changes in the two processes synchronize in time or even become decoupled.

Figure 12 shows the same time series at the annual time scale, with the year being defined as July–June for ΔT and February–January for $\Delta \ln[CO_2]$. The reason for this differentiation will be explained below. Here, it is more evident that, most of the time, the temperature change leads and that of CO_2 follows.

It is of interest here that the variability of global mean annual temperature is significantly influenced by the rhythm of ocean–atmosphere oscillations, such as ENSO, AMO, and IPO [72]. This mechanism may be a complicating factor, in turn influencing the link between temperature and CO_2 concentration. This is not examined here (except a short note in the end of the section) as, given the focus in examining just the connection of the latter two processes, it lies out of our present scope.

The climacograms of the differenced time series used (actually four of the six to avoid an overcrowded graph) are shown in Figure 13. It appears that the differenced temperature time series are consistent with the condition implied by stationarity, i.e., $H = 0$ for the differenced process. The same does not look to be the case for the CO_2 time series, particularly for the Mauna Loa time series, in which the Hurst parameter appears to be close to 1/2. Based on this, one would exclude stationarity for the Mauna Loa CO_2 time series. However, a simpler interpretation of the graph is that the data record is not long enough to reveal that $H = 0$ for the differenced process. Actually, all available data belong to a period in which $[CO_2]$ exhibits a monotonic increasing trend (as also verified by the fact that all values of $\Delta \ln[CO_2]$ in Figures 11 and 12 are positive, while stationarity entails a zero mean of the differenced process). Had the available database been broader, both positive and negative trends could appear. Indeed, a broader view of the [CO_2] process based on palaeoclimatic data (Figures 3 and 4) would justify a stationarity assumption.

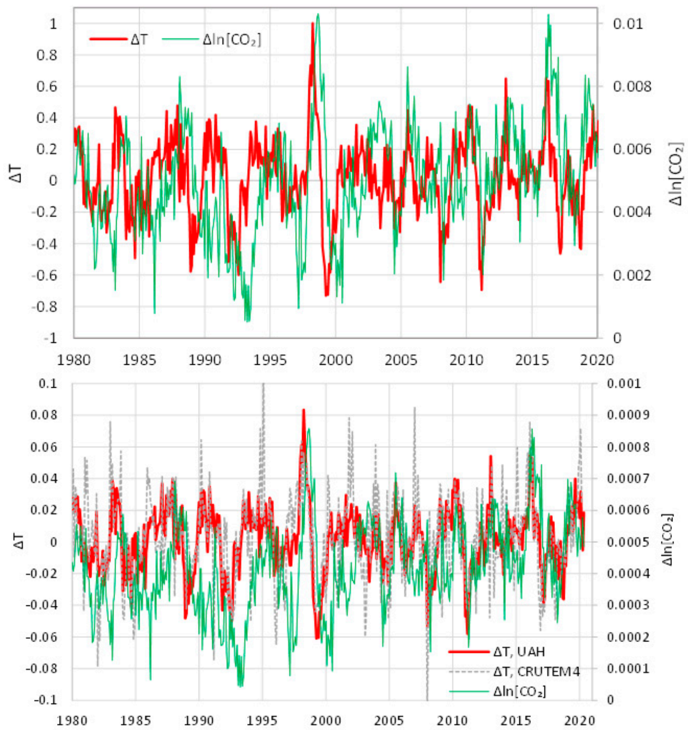


Figure 11. Differenced time series of UAH temperature and logarithm of CO_2 concentration at Mauna Loa at monthly scale. The graph in the upper panel was constructed in the manner described in the text. The graph in the lower panel is given for comparison and was constructed differently by taking differences of the values of each month with the previous month and then averaging over the previous 12 months (to remove periodicity); in addition, the lower graph includes the CRUTEM4 land temperature series.

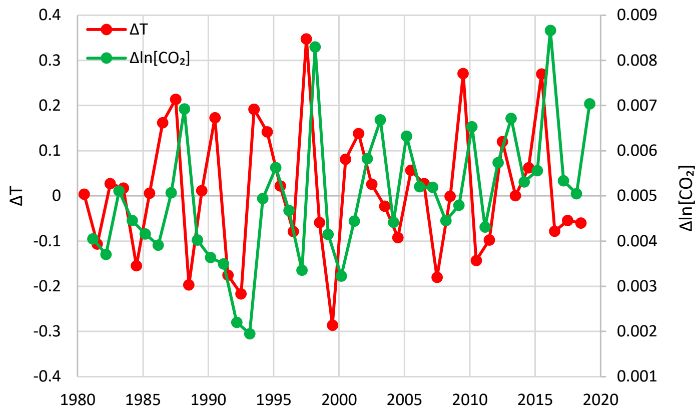


Figure 12. Annually averaged time series of differenced temperatures (UAH) and logarithms of CO_2 concentrations (Mauna Loa). Each dot represents the average of a one-year duration ending at the time of its abscissa.

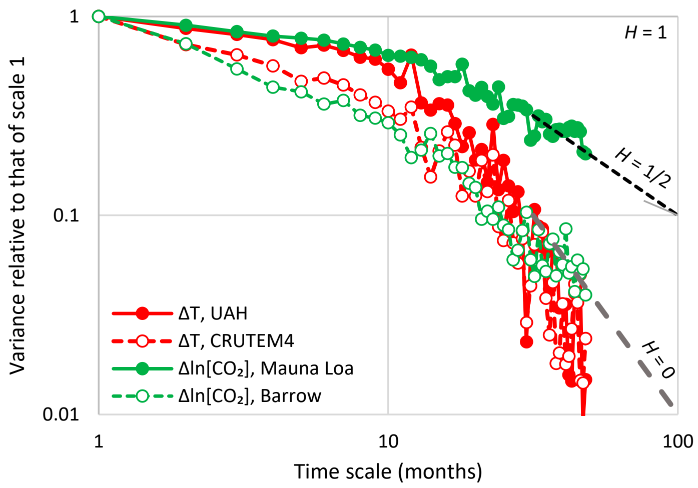


Figure 13. Empirical climacograms of the indicated differenced time series; the characteristic slopes corresponding to values of the Hurst parameter $H = 1/2$ (large-scale randomness), 0 (full antipersistence) and 1 (full persistence) are also plotted (note, $H = 1 + \text{slope}/2$).

The preliminary qualitative observation from graphical inspection of Figures 11 and 12 suggests that the temperature change very often precedes and the CO_2 change follows in the same direction. We note, though, that temperature changes alternate in sign while CO_2 changes are always positive.

A quantitative analysis based on the methodology in Section 4.1 requires the study of lagged cross-correlations of the two processes. Figure 14 shows the cross-correlogram between UAH temperature and Mauna Loa CO_2 concentration; the autocorrelograms of the two processes are also plotted for comparison. The fact that the cross-correlogram does not have values consistently close to zero at any of the semi-axes eliminates the possibility of an exclusive (unidirectional) causality and suggests consistency with "hen-or-egg" causality.

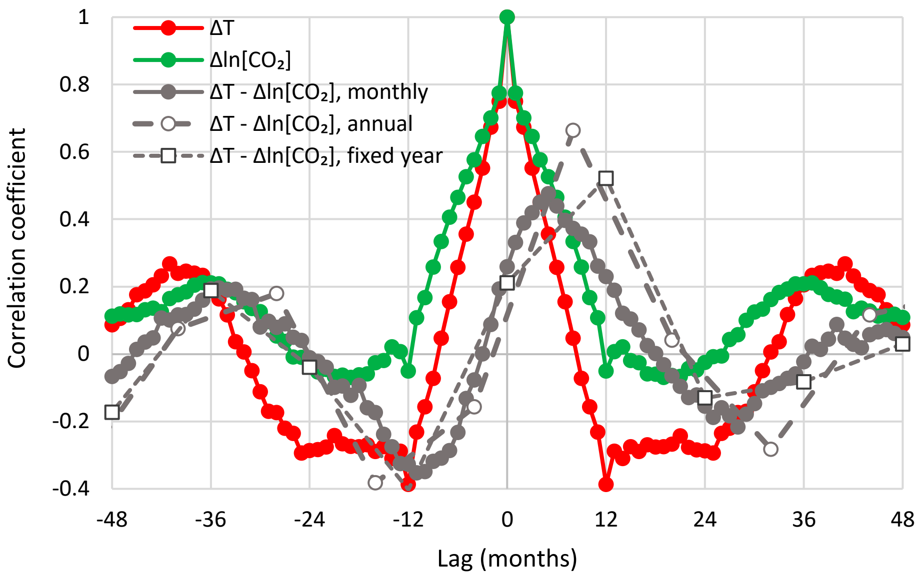


Figure 14. Auto- and cross-correlograms of the differenced time series of UAH temperature and Mauna Loa CO_2 concentration.

The maximum cross-correlation of the monthly series is 0.47 and appears at a positive lag, $\eta_1 = 5$ months, thus suggesting $T \rightarrow [CO_2]$, rather than $[CO_2] \rightarrow T$, as dominant causality direction. Similar are the graphs of the other combinations of temperature and CO_2 datasets, which are shown in Appendix A.5 (Figures A3–A7). In all cases, η_1 is positive, ranging from 5 to 11 months.

To perform similar analyses on the annual scale, we fixed the specification of a year for temperature for the period July–June, as already mentioned, and then slid the initial month specifying the beginning of a year for CO_2 concentration so as to find a specification that maximizes the cross-correlation at the annual scale. In Figure 14, maximization occurs when the year specification is February–January (of the next year), i.e., if the lag is 8 months. The maximum cross-correlation is 0.66. If we keep the specification of the year for CO_2 concentration the same as in temperature (July–June), then maximization occurs at lag one year (12 months) and the maximum cross-correlation is 0.52. Table 1 summarizes the results for all combinations examined. The lags are always positive. They vary between 8 and 14 months for a sliding window specification and are 12 months for the fixed window specification. Most interestingly, the opposite phase in the annual cycle of CO_2 concentration in the South Pole, with respect to the other three sites, does not produce any noteworthy differences in the shape of the cross-correlogram nor in the time lags maximizing the cross-correlations.

Table 1. Maximum cross-correlation coefficient (MCCC) and corresponding time lag in months. The annual window for temperature is July–June, while for CO_2, it is either different (sliding), determined so as to maximize MCCC, or the same (fixed).

Temperature—CO_2 Series	Monthly Time Series		Annual Time Series—Sliding Annual Window		Annual Time Series—Fixed Annual Window	
	MCCC	Lag	MCCC	Lag	MCCC	Lag
UAH—Mauna Loa	0.47	5	0.66	8	0.52	12
UAH—Barrow	0.31	11	0.70	14	0.59	12
UAH—South Pole	0.37	6	0.54	10	0.38	12
UAH—Global	0.47	6	0.60	11	0.60	12
CRUTEM4—Mauna Loa	0.31	5	0.55	10	0.52	12
CRUTEM4—Global	0.33	9	0.55	12	0.55	12

While, as explained in Sections 4.2 and 5.1, the Granger test has weaknesses that may not help in drawing conclusions, for completeness and as a confirmation, we list its results here:

- For the monthly scale and the causality direction $[CO_2] \to T$, the null hypothesis is not rejected at all usual significance levels for lag $\eta = 1$ and is rejected for significance level 1% for $\eta = 2$–8, with minimum attained p-value 1.8×10^{-4} for $\eta = 6$.
- For the monthly scale and the causality direction $T \to [CO_2]$, the null hypothesis is rejected at all usual significance levels for all lags η, with minimum attained p-value 2.1×10^{-8} for $\eta = 7$.
- For the monthly scale, the attained p-values in the direction $T \to [CO_2]$ are always smaller than in direction $[CO_2] \to T$ by about 4 to 5 orders of magnitude, thus clearly supporting $T \to [CO_2]$ as dominant direction.
- For the annual scale with fixed year specification and the causality direction $[CO_2] \to T$, the null hypothesis is not rejected at all usual significance levels for any lag η, thus indicating that this causality direction does not exist.
- For the annual scale with fixed year specification and the causality direction $T \to [CO_2]$, the null hypothesis is not rejected at significance level 1% for all lags $\eta = 1$–6, with minimum attained p-value 5% for lag $\eta = 2$, thus supporting this causality direction at this significance level.
- For the annual scale with fixed year specification, the attained p-values in the direction $T \to [CO_2]$ are always smaller than in direction $[CO_2] \to T$, again clearly supporting $T \to [CO_2]$ as the dominant direction.

We note that the test cannot be applied for the sliding time window case and, hence, we cannot provide results for this case.

We add a final remark, in view of a comment by Masters and Benestad [73] on the already mentioned study by Humlum et al. [45], in which they claim that "the inter-annual fluctuations in atmospheric CO_2 produced by ENSO can lead to a misdiagnosis of the long-term cause of the recent atmospheric CO_2 increase". Inspired by this comment, we have made a preliminary three-variable investigation using differenced temperatures (UAH), logarithms of CO_2 concentrations (Mauna Loa), and Equatorial South Oscillation index (SOI) characterizing ENSO. The investigation has been made on a monthly scale. $\Delta \ln[CO_2]$ has been linearly regressed with ΔT and the running average of SOI for the previous 12 months. At synchrony (without applying any time lag), the correlation of SOI with ΔT is 0.40, higher than that of ΔT and $\Delta \ln[CO_2]$ (0.24, as seen in Figure 14 at lag 0). The highest determination coefficient for the three regressed quantities is obtained when the time lag between $\Delta \ln[CO_2]$ and ΔT is again 5 months, as in the two-variable case (the optimal lag for SOI is 0, but the regression is virtually insensitive to the change of that lag). Its value is $r^2 = 0.23$, corresponding to $r = 0.48$, i.e., only slightly higher than the maximum cross-correlation coefficient of the two variable-case (which is 0.47 as seen in Table 1). In other words, by including ENSO in the modelling framework, the results do not change.

In brief, all above confirm the results of our methodology that the dominant direction of causality is $T \rightarrow [CO_2]$.

6. Physical Interpretation

The omnipresence of positive lags on both monthly and annual time scales and the confirmation by Granger tests reduce the likelihood that our results are statistical artefacts. Still, our results require physical interpretation which we seek in the natural process of soil respiration.

Soil respiration, R_s, defined to be the flux of microbially and plant-respired CO_2, clearly increases with temperature. It is known to have increased in the recent years [74,75]. Observational data of R_s (e.g., [76,77]; see also [78]) show that the process intensity increases with temperature. Rate of chemical reactions, metabolic rate, as well as microorganism activity, generally increase with temperature. This has been known for more than 70 years (Pomeroy and Bowlus [79]) and is routinely used in engineering design.

The Figure 6.1 of the latest report of the IPCC [75] provides a quantification of the mass balance of the carbon cycle in the atmosphere that is representative of recent years. The soil respiration, assumed to be the sum of respiration (plants) and decay (microbes), is 113.7 Gt C/year (IPCC gives a value of 118.7 including fire, which along with biomass burning, is estimated to be 5 Gt C/year by Green and Byrne [80]).

We can expect that sea respiration would also have increased. Moreover, outgassing from the oceans must also have increased as the solubility of CO_2 in water decreases with increasing temperature [14,81]. In addition, photosynthesis must have increased, as in the 21st century the Earth has been greening, mostly due to CO_2 fertilization effects [82] and human land-use management [83]. Specifically, satellite data show a net increase in leaf area of 2.3% per decade [83]. The sums of carbon outflows from the atmosphere (terrestrial and maritime photosynthesis as well as maritime absorption) amount to 203 Gt C/year. The carbon inflows to the atmosphere amount to 207.4 Gt C/year and include natural terrestrial processes (respiration, decay, fire, freshwater outgassing as well as volcanism and weathering), natural maritime processes (respiration) as well as anthropogenic processes. The latter comprise human CO_2 emissions related to fossil fuels and cement production as well as land-use change, and amount to 7.7 and 1.1 Gt C/year, respectively. The change in carbon fluxes due to natural processes is likely to exceed the change due to anthropogenic CO_2 emissions, even though the latter are generally regarded as responsible for the imbalance of carbon in the atmosphere.

7. Conclusions

The relationship between atmospheric concentration of carbon dioxide and the global temperature is widely recognized, and it is common knowledge that increasing CO_2 concentration plays a major role in enhancement of the greenhouse effect and contributes to global warming.

While the fact that these two variables are tightly connected is beyond doubt, the direction of the causal relationship needs to be studied further. The purpose of this study is to complement the conventional and established theory, that increased CO_2 concentration due to anthropogenic emissions causes an increase of temperature, by considering the concept of reverse causality. The problem is obviously more complex than that of exclusive roles of cause and effect, qualifying as a "hen-or-egg" ("ὄρνις ἢ ᾠὸν") causality problem, where it is not always clear which of two interrelated events is the cause and which the effect. Increased temperature causes an increase in CO_2 concentration and, hence, we propose the formulation of the entire process in terms of a "hen-or-egg" causality.

We examine the relationship of global temperature and atmospheric carbon dioxide concentration using the most reliable global data that are available—the data gathered from several sources, covering the common time interval 1980–2019, available at the monthly time step.

The results of the study support the hypothesis that both causality directions exist, with $T \rightarrow CO_2$ being the dominant, despite the fact that $CO_2 \rightarrow T$ prevails in public, as well as in scientific, perception. Indeed, our results show that changes in CO_2 follow changes in T by about six months on a monthly scale, or about one year on an annual scale.

The opposite causality direction opens a nurturing interpretation question. We attempted to interpret this mechanism by noting that the increase in soil respiration, reflecting the fact that the intensity of biochemical process increases with temperature, leads to increasing natural CO_2 emission. Thus, the synchrony of rising temperature and CO_2 creates a positive feedback loop. This poses challenging scientific questions of interpretation and modelling for further studies. In this respect, we welcome the review by Connolly [14], which already proposes interesting interpretations within a wider epistemological framework and in connection with a recent study [84]. In our opinion, scientists of the 21st century should have been familiar with unanswered scientific questions as well as with the idea that complex systems resist simplistic explanations.

Author Contributions: Conceptualization, D.K.; methodology, D.K.; software: D.K.; validation, Z.W.K.; formal analysis, D.K.; investigation, D.K. and Z.W.K.; data curation, D.K.; writing—original draft preparation, D.K. and Z.W.K.; writing—review and editing, D.K. and Z.W.K.; visualization, D.K. and Z.W.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding but was motivated by the scientific curiosity of the authors.

Acknowledgments: Some negative comments of two anonymous reviewers for an earlier submission of this manuscript in another journal (which we have posted online: manuscript at http://dx.doi.org/10.13140/RG.2.2.29154.15045/1, reviews at http://dx.doi.org/10.13140/RG.2.2.14524.87681) helped us improve the presentation and strengthen our arguments against their comments. We appreciate these reviewers' suggestions of relevant published works, which we were unaware of. We are grateful to the reviewers of the submission of our study to *Sci*, Yog Aryal, Ronan Connolly, and Stavros Alexandris, whose constructive comments helped us to improve the paper, mostly by adding a lot of additional information and clarification in appendices. The discussion of DK with Antonis Christofides helped to substantially improve an initial version of Appendix A.4, which he informally reviewed. We thank the journal *Sci* for the interesting experience it offered us.

Conflicts of Interest: The authors declare no conflict of interest.

Data Availability: The two temperature time series and the Mauna Loa CO_2 time series are readily available on monthly scale from http://climexp.knmi.nl. All NOAA CO_2 data are available from https://www.esrl.noaa.gov/gmd/ccgg/trends/gl_trend.html. The CO_2 data of Mauna Loa were retrieved from http://climexp.knmi.nl/data/imaunaloa_f.dat while the original measurements are in https://www.esrl.noaa.gov/gmd/dv/iadv/graph.php?code=MLO. The Barrow series is available (in irregular step) in https://www.esrl.noaa.gov/gmd/dv/iadv/graph.php?code=BRW, and the South Pole series in https://www.esrl.noaa.gov/gmd/dv/data/index.php?site=SPO. All these data were accessed (using the "Download data" link in the above sites) in June 2020. The global CO_2 series is accessed at https://www.esrl.noaa.gov/gmd/ccgg/trends/gl_data.html, of which the "Globally averaged marine surface monthly mean data" are used here. The palaeoclimatic data of Vostok CO_2 were retrieved from http://cdiac.ess-dive.lbl.gov/ftp/trends/co2/vostok.icecore.co2 (dated January 2003, accessed September 2018) and the temperature data from http://cdiac.ess-dive.lbl.gov/ftp/trends/temp/vostok/vostok.1999.temp.dat (dated January 2000, accessed September 2018).

Appendix A

Appendix A.1. On Early Non-Systematic Measurements of CO_2

This Appendix (not contained in Version 1 of our paper) addresses comments by all three reviewers of Version 1, Yog Aryal [85], Ronan Connolly [14], and Stavros Alexandris [86], about the reasons why we delimit our analysis to the period 1980–2019. The two latter reviewers suggested using earlier data compiled by Beck (2007), who referred to old chemical analyses of atmospheric concentration of CO_2.

We are sympathetic to the passion of the late Ernst-Georg Beck who, being a biology teacher, sacrificed a lot of time and effort to the exciting exercise of digging out old CO_2 measurements. Indeed, it could be worthwhile to have a critical look at the historical data and to try to make order in them and utilize them. However, this would certainly warrant an individual paper with this particular aim.

Historically, it was not the first review paper of this sort. For instance, in his Table 1, Beck [87] refers to old works by Letts and Blake (~1900; [88]), who considered 252 papers with data (all in 19th century), and to Stepanova [89], who considered 229 papers with data (130 in 19th century and 99 in 20th century). Beck himself [87] considered 156 papers with data (82 in 19th and 74 in 20th century).

As usual, it is instructive to consider the paper by Beck [87] jointly with critical commentaries published later in the journal where the original paper appeared [90,91]. In particular, R.F. Keeling [90] opined that the old chemical measurements examined by Beck [87] "exhibit far too much geographic and short-term temporal variability to plausibly be representative of the background. The variability of these early measurements must therefore be attributed to 'local or regional' factors or poor measurement practice". Keeling [90] also noted "basic accounting problems". "Beck's 11-year averages show large swings, including an increase from 310 to 420 ppm between 1920 and 1945 (Beck's Figure 11)". "To drive an increase of this magnitude globally requires the release of 233 billion metric tons of carbon to the atmosphere. The amount is equivalent to more than a third of all the carbon contained in land plants globally. [. . .] To make a credible case, Beck needed to offer evidence for losses or gains of carbon of this magnitude from somewhere. He offered none."

Meijer [91] expressed the opinion that Beck's work "contains major flaws, such that the conclusions are wrong". He also wrote: "The measurements presented in the paper are indeed useless for the purpose the author wants to use them, certainly in the way the author interprets them". There is a lack of interpretation of diurnal and seasonal variability (effects called the "diurnal" and the "seasonal" rectifier in the literature) and consideration of atmospheric mixing or lack thereof. Meijer also criticized the lack of meta-data: "The necessary data to judge, namely measurement height, consecutive length of a record and especially temporal resolution, are lacking in [Beck's] Table 2. In the light of the above, the whole 'Discussion and Conclusion' section is invalid, including [Beck's] Figures 11–14". Indeed, the records mentioned in Beck's Table 2 were local and short-lasting, with the longest periods being 1920–1926. Beck's Figures 11 and 13 show concatenated short segments of data at different places.

There are some other puzzling elements in Beck's paper. For instance, in his Figure 5, referring to data from a meteorological station near Giessen, the variability of high amplitude seems suspicious and not physically realistic. In particular, from June to August 1940, the measured CO_2 concentration increases from 340 to 550 ppm (much more than in Beck´s Figure 5 discussed by Keeling [90] and Meijer [91] as quoted above), with weird seasonal behaviour. Beck himself admitted that the results for Giessen "need to be adjusted downwards to take account of anthropogenic sources of CO_2 from nearby city, an influence that has been estimated as lying between 10 and 70 ppm [. . .] by different authors".

The controversy and disputes among these authors extended beyond pure scientific issues. Thus, Beck [87] wrote "[t]he data accepted [. . .] had to be sufficiently low to be consistent with the greenhouse hypothesis of climate change controlled by rising CO_2 emissions from fossil fuel burning". On the other hand, Meijer [91] wrote: "The author even accuses the pioneers Callendar and [Charles David] Keeling of selective data use, errors or even something close to data manipulation". In addition, [R.F.] Keeling [90] noted: "Beck is [. . .] wrong when he asserts that the earlier data have been discredited only because they don't fit a preconceived hypothesis of CO_2 and climate. [. . .] Instead, the data have been ignored because they cannot be accepted as representative without violating our understanding of how fast the atmosphere mixes".

In view of the above questions about the data reliability as well as the controversies and disputes, we decided to limit the period of our study in 1980–2019 in which the measurements are systematic and verifiable because they are made in several locations simultaneously.

Appendix A.2. Some Notes on the Averaged Differenced Process

The cumulative process of the differenced process $\underline{\tilde{x}}_{\tau,\nu}$ will be

$$\begin{aligned}\underline{\tilde{X}}_{\kappa,\nu} := \underline{\tilde{x}}_{1,\nu} + \underline{\tilde{x}}_{2,\nu} + \ldots + \underline{\tilde{x}}_{\kappa,\nu} &= \underline{x}_{1+\nu} - \underline{x}_1 + \underline{x}_{2+\eta} - \underline{x}_2 + \ldots + \underline{x}_{\kappa+\nu} - \underline{x}_\kappa \\ &= \underline{X}_{\kappa+\nu} - \underline{X}_\nu - \underline{X}_\kappa.\end{aligned} \tag{A1}$$

Note that for $\eta = 1$, this simplifies to

$$\underline{\tilde{X}}_{\kappa,1} = \underline{X}_{\kappa+1} - \underline{X}_1 - \underline{X}_\kappa = \underline{x}_{\kappa+1} - \underline{x}_1 = \underline{\tilde{x}}_{\kappa,1} =: \underline{\tilde{x}}_\kappa. \tag{A2}$$

Following Equation (7), the average differenced process at discrete time scale $\kappa = \eta$ will be

$$\underline{\tilde{x}}_\tau^{(\kappa)} = \frac{\underline{\tilde{X}}_{\tau\kappa,\kappa} - \underline{\tilde{X}}_{(\tau-1)\kappa,\kappa}}{\kappa} = \frac{(\underline{X}_{\tau\kappa+\kappa} - \underline{X}_\kappa - \underline{X}_{\tau\kappa}) - (\underline{X}_{(\tau-1)\kappa+\kappa} - \underline{X}_\kappa - \underline{X}_{(\tau-1)\kappa})}{\kappa} \tag{A3}$$

which, noting that in the rightmost part, the two terms $\underline{X}_\kappa$ cancel each other and by virtue of (7), simplifies to

$$\underline{\tilde{x}}_\tau^{(\kappa)} = \underline{x}_{\tau+1}^{(\kappa)} - \underline{x}_\tau^{(\kappa)} = \underline{\tilde{x}}_{\tau,1}^{(\kappa)}. \tag{A4}$$

In other words, the average differenced process equals the differenced average process in the case that the differencing time step η has been chosen as equal to the averaging time scale κ. For $\kappa = \eta = 1$, we have $\underline{\tilde{x}}_\tau^{(1)} \equiv \underline{\tilde{x}}_{\tau,1} \equiv \underline{\tilde{x}}_\tau$.

Appendix A.3. Some Notes on Time Directionality of Causal Systems

In a unidirectional causal system in continuous time t, in which the process $\underline{x}(t)$ is the cause of $\underline{y}(t)$, an equation of the form

$$\underline{y}(t) = \int_0^\infty \alpha(s)\underline{x}(t-s)\mathrm{d}s \tag{A5}$$

should hold [67], where $\alpha(t)$ is the impulse response function. The causality condition is thus

$$\alpha(t) = 0 \text{ for } t < 0. \tag{A6}$$

Here, we consider systems with positive dependence, in which $\alpha(t) \geq 0$ for $t \geq 0$, which are possibly also excited by another process $\underline{v}(t)$, independent of $\underline{x}(t)$. Working in discrete time, we write

$$\underline{y}_\tau = \sum\nolimits_{j=0}^{\infty} \alpha_j \underline{x}_{\tau-j} + \underline{v}_\tau. \tag{A7}$$

Assuming (without loss of generality) zero means for all processes, multiplying by $\underline{x}_{\tau-\eta}$, taking expected values, and denoting the cross-covariance function as $c_{xy}[\eta] := \mathrm{E}\left[\underline{x}_{\tau-\eta}\underline{y}_\tau\right]$ and the autocovariance function as $c_x[\eta] := \mathrm{E}\left[\underline{x}_{\tau-\eta}\underline{x}_\tau\right]$, we find

$$c_{xy}[\eta] = \sum\nolimits_{j=0}^{\infty} \alpha_j c_x[\eta - j]. \tag{A8}$$

For $\eta > 0$, using the property that $c_x[\eta]$ is an even function ($c_x[\eta] = c_x[-\eta]$), we get

$$c_{xy}[\eta] = \sum_{j=0}^{\infty} \alpha_j c_x[j-\eta] = \sum_{j=0}^{\eta-1} \alpha_j c_x[\eta - j] + \sum_{j=\eta}^{\infty} \alpha_j c_x[j-\eta], \tag{A9}$$

and for the negative part

$$c_{xy}[-\eta] = \sum_{j=0}^{\infty} \alpha_j c_x[j+\eta]. \tag{A10}$$

With intuitive reasoning, assuming that the autocovariance function is decreasing ($c_x[j'] < c_x[j]$ for $j' > j$), as usually happens in natural processes, we may see that the rightmost term of Equations (A9) and (A10) should be decreasing functions of η (as for $j' > j$ it will be $c_x[j' - \eta] < c_x[j - \eta]$ and $c_x[j' + \eta] < c_x[j + \eta]$). However, the term $\sum_{j=0}^{\eta-1} \alpha_j c_x[\eta - j]$ of Equation (A9) is not decreasing. Therefore, it should attain a maximum value at some positive lag $\eta = \eta_1$. Thus, a positive maximizing lag, $\eta = \eta_1 > 0$, is a necessary condition for causality direction from $\underline{x}_\tau$ to $\underline{y}_\tau$. Conversely, the condition that the maximizing lag is negative is a sufficient condition to exclude the causality direction exclusively from $\underline{x}_\tau$ to $\underline{y}_\tau$.

All above arguments remain valid if we standardize (divide) by the product of standard deviations of the processes $\underline{x}_\tau$ and $\underline{y}_\tau$ and, thus, we can replace cross-covariances $c_{xy}[\eta]$ with cross-correlations $r_{xy}[\eta]$ (or, in the case of differenced processes, $r_{\tilde{x}\tilde{y}}[\nu, \eta]$).

Appendix A.4. Some Notes on the Alternative Procedures on Causality

Reviewer Yog Aryal [85] opined that we missed referring to the recent relevant works by Hannart et al. [92] and Verbitsky et al. [93]. In response to this comment, we include this Appendix (not contained in Version 1 of our paper) explaining, in brief, why we do not compare our results with the ones of those studies, also noting that only the latter study contains material that is prima facie comparable to ours. The former study, focusing on the so-called causal counterfactual theory, is more theoretical and also much more interesting. While we, too, are preparing a theoretical study, in which we will discuss some theories in detail, in this Appendix, we give some key elements of our theoretical disagreements and a counterexample that illustrates the disagreements.

We first note that in order to define causality, Hannart et al. [92] refer to the work on the 18th century philosopher David Hume and, in particular, his famous book *Enquiry concerning Human Understanding* [94] first published in 1748. From this book, we wish to quote the following important passage, which emphasizes the difficulties even in defining causality:

> *Our thoughts and enquiries are, therefore, every moment, employed about this relation: Yet so imperfect are the ideas which we form concerning it, that it is impossible to give any just definition of cause, except what is drawn from something extraneous and foreign to it.*

Hannart et al. [92], while studying the probability of occurrence of an event Y, introduced the two-valued variable X_f to indicate whether or not a forcing f is present, and continue as follows:

> *The probability $p_1 = P(Y = 1 \middle| X_f = 1)$ of the event occurring in the real world, with f present, is referred to as factual, while $p_0 = P(Y = 1 \middle| X_f = 0)$ is referred to as counterfactual. Both terms will become clear in the light of what immediately follows. The so-called fraction of attributable risk (FAR) is then defined as*
>
> $$\mathrm{FAR} = 1 - \frac{p_0}{p_1} \tag{A11}$$
>
> *The FAR is interpreted as the fraction of the likelihood of an event that is attributable to the external forcing.*

They also show that under some conditions, FAR is a probability which they denote PN and call probability of necessary causality. They stress that it "is important to distinguish between necessary and sufficient causality" and they associate PN (or FAR) "with the first facet of causality, that of necessity". They claim to have "introduced its second facet, that of sufficiency, which is associated with the symmetric quantity $1-(1-p_1)/(1-p_0)$"; they denote it as PS, standing for probability of sufficient causality.

Central to the logical framework of Hannart et al. [92] is the notion of *intervention* of an experimenter, which is equivalent to experimentation with the ability to set the value of the assumed cause to a desired value. Clearly, this is feasible in laboratory experiments and infeasible in natural processes. The authors resort to the "so-called *in silico experimentation*" which, despite the impressive name chosen, is intervention in a mathematical model that represents the process. Hence, objectively, they examine the "causality" that is embedded in the model rather than the natural causality. One may argue that this it totally unnecessary. It would be better to inspect the model's equations or code to investigate what causality has been embedded in the model instead of running simulations and calculating probabilities. In particular, if the models used are climate models as in [92], their inability to effectively describe (perform in "prime time") the real-world processes [50,95–100] makes the entire endeavour futile. Another notion these authors use is *exogeneity*, which is related to the so-called *causal graph*, reflecting the assumed dependencies among the studied variables. Specifically, they state "a sufficient condition for X to be exogenous wrt any variable is to be a top node of a causal graph".

Here, we will use the simple example of Section 4.2, temperature–clothes weight–sweat, to show that using the quantities FAR (or PN) and PS may give spurious results that do not correspond to necessary or sufficient conditions for causality, at least with their meaning in our paper.

We use the two-valued random variables $\underline{x}, \underline{y}, \underline{z}$ to model the states of temperature, clothes weight, and sweat, respectively. We designate the following states:

$x = 1$: being hot *above* a threshold;
$y = 1$: wearing clothes with weight *below* a threshold;
$z = 1$: sweat quantity *above* a threshold;

and the opposite states with $x = 0, y = 0, z = 0$, respectively. We choose the threshold of temperature so that $P\{\underline{x} = 0\} = P\{\underline{x} = 1\} = 0.5$ and that of clothes weight so that $P\{\underline{y} = 0\} = P\{\underline{y} = 1\} = 0.5$. We choose a small probability, 0.05, of wearing light clothes when cold, or heavy clothes when hot, i.e., $P\{\underline{y} = 1 | \underline{x} = 0\} = P\{\underline{y} = 0 | \underline{x} = 1\} = 0.05$ (generally, we avoid choosing zero probabilities; rather the minimum value we choose is 0.05).

Using the definition of conditional probability,

$$P\{\underline{y} = y | \underline{x} = x\} = \frac{P\{\underline{y} = y, \underline{x} = x\}}{P\{\underline{x} = x\}}, \tag{A12}$$

we find the probability matrix A with elements $a_{ij} = P\{\underline{x} = i, \underline{y} = j\}$ as follows:

$$A = \begin{array}{cc} \begin{bmatrix} 0.475 & 0.025 \\ 0.025 & 0.475 \end{bmatrix} & \begin{array}{c} x = 0 \\ x = 1 \end{array} \\ \begin{array}{cc} y = 0 & y = 1 \end{array} & \end{array}. \tag{A13}$$

Now, we assign plausible values to the conditional probabilities of high sweat, $P\{\underline{z} = 1 | \underline{x} = x, \underline{y} = y\}$, as follows:

Cold, heavy clothes: $P\{\underline{z} = 1 | \underline{x} = 0, \underline{y} = 0\} = 0.2$

Cold, light clothes: $P\{\underline{z} = 1 | \underline{x} = 0, \underline{y} = 1\} = 0.1$

Hot, heavy clothes: $P\{\underline{z} = 1 | \underline{x} = 1, \underline{y} = 0\} = 0.95$

Hot, light clothes: $P\{\underline{z} = 1 | \underline{x} = 1, \underline{y} = 1\} = 0.80$

Again, we have avoided setting any of the conditional probabilities to 0 (or 1), and we have used multiples of 0.05 for all of them.

Using the definition of conditional probability in the form

$$P\left\{\underline{z} = z \middle| \underline{x} = x, \underline{y} = y\right\} = \frac{P\left\{\underline{z} = z, \underline{y} = y, \underline{x} = x\right\}}{P\left\{\underline{y} = y, \underline{x} = x\right\}}, \tag{A14}$$

we find the joint probabilities for each of the triplets $\{x, y, z\}$ that are shown in Table A1.

Table A1. Joint probabilities $P\left\{\underline{x} = x, \underline{y} = y, \underline{z} = z\right\}$ for all triplets $\{x, y, z\}$

x	y	$z = 0$	$z = 1$
0	0	0.38	0.095
0	1	0.0225	0.0025
1	0	0.00125	0.02375
1	1	0.095	0.38
	$P\left\{\underline{z} = z\right\} =$	0.49875	0.50125

Now, assume that we let an "artificial intelligence entity" (AIE) decide on causality based on the probability rules of the Hannart et al. [92] framework. Our AIE has access to numerous videos of people and is "trained" to assign accurate values of y and z, referring to clothes and sweat, based on the images in videos. In the video images, no thermometers are shown and, thus, our AIE cannot assign values of x, nor can it be aware of the notion of temperature. Our AIE tries to construct a causal graph putting, say, $\underline{y}$ as a top node and $\underline{z}$ as an end node; hence, it assumes that $\underline{y}$ is exogenous. Based on the huge information it can access, our AIE can (a) claim that it has constructed a prediction model based on one part of the data (e.g., using the so-called deep-learning technique) and, hence, is able to perform "in silico experimentation" (even though this is not absolutely necessary) and (b) accurately estimate the joint and conditional probabilities related to $\{y, z\}$ using either the model, the data, or both. Provided that the dataset is large enough, it will come up with the true values for the conditional probabilities, which are $b_{ij} = P\left\{\underline{y} = i, \underline{z} = j\right\}$ and $c_{ij} = P\left\{\underline{z} = j \middle| \underline{y} = i\right\}$, and form the matrices B and C, respectively, with values as follows:

$$B = \begin{array}{cc} \left[\begin{array}{cc} 0.38125 & 0.11875 \\ 0.1175 & 0.3825 \end{array}\right] & \begin{array}{c} y = 0 \\ y = 1 \end{array} \\ \begin{array}{cc} z = 0 & z = 1 \end{array} & \end{array}, \quad C = \begin{array}{cc} \left[\begin{array}{cc} 0.7625 & 0.2375 \\ 0.235 & 0.765 \end{array}\right] & \begin{array}{c} y = 0 \\ y = 1 \end{array} \\ \begin{array}{cc} z = 0 & z = 1 \end{array} & \end{array}. \tag{A15}$$

Here, the true values b_{ij} have been determined from the values of Table A1 noting that

$$b_{ij} = P\left\{\underline{y} = i, \underline{z} = j\right\} = P\left\{\underline{z} = j, \underline{y} = i, \underline{x} = 0\right\} + P\left\{\underline{z} = j, \underline{y} = i, \underline{x} = 1\right\} \tag{A16}$$

and the true values c_{ij} have been determined from the definition of conditional probability:

$$P\left\{\underline{z} = z \middle| \underline{y} = y\right\} = \frac{P\left\{\underline{z} = z, \underline{y} = y\right\}}{P\left\{\underline{y} = y\right\}}. \tag{A17}$$

Our AIE will then implement the causality conditions of sweat on clothes weight, assigning $p_0 = P\left\{\underline{z} = 1 \middle| \underline{y} = 0\right\} = 0.2375$ and $p_1 = P\left\{\underline{z} = 1 \middle| \underline{y} = 1\right\} = 0.765$. It will further calculate the probability of necessary causality as PN = 0.690, and the probability of sufficient causality even higher, PS = 0.692. Hence, our AIE will inform us that there is all necessary and sufficient evidence that light clothes cause high sweat.

Now, coming to the study by Verbitsky et al. [93], we notice that it assumes that "each time series is a variable produced by its hypothetical low dimensional system of dynamical equations" and uses the technique of distances of multivariate vectors for reconstructing the system dynamics. As demonstrated in Koutsoyiannis [101], such assumptions and techniques are good for simple toy models but, when real-world systems are examined, low dimensionality appears as a statistical artifact because the reconstruction actually needs an incredibly high number of observations to work, which are hardly available. The fact that the sums of multivariate vectors of distances is a statistical estimator with huge uncertainty is often missed in studies of this type, which treat data as deterministic quantities to obtain unreliable results. We do not believe that the Earth system and Earth processes (including global temperature and CO_2) are of low dimensionality, and we deem it unnecessary to discuss the issue further. We only note the fact that global temperature and CO_2 virtually behave as Gaussian, which enables reliable estimation of standard correlations and dismiss the need to use the overly complex and uncertain correlation sums.

Appendix A.5. Additional Graphical Depictions

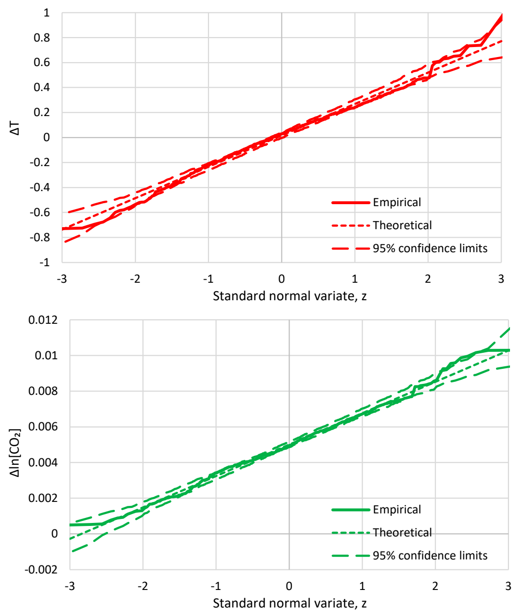


Figure A1. Normal probability plots of ΔT and $\Delta\ln[CO_2]$ where T is the UAH temperature and $[CO_2]$ is the CO_2 concentration at Mauna Loa at monthly scale.

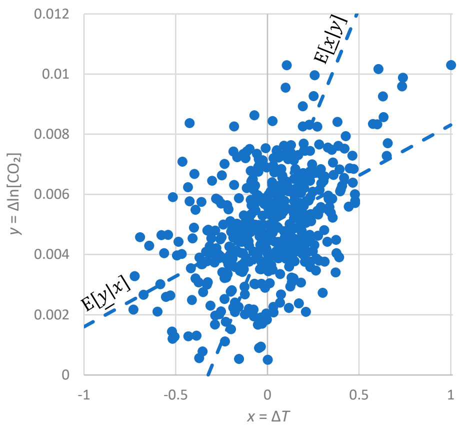


Figure A2. Scatter plot of ΔT and $\Delta\ln[CO_2]$ where T is the UAH temperature and $[CO_2]$ is the CO_2 concentration at Mauna Loa at monthly scale; the two quantities are lagged in time using the optimal lag of 5 months (Table 1). The two linear regression lines are also shown in the figure.

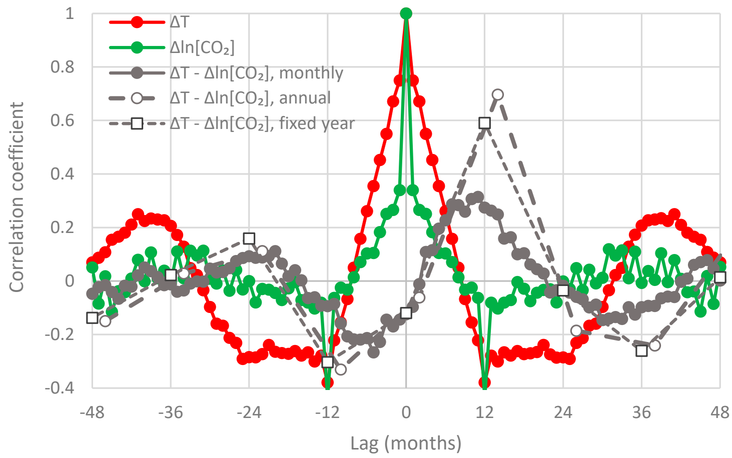


Figure A3. Auto- and cross-correlograms of the differenced time series of UAH temperature and Barrow CO_2 concentration.

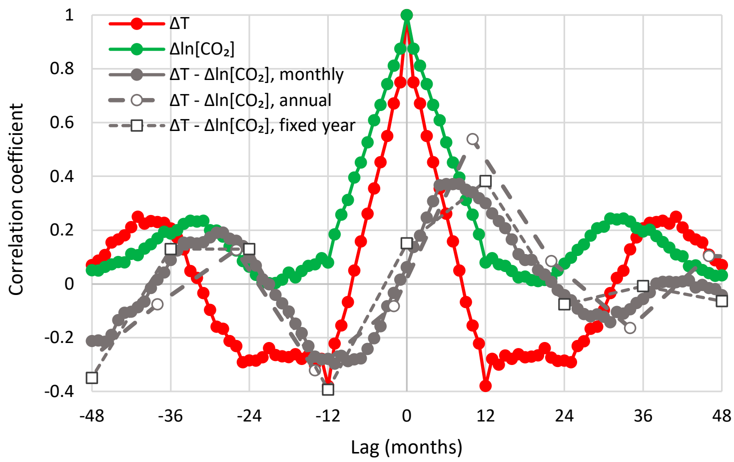


Figure A4. Auto- and cross-correlograms of the differenced time series of UAH temperature and South Pole CO_2 concentration.

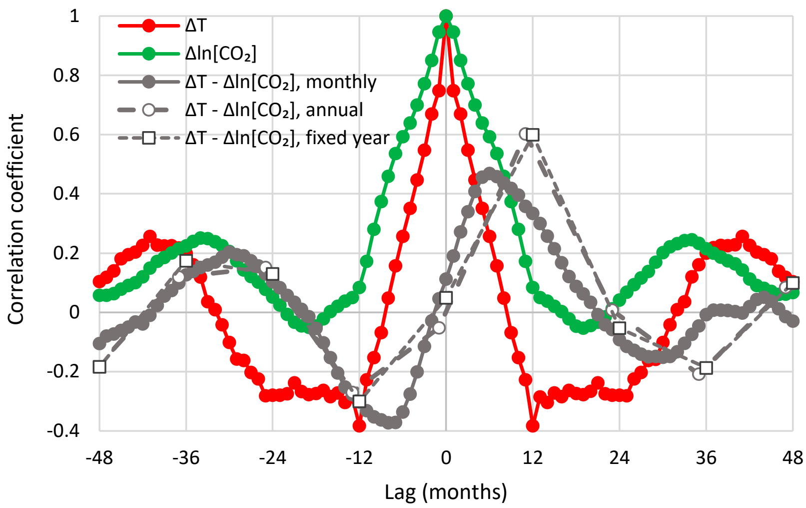


Figure A5. Auto- and cross-correlograms of the differenced time series of UAH temperature and global CO_2 concentration.

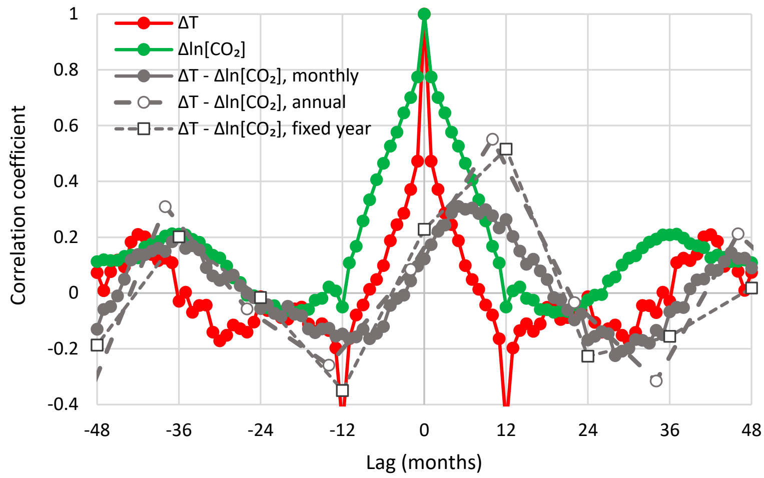


Figure A6. Auto- and cross-correlograms of the differenced time series of CRUTEM4 temperature and Mauna Loa CO_2 concentration.

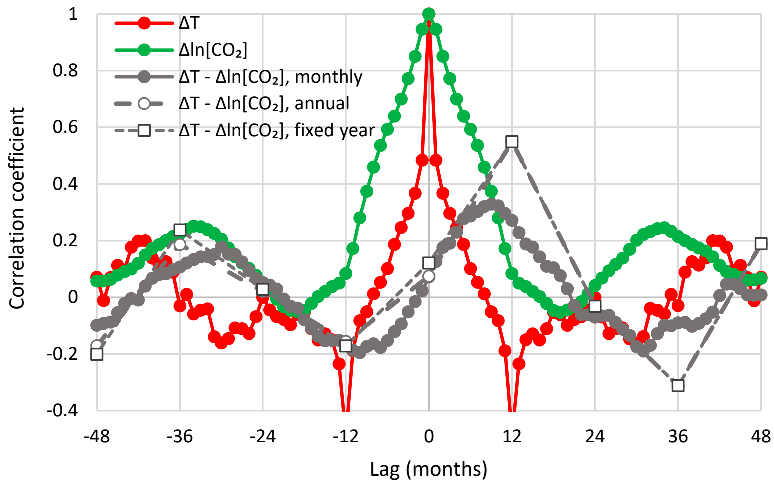


Figure A7. Auto- and cross-correlograms of the differenced time series of CRUTEM4 temperature and global CO_2 concentration.

References

1. Πλούταρχος, Συμποσιακά Β′ (Plutarch, Quaestiones Convivales B)—Βικιθήκη. Available online: https://el.wikisource.org/wiki/Συμποσιακά_Β′ (accessed on 15 October 2020).
2. Plutarchus-Quaestiones Convivales (612c–748d). Available online: http://www.poesialatina.it/_ns/Greek/html/map-en.html (accessed on 15 October 2020).
3. IEA (International Energy Agency). *Global Energy Review 2020*; IEA: Paris, French, 2020. Available online: https://www.iea.org/reports/global-energy-review-2020 (accessed on 1 July 2020).
4. Le Quéré, C.; Jackson, R.B.; Jones, M.W.; Smith, A.J.; Abernethy, S.; Andrew, R.M.; De-Gol, A.J.; Willis, D.R.; Shan, Y.; Canadell, J.G.; et al. Temporary reduction in daily global CO_2 emissions during the COVID-19 forced confinement. *Nat. Clim. Chang.* **2020**, *10*, 647–653. [CrossRef]
5. Global Carbon Atlas. Available online: http://www.globalcarbonatlas.org/en/CO2-emissions (accessed on 1 July 2020).
6. Our World in Data. Available online: https://ourworldindata.org/grapher/annual-co-emissions-by-region (accessed on 1 July 2020).
7. Peters, G.P.; Marland, G.; Le Quéré, C.; Boden, T.; Canadell, J.G.; Raupach, M.R. Rapid growth in CO_2 emissions after the 2008–2009 global financial crisis. *Nat. Clim. Chang.* **2012**, *2*, 2–4. [CrossRef]

8. Arrhenius, S. On the influence of carbonic acid in the air upon the temperature of the ground. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1896**, *41*, 237–276. [CrossRef]
9. De Marchi, L. *Le Cause dell'Era Glaciale*; Premiato dal R. Instituto Lombardo: Pavia, Italy, 1895.
10. Tyndall, J. *Heat a Mode of Motion*, 2nd ed.; Longmans: London, UK, 1865. Available online: https://archive.org/details/heatamodemotion05tyndgoog/ (accessed on 1 September 2020).
11. Foote, E. Circumstances affecting the heat of the sun's rays. *Am. J. Sci. Arts* **1856**, *22*, 382–383.
12. Jackson, R. Eunice Foote, John Tyndall and a question of priority. *Notes Rec.* **2020**, *74*, 105–118. [CrossRef]
13. Ortiz, J.D.; Jackson, R. Understanding Eunice Foote's 1856 experiments: Heat absorption by atmospheric gases. *Notes Rec.* **2020**. [CrossRef]
14. Connolly, R. Review of "Atmospheric temperature and CO_2: Hen-or-egg causality?" (Version 1). *Sci* **2020**. Available online: https://www.mdpi.com/2413-4155/2/4/77/review_report (accessed on 9 October 2020).
15. Schmidt, G.A.; Ruedy, R.A.; Miller, R.L.; Lacis, A.A. Attribution of the present-day total greenhouse effect. *J. Geophys. Res.* **2010**, *115*, D20106. [CrossRef]
16. Shaviv, N.J.; Veizer, J. Celestial driver of Phanerozoic climate? *Geol. Soc. Am. Today* **2003**, *13*, 4–10. [CrossRef]
17. Shaviv, N.J. On climate response to changes in the cosmic ray flux and radiative budget. *J. Geophys. Res. Space Phys.* **2005**, *110*, A08105. [CrossRef]
18. Rahmstorf, S.; Archer, D.; Ebel, D.S.; Eugster, O.; Jouzel, J.; Maraun, D.; Neu, U.; Schmidt, G.A.; Sever-Inghaus, J.; Weaver, A.J.; et al. Cosmic rays, carbon dioxide, and climate. *Eos* **2004**, *85*, 38–41. [CrossRef]
19. Shaviv, N.J.; Veizer, J. Comment on Cosmic rays, carbon dioxide, and climate. *Eos* **2004**, *85*, 510. [CrossRef]
20. Rahmstorf, S.; Archer, D.; Ebel, D.S.; Eugster, O.; Jouzel, J.; Maraun, D.; Neu, U.; Schmidt, G.A.; Sever-Inghaus, J.; Weaver, A.J.; et al. Reply to Comment [by Shaviv and Veizer]. *Eos* **2004**, *85*, 511. [CrossRef]
21. Royer, D.L.; Berner, R.A.; Montañez, I.P.; Tabor, N.J.; Beerling, D.J. CO_2 as a primary driver of phanerozoic climate. *Geol. Soc. Am. Today* **2004**, *14*, 4–10. [CrossRef]
22. Shaviv, N.J.; Veizer, J. Comment on "CO_2 as a primary driver of Phanerozoic climate" by Royer et al. *Geol. Soc. Am. Today* **2004**, *14*, 18.
23. Royer, D.L.; Berner, R.A.; Montañez, I.P.; Tabor, N.J.; Beerling, D.J. Reply to Comment on "CO_2 as a primary driver of Phanerozoic climate" by Royer et al. *Geol. Soc. Am. Today* **2004**, *14*, 18.
24. Roe, G. In defense of Milankovitch. *Geophys. Res. Lett.* **2006**, *33*. [CrossRef]
25. Jouzel, J.; Lorius, C.; Petit, J.R.; Genthon, C.; Barkov, N.I.; Kotlyakov, V.M.; Petrov, V.M. Vostok ice core: A continuous isotope temperature record over the last climatic cycle (160,000 years). *Nature* **1987**, *329*, 403–408. [CrossRef]
26. Petit, J.R.; Jouzel, J.; Raynaud, D.; Barkov, N.I.; Barnola, J.-M.; Basile, I.; Bender, M.; Chappellaz, J.; Davis, M.; Delayque, G.; et al. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **1999**, *399*, 429–436. [CrossRef]
27. Caillon, N.; Severinghaus, J.P.; Jouzel, J.; Barnola, J.M.; Kang, J.; Lipenkov, V.Y. Timing of atmospheric CO_2 and Antarctic temperature changes across Termination III. *Science* **2003**, *299*, 1728–1731. [CrossRef]
28. Soon, W. Implications of the secondary role of carbon dioxide and methane forcing in climate change: Past, present, and future. *Phys. Geogr.* **2007**, *28*, 97–125. [CrossRef]
29. Pedro, J.B.; Rasmussen, S.O.; van Ommen, T.D. Tightened constraints on the time-lag between Antarctic temperature and CO_2 during the last deglaciation. *Clim. Past* **2012**, *8*, 1213–1221. [CrossRef]
30. Koutsoyiannis, D. Time's arrow in stochastic characterization and simulation of atmospheric and hydrological processes. *Hydrol. Sci. J.* **2019**, *64*, 1013–1037. [CrossRef]
31. Beeman, J.C.; Gest, L.; Parrenin, F.; Raynaud, D.; Fudge, T.J.; Buizert, C.; Brook, E.J. Antarctic temperature and CO_2: Near-synchrony yet variable phasing during the last deglaciation. *Clim. Past* **2019**, *15*, 913–926. [CrossRef]
32. Liu, Z.; Huang, S.; Jin, Z. Breakpoint lead-lag analysis of the last deglacial climate change and atmospheric CO_2 concentration on global and hemispheric scales. *Quat. Int.* **2018**, *490*, 50–59. [CrossRef]
33. Scotese, C.R. Phanerozoic Temperatures: Tropical Mean Annual Temperature (TMAT), Polar Mean Annual Temperature (PMAT), and Global Mean Annual Temperature (GMAT) for the last 540 Million Years. In Proceedings of the Earth's Temperature History Research Workshop, Smithsonian National Museum of Natural History, Washington, DC, USA, 30–31 March 2018. Available online: https://www.researchgate.net/publication/324017003 (accessed on 8 September 2020).

34. Grossman, E.L. Oxygen isotope stratigraphy. In *The Geological Time Scale*; Gradstein, F.M., Ogg, J.G., Schmitz, M.D., Ogg, G.M., Eds.; Elsevier: Amsterdam, The Netherlands, 2012; pp. 181–206.
35. Veizer, J.; Prokoph, A. Temperatures and oxygen isotopic composition of Phanerozoic oceans. *Earth Sci. Rev.* **2015**, *146*, 92–104. [CrossRef]
36. Boucot, A.J.; Xu, C.; Scotese, C.; Morley, R.J. *Phanerozoic Paleoclimate: An Atlas of Lithologic Indications of Climate*; SEPM Concepts Sedimentol. Paleontol. No 11; SEPM Publication Editors: Tulsa, OK, USA, 2013; p. 30. ISBN 978-1-56576-282-4.
37. Davis, W.J. The relationship between atmospheric carbon dioxide concentration and global temperature for the last 425 million years. *Climate* **2017**, *5*, 76. [CrossRef]
38. Berner, R.A. Addendum to "Inclusion of the weathering of volcanic rocks in the GEOCARBSULF model" (R. A. Berner, 2006, v. 306, p. 295–302). *Am. J. Sci.* **2008**, *308*, 100–103. [CrossRef]
39. Ekart, D.D.; Cerling, T.E.; Montanez, I.P.; Tabor, N.J. A 400 million year carbon isotope record of pedogenic carbonate: Implications for paleoatmospheric carbon dioxide. *Am. J. Sci.* **1999**, *299*, 805–827. [CrossRef]
40. International Commission on Stratigraphy. Available online: https://stratigraphy.org/chart (accessed on 1 September 2020).
41. Kodra, E.; Chatterjee, S.; Ganguly, A.R. Exploring Granger causality between global average observed time series of carbon dioxide and temperature. *Theor. Appl. Climatol.* **2011**, *104*, 325–335. [CrossRef]
42. Stern, D.I.; Kaufmann, R.K. Anthropogenic and natural causes of climate change. *Clim. Chang.* **2014**, *122*, 257–269. [CrossRef]
43. Stips, A.; Macias, D.; Coughlan, C.; Garcia-Gorriz, E.; Liang, X.S. On the causal structure between CO_2 and global temperature. *Sci. Rep.* **2016**, *6*, 21691. [CrossRef]
44. Liang, X.S. Information flow and causality as rigorous notions ab initio. *Phys. Rev. E* **2016**, *94*, 052201. [CrossRef] [PubMed]
45. Humlum, O.; Stordahl, K.; Solheim, J.E. The phase relation between atmospheric carbon dioxide and global temperature. *Glob. Planet. Chang.* **2013**, *100*, 51–69. [CrossRef]
46. Richardson, M. Comment on "The phase relation between atmospheric carbon dioxide and global temperature" by Humlum, Stordahl and Solheim. *Glob. Planet. Chang.* **2013**, *107*, 226–228. [CrossRef]
47. Spencer, R.W.; Christy, J.R. Precise monitoring of global temperature trends from satellites. *Science* **1990**, *247*, 1558–1562. [CrossRef] [PubMed]
48. Christy, J.R.; Norris, W.B.; Spencer, R.W.; Hnilo, J.J. Tropospheric temperature change since 1979 from tropical radiosonde and satellite measurements. *J. Geophys. Res.* **2007**, *112*, D06102. [CrossRef]
49. Jones, P.D.; Lister, D.H.; Osborn, T.J.; Harpham, C.; Salmon, M.; Morice, C.P. Hemispheric and large-scale land surface air temperature variations: An extensive revision and an update to 2010. *J. Geophys. Res.* **2012**, *117*, D05127. [CrossRef]
50. Koutsoyiannis, D. Revisiting global hydrological cycle: Is it intensifying? *Hydrol. Earth Syst. Sci.* **2020**. [CrossRef]
51. Keeling, C.D.; Bacastow, R.B.; Bainbridge, A.E.; Ekdahl, C.A.; Guenther, P.R.; Waterman, L.S. Atmospheric carbon dioxide variations at Mauna Loa observatory, Hawaii. *Tellus* **1976**, *28*, 538–551.
52. Global Monitoring Laboratory—Carbon Cycle Greenhouse Gases. Available online: https://www.esrl.noaa.gov/gmd/ccgg/about/global_means.html (accessed on 1 September 2020).
53. Lawrance, A.J. Directionality and reversibility in time series. *Int. Stat. Rev.* **1991**, *59*, 67–79. [CrossRef]
54. Weiss, G. Time-reversibility of linear stochastic processes. *J. Appl. Probab.* **1975**, *12*, 831–836. [CrossRef]
55. Tong, H.; Zhang, Z. On time-reversibility of multivariate linear processes. *Stat. Sin.* **2005**, *15*, 495–504.
56. Koutsoyiannis, D. Generic and parsimonious stochastic modelling for hydrology and beyond. *Hydrol. Sci. J.* **2016**, *61*, 225–244. [CrossRef]
57. Kolmogorov, A.N. Wienersche spiralen und einige andere interessante Kurven im Hilbertschen Raum. *Dokl. Akad. Nauk SSSR* **1940**, *26*, 115–118.
58. Hurst, H.E. Long term storage capacities of reservoirs. *Trans. Am. Soc. Civ. Eng.* **1951**, *116*, 776–808.
59. Heller, M. Time, causality, and the quantum theory. *Rev. Metaphys.* **1983**, *37*, 408–409.
60. Kline, A.D. Are there cases of simultaneous causation? *PSA Proc. Bienn. Meet. Philos. Sci. Assoc.* **1980**, *1980*, 292–301, Philosophy of Science Association. [CrossRef]
61. Suppes, P. *A Probabilistic Theory of Causality*; North-Holland Publishing: Amsterdam, The Netherlands, 1970.
62. Granger, C.W. Testing for causality: A personal viewpoint. *J. Econ. Dyn. Control* **1980**, *2*, 329–352. [CrossRef]

63. Granger, C.; Newbold, P. *Forecasting Economic Time Series*; Academic Press: San Diego, CA, USA, 1986.
64. Gujarati, D.N.; Porter, D.C. *Basic Econometrics*, 5th ed.; McGraw Hill: Boston, MA, USA, 2009.
65. McGraw, M.C.; Barnes, E.A. Memory matters: A case for Granger causality in climate variability studies. *J. Clim.* **2018**, *31*, 3289–3300. [CrossRef]
66. Cohen, M.X. *Analyzing Neural Time Series Data: Theory and Practice*; MIT Press: Cambridge, MA, USA, 2014.
67. Papoulis, A. *Probability, Random Variables and Stochastic Processes*, 3rd ed.; McGraw-Hill: New York, NY, USA, 1991.
68. Paluš, M.; Krakovská, A.; Jakubík, J.; Chvosteková, M. Causality, dynamical systems and the arrow of time. *Chaos Interdiscip. J. Nonlinear Sci.* **2018**, *28*, 75307. [CrossRef] [PubMed]
69. Koutsoyiannis, D.; Yao, H.; Georgakakos, A. Medium-range flow prediction for the Nile: A comparison of stochastic and deterministic methods. *Hydrol. Sci. J.* **2008**, *53*, 142–164. [CrossRef]
70. Zaiontz, C. Real Statistics Using Excel. Available online: http://www.realstatistics.com/ (accessed on 1 September 2020).
71. Zaiontz, C. Real Statistics Examples Workbooks. Available online: http://www.real-statistics.com/free-download/real-statistics-examples-workbook/ (accessed on 1 September 2020).
72. Kundzewicz, Z.W.; Pińskwar, I.; Koutsoyiannis, D. Variability of global mean annual temperature is significantly influenced by the rhythm of ocean-atmosphere oscillations. *Sci. Total Environ.* **2020**, *747*, 141256. [CrossRef]
73. Masters, T.; Benestad, R. Comment on "The phase relation between atmospheric carbon dioxide and global temperature". *Glob. Planet. Chang.* **2013**, *106*, 141–142. [CrossRef]
74. Bond-Lamberty, B.; Thomson, A. Temperature-associated increases in the global soil respiration record. *Nature* **2010**, *464*, 579–582. [CrossRef] [PubMed]
75. IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013; p. 1535. Available online: http://www.climatechange2013.org/report/ (accessed on 14 February 2020).
76. Raich, J.W.; Schlesinger, W.H. The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. *Tellus B Chem. Phys. Meteorol.* **1992**, *44*, 81–99. [CrossRef]
77. Makita, N.; Kosugi, Y.; Sakabe, A.; Kanazawa, A.; Ohkubo, S.; Tani, M. Seasonal and diurnal patterns of soil respiration in an evergreen coniferous forest: Evidence from six years of observation with automatic chambers. *PLoS ONE* **2018**, *13*, e0192622. [CrossRef]
78. ÓhAiseadha, C.; Quinn, G.; Connolly, R.; Connolly, M.; Soon, W. Energy and Climate Policy—An Evaluation of Global Climate Change Expenditure 2011–2018. *Energies* **2020**, *13*, 4839. [CrossRef]
79. Pomeroy, R.; Bowlus, F.D. Progress report on sulfide control research. *Sew. Work. J.* **1946**, *18*, 597–640.
80. Green, C.; Byrne, K.A. Biomass: Impact on carbon cycle and greenhouse gas emissions. In *Encyclopedia of Energy*; Cleveland, C.J., Ed.; Elsevier: Amsterdam, The Netherlands, 2004; pp. 223–236. [CrossRef]
81. Weiss, R. Carbon dioxide in water and seawater: The solubility of a non-ideal gas. *Mar. Chem.* **1974**, *2*, 203–215. [CrossRef]
82. Zhu, Z.; Piao, S.; Myneni, R.B.; Huang, M.; Zeng, Z.; Canadell, J.G.; Ciais, P.; Sitch, S.; Friedlingstein, P.; Arneth, A.; et al. Greening of the earth and its drivers. *Nat. Clim. Chang.* **2016**, *6*, 791–795. [CrossRef]
83. Chen, C.; Park, T.; Wang, X.; Piao, S.; Xu, B.; Chaturvedi, R.K.; Fuchs, R.; Brovkin, V.; Ciais, P.; Fensholt, R.; et al. China and India lead in greening of the world through land-use management. *Nat. Sustain.* **2019**, *2*, 122–129. [CrossRef] [PubMed]
84. Connolly, R.; Connolly, M.; Carter, R.M.; Soon, W. How Much Human-Caused Global Warming Should We Expect with Business-As-Usual (BAU) Climate Policies? A Semi-Empirical Assessment. *Energies* **2020**, *13*, 1365. [CrossRef]
85. Aryal, Y. Review of Atmospheric Temperature and CO_2: Hen-or-Egg Causality? (Version 1). *Sci* **2020**. Available online: https://www.mdpi.com/2413-4155/2/4/77/review_report (accessed on 25 September 2020).
86. Alexandris, S. Review of Atmospheric Temperature and CO_2: Hen-or-Egg Causality? (Version 1). *Sci* **2020**. Available online: https://www.mdpi.com/2413-4155/2/4/77/review_report (accessed on 26 September 2020).
87. Beck, E.-G. 180 Years of atmospheric CO_2 gas analysis by chemical methods. *Energy Environ.* **2007**, *18*, 259–282. [CrossRef]
88. Letts, E.A.; Blake, R.F. The carbonic anhydride of the atmosphere. *Proc. R. Soc. Dublin* **1900**, *9*, 107–270.

89. Stepanova, N.A. A Selective Annotated Bibliography of Carbon Dioxide in the Atmosphere. *Meteorol. Abstr.* **1952**, *3*, 137–170.
90. Keeling, R.F. Comment on “180 Years of Atmospheric CO_2 Gas Analysis by Chemical Methods” by Ernst-Georg Beck. *Energy Environ.* **2007**, *18*, 637–639.
91. Meijer, H.A.J. Comment on “180 Years of Atmospheric CO_2 Gas Analysis by Chemical Methods” by Ernst-Georg Beck. *Energy Environ.* **2007**, *18*, 635–636.
92. Hannart, A.; Pearl, J.; Otto, F.E.L.; Naveau, P.; Ghil, M. Causal counterfactual theory for the attribution of weather and climate-related events. *Bull. Am. Meteorol. Soc.* **2016**, *97*, 99–110. [CrossRef]
93. Verbitsky, M.Y.; Mann, M.E.; Steinman, B.A.; Volobuev, D.M. Detecting causality signal in instrumental measurements and climate model simulations: Global warming case study. *Geosci. Model Dev.* **2019**, *12*, 4053–4060. [CrossRef]
94. Hume, D. *An Enquiry Concerning Human Understanding*; Oxford World’s Classics; Oxford University Press: New York, NY, USA, 2008.
95. Koutsoyiannis, D.; Efstratiadis, A.; Mamassis, N.; Christofides, A. On the credibility of climate predictions. *Hydrol. Sci. J.* **2008**, *53*, 671–684. [CrossRef]
96. Anagnostopoulos, G.G.; Koutsoyiannis, D.; Christofides, A.; Efstratiadis, A.; Mamassis, N. A comparison of local and aggregated climate model outputs with observed data. *Hydrol. Sci. J.* **2010**, *55*, 1094–1110. [CrossRef]
97. Kundzewicz, Z.W.; Stakhiv, E.Z. Are climate models “ready for prime time” in water resources management applications, or is more research needed? *Hydrol. Sci. J.* **2010**, *55*, 1085–1089. [CrossRef]
98. Koutsoyiannis, D.; Christofides, A.; Efstratiadis, A.; Anagnostopoulos, G.G.; Mamassis, N. Scientific dialogue on climate: Is it giving black eyes or opening closed eyes? Reply to “A black eye for the Hydrological Sciences Journal” by D. Huard. *Hydrol. Sci. J.* **2011**, *56*, 1334–1339. [CrossRef]
99. Tsaknias, D.; Bouziotas, D.; Koutsoyiannis, D. Statistical comparison of observed temperature and rainfall extremes with climate model outputs in the Mediterranean region. *ResearchGate* **2016**. [CrossRef]
100. Tyralis, H.; Koutsoyiannis, D. On the prediction of persistent processes using the output of deterministic models. *Hydrol. Sci. J.* **2017**, *62*, 2083–2102. [CrossRef]
101. Koutsoyiannis, D. On the quest for chaotic attractors in hydrological processes. *Hydrol. Sci. J.* **2006**, *51*, 1065–1091. [CrossRef]

Sci

MDPI

Article

On Hens, Eggs, Temperatures and CO_2: Causal Links in Earth's Atmosphere

Demetris Koutsoyiannis [1,*], Christian Onof [2], Zbigniew W. Kundzewicz [3] and Antonis Christofides [1]

1 Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, 15778 Zographou, Greece; anthony@itia.ntua.gr
2 Department of Civil and Environmental Engineering, Faculty of Engineering, Imperial College London, London SW7 2BX, UK; c.onof@imperial.ac.uk
3 Meteorology Lab, Department of Construction and Geoengineering, Faculty of Environmental Engineering and Mechanical Engineering, Poznan University of Life Sciences, 60-637 Poznań, Poland; kundzewicz@yahoo.com
* Correspondence: dk@itia.ntua.gr

Abstract: The scientific and wider interest in the relationship between atmospheric temperature (T) and concentration of carbon dioxide ([CO_2]) has been enormous. According to the commonly assumed causality link, increased [CO_2] causes a rise in T. However, recent developments cast doubts on this assumption by showing that this relationship is of the *hen-or-egg* type, or even unidirectional but opposite in direction to the commonly assumed one. These developments include an advanced theoretical framework for testing causality based on the stochastic evaluation of a potentially causal link between two processes via the notion of the impulse response function. Using, on the one hand, this framework and further expanding it and, on the other hand, the longest available modern time series of globally averaged T and [CO_2], we shed light on the potential causality between these two processes. All evidence resulting from the analyses suggests a unidirectional, potentially causal link with T as the cause and [CO_2] as the effect. That link is not represented in climate models, whose outputs are also examined using the same framework, resulting in a link opposite the one found when the real measurements are used.

Keywords: causality; causal systems; stochastics; impulse response function; geophysics; hydrology; climate

check for updates

Citation: Koutsoyiannis, D.; Onof, C.; Kundzewicz, Z.W.; Christofides, A. On Hens, Eggs, Temperatures and CO_2: Causal Links in Earth's Atmosphere. *Sci* **2023**, *5*, 35. https://doi.org/10.3390/sci5030035

Academic Editors: Anthony R. Lupo, João Miguel Dias, Claus Jacob and Ahmad Yaman Abdin

Received: 17 March 2023
Revised: 24 May 2023
Accepted: 5 September 2023
Published: 13 September 2023

Science is generated by and devoted to free inquiry: the idea that any hypothesis, no matter how strange, deserves to be considered on its merits. The suppression of uncomfortable ideas may be common in religion and politics, but it is not the path to knowledge; it has no place in the endeavor of science. We do not know in advance who will discover fundamental new insights.

Carl Sagan [1]

1. Introduction

A recent (2020) study [2] examining data from measurements of temperature (T) and atmospheric concentration of carbon dioxide ([CO_2]) challenged the conventional, and well established, wisdom that increased [CO_2] causes an increase in temperature. The study examined whether the commonly assumed causal chain is supported by data or, alternatively, whether a *hen-or-egg* (HOE) causal relationship is more plausible. The phrase "*hen or egg*" (originally in Greek, ὄρνις ἢ ᾠόν) was first used in a philosophical context by Plutarch [3] to describe situations in which it is not clear which of two interrelated events or processes is the cause and which the effect.

The study examined a case where the causal link is not between two events but between two processes, represented as stochastic processes. Denoting these processes as

$\underline{x}(t)$ and $\underline{y}(t)$ (where we follow the Dutch notational convention of underlining stochastic variables), in a typical causal system, denoted as $x \rightarrow y$, earlier realizations of $\underline{x}(t)$ affect the current realization of $\underline{y}(t)$. In an HOE causal system, earlier realizations of $\underline{x}(t)$ affect the current realization of $\underline{y}(t)$, but also earlier realizations of $\underline{y}(t)$ affect the current realization of $\underline{x}(t)$.

In terms of its applications, the study used global temperature data from satellites (University of Alabama in Huntsville—UAH) and ground-based (CRUTEM.4.6.0.0 global T 2 m land temperature) and [CO_2] data at several sites (Mauna Loa, HI, USA; Barrow, AK, USA; South Pole; global average) with monthly time steps for the period 1980–2019. An innovative element of this study was that it explained the reasons why using the original T and [CO_2] data series yielded spurious results, and it proposed using the changes (differences in time) thereof instead. We note that differencing is of very common use in economics literature (e.g., [4,5]). In particular, for the [CO_2] it proposed taking the logarithm before differencing (something resembling techniques used in economics [5]) and thus the time series that were correlated were ΔT and $\Delta \ln[CO_2]$, where the differences are taken over 12 months. By studying lagged correlations of the two, the study asserted that, while both causality directions exist, the results support the hypothesis that the dominant direction is $T \rightarrow CO_2$. Changes in [CO_2] follow changes in T by about six months on a monthly scale or about one year on an annual scale. In turn, the study attempted to interpret this mechanism by referring to biochemical reactions, since at higher temperatures soil respiration, and hence CO_2 emission, increases.

In a subsequent (2022) two-paper study, Koutsoyiannis et al. [6,7] developed a more complete theoretical framework by revisiting causality over the entire knowledge tree, from philosophy to science and to scientific and technological application. By reviewing various approaches to causality, the study located several problems in identifying causal links. Hence, the study developed the theoretical background of a stochastic approach to causality, with the objective of formulating necessary conditions that are operationally useful in identifying or falsifying causality claims. It also developed an effective algorithm applicable to large-scale open systems, which are neither controllable nor repeatable. The proposed framework was illustrated and showcased in a number of case studies, some of which were controlled synthetic examples and others real-world ones arising from interesting scientific problems in geophysics and, in particular, hydrology and climatology. The relationship of globally averaged temperature with [CO_2] concentration (again in terms of differences ΔT and $\Delta \ln[CO_2]$ over 12 months) was included in the thirty case studies presented. In brief, the related analyses pointed to the following (quoting from [7]):

> *Clearly, the results* [...] *suggest a (mono-directional) potentially causal system with T as the cause and [CO_2] as the effect. Hence the common perception that increasing [CO_2] causes increased T can be excluded as it violates the necessary condition for this causality direction.*
>
> [...] *in other words, it is the increase of temperature that caused increased CO_2 concentration. Though this conclusion may sound counterintuitive at first glance, because it contradicts common perception* [...], *in fact it is reasonable. The temperature increase began at the end of the Little Ice Period, in the early nineteenth century, when human CO_2 emissions were negligible* [...].

However, the scope of that study [6,7] was to formulate a general methodology for the detection of causality—in particular, necessary conditions thereof—rather than to study a specific system in detail. Therefore, no detailed modeling was made in the case studies, including the hydrological and climatic applications. However, given the enormous interest in the T-[CO_2] relationship, here we will go deeper into this.

Specifically, this paper, after summarizing the methodology (Section 2) and the data used (Section 3), focuses on the latter relationship with the following objectives:

1. To expand the time frame of the investigation backward and forward by exploiting the longest available data series (Section 4).
2. To check whether seasonality, as reflected in different phases of [CO_2] time series at different latitudes, plays any role that could modify or possibly reverse the detected causality relationship (Section 5).
3. To propose and apply a method for investigating the effect of the timescale in causality detection (Section 6).
4. To extend the methodology for disambiguating cases in which the type of causality, HOE or unidirectional, is not quite clear (Section 7).
5. To exploit the methodology in defining a type of data analysis that, regardless of the detection of causality per se, could shed light on modeling performance by comparing observational data with model results (Section 8).
6. To discuss possible extensions of the scope of the methodology, i.e., from detecting possible causality to building a more detailed model of stochastic type (Section 9).
7. To provide logical support for the findings by revisiting the carbon balance in the atmosphere (Appendix A.1) and investigating additional processes that may have caused increases in temperature (Appendices A.2–A.4).

2. Summary of the Stochastic Approach to Causality

The methodology in [6,7] is based on the impulse response function (IRF) between two processes $\underline{x}(t)$, $\underline{y}(t)$, denoted as $g(h)$ where h denotes time lag, based on the convolution

$$\underline{y}(t) = \int_{-\infty}^{\infty} g(h)\underline{x}(t-h)\mathrm{d}h + \underline{v}(t) \tag{1}$$

where $\underline{v}(t)$ is another stochastic process representing the part that is not explained by the causal link. To see that the function $g(h)$ is the impulse response function (IRF) of the system $(\underline{x}(t), \underline{y}(t))$, we set $\underline{v}(t) \equiv 0$ and $\underline{x}(t) = \delta(t)$ (the Dirac delta function, representing an impulse of infinite amplitude at $t = 0$ and attaining the value of 0 for $t \neq 0$), and we readily get $\underline{y}(t) = g(t)$.

On the other hand, if we set $g(h) = b\,\delta(h - h_0)$ (with constant b and h_0), which means that the IRF is zero for every lag except for the specific lag h_0, then Equation (1) becomes $\underline{y}(t) = b\underline{x}(t - h_0) + \underline{v}(t)$. This special case is equivalent to simply correlating $\underline{y}(t)$ with $\underline{x}(t - h_0)$ at any time instance t. It is easy to find (cf. linear regression) that in this case the multiplicative constant b is the correlation coefficient of $\underline{y}(t)$ and $\underline{x}(t - h_0)$ multiplied by the ratio of the standard deviations of the two processes. In general, however, we expect that the actual $g(h)$ is not a Dirac delta function but a continuous one over some domain. Thus, the IRF is a much more powerful tool than correlation as it integrates the correlations in the entire spectrum of lags.

For any two processes $\underline{x}(t)$ and $\underline{y}(t)$, Equation (1) has infinitely many solutions in terms of the function $g(h)$ and the process $\underline{v}(t)$. An obvious and trivial one is $g(h) \equiv 0, \underline{y}(t) \equiv \underline{v}(t)$. The sought solution is the one that corresponds to the minimum variance of $\underline{v}(t)$, called the least-squares solution. Equivalently, this solution maximizes the explained variance ratio:

$$e := 1 - \frac{\gamma_v}{\gamma_y} \tag{2}$$

where γ_v and γ_y denote the variances of the processes $\underline{v}(t)$ and $\underline{y}(t)$, respectively. (This is similar as in the correlation at a single time lag.) If the attained maximum e is close to zero, this will mean that the two processes are uncorrelated and thus no causality condition can exist between them (a non-causal system).

Otherwise, we may assume, without loss of generality, that processes $\underline{x}(t)$ and $\underline{y}(t)$ are positively correlated (i.e., an increase in $\underline{x}(t)$ would result in an increase in $\underline{y}(t)$). In the opposite case (if the processes are negatively correlated), by multiplying one of the two series by -1 we make the correlation positive. Therefore, we impose a nonnegativity constraint for the sought IRF,

$$g(h) \geq 0 \tag{3}$$

In the estimation of IRF, we may also impose a roughness constraint,

$$E \leq E_0 \tag{4}$$

where E is the roughness of the IRF determined in terms of the second derivative of $g(h)$:

$$E := \int_{-\infty}^{\infty} (g''(h))^2 \mathrm{d}h \tag{5}$$

Further justification for the two constraints is provided in [6].

In applications, the continuous time representation is replaced by a discrete time one, the IRF becomes a sequence of values g_j, where j denotes the time lag, the infinite range of the time lag h becomes a finite window of time lag j, specified in the interval $[-J, J]$, the integrals are replaced by sums, and the true values of statistics are replaced by estimates. Furthermore, the roughness E is standardized as

$$\varepsilon := \frac{E}{8\sum_{j=-J}^{J} g_j^2} \tag{6}$$

where ε ranges in (0,1) for nonnegative g_j. The determination of the IRF ordinates g_j is thus formulated as a constrained optimization problem, whose numerical solution is always possible, simple and fast, and can be attained even by commonly available solvers, e.g., in commercial or open source spreadsheet software.

We note that in applications, each of the directions $x \to y$ and $y \to x$ is investigated separately, as there is no symmetry (or antisymmetry) in the produced IRFs in the two directions. When we refer to direction $y \to x$ we mean that we interchange the time series x and y and still estimate the IRF in the same way, as described in our equations in which the direction $x \to y$ is assumed.

In applications investigating causality, we start by assuming a potentially hen-or-egg (HOE) causal model with a fixed number of weights g_j, $j = -J, \ldots, 0, \ldots J$, where in most of the cases studied in [7] the value $J = 20$ was adopted, and this is also followed here. Depending on the results of the estimation procedure, if e is non-negligible, the system is deemed:

- *Potentially HOE causal* if we have $g_j > 0$ for both some positive and some negative lags j,
- *Potentially causal* if $g_j = 0$ for all $j < 0$, and
- *Potentially anticausal* if $g_j = 0$ for all $j > 0$

Note that *anticausal* means that the actual causality direction is opposite to that assumed. These three cases are graphically illustrated in Figure 1. The adverb "potentially" in the above characterizations highlights the fact that the conditions tested provide necessary but not sufficient conditions for causality.

In a potentially causal (or anticausal) system, the time order is explicitly reflected in the above characterizations. In a potentially HOE causal system, the time order needs to be clarified by defining the principal direction. The most natural indices for this are: (a) the time lag h_c maximizing the absolute value of cross-covariance; (b) the mean (time average) μ_h of the function $g(h)$; and (c) the median $h_{1/2}$ of the function $g(h)$. We note that h_c, which was the basis in the original study [2], is completely independent from the $g(h)$. The other two, μ_h and $h_{1/2}$, depend on the $g(h)$ that is determined. However, extensive analyses in [7] showed that their estimation is quite robust; for example, the use of the roughness constraint, while affecting the resulting $g(h)$, practically does not affect the values of μ_h and

$h_{1/2}$. In general, the characteristic lags μ_h and $h_{1/2}$ do not differ substantially from each other, and any of them could be chosen for further use. Here we prefer to note both, as well as h_c, as they all provide useful information about the relationship of the two processes (like in the case of using both mean and median in the characterization of the probability distribution of a single variable).

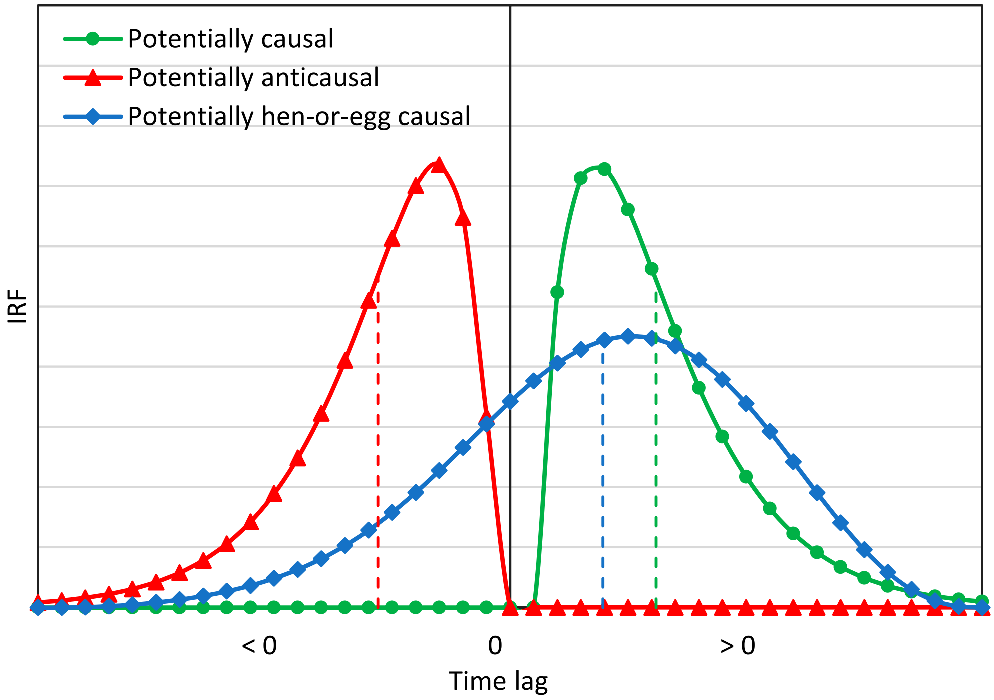


Figure 1. Explanatory sketch for the definition of the different potential causality types. For each graph, the mean μ_h is also plotted with a dashed line.

Needless to say, the literature offers a spectrum of alternative methods for estimating an IRF, using different tools such as auto- and cross-correlations functions [8,9], power spectra and cross-spectra based on the Fourier transform [10] or on a wavelet transform [11], as well as principal component analysis [12]. The method described above has some advantages over these alternatives, as it is a direct method that can work with time series of observations per se, rather than transformations thereof, being easily understandable and reproducible by any reader using simple computational means. Additionally, it is more reliable as it avoids using uncertain estimates of autocorrelation functions or their more uncertain transformations, such as the power spectrum, i.e., the Fourier transform of the autocorrelation function. Note that here we also used autocorrelation, but only to validate and confirm our results—not in the estimation procedure.

Additionally, as detailed in [6], our method differs conceptually and computationally from the so-called "Granger causality" [13,14] (a misnomer, as the method does not infer causality but prediction) and more recently reformulated in the study of Moraffah et al. [15], which also discusses other similar methods. Finally, our method has substantial differences from a framework proposed by Pearl and collaborators [16–18] as again discussed in detail in [6].

3. Data and Case Studies

To explore the T-[CO_2] relationship, case studies #23 and #24 in [7] used satellite-based temperature series (UAH) for the lower troposphere and [CO_2] data from Mauna Loa. Temperature data of the other two satellite levels for the troposphere were also examined, where the results were very similar to those reported for case studies #23 and #24.

Here we present additional case studies, listed in Table 1. In addition to the satellite-based temperature series in [7], here we use surface data (at 2 m) from the NCEP/NCAR Reanalysis 1 by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) [19], which are publicly available. The temporal coverage of the NCEP/NCAR Reanalysis 1 includes data collected four times daily to provide daily and monthly values from 1948 to the present at a horizontal resolution of 1.88° (∼210 km at the equator). It uses a state-of-the-art analysis and forecast system to perform data assimilation using observations and a numerical weather prediction model. The data assimilation and the model used are identical to the global system implemented operationally at NCEP, except for the horizontal resolution. A large subset of the data is available as daily and monthly averages.

Table 1. Main case studies and resulting summary indices. Δt is the time step of differencing; h_c is the time lag maximizing the cross-covariance $c_{yx}(h)$, or equivalently the cross-correlation $r_{yx}(h) := c_{yx}(h)/\sqrt{c_{xx}(0)c_{yy}(0)}$; μ_h is the mean (time average) of the function $g(h)$; $h_{1/2}$ is the median of the function $g(h)$; e is the explained variance ratio; and ε is the roughness ratio. The case studies #1 and #2 are contained in [7] as case studies #23 and #24 and are included in the table only for comparison.

Case System	#	Direction	h_c	μ_h	$h_{1/2}$	$r_{yx}(h_c)$	e	ε
Monthly timescale, varying Δt								
T: UAH; $[CO_2]$: Mauna Loa, 1979–2020 (from [7]), $\Delta t = 1$ year	1	$\Delta T \to \Delta\ln[CO_2]$	5	7.70	6.35	0.48	0.31	$1.3 \times 10^{-5\,*}$
	2	$\Delta\ln[CO_2] \to \Delta T$	–5	–5.67	–5.49	0.48	0.23	$7.3 \times 10^{-4\,*}$
T: NCEP/NCAR; $[CO_2]$: Mauna Loa, 1958–2021, $\Delta t = 1$ year	3	$\Delta T \to \Delta\ln[CO_2]$	8	7.75	6.86	0.56	0.34	$3.1 \times 10^{-4\,*}$
	4	$\Delta\ln[CO_2] \to \Delta T$	–8	−6.31	–6.30	0.56	0.23	$4.4 \times 10^{-3\,*}$
As #3 and #4, $\Delta t = 2$ years	5	$\Delta T \to \Delta\ln[CO_2]$	8	8.19	7.08	0.57	0.31	$3.4 \times 10^{-4\,*}$
	6	$\Delta\ln[CO_2] \to \Delta T$	–8	−6.31	–6.31	0.57	0.21	$4.5 \times 10^{-3\,*}$
As #3 and #4, $\Delta t = 4$ years	7	$\Delta T \to \Delta\ln[CO_2]$	9	10.65	10.32	0.53	0.29	$1.0 \times 10^{-4\,*}$
	8	$\Delta\ln[CO_2] \to \Delta T$	–9	−6.28	–6.28	0.53	0.14	$3.8 \times 10^{-3\,*}$
As #3 and #4, $\Delta t = 8$ years	9	$\Delta T \to \Delta\ln[CO_2]$	8	11.00	11.00	0.47	0.27	$5.6 \times 10^{-5\,*}$
	10	$\Delta\ln[CO_2] \to \Delta T$	–8	−6.55	–6.54	0.47	0.11	$3.6 \times 10^{-3\,*}$
As #3 and #4, $\Delta t = 16$ years	11	$\Delta T \to \Delta\ln[CO_2]$	6	11.74	12.15	0.45	0.31	$3.4 \times 10^{-5\,*}$
	12	$\Delta T \to \Delta\ln[CO_2]$	6	9.98	11.13	0.45	0.33	7.6×10^{-6}
	13	$\Delta\ln[CO_2] \to \Delta T$	–6	−6.33	–6.31	0.45	0.12	$7.7 \times 10^{-3\,*}$
T: NCEP/NCAR; $[CO_2]$: South Pole, 1975–2021, $\Delta t = 1$ year	14	$\Delta T \to \Delta\ln[CO_2]$	10	9.76	8.91	0.40	0.35	$2.0 \times 10^{-4\,*}$
	15	$\Delta\ln[CO_2] \to \Delta T$	–10	–8.51	–8.35	0.40	0.18	$1.1 \times 10^{-3\,*}$
Annual timescale, $\Delta t = 1$ year								
T: CMIP6 mean, SSP2-4.5; $[CO_2]$: SSP2-4.5, 1850–2100, w/o roughness constraint	16	$\Delta T \to \Delta\ln[CO_2]$	0	–3.75	–6.20	0.36	0.90	0.095
	17	$\Delta\ln[CO_2] \to \Delta T$	0	9.95	15.30	0.36	0.15	0.46
As #16 and #17 but for 1850–2021	18	$\Delta T \to \Delta\ln[CO_2]$	0	–6.23	–8.58	0.31	0.72	0.10
	19	$\Delta\ln[CO_2] \to \Delta T$	0	16.18	16.16	0.31	0.10	0.295
As #16 and #17 but with roughness constraint	20	$\Delta T \to \Delta\ln[CO_2]$	0	–3.65	–5.55	0.36	0.84	3.5×10^{-5}
	21	$\Delta\ln[CO_2] \to \Delta T$	0	6.86	1.63	0.36	0.13	7.7×10^{-3}
As #18 and #19 but with roughness constraint	22	$\Delta T \to \Delta\ln[CO_2]$	0	–7.34	–8.99	0.31	0.64	8.3×10^{-5}
	23	$\Delta\ln[CO_2] \to \Delta T$	0	11.26	14.77	0.31	0.13	9.4×10^{-3}

* The roughness was calculated without considering the second derivative at zero.

For CO_2 concentration, in addition to the Mauna Loa data set, which we updated to include the latest measurements of more than one year, we also added the South Pole data set compiled by the US National Oceanic and Atmospheric Administration (NOAA). The measurements began in 1975 and are taken for two cases, flask and in situ, of which we used the former on a mean monthly basis, except in a few cases of missing data where we filled in with in situ data.

Some of the analyses presented here refer to climate model outputs. Here we used the mean (CMIP6 mean) of the output series of the Coupled Model Intercomparison Project (CMIP6) averaged over the globe. The model outputs also go back to the past, extending over the time period of 1850–2100. When we studied past behaviors, we used the data up to 2021, as in the other case studies, but in some cases, we also used the entire data set up to 2100. In the latter case, among the scenarios provided, we used Scenario Shared Socio-Economic Pathways 245 (SSP2-4.5, [20]).

The $[CO_2]$ time series used in climate models for scenario SSP2-4.5 have also been retrieved and analyzed. We note, though, that these time series are given on an annual timescale, unlike all other data that are provided on a monthly scale.

To check whether the results of our methodology would change if we chose any particular member of the ensemble instead of the mean, we also retrieved outputs from a single model, namely the UK Earth System Model (UKESM1 [21]). For the sake of brevity of this paper, we give this latter analysis (whose results eventually do not differ from those of the CMIP6 mean) in the Supplementary Information (and therefore we do not list it in Table 1). The main case studies in which these data were used are summarized in Table 1, along with the summary indices of g_j that are related to potential causality. The details of the case studies are given in the following sections. In all of them, we started by assuming a potentially hen-or-egg (HOE) causal model with a fixed number of weights g_j, namely 41 (i.e., $J = 20$, as in [7]).

Additional case studies examining additional data sets have also been performed but are kept out of the body of the paper and contained in Appendices A.2–A.4. All data used are available online for free and the related links are given in the Data availability section.

4. Investigating the Maximum Time Span That Modern Data Allow

The longest time series of systematic $[CO_2]$ measurements is that of Mauna Loa, which began in 1958. The global temperature at 2 m of the NCEP/NCAR reanalysis series goes back to 1948 and thus allows studying the T-$[CO_2]$ relationship for the period 1958–2022 (two additional decades of data in comparison to those studied in [7]).

As seen in Table 1, this provided better characteristics than the UAH/Mauna Loa case examined in [7]: maximum cross correlation $r_{yx}(h_c) = 0.56$ against 0.48; explained variance e = 34% against 31%, at time lags greater than or equal to those in [7] (close to 8 months). As seen in Figure 2, again, we have a potentially causal system with the directionality being $\Delta T \to \Delta\ln[CO_2]$, while the possible causality $\Delta\ln[CO_2] \to \Delta T$ can be excluded as not satisfying the necessary condition of time precedence.

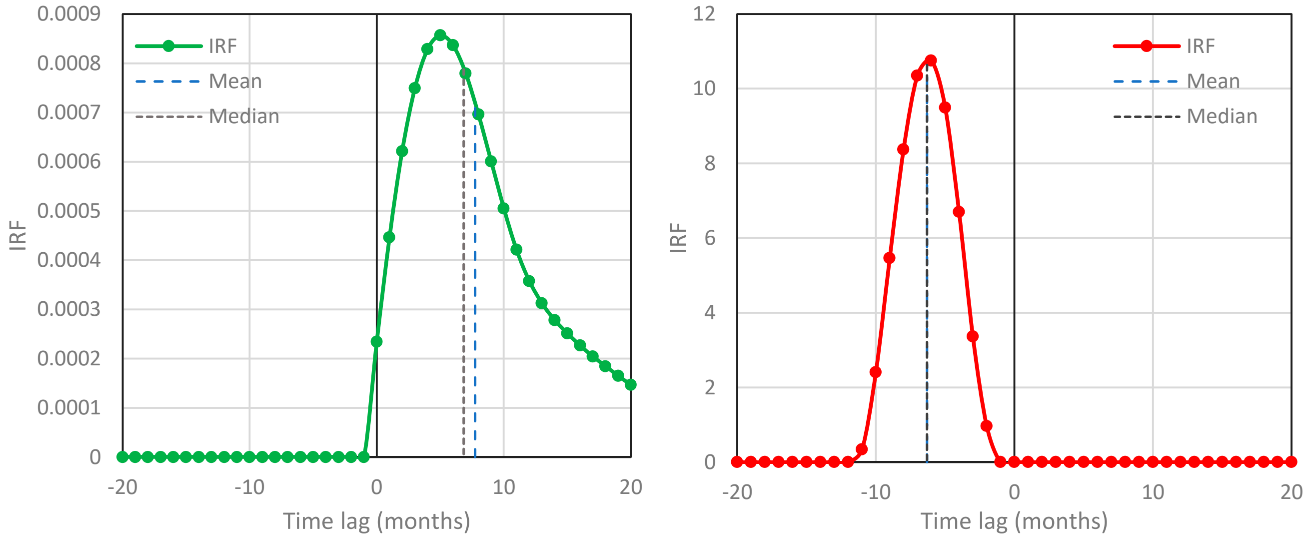


Figure 2. IRFs for temperature–CO_2 concentration based on the NCEP/NCAR Reanalysis temperature at 2 m and Mauna Loa $[CO_2]$ time series, respectively—case studies #3 (**left**; $\Delta T \to \Delta\ln[CO_2]$; potentially causal system) and #4 (**right**; $\Delta\ln[CO_2] \to \Delta T$; potentially anticausal system).

5. Investigating the Possible Effect of Seasonality

To enrich our results and also check whether seasonality, as reflected in different phases of [CO_2] time series at different latitudes, could modify or possibly reverse the detected causality relationship, we have conducted an additional analysis with the South Pole [CO_2] measurements, which began in 1975.

As seen in Table 1, this again provided better characteristics than the UAH/Mauna Loa case examined in [7] in terms of explained variance (e = 35% against 31%) and time lags higher than in [7] (close to 10 months). As seen in Figure 3, again, we have a potentially causal system with the directionality being $\Delta T \rightarrow \Delta\ln[\mathrm{CO_2}]$, while again the possible causality $\Delta\ln[\mathrm{CO_2}] \rightarrow \Delta T$ can be excluded as not satisfying the necessary condition of time precedence.

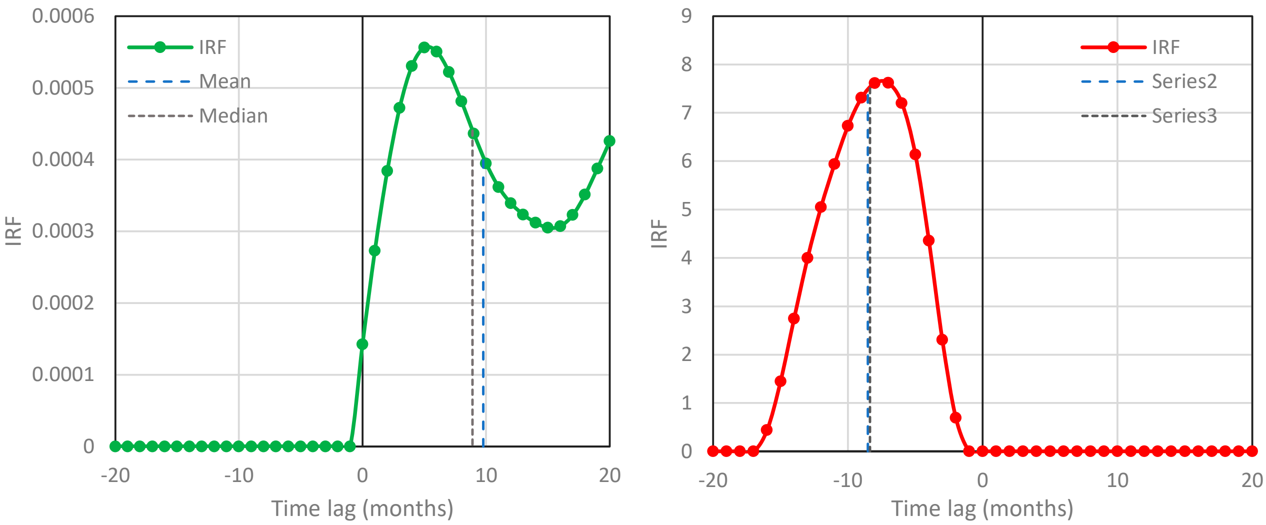


Figure 3. IRFs for temperature–CO_2 concentration based on the NCEP/NCAR Reanalysis temperature at 2 m and the South Pole time series, respectively—case studies #14 (**left**; $\Delta T \rightarrow \Delta\ln[\mathrm{CO_2}]$; potentially causal system) and #15 (**right**; $\Delta\ln[\mathrm{CO_2}] \rightarrow \Delta T$; potentially anticausal system).

Summarizing the two case studies in Sections 4 and 5—and similarly to what we found in [7]—we note that:

- The system T-[CO_2] appears to be potentially causal (unidirectional) in the direction $\Delta T \rightarrow \Delta\ln[\mathrm{CO_2}]$, rather than hen-or-egg causal.
- All characteristic time lags ($h_c, \mu_h,\ h_{1/2}$) are positive in the direction $\Delta T \rightarrow \Delta\ln[\mathrm{CO_2}]$ (ranging from about 7 to about 10 months), and negative in the direction $\Delta\ln[\mathrm{CO_2}] \rightarrow \Delta T$.
- The explained variance ratio is greater in the direction $\Delta T \rightarrow \Delta\ln[\mathrm{CO_2}]$ (about 1/3) than in the direction $\Delta\ln[\mathrm{CO_2}] \rightarrow \Delta T$ (around 1/5).

6. On the Timescale of Validity of Results

Overall, our results in this paper are those allowed by the available data at the time periods and timescales resolved by those data—more than 6 decades at the monthly scale. What would happen at other times—or if the data sets were longer and would resolve intermediate or even longer timescales—we cannot tell. The climate system is too complex to allow for hasty generalizations.

One may not exclude the case that the timescale of analysis is important in a detected potential causality relationship and that the latter would change if the timescale changed. This raises the question: up to which timescale does the validity of certain results hold? Certainly, this timescale is a fraction (not greater than 1/2) of the length of the time series. An indication can be obtained by inspecting the empirical cross-covariance function and how this compares with the theoretical one. As shown in [6], the latter ($c_{yx}(h)$:=

$\text{cov}\left[\underline{y}(t+h), \underline{x}(t)\right]$) for lag h, is related to the autocovariance function of $\underline{x}$, $c_{xx}(h) := \text{cov}[\underline{x}(t+h), \underline{x}(t)]$, by:

$$c_{yx}(h) = \int_{-\infty}^{\infty} g(a) c_{xx}(h-a) \mathrm{d}a \tag{7}$$

Thus, $c_{yx}(h)$ can be estimated from the IRF and the empirical $c_{xx}(h)$ from the data. Figure 4 shows the autocorrelation and cross-correlation functions (which are covariances standardized by standard deviations). For the $\Delta T \to \Delta\ln[\mathrm{CO_2}]$ case (i.e., case study #3; upper panels of Figure 4), the reconstructed cross-correlation function, calculated from the IRF (seen in Figure 2) and the empirical autocorrelation function of temperature (seen in the upper left panel of Figure 4) using the discretized version of equation (7), agrees well with the empirical function for time lags up to ± 10 years, i.e., spanning 20 years (1/3 of the series length). As time lag has an equivalence relationship with timescale [22], we may conclude that the potential causality relationship detected is valid for timescales of two decades. For comparison, the lower panels of Figure 4 show the reverse case, $\Delta\ln[\mathrm{CO_2}] \to \Delta T$, (case study #4), where there is no longer good agreement between empirical and reconstructed cross-correlation, which provides additional support for the claim that the true causality link is $\Delta T \to \Delta\ln[\mathrm{CO_2}]$.

A final observation in Figure 4 is the appearance of six peaks in 20 years, which may be interpreted as indicating a quasi-periodic behavior with an average period of 3.33 years, i.e., much different from the annual. However, this does not reflect periodicity but rather antipersistence imposed by the differencing operation, which necessarily results in some negative autocorrelations [23].

A more direct technique to deal with timescales is to average the time series on aggregate timescales and again investigate the causality relationships on these scales. This technique could detect whether a similar relationship holds for the larger timescales. In our case, since we are differencing the process, taking the average at a timescale k is equivalent to taking a difference for a time step k (notice that $(x_2 - x_1) + (x_3 - x_2) + \ldots + (x_{k+1} - x_k) = x_{k+1} - x_1$). Therefore, to increase the timescales it suffices to increase the time step of differencing. In Figure 2, this was 1 year. Now we increase the time step of differencing, replacing the 1-year step with 2, 4, 8 and 16 years. The results are shown in Figure 5 and in Table 1 (case studies #5 to #13). They are essentially the same, except that, as the differencing time step increases, the explained variance slightly decreases (from 0.34 down to 0.27) in the direction $\Delta T \to \Delta\ln[\mathrm{CO_2}]$, and is again much higher than that in the direction $\Delta\ln[\mathrm{CO_2}] \to \Delta T$. The time lags increase in the former direction and are again negative in the latter direction.

A characteristic pattern is that, as the time step increases, so does the rightmost ordinate of the IRF, g_{20}. This behavior, i.e., the increasing limb of g_j beyond some time lag, has been explained in [7] and is an artifact of an insufficient (small) window of time lag for determining IRF. Thus, it suggests that a higher J should be used. It is quite reasonable to expect that if the differencing time step increases, so should the window size do. In particular, it is interesting to observe that in the lowest panel of Figure 5, corresponding to a differencing time step of 16 years, while the time window of IRF is only 40 months (a little more than 3 years), the IRF is a monotonously increasing function. Apparently, this is an artifact due to the too small window of time lag. In such a case, it is also reasonable to expect some nonnegative estimates of IRF ordinates for negative lags, even if the system is unidirectionally causal. Indeed, this has appeared in the case of a differencing time step of 16 years, even though the lowest panel of Figure 5 shows the purely unidirectional version of the IRF. This may create some ambiguity in the identification of causality, which we analyze and resolve in the next section.

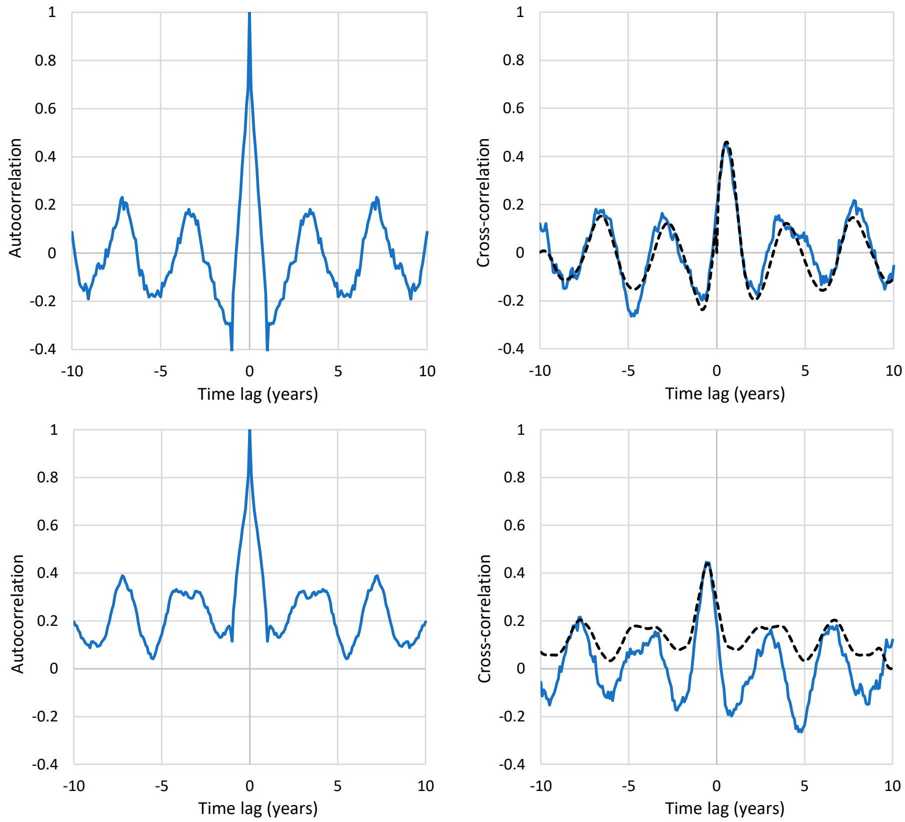


Figure 4. (**Left column**) Empirical autocorrelation function for the period 1958–2021 and for monthly timescale of (**upper**) the NCEP/NCAR ΔT time series and (**lower**) the $\Delta\ln[CO_2]$ time series for Mauna Loa. (**Right column**) Empirical cross-correlation function of the two time series (continuous lines in blue), compared with those reconstructed from the IRF and the autocorrelation function on the left panel using the discretized version of Equation (7) (dashed line), for case studies (**upper**) #3 and (**lower**) #4.

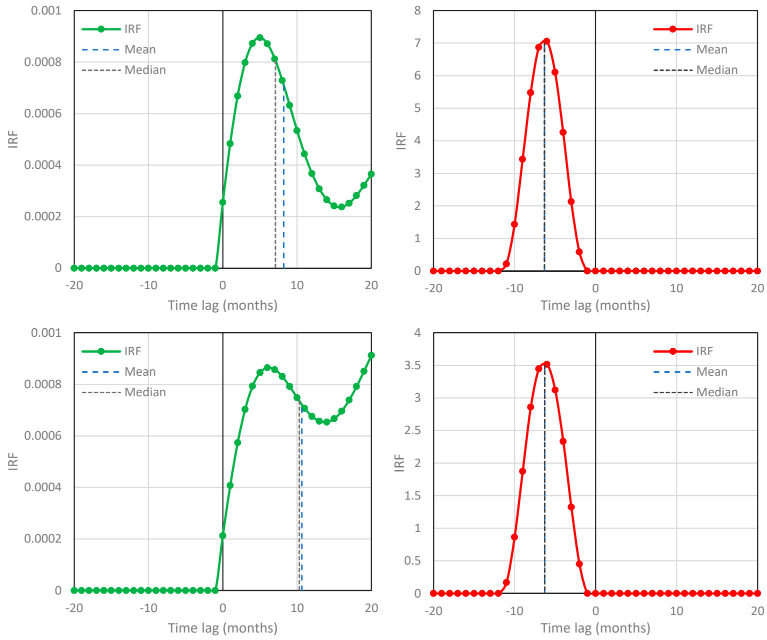


Figure 5. *Cont.*

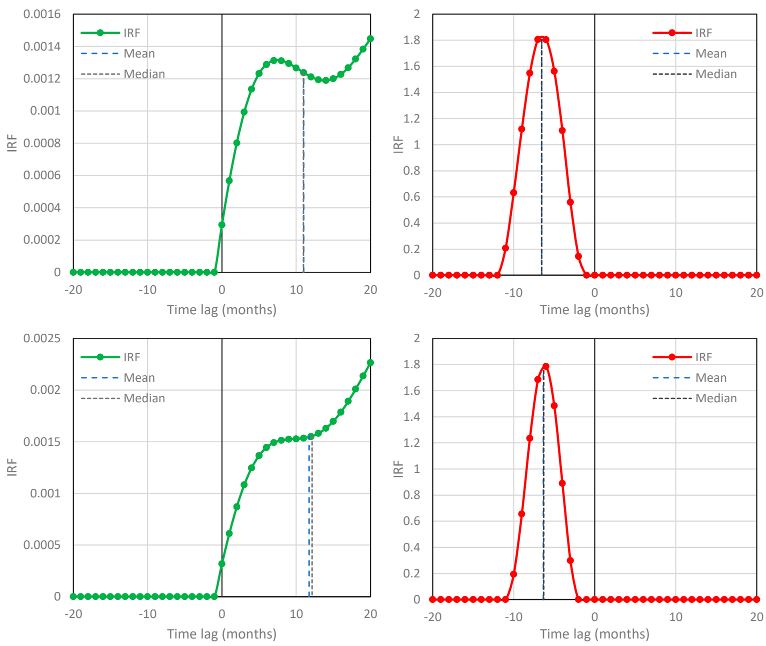


Figure 5. IRFs for temperature–CO_2 concentration based on the NCEP/NCAR Reanalysis temperature at 2 m and Mauna Loa time series, respectively, as in Figure 2, but for differencing time steps equal (from upper to lower) 2, 4, 8 and 16 years; **left**: $\Delta T \rightarrow \Delta \ln[CO_2]$ (potentially causal system); **right**: $\Delta \ln[CO_2] \rightarrow \Delta T$ (potentially anticausal system).

7. Possible Ambiguities and Disambiguation

In some applications, the detection of the type of causality, unidirectional or HOE, is direct, but in other cases, it may be more difficult. This is illustrated in Figure 6. The left panel is equivalent to that of Figure 2 (left) but not neglecting some very small ordinates g_j that originally the algorithm produced for negative j. The right panel is equivalent to that of Figure 5 (bottom left) but letting the optimization algorithm produce g_j for negative j, which in this particular case seem not negligible. Both figure panels refer to the same processes with the differencing time step being 1 and 16 years for the left and right panels, respectively. For the 16-year step, in comparison to the IRF of Figure 5, which explains a fraction 0.31 of the variance, that of Figure 6 yields a slightly higher explained variance, 0.325, while it has some small weights at negative lags. Should we conclude then that it indicates a potential HOE, rather than unidirectional, causality? Even if our reply is affirmative, it is important to note that the characteristic lags are again positive, suggesting a principal direction $\Delta T \rightarrow \Delta \ln[CO_2]$.

However, the reply is not necessarily affirmative. The explained variance of 0.31 is associated with 21 IRF ordinates, while that of 0.325 is associated with 41 IRF ordinates. Is it reasonable to accept 20 additional parameters for an increase in the explained variance of 0.015?

Arguably, it is more reasonable in this case to change from a symmetric to a non-symmetric window of time lag. Hence, we use a window of length 21 and slide it so that the lowest nonzero time lag vary from -20 to 0. Only the case where this is 0 denotes a potential unidirectional causality, where all 20 other cases correspond to potential HOE causality. The resulting IRFs are plotted in Figure 7, while the fraction of explained variance is plotted in Figure 8 as a function of the lowest nonzero time lag. It may be noticed that the curve for the lowest time lag 0 is different from Figure 5 (bottom left). In particular, the ordinate at 0 is higher in Figure 7, thus producing a concave shape of IFR. This is a result of the fact that the window length is fixed, while the lowest ordinate no longer contributes to the roughness (no second derivative can be determined for the lowest point and thus the algorithm can increase the ordinate at 0 at no cost). In turn, the explained variance further increased to 0.327.

Clearly, the result of this investigation is a unidirectional, rather than HOE, potential causality, as the explained variance reaches its maximum when the lowest j is 0.

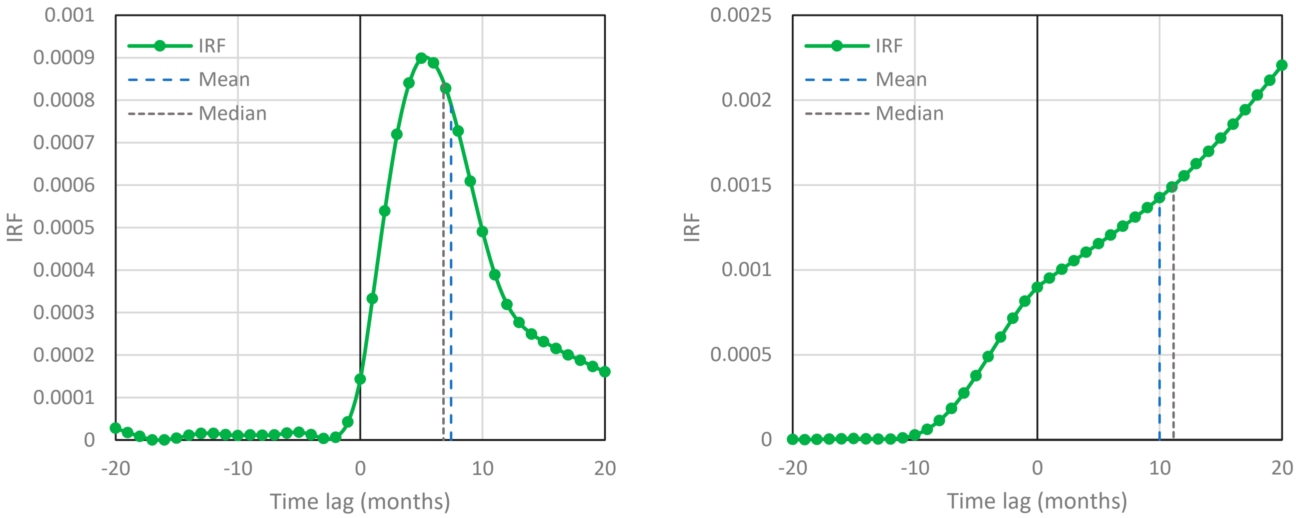


Figure 6. IRFs for temperature–CO_2 concentration based on the NCEP/NCAR Reanalysis temperature at 2 m and Mauna Loa time series, respectively, also enabling negative lags (HOE) for causality direction $\Delta T \rightarrow \Delta \ln[CO_2]$ and for differencing time step of 1 year (**left**, corresponding to Figure 2, left) and 16 years (**right**, corresponding to Figure 5, bottom left).

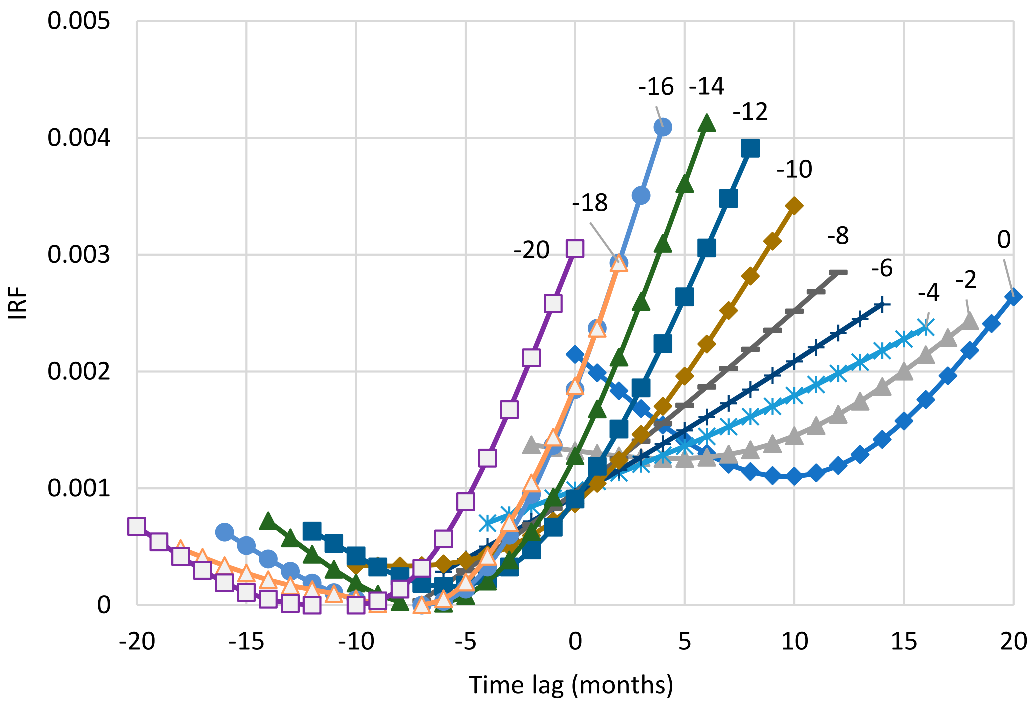


Figure 7. IRFs for temperature–CO_2 concentration based on the NCEP/NCAR Reanalysis temperature at 2 m and Mauna Loa time series, for 21 time lags, differencing time step of 16 years and direction $\Delta T \rightarrow \Delta \ln[CO_2]$. The lowest nonzero lag of each IRF is marked at the upper-right end of its curve.

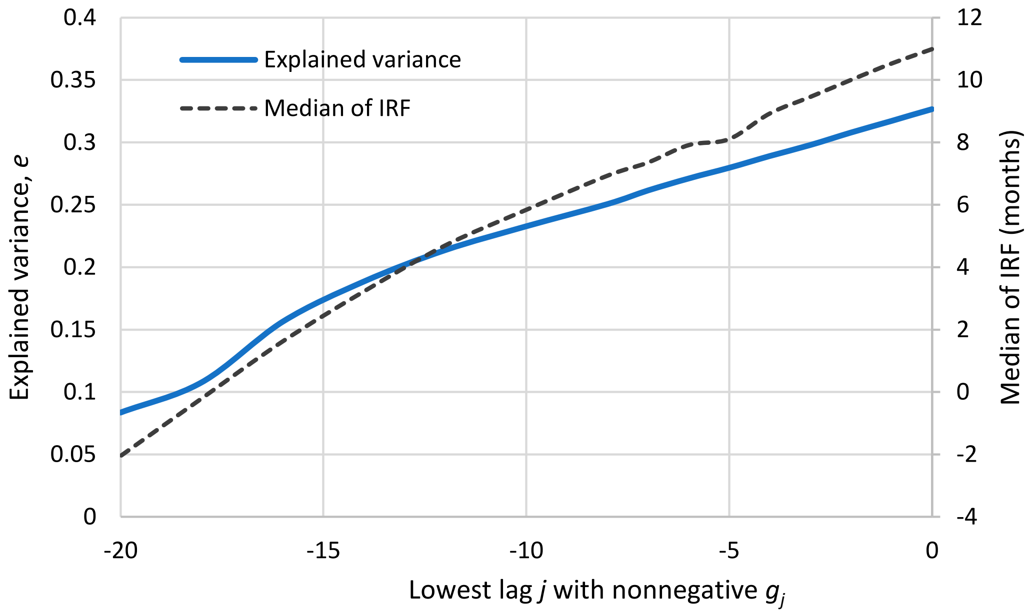


Figure 8. Explained variance and median of IRF as a function of the lowest nonzero lag of the IRFs in Figure 7 for the investigation of Section 7.

8. Comparing Observational Data with Model Results

Investigation of causality in systems that can be isolated from the environment is generally based on experiments. These are typically performed in laboratories and presuppose control actions (intervention) by the experimenter. In the absence of such intervention, it has long been regarded that we cannot say what causes what (Strotz and Wold, 1960 [24]). In complex systems, such as Earth's climatic system, experimentation is impossible. Yet it is a widespread belief that climate models are faithful representations of the climatic system and hence offer the possibility of so-called *in silico experimentation* (Hannart et al. [25]). Furthermore, it has been claimed (Hannart and Naveau [26]) that "in silico *experimentation* [is] *the only option*" and that "*the increasing realism of climate system models renders such an* in silico *approach plausible*". Such claims are epistemologically problematic. A hypothetical "causality" that is incorporated in any model, particularly of a complex system, is not necessarily identical to natural causality. In addition, the agreement of climate model outputs with reality has been questioned (e.g., [27–31]).

Our methodology can help with this epistemological problem in two ways. First, it provides a different option to test causality, showing that the so-called in silico experimentation is not the only option as claimed. Second, it can additionally test whether there is indeed realism in the representation of causality of the climatic system by the climate models. As already stated in the Introduction, our methodology, regardless of the detection of causality per se, can define a type of data analysis that could shed light on modeling performance by comparing observational data with model results. This is particularly useful in the case of climate modeling. In other words, it could help in verifying or falsifying the commonly accepted theory, which is incorporated in the climate models.

Specifically, we can test whether the climate model results are consistent with the findings of our T-[CO_2] causality analysis, which is based on measurements. To this aim, we use climate model outputs as specified in Section 3, in the case studies #16 to #23 detailed in Table 1. Numerical results of our analysis are shown in Table 1, and graphical depictions of IRFs are shown in Figure 9 for the cases in which no roughness constraint is used and in Figure 10 for the cases in which the roughness constraint is used.

Unfortunately, and unlike the [CO_2] time series of measurements, which are available on a monthly scale, the SSP2 [CO_2] data series is provided on an annual scale. Therefore, case studies #16 to #23 had to be made on an annual scale. If we did the analysis for the period 1958–2021, as in case studies #3 and #4 (NCEP/NCAR Reanalysis temperature at

2 m; and Mauna Loa [CO_2] time series), the annual data would be too few to support estimation of the IRFs (63 data values to estimate 41 coefficients). Therefore, in case studies #16 to #23 we extended the period back to 1850, which is covered by the climate model outputs. We performed separate analyses for the periods 1850–2100 (entire period covered by climate models) and 1850–2021 (only the past).

The results of these case studies allow us to make the following observations.

1. There is no essential difference between the results for the periods 1850–2100 and 1850–2021.
2. While, as expected, the IRFs differ if they are calculated with or without constraining roughness, the characteristic lags are similar in the two cases (with the exception of $h_{1/2}$ in cases #17 and #21).
3. In all cases, the lags are negative in the direction $\Delta T \to \Delta \ln[CO_2]$ and positive in the direction $\Delta \ln[CO_2] \to \Delta T$, suggesting a HOE causality with principal direction $\Delta \ln[CO_2] \to \Delta T$.
4. Clearly, the finding in point 3, resulting from climate model outputs, is opposite to the results found when real measurements are used (NCEP/NCAR Reanalysis temperature and Mauna Loa [CO_2] time series).
5. Oddly, while the principal direction suggested by the models is $\Delta \ln[CO_2] \to \Delta T$, the explained variance is impressively low (10–15%) in this direction and impressively high (reaching 90%) in the opposite direction, $\Delta T \to \Delta \ln[CO_2]$.

One may argue that the main result of this analysis, i.e., the point 4 above, may be affected by the difference in the study periods, i.e., 1958–2021 for the real measurements and 1850–2021 for the model outputs. To examine whether the origin of the different behavior is the time period or the system dynamics (actual vs. modeled), we performed an additional analysis, graphically depicted in Figures 11 and 12.

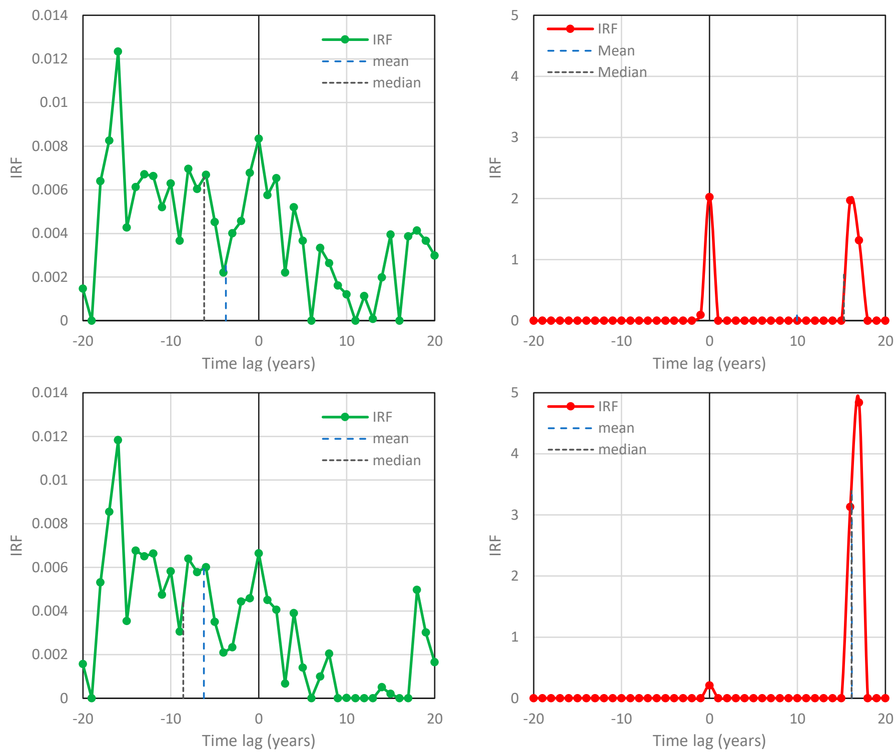


Figure 9. IRFs for temperature–CO_2 concentration based on the CMIP6 mean temperature and SSP2-4.5 CO_2 time series, respectively, calculated without using the roughness constraint; **upper row**: period 1850–2100—case studies #16 (**left**; $\Delta T \to \Delta \ln[CO_2]$) and #17 (**right**; $\Delta \ln[CO_2] \to \Delta T$); **lower row**: period 1850–2021—case studies #18 (**left**; $\Delta T \to \Delta \ln[CO_2]$) and #19 (**right**; $\Delta \ln[CO_2] \to \Delta T$).

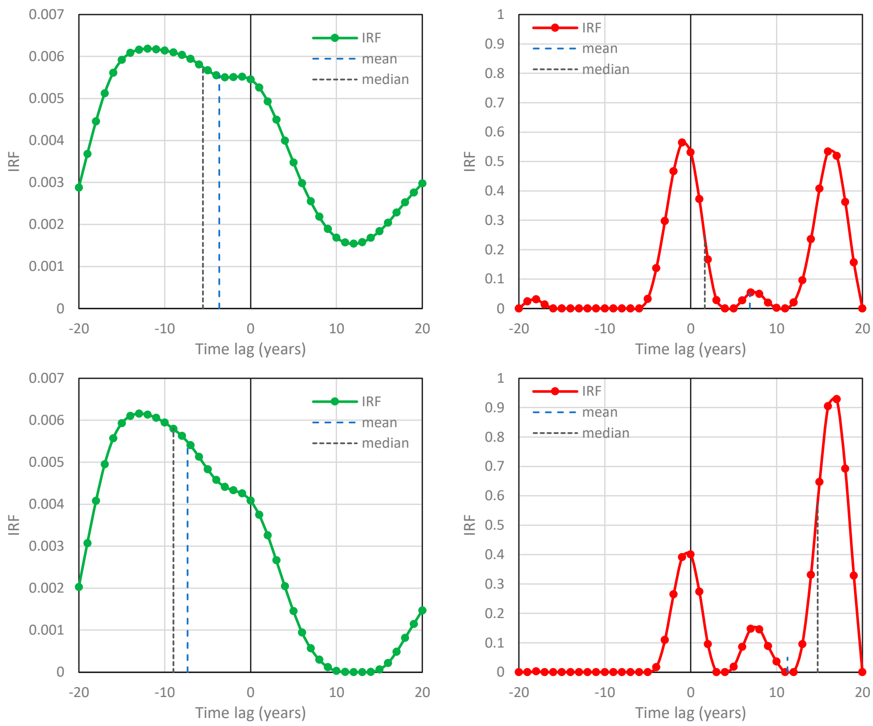


Figure 10. IRFs for temperature–CO_2 concentration based on the CMIP6 mean temperature and SSP2-4.5 CO_2 time series, respectively, as in Figure 9 but calculated using the roughness constraint; **upper row**: period 1850–2100—case studies #20 (**left**; $\Delta T \rightarrow \Delta\ln[CO_2]$) and #21 (**right**; $\Delta\ln[CO_2] \rightarrow \Delta T$); **lower row**: period 1850–2021—case studies #22 (**left**; $\Delta T \rightarrow \Delta\ln[CO_2]$) and #23 (**right**; $\Delta\ln[CO_2] \rightarrow \Delta T$).

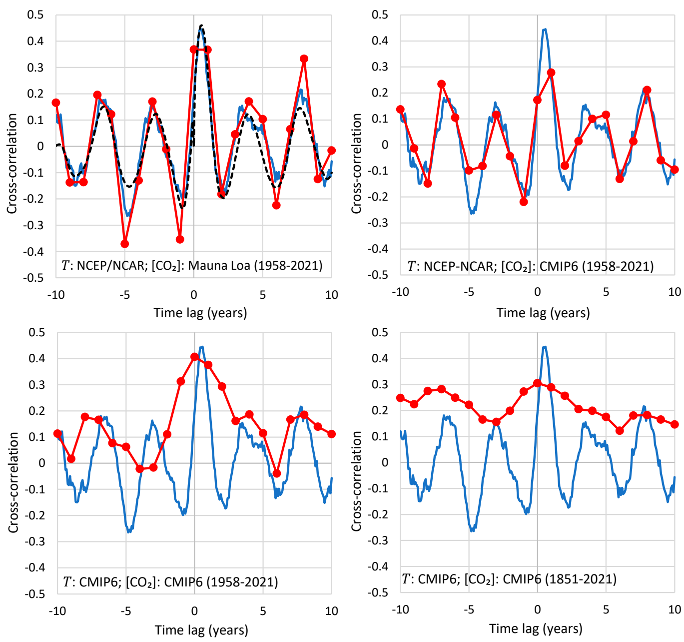


Figure 11. Empirical cross-correlation functions for monthly and annual timescales (continuous lines in blue without markers and red lines with circles, respectively) for the data sets indicated in each panel. In all panels, the plot for the monthly scale is that of the NCEP/NCAR data for T and Mauna Loa data for $[CO_2]$, for the period 1958–2021. The upper-left panel also shows the cross-correlation function reconstructed from the IRF and the autocorrelation function using the discretized version of Equation (7) (dashed line).

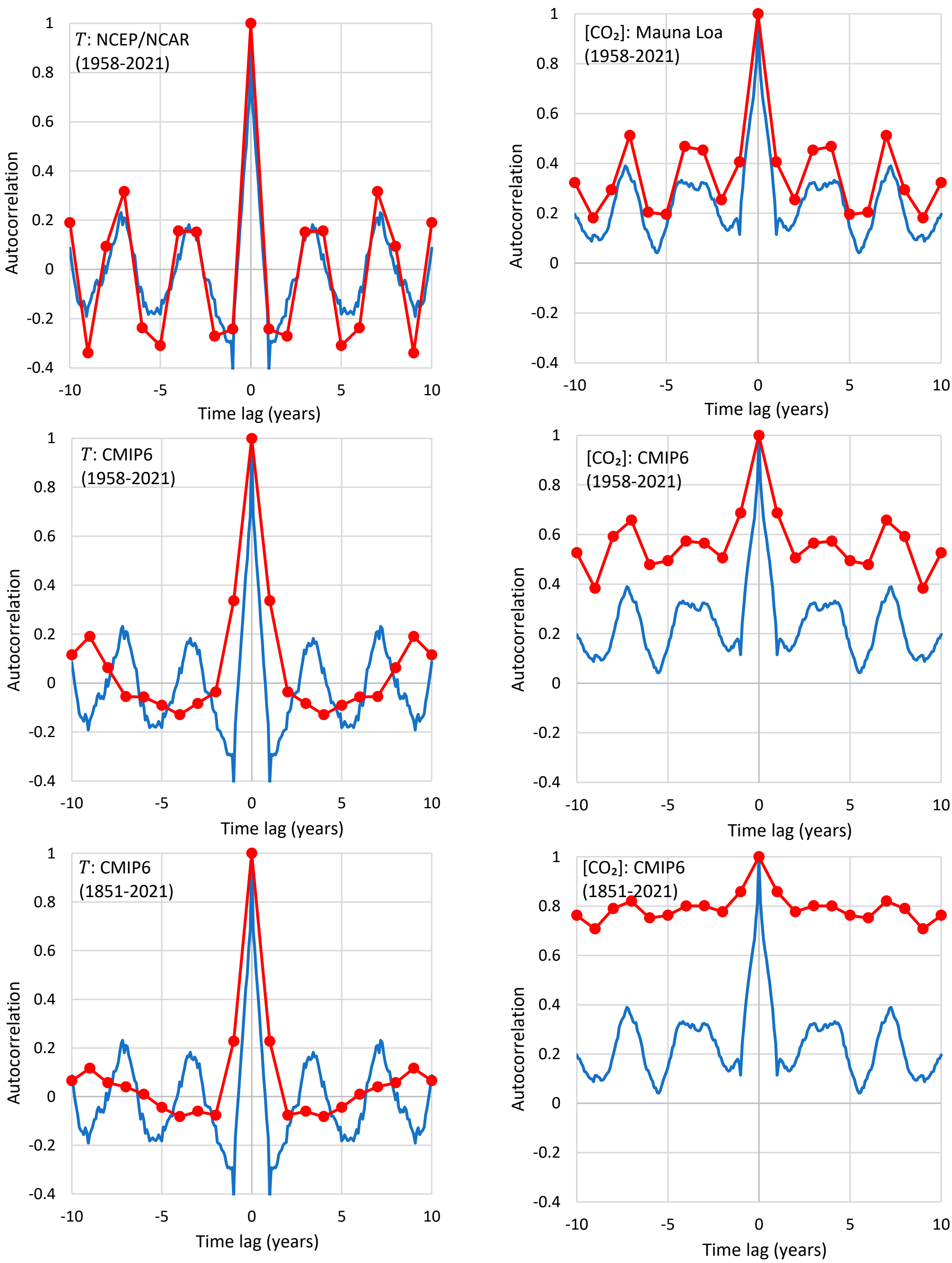


Figure 12. Empirical autocorrelation functions for monthly and annual timescales (continuous lines in blue without markers and red lines with circles, respectively) for the data sets indicated in each panel. In all panels, the plot for the monthly scale is that of the NCEP/NCAR data for T and Mauna Loa data for $[CO_2]$, for the period 1958–2021.

The upper left panel of Figure 11 is similar to that in Figure 4 (upper right), where, in addition, we have plotted the empirical cross-correlation function for the same data but averaged at the annual scale. This agrees relatively well with the empirical function at the monthly scale, which provides a basis for the comparisons that follow.

The empirical cross-correlation function at the monthly scale is copied in all other panels of Figure 11 and serves as a basis for the comparisons. In the upper right panel, where we replaced the Mauna Loa time series for [CO_2] with the CMIP6 time series for the same period, 1958–2021, while keeping the NCEP/NCAR time series for T, there is still agreement of the annual with the monthly cross-correlations. However, when we also replace the NCEP/NCAR time series for T with the CMIP6 series (lower left plot), the two plots decouple. The decoupling is even more prominent if we go to the longer period 1850–2021 (lower right panel).

Similar observations can be made about autocorrelations in Figure 12. In particular, the CMIP6 autocorrelation of T decouples from the actual one for both periods 1958–2021 and 1850–2021, while the two latter do not differ substantially from each other. These observations allow us to assert that the main cause of disagreement has to do with problems with the modeling of system dynamics rather than the time period of study.

9. Discussion and Further Results

The mainstream assumption of the causality direction [CO_2] $\rightarrow T$ makes a compelling narrative, as everything is blamed on a single cause, the human CO_2 emissions. Indeed, this has been the popular narrative for decades. However, popularity does not necessarily mean correctness, and here we have provided strong arguments against this assumption. Since we have identified atmospheric temperature as the cause and atmospheric CO_2 concentration as the effect, one may be tempted to ask the question: What is the cause of the modern increase in temperature? Apparently, this question is much more difficult to reply to, as we can no longer attribute everything to any single agent.

We do not claim to have the answer to this question, whose study is far beyond the article's scope. Neither do we believe that mainstream climatic theory, which is focused upon human CO_2 emissions as the main cause and regards everything else as feedback of the single main cause, can explain what happened on Earth for 4.5 billion years of changing climate.

Nonetheless, as a side product, in the Appendices to the paper, we provide several indications of the following:

1. The dependence of the carbon cycle on temperature is quite strong and indeed major increases of [CO_2] can emerge as a result of temperature rise. In other words, we show that the natural [CO_2] changes due to temperature rise are far larger (by a factor > 3) than human emissions (Appendix A.1).
2. There are processes, such as the Earth's albedo (which is changing in time as any other characteristic of the climate system), the El Niño–Southern Oscillation (ENSO) and the ocean heat content in the upper layer (represented by the vertically averaged temperature in the layer 0–100 m), which are potential causes of the temperature increase, unlike what is observed with [CO_2], their changes precede those of temperature (Appendices A.2–A.4).
3. On a large timescale, the analysis of paleoclimatic data supports the primacy of the causal direction $T \rightarrow$ [CO_2], even though some controversy remains about this issue (Appendix A.5).

In terms of the carbon cycle (point 1 above), several physical, chemical, biochemical and human processes are involved in it. The human CO_2 emissions due to the burning of fossil fuels have largely increased since the beginning of the industrial age. However, the global temperature increase began succeeding the Little Ice Period, at a time when human CO_2 emissions were very low. To cast light on the problem, we examine the issue of CO_2 emissions vs. atmospheric temperature further in the Supplementary Information, where we provide evidence that they are not correlated with each other. The outgassing from the sea is also highlighted sometimes in the literature among the climate-related mechanisms. On the other hand, the role of the biosphere and biochemical reactions is often downplayed, along with the existence of complex interactions and feedback. This role can be summarized in the following points, examined in detail and quantified in Appendix A.1.

- Terrestrial and maritime respiration and decay are responsible for the vast majority of CO_2 emissions [32], Figure 5.12.
- Overall, natural processes of the biosphere contribute 96% to the global carbon cycle, the rest, 4%, being human emissions (which were even lower in the past [33]).
- The biosphere is more productive at higher temperatures, as the rates of biochemical reactions increase with temperature, which leads to increasing natural CO_2 emission [2].
- Additionally, a higher atmospheric CO_2 concentration makes the biosphere more productive via the so-called carbon fertilization effect, thus resulting in greening of the Earth [34,35], i.e., amplification of the carbon cycle, to which humans also contribute through crops and land-use management [36].

In addition to the biosphere, there are other factors that drive the Earth's climate in periodic and non-periodic way. Orbital parameters of Earth's revolution change quasi-cyclically in a multi-millennial scale (variations in eccentricity, axial tilt, and precession of Earth's orbit), as interpreted by Milanković [37–41], and changes in the orbit geometry influence the amount of insolation. The non-periodic drivers of the Earth's climate variability include volcanic eruptions and collisions with large extraterrestrial objects, e.g., asteroids. An important climate driver is water in its three phases [33]. Another apparent factor is solar activity (including solar cycles) and the solar radiation (im)balance on Earth (e.g., albedo changes; see [33] and Appendix A.2). Notably, a recent study [42], by assessing 20 years of direct observations of energy imbalance from Earth-orbiting satellites, showed that the global changes observed appear largely from reductions in the amount of sunlight scattered by Earth's atmosphere.

ENSO and ocean heating, both of which affect temperature, are examined in Appendices A.3 and A.4, respectively. The results of Appendices A.2–A.4 are summarized in the schematic of Figure 13. Changes in all three examined processes, albedo, ENSO and the upper ocean heat, precede in time the changes in temperature and even more so those in $[CO_2]$. Generally, the time lags shown in Figure 13 complete a consistent picture of potential causality links among climate processes and always confirm the $T \rightarrow [CO_2]$ direction.

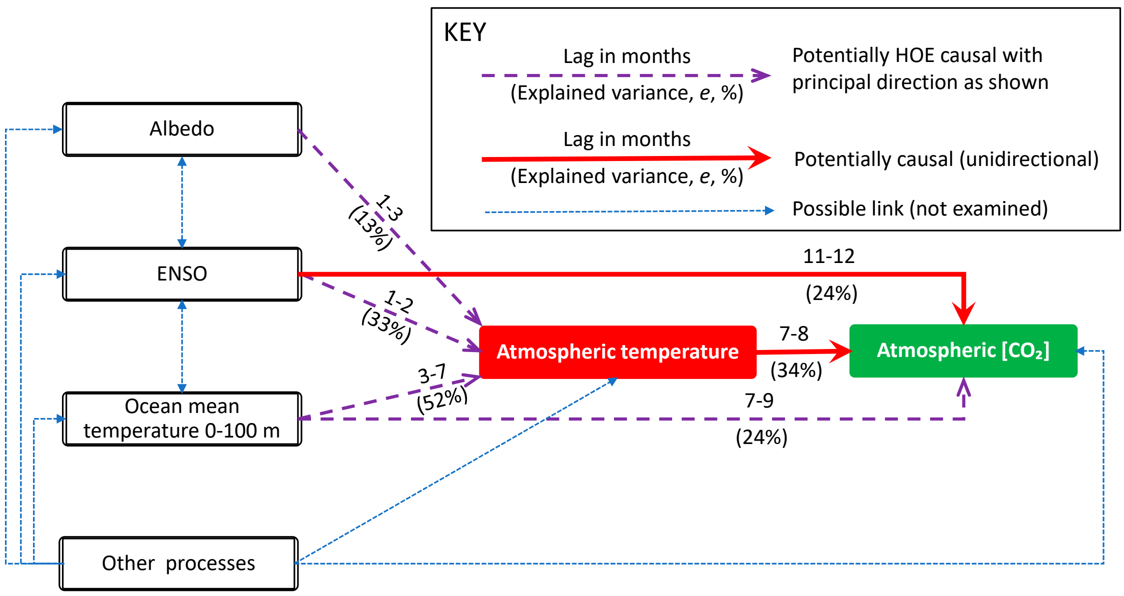


Figure 13. Schematic of the examined possible causal links in the climatic system, with noted types of potential causality, HOE or unidirectional, and its direction. Other processes, not examined here, could be internal of the climatic system or external.

The examined processes in the Appendices are internal to the climatic system. Other processes affecting T, not examined here, could also be external (e.g., solar and astronomical [43,44] and geological [45–49]). Generally, in complex systems, an identified causal link, even though it gives some explanation of a phenomenon, raises additional questions, e.g., what caused the change in the identified cause, etc. In turn, causal links in complex systems may form endless sequences. For this reason, it is naïve to expect complete answers to problems related to complex systems or to assume that a complex system is in permanent equilibrium and that an external agent is needed to "kick" it out of the equilibrium and produce change. Yet the investigation of a single causal link between two processes, as is the focus of this paper, provides useful information, with possible significant scientific, technical, practical, epistemological and philosophical implications. These are not covered in this paper. Readers interested in epistemological and philosophical aspects of causality are referred to Koutsoyiannis et al. [6], while those interested in the perennial changes in complex systems are referred to Koutsoyiannis [50,51].

As already clarified, the scope of our work is not to provide detailed modeling of the processes studied but to check causality conditions. We highlight the fact that the relationship we established explains only about 1/3 of the actual variance of $\Delta\ln[CO_2]$. This is not negligible for investigating causality, but also leaves a margin for many other climatic factors to act.

Nonetheless, our results can certainly be improved if we change our scope to more detailed modeling. To illustrate this, we provide the following toy model. Based on our result that the T-$[CO_2]$ system is potentially causal with direction $\Delta T \to \Delta\ln[CO_2]$, we estimate $\Delta\ln[CO_2]$ as

$$\Delta\ln[\mathrm{CO_2}] = \sum_{j=0}^{20} g_j \Delta T_{\tau-j} + \mu_v \tag{8}$$

and we proceed a step further, assuming that the mean μ_v also depends on past temperature, averaged at timescale m, i.e.,

$$\mu_v = \alpha(T_m - T_0) \tag{9}$$

where T_m is the average temperature of the previous m years, and α and T_0 are constants (parameters). Such a simple linear relationship is supported by the above-listed points related to the productivity of the biosphere. Equation (9) will result in negative values μ_v if $T_m < T_0$ and positive otherwise.

By re-evaluating the IRF coordinates g_j simultaneously with the parameters of Equation (9), we find the modified version of the IRF plotted in Figure 14. The optimized additional parameters are $m = 4$ (years), $\alpha = 0.0034$, $T_0 = 285.84$ K. Similarly to [6], we used a common spreadsheet software solver to perform the optimization, adding the two parameters α and T_0 to the unknown coordinates g_j of the IRF and performing the (nonlinear) optimization for all unknowns (m was found by trial-and-error). A graphical comparison of the actual $\Delta\ln[CO_2]$ and $[CO_2]$ with those simulated by the model of Equations (8) and (9) is given in Figure 15. The explained variance for $\Delta\ln[CO_2]$ was drastically increased from 34% to 55.5% and that for $[CO_2]$ is an impressive 99.9%.

For the convenience of the readers who are interested in repeating the calculations, we also give a parametric expression of IRF and summarize the toy model as:

$$\begin{gathered}\Delta\ln[\mathrm{CO_2}] = \sum_{j=0}^{20} g_j \Delta T_{\tau-j} + \mu_v, \\ g_j = 0.00076\, j^{0.67} e^{-0.2j}/\mathrm{K}, \mu_v = 0.0034\,(T_4/\mathrm{K} - 285.84)\end{gathered} \tag{10}$$

(where K is the unit of kelvin).

We emphasize, however, that we do not exploit the impressive result of explained variance of 99.9% to assert that we have built a decent model, even though this toy model is both accurate (in the lower panel of Figure 15, the simulated curve is indistinguishable from the actual) and parsimonious (the model expression in (10) contains only 5 fitted parameters). We prefer to highlight the fact that the hugely complex climate system entails high uncertainty, and its study needs reliable data that provide the basis for the modeling and testing of hypotheses.

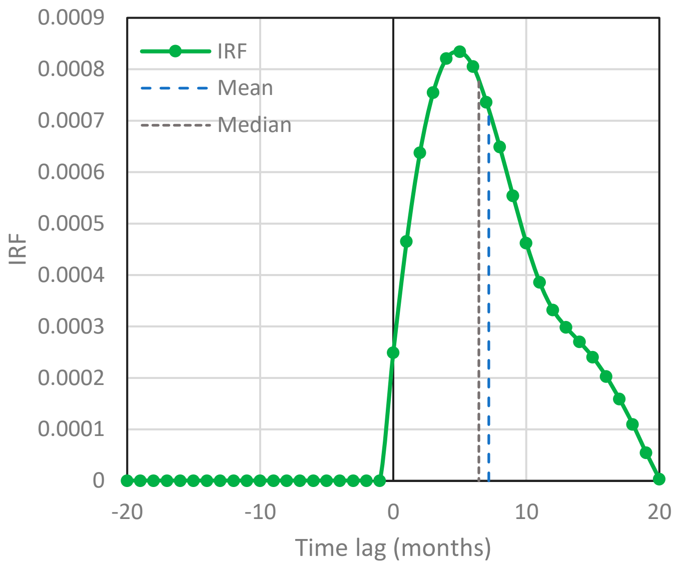


Figure 14. Modified IRF for temperature–CO_2 concentration based on the NCEP/NCAR Reanalysis temperature at 2 m and Mauna Loa time series, respectively, similar to Figure 2 but with IRF coordinates simultaneously optimized with the parameters of Equation (9).

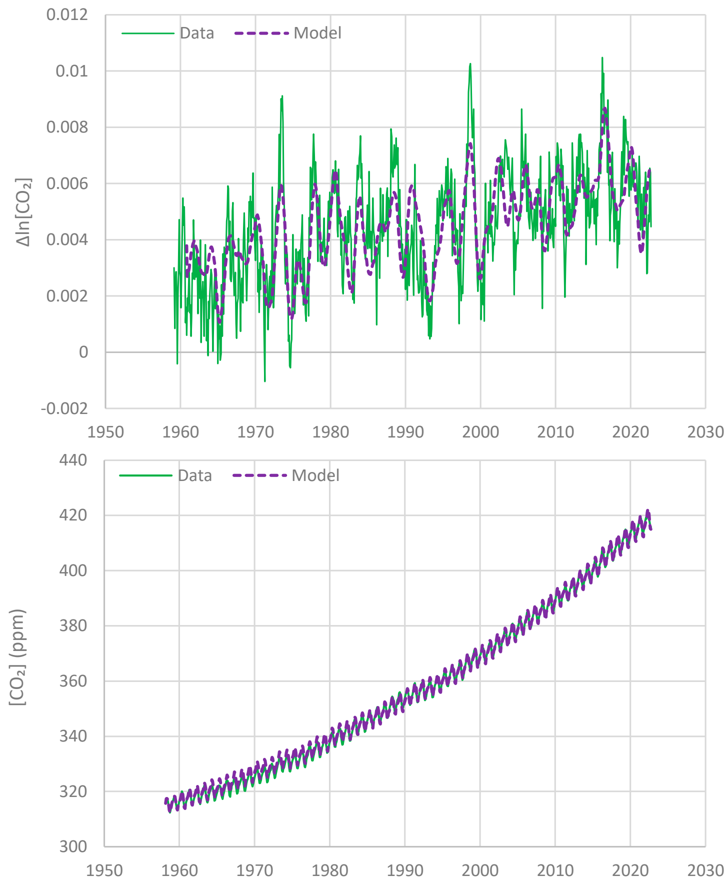


Figure 15. Comparison of the actual $\Delta\ln[CO_2]$ (**upper**) and $[CO_2]$ (**lower**) with those simulated by the model of Equations (8) and (9).

10. Conclusions

With reference to points 1–7 of the Introduction setting the paper's scope, the results of our analyses can be summarized as follows.

1. All evidence resulting from the analyses of the longest available modern time series of atmospheric concentration of $[CO_2]$ at Mauna Loa, Hawaii, along with that of globally averaged T, suggests a unidirectional, potentially causal link with T as the cause and $[CO_2]$ as the effect. This direction of causality holds for the entire period covered by the observations (more than 60 years).
2. Seasonality, as reflected in different phases of $[CO_2]$ time series at different latitudes, does not play any role in potential causality, as confirmed by replacing the Mauna Loa $[CO_2]$ time series with that in South Pole.
3. The unidirectional $T \to \ln[CO_2]$ potential causal link applies to all timescales resolved by the available data, from monthly to about two decades.
4. The proposed methodology is simple, flexible and effective in disambiguating cases where the type of causality, HOE or unidirectional, is not quite clear.
5. Furthermore, the methodology defines a type of data analysis that, regardless of the detection of causality per se, assesses modeling performance by comparing observational data with model results. In particular, the analysis of climate model outputs reveals a misrepresentation of the causal link by these models, which suggest a causality direction opposite to the one found when the real measurements are used.
6. Extensions of the scope of the methodology, i.e., from detecting possible causality to building a more detailed model of stochastic type, are possible, as illustrated by a toy model for the T-$[CO_2]$ system, with explained variance of $[CO_2]$ reaching an impressive 99.9%.
7. While some of the findings of this study seem counterintuitive or contrary to mainstream opinions, they are logically and computationally supported by arguments and calculations given in the Appendices.

Overall, the stochastic notion of a causal system, based on the concept of the impulse response function, proved to be very effective in studying demanding causality problems. A crucial characteristic of our methodology is its direct use of the data per se, in contrast with other methodologies that are based on uncertain estimates of autocorrelation functions or on the more uncertain tool of the power spectrum, i.e., the Fourier transform of the autocorrelation function. The methodology has the potential for further advances, as we first reported here (e.g., the asymmetric time lag window, the definition of a type of data analysis to be used in assessing modeling performance, and the extensions of its scope from detecting possible causality to building a more detailed model).

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/sci5030035/s1: Section SI1, Additional analysis of climate model behavior; Section SI2, On correlations of temperature with CO_2 emissions.

Author Contributions: Conceptualization, D.K.; methodology, D.K., C.O., Z.W.K. and A.C.; software, D.K.; validation, D.K., C.O., Z.W.K. and A.C.; formal analysis, D.K., C.O., Z.W.K. and A.C.; investigation, D.K., C.O., Z.W.K. and A.C.; data curation, D.K.; writing—original draft preparation, D.K., C.O., Z.W.K. and A.C; writing—review and editing, D.K., C.O., Z.W.K. and A.C.; visualization, D.K.; supervision, D.K., C.O., Z.W.K. and A.C.; project administration, D.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding but was motivated by the scientific curiosity of the authors.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data used in the paper are available online for free. The NCEP/NCAR temperature time series and the Mauna Loa CO_2 time series are readily available on a monthly scale from the climexp (http://climexp.knmi.nl/ (accessed on 1 January 2023) platform, namely from http://climexp.knmi.nl/data/inair_0-360E_-90-90N_n.dat and http://climexp.knmi.nl/getindices.cgi?WMO=CDIACData/maunaloa_f&STATION=Mauna_Loa_CO2 (accessed on 1 January 2023). The South Pole CO_2 concentration data are provided by the Global Monitoring Laboratory of the USA's National Oceanic and Atmospheric Administration (NOAA) at https://gml.noaa.gov/dv/data/index.php?parameter_name=Carbon%2BDioxide&site=SPO (accessed on 1 January 2023). The data retrieved are "Monthly Averages" for Type "Flask" and "Insitu". The CO_2 time series used in climate models have been downloaded from https://gmd.copernicus.org/articles/13/3571/2020/gmd-13-3571-2020-supplement.zip (accessed on 1 January 2023); from the Excel file provided, the data from the column "CO_2 ppm World" of the tabs "T2—History Year 1750 to 2014" and "T5—SSP2-4.5" have been retrieved. The climate model outputs were downloaded from the climexp platform, http://climexp.knmi.nl/selectfield_cmip6.cgi (accessed on 1 January 2023); specifically, from the "Monthly CMIP6 scenario runs", the globally averaged time series on "CMIP6 mean over all members" and "UKESM1-0-LL f2" have been derived through the platform. The CERES data were downloaded from https://ceres-tool.larc.nasa.gov/ord-tool/jsp/SSF1degEd41Selection.jsp (accessed on 17 March 2023). The SOI data were downloaded from https://www.ncdc.noaa.gov/teleconnections/enso/indicators/soi/ (accessed on 17 March 2023). The data on monthly global upper ocean mean temperature were downloaded from http://climexp.knmi.nl/getindices.cgi?WMO=NODCData/temp100_global&STATION=global_upper_ocean_mean_temperature (accessed on 17 March 2023).

Acknowledgments: We acknowledge a plethora of comments about our earlier papers [2,6,7]. Although few engaged with the methodology, they increased our confidence in our results. We also acknowledge the comments in five reviews in the two rounds of the review process of this paper, which triggered an expansion of the study by adding the Appendices, not contained in the original version, and their discussion in the body of the paper. We believe that the revised version is more complete.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A.

Appendix A.1. Notes on Carbon Cycle and its Dependence on Temperature

The carbon cycle is typically presented as a system with components that are in permanent equilibrium, except for the perturbation caused by anthropic activities. For example, the recent (2022) comprehensive study by Friedlingstein et al. [52], in its Figure 2 shows an absolute equilibrium in both the terrestrial and maritime parts of Earth, with inputs and outputs matching each other (±130 Gt C/year for the terrestrial part and ±80 Gt C/year for the maritime part), to which a human perturbation (9.6 ± 0.5 Gt C/year from the fossil fuel CO_2 emissions and 1.2 ± 0.7 Gt C/year from land-use change) is imposed and then distributed into several components. This representation is misleading, missing the large changes (by orders of magnitude) in the historical evolution of the abundance of CO_2 in Earth's atmosphere. A different approach and account has recently been provided by Stallinga [53], according to whom humans add 38 Gt of per year to the atmosphere-ocean system, a quantity equivalent to 10.4 Gt C/year.

Here, we follow the IPCC's [32] account in its recent (2021) Assessment Report (AR6). Its schematic (Figure 5.12 in that Report) does not hide (a) the imbalances in the different parts of Earth and (b) the fact that the natural carbon inputs and outputs in the atmosphere change over time—even though the IPCC's schematic implicitly assumes that "natural" is the budget that occurred in the preindustrial age (1750) and that any change that occurred since is anthropogenic. Interestingly, in an alternative view by Hansen et al. [54], civilization always produced greenhouse gases and aerosols, and humans likely contributed to the increase of both in the past 6000 years, thus resulting in climate forcings.

Based on the IPCC's representation, we have summarized in Figure A1 the information given in the IPCC's schematic, in terms of annual carbon balance. When seen in the entire picture, the human emissions due to fossil fuel combustion (9.4 Gt C/year including cement production) is a small part (4%) of the total CO_2 inflows to the atmosphere.

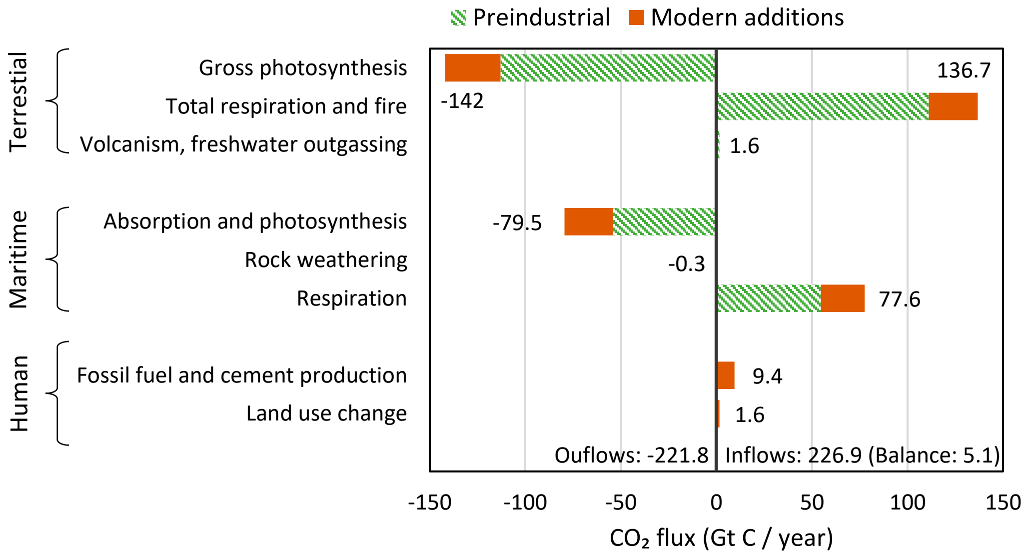


Figure A1. Annual carbon balance in the Earth's atmosphere in Gt C/year, based on the IPCC [32] estimates. The balance of 5.1 Gt C/year is the annual accumulation of carbon (in the form of CO_2) in the atmosphere.

The greatest part of the inflows is due to the respiration of the biosphere, i.e., the biochemical reaction whereby living organisms convert organic matter (e.g., glucose) to CO_2, releasing energy and consuming molecular oxygen [32]. As seen in Figure A1 (and in several publications, e.g., [55]), respiration has increased in recent years, the main reason for this being the increased temperature. Photosynthesis, the biochemical process that removes CO_2 from the atmosphere, producing carbohydrates in plants, algae and bacteria using the energy of light [32], has also increased, resulting in the greening of Earth [34–36] due to the increased atmospheric concentration of CO_2, which is plants' food.

It is not difficult to quantify the increase in respiration due to the temperature rise. The mechanism has been known in chemistry for more than a century. The rate of a chemical reaction k_T at temperature T is an increasing function of T, given by the Arrhenius equation [56]:

$$k_T = A\exp\left(-\frac{a}{R_* T}\right) \tag{A1}$$

where A and a are free parameters and R_* is the universal gas constant. Typically, the rate is measured in moles per unit volume, but it can readily be expressed in mass units. Expressing the relationship at a reference temperature T_0 and dividing with (A1), we obtain:

$$\frac{k_T}{k_0} = \exp\left(-\frac{a}{R_*}\left(\frac{1}{T} - \frac{1}{T_0}\right)\right) \tag{A2}$$

Taking the logarithms and setting $\Delta T := T - T_0$ we find

$$\begin{aligned} \ln\left(\tfrac{k_T}{k_0}\right) &= -\tfrac{a}{R_*}\left(\tfrac{1}{T} - \tfrac{1}{T_0}\right) = \tfrac{a}{R_* T_0}\left(1 - \tfrac{T_0}{T}\right) = \tfrac{a}{R_* T_0}\left(\tfrac{\Delta T}{T_0 + \Delta T}\right) \\ &= \tfrac{a}{R_* T_0}\left(\tfrac{\Delta T}{T_0} - \left(\tfrac{\Delta T}{T_0}\right)^2 + \left(\tfrac{\Delta T}{T_0}\right)^3 - \ldots\right) \end{aligned} \tag{A3}$$

and assuming that $\Delta T / T_0$ is small (nb., T_0 is of the order of 300 K, while typical values of ΔT is of the order of 1–10 K). We can neglect all terms beyond first order and find:

$$\frac{k_T}{k_{T0}} = \exp\left(\frac{a}{R_* T_0^2}\Delta T\right) = \left(\exp\left(\frac{a}{R_* T_0^2}\right)\right)^{\Delta T} = Q_1^{\Delta T} = Q_{10}^{\Delta T/10} \tag{A4}$$

where

$$Q_1 := \exp\left(\frac{a}{R_* T_0^2}\right), Q_{10} := Q_1^{10} \tag{A5}$$

Notice that both Q_1 and Q_{10} are dimensionless numbers > 1. The exponential expression in which Q_{10} is raised to power $\Delta T/10$ is known as the Q_{10} model [57].

The exponential increase of the process rate with temperature is a general chemical behavior, also extending to biochemical reactions. This is not a hypothesis or speculation but a proven fact that is widely used in engineering. For example, the metabolic rate in wastewater and sewer systems is expressed by the so-called effective BOD (EBOD, with BOD standing for biochemical oxygen demand). It has been known for more than 75 years that the metabolic rate increases with temperature, as microorganism activity generally increases with temperature. The relationship of EBOD with temperature has been expressed by Pomeroy and Bowlus [58] as $[\text{EBOD}] = [\text{BOD}]\ (1.07)^{T-20}$, which is similar to (A4), where the reference temperature is $T_0 = 20\ ^\circ\text{C}$ and $Q_1 = 1.07\ (Q_{10} = 2.0)$. The latter relationship is the standard of engineering design in sewer systems.

To apply this framework to find the increase of respiration in the last 65-year period investigated in our study, we first retrieved the global temperature separately for land and sea from the NCEP/NCAR Reanalysis data set. These are shown on an annual timescale in Figure A2. The resulting linear trends, also shown in Figure A2, yield an increase $\Delta T = 1.69\ ^\circ\text{C}$ for the terrestrial and $0.78\ ^\circ\text{C}$ for the maritime part for the 65-year period.

Now the literature gives representative average Q_{10} values of 3.05 for terrestrial respiration [57] and 4.07 for maritime respiration [59]. If R_B and R_E denote the respiration rate at the beginning and the end of the 65-year period, and $\Delta R := R_\text{E} - R_\text{B}$, then according to (A4),

$$\frac{R_\text{E}}{R_\text{B}} = Q_{10}^{\Delta T/10} \tag{A6}$$

and hence

$$\Delta R = R_\text{E}\left(1 - \frac{1}{Q_{10}^{\Delta T/10}}\right) \tag{A7}$$

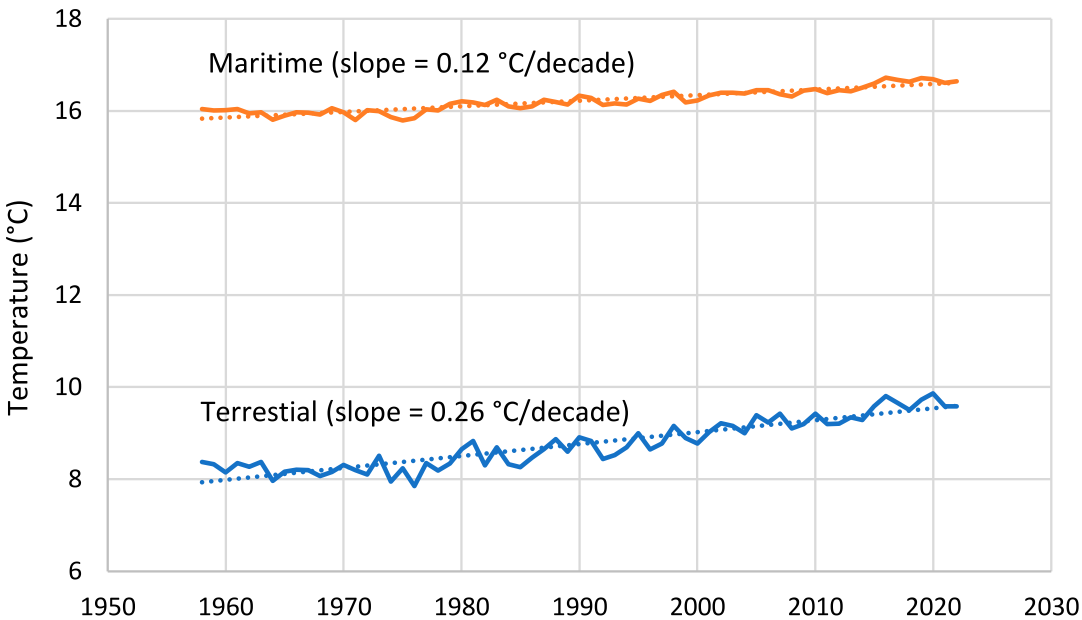


Figure A2. Evolution of global land (terrestrial) and sea (maritime) temperature at 2 m from the NCEP/NCAR Reanalysis data set, retrieved from the ClimExp platform, and resulting slopes of linear trends.

For the above given values of Q_{10} and ΔT, the expression in parentheses becomes 0.172 for the terrestrial part and 0.104 for the maritime part. Multiplying these by the R_E values shown in Figure A1, i.e., 136.7 and 77.6 Gt C/year, respectively, we find $\Delta R = 23.5$ and 8.1 Gt C/year, respectively, i.e., a total global increase in the respiration rate of $\Delta R = 31.6$ Gt C/year. This rate, which is a result of natural processes, is 3.4 times greater than the CO_2 emission by fossil fuel combustion (9.4 Gt C /year including cement production).

Appendix A.2. Investigation of Causality between Albedo and Atmospheric Temperature

There are several factors causing changes to the Earth's temperature. Solar radiation is a principal one, yet the changes in it have not been substantial at timescales of a few years. However, the Earth's response to solar radiation may change on such scales. Here, we investigate the changes of Earth's albedo. In the 21st century, the albedo at the top of the atmosphere (TOA) can be estimated from satellite data. Specifically, this can be done using the data from the ongoing project Clouds and the Earth's Radiant Energy System (CERES). This is part of NASA's Earth Observing System, designed to measure both solar-reflected and Earth-emitted radiation from the TOA to the Earth's surface. The data we used here are from the TERRA platform for the monthly timescale and are available online [60] for the period of March 2000 to date. The global TOA albedo was calculated as the ratio for each month of the globally integrated observed TOA shortwave flux (all-sky) over the globally observed TOA solar insolation flux. The resulting time series is shown in Figure A3. A falling linear trend of –0.0019/decade is also shown in the figure. A falling trend means that less solar radiation is reflected by the Earth, which may result in an increase in temperature. For the entire period, the decline of the albedo is about 0.004. As the average incoming solar radiation (insolation), according to the same data set, is 340 W/m^2, this implies a difference (imbalance) of received energy by Earth of $0.004 \times 340 = 1.4$ W/m^2. This result does not disagree with that of Hansen et al. [54], who found that in the period January 2015 through March 2022, the global absorbed solar energy is +1.01 W/m^2 higher than the mean for the first 10 years of data (2000–2009). These figures are greater than the average imbalance (net absorbed energy) of the Earth, which, if calculated from the ocean heat content data, is about 0.4 W/m^2 [33]. According to mainstream science, this imbalance is attributed to the increase of [CO_2], but the analyses in this study do not support this hypothesis. Moreover, it is hard to see how the albedo fall could be caused by increased [CO_2] (and for this reason, it is usually blamed on aerosols).

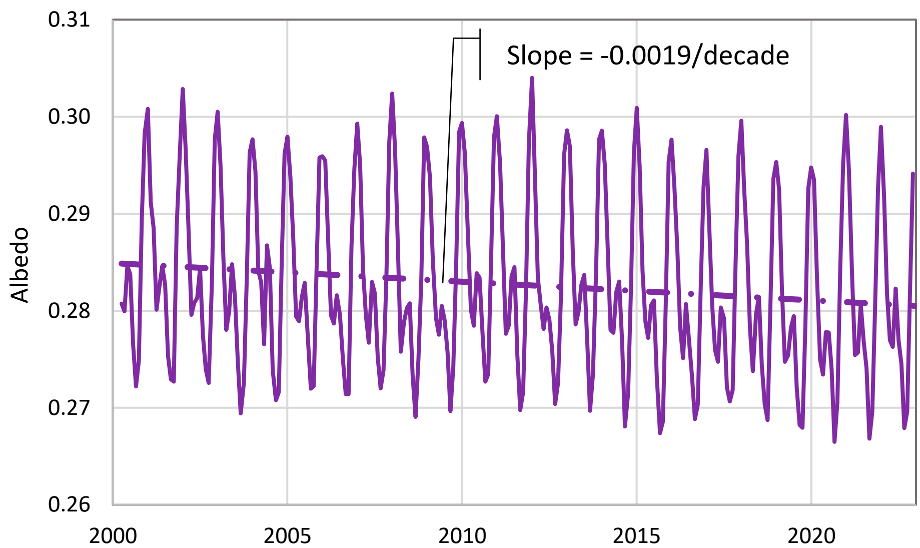


Figure A3. TOA albedo time series (continuous line), as provided by NASA's Clouds and the Earth's Radiant Energy System (CERES), along with linear trend (dashed line).

However, the potential causal link of albedo (α) with atmospheric T can be more thoroughly investigated by the stochastic framework discussed here. The resulting characteristics are shown in Table A1 and the resulting IRFs are shown in Figure A4. Notice that, as an increase of α is expected to cause a decrease of T, we changed the sign in albedo differences ($-\Delta\alpha$) so as to have a positive correlation with temperature differences (ΔT). The IRF determined suggests a potential HOE causation, with principal direction $\alpha \to T$ and time lags of 1–3 months. However, the explained variance is small, 13%.

Table A1. Summary indices of the case studies related to albedo. Data are on a monthly timescale and the time step of differencing is 1 year; for explanation of symbols see Table 1.

Case System	#	Direction	h_c	μ_h	$h_{1/2}$	$r_{yx}(h_c)$	e	ε
Albedo, α: CERES, TERRA;	24	$-\Delta\alpha \to \Delta T$	3	1.08	2.90	0.24	0.13	9.1×10^{-4}
T: NCEP/NCAR; period: 2000–2022	25	$\Delta T \to -\Delta\alpha$	–3	–0.31	–2.46	0.24	0.06	3.6×10^{-4}

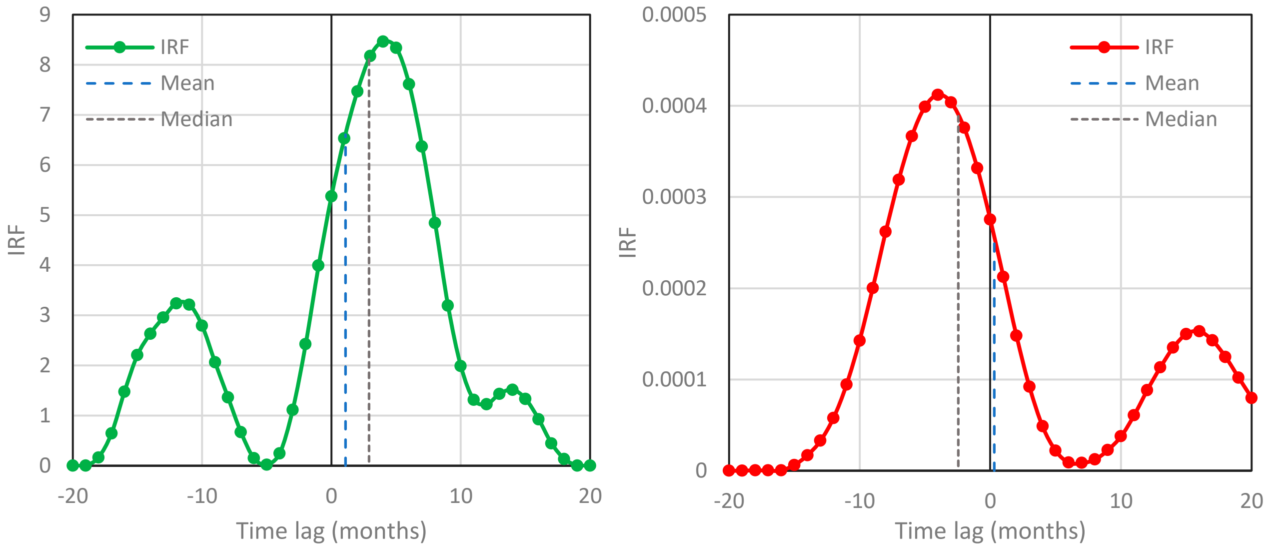


Figure A4. IRFs for albedo–temperature based on the CERES albedo time series and the NCEP/NCAR Reanalysis temperature at 2 m, respectively—case studies #24 (**left**; $-\Delta\alpha \to \Delta T$];) and #25 (**right**; $\Delta T \to -\Delta\alpha$).

Appendix A.3. Investigation of Causality between El Niño, Atmospheric Temperature and CO_2

A second process that is known to affect atmospheric temperature globally is the El Niño–Southern Oscillation (ENSO) (see [12,61] and a recent study by Kundzewicz et al. [62] for details). ENSO is associated with irregular variations in sea surface temperature and air pressure over the tropical Pacific Ocean. Several indices associated with ENSO are used in climatic studies, among which the most popular is US NOAA's Southern Oscillation Index (SOI), the time series of which is plotted in Figure A5.

Our stochastic methodology was previously applied with SOI along with satellite temperature data for the period 1979–2021 in [7]. Here we repeat the investigation replacing the temperature data with those of the NCEP/NCAR reanalysis and expanding the data back to 1951 (the beginning of the availability of SOI data) and forth to 2022. We also examined the potential causality between SOI and [CO_2]. In both cases, we tested differences with a time step of differencing of 1 year (thus reducing the effect of seasonality) and to make the correlation positive, we used $-\Delta$SOI (as in the albedo case).

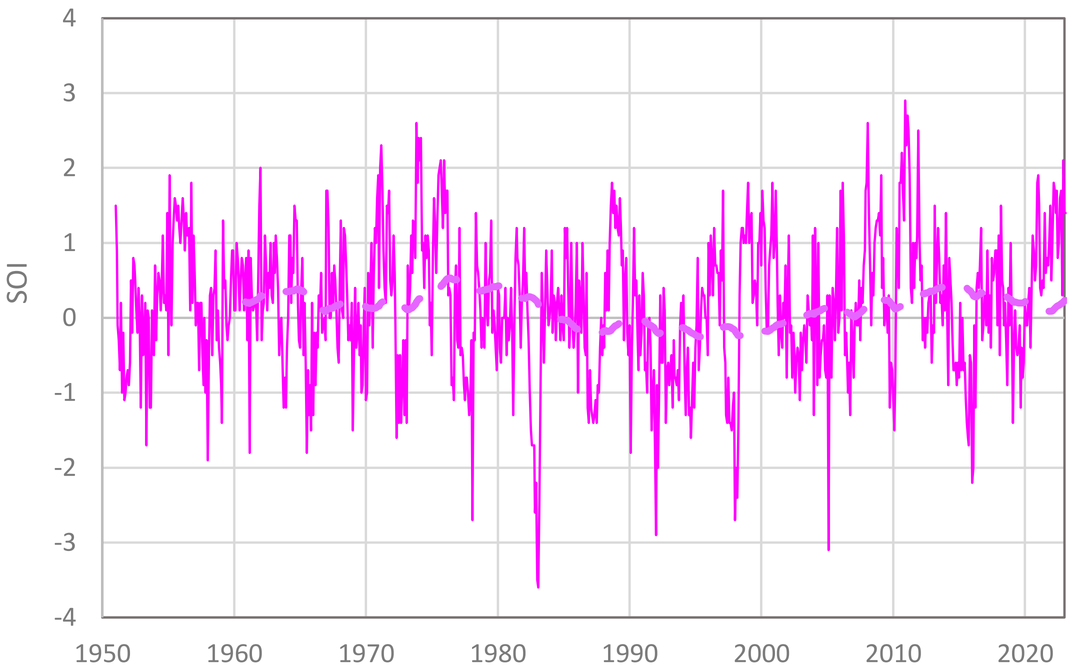


Figure A5. SOI time series (continuous line) along with rolling (right-aligned) 10-year average (dashed line). Negative and positive values indicate the El Niño and La Niña phases, respectively.

The results are shown in Table A2, Figures A6 and A7. The principal directions are SOI $\to T$ and SOI $\to [CO_2]$. In the former case, the explained variance is 33%, and the causality type is HOE but very close to unidirectional with a time lag of 4 months. In the latter case, the explained variance is 30%, and the causality type is unidirectional with a lag of about a year.

Table A2. Summary indices of the case studies related to ENSO. Data are on a monthly timescale and the time step of differencing is 1 year; for explanation of symbols see Table 1.

Case System	#	Direction	h_c	μ_h	$h_{1/2}$	$r_{yx}(h_c)$	e	ε
SOI: NOAA; T: NCEP/NCAR; period: 1951–2022	26	$-\Delta$SOI $\to \Delta T$	3	4.14	3.85	0.46	0.33	8.1×10^{-4}
	27	$\Delta T \to -\Delta$SOI	3	–2.15	–0.93	0.46	0.30	2.3×10^{-3}
SOI: NOAA; $[CO_2]$: Mauna Loa, period: 1958–2022	28	$-\Delta$SOI $\to \Delta\ln[CO_2]$	11	11.62	11.15	0.32	0.24	6.6×10^{-4}
	29	$\Delta\ln[CO_2] \to -\Delta$SOI	–11	–3.73	–3.84	0.32	0.08	0

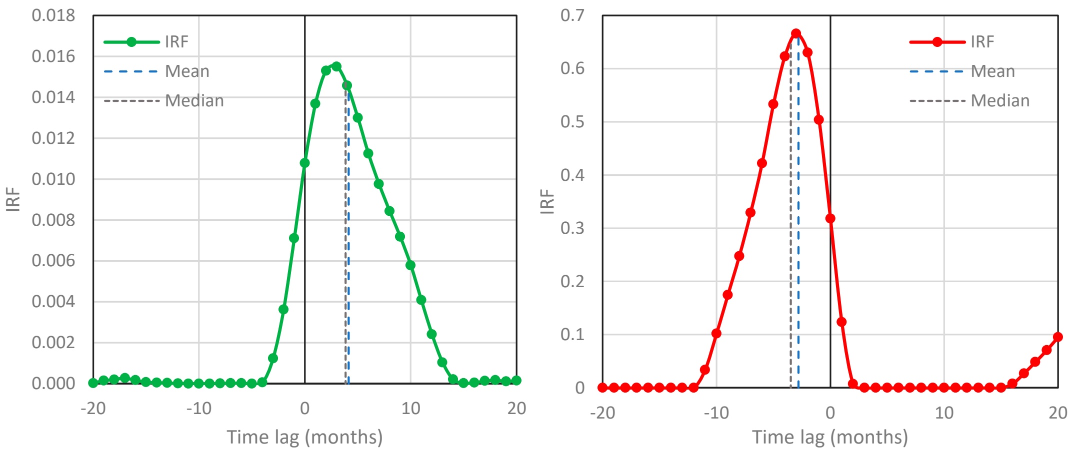


Figure A6. IRFs for ENSO–temperature based on the SOI time series and the NCEP/NCAR Reanalysis temperature at 2 m, respectively—case studies #26 (**left**; $-\Delta$SOI $\to \Delta T$];) and #27 (**right**; $\Delta T \to -\Delta$SOI).

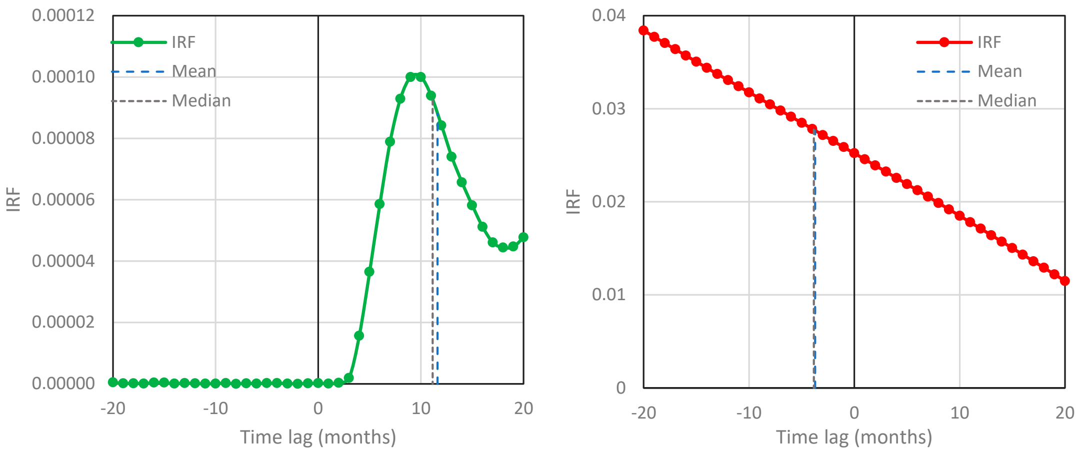


Figure A7. IRFs for ENSO–[CO_2] based on the SOI and the Mauna Loa time series, respectively—case studies #28 (**left**; $-\Delta\text{SOI} \rightarrow \Delta\ln[\text{CO}_2]]$;) and #29 (**right**; $\Delta\ln[\text{CO}_2] \rightarrow -\Delta\text{SOI}$).

Appendix A.4. Investigation of Causality between Ocean Heat Content, Atmospheric Temperature and CO_2

A third process examined in connection to both T and [CO_2] is the heat content of the upper-layer ocean. This is indirectly represented by data of the upper ocean mean temperature of the layer 0–100 m [63] (OMT0–100), also known as Vertically Averaged Temperature Anomaly (0–100 m layer) [64]. These are based on historical ocean temperature data, bathythermograph data and Argo data [65] and are made available at the timescale of three months by the National Oceanographic Data Center of the US NOAA [64], as well as by the ClimExp platform, which we used to download the data. The time series, starting in 1955, is plotted in Figure A8.

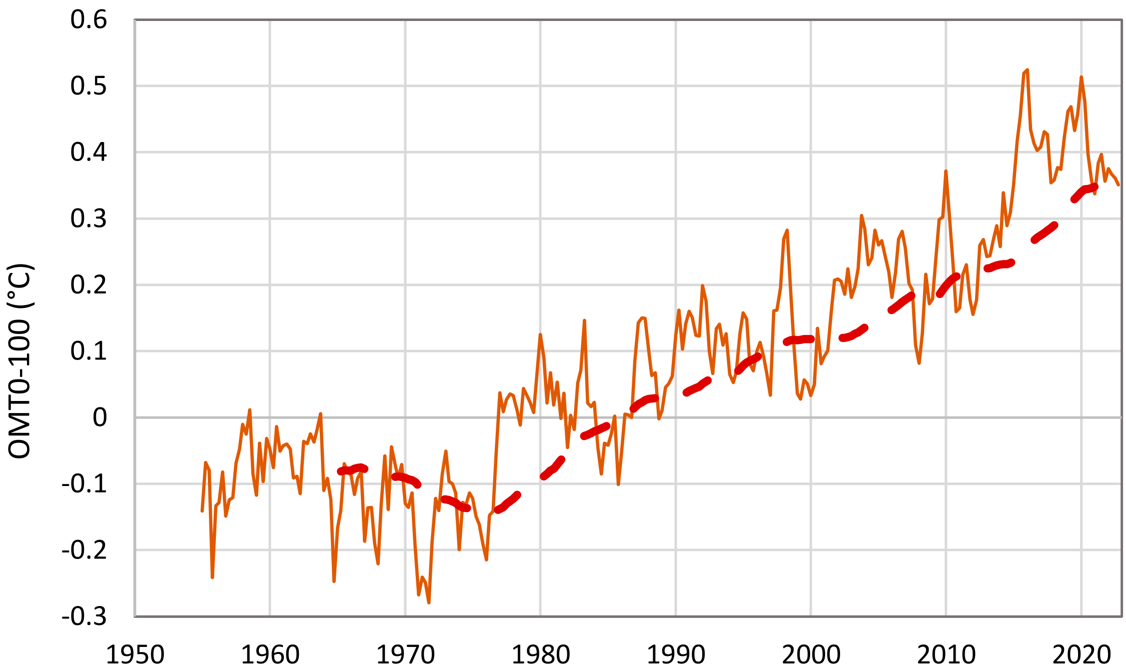


Figure A8. OMT0–100 time series (continuous line) along with rolling (right-aligned) 10-year average (dashed line).

The results of the stochastic analyses are shown in Table A3, Figures A9 and A10. The principal directions are OMT0–100 $\rightarrow T$ and OMT0–100 $\rightarrow$ [CO_2] and the causality type is HOE. In the former case, the explained variance is substantial, 53%, and the time lag is 3–7 months, depending on the metric used (nb., the time lags in Table A3 are given in 3-monthly steps). In the latter case, the explained variance is 35%, and the time lag is greater, 7–9 months.

Table A3. Summary indices of the case studies related to OMT0–100. Data are on a three-month timescale and the time step of differencing is 1 year; for explanation of symbols see Table 1.

Case System	#	Direction	h_c	μ_h	$h_{1/2}$	$r_{yx}(h_c)$	e	ε
OMT0–100: NOAA; T: NCEP/NCAR; period: 1955–2022	30 31	ΔOMT0–100 $\rightarrow \Delta T$ $\Delta T \rightarrow \Delta$OMT0–100	0 0	2.42 –2.15	0.98 –0.93	0.68 0.68	0.53 0.52	7.1×10^{-3} 3.8×10^{-3}
OMT0–100: NOAA; [CO_2]: Mauna Loa; period: 1958–2022	32 33	ΔOMT0–100 $\rightarrow \Delta\ln[CO_2]$ $\Delta\ln[CO_2] \rightarrow \Delta$OMT0–100	2 –2	2.22 –2.73	2.93 –2.82	0.46 0.46	0.35 0.21	5.8×10^{-4} 5.6×10^{-3}

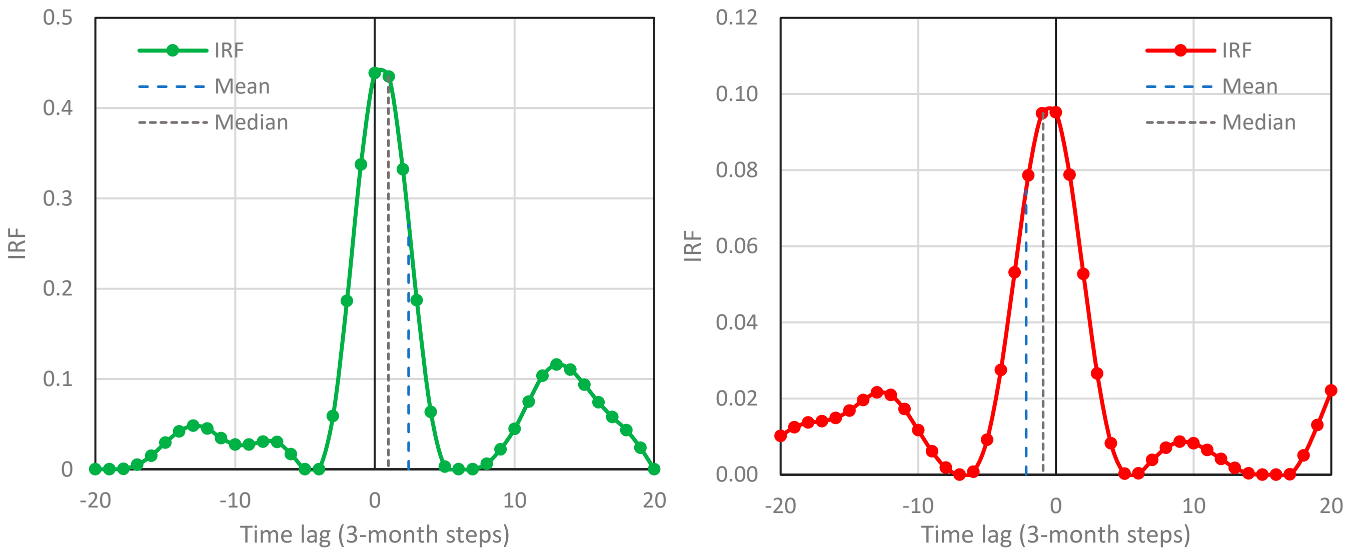


Figure A9. IRFs for upper ocean temperature—atmospheric temperature based on the OMT0–100 and the NCEP/NCAR Reanalysis data, respectively—case studies #30 (**left**; ΔOMT0–100 $\rightarrow \Delta T$];) and #31 (**right**; $\Delta T \rightarrow \Delta$OMT0–100).

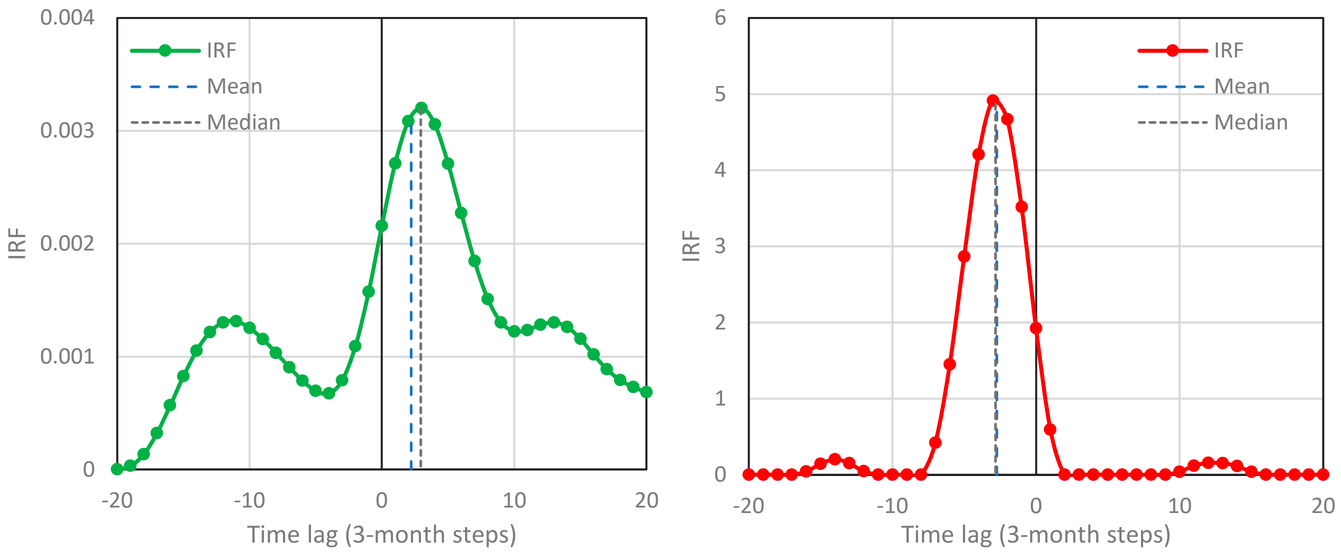


Figure A10. IRFs for upper ocean temperature—[CO_2] based on the OMT0–100 and the NCEP/NCAR Reanalysis data, respectively—case studies #32 (**left**; ΔOMT0–100 $\rightarrow \Delta\ln[CO_2]]$;) and #33 (**right**; $\Delta\ln[CO_2] \rightarrow \Delta$OMT0–100).

Appendix A.5. Notes on the T-[CO_2] *Relationship on Large Timescales*

While the scope of this study is the modern period covered by reliable CO_2 concentration measurements (about six decades), it may be useful to refer to studies that used paleoclimatic proxies to assess the T-[CO_2] relationship on much larger timescales. Berner and Kothavala [66] studied the entire phanerozoic (the last 530 million years) and asserted that "*over the long term there is indeed a correlation between CO_2 and paleotemperature*", while their Figure 13 showed that the atmospheric [CO_2] was much higher (up to 27 times) than the current one for most of the time during the phanerozoic. They also emphasized the "*importance of considering ALL factors affecting CO_2 when modelling the long term carbon cycle and not concentrating* [on] *only one cause*". On the other hand, Veizer et al. [67] presented evidence for decoupling atmospheric CO_2 and global climate during the phanerozoic, questioning the role of the (partial pressure of) CO_2 as the main driving force of past global (long-term) climate changes, at least during two of the four main cool climate modes of the phanerozoic.

Several studies, based on paleoclimatic reconstructions and mostly the Vostok ice cores [68,69] covering the most recent ~400 thousand years, have identified the change of T as the cause and that in [CO_2] as the effect, with estimates of the time lag varying from 50 to 1000 years, depending on the time period and the particular study [23,70–72]. Claims that CO_2 concentration leads (i.e., a negative lag) have not generally been made in these studies. At most, a synchrony claim has been sought on the basis that the estimated positive lags are often within the 95% uncertainty range [73], while in one of them [72], it has been asserted that a "*short lead of CO_2 over temperature cannot be excluded*".

Synchrony was also claimed for the last deglacial warming in another study by Parrenin et al. [74], who stated that they found no significant asynchrony between Antarctic temperature and atmospheric [CO_2]. For the same period, another study by Shakun et al. [75] gave different lead-lag relationships for the north and south hemispheres. Generally, it appears that the issue remains controversial, as illustrated, for example, in a recent report (2021) by NOAA (in the framework of Paleo Perspectives [76]), which states: "*While it might seem simple to determine cause and effect between carbon dioxide and climate from which change occurs first*, [. . .] *the determination of cause and effect remains exceedingly difficult*."

On the other hand, a convincing explanation about why, in the long run, change in temperature leads and in CO_2 concentration follows has been given by Roe [40], who demonstrated that in the Quaternary it is the effect of Milanković cycles, rather than of atmospheric CO_2 concentration, that explains the glaciation process. Specifically, he found that "*variations in atmospheric CO_2 appear to lag the rate of change of global ice volume. This implies only a secondary role for CO_2—variations in which produce a weaker radiative forcing than the orbitally-induced changes in summertime insolation—in driving changes in global ice volume.*"

The Vostok ice cores covering the most recent ~400 thousand years have also been examined, applying the same method as in this paper, by Koutsoyiannis et al. [7], who concluded that "*the causal relationship of atmospheric T and CO_2 concentration, as obtained by proxy data, appears to be of HOE type with principal direction $T \to [CO_2]$*", same as in the records of the modern period, but with a much higher time lag, of the order of 1000 years.

References

1. Sagan, C. *Cosmos*; Ballantine Books: New York, NY, USA, 1985.
2. Koutsoyiannis, D.; Kundzewicz, Z.W. Atmospheric temperature and CO_2: Hen-or-egg causality? *Sci* **2020**, *2*, 83. [CrossRef]
3. Πλούταρχος, Συμποσιακά Β′ (Plutarch, Quaestiones Convivales B)—Βικιθήκη. Available online: https://el.wikisource.org/wiki/Συμποσιακά_Β′ (accessed on 5 February 2023).
4. Chan, K.H.; Hayya, J.C.; Ord, J.K. A note on trend removal methods: The case of polynomial regression versus variate differencing. *Econometrica* **1977**, *45*, 737–744. [CrossRef]
5. Estrella, A. Why does the yield curve predict output and inflation? *Econ. J.* **2005**, *115*, 722–744. [CrossRef]
6. Koutsoyiannis, D.; Onof, C.; Christofides, A.; Kundzewicz, Z.W. Revisiting causality using stochastics: 1. Theory. *Proc. R. Soc. A* **2022**, *478*, 20210836. [CrossRef]
7. Koutsoyiannis, D.; Onof, C.; Christofides, A.; Kundzewicz, Z.W. Revisiting causality using stochastics: 2. Applications. *Proc. R. Soc. A* **2022**, *478*, 20210835. [CrossRef]

8. Young, P.C. *Recursive Estimation and Time Series Analysis*; Springer: Berlin/Heidelberg, Germany, 2011.
9. Young, P.C. Refined instrumental variable estimation: Maximum likelihood optimization of a unified Box-Jenkins model. *Automatica* **2015**, *52*, 35–46. [CrossRef]
10. Papoulis, A. *Probability, Random Variables and Stochastic Processes*, 3rd ed.; McGraw-Hill: New York, NY, USA, 1991.
11. Kestin, T.S.; Karoly, D.J.; Yang, J.I.; Rayner, N.A. Time-frequency variability of ENSO and stochastic simulations. *J. Clim.* **1998**, *11*, 2258–2272. [CrossRef]
12. Wills, R.C.; Schneider, T.; Wallace, J.M.; Battisti, D.S.; Hartmann, D.L. Disentangling global warming, multidecadal variability, and El Niño in Pacific temperatures. *Geophys. Res. Lett.* **2018**, *45*, 2487–2496. [CrossRef]
13. Granger, C.W. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **1969**, *37*, 424–438. [CrossRef]
14. Granger, C.W. Testing for causality: A personal viewpoint. *J. Econ. Dyn. Control.* **1980**, *2*, 329–352. [CrossRef]
15. Moraffah, R.; Sheth, P.; Karami, M.; Bhattacharya, A.; Wang, Q.; Tahir, A.; Raglin, A.; Liu, H. Causal inference for time series analysis: Problems, methods and evaluation. *Knowl. Inf. Syst.* **2021**, *63*, 3041–3085. [CrossRef]
16. Pearl, J. Causal inference in statistics: An overview. *Stat. Surv.* **2009**, *3*, 96–146. [CrossRef]
17. Pearl, J.; Glymour, M.; Jewell, N.P. *Causal Inference in Statistics: A Primer*; Wiley: Chichester, UK, 2016.
18. Pearl, J. and Mackenzie, D., The Book of Why, The New Science of Cause and Effect, Basic Books: New York, NY, USA, 2018.
19. Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 437–472. [CrossRef]
20. Meinshausen, M.; Nicholls, Z.R.J.; Lewis, J.; Gidden, M.J.; Vogel, E.; Freund, M.; Beyerle, U.; Gessner, C.; Nauels, A.; Bauer, N.; et al. The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geosci. Model Dev.* **2020**, *13*, 3571–3605. [CrossRef]
21. Sellar, A.A.; Jones, C.G.; Mulcahy, J.P.; Tang, Y.; Yool, A.; Wiltshire, A.; O'Connor, F.M.; Stringer, M.; Hill, R.; Palmieri, J.; et al. UKESM1: Description and evaluation of the UK Earth System Model. *J. Adv. Model. Earth Syst.* **2019**, *11*, 4513–4558. [CrossRef]
22. Koutsoyiannis, D. *Stochastics of Hydroclimatic Extremes—A Cool Look at Risk*, 2nd ed.; Kallipos Open Academic Editions: Athens, Greece, 2022; 346p, ISBN 978-618-85370-0-2. [CrossRef]
23. Koutsoyiannis, D. Time's arrow in stochastic characterization and simulation of atmospheric and hydrological processes. *Hydrol. Sci. J.* **2019**, *64*, 1013–1037. [CrossRef]
24. Strotz, R.H.; Wold, H.O.A. Recursive vs. nonrecursive systems: An attempt at synthesis (Part I of a triptych on causal chain systems). *Econometrica* **1960**, *28*, 417–427. Available online: https://www.jstor.org/stable/1907731 (accessed on 15 March 2023). [CrossRef]
25. Hannart, A.; Pearl, J.; Otto, F.E.L.; Naveau, P.; Ghil, M. Causal counterfactual theory for the attribution of weather and climate-related events. *Bull. Am. Met. Soc.* **2016**, *97*, 99–110. [CrossRef]
26. Hannart, A.; Naveau, P. Probabilities of causation of climate changes. *J. Clim.* **2018**, *31*, 5507–5524. [CrossRef]
27. Koutsoyiannis, D.; Efstratiadis, A.; Mamassis, N.; Christofides, A. On the credibility of climate predictions. *Hydrol. Sci. J.* **2008**, *53*, 671–684. [CrossRef]
28. Anagnostopoulos, G.G.; Koutsoyiannis, D.; Christofides, A.; Efstratiadis, A.; Mamassis, N. A comparison of local and aggregated climate model outputs with observed data. *Hydrol. Sci. J.* **2010**, *55*, 1094–1110. [CrossRef]
29. Koutsoyiannis, D.; Christofides, A.; Efstratiadis, A.; Anagnostopoulos, G.G.; Mamassis, N. Scientific dialogue on climate: Is it giving black eyes or opening closed eyes? Reply to "A black eye for the Hydrological Sciences Journal" by D. Huard. *Hydrol. Sci. J.* **2011**, *56*, 1334–1339. [CrossRef]
30. Tyralis, H.; Koutsoyiannis, D. On the prediction of persistent processes using the output of deterministic models. *Hydrol. Sci. J.* **2017**, *62*, 2083–2102. [CrossRef]
31. Scafetta, N. CMIP6 GCM validation based on ECS and TCR ranking for 21st century temperature projections and risk assessment. *Atmosphere* **2023**, *14*, 345. [CrossRef]
32. Masson-Delmotte, V.; Zhai, P.; Pirani, A.; Connors, S.L.; Péan, C.; Berger, S.; Caud, N.; Chen, Y.; Goldfarb, L.; Gomis, M.I.; et al. (Eds.) *IPCC, Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2021; 2391p. [CrossRef]
33. Koutsoyiannis, D. Rethinking climate, climate change, and their relationship with water. *Water* **2021**, *13*, 849. [CrossRef]
34. Zhu, Z.; Piao, S.; Myneni, R.B.; Huang, M.; Zeng, Z.; Canadell, J.G.; Ciais, P.; Sitch, S.; Friedlingstein, P.; Arneth, A.; et al. Greening of the Earth and its drivers. *Nature Climate Change* **2016**, *6*, 791–795. [CrossRef]
35. Li, Y.; Li, Z.L.; Wu, H.; Zhou, C.; Liu, X.; Leng, P.; Yang, P.; Wu, W.; Tang, R.; Shang, G.F.; et al. Biophysical impacts of earth greening can substantially mitigate regional land surface temperature warming. *Nat. Commun.* **2023**, *14*, 121. [CrossRef]
36. Chen, C.; Park, T.; Wang, X.; Piao, S.; Xu, B.; Chaturvedi, R.K.; Fuchs, R.; Brovkin, V.; Ciais, P.; Fensholt, R.; et al. China and India lead in greening of the world through land-use management. *Nat. Sustain.* **2019**, *2*, 122–129. [CrossRef]
37. Milanković, M. *Nebeska Mehanika*; Udruženje "Milutin Milanković": Beograd, Serbia, 1935.
38. Milanković, M. *Kanon der Erdbestrahlung und Seine Anwendung auf das Eiszeitenproblem*; Koniglich Serbische Akademice: Beograd, Serbia, 1941.
39. Milanković, M. *Canon of Insolation and the Ice-Age Problem*; Agency for Textbooks: Belgrade, Serbia, 1998.

40. Roe, G. In defense of Milankovitch. *Geophys. Res. Lett.* **2006**, *33*, L24703. [CrossRef]
41. Markonis, Y.; Koutsoyiannis, D. Climatic variability over time scales spanning nine orders of magnitude: Connecting Milankovitch cycles with Hurst–Kolmogorov dynamics. *Surv. Geophys.* **2013**, *34*, 181–207. [CrossRef]
42. Stephens, G.L.; Hakuba, M.Z.; Kato, S.; Gettelman, A.; Dufresne, J.-L.; Andrews, T.; Cole, J.N.S.; Willen, U.; Mauritsen, T. The changing nature of Earth's reflected sunlight. *Proc. R. Soc. A* **2022**, *478*, 1–37. [CrossRef]
43. Connolly, R.; Soon, W.; Connolly, M.; Baliunas, S.; Berglund, J.; Butler, C.J.; Cionco, R.G.; Elias, A.G.; Fedorov, V.M.; Harde, H.; et al. How much has the Sun influenced Northern Hemisphere temperature trends? An ongoing debate. *Res. Astron. Astrophys.* **2021**, *21*, 131.1–131.68. [CrossRef]
44. Scafetta, N.; Bianchini, A. The planetary theory of solar activity variability: A review. *Front. Astron. Space Sci.* **2022**, *9*, 937930. [CrossRef]
45. Kamis, J.E. The Plate Climatology Theory: How Geological Forces Influence, Alter, or Control Earth's Climate and Climate Related Events. Available online: https://books.google.gr/books/?id=7lRqzgEACAAJ (accessed on 10 March 2023).
46. Chakrabarty, D. *The Climate of History in a Planetary Age*; University of Chicago Press: Chicago, IL, USA, 2021; Available online: https://books.google.gr/books?id=ETQXEAAAQBAJ (accessed on 10 March 2023).
47. Davis, E.; Becker, K.; Dziak, R.; Cassidy, J.; Wang, K.; Lilley, M. Hydrological response to a seafloor spreading episode on the Juan de Fuca ridge. *Nature* **2004**, *430*, 335–338. [CrossRef] [PubMed]
48. Urakawa, L.S.; Hasumi, H. A remote effect of geothermal heat on the global thermohaline circulation. *J. Geophys. Res. Ocean.* **2009**, *114*, C07016. [CrossRef]
49. Patara, L.; Böning, C.W. Abyssal ocean warming around Antarctica strengthens the Atlantic overturning circulation. *Geophys. Res. Lett.* **2014**, *41*, 3972–3978. [CrossRef]
50. Koutsoyiannis, D. A random walk on water. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 585–601. [CrossRef]
51. Koutsoyiannis, D. Hydrology and change. *Hydrol. Sci. J.* **2013**, *58*, 1177–1197. [CrossRef]
52. Friedlingstein, P.; O'Sullivan, M.; Jones, M.W.; Andrew, R.M.; Gregor, L.; Hauck, J.; Le Quéré, C.; Luijkx, I.T.; Olsen, A.; Peters, G.P.; et al. Global Carbon Budget 2022. *Earth Syst. Sci. Data* **2022**, *14*, 4811–4900. [CrossRef]
53. Stallinga, P. Residence time vs. adjustment time of carbon dioxide in the atmosphere. *Entropy* **2023**, *25*, 384. [CrossRef]
54. Hansen, J.E.; Sato, M.; Simons, L.; Nazarenko, L.S.; von Schuckmann, K.; Loeb, N.G.; Osman, M.B.; Kharecha, P.; Jin, Q.; Tselioudis, G.; et al. Global warming in the pipeline. *arXiv* **2022**, arXiv:2212.04474.
55. Bond-Lamberty, B.; Thomson, A. Temperature-associated increases in the global soil respiration record. *Nature* **2010**, *464*, 579. [CrossRef] [PubMed]
56. Arrhenius, S.A. Über die Dissociationswärme und den Einfluß der Temperatur auf den Dissociationsgrad der Elektrolyte. *Z. Phys. Chem.* **1889**, *4*, 96–116. [CrossRef]
57. Patel, K.F.; Bond-Lamberty, B.; Jian, J.L.; Morris, K.A.; McKever, S.A.; Norris, C.G.; Zheng, J.; Bailey, V.L. Carbon flux estimates are sensitive to data source: A comparison of field and lab temperature sensitivity data. *Environ. Res. Lett.* **2022**, *17*, 113003. [CrossRef]
58. Pomeroy, R.; Bowlus, F.D. Progress report on sulfide control research. *Sew. Work. J.* **1946**, *18*, 597–640.
59. Robinson, C. Microbial respiration, the engine of ocean deoxygenation. *Front. Mar. Sci.* **2019**, *5*, 533. [CrossRef]
60. CERES Data Products. SSF1deg—Level 3, Gridded Daily and Monthly Averages of the SSF Product by Instrument. Available online: https://ceres-tool.larc.nasa.gov/ord-tool/jsp/SSF1degEd41Selection.jsp (accessed on 12 March 2023).
61. McPhaden, M.J.; Zebiak, S.E.; Glantz, M.H. ENSO as an integrating concept in earth science. *Science* **2006**, *314*, 1740–1745. [CrossRef]
62. Kundzewicz, Z.W.; Pinskwar, I.; Koutsoyiannis, D. Variability of global mean annual temperature is significantly influenced by the rhythm of ocean-atmosphere oscillations. *Sci. Total Environ.* **2020**, *747*, 141256. [CrossRef]
63. Levitus, S.; Antonov, J.I.; Boyer, T.P.; Baranova, O.K.; Garcia, H.E.; Locarnini, R.A.; Mishonov, A.V.; Reagan, J.R.; Seidov, D.; Yarosh, E.S.; et al. World Ocean heat content and thermosteric sea level change (0–2000 m). *Geophys. Res. Lett.* **2012**, *39*, L10603. [CrossRef]
64. National Oceanographic Data Center, NOAA, Global Ocean Heat and Salt Content. Available online: https://www.ncei.noaa.gov/access/global-ocean-heat-content/index3.html (accessed on 12 March 2023).
65. Roemmich, D.; Johnson, G.C.; Riser, S.; Davis, R.; Gilson, J.; Owens, W.B.; Garzoli, S.L.; Schmid, C.; Ignaszewski, M. The Argo Program: Observing the global ocean with profiling floats. *Oceanography* **2009**, *22*, 34–43. [CrossRef]
66. Berner, R.A.; Kothavala, Z. GEOCARB III: A revised model of atmospheric CO_2 over Phanerozoic time. *Am. J. Sci.* **2001**, *301*, 182–204. [CrossRef]
67. Veizer, J.; Godderis, Y.; François, L.M. Evidence for decoupling of atmospheric CO_2 and global climate during the Phanerozoic eon. *Nature* **2000**, *408*, 698–701. [CrossRef] [PubMed]
68. Jouzel, J.; Lorius, C.; Petit, J.R.; Genthon, C.; Barkov, N.I.; Kotlyakov, V.M.; Petrov, V.M. Vostok ice core: A continuous isotope temperature record over the last climatic cycle (160,000 years). *Nature* **1987**, *329*, 403–408. [CrossRef]
69. Petit, J.R.; Jouzel, J.; Raynaud, D.; Barkov, N.I.; Barnola, J.-M.; Basile, I.; Bender, M.; Chappellaz, J.; Davis, M.; Delayque, G.; et al. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **1999**, *399*, 429–436. [CrossRef]

70. Caillon, N.; Severinghaus, J.P.; Jouzel, J.; Barnola, J.M.; Kang, J.; Lipenkov, V.Y. Timing of atmospheric CO_2 and Antarctic temperature changes across Termination III. *Science* **2003**, *299*, 1728–1731. [CrossRef]
71. Soon, W. Implications of the secondary role of carbon dioxide and methane forcing in climate change: Past, present, and future. *Phys. Geogr.* **2007**, *28*, 97–125. [CrossRef]
72. Pedro, J.B.; Rasmussen, S.O.; van Ommen, T.D. Tightened constraints on the time-lag between Antarctic temperature and CO_2 during the last deglaciation. *Clim. Past* **2012**, *8*, 1213–1221. [CrossRef]
73. Chowdhry Beeman, J.; Gest, L.; Parrenin, F.; Raynaud, D.; Fudge, T.J.; Buizert, C.; Brook, E.J. Antarctic temperature and CO_2: Near-synchrony yet variable phasing during the last deglaciation. *Clim. Past* **2019**, *15*, 913–926. [CrossRef]
74. Parrenin, F.; Masson-Delmotte, V.; Köhler, P.; Raynaud, D.; Paillard, D.; Schwander, J.; Barbante, C.; Landais, A.; Wegner, A.; Jouzel, J. Synchronous change of atmospheric CO_2 and Antarctic temperature during the last deglacial warming. *Science* **2013**, *339*, 1060–1063. [CrossRef]
75. Shakun, J.D.; Clark, P.U.; He, F.; Marcott, S.A.; Mix, A.C.; Liu, Z.; Otto-Bliesner, B.; Schmittner, A.; Bard, E. Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* **2012**, *484*, 49–54. [CrossRef]
76. NOAA National Centers for Environmental Information. Temperature Change and Carbon Dioxide Change; 2021. Available online: https://www.ncei.noaa.gov/sites/default/files/2021-11/8%20-%20Temperature%20Change%20and%20Carbon%20Dioxide%20Change%20-%20FINAL%20OCT%202021.pdf (accessed on 12 January 2023).