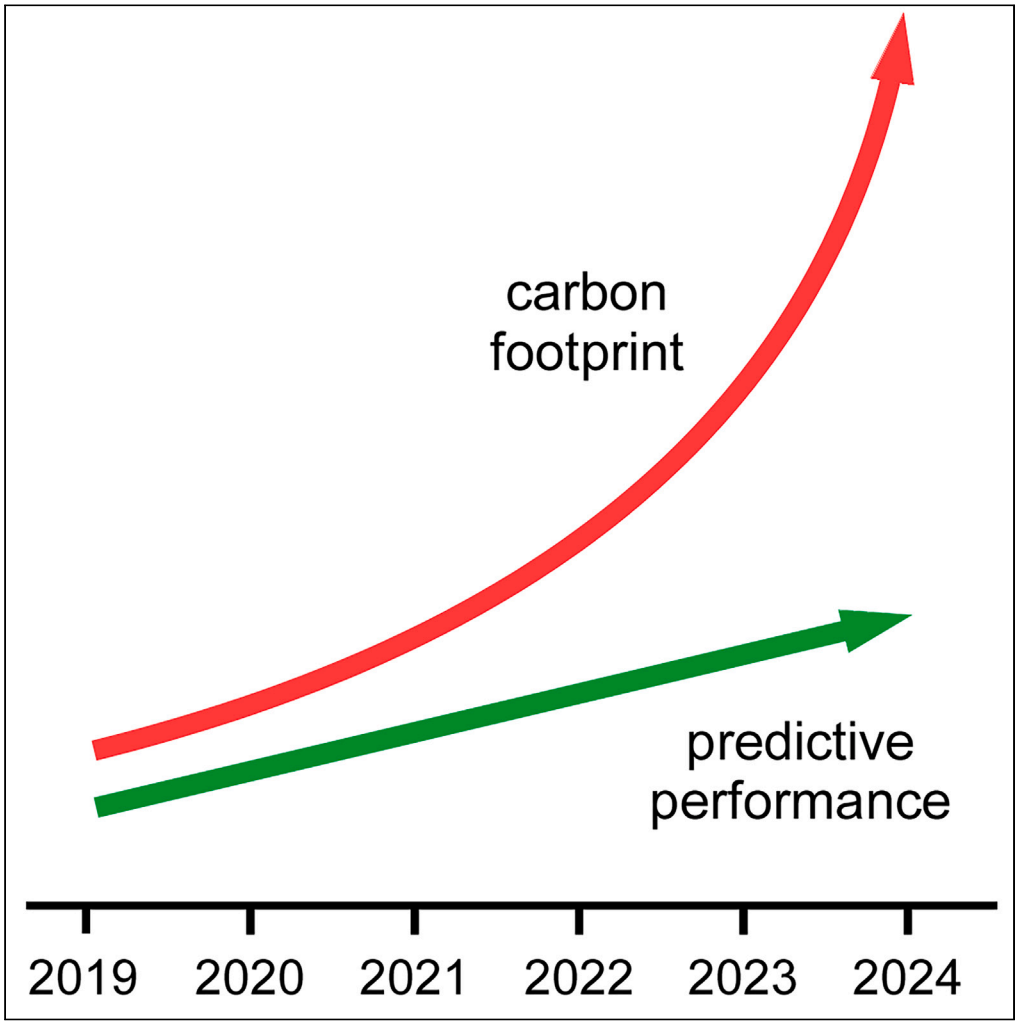
iScience

CellPress
OPEN ACCESS

Article

The carbon footprint of predicting CO_2 storage capacity in metal-organic frameworks within neural networks



Vadim Korolev,
Artem Mitrofanov

korolev@colloid.chem.msu.ru

Highlights

A diverse set of NNs for predicting CO_2 capacity in MOFs was trained

A disproportionate increase in model complexity and carbon footprint was noted

Environmental effects of AI in materials science need more quantification

Korolev & Mitrofanov, iScience *27*, 109644
May 17, 2024
https://doi.org/10.1016/j.isci.2024.109644

iScience

CellPress
OPEN ACCESS

Article

The carbon footprint of predicting CO_2 storage capacity in metal-organic frameworks within neural networks

Vadim Korolev[1,2,3,*] and Artem Mitrofanov[1,2]

SUMMARY

While artificial intelligence drives remarkable progress in natural sciences, its broader societal implications are mostly disregarded. In this study, we evaluate environmental impacts of deep learning in materials science through extensive benchmarking. In particular, a set of diverse neural networks is trained for a given supervised learning task to assess greenhouse gas (GHG) emissions during training and inference phases. A chronological perspective showed diminishing returns, manifesting themselves as a 28% decrease in mean absolute error and nearly a 15,000% increase in the carbon footprint of model training in 2016–2022. By means of up-to-date graphics processing units, it is possible to partially offset the immense growth of GHG emissions. Nonetheless, the practice of employing energy-efficient hardware is overlooked by the materials informatics community, as follows from a literature analysis in the field. On the basis of our findings, we encourage researchers to report GHG emissions together with standard performance metrics.

INTRODUCTION

Climate change is one of the greatest challenges that humanity has faced and that must be mitigated for a sustainable future.[1] Global warming caused by anthropogenic activities has already reached approximately 1°C above preindustrial levels.[2] Moreover, an increase by 2.7°C—far beyond the goals of the Paris Agreement—is expected to be reached in 2100 under the scenario with air pollutant emissions following current trends.[3] An immediate and deep reduction in greenhouse gas (GHG) emissions is necessary to limit global warming and avoid further escalation of the climate change crisis; global net zero CO_2 emissions need to be reached in the early 2050s.[4]

Substantial environmental impacts of scientific studies were shown recently. In particular, work-related GHG emissions per astronomer are 40%–60% higher than those per average citizen of the same country.[5,6] Electricity consumption takes the first and second place among the relevant types of activity, as follows from the assessments provided by the Australian astronomical community[5] and Max Planck Institute for Astronomy,[6] respectively. The corresponding GHG emissions are mainly associated with the use of (super)computing facilities, thus raising awareness of the carbon footprint deriving from the third pillar of modern science: simulation and modeling. Several attempts to estimate GHG emissions in other computing-related fields, including bioinformatics[7] and particle physics,[8] have been carried out as well. In this context, special attention should be paid to approaches under the umbrella of artificial intelligence (AI). Deep learning algorithms,[9] being one of the most exciting AI subfields, are rapidly developing; training compute has doubled every 8–17 months since 2015.[10] GHG emissions of large-scale models (e.g., T2T,[11] GPT-3; [12] and BigScience Large Open-science Open-access Multilingual Language Model [BLOOM][13]) have reached the level of tons of carbon dioxide equivalents, tCO_2e. Nevertheless, the focus of the AI community is still on achieving state-of-the-art (SOTA) results on benchmarks, whereas other aspects, including environmental impacts, are mostly ignored. This tendency, known as Red AI,[14] has been identified for deep learning models in natural language processing (NLP), but similar concerns hold true for other AI-related fields.[15,16]

In materials science, the tetrahedron that incorporates four basic concepts—processing, structure, properties, and performance—has lately been accompanied by a "digital twin" rooted in information science.[17] The edge of the tetrahedron joining structure and properties have been explored through the lens of AI algorithms, resulting in a paradigm shift toward data-driven materials science.[18] Unfortunately, these algorithms also follow the Red AI tendency by prioritizing accuracy above all other factors. In particular, there is no quantitative assessment of the environmental impacts of AI in materials science to the best of our knowledge. To fill this gap, we examine materials property prediction as a representative use case. A detailed consideration of temporal evolution in terms of model complexity and GHG emissions is performed for a series of neural networks (NNs) trained in a uniform manner. The revealed trends confirm our preliminary concerns: the carbon footprint of AI in materials science is no longer negligible and must be taken into account by AI practitioners in

[1]Department of Chemistry, Lomonosov Moscow State University, Moscow 119991, Russia
[2]MSU Institute for Artificial Intelligence, Lomonosov Moscow State University, Moscow 119192, Russia
[3]Lead contact
*Correspondence: korolev@colloid.chem.msu.ru
https://doi.org/10.1016/j.isci.2024.109644

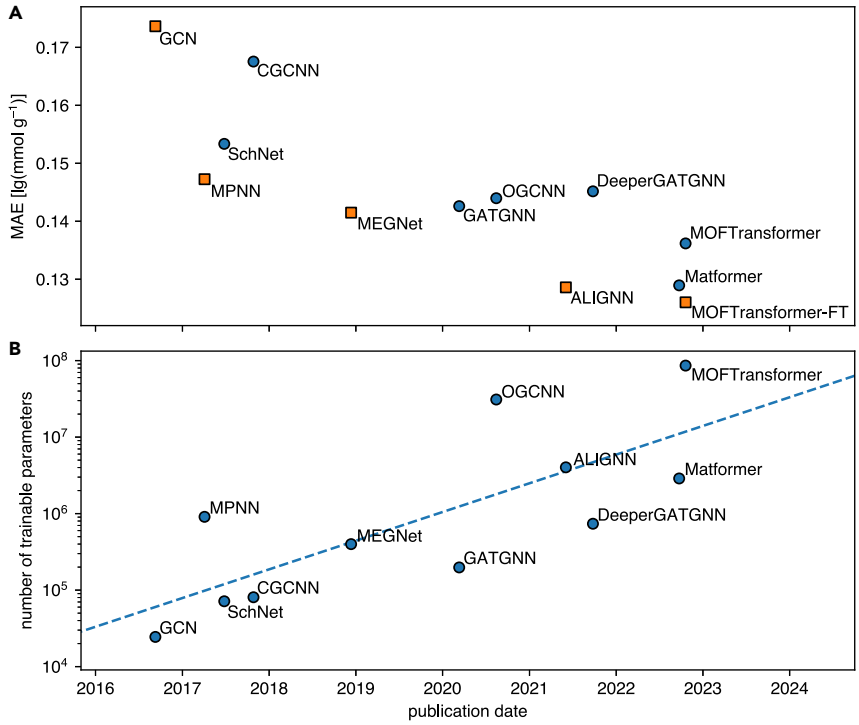


Figure 1. Temporal evolution of NNs in terms of performance and model complexity

(A) MAE in predicting CO_2 working capacity as a function of the publication date. The first appearance of a model, including preprints, is taken into consideration. The models with state-of-the-art performance are labeled orange.

(B) The number of trainable parameters as a function of the publication date. The dashed blue line corresponds to a log-linear fit.

the field. Nevertheless, an analysis of the literature indicates that the materials informatics community has yet to embrace a systematic evaluation of environmental impacts.

RESULTS AND DISCUSSION

We evaluate the performance and carbon footprint of training a set of diverse NNs[19–29] that predict CO_2 working capacity in metal-organic frameworks[30] (MOFs). In our opinion, the selected approach, which centers on one target property and a specific set of supervised learning algorithms, seems appropriate owing to the lack of comparable assessments in materials science. At this point, an analysis of the most general trends (even at a semiquantitative level) rather than comprehensive benchmarking is necessary. It should also be emphasized that most of the analyzed models are message-passing graph NNs, which are now the mainstream approach to achieving SOTA performance in materials property prediction.[31] Hence, the ensuing results will be primarily interesting to researchers who intensively use and develop these promising models; recent transformer-based architectures are also taken into account.[28,29] On the other hand, prediction of CO_2 capacity should serve as an emblematic example of a computational task that has a second-order positive impact on GHG emissions.

Mean absolute error (MAE) values as a measure of performance are depicted in Figure 1A. The models are organized in chronological order by publication date. SOTA MAE decreased by 28% in six years (2016–2022): from 0.174 mmol g^{-1} for graph convolutional network[19] (GCN) to 0.126 mmol g^{-1} for the fine-tuned version of the multi-modal transformer encoder pretrained with 1 million hypothetical MOFs[29] (MOFTransformer-FT). Starting from message-passing neural network[20], SOTA MAE seems to diminish linearly, but this estimate is rather speculative because of the small number of relevant models. In general, an incremental improvement in performance is achieved by adopting advanced NNs (e.g., transformer-based MOFTransformer[29]) and incorporating physics-informed features (e.g., many-body interactions in ALIGNN[26]). The diversity of NN architectures hinders a quantitative comparison of their complexity; therefore, here we apply a simple criterion: the number of trainable parameters (Figure 1B). For the same period (2016–2022), the quantity increased by more than three orders of magnitude: from 24.5 thousand for GCN to 86 million for MOFTransformer(-FT); the number of parameters doubled every 9.6 months according to a log-linear fit (dashed line in Figure 1B). Therefore, the performance gain is offset by the exponential cost in terms of model size, in line with the diminishing returns[32] rule in deep learning. It is also interesting to note that the deduced trend matches the doubling time of 9–10 months intrinsic in large-scale models surprisingly well.[10]

Next, we examine GHG emissions associated with training NNs (Figure 2A). As in the case of model size, a significant increase in the carbon footprint during the considered period (2016–2022) was observed: from 27.0 gCO_2eq for GCN to 4.07 kgCO_2eq for Matformer. Accordingly,

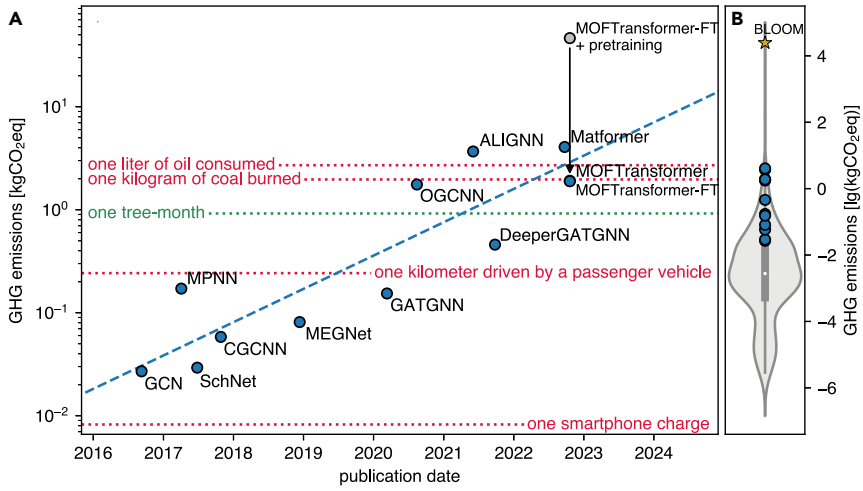


Figure 2. Greenhouse gas (GHG) emissions of model training in different domains

(A) GHG emissions of model training as a function of the publication date. The dashed blue line represents a log-linear fit. The horizontal dotted green and red lines indicate GHG emissions associated with common activities. The gray circle shows cumulative GHG emissions of MOFTransformer-FT pretraining and fine-tuning.

(B) GHG emissions of the models deposited on the Hugging Face Hub. The blue circles denote the models depicted in a subplot (A). The gold star indicates GHG emissions of training the BLOOM.

GHG emissions doubled every 11.2 months, as follows from a log-linear fit (dashed line in Figure 2A). A similar rate was noted in the case of GHG emissions per epoch (11.7 months, Figure S1). If the trend continues, the carbon footprint of training a single model will exceed 10 $kgCO_2eq$ and 100 $kgCO_2eq$ in mid-2024 and mid-2027, respectively. To put the presented estimates into context, GHG emissions of common activities (https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator) are shown in Figure 2A. For instance, training the GCN model has a carbon footprint equivalent to three smartphone charges, whereas the GHG emissions of the Matformer training phase are comparable to driving an average passenger vehicle for 16.8 km. On the other hand, the time taken by a mature tree to sequester the carbon dioxide equivalents[33] emitted during the training of MOFTransformer-FT is 2.1 months, whereas this value is expected to be 9.1 years for the presented hypothetical model in mid-2027. Another approach to assessing GHG emissions of the models in question is to compare with those of some models from a different domain. An extensive collection of climate performance cards is available on the Hugging Face Hub (https://huggingface.co/blog/carbon-emissions-on-the-hub). The distribution of carbon dioxide equivalents for training NLP models (released from early 2022 to 30 April 2023) is shown in Figure 2B; the data on materials property predictors are also provided for comparison. The values at the 5% and 95% quantiles are 7.4 $mgCO_2eq$ and 271.1 gCO_2eq, respectively. Surprisingly, even the GCN model has a carbon footprint higher than 81.1% of NLP models; this ratio for Matformer equals 99.3%. At the same time, the GHG emissions of SOTA models predicting CO_2 capacity are much lower than those from large language models (LLMs), such as the BLOOM.[34] The difference between the carbon footprints of training BLOOM and Matformer turned out to be 6,074-fold.

If we speculate about the mechanism of further increases in GHG emissions, the implementation of advanced LLMs seems to be a likely candidate. The MOFTransformer model is based on BERT (Bidirectional Encoder Representations from Transformers[35]), presented by Google in 2018. In our recent study,[36] BERT was implemented to achieve SOTA performance on predicting properties of inorganic crystals. More advanced LLMs, including Generative Pretrained Transformers[37,38] (formed GPT-*n* series) from OpenAI, exceed BERT in terms of model size by orders of magnitude and have yielded promising results on predicting chemical properties, according to preliminary attempts.[39–41] Nevertheless, the avalanche-like growth of GHG emissions caused by the use of ever-growing LLMs can be partially amortized by large learning capacity: a prominent feature of these models. Diverse pretraining strategies, including transfer learning and self-supervised learning,[42] make it possible to achieve SOTA performance in downstream tasks at relatively low computational costs. MOFTransformer-FT should serve as a vivid example: 1.89 $kgCO_2eq$ was emitted during the model training, whereas the pretraining phase would result in additional 44.5 $kgCO_2eq$ (Figure 2A). This rough estimate is based on the assumption that GHG emissions per example in the pretraining phase are equal to those at the fine-tuning step. It should be underscored that the pretrained MOFTransformer can be reused multiple times without GHG emissions, which occur during the pretraining. Still, inference-related GHG emissions, though negligible per sample (see Figure S2), can reach significant levels because of intensive use of the corresponding model; trillions of hypothetical MOFs are potentially accessible for screening within predictive algorithms.[43] In a broader context, similar concerns are most relevant to generative AI: annual GHG emissions from deployed LLM-powered applications (e.g., ChatGPT) can be many times greater than those from model training.[44]

The aforementioned results on GHG emissions were obtained through retroactive computations, which have inherent drawbacks. All training experiments were conducted in a consistent manner using the same hardware, whereas energy efficiency of graphics processing units (GPUs)—workhorses of deep learning[45]—consistently improved during the period of interest. For instance, single-precision performance of

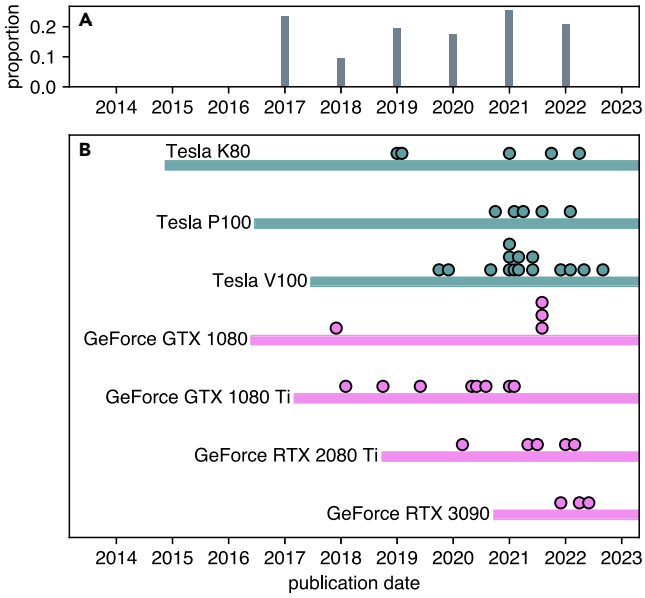


Figure 3. Graphics processing units (GPUs) employed in materials informatics studies

(A) The proportion of publications (by year) that mention a specific GPU model used.

(B) Timelines for selected GPU accelerators. The circles indicate the publication date with a time step of a month; the vertical arrays show that several preprints were submitted in the same month. Enthusiast-class and professional GPUs are marked violet and blue, respectively.

the NVIDIA Tesla series increased from 37.2 GFLOPS watt^{-1} for P100 (2016) to 145.7 GFLOPS watt^{-1} for H100 (2022). We suppose that if the newly released GPUs are used during the NN training, then the GHG emission ratio for Matformer (2022) and GCN (2016) will decrease from 150.8 (Figure 2A) to 38.5. The choice in favor of the recent GPUs and other AI-specific hardware—tensor processing units—is recognized as an effective way to lower the environmental impacts of deep learning.[46] At the same time, little is known about the practical implementation of such guidelines, especially in the field of materials informatics. To clarify this aspect, we conducted a survey of preprints on arXiv that deal with the application of NNs in materials science. The arXiv repository is an open-access platform for sharing findings that push the frontiers of materials informatics before publication in peer-reviewed journals. In particular, 10 of the 11 tested NNs have a corresponding preprint on arXiv. Among 317 manuscripts examined, only 62 (20%) contain a direct mention of GPU; the proportion ranged from 10% (2017) to 26% (2021) over the years (Figure 3A). Upon closer examination, it became apparent that there is an environmentally unfriendly tendency: previous-generation GPUs are still in use alongside the newer devices (Figure 3B). For instance, Tesla K80 (launched in 2014) is mentioned as a GPU accelerator in the manuscripts published before 2022 despite the availability of more energy-efficient successors, including P100 (2016) and V100 (2017). Other concerns are related to the abundance of enthusiast-class GPUs (GeForce 10, 20, and 30 series) that are used almost as often as professional devices (Tesla series). Both trends suggest that the materials informatics community has mainly overlooked the importance of utilizing the most advanced hardware, including GPUs, for reducing environmental impacts. Unfortunately, much broader trends, such as the de-democratization of AI,[47] tend to limit the access of academic institutions (primarily in low- and middle-income countries) to computing facilities necessary to implement SOTA models. Moreover, researchers generally do not consider the carbon footprint of model training. While 55 manuscripts (17%) mention computation time, only one[48] (0.3%) includes some estimates of GHG emissions. Therefore, we should conclude that the negative environmental impacts of AI-related applications in materials science are not even recognized as a problem by the community.

The current level of operational GHG emissions (in kgCO_2eq) from model training and the (exponential) rate at which they grow should raise concerns, whereas taking into account the carbon footprint of the entire model life cycle will reveal that the situation is even more frightening. Specifically, model tuning (via hyperparameter optimization) aimed at improving accuracy typically requires training hundreds of models with various architectures,[49] causing GHG emissions to go up proportionally. Hence, the carbon footprint of SOTA models involved in materials informatics tasks can reach hundreds of kgCO_2eq during model development. At the inference stage, GHG emissions per test point seem to be negligible (Figure S2). The practical use of a deployed model in industry-scale machine learning (ML) use cases may lead to GHG emissions greater than those of the preceding stages[50]; there are no such estimates for materials informatics applications to date. From a more general perspective,[51] the effects of ML on GHG emissions have been classified into three categories: computing-related impacts, immediate impacts of applying ML, and system-level impacts. All the aforementioned model-related emissions fall into the first category in conjunction with the intrinsic emissions of hardware manufacturing. The last two categories are difficult to differentiate and estimate appropriately. It is interesting to note that these impacts can be either positive or negative, depending on an application. The presented use case has good potential to influence climate change mitigation by accelerating the design of efficient CO_2 adsorbers, whereas system-level

impacts beyond immediate application-level impacts are not so obvious. To sum up, we expect that evaluating the environmental impacts of AI in materials science holistically will require a community-wide effort via a shift of the focus from the performance-centric point of view.

The carbon footprint of AI in materials science definitely deserves much more attention than it has attracted so far. The confirmed tendencies indicate environmental impacts that should not be ignored and are expected to grow. By presenting this study, we hope to inspire further efforts on evaluation of GHG emissions in the field. Currently, there are several ways to make quantified GHG emissions accessible to the scientific community, including impact statements of data science conferences and metadata of model cards in open-access repositories, such as the Hugging Face Hub. Although the proposed options meet only some specific needs of AI practitioners in materials informatics, they can still be a stepping stone toward a more sustainable and greener AI-powered materials science.

Limitations of the study

The emission intensity coefficient at a specific location can vary throughout the day, affecting the conversion of energy consumption to CO_2 emissions; assuming an insignificant influence of this factor, we excluded it from our consideration.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Dataset
  - Models
  - Literature analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.109644.

ACKNOWLEDGMENTS

V.K. was supported by a Fellowship from Non-commercial Foundation for the Advancement of Science and Education INTELLECT.

AUTHOR CONTRIBUTIONS

Conceptualization, V.K. and A.M.; methodology, V.K.; software, V.K.; investigation, V.K. and A.M.; writing – original draft, V.K.; writing – review and editing, A.M.; supervision, A.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 10, 2024
Revised: February 28, 2024
Accepted: March 27, 2024
Published: March 29, 2024

REFERENCES

1. Ripple, W.J., Wolf, C., Gregg, J.W., Levin, K., Rockström, J., Newsome, T.M., Betts, M.G., Huq, S., Law, B.E., Kemp, L., et al. (2022). World Scientists' Warning of a Climate Emergency 2022. Bioscience *72*, 1149–1155.
2. Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., and Shukla, P.R.; others (2022). Global Warming of 1.5° C: IPCC Special Report on Impacts of Global Warming of 1.5° C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty (Cambridge University Press).
3. 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. (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 (Cambridge University Press).
4. Pörtner, H.-O., Roberts, D.C., Adams, H., Adler, C., Aldunce, P., Ali, E., Begum, R.A., Betts, R., Kerr, R.B., Biesbroek, R., et al. (2022). Climate Change 2022: Impacts, Adaptation, and Vulnerability. In Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC Geneva).
5. Stevens, A.R.H., Bellstedt, S., Elahi, P.J., and Murphy, M.T. (2020). The imperative to reduce carbon emissions in astronomy. Nat. Astron. *4*, 843–851.
6. Jahnke, K., Fendt, C., Fouesneau, M., Georgiev, I., Herbst, T., Kaasinen, M.,

Kossakowski, D., Rybizki, J., Schlecker, M., Seidel, G., et al. (2020). An astronomical institute's perspective on meeting the challenges of the climate crisis. Nat. Astron. *4*, 812–815.
7. Grealey, J., Lannelongue, L., Saw, W.-Y., Marten, J., Méric, G., Ruiz-Carmona, S., and Inouye, M. (2022). The carbon footprint of bioinformatics. Mol. Biol. Evol. *39*, msac034.
8. Bloom, K., Boisvert, V., Britzger, D., Buuck, M., Eichhorn, A., Headley, M., Lohwasser, K., and Merkel, P. (2022). Climate impacts of particle physics. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2203.12389.
9. Lecun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. Nature *521*, 436–444. https://doi.org/10.1038/nature14539.
10. Sevilla, J., Heim, L., Ho, A., Besiroglu, T., Hobbhahn, M., and Villalobos, P. (2022). Compute trends across three eras of machine learning. In 2022 International Joint Conference on Neural Networks (IJCNN), pp. 1–8.
11. Strubell, E., Ganesh, A., and McCallum, A. (2020). Energy and policy considerations for modern deep learning research. In Proceedings of the AAAI Conference on Artificial Intelligence, pp. 13693–13696.
12. Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D., Texier, M., and Dean, J. (2021). Carbon emissions and large neural network training. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2104.10350.
13. Luccioni, A.S., Viguier, S., and Ligozat, A.-L. (2022). Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2211.02001.
14. Schwartz, R., Dodge, J., Smith, N.A., and Etzioni, O. (2020). Green AI. Commun. ACM *63*, 54–63.
15. Fujinuma, N., DeCost, B., Hattrick-Simpers, J., and Lofland, S.E. (2022). Why big data and compute are not necessarily the path to big materials science. Commun. Mater. *3*, 59.
16. Probst, D. (2023). Aiming beyond slight increases in accuracy. Nat. Rev. Chem *7*, 227–228.
17. Deagen, M.E., Brinson, L.C., Vaia, R.A., and Schadler, L.S. (2022). The materials tetrahedron has a "digital twin.". MRS Bull. *47*, 379–388.
18. Agrawal, A., and Choudhary, A. (2016). Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. Apl. Mater. *4*, 053208.
19. Kipf, T.N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv1609.02907.
20. Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., and Dahl, G.E. (2017). Neural message passing for quantum chemistry. In International conference on machine learning, pp. 1263–1272.
21. Schütt, K., Kindermans, P.-J., Sauceda Felix, H.E., Chmiela, S., Tkatchenko, A., and Müller, K.-R. (2017). Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. Adv. Neural Inf. Process. Syst. *30*.
22. Xie, T., and Grossman, J.C. (2018). Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. Phys. Rev. Lett. *120*, 145301.
23. Chen, C., Ye, W., Zuo, Y., Zheng, C., and Ong, S.P. (2019). Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. Chem. Mater. *31*, 3564–3572.
24. Louis, S.-Y., Zhao, Y., Nasiri, A., Wang, X., Song, Y., Liu, F., and Hu, J. (2020). Graph convolutional neural networks with global attention for improved materials property prediction. Phys. Chem. Chem. Phys. *22*, 18141–18148.
25. Karamad, M., Magar, R., Shi, Y., Siahrostami, S., Gates, I.D., and Barati Farimani, A. (2020). Orbital graph convolutional neural network for material property prediction. Phys. Rev. Mater. *4*, 093801.
26. Choudhary, K., and DeCost, B. (2021). Atomistic Line Graph Neural Network for improved materials property predictions. npj Comput. Mater. *7*, 185–188.
27. Omee, S.S., Louis, S.-Y., Fu, N., Wei, L., Dey, S., Dong, R., Li, Q., and Hu, J. (2022). Scalable deeper graph neural networks for high-performance materials property prediction. Patterns *3*, 100491.
28. Yan, K., Liu, Y., Lin, Y., and Ji, S. (2022). Periodic graph transformers for crystal material property prediction. Adv. Neural Inf. Process. Syst. *35*, 15066–15080.
29. Kang, Y., Park, H., Smit, B., and Kim, J. (2023). A multi-modal pre-training transformer for universal transfer learning in metal–organic frameworks. Nat. Mach. Intell. *5*, 309–318.
30. Burner, J., Luo, J., White, A., Mirmiran, A., Kwon, O., Boyd, P.G., Maley, S., Gibaldi, M., Simrod, S., Ogden, V., and Woo, T.K. (2023). ARC–MOF: a diverse database of metal-organic frameworks with DFT-derived partial atomic charges and descriptors for machine learning. Chem. Mater. *35*, 900–916.
31. Reiser, P., Neubert, M., Eberhard, A., Torresi, L., Zhou, C., Shao, C., Metni, H., van Hoesel, C., Schopmans, H., Sommer, T., and Friederich, P. (2022). Graph neural networks for materials science and chemistry. Commun. Mater. *3*, 93.
32. Thompson, N.C., Greenewald, K., Lee, K., and Manso, G.F. (2021). Deep learning's diminishing returns: The cost of improvement is becoming unsustainable. IEEE Spectr. *58*, 50–55.
33. Akbari, H. (2002). Shade trees reduce building energy use and CO2 emissions from power plants. Environ. Pollut. *116*, S119–S126.
34. Scao, T.L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., Hesslow, D., Castagné, R., Luccioni, A.S., Yvon, F., Gallé, M., et al. (2022). Bloom: A 176b-parameter open-access multilingual language model. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2211.05100.
35. Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv1810.04805.
36. Korolev, V., and Protsenko, P. (2023). Accurate, interpretable predictions of materials properties within transformer language models. Patterns *4*, 100803.
37. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I.; others (2019). Language models are unsupervised multitask learners. OpenAI Blog *1*, 9.
38. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. Adv. Neural Inf. Process. Syst. *33*, 1877–1901.
39. Castro Nascimento, C.M., and Pimentel, A.S. (2023). Do Large Language Models Understand Chemistry? A Conversation with ChatGPT. J. Chem. Inf. Model. *63*, 1649–1655.
40. Boiko, D.A., MacKnight, R., and Gomes, G. (2023). Emergent autonomous scientific research capabilities of large language models. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2304.05332.
41. Guo, T., Guo, K., Liang, Z., Guo, Z., Chawla, N.V., Wiest, O., and Zhang, X.; others (2023). What indeed can GPT models do in chemistry? A comprehensive benchmark on eight tasks. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2305.18365.
42. Yang, X., He, X., Liang, Y., Yang, Y., Zhang, S., and Xie, P. (2020). Transfer learning or self-supervised learning? A tale of two pretraining paradigms. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2007.04234.
43. Lee, S., Kim, B., Cho, H., Lee, H., Lee, S.Y., Cho, E.S., and Kim, J. (2021). Computational screening of trillions of metal–organic frameworks for high-performance methane storage. ACS Appl. Mater. Interfaces *13*, 23647–23654.
44. Chien, A.A., Lin, L., Nguyen, H., Rao, V., Sharma, T., and Wijayawardana, R. (2023). Reducing the Carbon Impact of Generative AI Inference (today and in 2035). In Proceedings of the 2nd Workshop on Sustainable Computer Systems, pp. 1–7.
45. Pandey, M., Fernandez, M., Gentile, F., Isayev, O., Tropsha, A., Stern, A.C., and Cherkasov, A. (2022). The transformational role of GPU computing and deep learning in drug discovery. Nat. Mach. Intell. *4*, 211–221.
46. Patterson, D., Gonzalez, J., Hölzle, U., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D.R., Texier, M., and Dean, J. (2022). The carbon footprint of machine learning training will plateau, then shrink. Computer *55*, 18–28.
47. Ahmed, N., and Wahed, M. (2020). The De-democratization of AI: Deep learning and the compute divide in artificial intelligence research. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2010.15581.
48. Korovin, A.N., Humonen, I.S., Samtsevich, A.I., Eremin, R.A., Vasilyev, A.I., Lazarev, V.D., and Budennyy, S.A. (2022). Boosting Heterogeneous Catalyst Discovery by Structurally Constrained Deep Learning Models. Preprint at arXiv *1*. https://doi.org/10.48550/arXiv2207.05013.
49. Baird, S.G., Liu, M., and Sparks, T.D. (2022). High-dimensional Bayesian optimization of 23 hyperparameters over 100 iterations for an attention-based network to predict materials property: A case study on CrabNet using Ax platform and SAASBO. Comput. Mater. Sci. *211*, 111505.
50. Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G., Aga, F., Huang, J., Bai, C., et al. (2022). Sustainable ai: Environmental implications, challenges and opportunities. Proc. Mach. Learn. Syst. *4*, 795–813.
51. Kaack, L.H., Donti, P.L., Strubell, E., Kamiya, G., Creutzig, F., and Rolnick, D. (2022). Aligning artificial intelligence with climate change mitigation. Nat. Clim. Change *12*, 518–527.
52. Jablonka, K.M., Rosen, A.S., Krishnapriyan, A.S., and Smit, B. (2023). An ecosystem for digital reticular chemistry. ACS Cent. Sci. *9*, 563–581.
53. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). Pytorch: An imperative style, high-performance deep learning library. Adv. Neural Inf. Process. Syst. *32*, 8026–8037.

54. Budennyy, S.A., Lazarev, V.D., Zakharenko, N.N., Korovin, A.N., Plosskaya, O.A., Dimitrov, D.V., Akhripkin, V.S., Pavlov, I.V., Oseledets, I.V., Barsola, I.S., et al. (2022). Eco2ai: carbon emissions tracking of machine learning models as the first step towards sustainable ai. In Doklady Mathematics, pp. S118–S128.

55. Janet, J.P., and Kulik, H.J. (2017). Resolving transition metal chemical space: Feature selection for machine learning and structure–property relationships. J. Phys. Chem. A *121*, 8939–8954.

56. Moosavi, S.M., Nandy, A., Jablonka, K.M., Ongari, D., Janet, J.P., Boyd, P.G., Lee, Y., Smit, B., and Kulik, H.J. (2020). Understanding the diversity of the metal-organic framework ecosystem. Nat. Commun. *11*, 1–11.

57. Georgiou, S., Kechagia, M., Sharma, T., Sarro, F., and Zou, Y. (2022). Green ai: Do deep learning frameworks have different costs? In Proceedings of the 44th International Conference on Software Engineering, pp. 1082–1094.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
ARC–MOF database	Burner et al.[30]	https://doi.org/10.5281/zenodo.6908727
Source code and data	This study	https://doi.org/10.5281/zenodo.10726802
Software and algorithms		
Python version 3.10	Python Software Foundation	https://www.python.org
Mofdscribe v0.0.7	Jablonka et al.[52]	https://github.com/kjappelbaum/mofdscribe
PyTorch v1.13.1	Paszke et al.[53]	https://github.com/pytorch/pytorch
Eco2AI v0.3.9	Budennyy et al.[54]	https://github.com/sb-ai-lab/Eco2AI
GCN	Kipf et al.[19]	https://github.com/Fung-Lab/MatDeepLearn
MPNN	Gilmer et al.[20]	https://github.com/Fung-Lab/MatDeepLearn
SchNet	Schütt et al.[21]	https://github.com/Fung-Lab/MatDeepLearn
CGCNN	Xie et al.[22]	https://github.com/Fung-Lab/MatDeepLearn
MEGNet	Chen et al.[23]	https://github.com/Fung-Lab/MatDeepLearn
GATGNN	Louis et al.[24]	https://github.com/usccolumbia/deeperGATGNN
OGCNN	Karamad et al.[25]	https://github.com/RishikeshMagar/OGCNN
ALIGNN	Choudhary et al.[26]	https://github.com/usnistgov/alignn
DeeperGATGNN	Omee et al.[27]	https://github.com/usccolumbia/deeperGATGNN
Matformer	Yan et al.[28]	https://github.com/YKQ98/Matformer
MOFTransformer	Kang et al.[29]	https://github.com/hspark1212/MOFTransformer

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Vadim Korolev (korolev@colloid.chem.msu.ru).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The data accompanying this study were made available on Zenodo at https://doi.org/10.5281/zenodo.10726802.
- The source code accompanying this study were made available on Zenodo at https://doi.org/10.5281/zenodo.10726802.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Dataset

A subset of diverse crystal structures was taken from the *ab initio* REPEAT charge MOF[30] (ARC−MOF) database; the following procedure was applied to select appropriate entities. First, we decided to consider the structures for which the corresponding target value—CO_2 working capacity in a post-combustion vacuum swing adsorption (VSA) process—was greater than 0.05 mmol g^{-1}. By doing this, we eliminated unphysically calculated values (negative working capacity) and outliers on a logarithmic scale; the distribution of the target property was initially lognormal (Figure S3). Detailed information on relevant grand canonical Monte Carlo (GCMC) simulations was provided in the original study.[30] To reduce the computational expense, we did not consider structures that contain more than 200 atoms per unit cell. Second, revised

autocorrelations[55,56] (RACs), as implemented in the mofdscribe library,[52] were calculated for each valid structure. Finally, the maximin greedy algorithm was used for the subset selection: starting with an empty set, a new entity was added to maximize the distance in feature space (formed by 480-dimensional RAC vectors) from the current entities. The final ARC−MOF subset containing 20,000 structures was split into training (80%), validation (10%), and test (10%) sets. The base-10 logarithm of CO_2 working capacity was predicted by the supervised learning algorithms of interest.

Models

The following supervised learning algorithms were benchmarked: Graph Convolutional Network[19] (GCN), Message Passing Neural Network[20] (MPNN), SchNet,[21] Crystal Graph Convolutional Neural Network[22] (CGCNN), MatErials Graph Network[23] (MEGNet), Global ATtention-based Graph convolutional Neural Network[24] (GATGNN), Orbital Graph Convolutional Neural Network[25] (OGCNN), Atomistic Line Graph Neural Network[26] (ALIGNN), Global ATtention-based GNN with differentiable group normalization and residual connection[27] (DeeperGATGNN), Matformer,[28] and multi-modal Transformer encoder pretrained with 1 million hypothetical MOFs[29] (MOFTransformer). All models were trained with hyperparameters provided in the original articles. We used PyTorch-based[53] implementations exclusively, taking into account the substantial effect of a framework (PyTorch vs. TensorFlow) on energy and runtime performance.[57] All models (except MOFTransformer) were trained for 1000 (300) epochs with early stopping at 100. GHG emissions of model training and inference were measured by means of the Eco2AI package,[54] recommending the emission intensity coefficient of 240.56 $kgCO_2e$ per MW·h (Moscow). All calculations were carried out on a workstation equipped with two Intel Xeon CPUs E5-2695 v4 @ 2.10 GHz, 144 GB RAM, and NVIDIA GeForce RTX 3090 Ti.

Literature analysis

We examined manuscripts submitted to arXiv from 1 January 2016 to 25 April 2023 that satisfied the following conditions: ("cond-mat.mtrl-sci" IN Subjects) AND ("deep learning" IN Title OR "deep learning" IN Abstract OR "neural network?" IN Title OR "neural network?" IN Abstract). The deposited preprints were not peer-reviewed beforehand, and therefore, to ensure scientific validity, we included only preprints that were published later in a journal. A total of 317 manuscripts met all the aforementioned criteria. Both authors of the present study independently curated the chosen preprints to examine the following aspects: the model of a GPU accelerator used, estimates of training time, and the level of GHG emissions.