Tackling Climate Change with Machine Learning

David Rolnick1*, Priya L. Donti2, Lynn H. Kaack3, Kelly Kochanski4, Alexandre Lacoste5, Kris Sankaran6,7, Andrew Slavin Ross8, Nikola Milojevic-Dupont9,10, Natasha Jaques11, Anna Waldman-Brown11, Alexandra Luccioni6,7, Tegan Maharaj6,7, Evan D. Sherwin2, S. Karthik Mukkavilli6,7, Konrad P. Kording1, Carla Gomes12, Andrew Y. Ng13, Demis Hassabis14, John C. Platt15, Felix Creutzig9,10, Jennifer Chayes16, Yoshua Bengio6,7

1University of Pennsylvania, 2Carnegie Mellon University, 3ETH Zürich, 4University of Colorado Boulder, 5Element AI, 6Mila, 7Université de Montréal, 8Harvard University, 9Mercator Research Institute on Global Commons and Climate Change, 10Technische Universität Berlin, 11Massachusetts Institute of Technology, 12Cornell University, 13Stanford University, 14DeepMind, 15Google AI, 16Microsoft Research

Abstract

Climate change is one of the greatest challenges facing humanity, and we, as machine learning experts, may wonder how we can help. Here we describe how machine learning can be a powerful tool in reducing greenhouse gas emissions and helping society adapt to a changing climate. From smart grids to disaster management, we identify high impact problems where existing gaps can be filled by machine learning, in collaboration with other fields. Our recommendations encompass exciting research questions as well as promising business opportunities. We call on the machine learning community to join the global effort against climate change.

Introduction

The effects of climate change are increasingly visible.2 Storms, droughts, fires, and flooding have become stronger and more frequent. Global ecosystems are changing, including the natural resources and agriculture on which humanity depends. The 2018 intergovernmental report on climate change estimated that the world will face catastrophic consequences unless global greenhouse gas emissions are eliminated within thirty years. Yet year after year, these emissions rise.

Addressing climate change involves mitigation (reducing emissions) and adaptation (preparing for unavoidable consequences). Both are multifaceted issues. Mitigation of greenhouse gas (GHG) emissions requires changes to electricity systems, transportation, buildings, industry, and land use. Adaptation requires climate modeling, risk prediction, and planning for resilience and disaster management. Such a diversity of problems can be seen as an opportunity: there are many ways to have an impact.

In recent years, machine learning (ML) has been recognized as a broadly powerful tool for technological progress. Despite the growth of movements applying ML and AI to problems of societal and global good2
Table 1: Climate change solution domains, along with areas of ML that are relevant to each. Rows of the table correspond to sections of this paper. This table should not be seen as comprehensive.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Computer Vision</th>
<th>NLP</th>
<th>Time-series analysis</th>
<th>Unsupervised learning</th>
<th>RL &amp; Control</th>
<th>Causal Inference</th>
<th>Uncertainty Quantification</th>
<th>Transfer Learning</th>
<th>Interpretable ML</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electricity Systems</td>
<td>1</td>
<td>1.1</td>
<td>1.1</td>
<td>1</td>
<td>1.1</td>
<td>1.2</td>
<td>1.2</td>
<td>1.3</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>Transportation</td>
<td>2.1</td>
<td>2.2</td>
<td>2</td>
<td>2.1</td>
<td>2.1</td>
<td>2.4</td>
<td>2.4</td>
<td>2.1</td>
<td>2.4</td>
<td>2</td>
</tr>
<tr>
<td>Buildings &amp; Cities</td>
<td>3.2</td>
<td>3.3</td>
<td>3.5</td>
<td>3.1</td>
<td>3.1</td>
<td>3.3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Industry</td>
<td>4.1</td>
<td>4.3</td>
<td>4.3</td>
<td>4</td>
<td>4.2</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td>Farms &amp; Forests</td>
<td>5.1</td>
<td>5.3</td>
<td>5.2</td>
<td>5.4</td>
<td>5.2</td>
<td>5.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO₂ Removal</td>
<td>6.3</td>
<td></td>
<td></td>
<td></td>
<td>6.3</td>
<td>6.3</td>
<td>6.3</td>
<td>6.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate Prediction</td>
<td>7.1</td>
<td>7</td>
<td></td>
<td></td>
<td>7.3</td>
<td>7.3</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Societal Impacts</td>
<td>8.1</td>
<td>8.4</td>
<td>8.3</td>
<td>8.2</td>
<td>8.2</td>
<td>8.2</td>
<td>8.2</td>
<td>8.1</td>
<td>8.3</td>
<td></td>
</tr>
<tr>
<td>Solar Geoengineering</td>
<td>9.3</td>
<td>9.3</td>
<td>9.4</td>
<td>9.3</td>
<td>9.3</td>
<td>9.4</td>
<td>9.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tools for Individuals</td>
<td>10.1</td>
<td>10.1</td>
<td>10.2</td>
<td>10.3</td>
<td>10.2</td>
<td>10.1</td>
<td>10.2</td>
<td>10.2</td>
<td></td>
<td>10.2</td>
</tr>
<tr>
<td>Tools for Society</td>
<td>11.1</td>
<td>11.2</td>
<td>11.3</td>
<td>11.2</td>
<td>11.1</td>
<td>11.1</td>
<td>11.1</td>
<td>11.1</td>
<td>11.1</td>
<td>11.1</td>
</tr>
<tr>
<td>Education</td>
<td>12.2</td>
<td>12.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12.1</td>
</tr>
<tr>
<td>Finance</td>
<td>13.2</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
</tr>
</tbody>
</table>

There remains the need for a concerted effort to identify how these tools may best be applied to climate change. Many ML practitioners wish to act, but are uncertain how. On the other side, many fields have begun actively seeking input from the ML community.

This paper aims to provide an overview of where machine learning can be applied with high impact in the fight against climate change, through either effective engineering or innovative research. The solutions we highlight include climate mitigation and adaptation, as well as meta-level tools that enable other solutions. In order to maximize the relevance of our recommendations, we have consulted experts across many fields (see Acknowledgments) in the preparation of this paper.

movement (e.g. [8–12]). Faghmous and Kumar presented an overview of climate change problems from the perspective of big data [13], and Kaack recently presented an overview of ML applications to climate mitigation [14]. Climate informatics specifically considers the problem of applying ML to climate modeling [15, 16], which we consider in §7. Ford et al. also call for applications of big data to climate change adaptation domains including vulnerability assessment, early warning, and monitoring and evaluation [17], topics which we consider in §8.
Who is this paper written for?

We believe that our recommendations will prove valuable to several different audiences (detailed below). In our writing, we have assumed some familiarity with basic terminology in machine learning, but do not assume any prior familiarity with application domains (such as agriculture or electric grids).

Researchers and engineers: We identify many problems that require conceptual innovation and can advance the field of ML, as well as being highly impactful. For example, we highlight how weather models afford an exciting domain for interpretable ML (see §7.1). We encourage researchers and engineers across fields to use their expertise in solving urgent problems relevant to society.

Entrepreneurs and investors: We identify many problems where existing ML techniques could have a major impact without further research, and where the missing piece is deployment. We realize that some of the recommendations we offer here will make valuable startups and nonprofits. For example, we highlight techniques for providing fine-grained solar forecasts for power companies (see §1.1.1), tools for helping reduce personal energy consumption (see §10.3), and predictions for the financial impacts of climate change (see §13). We encourage entrepreneurs and investors to fill what is currently a wide-open space.

Corporate leaders: We identify problems where ML can lead to massive efficiency gains if adopted at scale by corporate players. For example, we highlight means of optimizing supply chains to reduce waste (see §4.1) and software/hardware tools for precision agriculture (see §5.2). We encourage corporate leaders to take advantage of opportunities offered by ML to benefit both the world and the bottom line.

Local and national governments: We identify problems where ML can improve public services, help gather data for decision-making, and guide plans for future development. For example, we highlight intelligent transportation systems (see §2.1, 2.4), techniques for automatically assessing the energy consumption of buildings in cities (see §3.2), and tools for improving disaster management (see §8.4). We encourage governments to consult ML experts while planning infrastructure and development, as this can lead to better, more cost-effective outcomes. We further encourage public entities to release data that may be relevant to climate change mitigation and adaptation goals.

How to read this paper

The paper is broken into sections according to application domain (see Table 1). To help the reader, we have also included the following flags at the level of individual solutions.

- **High Leverage** denotes bottlenecks that domain experts have identified in climate change mitigation or adaptation and that we believe to be particularly well-suited to tools from ML. These solutions may be especially fruitful for ML practitioners wishing to have an outsized impact, though applications not marked with this flag are also valuable and should be pursued.

- **Long-term** denotes solutions that will have their primary impact after 2040. Such solutions are neither more nor less important than short-term solutions – both are necessary.

- **High Risk** denotes solutions that are risky in one of the following ways: (i) the technology involved is uncertain and may ultimately not succeed, (ii) there is uncertainty as to the impact on GHG emissions (for example, the Jevons paradox may apply), or (iii) there is the potential for unwanted side effects (negative externalities).

---

3 Approximate cost-benefit analyses for some of these are considered e.g. in [18].

4 The Jevons paradox in economics refers to a situation where increased efficiency nonetheless results in higher overall demand. For example, autonomous vehicles could cause people to drive far more, so that overall GHG emissions could increase even if each ride is more efficient. In such cases, it becomes especially important to make use of specific policies, such as carbon pricing, to direct new technologies and the ML behind them. See also the literature on rebound effects and induced demand.
These flags should not be taken as definitive; they represent our understanding of more rigorous analyses within the domains we consider, combined with our subjective evaluation of the potential role of ML in these various applications.

Despite the length of the paper, we cannot cover everything. There will certainly be many solutions that we have not considered, or that we have erroneously dismissed. We look forward to seeing where future work leads.

A call for collaboration

All of the problems we highlight in this paper require collaboration across fields. As the language used to refer to problems often varies between disciplines, we have provided keywords and background reading within each section of the paper. Finding collaborators and relevant data can sometimes be difficult; for additional resources, please visit the website that accompanies this paper (https://www.climatechange.ai/).

Collaboration makes it easier to develop effective solutions. Working with domain experts reduces the chance of using powerful tools when simple tools will do the job, of working on a problem that isn’t actually relevant to practitioners, of overly simplifying a complex issue, or of failing to anticipate risks.

Collaboration can also help ensure that new work reaches the audience that will use it. To be impactful, ML code should be accessible and published using a language and a platform that are already popular with the intended users. For maximal impact, new code can be integrated into an existing, widely used tool.

We emphasize that machine learning is not a silver bullet. The applications we highlight are impactful, but no one solution will “fix” climate change. There are also many areas of action where ML is inapplicable, and we omit these entirely. Furthermore, technology alone is not enough – technologies that would reduce climate change have been available for years, but have largely not been adopted at scale by society. While we hope that ML will be useful in reducing the costs associated with climate action, humanity also must decide to act.
Mitigation

1 Electricity Systems by Priya L. Donti

AI has been called the new electricity, due to its potential to transform entire industries. Interestingly, electricity itself is one of the industries that AI is poised to transform. Many electricity systems are awash in data, and the industry has begun to envision next-generation systems (smart grids) driven by AI and ML.

Electricity systems are currently responsible for about a quarter of human-caused greenhouse gas emissions. Moreover, as buildings, transportation, and other sectors seek to reduce their emissions impacts by replacing traditional fuels, demand for low-carbon electricity will grow even further. To reduce the impact of electricity systems across the globe, society must

- Rapidly transition to low-carbon electricity sources (such as solar, wind, hydro, and nuclear) and phase out carbon-emitting sources (such as coal, natural gas, and other fossil fuels).
- Reduce emissions from existing carbon-emitting power plants, since the transition to low-carbon fuels will not happen overnight.
- Implement these changes across all countries and contexts, as electricity systems are everywhere.

ML can contribute on all fronts by informing the research, deployment, and operation of electricity system technologies. Such contributions include accelerating the development of clean energy technologies, improving forecasts of demand and clean energy, improving system optimization and management, and enhancing system monitoring. These solutions require a variety of ML paradigms and techniques, and warrant working closely with domain experts to integrate insights from operations research, electrical engineering, physics, chemistry, the social sciences, and other fields.

1.1 Enabling low-carbon electricity

Low-carbon electricity sources are essential to fighting climate change. These sources come in two forms: variable and controllable. Variable sources fluctuate based on external factors: for instance, solar panels produce power only when the sun is shining, and wind turbines only when the wind is blowing. On the other hand, controllable sources such as nuclear or geothermal plants can be turned on and off (though they are subject to physical constraints on how quickly they can change their power output). These two types of sources have different implications for how electricity systems are run, and thus present distinct opportunities for ML techniques.

---

5This oft-cited analogy can be found e.g. in the following talk at the Stanford Graduate School of Business in 2017: https://www.youtube.com/watch?v=21E1KfQYzxc.

6Throughout this section, we use the term “electricity systems” to refer to the procurement of fuels and raw materials for electric grid components; the generation and storage of electricity; and the delivery of electricity to end-use consumers. For primers on these topics, see [22, 23].

7We use the term “low-carbon” here instead of “renewable” because of this paper’s explicit focus on climate change goals. Renewable energy is produced from inexhaustible energy sources such as the sun, wind, or water, but need not necessarily be carbon-free (as in the case of biomass). Similarly, not all low-carbon energy is renewable (as in the case of nuclear energy).

8While technically controllable, nuclear power plants are often viewed as inflexible in practice: since they can take hours or days to turn on or off, they are often left on (at full capacity) to operate as base load. That said, nuclear power plants may have some flexibility to change their power generation for load-following and other grid services, as in the case of France [27].
1.1.1 Variable sources

Most electricity is delivered to consumers using a physical network called the electric grid, where the power generated must equal the power consumed at every moment. This implies that for every solar panel, wind turbine, or other variable electricity generator, there is some mix of natural gas plants, storage, or other controllable sources ready to buffer unexpected changes in its output (which occur, for example, when there are unexpected clouds blocking the sun or the wind blows less strongly than was forecast). Today, this buffer is often provided by natural gas plants run in a standby mode that causes them to release CO$_2$ even when not producing any power. In (future) low-carbon systems with a high reliance on variable resources, much of this buffer may need to be provided by energy storage technologies such as batteries, pumped hydro, or power-to-gas. However, managing such highly variable systems is complex, and system operators may be unable to transition towards this low-carbon future without improvements in key technologies. ML can both reduce emissions from today’s standby generators and enable the transition to carbon-free systems by helping improve necessary technologies (namely forecasting, scheduling, and control) and by helping create advanced electricity markets that accommodate both variable electricity and flexible demand.

Generation and demand forecasting

Since variable generation and electricity demand both fluctuate, they must be forecast ahead of time to inform real-time electricity scheduling and longer-term system planning. Better short-term forecasts can improve electricity scheduling, enabling operators to both reduce their reliance on polluting standby plants and proactively manage increasing amounts of variable sources. Better long-term forecasts can improve system planning, helping operators understand where and how many variable plants should be built. While many system operators today use basic forecasting techniques, forecasts will need to become increasingly accurate, span multiple horizons in time and space, and better quantify uncertainty to support these use cases.

To date, many ML and deep learning methods have been applied to power generation and demand forecasting. These methods have employed historical data, physical model outputs, images, and even video data to create short- to medium-term forecasts of solar power, wind power, hydro power, demand, or more than one of these at aggregate spatial scales. These methods span various types of supervised machine learning, fuzzy logic, and hybrid physical models, and take different approaches to quantifying (or not quantifying) uncertainty. At a more spatially granular level, some demand forecasting work has attempted to understand specific categories of demand, for instance by using clustering techniques on households or using game theory, optimization, regression, and/or online learning to predict disaggregated quantities from aggregate electricity signals.

While much of this previous work has used domain-agnostic techniques, ML algorithms of the future will need to meaningfully incorporate domain-specific insights. For instance, since weather fundamentally drives both variable generation and electricity demand, ML algorithms forecasting these quantities should draw from innovations in climate modeling and weather forecasting and in hybrid physics-plus-ML modeling techniques. Such techniques can help improve short- to medium-term forecasts, and are also necessary for ML to contribute to longer-term (e.g. year-scale) forecasts since weather distributions shift over time. In addition to incorporating system physics, ML models should also directly optimize for system goals. For instance, the authors of use a deep neural network to produce demand forecasts that optimize for electricity scheduling costs rather than forecast accuracy (assuming scheduling is automated); this notion could be extended to produce forecasts that minimize GHG emissions. In non-automated settings where power system control engineers (partially) determine how much power each gen-

---

9This standby mode is called spinning reserve.

10It is worth noting that in systems with many fossil fuel plants, storage can actually increase emissions depending on how it is operated.29,30.
erator should produce, explainable ML and automated visualization techniques could help engineers better understand forecasts and thus improve how they schedule low-carbon generators. More broadly, understanding the domain value of improved forecasts is an interesting challenge for ML. For example, previous work has characterized the benefits of specific solar forecast improvements in a region of the United States [61]; further study in different contexts and for different types of improvements could help better direct the attention of the ML community within the forecasting space.

**Improving scheduling and flexible demand**

When balancing electricity systems, system operators use a process called *scheduling and dispatch* to determine how much power each generator should produce. This process is slow and complex, as it is governed by NP-hard optimization problems that need to be coordinated across multiple time scales (from sub-second to days ahead). However, scheduling becomes even more complex in a system with variable generators, storage, and *flexible demand*, since operators will need to manage even more system components while simultaneously solving scheduling problems more quickly to account for real-time variations in electricity production. Scheduling processes must therefore improve significantly for operators to manage systems with a high reliance on variable sources.

ML can help improve the existing (centralized) process of scheduling and dispatch by speeding up power system optimization problems. A great deal of work primarily in optimization, but also using techniques such as neural networks, genetic algorithms, and fuzzy logic [62], has focused on improving the tractability of power system optimization problems. ML could also be used to fit fast function approximators to existing optimization problems or to find good starting points for optimization. Dynamic scheduling [63,64] and safe reinforcement learning could also be used to balance the electric grid in real time to accommodate variable generation or demand; in fact, some electricity system operators have started to pilot similar methods at small, test case-based scales.

While many modern electricity systems are centrally coordinated, recent work has examined how to (at least partially) *decentralize* scheduling and dispatch, primarily through advanced electricity markets. In particular, storage and flexible demand can balance the electricity system by responding to real-time prices that reflect (for example) how much variable electricity is available. ML or potentially even simpler techniques can enable flexible demand by helping storage and smart devices automatically respond to electricity prices; previous work has optimized storage and/or flexible demand using techniques such as agent-based models [68–71], online optimization [72], and dynamic programming [73]. To provide appropriate signals for flexible demand, system operators can design electricity prices based on e.g. forecasts of variable electricity (see the discussion above) or grid emissions (see §1.2). Previous work has used dynamic programming to set real-time electricity prices that maximize revenue [74], and similar techniques could be applied to create prices that instead optimize for GHG emissions. In general, much more work is needed to test and scale existing decentralized solutions; barring deployment on real systems, platforms such as

---

11. This *scheduling and dispatch* process is used in both *regulated* electricity networks (where a central entity more-or-less completely determines which power plants will produce power) and *deregulated* electricity networks (where power plants bid to determine who will produce power, but oversight for balancing the electricity system ultimately falls upon a system operator).

12. For instance, *unit commitment* is a (large) mixed integer program that includes nonlinear power flow constraints. System operators often use simpler variants of this problem, but even these are often computationally complex.

13. For instance, researchers and entrepreneurs have proposed distributed energy trading or management systems that enable variable low-carbon energy. Some of these suggested distributed systems include transactive energy systems [63,64] and peer-to-peer energy trading systems [67]. Some companies also propose to run truly distributed markets within existing electricity market structures (see e.g. Camus Energy: [https://camus.energy/]).

14. Some modern electricity operators are already beginning to procure autonomous flexible demand for grid-balancing purposes. For instance, the Southern California Edison electricity supply company has procured 50 MW of *demand response* from Nest smart thermostats to prevent electricity blackouts. See [https://www.greentechmedia.com/articles/read/inside-nest-50000-home-virtual-power-plant-for-southern-california-edison]
PowerTAC\textsuperscript{15} can provide large-scale simulated electricity markets on which to perform these tests.

**Accelerated science for materials**

Scientists are working to develop new materials that can better store or otherwise harness energy from variable natural resources. For instance, creating solar fuels (synthetic fuels produced from sunlight or solar heat) could allow us to capture solar energy when the sun is shining and then store this energy for later use. However, the process of discovering new materials can be slow and imprecise; the physics behind materials are not completely understood, so human experts often manually apply heuristics to understand a proposed material’s physical properties \cite{75, 76}. ML techniques can automate this process by combining existing heuristics with experimental data, physics, and reasoning to apply and even extend existing physical knowledge. For instance, recent work has used tools from ML, AI, optimization, and physics to figure out a proposed material’s crystal structure, with the goal of accelerating materials discovery for solar fuels \cite{76–78}. Other work seeking to improve battery storage technologies has combined first-principles physics calculations with support-vector regression to design conducting solids for lithium-ion batteries \cite{79}. (Additional applications of ML to batteries are discussed in \textsection 2.3.)

More generally in materials science, ML techniques including supervised learning, active learning, and generative models have been applied to the synthesis, characterization, modeling, and design of new and existing materials, as described in reviews \cite{75, 80} and more recent work \cite{81}. As discussed in \cite{75}, novel challenges for ML in materials science include coping with moderately sized datasets and producing interpretable predictions that shed light on the physical laws learned by a model; for example, the authors of \cite{82} analyze the gradients of a convolutional neural network trained on materials data to understand the rules it learned. ML can also help enable accelerated materials science by informing relevant innovation policies; for instance, previous work has applied NLP to patent data to understand the solar panel innovation process \cite{83}. We note that while our focus here has been on electricity system applications, ML for accelerated science may also have significant impacts outside electricity systems, e.g. by helping design alternatives to cement (\textsection 4.2) or create better CO\textsubscript{2} sorbents (\textsection 6.2).

**Additional applications**

While the applications discussed above will likely provide some of the larger gains for variable low-carbon sources, there are a multitude of additional applications for ML in this space. For instance, it is important to ensure that low-carbon variable generators capture ambient energy as efficiently as possible. Prior work has attempted to maximize electricity production by controlling movable solar panels \cite{84, 85} or wind turbine blades \cite{86} using RL or Bayesian optimization. Other work has used graphical models to detect faults in rooftop solar panels \cite{87}, genetic algorithms to optimally place wind turbines within a wind farm \cite{88}, and multi-objective optimization to place hydropower dams in a way that satisfies both energy and ecological objectives \cite{89}.

ML can also help integrate rooftop solar panels into the electric grid, particularly in the United States and Europe. Rooftop solar panels are connected to a part of the electric grid called the distribution grid, which traditionally did not have many sensors because it was only used to deliver electricity “one-way” from centralized power plants to consumers. However, rooftop solar and other distributed energy resources have created a “two-way” flow of electricity on distribution grids. Since the locations and sizes of rooftop solar panels are not often known to electricity system operators, previous work has used computer vision techniques on satellite imagery to generate size and location data for rooftop solar panels \cite{90, 91}. Further, since electricity system operators need to ensure that the distribution system is stable despite lack of sensing, recent work has employed techniques such as matrix completion and deep neural networks for distribution system state estimation and forecasting \cite{92–94}.

\textsuperscript{15}https://powertac.org/
1.1.2 Nuclear fission and fusion

Low-carbon electricity sources such as nuclear fission and nuclear fusion can help achieve climate change goals while requiring very few changes to how the electric grid is run. However, nuclear fission faces many practical challenges, and nuclear fusion is not yet viable. ML can support these technologies by helping mitigate some challenges faced by fission plants while helping accelerate the development of fusion plants.

Nuclear power plants

Some argue that nuclear fission reactors (also known as nuclear power plants) are essential to meeting climate change goals [95], but these technologies face significant challenges including public safety, waste disposal, slow technological learning [96, 97], and high costs [95]. ML can help with a small piece of the latter problem by reducing maintenance costs; specifically, deep networks can speed up inspections by detecting cracks and anomalies from image and video data [98] or by preemptively detecting faults from high-dimensional sensor and simulation data [99]. The authors of [100] speculate that ML and high performance computing could also be used to help design next-generation nuclear reactors or simulate nuclear waste disposal options.

Nuclear fusion

Nuclear fusion reactors have the potential to produce safe and carbon-free electricity using a virtually limitless hydrogen fuel supply, but currently consume more energy than they produce [101]. While considerable scientific and engineering research is still needed, ML can help accelerate this work by guiding experimental design and monitoring physical processes. Fusion reactors require intelligent experimental design because they have a large number of tunable parameters; ML can help prioritize which parameter configurations should be explored during physical experiments. For instance, Google and TAE Technologies have developed a human-in-the-loop experimental design algorithm enabling rapid parameter exploration for TAE’s reactor [102].

Physically monitoring fusion reactors is also an important application for ML. Modern reactors attempt to super-heat hydrogen into a plasma state and then stabilize it, but during this process, the plasma may experience rapid instabilities that damage the reactor. Prior work has tried to preemptively detect disruptions for tokamak reactors, using supervised learning methods such as support-vector machines, adaptive fuzzy logic, decision trees, and deep learning [103–108] on previous disruption data. While many of these methods are tuned to work on individual reactors, recent work has shown that deep learning may enable insights that generalize to multiple reactors [108]. More generally, rather than simply detecting disruptions, scientists need to understand how plasma’s state evolves over time, e.g. by finding the solutions of time-dependent magnetohydrodynamic equations [109]; speculatively, ML could help characterize this evolution and even help steer plasma into safe states through reactor control. ML models for such fusion applications would likely employ a combination of simulated and experimental data, and would need to account for the different physical characteristics, data volumes, and simulator speeds or accuracies associated with different reactor types.

1.2 Reducing current-system climate impacts

While switching to low-carbon electricity sources will be essential, in the meantime, it will also be important to mitigate emissions from the electricity system as it currently stands. Some methods for mitigating current-

---

Footnotes:

16For an overview of nuclear fusion, please see this collection of articles in Nature Physics: https://www.nature.com/collections/bccqhmkbwyw

17Plasma simulation frameworks for tokamak reactors include RAPTOR [110, 111], ASTRA [112], CRONOS [113], PTRANSP [114], and IPS [115].
system impacts include cutting emissions from fossil fuels, reducing waste from electricity delivery, and flexibly managing demand to minimize its emissions impacts.

Reducing life-cycle fossil fuel emissions

Reducing emissions from fossil fuel power generation is a necessary stop gap while society transitions towards low-carbon electricity. In particular, ML can help prevent the leakage of methane (an extremely potent greenhouse gas) from natural gas pipelines and compressor stations. Previous and ongoing work has used sensor and/or satellite data to proactively suggest pipeline maintenance\(^{18}\) or detect existing leaks (see \(^{19}\) the SLED project\(^{19}\) and Bluefield Technologies\(^{20}\)), and there is a great deal of opportunity in this space to improve and scale existing solutions. In addition to leak detection, ML can help reduce emissions from freight transportation of solid fuels (\(^{12}\)) and may also have applications in the sequestration of CO\(_2\) from power plant flue gas (\(^{6.3}\)). In all these cases, solutions should be pursued with great care so as not to impede or prolong the transition to a low-carbon electricity system.

Reducing system waste

As electricity gets transported from generators to consumers, some of it gets lost as resistive heat on electricity lines. While some of these losses are unavoidable, others can be significantly mitigated to reduce waste and emissions. ML can help prevent avoidable losses through predictive maintenance, i.e. by suggesting proactive electricity grid upgrades. Prior work has performed predictive maintenance using LSTMs\(^{118}\), bipartite ranking\(^{119}\), and neural network-plus-clustering techniques\(^{120}\) on electric grid data, and future work will need to improve and/or localize these solutions to different contexts.

Modeling emissions

Flexibly managing household, commercial, industrial, and electric vehicle demand (as well as energy storage) can help minimize electricity-based emissions (\(^{2,3,4,10}\)), but doing so involves understanding what the emissions on the electric grid actually are at any moment. Specifically, marginal emissions factors capture the emissions effects of small changes in demand at any given time. To inform consumers about marginal emissions factors, WattTime\(^{21}\) uses regression-based techniques to estimate these factors in real time for the US, and the electricityMap project\(^{22}\) employs ensemble models on electricity and weather data to forecast these factors a few days ahead for Europe. Great Britain’s National Grid ESO also uses ensemble models to forecast average emissions factors, which measure the aggregate emissions intensity of all power plants\(^{23}\). There is still much room to improve the performance of these methods, as well as to forecast related quantities such as electricity curtailments (i.e. the wasting of usually low-carbon electricity for grid balancing purposes). As most existing factor estimates are point estimates, it would also be important to quantify the uncertainty of these estimates to ensure that load-shifting techniques indeed decrease (rather than increase) emissions.

---

\(^{18}\)See also https://www.oilandgaseng.com/articles/how-machine-learning-contributes-to-smaller-pipeline-maintenance/


\(^{20}\)http://bluefield.co/

\(^{21}\)https://www.watttime.org/

\(^{22}\)https://www.electricitymap.org

\(^{23}\)https://carbonintensity.org.uk/
1.3 Empowering developing and low-data settings

Much of the discussion around electricity systems often focuses on settings such as the United States with universal electricity access and relatively abundant data. However, many places that do not share these attributes are still integral to the fight against climate change \[26\]. While ML applications for climate change mitigation are largely uncharted in such places, they warrant serious consideration from the ML community. Such applications in this area include improving electricity access and translating electricity system insights from data-abundant to low-data contexts.

Improving electricity access

Electricity is critical to economic and social development, and can also help address climate change. Specifically, promoting clean electricity via electric grids, microgrids, or off-grid methods can displace diesel generators, wood-burning stoves, and other carbon-emitting electricity sources, and can also increase educational outcomes \[121, 122\] (see §12). Figuring out what electrification methods are best for different areas can require intensive, boots-on-the-ground surveying work, but ML can help provide input to this process in a scalable manner; for instance, previous work has used image processing, clustering, and optimization techniques on satellite imagery as an input to planning electrification \[123\]. ML can also aid rural microgrid operation by accurately forecasting demand and power production (from e.g. solar panels), since small microgrids are even harder to balance than country-scale electric grids; for example, recent work used a hybrid LSTM and neural network architecture to model electricity load in rural microgrids \[124\]. Generating data to aid energy access policy and better managing energy access solutions are therefore two areas in which ML may have promising applications.

Low-data settings

While ML methods have often been applied to grids with widespread sensing, system operators in many countries do not collect or share system data. Although these data availability practices may evolve, it may meanwhile be beneficial to use techniques such as transfer learning to translate insights from data-abundant to low-data settings (especially since all electric grids share the same underlying system physics). Low-data ML techniques may also be beneficial in this setting; for instance, in \[125\], the authors enforce physical or other domain-specific constraints on weakly supervised ML models, allowing these models to learn from very little labeled data. ML techniques can also help generate information about low-data settings. For instance, recent work has used satellite image recognition (along with graph search techniques) to estimate the layout of electricity grids in regions where they may not be explicitly mapped \[126\], and companies have also proposed to use satellite imagery to measure power plant CO\(_2\) emissions \[24\] (also see §5.1). Other recent work has used regression-based techniques on cellular network data to model electricity consumption \[127\], which may prove useful in settings with many cellular towers but few electric grid sensors. Although low-data settings are generally under-explored by the ML community, electricity systems research in these settings presents a great opportunity for both ML and climate change.

1.4 Discussion

Data-driven and critical to climate change, electricity systems hold many opportunities for ML. At the same time, applications in this space hold many potential pitfalls; for instance, innovations that seek to reduce GHG emissions in the oil and gas industries could actually increase emissions by making them cheaper to emit \[21\]. Given these domain-specific nuances, working in this area requires close collaborations with electricity system decision-makers and with practitioners in fields including electrical engineering,
the natural sciences, and the social sciences. Explainable ML could facilitate interactions with domain experts by allowing them to interpret and appropriately apply ML model outputs to real-world settings. Similarly, it will be important to develop hybrid ML models that explicitly account for system physics (e.g. see [125, 128–130]), directly optimize for domain-specific goals [58–60], or otherwise incorporate or scale existing domain knowledge. Finally, since most modern electric grids are not data-abundant (although they may be data-driven), understanding how to apply data-driven insights to these grids may be the next grand challenge for ML in electricity systems.
2 Transportation

Transportation systems form a complex web that is fundamental to an active and prosperous society. Overall, the transportation sector accounts for about a quarter of global energy-related CO\(_2\) emissions [4]. In contrast to the electricity sector, transportation has not made significant progress to lower its CO\(_2\) emissions [131] and much of the transportation sector is regarded as hard to decarbonize [132]. This is because of the high energy density of fuels required for many types of vehicles, which constrains low-carbon alternatives, and because transport policies directly impact end-users and are thus more likely to be controversial.

Passenger and freight transportation are each responsible for about half of transport GHG emissions [133]. Both freight and passengers can travel by road, by rail, by water, or by air (referred to as transport modes). Different modes carry vastly different carbon emission intensities.\(^{25}\) At present, more than two-thirds of transportation emissions are from road travel [133], but air travel has the highest emission intensity and is responsible for an increasingly large share. Strategies to reduce GHG emissions\(^{26}\) from transportation include [133]:

- Decreasing transportation activity.
- Increasing vehicle efficiency.
- Reducing the carbon impact of fuel.
- Shifting to lower-carbon options, like rail.

Each of these mitigation strategies offers opportunities for ML. While many of us probably think of autonomous vehicles and shared mobility when we think of transport and ML, these technologies can help to reduce but also might increase GHG emissions [137]. Here, we discuss these disruptive technologies (§2.1) but show that ML can play a role for decarbonizing transportation that goes much further. ML can improve vehicle engineering, enable intelligent infrastructure, and provide policy-relevant information. Many interventions that reduce GHG emissions in the transportation sector require changes in planning, maintenance, and operations of transportation systems, even though the GHG reduction potential of those measures might not be immediately apparent. ML that is concerned with improving these tasks, for example by providing better demand forecasts, can make transportation more efficient. Typically, ML solutions are most effective in tandem with strong public policies. While we do not cover all ML applications in the transportation sector, we aim to include those areas which can conceivably reduce GHG emissions.

2.1 Reducing transport activity

A colossal amount of transport occurs each day across the world, but much of this mileage is used inefficiently, resulting in needless GHG emissions. With the help of ML, the number of vehicle-miles traveled can be reduced by making long trips less necessary, increasing loading, and optimizing vehicle routing. Here, we discuss the first two in depth – for a discussion of ML and routing, see for example [138].

Understanding transportation data

Many areas of transportation lack data, and decision-makers often plan infrastructure and policy based on uncertain information. In recent years, new types of sensors have become available, and ML can provide relevant information from these data. Traditionally, traffic is monitored with ground-based counters that are installed on a selected number of roads. A variety of technologies are used, such as inductive loop detectors

\(^{25}\)Carbon intensity is measured in grams of CO\(_2\)-equivalent per person-km or per ton-km, respectively.

\(^{26}\)For general resources on how to decarbonize the transportation sector, see the AR5 chapter on transportation [133], and [134–136].
or pneumatic tubes. In particular when counting pedestrians and cyclists, traffic is monitored with video systems, which can be automated with computer vision [139]. Since counts on most roads are often available only over short time frames, these roads are modeled by looking at known traffic patterns for similar roads. ML methods, such as SVMs and neural networks, have made it easier to classify roads with similar traffic patterns [140–142]. As ground-based counters require costly installation and maintenance, many countries do not have such systems. Vehicles can also be detected in high-resolution satellite images with high accuracy [143–146], and image counts can serve to estimate average vehicle traffic [147]. Similarly, ML methods can help with imputing missing data for precise bottom-up estimation of GHG emissions [148] and they are also applied in simulation models of vehicle emissions [149].

**Modeling demand**

*High Leverage*

By discouraging sprawl and creating new transportation links, modeling demand and planning new infrastructure can significantly shape how long trips are and which transport modes are chosen by passengers and shippers. ML can provide information about mobility patterns – which is directly necessary for agent-based travel demand models, one of the main transport planning tools [150]. For example, ML makes it possible to estimate origin-destination demand from traffic counts [151], and it offers new methods for spatio-temporal road traffic forecasting – which do not always outperform other statistical methods [152] but may transfer well between areas [153]. Also, short-term forecasting of public transit ridership can improve with ML; see for example [154, 155]. ML is particularly relevant for deducing information from novel data – for example, learning about the behavior of public transit users from smart card data [156, 157]. Also, mobile phone sensors provide new means to understand personal travel demand and the urban topology, such as walking route choices [158]. Similarly, ML-based modeling can help to mitigate climate change by improving operational efficiency of modes that emit significant CO₂. In the aviation sector, interventions that reduce aircraft taxi time and congestion on the runway bring fuel consumption down [159]. ML can help, for example, by predicting taxi time for efficient runway scheduling [160].

**Shared mobility**

*High Risk*

In the passenger sector, shared mobility (such as on-demand ride services or vehicle-sharing [27]), is undoubtedly disrupting the way people travel and think about vehicle ownership, and ML plays an integral part in optimizing these services (e.g. [161]). However, it is largely unclear what the impact of this development will be. For example, shared cars can actually cause more people to travel by car, as opposed to using public transportation. Similarly, on-demand taxi services add mileage when traveling without a customer, possibly negating any GHG emission savings [162]. On the other hand, shared mobility can lead to higher utilization of each vehicle, which means a more efficient use of materials [163]. The use of newer and more efficient vehicles, ideally electric ones, could increase with vehicle sharing concepts, reducing GHG emissions. Some of the issues raised above could also perhaps be overcome by making taxis autonomous. Such vehicles also might integrate better with public transportation, and offer new concepts for pooled rides, which substantially reduce the emissions per person-mile.

ML methods can help to understand the energy impact of shared mobility concepts, for example what results in more energy-efficient customer behavior, such as ride sharing [164]. It would also be very important for decision-makers to have access to timely location-specific empirical analysis to understand if a ride share service is taking away customers from low-carbon transit modes and increasing the use of cars.

Car-sharing services using autonomous vehicles could yield GHG emission savings when they encourage people to use public transit for part of the journey [165] or with autonomous electric vehicles [166]. However, using autonomous shared vehicles alone could increase the total vehicle-miles traveled and therefore do not necessarily lead to lower emissions when based on internal combustion engines (or electrical

---

27 In this section, we discuss shared cars; see [27] for bike shares and electric scooters.
vehicles and a “dirty” electrical grid) [167, 168]. We see the intersection of shared mobility, autonomous and electric vehicles, and smart public transit as a path where ML can make a contribution to shaping future mobility. See also §2.2 for more on autonomous vehicles.

When designing and promoting new mobility services, industry and public policy should prioritize lowering GHG emissions. Misaligned incentives in the early stages of technological development could result in the lock-in to a service with high GHG emissions [169, 170].

Freight routing and consolidation

Bundling shipments together, which is referred to as freight consolidation, dramatically reduces the number of trips (and therefore GHG emissions). The same is true for changing routing so that trucks do not have to return empty. As rail and water modes require much larger loads than trucks, consolidation also enables shipments to use these modes for part of the journey [147]. Freight consolidation and routing decisions are often taken by third-party logistics service providers and other freight forwarders, such as in the less-than-truckload market, which deals with shipments of smaller sizes. ML offers opportunities to optimize this complex interaction of shipment sizes, modes, origin-destination pairs, and service requirements. There are many proposed and deployed solutions using ML, for example predicting arrival times or demand, identifying and planning around transportation disruptions [171], or clustering suppliers by their geographical location and common shipping destinations. Designing allocation algorithms and freight auctions are proposed planning approaches, and ML has for example been shown to help pick good algorithms and parameters to solve auction markets [172].

Alternatives to transport

Disruptive technologies that are based on ML could replace or reduce transportation demand. For example, additive manufacturing (AM, or 3-D printing) has (limited) potential to reduce freight transport by producing lighter goods and enabling production closer to the consumer [136]. ML can be a valuable tool for improving AM processes [173]. ML can also help to improve virtual communication [174]. If passenger trips are replaced by telepresence, travel demand can be reduced, as has been shown for example in public agencies [175] and for scientific teams [176]. However, it is uncertain to what extent virtual meetings replace physical travel, or if they may actually give rise to more face-to-face meetings [177].

2.2 Vehicle efficiency

Most vehicles are not very efficient compared to what is technically possible: for example, through fleet turnover alone, aircraft carbon intensity can decline by more than a third with respect to 2012 [178]. Both the design of the vehicle and the way it is operated can increase the fuel economy. Here, we discuss how ML can help constructing more efficient vehicles and what kind of impact autonomous driving could have on GHG emissions. Encouraging drivers to adopt more efficient vehicles is also a priority; while we do not focus on this here, ML plays a role in studying consumer preferences in vehicle markets [179].

Designing for efficiency

There are many ways to reduce the energy a vehicle uses – such as more efficient engines, improved aerodynamics, hybrid electric engines, and reducing the vehicle’s weight or tire resistance. These different strategies require a broad range of engineering techniques, many of which can benefit from ML. For example, ML is applied in advanced combustion engine design [180]. Hybrid electric vehicles, which are more efficient than combustion engines alone, rely on power management methods that can be improved with ML [181]. Aerodynamic efficiency improvements need turbulence modeling that is often computationally intensive and relies heavily on ML-based surrogate models [182]. Aerodynamic improvements can not only
be made by vehicle design but also by rearranging load. Lai et al. [183] use computer vision to detect aerodynamically inefficient loading on freight trains. Additive manufacturing (3-D printing) can produce lighter parts in vehicles, such as road vehicles and aircraft, that reduce energy consumption [147, 163]. ML is applied to improve those processes, for example through failure detection [184, 185] or material design [186].

**Automobiles**

Machine learning is essential in the development of autonomous vehicles (AVs), including in such basic tasks as following the road and detecting obstacles [187]. While AVs could reduce energy consumption – for example, by reducing traffic congestion and inducing eco-driving – it is also possible that AVs will lead to an increase in overall road traffic that nullifies efficiency gains. (For an overview of possible energy impacts of AVs see [137, 188] and for broader impacts on mobility see [189].) Two advantages of AVs in the freight sector promise to cut GHG emissions: First, small autonomous vehicles, such as delivery robots and drones, could reduce the energy consumption of last-mile delivery [190], though they come with regulatory challenges [191]. Second, trucks can reduce energy consumption by **platooning** (driving very close together to reduce air resistance), thereby alleviating some of the challenges that come with electrifying long-distance road freight [192]. Platooning relies on autonomous driving and communication technologies that allow vehicles to brake and accelerate simultaneously.

ML can help to develop AV technologies specifically aimed at reducing energy consumption. For example, Wu et al. [193, 194] develop AV controllers based on reinforcement learning to smooth out traffic involving non-autonomous vehicles, reducing congestion-related energy consumption. ML methods can also help to understand driving practices that are more energy efficient. For example, Jiménez et al. [195] use data from smartphone sensors to identify driving behavior that led to higher energy consumption in electric vehicles.

2.3 Alternative fuels and electrification

**Electric vehicles**

Electrifying vehicles is regarded as a primary means to decarbonize transport. Electric vehicle (EV) technologies rely on batteries, hydrogen fuel cells, or electrified roads and railways, and can have very low GHG emissions – assuming, of course, that the electricity is generated with mostly low-carbon generators. ML is vital for a range of different problems related to EVs. Rigas et al. [196] detail methods by which ML can improve charge scheduling, congestion management, and vehicle-to-grid algorithms. ML methods have also been applied to battery energy management (for example charge estimation [197] or optimization in hybrid EVs [181]), and to detect faults and lateral misalignment in wireless charging of EVs [198].

As more people drive EVs, understanding their use patterns will become more important. Modeling charging behavior will be useful for grid operators looking to predict electric load. For this application, it is possible to analyze residential EV charging behavior from aggregate electricity load (**energy disaggregation**, see also [3.1, 199]). Also, in-vehicle sensors and communication data are increasingly becoming available and offer an opportunity to understand travel and charging behavior of EV owners, which can for example inform the placement of charging stations [200].

Battery electric vehicles are typically not used for more than a fraction of the day, allowing them to act as energy storage for the grid at other times, where charging and discharging is controlled for example by price signals [201] (see [1.1, 1.2]). There is much potential for ML to improve such vehicle-to-grid technology, for example with reinforcement learning [202], which can reduce GHG emissions from electricity generation. Vehicle-to-grid technology comes with private and social financial benefits. However, consumers are

---

28Providing details on the role of ML for AVs is beyond the scope of this paper.
expected to be reluctant to agree to such services, as they might not want to compromise their driving range.

Finally, ML can also play a role in the research and development of batteries, a decisive technology for EV costs and usability. Work in this area has focused on predicting battery state, degradation, and remaining lifetime using supervised learning techniques, fuzzy logic, and clustering [204–211]. However, many models developed in academia are based on laboratory data that do not account for real-world factors such as environmental conditions [204, 206]. By contrast, industry lags behind in ML modeling, but real-world operational data are readily available. Merging these two perspectives could yield significant benefits for the field.

**Alternative fuels**

Much of the transportation sector is highly dependent on liquid fossil fuels. Aviation, long-distance road transportation, and ocean shipping require fuels with high energy density and thus are not conducive to electrification [132]. Electrofuels [212], biofuels [213], hydrogen [214, 215], and perhaps natural gas [216] offer alternatives, but the use of these fuels is constrained by factors such as cost, land-use, and (for hydrogen and natural gas) incompatibility with current infrastructure [132]. Electrofuels and biofuels have the potential to serve as low-carbon drop-in fuels that retain the properties of fossil fuels, such as high energy density, while retaining compatibility with the existing fleet of vehicles and the current fuel infrastructure [147]. Fuels such as electrofuels and hydrogen can be produced using electricity-intensive processes and can be stored at lower cost than electricity. Thus, these fuels could provide services to the electricity grid by using electricity flexibly and balancing variable electricity generators (§1.1.1). Given their relative long-term importance and early stage of development, they present a critical opportunity to mitigate climate change. ML techniques may present opportunities for improvement at various stages of research and development of alternative fuels (similar to applications in §1.1.1).

### 2.4 Modal shift

Shifting passengers and freight to low carbon-intensity modes is one of the most important means to decarbonize transport. This *modal shift* in passenger transportation can for example involve providing people with public transit, which requires analyzing mode choice and travel demand data. ML can also make low carbon-intensive freight modes more competitive by helping to coordinate intermodal transport.

**Passenger preferences**

ML can improve our understanding about passengers’ travel mode choices, which in turn informs transportation planning, such as where public transit should be built. Some recent studies have shown that supervised ML based on survey data can improve passenger mode choice models [217–219]. Seo et al. propose to conduct long-term travel surveys with online learning, which reduces the demand on respondents, while obtaining high data quality [220]. Sun et al. [221] use SVMs and neural networks for analyzing preferences of customers traveling by high speed rail in China. There is also work on inferring people’s travel modes and destinations from social media or various mobile phone sensors such as GPS (*transportation mode detection*), e.g. [222, 223]. Also in the freight sector, ML has been applied to analyze modal trade-offs, for example by imputing data on counterfactual mode choices [224].

**Improving low-carbon options**

In order to incentive more users to choose low-carbon transport modes, their costs and service quality can be improved. Many low-carbon modes must be integrated with other modes of transportation to deliver the same level of service. For example, when traveling by train, the trip to and from the station will often be
by car, taxi, bus, or bike. There are many opportunities for ML to facilitate a better integration of modes, both in the passenger and freight sectors. ML can also help to improve the operation of low-carbon modes, for example by reducing the operations and maintenance costs of rail \cite{225} and predicting track degradation \cite{226}.

Bike sharing and electric scooter services can offer low-carbon alternatives for urban mobility that do not require ownership and integrate well with public transportation. ML studies help to understand how usage patterns for bike stations depend on their immediate urban surroundings \cite{227}. ML can also help solve the bike sharing rebalancing problem by improving forecasts of bike demand and inventory \cite{228}. Singla et al. \cite{229} propose a pricing mechanism based on online learning to provide monetary incentives for bike users to help rebalancing. By producing accurate travel time estimates, ML can provide tools that help to integrate bike shares with other modes of transportation \cite{230}. Many emerging bike and scooter sharing services are dockless, which means that they are parked anywhere in public space and can block sidewalks \cite{231}. ML has been applied to monitor public sentiment about such bike shares via tweets \cite{232}. ML could also provide tools and information for regulators to ensure that public space can be used by everyone \cite{233}.

Coordination between modes resulting in faster and more reliable transit times could increase the amount of people or goods traveling on low-carbon modes such as rail. ML algorithms could be applied to make public transportation faster and easier to use. For example, there is a rich literature exploring ML methods to predict bus arrival times and their uncertainty \cite{234, 235}. Often freight is packaged so that it can switch between different modes of transport easily. Such intermodal transportation relies on low-carbon modes such as rail and water for part of the journey \cite{136}. ML can provide solutions by improving predictions of the estimated time of arrival (for example of freight trains \cite{236}) or the weight or volume of expected freight (for example for roll-on/roll-off transport – often abbreviated as Ro-Ro \cite{237}). Intelligent transport systems of different modes could be combined and enable more efficient multimodal freight transportation \cite{136}.

Lax enforcement of regulation can make modes with high GHG emissions, such as trucks, competitive \cite{136}. ML can assist public institutions with enforcing their regulations. For example, image recognition can help law enforcement detect overloading of trucks \cite{238}.

2.5 Discussion

Decarbonizing transport will likely become a priority in the near future, and there are numerous applications where ML can make an impact. This is because transportation causes a large share of energy-related GHG emissions, but reducing them has been slow and complex. Solutions are likely very technical, are highly dependent on existing infrastructure, and require detailed understanding of passengers’ and freight companies’ behavior. ML can help by providing data, gaining knowledge from data, planning, and automation. Moreover, ML is fundamental to shared mobility, AVs, EVs, and smart public transit, which, with the right incentives, can be used to enable significant reductions in GHG emissions.
3 Buildings & Cities by Nikola Milojevic-Dupont and Lynn H. Kaack

Buildings offer some of the lowest-hanging fruit when it comes to reducing GHG emissions. While the energy consumed in buildings is responsible for a quarter of global energy-related emissions [4], a combination of easy-to-implement fixes and state-of-the-art solutions [29] could reduce emissions for existing buildings by 90% [240]. It is possible today for buildings to consume almost no energy [241]. Many of these energy efficiency measures actually result in overall cost savings [242] and simultaneously yield other benefits, such as cleaner air for occupants. This potential can be achieved while maintaining the services that buildings provide – and even while extending them to more people, as climate change will necessitate. For example, with the changing climate, more people will need access to air conditioning in regions where deadly heat waves will become common [243, 244].

Two major challenges are heterogeneity and inertia. Buildings vary according to age, construction, usage, and ownership, so optimal solutions vary widely depending on the context. For instance, buildings with access to cheap, low-carbon electricity may have less need for expensive features such as intelligent light bulbs. Buildings also have very long lifespans; thus, it is necessary both to create new, energy-efficient buildings, and to retrofit old buildings to be as efficient as possible [245]. Urban planning and public policy can play a major role in reducing emissions by providing infrastructure, financial incentives, or energy standards for buildings [31].

Machine learning provides critical tools both for managing buildings and for designing policies surrounding them. At the level of building management, ML can help select solutions that are tailored to individual buildings, and can also contribute to implementing those solutions via smart control systems (§3.1). At the level of urban planning, ML can be used to gather and make sense of data to inform policy makers (§3.2) [32].

3.1 Optimizing buildings

In designing new buildings and improving existing ones, there are numerous technologies that can reduce both costs and GHG emissions. ML can accelerate these solutions by (i) modeling data on energy consumption and (ii) optimizing energy use (in smart buildings).

Energy use models

An essential step towards efficiency is making sense of the increasing amounts of data produced by meters and home energy monitors [33]. This can take the form of energy demand forecasts for individual buildings, which are useful both for power companies (§1.1.1) and in evaluating building design and operation strategies [246]. Traditional energy demand forecasts model the physical structure of a building, and are essentially massive thermodynamics computations. ML has the potential to speed up these computations greatly, either by ignoring physical knowledge of the building entirely [247, 248], by incorporating it into the computation [249], or by learning to approximate the physical model to reduce the need for expensive

---

29 The IPCC classifies mitigation actions in buildings into four categories: carbon efficiency (switching to low-carbon fuels or to natural refrigerants); energy efficiency (reducing energy waste through insulation, efficient appliances, better heating and ventilation, or other similar measures); system and infrastructure efficiency (e.g. passive house standards, urban planning, and district cooling and heating); and service demand reduction (behavioral and lifestyle changes) [239].

30 There are even high-rise buildings, e.g. the Tower Raiffeisen-Holding NO-Vienna office, or large university buildings, e.g. the Technical University also in Vienna, that achieve such performance.

31 For example, see the case of New York City, which mandated that building owners collectively reduce their emissions by 40% by 2040, https://www.nytimes.com/2019/04/17/nyregion/nyc-energy-laws.html.

32 For example, the startup nam.R is developing a database of all school buildings in France, harmonizing vast amounts of open and proprietary data with ML. See project [rees at https://namr.com/#.

33 For example, see https://sense.com
simulation (surrogate models) [250]. Learning how to transfer the knowledge gained from modeling one building to another can make it easier to render precise estimations of more buildings. For instance, Mocanu et al. [251] modeled building load profiles with reinforcement learning and deep belief networks using data on commercial and residential buildings; they then used approximate reinforcement learning and transfer learning to make predictions about new buildings, enabling the transfer of knowledge from commercial to residential buildings, and from gas- to power-heated buildings.

Within a single building, understanding which appliances drive energy use (energy disaggregation) is crucial for targeting efficiency measures and enabling behavioral changes. Promising ML approaches to this problem include hidden Markov models [252], sparse coding algorithms for structured prediction [253], harmonic analysis that picks out the “signatures” of individual appliances [254], and deep neural networks [255].

To verify the success or failure of energy efficiency interventions, statistical ML offers methods for causal inference. For example, Burlig et al. [256] used Lasso regression on hourly electricity consumption data from schools in California to find that energy efficiency interventions fall short of the expected savings. Such problems could represent a useful application of deep learning methods for counterfactual prediction [257].

Smart buildings
Intelligent control systems in buildings can decrease the carbon footprint both by reducing the energy consumed and by providing means to integrate lower-carbon sources into the electricity mix [258]. Specifically, ML can reduce energy usage by allowing devices and systems to adapt to usage patterns. Further, buildings can respond to signals from the electricity grid, providing flexibility to the grid operator and lowering costs to the consumer (§1.1.1).

Many critical systems inside buildings can be made radically more efficient. While this is also true for small appliances such as refrigerators and lightbulbs, we use the example of heating and cooling (HVAC) systems, both because they are notoriously inefficient and because they account for more than half of the energy consumed in buildings [239]. There are several promising ways to enhance HVAC operating performance, each providing significant opportunities for using ML: forecasting what temperatures are needed throughout the system, better control to achieve those temperatures, and fault detection. Forecasting temperatures, as with modeling energy use in buildings, has traditionally been performed using detailed physical models of the system involved; however, ML approaches such as deep belief networks can potentially increase accuracy with less computational expense [259, 260] (see also §4.3). For control, Kazmi et al. [261] used deep reinforcement learning to achieve a scalable 20% reduction of energy while requiring only three sensors: air temperature, water temperature, and energy use (see also §4.3 for similarly substantial gains in datacenter cooling). Finally, ML can automate building diagnostics and maintenance through fault-detection. For example, the energy efficiency of cooling systems can degrade if refrigerant levels are low [262]; ML approaches are well-suited to detect faults in these systems. Wang et al. [263] treated HVAC fault-detection as a one-class classification problem, using only temperature readings for their predictions. Deep autoencoders can be used to simplify information about machine operation so that deep neural networks can then more easily predict multiple kinds of faults [264].

Many systems within buildings – such as lights and heating – can also adjust how they operate based on whether a building or room is occupied, thereby improving both occupant comfort and energy use [265]. ML can help these systems dynamically adapt to changes in occupancy patterns [266]. Moreover, occupancy detection itself represents an opportunity for ML algorithms, ranging from decision trees [267, 268] to deep neural networks [269] that take input from occupancy sensors [267], WiFi signals [269, 270], or appliance power consumption data [268].

In §1.1.1 we discussed how using variable low-carbon energy can mean that the supply and price of
electricity varies over time. Thus, energy flexibility in buildings is increasingly useful to schedule consumption when supply is high [271]. For this, automated demand-side response [272] can respond to electricity prices, smart meter signals, or learned user preferences [273]. Edge computing can be used to process data from distributed sensors and other Internet of Things devices, and deep reinforcement learning can then use this data to efficiently schedule energy use [274].

While smart building technologies have the capability to significantly increase efficiency, we should note that there are potential drawbacks [275]. First, smart building devices and connection networks, like wireless sensor networks, consume energy themselves; however, deep neural networks can be used to monitor and optimize their operations [276]. Second, rebound effects are likely to happen in certain cases [277], leading to additional consumption typically ranging between 10 and 20% [278] for buildings in general. If control systems optimize for costs, interventions do not necessarily translate into energy efficiency measures or GHG reductions. Therefore, public policies are needed to mandate, support and complement actions for individual building managers [239]. Another concern in case of widespread adoption of smart meters is the impact on mineral use and embodied energy use arising from their production [279]. Finally, smart home applications present security and privacy risks [280] that require adequate regulation.

3.2 Urban planning

For many impactful mitigation strategies – such as district heating and cooling, neighborhood planning, and large-scale retrofitting of existing buildings – coordination at the district and city level is essential. Policy-makers use instruments such as building codes, retrofitting subsidies, investments in public utilities, and public-private partnerships in order to reduce GHG emissions without compromising equity. While infrastructure models have yet to be adopted at scale, they can be highly impactful in informing policy makers about heterogeneities among buildings, the energy impact of policies, and aggregated GHG emission estimates and forecasts. This can for example be used for planning and operating district heating and cooling, where a central plant supplies many households in a district. District heating and cooling reduces HVAC energy consumption and can provide flexible load [281], but it needs large amounts of data at the district level for implementation and operation.

However, district-level data is often not available. ML can help in obtaining it in two ways: Where energy-use data on individual buildings exists, ML can be used to derive higher-level patterns. Where data on energy use and infrastructure is completely lacking, ML can infer it.

District-level energy use

Bottom-up multi-building energy models are expected to become fundamental for enabling localized action by city planners [282]. ML can learn from available energy use data to extrapolate building energy use predictions to the city level. Based on energy data disclosed by residents of New York City, Kontokosta and colleagues used various ML methods to predict the energy use of the city’s 1.1 million buildings [283], analyzed the effect of energy disclosure on the demand [284], and developed a system for ranking buildings based on energy efficiency [285]. Zhang et al. [286] matched residential energy consumption survey data with public use microdata samples to estimate residential energy consumption at the neighborhood level. Robinson et al. [287] showed that using a simple gradient boosting technique can predict commercial building energy use across large American cities, using simple features of individual buildings.

Gathering infrastructure data

Many regions of the world have almost no energy consumption data, which can make it difficult to design targeted mitigation strategies. ML is uniquely capable of predicting energy consumption and GHG mitigation potential at scale from other types of available data. Information about building footprint, usage, material, roof type, immediate surroundings etc. can be predictive of energy consumption. For example,
Kolter and Ferreira used Gaussian process regression to predict energy use from features such as property class or the presence of central AC [288]. ML can be used to pinpoint which buildings have the highest retrofit potential using simple building characteristics and surrounding environmental factors [289, 290] – both potentially available at scale.

Specifics about building infrastructure can also often be predicted using ML techniques, providing data for energy planning. Remote sensing is key to this process [91, 291–295] as satellite data offers a source of information that is globally available and largely consistent worldwide. For example, using remote sensing data, Geiß et al. [297] clustered buildings into types to assess the potential of district heat in a German town.

The resolution of infrastructure models ranges from coarse localization of all buildings at the global scale [291], to precise 3D reconstruction of a neighborhood [295]. It is possible to produce a global map of human settlement footprints at meter-level resolution from satellite radar images [291]. For this, Esch et al. used highly automated learners, which make classification at such scale possible by retraining locally. Segmentation of high-resolution satellite images can now generate exact building footprints at a national scale [292]. Energy-relevant building attributes, such as the presence of photovoltaic panels, can also be retrieved from these images [91] (see §1.1.1). To generate 3D models, LiDAR data is often used to retrieve heights or classify buildings at city scale [293, 294], but its collection is expensive. Recent research showed that heights can be predicted even without such elevation data, as demonstrated by [298], who predicted these from real estate records and census data. Studies, which for now are small scale, aim for complete 3D reconstruction with class labels for different components of buildings [295].

3.3 The future of cities

Since most of the resources of the world are ultimately channeled into cities, municipal governments have a unique opportunity to mitigate climate change. City governments regulate (and sometimes operate) transport, buildings, and economic activity. They take care of such diverse issues as water, waste, energy, crime, health, and noise. Recently, data and ML have become more common for improving efficiency in such areas, giving rise to the notion of smart city. While the phrase smart city encompasses a wide array of technologies [299], here we discuss only applications that are relevant reducing GHG emissions.

Data for smart cities

Increasingly, important aspects of city life come with digital information that can make the city function in a more coordinated way. Habibzadeh et al. [300] differentiate between hard-sensing, i.e., fixed-location-dedicated sensors like traffic cameras, and soft-sensing, for example from mobile devices. Hard sensing is the primary data collection paradigm in many smart city applications, as it is adapted to precisely meet the application requirements. However, there is a growing volume of data coming from soft sensing, due to the widespread adoption of personal devices like smartphones that can provide movement data and geotagged pictures. Jiang provides a review of urban computing for mobile phone traces [302]. Relevant information on the urban space can also be learned from social media activity, like Twitter, as reviewed in [303, 304]. There are, however, numerous obstacles to making sense of this wealth of data (see [305]).

First, cities need to obtain relevant data on activities that directly or indirectly consume energy. Data are often proprietary. To obtain these data, the city of Los Angeles now requires all mobility as a service providers, i.e. vehicle-sharing companies, to use an open-source API. Data on location, use, and condition of all those vehicles, which can be useful in guiding regulation, are thus transmitted to the city [306]. ML can also distill information on urban issues related to climate change through web-scraping and text-mining, e.g. [232]. As discussed above (§3.2), ML can also be used to infer infrastructure data.

34See [296] for a review of different sources of data and deep learning methods for processing them.
35Note that management of any such private data, even if they are anonymized, poses challenges [301].
Second, smart city applications must transmit high volumes of data in real-time. ML is key to preprocessing large amounts of data in large sensor networks, allowing only what is relevant to be transmitted, instead of all the raw data that is being collected [307, 309]. Similar techniques also help to reduce the amount of energy consumed during transmission itself [310].

Third, urban policy-making based on intelligent infrastructure faces major challenges with data management [311]. Smart cities require the integration of multiple large and heterogeneous sources of data, for which ML can be a valuable tool, which includes data matching [312, 313], data fusion [314], and ensemble learning [315].

**Low-emission infrastructure**

When smart city projects are properly integrated into urban planning, they can make cities more sustainable and foster low-carbon lifestyles (see [310, 316, 317] for extensive reviews on this topic). Because of the feedback between different sectors, many mitigation options need to be planned by one entity – the local government. For instance, urban sprawl influences the energy use from transport, as wider cities tend to be more car-oriented [318–320]. ML-based analysis has shown that the development of efficient public transportation is dependent on both the extent of urban sprawl and the local development around transportation hubs [321, 322]. This shows how much buildings influence transportation systems, and vice versa.

Cities also can reduce GHG emissions by coordinating between infrastructure sectors and better adapting services to the needs of the inhabitants. Smart applications based on ML and AI can coordinate, for example, district heating and cooling networks, solar power generation, and charging stations for electric vehicles and bikes [317]. ML predictions are often useful, e.g. for improving public lighting systems by regulating light intensity based on historical patterns of foot traffic [323]. Due to inherent variability in energy demand and supply, there is a need for uncertainty estimation, e.g. using Markov chain Monte Carlo methods or Gaussian processes [317].

Currently, most smart city projects for urban climate change mitigation are implemented in wealthier regions such as the United States, China, and the EU [36]. The literature on city-scale solutions is also strongly biased towards the global North [324], while key mitigation challenges are expected to arise from the global South [325]. Infrastructure models described in [322] could be used to plan low-carbon neighborhoods without relying on advanced smart city technologies. To transfer solutions across cities, it is possible to cluster similar cities based on climate-relevant dimensions [319, 326, 327]. Creutzig et al. [319], for example, related the energy use of 300 cities worldwide to historical structural factors such as fuel taxes (which have a strong impact on urban sprawl). Other relevant applications include groupings of transportation systems [326] using a latent class choice model, or of street networks [327] to identify common patterns in urban development using hierarchical clustering.

### 3.4 Discussion

We have shown many different ways that ML can help to reduce GHG emissions from buildings and cities. A central challenge in this sector is the availability of high-quality data for training the algorithms, which rarely go beyond main cities or represent the full spectrum of building types. Techniques for obtaining these data, however, can themselves be an important application for ML (e.g. via computer vision algorithms to parse satellite imagery). In cities, integrating smart systems beyond gadget-level solutions and activating the potential of data-driven urban planning can play key roles for mitigation, while improving the well-being of citizens [240, 245, 328].
4 Industry

by Anna Waldman-Brown

Industrial production, logistics, and building materials are leading causes of difficult-to-eliminate GHG emissions [132]. Luckily for ML researchers, the global industrial sector spends billions of dollars annually gathering data on factories and supply chains [329] – aided by improvements in the cost and accessibility of sensors and other data-gathering mechanisms (such as QR codes and image recognition). The availability of large quantities of data, combined with affordable cloud-based storage and computing, indicates that industry may be an excellent place for ML to make a positive climate impact. ML researchers can potentially reduce global emissions by helping to streamline supply chains, improve production quality, predict machine breakdowns, optimize heating and cooling systems, and prioritize the use of clean electricity over fossil fuels [330–333].

Nonetheless, two key challenges stand in the way of harnessing ML for GHG mitigation within industry. Most importantly, the Jevons paradox indicates that efficiency increases may lead to rebound effects, resulting in increased goods production with even greater GHG emissions unless industrial actors have sufficient incentives to reduce overall emissions [334]. Secondly, despite the proliferation of industrial data, much of the information is proprietary, low-quality, or very specific to individual machines or processes; practitioners estimate that 60-70% of industrial data goes unused [329, 335]. Before investing in extensive ML research, researchers should be sure that they will be able to eventually access and clean any data needed for their algorithms.

4.1 Supply chains

In 2006, at least two Scottish seafood firms flew hundreds of metric tons of shrimp from Scotland to China and Thailand for peeling, then back to Scotland for sale – because they could save on labor costs [336]. This indicates the complexity of today’s globalized supply chains, defined as the processes and systems of organizations and the shipping networks that are required to bring a product from producer to final consumer. While ML can help minimize emissions by optimizing shipping routes [2.1, 2.3], reducing waste, and helping firms identify local producers and suppliers, firms’ financial incentives must also align with climate change mitigation through carbon pricing or other policy mechanisms. ML could reduce emissions in supply chains by intelligently predicting supply and demand, identifying lower-carbon products, and optimizing shipping routes. For details on shipping and delivery optimization, see [2].

The production, shipment, and climate-controlled warehousing of excess products is a major source of industrial GHG emissions, particularly for time-dependent goods such as perishable food or retail goods that quickly fall out of fashion [337]. In 2011, a survey of corporate sales estimates diverged from actual sales by an average of 40%, and the Council of Supply Chain Management Professionals estimated global excess inventory to be $8 trillion worth of goods [338]. ML may be able to mitigate these issues of overproducing and/or overstocking goods by improving demand forecasting [339, 340]. For example, the clothing industry sells an average of only 60% of its wares at full price, but some brands can sell up to 85% due to just-in-time manufacturing and clever intelligence networks [341]. As online shopping and just-in-time manufacturing become more prevalent and websites offer more product types than physical storefronts, better demand forecasts will be needed on a regional level to efficiently distribute inventory without letting unwanted goods travel long distances only to languish in warehouses [342]. Nonetheless, negative side effects can be significant depending on the type of product and regional characteristics; just-in-time manufacturing and online shopping are often responsible for smaller and faster shipments of goods, mostly on road, that lack the energy efficiency of freight aggregation and slower shipping methods such as rail [342, 343].

For ML to help reduce supply chain emissions, industry will need to make substantial improvements in data as well as more domain-specific models, such as task-based deep learning frameworks that could allow for the prediction of demand to specifically reduce the cost (or GHG emissions) associated with inventory.
stocking. Firms may then be able both to manufacture fewer products overall and to minimize the GHG needed for climate control during storage. Researchers are training deep learning systems to minimize counterfactual prediction error, which can theoretically lead to more accurate demand algorithms. Yet again, industrial incentives must be aligned with reducing emissions.

Recommender systems can potentially direct consumers and purchasing firms toward climate-friendly options, as long as one can obtain information about GHG emissions throughout the entire life-cycle of some product. The challenge here lies in hunting down usable data on every relevant material and production process from metal ore extraction through production, shipping, and eventual use and disposal of a product. One must also convince companies to share proprietary data to help other firms learn from best practices. If these datasets can be acquired, ML algorithms could hypothetically assist in identifying the cleanest options.

Optimized supply chains, combined with improved heating and cooling systems (see §4.3), can also play a role in reducing food waste, which ranks as #3 on Project Drawdown’s list of climate change solutions. Globally, society loses or wastes 1.3 billion metric tons of food each year, which translates to one-third of all food produced for human consumption. Developing countries lose 40% of this food between harvest and processing or retail, while over 40% of food waste in industrialized nations occurs at the end of supply chains, in retail outlets, restaurants, and consumers’ homes. ML can help reduce food waste in all these contexts. The highest impacts will come from mitigating post-harvest losses through optimized delivery routes and more affordable climate control systems, as well as tackling retail and consumer losses through better demand forecasting at the point of sale. ML can also potentially assist with other issues related to food waste, such as helping develop sensors to identify when produce is about to spoil, so it can be sold quickly, or removed from a storage crate before it ruins the rest of the shipment.

4.2 Materials and construction

Cement and steel production together account for approximately 9% of all global GHG emissions; the cement industry alone emits more GHG than every country except the US and China. Several studies indicate that ML may be able to help minimize these emissions by reducing the need for carbon-intensive materials, by transforming industrial processes to run on low-carbon energy, or even by replacing cement and steel altogether with more efficient structural materials.

To reduce the use of cement and steel, researchers have combined ML with generative design to develop structural products that require less raw material, thus minimizing the resulting GHG emissions. Novel manufacturing techniques such as 3D printing allow for the production of unusual shapes that use less material but may be impossible to produce through traditional metal-casting or poured concrete; Baturynska et al. used ML and finite element modeling to better simulate the physical processes of 3D printing in order to improve the quality of finished products.

Assuming future advances in materials science, ML researchers could potentially draw upon databases such as the Materials Project to help invent new, climate-friendly materials with desirable chemical properties. Researchers are also experimenting with supervised learning and thermal imaging systems to rapidly identify promising catalysts and chemical reactions, as described in §1.1.1. Firms are unlikely to adopt new materials or change existing practices without financial incentives, so widespread adoption might require subsidies for low-carbon alternatives or penalties for high GHG emissions.

4.3 Production and energy

ML can potentially assist in reducing overall electricity consumption; streamlining factories’ heating, ventilation, and air conditioning (HVAC) systems; and redesigning some types of industrial processes to run on low-carbon energy instead of coal, oil, or gas. Again, the higher the incentives for reducing carbon emis-
sions, the more likely that firms will optimize for low-carbon energy use. New factory equipment can be very expensive to purchase and set up, so many firms’ cost-benefit calculations may dissuade them from retrofitting existing factories to run using low-carbon electricity or to save a few kilowatts.

Ammonia production for fertilizer use relies upon natural gas to heat up and catalyze the reaction, and accounts for around 2% of global energy consumption. To develop cleaner ammonia, chemists may be able to invent electrochemical strategies for lower-temperature ammonia production. Given the potential of ML for predicting chemical reactions, ML may also be able to help with the discovery of new materials for electrocatalysts and/or proton conductors to facilitate ammonia production.

On the production side, ML can potentially improve the efficiency of HVAC systems and other industrial control mechanisms—given necessary data about all relevant processes. Deep neural networks could be used for adaptive control in a variety of industrial networking applications, enabling energy savings through self-learning about devices’ surroundings. To reduce GHG emissions from HVAC systems, researchers suggest combining optimization-based control algorithms with ML techniques such as image recognition, regression trees, and time delay neural networks (see also §3.1). DeepMind has used reinforcement learning to optimize cooling centers for Google’s internal servers by predicting and optimizing the power usage effectiveness (PUE), thus lowering emissions and reducing cooling costs.

Air conditioning systems have an outsized impact on climate, as the hydrofluorocarbons (HFCs) they contain are extremely potent greenhouse gases; proper management of such refrigerants ranks #1 on Project Drawdown’s solutions list. Computer vision and deep learning can potentially assist factories and regulators in better detecting HFC leaks and improving disposal management, especially in combination with satellite imagery and data related to leak-prone equipment.

ML could also contribute to predictive maintenance by more accurately modelling the wear and tear of machinery that is currently in use, and interpretable ML could assist factory owners in developing a better understanding of how best to minimize GHG emissions for specific equipment and processes. For example, creating a digital twin model of some industrial equipment or process could enable a manufacturer to identify and prevent undesirable scenarios, as well as virtually test out a new piece of code before uploading it to the actual factory floor—thus potentially increasing the GHG efficiency of industrial processes. Digital twins can also reduce production waste by identifying broken or about-to-break machines before the actual factory equipment starts producing damaged products. Industrial systems can employ similar models to predict which pipes are liable to spring leaks, in order to minimize the direct release of GHGs such as HFCs and natural gas.

ML may be particularly useful for enabling more flexible operation of industrial electrical loads, through optimizing a firm’s demand response to electricity prices as addressed in §1. Such optimization can contribute to cutting GHG emissions as long as firms have a financial incentive to optimize for minimal emissions, maximal low-carbon energy, or minimum overall power usage. Demand response optimization algorithms can help firms adjust the timing of energy-intensive processes such as cement crushing and powder-coating to take advantage of electricity price fluctuations, although published work on the topic has to date used relatively little ML. Online algorithms for optimizing demand response can reduce overall power usage for computer servers by dynamically shifting the internet traffic load of data providers to underutilized servers, although most of this research, again, has focused on minimizing costs rather than GHG emissions. Berral et al. proposed a framework that demonstrates how such optimization algorithms might be combined with RL, digitized controls, and feedback systems to enable the autonomous control of industrial processes.

4.4 Discussion

There are a number of potentially useful applications for using ML to reduce GHG emissions in industry, although there has been little published research focused specifically on GHG reductions rather than cost...
savings. On the logistics side, ML may help optimize shipping routes and improving market demand predictions in order to produce fewer unwanted goods. On the design and production side, the strategic use of ML can potentially reduce the need for GHG-intensive materials as well as optimizing industrial processes to use less energy and/or run on low-carbon fuels. Before applying ML to industry applications, however, researchers must consider questions such as the following: Who has access to the requisite data for training? Under what circumstances does ML actually outperform regressions, naive forecasting algorithms, and other optimization solutions that require less data and may be simpler to design and implement? Even if ML can improve some industrial system, can firms feasibly implement the solution? Will this improvement actually result in GHG mitigation, or are firms liable to respond with increased production and even greater GHG emissions as indicated by the Jevons paradox?

In conclusion, ML demonstrates considerable potential for reducing industrial GHG emissions under the following circumstances:

- When there are large amounts of accessible, high-quality data around specific processes or transport routes.
- When firms have an incentive to share their proprietary data and/or algorithms with researchers and other firms.
- When aspects of production or shipping can be readily fine-tuned or adjusted, or when the benefits of overhauling old systems exceed the costs.
- When firms’ incentives align with reducing emissions (for example, through efficiency gains, regulatory compliance, or high GHG prices).

Given the globalized nature of international trade and the urgency of climate change, decarbonizing the industrial sector must become a key priority for both policy-makers and factory owners worldwide. If firms can make more money by reducing their GHG emissions, market competition will drive companies towards cleaner and more efficient production and distribution – and, given the right datasets, ML researchers can help pave the way.
Plants and algae have been accumulating and sequestering carbon through photosynthesis for millions of years. While some carbon is stored deep underground as coal and oil, a large amount of carbon continues to be held in trees, peat bogs, and soil. Our current economy encourages practices that are freeing large amounts of this sequestered carbon through deforestation and unsustainable agriculture. In addition, cattle farming in particular generates methane, a more potent greenhouse gas than CO₂ itself. Overall, land use is estimated to be responsible for about a quarter of global GHG emissions (26) (and this may be an underestimate [370]). In addition to this direct release of carbon through human actions, the permafrost is now melting and forest fires are becoming more frequent as a consequence of climate change itself – which releases yet more carbon [371]. Permafrost alone is expected to release large amounts of GHG in the next few decades, causing an additional 12-17% increase in global emissions [372].

The large scale of this problem allows for a similar scale of positive impact, and ML will play an important role in many of these solutions. Precision agriculture can reduce carbon release from the soil and improve crop yield, which in turn reduces the need for deforestation. With satellite images, we can estimate the amount of carbon sequestered in every parcel of land, as well as how much GHG it releases. We can monitor the health of forests and peatlands and predict the risk of fire. We can also use ML in sustainable forestry and use drones to plant seeds. Overall, according to Project Drawdown [18], around a third of the potential impact of promising climate change solutions involves better land management and agriculture.

5.1 Remote sensing of emissions

Having real-time maps of GHG emissions could help us quantify emissions from agriculture practices and detect areas where vegetative decay emits more GHG than is absorbed, as well as monitor emissions from other sectors ([1.2]). Such information would be invaluable in guiding regulations or incentives that could steer land usage towards better practices. More concretely, we could put a price on emissions and pinpoint those responsible for it.

While greenhouse gases are invisible to our eyes, they must by definition interact with sunlight. This means that we can observe these compounds with hyperspectral cameras [373, 374]. These cameras can contain up to several hundred contiguous channels (instead of simply RGB), providing information on the interaction between light and individual chemicals. Many satellites are equipped with such cameras and can perform, to some extent, estimations of CO₂, CH₄ (methane), H₂O, and N₂O (nitrous oxide) emissions [375, 376]. While extremely useful for studying climate change, most of these satellites have very coarse spatial resolution and large temporal and spatial gaps, making them unsuitable for precise tracking of emissions. Standard satellite imagery provides RGB images with much higher resolution, which could be used in an ML algorithm to fill the gaps in hyperspectral data and obtain more precise information about emissions. Some preliminary work [375] has studied this possibility, but there are no clear results as of yet. This is therefore an open problem with high potential impact [37].

5.2 Precision agriculture

Overall, agriculture is responsible for 14% of GHG emissions [26]. This might come as a surprise, since one might expect that growing plants would take up CO₂ from the air. However, modern industrial agriculture involves more than just growing plants. First, the land is stripped of trees, releasing carbon sequestered...
there. Second, the process of tilling exposes topsoil to the air, thereby releasing carbon that had been bound in soil aggregates. Finally, because such farming practices strip soil of nutrients, nitrogen-based fertilizers must be added back to the system. While some of this nitrogen is absorbed by plants or retained in the soil, some is converted to nitrous oxide, a greenhouse gas that is 300 times more potent than CO₂.

Our current approach to agriculture is based on making farmland more uniform and predictable. This allows it to be managed at scale using basic automation tools like tractors, but can be both more destructive and less productive than approaches that work with the natural heterogeneity of land and crops. Increasingly, there is demand for sophisticated tools which would allow farmers to work at scale, but adapt to what the land needs. This approach is often known as “precision agriculture.”

On the hardware side, smarter robotic tools are being developed. RIPP A [378], a robot under development at the University of Sydney, is equipped with a hyperspectral camera and has the capacity to do mechanical weeding, targeted pesticide application, and vacuuming of pests. It can cover 5 acres per day on solar energy and collect large datasets [379] for continual improvement. Many other robotic platforms [383] likewise offer opportunities for developing new ML algorithms. There remains significant room for development in this space, since current robots still sometimes get stuck, are optimized only for certain types of crops, and rely on ML algorithms that may be highly sensitive to changes of environment.

The potential for precision agriculture goes beyond robots in the field. Simple macroeconomic models can help farmers predict crop demand and decide what to plant at the beginning of the season [380]. More intelligent irrigation systems can save large amounts of water while reducing pests that thrive under excessive moisture [18]. Overall, ML can improve crop yield prediction [381], disease detection, weed detection, and soil sensing [382–384]. These problems often have minimal hardware requirements, as devices such as Unmanned Aerial Vehicles (UAVs) with hyperspectral cameras can be used for all of these tasks.

Globally, agriculture constitutes a $2.4 trillion industry [385], and there is already a significant economic incentive to increase efficiency. However, efficiency gains do not necessarily translate into reduced GHG emissions (for example, via the Jevons paradox). Moreover, significantly reducing emissions may require a shift in agricultural paradigms – for example, widespread adoption of regenerative agriculture, silvopasture, and tree intercropping [18]. ML tools for policy makers and agronomists [386] could potentially encourage climate-positive action: for example, remote sensing with UAVs and satellites could perform methane detection and carbon stock estimation, which could be used to incentivize farmers to sequester more carbon and reduce emissions.

5.3 Protecting peatlands

Peatlands (a type of wetland ecosystem) cover only 3% of the Earth’s land area, yet hold twice the total carbon in all the world’s forests, making peat the largest source of sequestered carbon on Earth [387]. When peat dries, however, it releases carbon through decomposition and also becomes susceptible to fire [387,388]. A single peat fire in Indonesia in 1997 is reported to have released emissions comparable to 20-50% of global fossil fuel emissions [389] during the same year.

Monitoring peatlands and protecting them from artificial drainage or droughts is essential to preserve the carbon sequestered in them [390,391]. In [392], ML was applied to features extracted from remote sensing data to estimate the thickness of peat and assess the carbon stock of tropical peatlands. A more precise peatlands map is expected by 2020 using specialized satellites [393]. Advanced ML could potentially help develop precise monitoring tools at low cost and predict the risk of fire.

---

38 Some fertilizer additionally often ends up in waterways, which can contaminate drinking water and induce blooms of toxic algae [377].

39 Examples include sagarobotics.com, ecorobotix.com, and farm.bot.
5.4 Forests

Carbon stock estimation
Modeling (and pricing) carbon stored in forests requires us to assess how much is being sequestered or released across the planet. Since most of a forest’s carbon is stored in above-ground biomass [394], tree heights and tree types are a good indicator of the carbon stock.

The height of trees can be estimated fairly accurately with LiDAR devices mounted on UAVs. However, this technology is not scalable and many areas are closed to UAVs. To address this challenge, ML can be used to predict the LiDAR’s outcome from satellite imagery [394, 395]. From there, the learned estimator can perform predictions at the scale of the planet. Despite progress in this area, there is still significant room for improvement. For example, LiDAR collects height of trees only in certain regions, seasons, or climate. Hence domain adaptation and transfer learning techniques may help algorithms to generalize better.

Automated afforestation
Crowther and colleagues report that, theoretically, there is the capacity for 1.2 trillion additional trees that can be planted in existing forests and abandoned lands [396]. This would have the potential to cancel out a decade of global CO$_2$ emissions. Clearly, planting billions of trees requires a large amount of physical labor, for which automation might be useful. Startups like BioCarbon Engineering [40] and Droneseed [41] are developing drones capable of planting seed packets. Such drones can be 150 times faster than traditional methods and 10 times cheaper [397]. Working with local species to promote natural ecosystems, they use ML to locate appropriate planting sites, monitor plant health, assess weeds, and analyze trends.

Forest fire management
Besides their potential for harming people and property, forest fires release CO$_2$ into the atmosphere (which in turn increases the rate of forest fires [398]). On the other hand, small forest fires are part of natural forest cycles. Preventing them causes biomass to accumulate on the ground and increases the chances of large fires, which can then burn all trees to the ground and erode top soil, resulting in high CO$_2$ emissions, biodiversity loss, and a long recovery time [399]. Drought forecasting [400] is helpful in predicting regions that are more at risk, as does estimating the water content in the tree canopy [401]. In [402, 403], reinforcement learning is used to predict the spatial progression of fire. This helps firefighters decide when to let a fire burn and when to stop it [404]. With good tools to evaluate regions that are more at risk, firefighters can perform controlled burns and cut select areas to prevent progression of fires.

Forestry
Only 17% of the world’s forests are protected [405]. The rest are subject to deforestation, which contributes to approximately 10% of global GHG emissions [26] as vegetation is burned or decays. While some deforestation is the result of expanding agriculture or urban developments, most of it comes from the logging industry. Clearcutting, which has a particularly ruinous effect upon ecosystems and the carbon they sequester, remains a widespread practice across the world.

Providing tools for tracking deforestation can help provide valuable data for informing policy-makers, as well as being a tool for law-enforcement in case where deforestation may be conducted illegally. ML can be used to differentiate selective cutting from clearcutting using remote sensing imagery [406–409]. Taking another approach, Rainforest Connection [410] has installed old smart-phones powered by solar panels in the forest. Then, an ML algorithm can detect chainsaw sounds within a radius of a kilometer and report them to a nearby cellphone antenna.

40 www.biocarbonengineering.com
41 www.droneseed.co
Logistics and transport still dominate the cost of wood harvesting, which often motivates clearcutting. Increasingly, ML tools [411] are becoming available to help foresters decide when to harvest, where to fertilize, and what roads to build. However, once more, the Jevons paradox is at play. Making forestry more efficient can have a negative effect. On the other hand, developing the right combination of tools for regulation and selective cutting could have a significant positive impact.

5.5 Discussion

Farms and forests make up a large portion of global GHG emissions, but reducing these emissions is challenging for two reasons: the scope of the problem is highly globalized, but the necessary actions are highly localized. A major challenge in many land use applications is the diversity of stakeholders involved. Agriculture, for example, involves a complex mix of large-scale farming interests, small-scale farmers, agricultural equipment manufacturers, and chemical companies. Each stakeholder has different interests, and each often has access to a different portion of the data that would be useful for impactful ML applications. Interfacing between these different stakeholders is a practical challenge for meaningful work in this area.
6 Carbon Dioxide Removal

by Andrew S. Ross and Evan D. Sherwin

Even if we could cut emissions to zero today, we would still face significant climate consequences from greenhouse gases already in the atmosphere. Eliminating emissions entirely may also be tricky, given the sheer diversity of sources (such as airplanes and cows). Instead, many experts argue that to meet critical climate goals, global emissions must become net-negative—that is, we must remove more CO$_2$ from the atmosphere than we release [412, 413]. Although there has been significant progress in negative emissions research [414–418], the actual CO$_2$ removal industry is still in its infancy. As such, many of the ML applications we outline in this section are highly speculative.

6.1 Natural or semi-natural methods

Many of the primary candidate technologies for CO$_2$ removal directly harness the same natural processes which have (pre-)historically shaped our atmosphere. One of the most promising methods is simply allowing or encouraging more natural uptake of CO$_2$ by plants (whose ML applications we discuss in §5). Other plant-based methods include bioenergy with carbon capture and biochar, where plants are grown specifically to absorb CO$_2$ and then burned in a way that sequesters it (while creating energy or fertilizer as a useful byproduct) [414, 419, 420]. Finally, the way most of Earth’s CO$_2$ has been removed over geologic timescales is the slow process of mineral weathering, which also initiates further CO$_2$ absorption in the ocean due to alkaline runoff [421]. These processes can both be massively accelerated by human activity to achieve necessary scales of CO$_2$ removal [414]. However, although these biomass, mineral, or ocean-based methods are all promising enough as techniques to merit mention, they may have drawbacks in terms of land use and potentially serious environmental impacts, and (more relevantly for this paper) they would not likely benefit significantly from ML.

6.2 Direct air capture (DAC)

A more technological approach is to build facilities to extract CO$_2$ from power plant exhaust, industrial processes, or even directly from ambient air [422]. While this approach faces technical hurdles, it requires little land and has, according to current understanding, minimal negative environmental impacts [423]. The basic idea behind direct air capture is to blow air onto CO$_2$ sorbents (essentially like sponges, but for gas), which are either solid or in solution, then use heat-powered chemical processes to release the CO$_2$ in purified form for sequestration [414, 415]. Several companies have recently been started to pilot these methods [42, 43, 44].

While CO$_2$ sorbents are improving significantly [424, 425], issues still remain with efficiency and degradation over time, offering potential (though still speculative) opportunities for ML. ML could be used (as in §1.1.1) to accelerate the materials discovery process [75, 80, 81, 426] to maximize sorbent reusability and CO$_2$ uptake while minimizing the heat required for CO$_2$ release. ML might also help to develop corrosion-resistant components capable of withstanding high temperatures, as well as optimize their geometry for air-sorbent contact (which strongly impacts efficiency [427]).

6.3 Sequestering CO$_2$

Captured CO$_2$ will ultimately be released back into the atmosphere unless it is permanently stored. The best-understood form of CO$_2$ sequestration is direct injection into geologic formations such as saline aquifers, which are generally similar to oil and gas reservoirs [414]. A Norwegian oil company has successfully

---

https://carbonengineering.com/
https://www.climeworks.com/
https://globalthermostat.com/
sequestered CO\textsubscript{2} from an offshore natural gas field in a saline aquifer for more than twenty years \cite{428}. Another promising option is to sequester CO\textsubscript{2} in volcanic basalt formations, which is being piloted in Iceland \cite{429}.

Machine learning may be able to help with many aspects of CO\textsubscript{2} sequestration. First, ML can help identify and characterize potential storage locations. Oil and gas companies have had promising results using ML for subsurface imaging based on raw seismograph traces \cite{430}. These models and the data behind them could likely be repurposed to help trap CO\textsubscript{2} rather than release it. Second, ML can help monitor and maintain active sequestration sites. Noisy sensor measurements must be translated into inferences about subsurface CO\textsubscript{2} flow and remaining injection capacity \cite{431}; recently, \cite{432} found success using convolutional image-to-image regression techniques for uncertainty quantification in a global CO\textsubscript{2} storage simulation study. Additionally, it is important to monitor for CO\textsubscript{2} leaks \cite{433}. ML techniques have recently been applied to monitoring potential CO\textsubscript{2} leaks from wells \cite{434}; computer vision approaches for emissions detection (see \cite{435} and §5.1) may also be applicable.

6.4 Discussion

Given limits on how much more CO\textsubscript{2} humanity can safely emit and the difficulties associated with eliminating emissions entirely, CO\textsubscript{2} removal may have a critical role to play in tackling climate change. Promising applications for ML in CO\textsubscript{2} removal include informing research and development of novel component materials, characterizing geologic resource availability, and monitoring underground CO\textsubscript{2} in sequestration facilities. Although many of these applications are speculative, the industry is growing, which will create more data and more opportunities for ML approaches to help.
7 Climate Prediction

by Kelly Kochanski

The first global warming prediction was made in 1896, when Arrhenius estimated that burning fossil fuels could eventually release enough CO$_2$ to warm the Earth by $5^\circ$C. The fundamental physics underlying Arrhenius's calculations has not changed, but our predictions have become far more detailed and precise. The predominant predictive tools are climate models, known as general circulation models (GCMs) or Earth system models (ESMs). These models inform local and national government decisions (see IPCC reports [4, 26, 437]), help individuals calculate their climate risks (see §10) and allow us to estimate the potential impacts of solar geoengineering (see §9).

Recent trends have created opportunities for ML to advance the art of climate prediction. First, new and cheaper satellites are creating petabytes of climate observation data. Second, massive climate modeling projects are generating petabytes of simulated climate data. Third, climate forecasts are computationally expensive (the simulations in [440] took three weeks to run on NCAR supercomputers), but ML applications are driving the design of next-generation supercomputers that could ease current computational bottlenecks. As a result, climate scientists have recently begun to explore ML techniques, and are starting to team up with computer scientists to build new and exciting applications.

7.1 Uniting data, ML, and climate science

Climate models represent our understanding of Earth and climate physics. We can learn about the Earth by collecting data. To turn that data into useful predictions, we need to condense it into coherent, computationally tractable models. ML models are likely to be more accurate or less expensive than other models where: (1) there is plentiful data, but it is hard to model systems with traditional statistics, or (2) there are good models, but they are too computationally expensive to use in production.

When data is plentiful, climate scientists build many data-driven models. These models are mostly built by solving regression and classification problems, and new ML techniques may solve many problems that were previously challenging. For example, the authors of [442–444] use ML to calibrate satellite sensors, classify crop cover, and identify pollutant sources. More applications like these are likely to appear as satellite databases grow. This year, Reichstein et al. proposed that deep learning could be used extensively for pattern recognition, super-resolution, and short-term forecasting in climate models [445], and Mukkavilli proposed to compile a new labelled dataset of environmental imagery, called EnviroNet, that would accelerate ML work in environmental science [446]. We recommend that modellers who seek to learn directly from data see [447] for specific advice on fitting and over-fitting climate data.

Many climate prediction problems are irremediably data-limited. No matter how many weather stations we construct, how many field campaigns we run, or how many satellites we deploy, the Earth will generate at most one year of new climate data per year. Existing climate models deal with this limitation by relying heavily on physical laws, such as thermodynamics. ML models can leverage existing physics-based models as data sources to solve important climate problems.

Recent work has shown how deep neural networks and existing thermodynamics knowledge could be combined to fix the largest source of uncertainty in current climate models: clouds. Bright clouds block sunlight and cool the Earth; dark clouds catch outgoing heat and keep the Earth warm [437, 448]. These...
effects are controlled by small-scale processes such as cloud convection and atmospheric aerosols (see uses of aerosols for cloud seeding and solar geoengineering in §9). Physical models of these processes are far too computationally expensive to include in global climate models — but ML models are not. Gentile et al. trained a deep neural network to emulate the behavior of a high-resolution cloud simulation, and found that the network gave similar results for a fraction of the cost [449] and was stable in a simplified global model [450]. Existing scientific models have fixed trade-offs between cost and accuracy, and sometimes these trade-offs do not include any great solutions. Neural networks trained on those scientific models produce similar predictions, but offer an entirely new set of compromises between training cost, production cost, and accuracy. Replacing select climate model components with neural network approximators may thus improve both the cost and the accuracy of global climate models. Additional work is needed to optimize the cloud model above; to identify more climate model components that could be replaced by neural networks (we highlight other impactful components below); to train neural networks that replace those components; and to build pipelines that re-train these neural networks in response to errors or extrapolation (example workflow in §4.5 of [442]).

The next most important targets for climate model improvements are ice sheet dynamics and sea level rise. The Arctic and Antarctic are warming faster than anywhere else on Earth, and their climates control the future of global sea level rise and many vulnerable ecosystems [4, 26]. Unfortunately, these regions are dark and cold, and until recently they were difficult to observe. In the past few years, however, new satellite campaigns have illuminated them with hundreds of terabytes of data. These data could make it possible to use ML to solve some of the field’s biggest outstanding questions. In particular, models of mass loss from the Antarctic ice-sheet are highly uncertain [451] and models of the extent of Antarctic sea ice do not match reality well [452]. The most uncertain parts of these models, and thus the best targets for improvement, are snow reflectivity, sea ice reflectivity, ocean heat mixing and ice sheet grounding line migration rates [447, 451, 453]. Computer scientists who wish to work in this area could build models that learn snow and sea ice properties from satellite data, or use new video prediction techniques (e.g. [454]) to predict short-term changes in the extent of sea ice.

ML could also improve climate model efficiency by identifying and leveraging relationships between climate variables. For example, Nowack et al. demonstrated that ozone concentrations could be computed as a function of temperature, rather than physical transport laws, which led to considerable computational savings [455]. Pattern recognition and feature extraction techniques could allow us to identify more useful connections in the climate system, and regression models could allow us to quantify non-linear relationships between connected variables.

In the further future, the Climate Modeling Alliance has proposed to build an entirely new climate model that learns continuously from data and from high-resolution simulations [456]. The proposed model would be written in Julia, in contrast to existing models which are mostly written in C++ and inherited Fortran. At the cost of a daunting translation workload, they aim to build a model that is more accessible to new developers and more compatible with ML libraries.

Finally, the best climate predictions are synthesized from ensembles of 20+ climate models [457]. Making good ensemble predictions is an excellent ML problem. Monteleoni et al. proposed that online ML algorithms (e.g. [458]) could select the best-performing model at any given point in time [459]; this idea has been refined in further work [460, 461]. More recently, Anderson and Lucas used random forests to make high-resolution predictions from a mix of high- and low-resolution models, thereby reducing the costs of building multi-model ensembles [462]. These studies leave room for the development of more specialized and sophisticated ensemble methods. For example, climate models serve many users with different objectives. The model in [459] optimizes the ensemble to predict global temperature; however, their solution is not necessarily optimal for users who need predictions of local temperatures, local rainfall, or the dates the

---

48See e.g. icebridge.gsfc.nasa.gov and pgc.umn.edu/data/arcticdem
Northwest Passage will open for shipping.

7.2 Forecasting extreme events

For most people, extreme event prediction means the local weather forecast and a few days’ warning to stockpile food, go home, and lock the shutters. Weather forecasts are shorter-term than climate forecasts, but they produce abundant data that makes them amenable to some ML techniques that would not work in climate models. Weather models are optimized to track the rapid, chaotic changes of the atmosphere; since these changes are fast, tomorrow’s weather forecast is made and tested every day. Climate models, in contrast, are chaotic on short time scales, but their long-term trends are driven by slow, predictable changes of ocean, land, and ice (see [463]). As a result, climate model output can only be tested against long-term observations (at the scale of years to decades). Intermediate time scales, of weeks to months, are exceptionally difficult to predict, although Cohen et al. [464] argue that machine learning could bridge that gap by making good predictions on four to six week timescales [465]. Thus far, however, weather modelers have had hundreds of times more test data than climate modelers, and began to adopt ML techniques earlier. Numerous ML weather models are already running in production. For example, Gagne et al. recently used an ensemble of random forests to improve hail predictions within a major weather model [466].

Climate models do predict changes in long-term trends like drought frequency and storm intensity, although they cannot predict the specific dates of future events. These trends help individuals, corporations and towns to make informed decisions about infrastructure, asset valuation and disaster response plans (see also §8.4). Identifying extreme events in climate model output, however, is a classification problem with a twist: all of the available data sets are strongly skewed because extreme events are, by definition, rare. ML has been used successfully to classify some extreme weather events. Liu et al. used deep convolutional neural networks to count cyclones and weather fronts in climate data sets [467], and Lakshmanan has devised a series of techniques to track storms and tornadoes (e.g. [468]). Tools for more event types would be useful, as would online tools that work within climate models, and statistical tools that quantify the uncertainty in new extreme event forecasts.

Forecasts are most actionable if they are specific and local. ML is widely used to make local forecasts from coarse 10–100 km climate or weather model predictions; various authors have attempted this using support vector machines, autoencoders, Bayesian deep learning, and super-resolution convolutional neural networks (e.g. [469]). Several groups are now working to translate high-resolution climate forecasts into risk scenarios. For example, ML can predict localized flooding patterns from past data [470], which could inform individuals buying insurance or homes. Currently, flood maps from the U.S. Federal Emergency Management Agency (FEMA) (part of the National Flood Insurance Program) do not account for the effects of climate change on flooding [471]. Since ML methods like neural networks are effective at predicting local flooding during extreme weather events [472], these could be used to update local flood risk estimates to benefit individuals. The start-up Jupiter Intelligence [50] is working to make climate predictions more accessible and actionable to companies and local governments, by translating climate forecasts into localised flood and temperature risk scores.

A full review of the applications of ML for extreme weather forecasting is beyond the scope of this article. Fortunately, that review has already been written: see [473]. The authors describe ML systems that correct bias, recognize patterns, and predict storms. Moving forward, they envision human experts working in sync with automated forecasts.

49 This is one of several reasons why climate models produce accurate long-term predictions in spite of the chaotic nature of the atmosphere.

50 https://jupiterintel.com/
7.3 Discussion

ML may change the way that scientific modeling is done. The examples above have shown that many components of large climate models can be replaced with ML models at lower computational costs. From an ML standpoint, learning from an existing model has many advantages: modelers can generate new training and test data on-demand, and the new ML model inherits some community trust from the old one. This is an area of active ML research. Recent papers have explored data-efficient techniques for learning dynamical systems [474], including physics-informed neural networks [475] and neural ordinary differential equations [128]. In the further future, researchers are developing ML solutions for a wide range of scientific modeling challenges, including crash prediction [476], adaptive numerical meshing [477], uncertainty quantification [478, 479] and performance optimization [480]. If these solutions are effective, they may solve some of the largest structural challenges facing current climate models.

New ML models for climate will be most successful if they are closely integrated into existing scientific models. This has been emphasized, again and again, by authors who have laid future paths for artificial intelligence within climate science [450, 456, 473, 481]. New models need to leverage existing knowledge to make good predictions with limited data. In ten years, we will have more satellite data, more interpretable ML techniques, hopefully more trust from the scientific community, and possibly a new climate model written in Julia. For now, however, ML models must be creatively designed to work within existing climate models. The best of these models are likely to be built by close-knit teams including both climate and computational scientists.
8 Societal Impacts

Changes in the atmosphere have impacts on the ground. The expected societal impacts of climate change include prolonged ecological and socioeconomic stresses as well as brief, but severe, societal disruptions. For example, impacts could include both gradual decreases in crop yield and localized food shortages. If we can anticipate climate impacts well enough, then we can prepare for them by asking:

- How do we reduce vulnerability to climate impacts?
- How do we support rapid recovery from climate-induced disruptions?

A wide variety of solutions have been put forward, from robust power grids to food shortage prediction, and while this is good news for society, it can be overwhelming for an ML researcher hoping to contribute. Fortunately, a few critical needs tend to recur across solution strategies – it is by meeting these needs that ML has the greatest potential to support societal adaptation [8, 17, 482]. From a high-level, these involve

- Sounding alarms: Identifying and prioritizing the areas of highest risk, by using evidence of risk from historical data.
- Providing annotation: Extracting actionable information or labels from unstructured raw data.
- Promoting exchange: Making it easier to share resources and information to pool and reduce risk.

These unifying threads will appear repeatedly in the sections below, where we review strategies to help ecosystems, infrastructure, and societies adapt to climate change, and explain how ML supports each strategy.

8.1 Ecology

Changes in climate are increasingly affecting the distribution and composition of ecosystems. This has profound implications for global biodiversity, as well as agriculture, disease, and natural resources such as wood and fish. ML can help by supporting efforts to monitor ecosystems and biodiversity.

Ecosystem monitoring

To preserve ecosystems, it is important to know which are most at risk. This has traditionally been done via manual, on-the-ground observation, but the process can be accelerated by annotation of remote sensing data [483–486] (see also §5.1). For example, tree cover can be automatically extracted from aerial imagery to characterize deforestation [487, 488]. At the scale of regions or biomes, analysis of large-scale simulations can illuminate the evolution of ecosystems across potential climate futures [489, 490]. A more direct source of data is offered by environmental sensor networks, made from densely packed but low-cost devices [12, 491, 492].

For a system to have the most real-world impact, regardless of the underlying data source, it is necessary to “personalize” predictions across a range of ecosystems. A model trained on the Sahara would almost certainly fail if deployed in the Amazon. Hence, these applications may motivate ML researchers interested in heterogeneity, data collection, transfer learning, and rapid generalization. In sensor networks, individual nodes fail frequently, but are redundant by design – this is an opportunity for research into anomaly detection and missing data imputation [493, 494]. Finally, beyond aiding adaptation by prioritizing at-risk environments, the design of effective methods for ecosystem monitoring will support the basic science necessary to shape adaptation in the long-run [11, 13, 495].
Biodiversity monitoring

Accurate estimates of species populations are the foundation on which conservation efforts are built. Camera traps and aerial imagery have increased the richness and coverage of sampling efforts. The idea of a camera trap is to have photos taken automatically whenever a motion sensor is activated – computer vision can be used to classify the species that pass by, supporting a real-time, less labor-intensive species count [496, 497]. It is also possible to use aerial imagery for estimating the size of large herds [498]. Alternatively, citizen science can enable dataset collection at a scale impossible in individual studies [499–502]. For example, by leveraging the general public’s enthusiasm for birdwatching, eBird has logged more than 140 million observations [499], which have been used for population and migration studies [503].

It is worth noting that these are not routine ML applications. Camera trap data have wide class imbalances, and species are related hierarchically, separated by fine-grained differences. Further, given these types of data, it is possible to formulate resource allocation policies in a quantitative way, either to protect rare species or to control invasive ones [504–507], and it can be important to make these policies be robust to poaching activity [508]. Citizen science poses unique challenges, as researchers have no control over where samples come from. To incentivize observations from undersampled regions, mechanisms from game theory can be applied [509], and even when sampling biases persist, estimates of dataset shift can minimize their influence [510].

8.2 Infrastructure

Physical infrastructure is so tightly woven into the fabric of everyday life – like the buildings we inhabit and lights we switch on – that it is easy to forget that it exists (see §3). The fact that something so basic will have to be rethought in order to adapt to climate change can be unsettling, but viewed differently, the sheer necessity of radical redesign can inspire creative thinking.

We first consider the impacts of climate change on the built environment. Shifts in weather patterns are likely to put infrastructure under more persistent stress. Heat and wind damage roads, buildings, and power lines. Rising water tables near the coast will lead to faults in pipelines. Urban heat islands will be exacerbated and flooding caused by heavy rain or coastal inundations will become more routine, along with resulting property damage and traffic blockages.

A clear target is construction of physical defenses – for example, “climate proofing” cities with new coastal embankments and increased storm drainage capacity. However, focusing solely on defending existing structures can stifle proactive thinking about urban and social development – for example, floating buildings are being tested in Rotterdam – and one may alternatively consider resilience and recovery more broadly [511, 512]. From this more general perspective of improving social processes, ML can support two types of activities: Design and maintenance.

Design

How can infrastructure be (re)designed to dampen climate impacts? In road networks, it is possible to incorporate flood hazard and traffic information in order to uncover vulnerable stretches of road, especially those with few alternative routes [513]. If traffic data are not directly available, it is possible to construct proxies from mobile phone usage and city-wide CCTV streams – these are promising in rapidly developing urban centers [514, 515]. More than drawing from flood hazard maps, it is possible to use data from real-world flooding events [516], and to send localized predictions to those at risk [517]. For electrical, water, and waste collection networks, the same principle can guide investments in resilience – using proxy or historical data about disruptions to anticipate vulnerabilities [518–521]. Robust components can replace those at risk; for example, *adaptive islands*, parts of an energy grid that continue to provide power even when disconnected from the network, prevent cascading outages in power distribution [522].
Infrastructure is long-lived, but the future is uncertain, and planners must weigh immediate resource costs against future societal risks [523]. One area that urgently needs adaptation solutions is the consistent access to drinking water, which can be jeopardized by climate variability [524]. Investments in water infrastructure can be optimized; for example, a larger dam might cost more up front, but would have a larger storage capacity, giving a stronger buffer against drought. To delay immediate decisions, infrastructure can be upgraded in phases – the technical challenge is to discover policies that minimize a combination of long-term resource and societal costs under plausible climate futures, with forecasts being updated as climates evolve [525, 526].

Maintenance
What types of systems can keep infrastructure functioning well under increased stress? Two strategies for efficiently managing limited maintenance resources are predictive maintenance and anomaly detection; both can be applied to electrical, water, and transportation infrastructure. In predictive maintenance, operations are prioritized according to the predicted probability of a near-term breakdown [119, 120, 527, 528]. For anomaly detection, failures are discovered as soon as they occur, without having to wait for inspectors to show up, or complaints to stream in [529, 530].

The systems referenced here have required the manual curation of data streams, structured and unstructured. The data are plentiful, just difficult to glue together. Ideas from the missing data, multimodal data, and AutoML communities have the potential to resolve some of these issues.

8.3 Social systems
While less tangible, the social systems we construct are just as critical to the smooth functioning of society as any physical infrastructure, and it is important that they adapt to changing climate conditions. First, consider what changes these systems may encounter. Decreases in crop yield for many environments, due to desertification and drought, will pose a threat to food security, as already evidenced by long periods of drought in North America, West Africa and East Asia [531]. More generally, communities dependent on ecosystem resources will find their livelihoods at risk, and this may result in mass migrations, as people seek out more supportive environments.

At first, these problems may seem beyond the reach of algorithmic thinking, but investments in social infrastructure can increase resilience. ML can amplify the reach and effectiveness of this infrastructure.

Food security
Data can be used to monitor the risk of food insecurity in real time, to forecast near-term shortages, and to identify areas at risk in the long-term, all of which can guide interventions. For real-time and near-term systems, it is possible to distill relevant signals from mobile phones, credit card transactions, and social media data [532–534]. These have emerged as low-cost, high-reach alternatives to manual surveying. The idea is to train models that link these large, but decontextualized, data with ground truth consumption or survey information, collected on small representative samples. This process of developing proxies to link small, rich datasets with large, coarse ones can be viewed as a type of semi-supervised learning, and is fertile ground for research.

For longer-term warnings, spatially localized crop yield predictions are needed. These can be generated by aerial imagery or meteorological data (see 5.2), if they can be linked with historical yield data [535, 536]. On the ground, it is possible to perform crop-disease identification from plant photos – this can alert communities to disease outbreaks, and enhance the capacity of agricultural inspectors. For even longer-run risk evaluation, it is possible to simulate crop yield via biological and ecological models [537–539], presenting another opportunity for blending large scale simulation with ML [540, 541].
Beyond sounding alarms, ML can improve resilience of food supply chains. As detailed in §4, ML can reduce waste along these chains; we emphasize that for adaptation, it is important that supply chains also be made robust to unexpected disruptions [542–545].

**Resilient livelihoods**

Individuals whose livelihoods depend on one activity, and who have less access to community resources, are those who are most at risk [546, 547]. Resilient livelihoods can be promoted through increased diversification, cooperation, and exchange, all of which can be facilitated by ML systems. For example, they can guide equipment and information sharing in farming cooperatives, via growers’ social networks [548]. Mobile money efforts can increase access to liquid purchasing power; they can also be used to monitor economic health [549, 550]. Skill-matching programs and online training are often driven by data, with some programs specifically aiming to benefit refugees [551–553].

**Migration**

Human populations move in response to threats and opportunities, and ML can be used to predict large-scale migration patterns. Work in this area has relied on accessible proxies, like social media, where users’ often self-report location information, or aerial imagery, from which the extent of informal settlement can be gauged [554–557]. More than quantifying migration patterns, there have been efforts directly aimed at protecting refugees, either through improving rescue operations [558, 559] or monitoring negative public sentiment [560]. It is worth cautioning that immigrants and refugees are vulnerable groups, and systems that surveil them can easily be exploited by bad actors. Designing methodology and governance mechanisms that allow vulnerable populations to benefit from such data, without putting them at additional risk, should be a research priority.

Across social applications, there are worthwhile research challenges – guiding interventions based on purely observational, potentially unrepresentative data poses risks. In these contexts, transparency is necessary, and ideally, causal effects of interventions could be estimated, to prevent feedback loops in which certain subgroups are systematically ignored from policy interventions.

**8.4 Crisis**

Perhaps counterintuitively, natural disasters and health crises are not entirely unpredictable – they can be prepared for, risks can be reduced, and coordination can be streamlined. Furthermore, while crises may be some of the most distressing consequences of climate change, disaster response and public health are mature disciplines in their own right, and have already benefited extensively from ML methodology [561–563].

**Epidemics**

Disease surveillance and outbreak forecasting systems can be built from web data and specially-designed apps, in addition to traditional surveys [564–566]. While non-survey proxies are observational and self-reported, current research attempts to address these issues [567, 568]. Beyond surveillance, point-of-care diagnostics have enjoyed a renaissance, thanks in part to ML [482, 569]. These are tools that allow health workers to make diagnoses when specialized lab equipment is inaccessible. An example is malaria diagnosis based on photos of prepared pathology slides taken with a mobile phone [570]. Ensuring that these systems reliably and transparently augment extension workers, guiding data collection and route planning when appropriate, are active areas of study [571, 572].

**Disaster relief**

In disaster preparation and response, two types of ML tasks have proven useful: creating maps from aerial
imagery and performing information retrieval on social media data. Accurate and well-annotated maps can inform evacuation planning, retrofitting campaigns, and delivery of relief [573, 574]. Further, this imagery can assist damage assessment, by comparing scenes immediately pre- and post-disaster [575]. Social media data can contain kernels of insight – places without water, clinics without supplies – which can inform relief efforts. ML can help properly surface these insights, compressing large volumes of social media data into the key takeaways, which can be acted upon by disaster managers [562, 576, 577].

8.5 Discussion

Climate change will have profound effects on the planet, and the ML community can support efforts to minimize the damage it does to ecosystems and the harm it inflicts on people. This section has suggested areas of research that may help societies adapt more effectively to these ever changing realities. We have identified a few recurring themes, but also emphasized the role of understanding domain-specific needs. The use of ML to support societal resilience would be a noble goal at any time, but the need for tangible progress towards it may never have been so urgent as it is today, in the face of the wide-reaching consequences of climate change.
9 Solar Geoengineering

Airships floating through the sky, spraying aerosols; robotic boats crisscrossing the ocean, firing vertical jets of spray; arrays of mirrors carefully positioned in space, micro-adjusted by remote control: these images seem like science fiction, but they are actually real proposals for solar radiation management, commonly called solar geoengineering [578–581].

Solar geoengineering, much like the greenhouse gases causing climate change, shifts the balance between how much heat the Earth absorbs and how much it releases. The difference is that it is done deliberately, and in the opposite direction. The most common umbrella strategy is to make the Earth more reflective, keeping heat out, though there are also methods of helping heat escape (besides CO₂ removal, which we discuss in §5 and §6).

Solar geoengineering generally comes with a host of potential side effects and governance challenges. Moreover, unlike CO₂ removal, it cannot simply reverse the effects of climate change (average temperatures may return to pre-industrial levels, but location-specific climates still change), and also comes with the risk of termination shock (fast, catastrophic warming if humanity undertakes solar geoengineering but stops suddenly) [582]. Because of these and other issues, it is not within the scope of this paper to evaluate or recommend any particular technique. However, the potential for solar geoengineering to moderate some of the most catastrophic hazards of climate change is well-established [583], and it has received increasing attention in the wake of societal inaction on mitigation. Although [581] argue that the “hardest and most important problems raised by solar geoengineering are non-technical,” there are still a number of important technical questions that machine learning may be able to help us study.

9.1 Overview of methods

The primary candidate methods for geoengineering are marine cloud brightening [584] (making low-lying clouds more reflective), cirrus thinning [585] (making high-flying clouds trap less heat), and stratospheric aerosol injection [586] (which we discuss below). Other candidates (which are either less effective or harder to implement) include “white-roof” methods [587] and even launching sunshades into space [588].

Injecting sulfate aerosols into the stratosphere is considered a leading candidate for solar geoengineering both because of its economic and technological feasibility [589, 590] and because of a reason that should resonate with the ML community: we have data. (This data is largely in the form of temperature observations after volcanic eruptions, which release sulfates into the stratosphere when sufficiently large [591].) Once injected, sulfates circulate globally and remain aloft for 1 to 2 years. As a result, the process is reversible, but must also be continually maintained. Sulfates come with a well-studied risk of ozone loss [592], and they make sunlight slightly more diffuse, which can impact agriculture [593].

9.2 Designing aerosols

The effects and side-effects of aerosols in the stratosphere (or at slightly lower altitudes for cirrus thinning [594]) vary significantly with their optical and chemical properties. Although sulfates are the best understood due to volcanic eruption data, many others have been studied, including zirconium dioxide, titanium dioxide, calcite (which preserves ozone), and even synthetic diamond [595]. However, the design space is far from fully explored. Machine learning has had recent success in predicting or even optimizing for specific chemical and material properties [75, 80, 81, 426]. Although speculative, it is conceivable that ML could accelerate the search for aerosols that are chemically nonreactive but still reflective, cheap, and easy to keep aloft.
9.3 Aerosol modeling

One reason that sulfates have been the focus for aerosol research is that atmospheric aerosol physics is not perfectly captured by current climate models, so having natural data is important for validation. Furthermore, even if current aerosol models are correct, their best-fit parameters must still be determined (using historical data), which comes with uncertainty and computational difficulty. ML may offer tools here, both to help quantify and constrain uncertainty, and to manage computational load. As a recent example, [596] use Gaussian processes to emulate climate model outputs based on nine possible aerosol parameter settings, allowing them to establish plausible parameter ranges (and thus much better calibrated error-bars) with only 350 climate model runs instead of >100,000. Although this is important progress, ideally we want uncertainty-aware aerosol simulations with a fraction of the cost of one climate model run, rather than 350. ML may be able to help here too (see §7 for more details).

9.4 Engineering a planetary control system

Efficient emulations and error-bars will be essential for what MacMartin and Kravitz [597] call “The Engineering of Climate Engineering.” According to [597], any practical deployment of geoengineering would constitute “one of the most critical engineering design and control challenges ever considered: making real-time decisions for a highly uncertain and nonlinear dynamic system with many input variables, many measurements, and a vast number of internal degrees of freedom, the dynamics of which span a wide range of timescales.” Bayesian and neural network-based approaches could facilitate the fast, uncertainty-aware nonlinear system identification this challenge might require. Additionally, there has been recent progress in reinforcement learning for control [598–600], which could be useful for fine-tuning geoengineering interventions. For an initial attempt at analyzing stratospheric aerosol injection as a reinforcement learning problem (using a neural network climate model emulator), see [601].

9.5 Impact modeling

Of course, optimizing interventions requires defining objectives, and the choices here are far from clear. Although it is possible to stabilize global mean temperature and even regional temperatures through geoengineering, it is most likely impossible to preserve all relevant climate characteristics in all locations. Furthermore, climate model outputs do not tell the full story; ultimately, the goal of climate engineering is to minimize harm to people, ecosystems, and society. It is therefore essential to develop robust tools for estimating the extent and distribution of these potential harms.

The field of Integrated Assessment Modeling [602, 603] aims to map the outputs of a climate model to societal impacts, but such approaches have significant limitations [604] and suffer from the ever-present problem of “garbage in, garbage out.” As such, opportunities for improving them with ML seem limited. However, there has been recent progress using ML to predict a more limited but well-defined set of geoengineering or aerosol impacts. For example, [605] use deep neural networks to estimate the effects of aerosols on human health, while [606] use them to estimate the effects of solar geoengineering on agriculture.

9.6 Discussion

Any consideration of solar geoengineering raises many moral questions. It may help certain regions at the expense of others, introduce risks like termination shock, and serve as a “moral hazard”: widespread awareness of its very possibility may undermine mainstream efforts to cut emissions [607]. Because of these issues, there has been significant debate about whether it is ethically responsible to research this topic [608, 609]. However, although it creates new risks, solar geoengineering could actually be a moderating force against the terrifying uncertainties climate change already introduces [583, 610], and ultimately...
many environmental groups and governmental bodies have come down on the side of supporting further re-
search.\textsuperscript{51, 52, 53} In this section, we have attempted to outline some of the technical challenges in implement-
ing and evaluating solar geoengineering. We hope the ML community can help geoengineering researchers
tackle these challenges.

\textsuperscript{51} https://www.edf.org/climate/our-position-geoengineering
\textsuperscript{52} https://www.nrdc.org/media/2015/150210
\textsuperscript{53} https://www.ucsusa.org/sites/default/files/attach/2019/gw-position-Solar-Geoenginee-
ing-022019.pdf
Tools for Action

10 Tools for Individuals

by Natasha Jaques

It is a common misconception that individuals cannot take meaningful action on climate change. However, over thirty climate change mitigation solutions that involve behavior change on an individual level have been identified, which collectively could mitigate 20–37% of global emissions from 2020 to 2050 [18, 611].

10.1 Understanding and reducing personal carbon footprint

We are constantly confronted with decisions that affect our carbon footprint, but we may lack the data and knowledge to know which decisions are most impactful. Fortunately, ML can help determine an individual’s carbon footprint from their personal and household data. For example, natural language processing can be used to extract the flights a person takes from their email, or determine specific grocery items purchased from a bill, making it possible to predict the associated emissions. Systems that combine this information with data obtained from the user’s smartphone (e.g. from a ride-sharing app) can then help consumers identify which behaviors result in the highest emissions. Given such a ML model, counterfactual reasoning can potentially be used to demonstrate to consumers how much their emissions would be reduced for each behavior they changed. As a privacy-conscious alternative, emissions estimates could be directly incorporated into grocery labels or interfaces for purchasing flights. Such information can empower people to understand how they can best help mitigate climate change through behavior change.

10.2 Calculating and reducing household energy impact

Residences are responsible for approximately 30% of global electricity consumption [613]. A large meta-analysis found that significant residential energy savings can be achieved [614], by targeting the right interventions to the right households [615–617]. ML can predict a household’s emissions in transportation, energy, water, waste, foods, goods, and services, as a function of its characteristics [618]. These predictions can be used to tailor customized interventions for high-emissions households [619]. Changing behavior both helps mitigate climate change and benefits individuals; studies have shown that many carbon mitigation strategies also provide cost savings to consumers [618].

Household energy disaggregation provides consumers with appliance-level energy use data from the aggregate electricity consumption signal (see also §3.1) [621], which can help facilitate behavior change [621]. For example, it can be used to inform consumers of high-energy appliances of which they were previously unaware. This alone could have a significant impact, since many devices consume a large amount of electricity even when not in use; standby power consumption accounts for roughly 8% of residential electricity demand [622]. A variety of ML techniques have been used to effectively disaggregate household energy, such as spectral clustering, Hidden Markov Models, and neural networks [620].

ML can also be used to predict the marginal emissions of energy consumption in real time, on a scale of hours, potentially allowing consumers to effectively schedule activities such as charging an electric vehicle when the emissions (and prices [623]) will be lowest [624]. Combining these predictions with disaggregated energy data allows for the efficient automation of household energy consumption, ideally through products that present interpretable insights to the consumer (e.g. [625, 626]). Methods like reinforcement learning can be used to learn how to optimally schedule household appliances to consume energy more efficiently.

---

See e.g. [https://www.tmrow.com/](https://www.tmrow.com/)
[https://www.watttime.org/](https://www.watttime.org/)
and sustainability [627, 628]. Multi-agent learning has also been applied to this problem, to ensure that groups of homes can coordinate to balance energy consumption to keep peak demand low [68, 71].

10.3 Modeling consumer behavior and facilitating behavior change

ML is highly effective at modeling human preferences, and this can be leveraged to help mitigate climate change. Using ML, we can model and cluster individuals based on their climate knowledge, preferences, demographics, and consumption characteristics (e.g. [629–633]), and thus predict who will be most amenable to new technologies and sustainable behavior change. Such techniques have improved the enrollment rate of customers in an energy savings program by 2-3x [615]. Other authors have used ML to predict how much consumers are willing to pay to avoid potential environmental harms of energy consumption [634], finding that some groups were totally insensitive to cost and would pay the maximum amount to mitigate harm, while other groups were willing to pay nothing. Given such disparate types of consumers, targeting interventions toward particular households may be particularly worthwhile; all the more so because data show that the size and composition of household carbon footprints varies dramatically across geographic regions and demographics [618].

Understanding individual behaviour can help signal how it can be nudged. For example, path analysis has shown that an individual’s psychological distance to climate change (on geographic, temporal, social, and uncertainty dimensions) fully mediates their level of climate change concern [635]. This suggests that interventions minimizing psychological distance to the effects of climate change may be most effective. Similarly, ML has revealed that cross-cultural support for international climate programs is not reduced, even when individuals are exposed to information about other countries’ climate behavior [636]. To make the effects of climate change more real for consumers, and thus encourage climate support, image generation techniques such as CycleGANs have been used to visualize the potential consequences of extreme weather events on houses and cities [637]. Gamification via deep learning has been proposed to further push individuals to consider their personal energy usage [638]. All of these programs may be an incredibly cost-effective way to reduce energy consumption; behavior change programs can cost as little as 3 cents to save a kilowatt hour of electricity, whereas generating one kWh would cost 5-6 cents with a coal or wind power plant, and 10 cents with solar [639, 640].

10.4 Discussion

While individuals can sometimes feel that their contributions to climate change are dwarfed by other factors, in reality individual actions can have a significant impact in mitigating climate change. ML can aid this process by empowering consumers to understand which of their behaviors lead to the highest emissions, automatically scheduling energy consumption, and providing insights into how to facilitate behavior change.
11 Tools for Society

Like public safety or roads, a stable climate is a good that is ultimately shared across all of society. In managing shared goods, we often encounter public goods problems such as the tragedy of the commons [641], wherein the short-term incentives of individuals are not aligned with desirable long-term outcomes – i.e. people can’t, won’t, or don’t know how to act in a way that will sustain the public good for the benefit of all. When we make decisions at a societal level – e.g. via policy, markets, or large-scale planning – we often deal with public goods problems, and ML/AI can provide tools for action and decision-making. It is worth noting that many methods of addressing public goods problems require little to no ML and/or are not applicable to problems of climate change, and while impactful are thus outside the scope of this work.

11.1 Policy design

Decision-makers may wish to use ML for creating socially beneficial policies, which involves understanding how these policies will pan out when they are applied in the real-world. Tools from game theory and incentive/mechanism design have been applied to develop climate policy [642–644], but there are many opportunities for machine learning in this area, including exploration of incentive design, application of multi-agent reinforcement learning (RL) to planning and coordination in climate change policy or mitigation, decision support systems, and data visualization tools.

Statistical and ML tools can help model and understand peoples’ preferences and behaviors in order to design better policy. For example, ML can be used to analyze social norms, behavior, and public opinion about climate change [616, 644–646], as well as the effectiveness of imagery for climate-change engagement [647, 648]. Some works have also used agent-based models to better understand behavior from a climate economics perspective [617, 649, 650]. Such analyses are important for informing policy, engagement, and education efforts. Additionally, methods for semi-supervised learning and tools for interpretability and explainability of learned models could provide valuable insights in this area.

11.2 Market design

Decision-makers may wish to design market prices that optimize for social good (e.g. reduce GHG emissions). Following seminal research on modeling pricing in markets [65] as a bandit problem [652], many works have applied bandit and other RL algorithms to determine prices or other market values. For example, RL has been applied to predict bids [653] and market power [654] in electricity markets, to help solve auctions in supply chains [172], and to set dynamic pricing of carbon [655, 656] and electricity [657] (see also §1 and §2.1).

To optimally set prices, it may be valuable for decision-makers to understand the price at which individuals value goods. Economists have begun to apply ML to hedonic pricing, which uses consumer behavior to understand the value consumers place on goods such as housing [658–660]. In parallel, environmental economists have used hedonic pricing to understand how humans value “bads” such as air pollution [661]. Speculatively, as the impacts of climate change become more visible, governments could employ ML-based hedonic pricing techniques to understand how society values the impacts of climate change and carbon regulation.

56Markets here refers to the term in economics: a model or platform in which goods and services are traded. See [65] for an introduction to economics which covers markets and other terms used in this section.
11.3 Multi-objective optimization and decision-making

Tools from the fields of *multi-criteria decision-making* and *multi-objective optimization* are broadly useful for managing trade-offs related to climate change. These are rich fields in which relatively off-the-shelf ML algorithms can provide considerable value in addressing problems of climate change. There are many opportunities for collaborations between ML researchers and experts on a particular problem to produce more sophisticated, tailored solutions or deeper understanding of that problem. This overlap has already been fruitful; numerous papers have considered problems relevant to climate change from the perspective of reconciling conflicting objectives and minimizing side-effects (*negative externalities*): for instance, in energy planning [662–664] and generation (e.g. [665]), computational sustainability for design or planning of infrastructure [89, 666, 667] (see also §3) and manufacturing [668], waste management [669], decisions about land use (e.g. [670, 671]; see also §5), supply-chain management [672], air-quality measurement [673], and policy development [674]. Many such methods use bio-inspired algorithms such as particle swarm, genetic, or evolutionary algorithms to search for or compute Pareto-optimal solutions. These methods can provide appealingly pragmatic solutions for businesses or governments transitioning to lower-carbon technology in balancing the interests of different stakeholders in policy decisions.
High-quality education is paramount to helping different populations understand and address the causes and consequences of climate change, and to acquire skills and tools that will help them adapt to its impacts. More specifically, education can both improve the resilience of communities in developing countries that will be disproportionately affected by climate change \cite{675} and empower learners from developed countries to adopt more sustainable lifestyles \cite{676}. Paradoxically, climate change itself may reduce educational outcomes for some populations, due to its negative effects on agricultural productivity and household income \cite{677, 678}, making high-quality educational interventions all the more important.

12.1 AI in Education

Leveraging ML in education allows teaching to be personalized and scalable, giving increasing numbers of learners access to high-quality, interactive and relevant educational content. ML-based technology can be used either as a standalone tool for independent learners or as an educational resource that frees up teachers to have more one-on-one time with students. The field of AIED (Artificial Intelligence in Education) has existed for over 30 years, previously relying on explicitly modeling content, learners and tutoring strategies, but increasingly incorporating ML in research. For instance, Intelligent Tutoring Systems (ITS) have been a key tool in the field of AIED, with much research being done to create digital environments that can adapt their behavior in real time according to the needs of individuals or to support collaborative learning \cite{679}. While ITSs are traditionally defined and constructed by hand based on cognitive theories of learning, some recent approaches have been more data-driven, applying Multi-Arm Bandit techniques to adaptively personalize sequences of learning activities \cite{680}, LSTMs to generate questions to evaluate language comprehension \cite{681}, and RL to improve the strategies used within the ITS \cite{682, 683, 57}.

AIED also has the potential to improve educational outcomes in developing countries: for instance, creating scalable, adaptive online courses could give hundreds of thousands of learners in the developing world access to learning resources that they would not usually have in their local educational facilities \cite{687}. Furthermore, giving teachers guidance derived from computational teaching algorithms or heuristics could help them design better educational curricula and improve student learning outcomes \cite{688}. Key considerations for creating AIED tools in developing countries include adapting to local technological and cultural needs, addressing barriers such as access to electricity and internet \cite{121, 122}, and taking into account students’ computing skills, language, and culture \cite{689, 690}.

12.2 Learning about Climate

Research has shown that simple learning activities around concepts like ecological footprint, greenhouse gases and global warming can personally engage learners in understanding the connection between personal energy use and climate change, guiding them towards making sustainable lifestyle changes \cite{691}. There have also been proposals for active-response websites explaining climate science as well as educational interventions focusing on local and actionable aspects of sustainable development \cite{692}; both of these could benefit greatly from ML approaches such as NLP techniques (for creating conversational agents) \cite{693} and adaptive learning approaches \cite{694}.

\footnote{For further background on this area, see \cite{684, 686}.}
The rise and fall of financial markets is linked to many events, both sporadic (e.g. the 2008 global financial crisis) and cyclical (e.g. the price of gas over the years), with profits and losses that can be measured in the billions of dollars and can have global consequences. While the total cumulative impact of climate change on global financial assets is estimated to be $2.5 trillion \[695\], its concrete consequences are unpredictable: it is hard to forecast where, how, or when it is going to impact the stock price of a given company, or even the debt of an entire nation. While financial analysts and investors focus on pricing risk and forecasting potential earnings, the majority of the current financial system is based on quarterly or yearly performance. This fails to incentivize the prediction of medium or long-term risks, which include most climate change-related exposures such as its physical impacts on assets or distribution chains, its legislative impacts on profit generation, and indirect market consequences such as supply and demand.\[58\]

### 13.1 Climate Investment

*Climate investment*, the current dominant approach in climate finance, involves investing money in sustainable technologies \[699\]. The dominant ways in which major financial institutions take this approach are by creating “green” financial indexes that focus on low-carbon energy, clean technology, and/or environmental services \[700\] or by designing carbon-neutral investment portfolios that remove or under-weigh companies with relatively high carbon footprints \[701\]. This investment strategy is creating major shifts in certain sectors of the market (e.g. utilities and energy) towards renewable energy alternatives, which are seen as having a greater growth potential than traditional energy sources such as oil and gas \[702\]. While this approach currently does not utilize ML directly, we see the potential in applying deep learning both for portfolio selection (based on features of the stocks involved) and investment timing (using historical patterns to predict future demand), to maximize both the impact and scope of climate investment strategies.

### 13.2 Climate Analytics

The other main approach to climate finance is *climate analytics*, which aims to predict the financial effects of climate change, and is still gaining momentum in the mainstream financial community \[699\]. Since this is a predictive approach to addressing climate change from a financial perspective, it is one where ML can potentially make a greater impact. Climate analytics involves analyzing investment portfolios, funds and companies in order to pinpoint areas with heightened risk due to climate change, such as timber companies that could be bankrupted by wildfires or water extraction initiatives that could see their source polluted by shifting landscapes. Approaches used in this field include: using NLP techniques for identifying climate risks and investment opportunities in disclosures made by companies \[703\] as well as for analyzing the evolution of media coverage of climate change to dynamically hedge climate change risk \[704\]; using econometric approaches for developing arbitrage strategies that take advantage of the carbon risk factor in financial markets \[705\]; and ML approaches for forecasting the price of carbon in emission exchanges \[707, 708\].

To date, the field of climate finance has been largely neglected within the larger scope of financial research and analysis. This leaves many directions for improvement, such as (1) improving existing traditional portfolio optimization approaches using deep learning; (2) more in-depth modeling of the types of variables that constitute climate risk; (3) designing a statistical climate factor which can be used for projecting the variation of stock prices given a compound set of events; and (4) identifying direct and indirect climate risk exposure in annual company reports describing their financial performance, among other directions. ML

---

\[58\] For further reading regarding the impact of climate change on financial markets, see \[696–698\].

\[59\] Carbon pricing, e.g. via CO\(_2\) cap-and-trade or a carbon tax, is a commonly-suggested policy approach for getting firms to price future climate change impacts into their financial calculations. For an introduction to these topics, see \[706\].
plays a central role in these directions, and can play a big role in avoiding, or at least mitigating, global climate-related financial impacts.
Conclusion

Machine learning, like all technology, does not always make the world a better place – but it can. In the fight against climate change, we have seen that ML has significant contributions to offer across domain areas. ML can make systems more efficient (e.g. prevent electricity loss during transmission, consolidate freight, and reduce food waste). It can enable remote sensing and automatic monitoring (e.g. pinpoint deforestation, gather data on buildings, and track personal energy use). ML can provide fast approximations to time-intensive simulations (e.g. climate models and energy scheduling models), and it has the potential to lead to interpretable or causal models (e.g. for understanding weather patterns, informing policy makers, and planning for disasters). In all these cases, we emphasize that ML is only one part of the solution; it is a tool that enables other tools across fields.

Addressing the problems of climate change has the potential to both benefit society and present new directions for the field of machine learning. The solutions we envision require dialogue with fields outside and within computer science, which will lead not only to novel application domains but also to new methodological insights applicable across ML. For example, there is a need for research bringing together ideas from ML, symbolic AI, optimization, and dynamical systems, among other disciplines. Moreover, many of the problems we have discussed here highlight challenges topical to ML as a whole, such as interpretability, uncertainty quantification, and integration of prior knowledge and constraints.

The nature of the data involved also poses challenges and opportunities. For many of the applications we identify, data can be proprietary or include sensitive personal information. Where datasets exist, they are often open-ended, unlike typical ML benchmarks, where data have been organized with a specific task in mind. Datasets may include information from multiple heterogeneous sources, which must be integrated using domain knowledge. Moreover, the available data may not be representative of global use cases. For example, electricity systems in the US, where data are abundant, are very different from those in India, where data can be scarce. Tools from transfer learning and domain adaptation will likely prove essential in low-data settings. For some tasks, it may also be feasible to augment learning with carefully simulated data. Of course, the best option if possible is always more real data, and we strongly encourage public and private entities to release datasets and to solicit involvement from the ML community.

For those who want to apply ML to climate change, we provide a roadmap:

- **Learn.** Identify how your skills may be useful – we hope this paper provides a starting point.

- **Collaborate.** Find collaborators, who may be researchers, entrepreneurs, established companies, or policy-makers. Remember that for every domain we have discussed here, there are experts in that area who understand its opportunities and pitfalls, even if they do not necessarily understand ML.

- **Listen.** Listen to what your collaborators say is needed, and gather input more broadly as well to make sure your work will have the desired impact. Groundbreaking technologies have an impact, but so do well-constructed solutions to mundane problems.

- **Deploy.** Ensure that your work is deployed where its impact can be realized.

We call upon the machine learning community to use its skills as part of the global effort against climate change.
Acknowledgments

Electricity systems. We thank James Kelloway (National Grid ESO), Jack Kelly (Open Climate Fix), Zico Kolter (CMU), and Henry Richardson (WattTime) for their help and ideas in shaping this section. We also thank Samuel Buteau (Dalhousie University) and Marc Cormier (Dalhousie University) for their inputs on accelerated science and battery storage technologies; Julian Kates-Harbeck (Harvard) and Melrose Roderick (CMU) for their extensive inputs and ideas on nuclear fusion; and Alasdair Bruce (formerly National Grid ESO) for inputs on emissions factor forecasting and automated dispatch. Finally, we thank Lea Boche (EPRI), Guannan He (CMU), Jeremy Keen (CMU), Zico Kolter (CMU), Luke Lavin (CMU), Sanam Mirzazad (EPRI), and Matt Wytock (Gridmatic, Inc.) for helpful comments and feedback.

Transportation. We are grateful for advice from Alan T. Jenn (UC Davis) and Prithvi S. Acharya (CMU) on electric vehicles, Alexandre Jacquillat (CMU) on decarbonizing aviation, Michael Whiston (CMU) on hydrogen fuel cells, Evan Sherwin (CMU) on alternative fuels, and Samuel Buteau (Dalhousie University) on batteries.

Buildings and Cities. We thank Duccio Piovani (nam.R) and Jack Kelly (Open Climate Fix).

Industry. We appreciate Angela Acocella (MIT) and Kevin McCloskey’s (Google) constructive feedback, and Kipp Bradford’s (Yale) recommendations on embodied energy and refrigeration. Thanks to Allie Schwertner (Rockwell Automation), Greg Kochanski (Google), and Paul Weaver (Abstract) for their suggestions around optimizing industrial processes for low-carbon energy.

Land Use. We would like to give thanks to David Marvin (Salo) and Remi Charpentier (Tesselo) on remote sensing for land use. Max Nova (SilviaTerra) provided insight on forestry, Mark Crowley (University of British Columbia) on forest fire management, Benjamin Deleener (ChrysaLabs) on precision agriculture, and Lindsay Brin (Element AI) on soil chemistry.

Climate prediction. We thank Ben Kravitz (PNNL), Ghaleb Abdulla (LLNL), and David John Gagne II (UCAR) for enlightening conversations, and Ben Kravitz and Goodwin Gibbons (Imperial College London) for detailed feedback on this manuscript. This research was supported by a Department of Energy Computational Science Graduate Fellowship (DE-FG02-97ER25308).

Societal adaptation. We thank Loubna Benabbou (UQAR), Mike Schäfer (University of Zurich), Andrea Garcia Tapia (Stevens Tech), and Sarah M. Fletcher (MIT) for valuable conversations on the social aspects of climate change.

Solar geoengineering. We thank David Keith (Harvard), Peter Irvine (Harvard), Zhen Dai (Harvard), Colleen Golja (Harvard), Ross Boczar (UC Berkeley), Jon Proctor (UC Berkeley), Ben Kravitz (Indiana University), Andrew Lockley (University College London), Trude Storelvmo (University of Oslo), and Simon Gruber (University of Oslo) for help and useful feedback.

Tools for individuals. We thank Priyanka deSouza (MIT), Olivier Corradi (Tomorrow), Jack Kelly (Open Climate Fix), Ioana Marinescu (UPenn), and Aven Satre-Meloy (Oxford).

Education. We appreciated the constructive feedback received by Jacqueline Bourdeau (TÉLUQ University), who gave us valuable insights regarding the field of AIED.
Finance. We thank Himanshu Gupta (ClimateAI), and Bjarne Steffen (ETH Zürich) for the constructive discussions and the valuable feedback.

The authors gratefully acknowledge support from National Science Foundation grant 1803547, US Department of Energy contract DE-FG02-97ER25308, the Natural Sciences and Engineering Research Council of Canada, and the MIT Media Lab Consortium.
References


[187] Mariusz Bojarski, Davide Del Testa, Daniel Dworakowski, Bernhard Firner, Beat Flepp, Prasoon Goyal, Lawrence D. Jackel, Mathew Monfort, Urs Muller, Jiakai Zhang, Xin Zhang, Jake Zhao, and Karol Zieba. End to end learning for self-driving cars. CoRR, abs/1604.07316, 2016.


J Rhee, J Im, and S Park. Drought forecasting based on machine learning of remote sensing and long-range forecast data. *APEC Climate Center, Republic of Korea*, 2016.


[425] Veronica B Cashin, Daniel S Eldridge, Aimin Yu, and Dongyuan Zhao. Surface functionalization and manipu-
lation of mesoporous silica adsorbents for improved removal of pollutants: a review. *Environmental Science:

[426] Paul Raccuglia, Katherine C Elbert, Philip DF Adler, Casey Falk, Malia B Wenny, Aurelio Mollo, Matthias
Zeller, Sorelle A Friedler, Joshua Schrier, and Alexander J Norquist. Machine-learning-assisted materials dis-


[429] Sandra Ö Snæbjörnsdóttir and Sigurdur R Gislason. CO2 storage potential of basaltic rocks offshore Iceland.


emphasis on modeling approaches and practical simulations. *Water Resources Research*, 51(9):6846–6892,
2015.

[432] Shaoxing Mo, Yinhao Zhu, Nicholas Zabaras, Xiaoqing Shi, and Jichun Wu. Deep convolutional encoder-
decoder networks for uncertainty quantification of dynamic multiphase flow in heterogeneous media. *Water

[433] Dylan Moriarty, Laura Dobeck, and Sally Benson. Rapid surface detection of CO2 leaks from geologic seque-

[434] Bailian Chen, Dylan R Harp, Youzu Lin, Elizabeth H Keating, and Rajesh J Pawar. Geologic co 2 sequestra-
tion monitoring design: A machine learning and uncertainty quantification based approach. *Applied energy*,

Savarese, and Adam R Brandt. Machine vision for natural gas methane emissions detection using an infrared

[436] H Goosse, P Barriat, W Lefebvre, M Loutre, and V Zunz. *Introduction to climate dynamics and climate

Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri and


[533] UN Global Pulse. Using mobile phone data and airtime credit purchases to estimate food security. *New York: UN World Food Programme (WFP), Université Catholique de Louvain, Real Impact Analytics, Pulse Lab New York*, 2015.


